



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2023년05월08일  
(11) 등록번호 10-2530118  
(24) 등록일자 2023년05월03일

- (51) 국제특허분류(Int. Cl.)  
C12N 15/63 (2006.01) A61K 48/00 (2006.01)  
C12N 15/62 (2006.01) C12N 15/90 (2006.01)
- (52) CPC특허분류  
C12N 15/635 (2013.01)  
A61K 48/00 (2013.01)
- (21) 출원번호 10-2015-7004939
- (22) 출원일자(국제) 2013년07월21일  
심사청구일자 2018년07월20일
- (85) 번역문제출일자 2015년02월25일
- (65) 공개번호 10-2015-0056539
- (43) 공개일자 2015년05월26일
- (86) 국제출원번호 PCT/US2013/051418
- (87) 국제공개번호 WO 2014/018423  
국제공개일자 2014년01월30일
- (30) 우선권주장  
61/675,778 2012년07월25일 미국(US)  
(뒷면에 계속)
- (56) 선행기술조사문헌  
Nature Methods (2013.07.25.) 10(10):973-976  
(뒷면에 계속)

- (73) 특허권자  
더 브로드 인스티튜트, 인코퍼레이티드  
미국 02142 매사추세츠주 캠프리지 메인 스트리트 415  
메사추세츠 인스티튜트 오브 테크놀로지  
미국 02139 매사추세츠 캠프리지 메사추세츠 애브뉴 77  
프레지던트 앤드 펠로우즈 오브 하바드 칼리지  
미합중국, 메사추세츠 02138, 캠프리지, 퀸시스트리트17
- (72) 발명자  
장 평  
미국 02142 매사추세츠주 캠프리지 메인 스트리트 415 더 브로드 인스티튜트, 인코퍼레이티드 씨/오브리검 마크  
미국 01464 매사추세츠주 설리 빌리지 드라이브 13  
(뒷면에 계속)
- (74) 대리인  
특허법인코리아나

전체 청구항 수 : 총 24 항

심사관 : 김민정

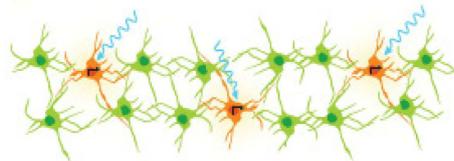
(54) 발명의 명칭 유도 DNA 결합 단백질 및 계놈 교란 도구 및 이의 적용

(57) 요약

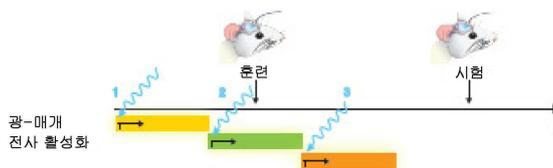
본 발명은 일반적으로 유도 전사 효과기를 사용할 수 있는 유전자 발현의 공간적이고 시간적인 제어에 사용되는 방법 및 조성물에 관한 것이다. 본 발명은 특히 세포에서의 관심 계놈 유전자좌의 발현을 변형하거나 교란하는 유도 방법에 관한 것이며, 이때 상기 계놈 유전자좌는 데옥시리보핵산(DNA) 결합 폴리펩티드를 포함하는 비-자연발생적 또는 조작된 조성물과 접촉될 수 있다.

대표도 - 도1

공간적 제어



시간적 제어



- (52) CPC특허분류  
*C12N 15/62* (2013.01)  
*C12N 15/63* (2013.01)  
*C12N 15/902* (2013.01)
- (72) 발명자  
**콩 레**  
 미국 02142 매사추세츠주 캠브리지 메모리얼 드라이브 100 아파트먼트 8-21비  
**코네르만 질파나**  
 미국 02139 매사추세츠주 캠브리지 매사추세츠 애비뉴 77 매사추세츠 인스티튜트 오브 테크놀로지 씨/오  
**샌자나 네빌 에스피**  
 미국 02139 매사추세츠주 캠브리지 매사추세츠 애비뉴 77 매사추세츠 인스티튜트 오브 테크놀로지 씨/오
- (56) 선행기술조사문헌  
 Acta Nature (2014) 6(3)22:19-40  
 Nature (2013.07.23.) 500:472-475  
 Cell (2013) 154(2):442-451  
 Nature Methods (2010) 7(12):973-975  
 FEBS Letters (2012.03.03.) 586:2084-2096  
 Molecular Therapy (2012.05.) 20(Supplement 1):S194-S195  
 US20110145940 A1\*  
 US20110301073 A1\*  
 WO2013082519 A2  
 KR101844123 B1  
 KR1020150101476 A  
 KR1020150016588 A  
 KR101656236 B1  
 WO2014018423 A2  
 \*는 심사관에 의하여 인용된 문헌
- (30) 우선권주장  
 61/721,283 2012년11월01일 미국(US)  
 61/736,465 2012년12월12일 미국(US)  
 61/794,458 2013년03월15일 미국(US)  
 61/835,973 2013년06월17일 미국(US)
-

**명세서**

**청구범위**

**청구항 1**

하기를 포함하는, 균집화된 규칙적 공간화 단길이 회귀성 반복물(CRISPR)-Cas 복합체로서:

- Cas9, 여기에서 Cas9 은 둘 이상의 핵 국부화 신호를 갖고;
- tracr mate 서열에 연결된 안내 서열, 및
- tracr mate 서열의 전부 또는 일부와 하이브리드화 가능한 tracr 서열;

여기에서, 안내 서열은 진핵세포에서의 표적 서열과 상보성을 갖도록 설계되고, 진핵세포에서의 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 유도할 수 있는 CRISPR-Cas 복합체.

**청구항 2**

하기를 포함하는, 균집화된 규칙적 공간화 단길이 회귀성 반복물(CRISPR)-Cas 벡터 시스템으로서:

- a) (1) tracr mate 서열에 연결된 안내 서열 및 (2) tracr mate 서열의 전부 또는 일부와 하이브리드화 가능한 tracr 서열;을 인코딩하는 하나 이상의 서열에 작동가능하게 연결된 제 1 조절 요소;
- b) Cas9 을 인코딩하는 서열에 작동가능하게 연결된 제 2 조절 요소, 여기에서 Cas9 은 둘 이상의 핵 국부화 신호를 갖고;

이때, 성분 a) 및 b) 는 시스템의 동일하거나 상이한 벡터 상에 위치하고, 안내 서열은 진핵세포에서의 표적 서열과 상보성을 갖도록 설계되는 CRISPR-Cas 벡터 시스템.

**청구항 3**

제 2 항에 있어서, Cas9 를 인코딩하는 서열이 진핵세포에서의 발현을 위해 최적화된 코돈인 CRISPR-Cas 벡터 시스템.

**청구항 4**

제 2 항에 있어서, 둘 이상의 안내 서열을 포함하는 CRISPR-Cas 벡터 시스템.

**청구항 5**

제 2 항에 있어서, 상기 동일하거나 상이한 벡터는 하나 이상의 바이러스 벡터인 CRISPR-Cas 벡터 시스템.

**청구항 6**

제 5 항에 있어서, 상기 하나 이상의 바이러스 벡터는 하나 이상의 레트로바이러스, 아데노바이러스, 아데노-연관 알파바이러스 또는 단순포진 바이러스 벡터인 CRISPR-Cas 벡터 시스템.

**청구항 7**

제 6 항에 있어서, 상기 레트로바이러스는 렌티바이러스인 CRISPR-Cas 벡터 시스템.

**청구항 8**

제 6 항에 있어서, 상기 알파바이러스는 신드비스 바이러스 (Sindbis virus), 쉘리키 삼림 바이러스 (Semliki Forest virus) 또는 베네수엘라 말 뇌염 바이러스 (Venezuelan equine encephalitis virus) 인 CRISPR-Cas 벡터 시스템.

**청구항 9**

제 2 항에 있어서, 성분 a) 및 b) 는 동일한 벡터 상에 위치한 CRISPR-Cas 벡터 시스템.

**청구항 10**

제 2 항 내지 제 9 항 중 어느 한 항에 있어서, tracr mate 서열은 tracr 서열에 연결되어 키메라 RNA (chiRNA) 를 형성하는 CRISPR-Cas 벡터 시스템.

**청구항 11**

제 2 항 내지 제 9 항 중 어느 한 항에 있어서, Cas9 은 표적 서열을 함유하는 표적 폴리뉴클레오티드의 둘 모두의 가닥의 절단을 유도하는 CRISPR-Cas 벡터 시스템.

**청구항 12**

제 2 항 내지 제 9 항 중 어느 한 항에 있어서, Cas9 는 S. 피오게네스 Cas9 의 위치 번호와 관련된 D10A, H840A, N854A 및 N863A 로 이루어진 균으로부터 선택되는 돌연변이를 포함하는 CRISPR-Cas 벡터 시스템.

**청구항 13**

제 2 항 내지 제 9 항 중 어느 한 항에 있어서, Cas9 는 아미노-말단에서 하나 이상의 핵 국부화 신호, 및 카르복시-말단에서 하나 이상의 핵 국부화 신호를 포함하는 CRISPR-Cas 벡터 시스템.

**청구항 14**

제 2 항 내지 제 9 항 중 어느 한 항에 있어서, Cas9 는 S. 피오게네스 Cas9 인 CRISPR-Cas 벡터 시스템.

**청구항 15**

제 2 항 내지 제 9 항 중 어느 한 항에 있어서, Cas9 는 S. 썬모필루스 Cas9 인 CRISPR-Cas 벡터 시스템.

**청구항 16**

제 2 항 내지 제 9 항 중 어느 한 항에 있어서, 치료에 사용되는 CRISPR-Cas 벡터 시스템.

**청구항 17**

제 2 항 내지 제 9 항 중 어느 한 항에 있어서, 게놈 조작에 사용되는 CRISPR-Cas 벡터 시스템으로서, 상기 사용이 치료에 의한 동물 또는 인간 신체의 처치 방법이 아니고, 상기 사용이 인간의 생식-계열 유전적 동일성을 변형시키는 과정이 아닌 CRISPR-Cas 벡터 시스템.

**청구항 18**

제 1 항에 있어서, tracr mate 서열은 tracr 서열에 연결되어 키메라 RNA (chiRNA) 를 형성하는 CRISPR-Cas 복합체.

**청구항 19**

제 1 항에 있어서, Cas9 은 표적 서열을 함유하는 표적 폴리뉴클레오티드의 둘 모두의 가닥의 절단을 유도하는 CRISPR-Cas 복합체.

**청구항 20**

제 1 항, 제 18 항 및 제 19 항 중 어느 한 항에 있어서, Cas9 는 S. 피오게네스 Cas9 의 위치 번호와 관련된 D10A, H840A, N854A 및 N863A 로 이루어진 균으로부터 선택되는 돌연변이를 포함하는 CRISPR-Cas 복합체.

**청구항 21**

제 1 항, 제 18 항 및 제 19 항 중 어느 한 항에 있어서, Cas9 는 아미노-말단에서 하나 이상의 핵 국부화 신호, 및 카르복시-말단에서 하나 이상의 핵 국부화 신호를 포함하는 CRISPR-Cas 복합체.

**청구항 22**

제 1 항, 제 18 항 및 제 19 항 중 어느 한 항에 있어서, Cas9 는 S. 피오게네스 Cas9 인 CRISPR-Cas 복합체.

**청구항 23**

제 1 항, 제 18 항 및 제 19 항 중 어느 한 항에 있어서, Cas9 는 S. 썬모필루스 Cas9 인 CRISPR-Cas 복합체.

**청구항 24**

제 1 항, 제 18 항 및 제 19 항 중 어느 한 항에 있어서, 치료에 사용되는 CRISPR-Cas 복합체.

**청구항 25**

삭제

**청구항 26**

삭제

**청구항 27**

삭제

**청구항 28**

삭제

**청구항 29**

삭제

**청구항 30**

삭제

**청구항 31**

삭제

**청구항 32**

삭제

**청구항 33**

삭제

**청구항 34**

삭제

**청구항 35**

삭제

**청구항 36**

삭제

**청구항 37**

삭제

**청구항 38**

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

- 청구항 55
- 삭제
- 청구항 56
- 삭제
- 청구항 57
- 삭제
- 청구항 58
- 삭제
- 청구항 59
- 삭제
- 청구항 60
- 삭제
- 청구항 61
- 삭제
- 청구항 62
- 삭제
- 청구항 63
- 삭제
- 청구항 64
- 삭제
- 청구항 65
- 삭제
- 청구항 66
- 삭제
- 청구항 67
- 삭제
- 청구항 68
- 삭제
- 청구항 69
- 삭제
- 청구항 70
- 삭제

청구항 71

삭제

청구항 72

삭제

**발명의 설명**

**기술 분야**

- [0001] 관련 명세서 및 참고문헌
- [0002] 이 명세서는 미국 가출원 일련 번호 61/675,778 (2012 년 7 월 25 일 출원), 61/721,283 (2012 년 11 월 1 일 출원), 61/736,465 (2012 년 12 월 12 일 출원), 61/794,458 (2013 년 3 월 15 일 출원) 및 61/835,973 (2013 년 6 월 17 일 출원) (발명의 명칭: 유도성 DNA 결합 단백질 및 게놈 교란 도구 및 이의 적용 (INDUCIBLE DNA BINDING PROTEINS AND GENOME PERTURBATION TOOLS AND APPLICATIONS THEREOF)) 에 우선권을 주장한다.
- [0003] 미국 가출원 번호 61/565,171 (2011 년 11 월 30 일 출원) 및 미국 출원 번호 13/554,922 (2012 년 7 월 30 일 출원) 및 13/604,945 (2012 년 9 월 6 일 출원) (발명의 명칭: 설계자 TAL 효과기에 대한 뉴클레오티드-특이적 인지 서열 (NUCLEOTIDE-SPECIFIC RECOGNITION SEQUENCES FOR DESIGNER TAL EFFECTORS)) 를 또한 참고한다.
- [0004] 미국 가출원 번호 61/736,527 (2012 년 12 월 12 일 출원); 61/748,427 (2013 년 1 월 2 일 출원); 61/757,972 (2013 년 1 월 29 일 출원), 61/768,959 (2013 년 2 월 25 일 출원) 및 61/791,409 (2013 년 3 월 15 일 출원) (발명의 명칭: 서열 조작을 위한 시스템 방법 및 조성물 (SYSTEMS METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION)) 를 또한 참고한다.
- [0005] 미국 가출원 번호 61/758,468 (2013 년 1 월 30 일 출원) 및 61/769,046 (2013 년 3 월 15 일 출원) (발명의 명칭: 서열 조작을 위한 시스템, 방법 및 조성물의 조작 및 최적화 (ENGINEERING AND OPTIMIZATION OF SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION)) 을 또한 참고한다.
- [0006] 미국 가출원 번호 61/835,931; 61/835,936; 61/836,080; 61/836,101; 61/836,123 및 61/836,127 (2013 년 6 월 17 일 출원) 을 또한 참고한다.
- [0007] 미국 가출원 번호 61/842,322 (2013 년 7 월 2 일 출원) (발명의 명칭: 유전자 생성물의 발현 변형을 위한 CRISPR-CAS 시스템 및 방법 (CRISPR-CAS SYSTEMS AND METHODS FOR ALTERING EXPRESSION OF 유전자 산물S)) 및 미국 가출원 번호 61/847,537 (2013 년 7 월 17 일 출원) (발명의 명칭: 서열 조작 및 적용을 위한 시스템, 방법 및 조성물의 전달, 조작 및 최적화 (DELIVERY, ENGINEERING AND OPTIMIZATION OF SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION AND APPLICATIONS)) 을 또한 참고한다.
- [0008] 진술한 출원 및 이에 인용되거나 그의 실행 동안 인용된 모든 문헌 ("출원 인용 문헌") 및 출원 인용 문헌에서의 인용되거나 참고된 모든 문헌, 및 본원에 인용되거나 참고된 모든 문헌 ("본원에 인용된 문헌"), 및 본원에 인용된 문헌에서 인용되거나 참고된 모든 문헌은 본원에 언급되거나 본원에 참고로 포함되는 임의의 문헌에서 언급된 임의의 제품에 대한 임의의 제조사 지시사항, 설명서, 제품 사양 및 제품 시트와 함께 본원에 참고로 포함되고, 본 발명의 실행에서 이용될 수 있다. 보다 특히, 모든 참고 문헌은 개별적 문헌이 참고로 포함되는 것으로 구체적이고 개별적으로 표시되는 바와 동일한 정도로 참고로 포함된다.
- [0009] 발명의 분야
- [0010] 본 발명은 일반적으로 유도 전사 효과기를 사용할 수 있는 게놈 교란과 같은 유전자 발현의 공간적이고 시간적인 제어에 사용되는 방법 및 조성물에 관한 것이다.
- [0011] 연방 자금 지원 설명
- [0012] 본 발명은 미국 국립 보건원에 의해 수여된 R01NS073124 및 개척자 상 (Pioneer Award) 1 DP IMH 100706 하 정부 지원으로 이루어졌다. 정부는 본 발명에 특정 권리를 갖는다.

**배경 기술**

- [0013] 정상 유전자 발현은 조심스럽게 조직된 시간적이고 공간적인 성분, 유기체의 정상 발생, 항상성 유지 및 항상에 필요한 정확도를 갖는 동적 프로세스이다. 결과적으로, 유전자 또는 유전자 집합의 증가, 감소 또는 변형된 기능에 의한 필요한 유전자 발현 패턴의 조절 장애는 광범위한 병리학과 연결된다. 시공간적으로 정확한 방식으로 유전자 발현을 조절할 수 있는 기술은 정상 생물학적 프로세스 및 질환 메커니즘을 책임지는 유전자적 신호의 설명을 가능하게 할 것이다. 이러한 기술적 필요성을 다루기 위해서, 출원인은 유전자 발현, 특히 내재적 유전자 발현의 광-매개 제어를 제공하는 광-유도 전사 효과기 (LITE) 를 조절할 수 있는 유도 분자 도구를 개발하였다.
- [0014] 유도 유전자 발현 시스템은 삽입된 개방 판독 프레임 또는 shRNA 서열의 화학적 유도 활성화가 가능하도록 통상 설계되어, 각각 유전자 과발현 또는 억제를 초래한다. 과발현에 대한 개방 판독 프레임을 사용하는 것의 불리한 점은 스플라이스 변이의 손실 및 유전자 크기 제한을 포함한다. RNA 간섭을 통한 유전자 발현은, 이의 인간 생물학에 있어서의 형질전환적 동력에도 불구하고, 복잡한 오프-타겟 (off-target) 효과에 의해 방해될 수 있다. 에스트로겐, 엑디손 및 FKBP12/FRAP 기반 시스템을 포함하는 특정 유도 시스템은 내재적 유전자의 오프-타겟을 활성화시키는 것으로 알려져있다. 장기간 항생제 처리의 잠재적으로 유해한 효과는 테트라사이클린 전사활성제 (TET) 기반 시스템의 사용을 복잡하게 할 수 있다. 이러한 화학적 유도 시스템의 시간적 정확도는 유도물질 흡수 및 제거의 속도에 따라 좌우된다. 또한, 유도물질이 일반적으로 전신 전달되기 때문에, 이러한 시스템의 공간적 정확도는 외재적 벡터 전달의 정확도에 의해 결부된다.
- [0015] 미국 특허 공개 번호 20030049799 는 사전선택된 자극에 반응하여 검출가능한 산출물이 생기도록 조작된 자극-반응성 스위치에 관한 것이다.
- [0016] 관심 게놈 유전자좌의 효율적이고 정확한 공간적 및 시간적 제어를 가능하게 하는 방법 및 조성물에 대한 명백한 필요성이 존재한다. 이들 방법 및 조성물은 생체내 및 시험관내 모두에서의 게놈 발현의 조절 및 조정을 제공할 수 있을 뿐 아니라 수많은 질환 병리학에 대한 신규한 치료 방법을 제공할 수 있다.
- [0017] 상기 출원에서의 임의의 문헌의 인용 또는 확인은 그러한 문헌이 본 발명에 대해 선행 기술로서 이용가능하다는 것을 허용하는 것은 아니다.

**발명의 내용**

- [0018] 한 양태에서 본 발명은 하나 이상의 스위치를 포함할 수 있는 비-자연 발생적 또는 조작된 TALE 또는 CRISPR-Cas 시스템을 제공하며, 이때 상기 TALE 또는 CRISPR-Cas 시스템의 활성화는 스위치에 대한 하나 이상의 유도물질 에너지원과의 접촉에 의해 제어된다. 본 발명의 구현예에서 상기 TALE 또는 CRISPR-Cas 시스템의 활성화 또는 하나 이상의 스위치에 관한 제어가 활성화되고, 증진되고, 종료되거나 억제될 수 있다. 하나 이상의 유도물질 에너지원과의 접촉은 제 1 효과 및 제 2 효과를 초래할 수 있다. 제 1 효과는 핵 이입, 핵 이출, 2 차 성분 (예컨대 효과기 분자) 의 영입, 형태적 변화 (단백질, DNA 또는 RNA 의), 절단, 카고 (cargo) (예컨대 감혀있는 분자 또는 보조-인자) 의 배출, 결합 또는 해리 중 하나 이상일 수 있다. 제 2 효과는 상기 TALE 또는 CRISPR-Cas 시스템의 활성화 또는 하나 이상의 스위치에 관한 제어의 활성화, 증진, 종료 또는 억제 중 하나 이상일 수 있다. 한 구현예에서 제 1 효과 및 제 2 효과는 캐스케이드로 발생할 수 있다.
- [0019] 본 발명의 또 다른 양태에서 TALE 또는 CRISPR-Cas 시스템은 하나 이상의 핵 국부화 신호 (NLS), 핵 이출 신호 (NES), 기능적 도메인, 유연성 링커, 돌연변이, 결실, 변형 또는 절두를 추가로 포함할 수 있다. NLS, NES 또는 기능적 도메인 중 하나 이상은 조건적으로 활성화되거나 불활성화될 수 있다. 또 다른 구현예에서, 돌연변이는 전사 인자 상동 부위에서의 돌연변이, DNA 결합 도메인에서의 돌연변이 (예컨대 염기성 헬릭스 루프 헬릭스의 염기성 잔기 돌연변이화), 내인성 NLS 에서의 돌연변이 또는 내인성 NES 에서의 돌연변이 중 하나 이상일 수 있다. 본 발명은 유도물질 에너지원이 열, 초음파, 전자기 에너지 또는 화학 물질일 수 있다는 것을 이해한다. 본 발명의 바람직한 구현예에서, 유도물질 에너지원은 항생제, 소분자, 호르몬, 호르몬 유도체, 스테로이드 또는 스테로이드 유도체일 수 있다. 보다 바람직한 구현예에서, 유도물질 에너지원은 아브시스산 (ABA), 독시사이클린 (DOX), 큐메이트, 라파마이신, 4-히드록시타복시펜 (4OHT), 에스트로겐 또는 엑디손일 수 있다. 본 발명은 항생제 기반 유도 시스템, 전자기 에너지 기반 유도 시스템, 소분자 기반 유도 시스템, 핵 수용체 기반 유도 시스템 및 호르몬 기반 유도 시스템으로 이루어지는 군에서 선택될 수 있는 하나 이상의 스위치를 제공한다. 보다 바람직한 구현예에서 하나 이상의 스위치는 테트라사이클린 (Tet)/DOX 유도 시스템, 광 유도 시스템, ABA 유도 시스템, 큐메이트 억제제/오퍼레이터 시스템, 4OHT/에스트로겐 유도 시스템, 엑디손-기반 유도 시스템 및 FKBP12/FRAP (FKBP12-라파마이신 복합체) 유도 시스템으로 이루어지는 군에서 선택

될 수 있다.

- [0020] 본 발명의 한 양태에서 유도물질 에너지원은 전자기 에너지이다. 전자기 에너지는 450nm-700nm 범위의 파장을 갖는 가시광의 성분일 수 있다. 바람직한 구현예에서 가시광의 성분은 450nm-500nm 범위의 파장을 가질 수 있으며 청색광일 수 있다. 청색광은  $0.2 \text{ mW/cm}^2$  이상, 또는 보다 바람직하게는  $4 \text{ mW/cm}^2$  이상의 강도를 가질 수 있다. 또 다른 구현예에서, 가시광의 성분은 620-700nm 범위의 파장을 가질 수 있으며 적색광이다.
- [0021] 본 발명은 하나 이상의 기능적 도메인이 하기로 이루어지는 군에서 선택될 수 있는 시스템을 포함한다: 트랜스포사아제 도메인, 인테그라아제 도메인, 재조합효소 도메인, 리졸바아제 도메인, 인버타아제 도메인, 프로테아제 도메인, DNA 메틸트랜스퍼라아제 도메인, DNA 히드록실메틸라아제 도메인, DNA 데메틸라아제 도메인, 히스톤 아세틸라아제 도메인, 히스톤 데아세틸라아제 도메인, 뉴클레아제 도메인, 억제제 도메인, 활성제 도메인, 핵-국부화 신호 도메인, 전사-조절 단백질 (또는 전사 복합체 영입) 도메인, 세포 흡수 활성 관련 도메인, 핵산 결합 도메인, 항체 제시 도메인, 히스톤 개질 효소, 히스톤 개질 효소의 영입제; 히스톤 개질 효소의 저해제, 히스톤 메틸트랜스퍼라아제, 히스톤 데메틸라아제, 히스톤 키나아제, 히스톤 포스파타아제, 히스톤 리보실라아제, 히스톤 데리보실라아제, 히스톤 유비퀴티나아제, 히스톤 데유비퀴티나아제, 히스톤 비오틴아아제 및 히스톤 테일 프로테아제.
- [0022] 본 발명은 또한 유전적 또는 비유전적 관심 유전자좌의 교란을 위한 시스템의 용도를 제공한다. 또한, 약화합물의 제조를 위한 시스템의 용도가 제공된다.
- [0023] 추가 양태에서, 본 발명은 하나 이상의 스위치를 포함하는 비-자연 발생적 또는 조작된 TALE 또는 CRISPR-Cas 시스템을 제공하는 것을 포함하는, 상기 TALE 또는 CRISPR-Cas 시스템의 제어 방법을 제공하며, 이때 상기 TALE 또는 CRISPR-Cas 시스템의 활성화는 스위치에 대한 하나 이상의 유도물질 에너지원과의 접촉에 의해 제어된다.
- [0024] 본 발명의 한 구현예에서, 본 발명은 상기 TALE 또는 CRISPR-Cas 시스템의 활성화 또는 하나 이상의 스위치에 관한 제어가 활성화되고, 증진되고, 완료되거나 억제될 수 있는 방법을 제공한다. 하나 이상의 유도물질 에너지원과의 접촉으로 제 1 효과 및 제 2 효과가 생길 수 있다. 제 1 효과는 핵 이입, 핵 이출, 2 차 성분 (예컨대 효과기 분자)의 유도, 형태적 변화 (단백질, DNA 또는 RNA 의), 카고 (예컨대 갇힌 분자 또는 보조-인자)의 절단, 배출, 결합 또는 해리 중 하나 이상일 수 있다. 제 2 효과는 상기 TALE 또는 CRISPR-Cas 시스템의 활성화 또는 하나 이상의 스위치에 관한 제어의 활성화, 증진, 완료 또는 억제 중 하나 이상일 수 있다. 한 구현예에서 제 1 효과 및 제 2 효과는 캐스케이드로 발생할 수 있다.
- [0025] 본 발명의 방법의 또 다른 양태에서 TALE 또는 CRISPR-Cas 시스템은 하나 이상의 핵 국부화 신호 (NLS), 핵 이출 신호 (NES), 기능적 도메인, 유연성 링커, 돌연변이, 결실, 변형 또는 절두를 추가로 포함할 수 있다. NLS, NES 또는 기능적 도메인 중 하나 이상은 조건적으로 활성화되거나 불활성화될 수 있다. 또 다른 구현예에서, 돌연변이는 전사 인자 상동 부위에서의 돌연변이, DNA 결합 도메인에서의 돌연변이 (예컨대 염기성 헬릭스 루프 헬릭스의 염기성 잔기의 돌연변이화), 내인성 NLS 에서의 돌연변이 또는 내인성 NES 에서의 돌연변이 중 하나 이상일 수 있다. 본 발명은 유도물질 에너지원이 열, 초음파, 전자기 에너지 또는 화학 물질일 수 있다는 것을 이해한다. 본 발명의 바람직한 구현예에서, 유도물질 에너지원은 항생제, 소분자, 호르몬, 호르몬 유도체, 스테로이드 또는 스테로이드 유도체일 수 있다. 보다 바람직한 구현예에서, 유도물질 에너지원은 아브시스산 (ABA), 독시사이클린 (DOX), 큐메이트, 라파마이신, 4-히드록시타목시펜 (4OHT), 에스트로겐 또는 엑디손일 수 있다. 본 발명은 항생제 기반 유도 시스템, 전자기 에너지 기반 유도 시스템, 소분자 기반 유도 시스템, 핵 수용체 기반 유도 시스템 및 호르몬 기반 유도 시스템으로 이루어지는 군에서 선택될 수 있는 하나 이상의 스위치를 제공한다. 보다 바람직한 구현예에서 하나 이상의 스위치는 테트라사이클린 (Tet)/DOX 유도 시스템, 광 유도 시스템, ABA 유도 시스템, 큐메이트 억제제/오퍼레이터 시스템, 4OHT/에스트로겐 유도 시스템, 엑디손-기반 유도 시스템 및 FKBP12/FRAP (FKBP12-라파마이신 복합체) 유도 시스템으로 이루어지는 군에서 선택될 수 있다.
- [0026] 본 발명의 방법의 한 양태에서 유도물질 에너지원은 전자기 에너지이다. 전자기 에너지는 450nm-700nm 범위의 파장을 갖는 가시광의 성분일 수 있다. 바람직한 구현예에서 가시광의 성분은 450nm-500nm 범위의 파장을 가질 수 있고 청색광일 수 있다. 청색광은  $0.2 \text{ mW/cm}^2$  이상, 또는 보다 바람직하게는  $4 \text{ mW/cm}^2$  이상의 강도를 가질 수 있다. 또 다른 구현예에서, 가시광의 성분은 620-700nm 범위의 파장을 가질 수 있고 적색광이다.
- [0027] 본 발명은 하나 이상의 기능적 도메인이 하기로 이루어지는 군에서 선택될 수 있는 방법을 포함한다: 트랜스포

사아제 도메인, 인테그라아제 도메인, 재조합효소 도메인, 리졸바아제 도메인, 인버타아제 도메인, 프로테아제 도메인, DNA 메틸트랜스퍼라아제 도메인, DNA 히드록실메틸라아제 도메인, DNA 데메틸라아제 도메인, 히스톤 아세틸라아제 도메인, 히스톤 데아세틸라아제 도메인, 뉴클레아제 도메인, 억제제 도메인, 활성제 도메인, 핵-국부화 신호 도메인, 전사-조절 단백질 (또는 전사 복합체 영입) 도메인, 세포 흡수 활성 관련 도메인, 핵산 결합 도메인, 항체 제시 도메인, 히스톤 개질 효소, 히스톤 개질 효소의 영입제; 히스톤 개질 효소의 저해제, 히스톤 메틸트랜스퍼라아제, 히스톤 데메틸라아제, 히스톤 키나아제, 히스톤 포스파타아제, 히스톤 리보실라아제, 히스톤 데리보실라아제, 히스톤 유비퀴티나아제, 히스톤 데유비퀴티나아제, 히스톤 비오틴아아제 및 히스톤 데일 프로테아제.

[0028] 본 발명의 추가 양태는 TALE 시스템이 하기를 포함하는 DNA 결합 폴리펩티드를 포함하는 본원에 기재한 바와 같은 시스템 또는 방법을 제공한다:

[0029] (i) 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 개 또는 그 이상의 전사 활성제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너와 결합할 수 있음), 및/또는

[0030] (ii) 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 상호작용 파트너에 결합함).

[0031] 본 발명의 시스템 및 방법은 하기를 포함하는 DNA 결합 폴리펩티드를 제공하며: (a) N-말단 캡핑 부위 (b) 관심 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 내지 40 개의 전사 활성제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 DNA 결합 도메인, 및 (c) C-말단 캡핑 부위, 여기서 (a), (b) 및 (c) 가 사전결정된 N-말단에서 C-말단의 배향으로 배열될 수 있고, 이때 게놈 유전자좌가 표적 DNA 서열 5'-T<sub>0</sub>N<sub>1</sub>N<sub>2</sub>...N<sub>z</sub> N<sub>z+1</sub>-3' (여기서 T<sub>0</sub> 및 N = A, G, T 또는 C 임) 을 포함하고, 상기 표적 DNA 서열은 DNA 결합 도메인에 결합하고, DNA 결합 도메인은 (X<sub>1-11</sub>-X<sub>12</sub>X<sub>13</sub>-X<sub>14-33</sub> 또는 34 또는 35)<sub>z</sub> (여기서 X<sub>1-11</sub> 은 11 개의 인접한 아미노산 사슬이고, X<sub>12</sub>X<sub>13</sub> 은 반복 가변 이중잔기 (RVD) 이고, X<sub>14-33</sub> 또는 34 또는 35 는 21, 22 또는 23 개의 인접한 아미노산 사슬이고, z 는 적어도 5 내지 40 일 수 있음) 를 포함할 수 있고, 이때 상기 폴리펩티드는 코돈 최적화된 핵산 분자에 의해 인코딩되고 이로부터 번역되어 폴리펩티드가 바람직하게는 관심 유전자좌의 DNA 에 결합할 수 있다.

[0032] 추가 구현예에서, 본 발명의 시스템 또는 방법은 야생형 N-말단 캡핑 부위의 147 개의 인접한 아미노산을 포함하는 N-말단 캡핑 부위 또는 이의 단편, 또는 야생형 C-말단 캡핑 부위의 68 개의 인접한 아미노산을 포함하는 C-말단 캡핑 부위 또는 이의 단편, 또는 야생형 N-말단 캡핑 부위의 136 개의 인접한 아미노산을 포함하는 N-말단 캡핑 부위 또는 이의 단편 및 야생형 C-말단 캡핑 부위의 183 개의 인접한 아미노산을 포함하는 C-말단 캡핑 부위 또는 이의 단편을 제공한다. 또 다른 구현예에서, 하나 이상의 RVD 는 (a) 구아닌 (G) 의 인지를 위한 HH, KH, NH, NK, NQ, RH, RN, SS, NN, SN, KN; (b) 아데닌 (A) 의 인지를 위한 NI, KI, RI, HI, SI; (c) 티민 (T) 의 인지를 위한 NG, HG, KG, RG; (d) 시토신 (C) 의 인지를 위한 RD, SD, HD, ND, KD, YG; (e) A 또는 G 의 인지를 위한 NV, HN; 및 (f) A 또는 T 또는 G 또는 C 의 인지를 위한 H\*, HA, KA, N\*, NA, NC, NS, RA, S\* (여기서 (\*) 는 X13 에서의 아미노산이 부재함을 의미함) 로 이루어지는 군에서 선택될 수 있다.

[0033] 또 다른 구현예에서 하나 이상의 RVD 는 (a) 구아닌 (G) 의 인지를 위한 HH, KH, NH, NK, NQ, RH, RN, SS; (b) 아데닌 (A) 의 인지를 위한 SI; (c) 티민 (T) 의 인지를 위한 HG, KG, RG; (d) 시토신 (C) 의 인지를 위한 RD, SD; (e) A 또는 G 의 인지를 위한 NV, HN 및 (f) A 또는 T 또는 G 또는 C 의 인지를 위한 H\*, HA, KA, N\*, NA, NC, NS, RA, S\* (여기서 (\*) 는 X13 에서의 아미노산이 부재함을 의미함) 로 이루어지는 군에서 선택될 수 있다. 바람직한 구현예에서, G 의 인지를 위한 RVD 는 RN, NH, RH 또는 KH 이거나; A 의 인지를 위한 RVD 는 SI 이거나; T 의 인지를 위한 RVD 는 KG 또는 RG 이고; C 의 인지를 위한 RVD 는 SD 또는 RD 이다. 또 다른 구현예에서, 하기 중 하나 이상이 존재한다: X<sub>1-4</sub> 에서의 [LTLD] 또는 [LTLA] 또는 [LTQV], 또는 위치 X<sub>30-33</sub> 또는 X<sub>31-34</sub> 또는 X<sub>32-35</sub> 에서의 [EQHG] 또는 [RDHG].

[0034] 본 발명의 한 양태에서 TALE 시스템은 AAV 또는 렌티바이러스 (lentivirus) 벡터 내로 패키징된다.

[0035] 본 발명의 추가 양태는 CRISPR 시스템이 하기를 포함하는 벡터 시스템을 포함할 수 있는, 본원에 기재한 바와

같은 시스템 또는 방법을 제공하며: a) 관심 유전자좌를 표적화하는 CRISPR-Cas 시스템 안내 RNA 에 작동가능하게 연결된 제 1 조절 요소, b) Cas 단백질에 작동가능하게 연결된 제 2 조절 유도 요소, 이때 성분 (a) 및 (b) 는 시스템의 동일하거나 상이한 벡터 상에 위치할 수 있고, 안내 RNA 는 관심 유전자좌의 DNA 를 표적화하고, Cas 단백질 및 안내 RNA 는 함께 자연적으로 발생하지 않는다. 본 발명의 바람직한 구현예에서, Cas 단백질 은 Cas9 효소이다. 본 발명은 또한 AAV 또는 렌티바이러스인 벡터를 제공한다.

[0036] 본 발명은 특히, 관심 게놈 유전자좌의 발현 변형의 유도 방법 및 게놈 관심 게놈 유전자좌 (게놈 유전자좌는 데옥시리보핵산 (DNA) 결합 폴리펩티드를 포함하는 비-자연 발생적 또는 조작된 조성물과 접촉할 수 있음) 의 발현을 유도가능하게 변형하는 조성물에 관한 것이다.

[0037] 이러한 폴리펩티드는 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 개 또는 그 이상의 전사 활성화제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 DNA 결합 도메인을 포함할 수 있다. 상기 에너지 민감성 단백질 또는 이의 단편은 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너에 결합할 수 있다. 폴리펩티드는 또한, 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 변이체 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인을 포함할 수 있으며, 이때 에너지 민감성 단백질 또는 이의 단편은 에너지원에 의한 유도시 상호작용 파트너에 결합할 수 있다. 상기 방법은 또한 에너지원을 적용하고 게놈 유전자좌의 발현이 변형되는 것을 측정하는 것을 포함할 수 있다. 본 발명의 바람직한 구현예에서 게놈 유전자좌는 세포 내에 있을 수 있다.

[0038] 본 발명은 또한 관심 게놈 유전자좌 발현 억제제의 유도 방법 및 관심 게놈 유전자좌의 발현을 유도가능하게 억제하는 조성물에 관한 것이며, 이때 상기 게놈 유전자좌는 DNA 결합 폴리펩티드를 포함하는 비-자연 발생적 또는 조작된 조성물과 접촉될 수 있다.

[0039] 폴리펩티드는 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 억제제 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 개 또는 그 이상의 전사 활성화제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 DNA 결합 도메인을 포함할 수 있다. 에너지 민감성 단백질 또는 이의 단편은 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너와 결합할 수 있다. 폴리펩티드는 또한, 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 변이체 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인을 포함할 수 있으며, 이때 에너지 민감성 단백질 또는 이의 단편은 에너지원에 의한 유도시 상호작용 파트너에 결합할 수 있다. 상기 방법은 또한 에너지원을 적용하고, 게놈 유전자좌의 발현이 억제되는 것을 측정하는 것을 포함할 수 있다. 본 발명의 바람직한 구현예에서 게놈 유전자좌는 세포 내에 있을 수 있다.

[0040] 본 발명은 또한 관심 게놈 유전자좌의 발현 활성화 유도 방법 및 관심 게놈 유전자좌의 발현을 유도가능하게 활성화하는 조성물에 관한 것이며, 여기서 게놈 유전자좌는 DNA 결합 폴리펩티드를 포함하는 비-자연 발생적 또는 조작된 조성물과 접촉될 수 있다.

[0041] 폴리펩티드는 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 활성화제 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 개 또는 그 이상의 TALE 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 DNA 결합 도메인을 포함할 수 있다. 에너지 민감성 단백질 또는 이의 단편은 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너에 결합할 수 있다. 폴리펩티드는 또한 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 변이체 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인을 포함할 수 있으며, 여기서 에너지 민감성 단백질 또는 이의 단편은 에너지원에 의한 유도시 상호작용 파트너에 결합할 수 있다. 상기 방법은 또한 에너지원을 적용하고, 게놈 유전자좌의 발현이 활성화되는 것을 측정하는 것을 포함할 수 있다. 본 발명의 바람직한 구현예에서 게놈 유전자좌는 세포 내에 있을 수 있다.

[0042] 본 발명의 또 다른 바람직한 구현예에서, 유도 효과기는 광 유도 전사 효과기 (LITE) 일 수 있다. LITE 시스템의 모듈성은 전사 모듈화에 임의의 효과기 도메인이 이용될 수 있게 한다.

[0043] 본 발명의 또 다른 바람직한 구현예에서, 유도 효과기는 화학 물질일 수 있다.

[0044] 본 발명은 또한 CRISPR (군집화 규칙적 공간화 단길이 회귀성 반복물 (clustered regularly interspaced short palindromic repeats))/Cas 시스템을 사용하는 유도 멀티플렉스 게놈 조작을 고려한다.

- [0045] 본 발명은 또한 본 발명의 폴리펩티드를 인코딩하는 핵산을 포함한다. 핵산은 프로모터, 유리하게는 인간 시냅신 I 프로모터 (hSyn) 를 포함할 수 있다. 특히 유리한 구현예에서, 핵산은 아데노 관련 바이러스 벡터 (AAV) 내에 패키징될 수 있다.
- [0046] 본 발명은 추가로 또한, 본원에 기재한 방법 및 조성물을 포함하는 치료 방법 또는 치료요법에 관한 것이다.
- [0047] 따라서, 본 발명 내에 임의의 이전에 공지된 생성물, 상기 생성물 제조 공정, 또는 상기 생성물 사용 방법을 포함하지 않아, 출원인이 권한을 남겨 둠으로써 임의의 이전에 공지된 생성물, 공정 및 방법의 권리불요구를 개시하는 것이 본 발명의 목적이다. 본 발명이 본 발명의 범주 내에, USPTO (35 U.S.C. § 112, 첫 번째 단락) 또는 EPO (EPC 의 제 83 조) 의 서면 설명 및 실시가능 필요사항을 충족하지 않는 임의의 생성물, 상기 생성물 제조 공정 또는 상기 생성물 사용 방법을 포함하는 것으로 의도되지 않아, 출원인이 권한을 남겨 둠으로써 이전에 기재된 생성물, 상기 생성물 제조 공정 또는 상기 생성물 사용 방법의 권리불요구를 개시하는 것에 또한 주목한다.
- [0048] 이 개시물 및 특히 특허청구범위 및/또는 단락에서, 용어 예컨대 "포함하다", "포함된", "포함하는" 등은 미국 특허법에 부여된 의미를 가질 수 있다는 것에 주목한다; 예를 들어 이는 "함유하다", "함유된", "함유하는" 등을 의미할 수 있고; 용어 예컨대 "~로 본질적으로 이루어지는" 및 "~로 본질적으로 이루어진다" 는 미국 특허법에서 기인한 의미를 갖고, 예를 들어 요소가 명백하게 인용되지 않게 하나, 선행 기술에서 발견되거나 본 발명의 기본적으로 신규한 특징에 영향을 주는 요소는 배제한다.
- [0049] 구현예 및 기타 구현예가 하기 상세한 설명으로부터 개시되거나 명백하고 그에 의해 포함된다.

**도면의 간단한 설명**

[0050] 예로써 주어졌으나 기재한 특정 구현예에만 본 발명을 제한하는 것으로 의도되지 않는 하기의 상세한 설명은 동반하는 도면과 함께 가장 최상으로 이해될 수 있다.

도 1 은 공간적 및 시간적 정확도의 필요성을 나타내는 모식도를 나타낸다.

도 2 는 전사 활성화제 유사 효과기 (TALE) 를 나타낸다. TALE 는 그의 서열의 코어에서 34 개 aa 반복물로 이루어진다. 각각의 반복물은 TALE 에 의해 결합되는 표적 DNA 에서의 염기에 상응한다. 반복물은 위치 12 및 13 에서 2 개의 가변적인 아미노산에 의해서만 상이하다. 이러한 대응의 코드가 설명된 바 있으며 (Boch, J et al. , Science, 2009 및 Moscou, M et al. , Science, 2009) 이 도면에서 나타낸다. 출원인은 이러한 코드를 통합하며 게놈 내에 선택 서열을 결합시킬 수 있는 설계자 TALE 의 합성 방법을 개발하였다 (Zhang, F et al. , Nature Biotechnology, 2011).

도 3 은 LITE: TALE/크립토크롬 전사 활성화의 설계를 나타낸다. 각각의 LITE 는 CRY2 에 융합된 TALE 및 VP64 (전사 인자) 에 융합된 크립토크롬 결합 파트너 CIB1 을 포함할 수 있는 2-성분 시스템이다. 불활성 상태에서, TALE 은 이의 융합된 CRY2 도메인을 관심 유전자의 프로모터 부위에 국부화시킨다. 이때, CIB1 은 CRY2 에 결합할 수 없고, CIB1-VP64 가 핵 공간 내에서 미결합된 채로 있다. 488 nm (청색) 광으로 자극 시, CRY2 는 형태적 변화를 거쳐, 이의 CIB1 결합 부위가 밝혀진다 (Liu, H et al. , Science, 2008). CIB1 의 신속한 결합은 융합된 VP64 도메인의 영입을 초래하여, 표적 유전자의 전사를 유도한다.

도 4 는 LITE 활성화에 대한 크립토크롬 이량체 절두의 효과를 나타낸다. CRY2 및 CIB1 의 활성을 변경시키는 것으로 알려져 있는 절두를 (Kennedy M et al, Nature Methods 2010) 전체 길이 단백질에 대해 비교하였다. Neurog2 의 프로모터에 대해 표적화된 LITE 를 도메인의 각 조합에 대해 Neuro-2a 세포에서 시험하였다. 488 nm 광으로 자극한 후, Neurog2 의 전사 수준을 자극된 샘플 및 미자극된 샘플에 대해 qPCR 을 사용하여 정량하였다.

도 5 는 KLF4 LITE 의 광-강도 의존적 반응을 나타낸다.

도 6 은 Neurog2 LITE 의 활성화 반응속도 및 Neurog2 LITE 의 불활성화 반응속도를 나타낸다.

도 7A 는 출원인의 RVD 스크리닝 시스템을 사용하여 측정된 바와 같은 다양한 RVD 의 염기-선택도를 나타낸다.

도 7B 는 출원인의 RVD 스크리닝 시스템을 사용하여 측정된 바와 같은 추가적인 RVD 의 염기-선택도를 나타낸다.

도 8A-D 는 다음을 나타낸다: (a) 잔토모나스 종 (Xanthomonas sp.) 에서 유래한 TALE 의 자연적 구조. 각

각의 DNA-결합 모듈은 34 개의 아미노산으로 이루어지는데, 각각의 반복물의 12 번째 및 13 번째 아미노산 위치에서의 RVD 는 부호 NG = T, HD = C, NI = A 및 NN = G 또는 A 에 따라 표적화되는 DNA 염기를 명시한다. DNA-결합 모듈은 비반복적 N 및 C 말단에 측면 위치하여, 전좌, 핵 국부화 (NLS) 및 전사 활성화 (AD) 도메인을 수반한다. N 말단 내 암호성 신호는 표적 부위의 제 1 염기로서 티민을 명시한다. (b) TALE 툴박스는 커스텀 TALE-TF 및 TALEN 의 신속하고 값싼 구성을 가능하게 한다. 키트는 총 12 개 플라스미드로 이루어진다: 4 개 단량체 플라스미드가 PCR 증폭용 주형으로서 사용되고, 4 개의 TALE-TF 및 4 개의 TALEN 클로닝 백본은 0.5 반복물에 의해 표적화된 4 개의 상이한 염기에 상응한다. CMV, 사이토메갈로바이러스 프로모터; N 말단, Hax3 TALE 로부터의 비반복적 N 말단; C 말단, Hax3 TALE 로부터의 비반복적 C 말단; BsaI, 커스텀 TALE DNA-결합 도메인의 삽입에 사용된 유형 II 제한 부위; ccdB + CmR, ccdB 음성 선택 유전자 및 클로람페니콜 저항성 유전자를 포함하는 음성 선택 카세트; NLS, 핵 국부화 신호; VP64, 단순 포진 바이러스의 VP 16 단백질에서 유래한 합성 전사 활성제; 2A, 2A 자체-절단 링커; EGFP, 증강된 녹색 형광 단백질; 폴리A 신호, 폴리아데닐화 신호; FokI, FokI 엔도뉴클레아제로부터의 촉매적 도메인. (c) TALE 은 커스텀 TALE-TF 를 생성시키고 게놈으로부터의 내인성 유전자의 전사를 조정하는데 사용될 수 있다. TALE DNA-결합 도메인은 합성 VP64 전사 활성제에 융합되는데, 이는 RNA 폴리머라아제 및 전사 개시에 필요한 기타 인자를 영입한다. (d) TALEN 은 비상동 쌍 또는 상동성 유도 쌍을 통해 게놈 편집을 촉진시키기 위해 위치-특이적 이중가닥 과단점을 생성시키는데 사용될 수 있다. 2 개의 TALEN 은 측면 위치 16-bp 스페이스의 결합 부위 쌍을 표적화한다. 좌측 및 우측 TALEN 은 각각 표적 부위의 상부 및 하부 가닥을 인지한다. 각각의 TALE DNA-결합 도메인은 FokI 엔도뉴클레아제의 촉매적 도메인에 융합되고; FokI 이 이량체화될 때, 이는 좌측 및 우측 TALEN-결합 도메인 사이의 부위에서 DNA 를 절단한다.

**도 9A-F** 는 단량체 서열 (위치 12 및 13 에서의 RVD 제외) 및 특정 서열을 갖는 단량체가 발생하는 빈도를 열거하는 표를 나타낸다.

**도 10** 은 TALE 활성화에 대한 비-RVD 아미노산의 효과 비교를 나타낸다.

**도 11** 은 VP64, p65 와 VP16 사이의 활성화 수준을 비교하는 활성화제 스크린을 나타낸다.

**도 12A-D** 는 TALE 전사 억제제 구조의 개발을 나타낸다. (a) TALE 억제제 스크리닝에 대한 SOX2 TALE 의 설계. 인간 게놈의 SOX2 유전자와 내의 14bp 서열을 표적화하는 TALE 을 합성하였다. (b) 스크리닝한 모든 억제제 및 그의 숙주 기원의 목록 (좌측). 8 개의 상이한 후보물 억제제 도메인은 SOX2 TALE 의 C-말단에 융합되었다. (c) 내인성 SOX2 mRNA 의 감소 배수를 모의 트랜스펙션 세포에서의 SOX2 mRNA 수준을, 각각의 후보 TALE 억제제로 트랜스펙션시킨 세포에서의 SOX2 mRNA 수준에 의해 나누어, qRT-PCR 을 사용하여 측정한다. (d) 내인성 CACNA1C 의 전사 억제. G-표적화 RVD 로서 NN, NK 및 NH 를 사용하는 TALE 을 구축하여 인간 CACNA1C 유전자와 내 18 bp 표적 부위를 표적화하였다. 각각의 TALE 을 SID 억제 도메인에 융합시킨다. NLS, 핵 국부화 신호; KRAB, 크루펠 (Kruppel)-관련 박스; SID, mSin 상호작용 도메인. 모든 결과를 HEK 293FT 세포에서 3 개의 독립적 실험으로부터 수집하였다. 오류 막대는 s.e.m. 을 나타내고; n = 3 이다. \* p < 0.05, 스튜던트 (Student) t 시험.

**도 13A-C** 는 SID 및 SID4X 를 사용하는 TALE 전사 억제제 구조의 최적화를 나타낸다. (a) TALE 억제제 구조의 시험을 위한 p11 TALE 의 설계. 마우스 (Mus musculus) 게놈의 p11 (s100a10) 유전자와 내의 20bp 서열을 표적화하는 TALE 을 합성하였다. (b) 내인성 마우스 p11 mRNA 의 전사 억제. 야생형 TALE 구조의 2 개 상이한 절두부를 갖는 마우스 p11 유전자좌를 표적화하는 TALE 을 x-축에 나타낸 바와 같이 상이한 억제제 도메인에 융합시켰다. 괄호 안의 값은 DNA 결합 반복물에 측면 위치하며 상기 구축물에서 사용한 억제제 도메인이 그 뒤에 있는 TALE DNA 결합 도메인의 N-말단 및 C-말단에서의 아미노산 수를 나타낸다. 내인성 p11 mRNA 수준을 qRT-PCR 을 사용하여 측정하고, GFP-인코딩 구축물로 트랜스펙션시킨 음성 대조군 세포에서의 수준으로 정규화시켰다. (c) 내인성 마우스 p11 의 전사 억제 배수. 내인성 p11 mRNA 의 감소 배수를, 음성 대조군 GFP 구축물로 트랜스펙션시킨 세포에서의 p11 mRNA 수준을 각각의 후보 TALE 억제제로 트랜스펙션시킨 세포에서의 p11 mRNA 수준에 의해 나누어, qRT-PCR 을 사용하여 측정한다. x-축에 따른 구축물의 표지는 이전 패널에서와 동일하다. NLS, 핵 국부화 신호; SID, mSin 상호작용 도메인; SID4X, 단길이 랩티드 링커에 의해 연결된 SID 도메인의 최적화된 4-회 탠덤 (tandem) 반복물. 모든 결과는 Neuro2A 세포에서의 3 개 독립적 실험에서 수집하였다. 오류 막대는 s.e.m. 을 나타내고; n = 3 이다. \*\*\* /p< 0.001, 스튜던트 t 시험.

**도 14** 는 2 가지 상이한 유형의 TALE 구조의 비교를 나타낸다.

**도 15A-C** 는 화학적 유도 TALE ABA 유도 시스템을 나타낸다. ABI (ABA 비-민감성 1) 및 PYL (PYL 단백질: 피라박틴 저항성 (PYR)/PYR1-유사 (PYL)) 은 식물 호르몬 아브시스산 (ABA) 의 결합시 이량체화하는 하기 열거한 2 개 단백질로부터의 도메인이다. 이러한 식물 호르몬은 출원인이 출원인의 유도 TALE 시스템에서 사용한 소분자 화학 물질이다. 이러한 시스템에서, TALE DNA-결합 폴리펩티드는 ABI 도메인에 융합되는 반면, VP64 활성화 도메인 또는 SID 억제제 도메인 또는 임의의 효과기 도메인은 PYL 도메인에 연결된다. 따라서 ABA 분자의 존재에 의해 유도시, 2 개 상호작용 도메인, ABI 및 PYL 은 이량체화되고 TALE 이 효과기 도메인에 연결되게 하여 표적 유전자 발현을 조절하는데 있어서의 이의 활성을 수행할 수 있게 한다.

**도 16A-B** 는 화학적 유도 TALE 4OHT 유도 시스템을 나타낸다.

**도 17** 은 LITE 기능성에 대한 크립토크롬2 이형이량체 배향의 효과를 나타낸다.

**도 18** 은 마우스 피질 뉴런 배양액에서의 mGluR2 LITE 활성을 나타낸다.

**도 19** 는 LITE AAV 벡터를 사용하는 1 차 마우스 뉴런의 형질도입을 나타낸다.

**도 20** 은 생체내 LITE 성분의 발현을 나타낸다.

**도 21** 은 사용한 특정 NES 펩티드가 LDLASLIL 인 구축물의 향상된 설계를 나타낸다.

**도 22** 는 4OH 타목시펜의 부재 및 존재 하 Sox2 mRNA 수준을 나타낸다.

**도 23A-E** 는 스트렙토코쿠스 피오게네스 (*Streptococcus pyogenes*) SF370 으로부터의 유형 II CRISPR 유전자좌가 DNA 의 표적화된 DSB 가 촉진되도록 포유동물 세포에서 재구성될 수 있다는 것을 나타낸다. (A) NLS 를 갖는 SpCas9 및 SpRNase III 의 조작은 포유동물 핵 내로의 이입을 가능하게 한다. (B) SpCas9 및 SpRNase III 의 포유동물 발현은 EF1a 프로모터에 의해 유도되는 반면, tracrRNA 및 예비-crRNA 어레이 (DR-스페이스-DR) 는 U6 프로모터에 의해 유도된다. PAM 을 갖는 인간 EMX1 유전자좌로부터의 프로토스페이스 (청색 하이 라이트) 를 예비-crRNA 어레이에서 스페이스에 대한 주형으로서 사용한다. (C) 표적 유전자좌와 EMX1-표적화 crRNA 사이의 염기 페어링의 모식적 표현. 적색 화살표는 추정상 절단 부위를 나타낸다. (D) SpCas9-매개 삽입-결실 (indel) 에 대한 SURVEYOR 검정. (E) 187 클론 암플리콘으로부터 확인된 돌연변이화된 대립형질의 대표 서열을 나타내는 예시적 크로마토그램. 적색 대시 기호는 결실된 염기를 나타내고; 적색 염기는 삽입 또는 돌연변이를 나타낸다. 스케일 바 = 10  $\mu$ m.

**도 24A-C** 는 SpCas9 가 포유동물 세포에서의 다수 게놈 유전자좌를 표적화하도록 재프로그래밍될 수 있다는 것을 나타낸다. (A) 청색 선으로 표시한 5 개 프로토스페이스의 위치를 나타내는 인간 EMX1 유전자좌의 모식도, 상응하는 PAM 은 자홍색으로 나타냄. (B) 예비-crRNA 및 tracrRNA 의 직접 반복 (회색) 부위 사이의 하이브리드화를 나타내는 예비-crRNA:tracrRNA 복합체 (상부) 의 모식도. 키메라 RNA 설계의 모식도 (M. Jinek et al., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816 (Aug 17, 2012)) (하부). tracrRNA 서열을 적색으로 나타내고 20bp 스페이스를 청색으로 나타낸다. (C) 인간 EMX1 유전자좌에서의 5 개 프로토스페이스에서의 Cas9-매개 절단의 효능을 비교하는 SURVEYOR 검정. 각각의 프로토스페이스를 처리된 예비-crRNA:tracrRNA 복합체 (crRNA) 또는 키메라 RNA (chiRNA) 를 사용하여 표적화한다.

**도 25A-D** 는 SpCas9 특이성의 평가 및 TALEN 과의 효율 비교를 나타낸다. (A) 단일 점 돌연변이를 갖는 EMX1-표적화 키메라를 생성시켜 스페이스-프로토스페이스 미스매치의 효과를 평가하였다. (B) 상이한 돌연변이체 키메라 RNA 의 절단 효율을 비교하는 SURVEYOR 검정. (C) TALEN 표적화 EMX1 의 설계를 나타내는 모식도. (D) TALEN 및 SpCas9 의 효율을 비교하는 SURVEYOR 겔 (N = 3).

**도 26A-G** 는 상동 재조합 및 멀티플렉스 게놈 조작을 위한 Cas9 의 적용을 나타낸다. (A) RuvC I 도메인의 돌연변이는 Cas9 를 절단 효소 (SpCas9n) 로 전환시킨다. (B) EMX1-표적화 키메라 RNA 와 SpCas9 와의 동시-발현은 삽입-결실을 유발하는 반면, SpCas9n 은 그렇지 않다 (N = 3). (C) 재조합 계획의 모식적 표시. 수복 주형은 제한 부위가 EMX1 유전자좌에 삽입되도록 설계된다. 변형된 부위를 증폭하는데 사용한 프라이머를 적색 화살표로 나타낸다. (D) 제한 단편 길이 다형성 겔 분석. 화살표는 HindIII 소화에 의해 생성된 단편을 나타낸다. (E) 성공적인 재조합을 나타내는 예시적 크로마토그램. (F) SpCas9 는 EMX1 및 PVALB 를 표적화하는 2 개 스페이스를 포함하는 crRNA 어레이를 사용하여 멀티플렉스 게놈 변형을 촉진시킬 수 있다. crRNA 어레이의 설계를 나타내는 모식도 (상부). 스페이스 모두 효율적인 프로토스페이스 절단을 매개한다 (하부). (G) SpCas9 를 사용하여 정확한 게놈 결실을 수행할 수 있다. EMX1 을 표적화하는 2

개 스페이서 (상부) 는 118bp 게놈 결실을 매개하였다 (하부).

**도 27** 은 유형 II CRISPR-매개 DNA 이중가닥 파단의 모식도를 나타낸다. 스트렙토코쿠스 피오게네스 SF370 으로부터의 유형 II CRISPR 유전자좌는 4 개 유전자, Cas9, Cas1, Cas2 및 Csn1 의 클러스터 뿐 아니라 2 개의 비-코딩 RNA 요소, tracrRNA, 및 비-반복 서열의 단길이 스트레치 (스페이서, 각각 30bp) 에 의해 공간이 남겨진 반복 서열의 특징적 어레이 (직접 반복물) 을 포함한다 (15-18, 30, 31). 각각의 스페이서는 통상 외래 유전적 물질 (프로토스페이서) 에서 유래하고, CRISPR-매개 핵산 절단의 특이성을 유도한다. 표적 핵산에서, 각각의 프로토스페이서는 그의 인지가 개별적 CRISPR 시스템에 특이적인 프로토스페이서 인접 모티프 (PAM) 와 연관된다 (22, 23). 유형 II CRISPR 시스템은 순차적 단계로 표적화 DNA 이중가닥 파단 (DSB) 을 실행한다 (M. Jinek et al, Science 337, 816 (Aug 17, 2012); Gasiunas, R. et al. Proc Natl Acad Sci USA 109, E2579 (Sep 25, 2012); J. E. Garneau et al, Nature 468, 67 (Nov 4, 2010); R. Saprunauskas et al, Nucleic Acids Res 39, 9275 (Nov, 2011); A. H. Magadan et al. PLoS One 7, e40913 (2012)). 먼저, 예비-crRNA 어레이 및 tracrRNA 를 CRISPR 유전자좌로부터 전사한다. 두 번째로, tracrRNA 를 예비-crRNA 의 직접 반복물에 하이브리드화하고 이중체로서 Cas9 와 연관시킨다 (예비-crRNA 의, 개별적인, 절두된 스페이서 서열을 포함하는 성숙 crRNA 로의 처리를 매개함). 세 번째로, 성숙 crRNA:tracrRNA 이중체는 Cas9 를 crRNA 의 스페이서 부위와 프로토스페이서 DNA 사이의 이형이중체 형성을 통해 프로토스페이서 및 필수적 PAM 으로 이루어지는 DNA 표적으로 유도한다. 마지막으로, Cas9 는 PAM 의 표적 DNA 업스트림의 절단을 매개하여 프로토스페이서 내에 DSB 를 생성시킨다.

**도 28A-C** 는 Cas9-매개 유전자 표적화를 위한 상이한 tracrRNA 전사체의 비교를 나타낸다. (A) 시험한 2 개 (단길이 및 장길이) tracrRNA 전사체의 설계 및 서열을 나타내는 모식도. 각각의 전사체는 U6 프로모터에 의해 주도된다. 전사 시작 부위를 +1 로 표기하고 전사 터미네이터는 표시한 바와 같다. 청색 선은 그의 역-상보 서열이 tracrRNA 검출용 노던 블롯 프로브 생성에 사용되는 부위를 나타낸다. (B) EMX1 유전자좌의 hSpCas9-매개 절단의 효율을 비교하는 SURVEYOR 검정. 2 개의 생물학적 레플리카를 각각의 tracrRNA 전사체에 대해 나타낸다. (C) 장길이 또는 단길이 tracrRNA 뿐 아니라 SpCas9 및 DR-EMX1(1)-DR 을 운반하는 U6 발현 구축물로 트랜스펙션된 293FT 세포로부터 추출한 총 RNA 의 노던 블롯 분석. 좌측 및 우측 패널은 각각 SpRNase III 의 부재 또는 존재 하에 트랜스펙션된 293FT 세포로부터의 것이다. U6 은 인간 U6 snRNA 를 표적화하는 프로브로 블롯팅된 로딩 대조군을 나타낸다. 단길이 tracrRNA 발현 구축물의 트랜스펙션은 tracrRNA 의 처리 형태의 과잉 수준을 이끌어낸다 (~75bp) (E. Deltcheva et al., Nature 471, 602 (Mar 31, 2011)). 매우 소량의 장길이 tracrRNA 를 노던 블롯에서 검출한다. 이들 실험의 결과로서, 출원인은 포유동물 세포에서의 적용을 위해 단길이 tracrRNA 사용을 선택하였다.

**도 29** 는 이중가닥 파단-유도 미세 삽입 및 결실의 검출을 위한 SURVEYOR 검정을 나타낸다 (D. Y. Guschin et al. Methods Mol Biol 649, 247 (2010)). Cas9-매개 절단 효율을 측정하는데 사용한 SURVEYOR 검정의 모식도. 먼저, 게놈 PCR (gPCR) 을 사용하여 변형 및 미변형 세포의 불균일 집단으로부터의 Cas9 표적 부위를 증폭하고, gPCR 생성물을 천천히 재어닐링하여 이형이중체를 생성한다. 재어닐링한 이형이중체를 SURVEYOR 뉴클레아제에 의해 절단하는 반면, 동종이중체는 미손상인 채로 둔다. Cas9-매개 절단 효율 (% 삽입-결실) 을 절단된 DNA 분획물을 기준으로 계산한다.

**도 30A-B** 는 포유동물 세포에서의 crRNA 처리의 노던 블롯 분석을 나타낸다. (A) 2 개 직접 반복물 (DR-EMX1(1)-DR) 에 의해 측면 위치한 단일 스페이서에 대한 발현 벡터를 나타내는 모식도. 인간 EMX1 유전자좌 프로토스페이서 1 을 표적화하는 30bp 스페이서 (표 1) 를 청색으로 나타내고 직접 반복물을 회색으로 나타낸다. 주황색 선은 그의 역-상보 서열을 사용하여 EMX1(1) crRNA 검출용 노던 블롯 프로브를 생성시키는 부위를 나타낸다. (B) DR-EMX1(1)-DR 을 운반하는 U6 발현 구축물로 트랜스펙션된 293FT 세포로부터 추출한 총 RNA 의 노던 블롯 분석. 좌측 및 우측 패널은 각각 SpRNase III 의 부재 또는 존재 하에 트랜스펙션된 293FT 세포로부터의 것이다. DR-EMX1(1)-DR 은 SpCas9 및 단길이 tracrRNA 의 존재 하에서만 성숙 crRNA 로 처리되었으며, SpRNase III 의 존재에 의해서는 좌우되지 않았다. 트랜스펙션된 293FT 총 RNA 로부터 검출된 성숙 crRNA 는 ~33bp 이고 S. 피오게네스 (S. pyogenes) 로부터의 39-42bp 성숙 crRNA 보다 더 짧은데 (E. Deltcheva et al., Nature 471, 602 (Mar 31, 2011)), 이는 인간 293FT 세포에서의 처리된 성숙 crRNA 가 S. 피오게네스에서의 박테리아 성숙 crRNA 와 상이할 것으로 예상된다는 것을 제시한다.

**도 31A-B** 는 Cas9 를 갖는 키메라 crRNA 또는 예비-crRNA 어레이에 대한 바이스트로닉 발현 벡터를 나타낸다. (A) 예비-crRNA 어레이에 대한 발현 벡터의 설계를 나타내는 모식도. 스페이서는 어닐링된 올리고뉴클레오타이드를 사용하여 2 개 BbsI 부위 사이에 삽입될 수 있다. 올리고뉴클레오타이드에 대한 서열 설계를 하기에

나타내며, 적절한 라이게이션 어댑터를 함께 표시하였다. (B) 키메라 crRNA 에 대한 발현 벡터의 모식도. 안내 서열은 어닐링된 올리고뉴클레오티드를 사용하여 2 개 BbsI 부위 사이에 삽입될 수 있다. 벡터는 이미 부분적 직접 반복물 (회색) 및 부분적 tracrRNA (적색) 서열을 포함한다. WPRE, 마뚝 (Woodchuck) 간염 바이러스 전사후 조절 요소.

**도 32A-B** 는 인간 PVALB 및 마우스 Th 유전자좌에서의 프로토스페이서의 선택을 나타낸다. 각각 PVALB 및 Th 유전자의 마지막 엑손 내의 3 개 프로토스페이서의 위치 및 인간 PVALB (A) 및 마우스 Th (B) 유전자좌의 모식도. 30bp 프로토스페이서를 흑색 선으로 나타내고 인접한 PAM 서열을 자홍색 막대로 나타낸다. 센스 및 안티-센스 가닥 상의 프로토스페이서를 각각 DNA 서열 위 및 아래에 나타낸다.

**도 33A-C** 는 인간 게놈에서의 PAM 서열의 발생을 나타낸다. 인간 게놈에서의 인접한 스트렙토코쿠스 피오케네스 SF370 유전자좌 1 PAM (NGG) (A) 와 스트렙토코쿠스 써모필루스 (*Streptococcus thermophilus*) LMD9 유전자좌 1 PAM (NNAGAAW) (B) 사이 거리의 히스토그램. (C) 염색체에 의한 각각의 PAM 에 대한 거리. Chr, 염색체. 인간 염색체 서열의 플러스 및 마이너스 가닥 모두를 사용하여 추정상 표적을 확인하였다.

일부 프로토스페이서 표적에서 절단 활성을 제한할 수 있는 크로마틴, DNA 메틸화-, RNA 구조 및 기타 인자를 고려하면, 실제 표적화 능력이 이러한 컴퓨터적 분석 결과보다 덜 할 수 있다는 것에 주목하는 것이 중요하다.

**도 34A-D** 는 스트렙토코쿠스 써모필루스 LMD-9 로부터의 유형 II CRISPR 가 또한 진핵세포에서 기능할 수 있다는 것을 나타낸다. (A) 스트렙토코쿠스 써모필루스 LMD-9 로부터의 CRISPR 유전자좌 2 의 모식도. (B) S. 써모필루스 (*S. thermophilus*) CRISPR 시스템에 대한 발현 시스템의 설계. 항시성 EF1a 프로모터를 사용하여 인간 코돈-최적화 hStCas9 를 발현시킨다. 정확한 전사 개시를 확실하게 하기 위해, U6 프로모터를 사용하여 tracrRNA 및 crRNA 의 성숙한 버전을 발현시킨다. 성숙 crRNA 및 tracrRNA 에 대한 서열을 나타낸다. crRNA 서열에서 소문자 "a" 로 나타낸 단일 염기를 사용하여, RNA Pol III 전사 터미네이터로서 역할하는 폴리U 서열을 제거하였다. (C) 인간 EMX1 유전자좌에서의 프로토스페이서 및 상응하는 PAM 서열 표적을 나타내는 모식도. 2 개의 프로토스페이서 서열을 하이라이트 표시하고 NNAGAAW 모티프를 만족시키는 그의 상응하는 PAM 서열을 자홍색 선으로 표시한다. 프로토스페이서 모두는 안티-센스 가닥을 표적화한다. (D) 표적 유전자좌에서의 StCas9-매개 절단을 나타내는 SURVEYOR 검정. RNA 안내 스페이서 1 및 2 는 각각 14% 및 6.4% 유도되었다. 이들 2 개 프로토스페이서 부위에서의 생물학적 레플리카 전체에 걸친 절단 활성의 통계적 분석을 표 1 에서 발견할 수 있다.

**도 35** 는 AAV-프로모터-TALE-효과기 구축물의 예를 나타내며, 여기서 hSyn = 인간 시냅신 1 프로모터, N+136 = TALE N-말단, AA +136 절두, C63 = TALE C-말단, AA +63 절두, vp = VP64 효과기 도메인, GFP = 녹색 형광 단백질, WPRE = 마뚝 간염 바이러스 전사후 조절 요소, bGH = 소 성장 호르몬 폴리A, ITR = AAV 역전 말단 반복물 및 AmpR = 암피실린 저항성 유전자이다.

**도 36A-C** 는 LITE 시스템의 설계 및 최적화를 나타낸다. (a) TALE DNA-결합 도메인은 CRY2 에 융합되고 전사 효과기 도메인은 CIB1 에 융합된다. 불활성 상태에서, TALE-CRY2 는 표적 유전자의 프로모터 부위에 결합하는 한편, CIB1-효과기는 핵 내에서 결합하지 않은 채로 남아있다. VP64 전사 활성제는 상기에 나타낸다. 청색으로 조명시, TALE-CRY2 및 CIB1-효과기는 빠르게 이량체화하고, CIB1-효과기를 표적 프로모터에 영입한다. 상기 효과기는 표적 유전자의 전사를 조정한다. (b) 그의 광-민감성 결합 파트너의 기능적 절두를 포함하는 LITE 를 갖는 내인성 표적 Neurog2 mRNA 의 광-의존적 상향조절. LITE-트랜스펙션된 Neuro-2a 세포를 7% 의 사용률로 5 mW/cm<sup>2</sup> 의 강도에서 466 nm 광으로 24 시간 동안 자극하였다 (0.066 Hz 에서의 1 초 펄스). (c) TALE-CRY2PHR 및 CIB1-VP64 LITE 에 의한 광-의존적 Neurog2 상향조절의 시간 경과.

LITE-트랜스펙션된 Neuro-2a 세포를 7% 의 사용률로 5 mW/cm<sup>2</sup> 의 강도에서 466 nm 광으로 자극하였고 (0.066 Hz 에서 1 초 펄스), Neurog2 mRNA 수준은 광 자극 6 시간 후 감소한다. 모든 Neurog2 mRNA 수준을 발현하는 GFP 대조군 세포에 대해 측정하였다 (평균 ± s.e.m.; n = 3-4) (\*, p < 0.05; 및 \*\*\*, p < 0.001).

**도 37A-F** 는 뉴런에서의 내인성 유전자좌를 표적화하는 시험관내 및 생체내 AAV-매개 TALE 전달을 나타낸다.

(a) AAV 내로 패키징된 항시성 TALE 전사 활성제의 일반적 모식도. 효과기 도메인 VP64 를 하이라이트 표시하였다. hSyn: 인간 시냅신 프로모터; 2A: 구제역-유래 2A 썩티드; WPRE: 마뚝 간염 전사후 반응 요소; bGH pA: 소 성장 호르몬 폴리-A 신호, (b) 1 차 피질 뉴런에서의 (a) 로부터의 AAV-TALE-VP64 구축물로의 형질 도입을 나타내는 대표 이미지. 세포를 GFP 및 뉴런 마커 NeuN 에 대해 염색하였다. 스케일 바 = 25 μm.

(c) 다양한 내인성 유전자좌를 표적화하는 AAV-TALE-VP64 구축물을 1 차 피질 뉴런에서 전사 활성화에 대해 스크리닝하였다 (\*, p < 0.05; \*\*, p < 0.01 ; p < 0.001). (d) AAV 에 의한 TALE-VP64 의 마우스 ILC 로

의 효율적 전달. 스케일 바 = 100  $\mu\text{m}$ . (Cg1 = 대상 피질, PLC = 변연전 피질, ILC = 변연계 아래 피질), (e) ILC 에서의 뉴런의 효율적 형질도입의 고배율 이미지. (f) ILC 에서의 생체내 TALE-VP64 에 의한 Grm2 mRNA 상향조절 (평균  $\pm$  s.e.m.; n = 조건 당 3 마리 동물), 300  $\mu\text{m}$  조직 편치를 사용하여 측정함.

**도 38A-I** 은 1 차 뉴런 및 생체내에서의 내인성 전사의 LITE-매개 광유전학적 조절을 나타낸다. (a) 교환된 CRY2PHR 및 CIB1 구조를 갖는 AAV-LITE 활성화제 구축물, (b) 1 차 뉴런에서의 AAV-전달 LITE 구축물의 동시-형질도입을 나타내는 대표적 이미지. GFP, HA-tag 및 DAPI 에 대해 세포를 염색하였다. (스케일 바 = 25  $\mu\text{m}$ ). (c) 0.8% 사용률로 466 nm 광 펄스로 (0.033Hz 에서 250 ms 펄스 또는 0.016Hz 에서 500 ms 펄스; 5  $\text{mW}/\text{cm}^2$ ) 자극 24 시간 후 1 차 뉴런에서의 Grm2 발현의 광-유도 활성화. (d) 4 시간 및 24 시간점에서의 광 자극의 존재 및 부재 하 1 차 피질 뉴런에서의 Grm2 mRNA 의 상향 조절. 발현 수준을 GFP 로만 형질도입된 뉴런에 대해 나타낸다. (e) GFP 단독 대조군 형질도입, LITE 를 갖는 미자극 뉴런, 및 LITE 를 갖는 광-자극 뉴런에서의 mGluR2 단백질 수준의 정량. 대표적인 웨스턴 블롯을 로딩 대조군으로서  $\beta$ -튜불린-III 을 사용하여 나타낸다. (f) LITE 시스템, 광학 섬유 이식물, 및 조직 단리에 사용한 0.35 mm 직경 뇌 편치를 사용한 형질도입을 나타내는 모식도. (g) 두 LITE 성분 모두로 동시-형질도입된 ILC 의 대표적 이미지. 염색을 HA-tag (적색), GFP (녹색) 및 DAPI (청색) 로 나타낸다. (스케일 바 = 25  $\mu\text{m}$ ). (h) ILC 로 형질도입된 LITE 를 사용하는 내인성 Grm2 발현의 광-유도 활성화. \*\*,  $p < 0.05$ ; 각 실험 조건에 대해 4 마리의 상이한 마우스로부터 데이터를 생성시켰다. (i) LITE 1.0 및 최적화 LITE 2.0 을 사용하는 Neurog2 발현의 광 유도 및 증가 배수. LITE2.0 은 높은 수준의 활성화를 유지하면서 최소의 배경을 제공한다. NLS<sub>α-임포틴</sub> 및 NLS<sub>sv40</sub>, 각각-임포틴 및 원숭이 바이러스 40 으로부터의 핵 국부화 신호; GS, Gly-Ser 링커; NLS\*, 표시한 잔기가 Ala 로 치환되어 핵 국부화 활성을 방지하는 돌연변이화된 NLS;  $\Delta$ 318-334; 고등 식물 헬릭스-루프-헬릭스 전사 인자 상동성 부위의 결실.

**도 39A-H** 는 TALE- 및 LITE-매개 후생적 변형을 나타낸다. (a) LITE 후생적 변형제 (epiLITE) 의 모식도. (b) AAV 벡터 내의 조작된 후생적 전사 억제제 SID4X 의 모식도. 효율적인 AAV 패키징을 확실히 하기 위해 GFP (800bp) 보다는 phiLOV2.1 (330bp) 을 형광 마커로서 사용하였다. (c) 광 자극의 존재 및 부재 하 1 차 피질 뉴런에서의 내인성 Grm2 발현의 epiLITE-매개 억제. GFP 단독으로 형질도입된 뉴런에 대해 하향 조절 배수를 나타낸다. (d) 광-자극의 존재 및 부재 하 Grm2 프로모터에서의 H3K9 히스톤 잔기 아세틸화에서의 epiLITE-매개 감소. (e, f) epiTALE-메틸트랜스퍼라아제 (epiTALE-KYP, -TgSET8 및 -NUE) 에 의한 Grm2 mRNA 의 감소 배수, 및 Grm2 프로모터에서의 히스톤 메틸화 마크 H3K9me1, H4K20me3 및 H3K27me3 의 상응하는 풍족. (g, h) epiTALE 히스톤 데아세틸라아제 (epiTALE-HDAC8, -RPD3, -Sir2a 및 -Sin3a) 에 의한 Grm2 mRNA 의 감소 배수, 및 Grm2 프로모터에서의 히스톤 잔기 아세틸화 마크 H4K8Ac 및 H3K9Ac 에 있어서의 상응하는 감소. 모든 패널에서 나타낸 값은 평균  $\pm$  s.e.m., n = 3-4 이다.

**도 40** 은 시험관내 CRY2 의 흡수 스펙트럼의 설명을 나타낸다. 크립토크롬 2 를 350-475 nm 광에 의해 최적 으로 활성화시켰다<sup>1</sup>. 흡수 및 활성화에서의 급격한 감소가 480 nm 초과에 대해 나타났다. 스펙트럼을 [Banerjee, R. et al. The Signaling State of Arabidopsis Cryptochrome 2 Contains Flavin Semiquinone. Journal of Biological Chemistry 282, 14916-14922, doi: 10.1074/jbc.M700616200 (2007)] 로 부터 조정하였다.

**도 41** 은 LITE-매개 유전자 발현에 대한 조명 사용률의 효과를 나타낸다. 가변적인 사용률 (총 시간의 % 로서의 조명) 을 사용하여 KLF4 유전자를 표적화하는 LITE 를 발현하는 293FT 세포를 자극하여, LITE 활성화에 대한 사용률의 효과를 조사하였다. KLF4 발현 수준을 GFP 만을 발현하는 세포와 비교하였다. 자극 매개변수는 다음과 같다: 466 nm, 5  $\text{mW}/\text{cm}^2$ , 24 시간 동안. 펄스를 0.067 Hz 에서 하기 지속 기간으로 수행하였다: 1.7% = 0.25 초 펄스, 7% = 1 초 펄스, 27% = 4 초 펄스, 100% = 연속 조명. (평균  $\pm$  s.e.m.; n = 3-4).

**도 42A-B** 는 LITE-매개 유전자 발현 및 세포 생존에 대한 광 강도의 효과를 나타낸다. (a) CRY2PHR :: CIB1 LITE 의 전사 활성화는 466 nm 청색광의 강도에 따라 가변적이라는 것이 발견되었다. Neuro 2a 세포를 24 시간 동안 7% 사용률 (0.066Hz 에서 1 초 펄스) 에서 자극하였다. (b) 적색-형광 에티디움 동중이량체-1 에 대해 양성인 세포 대 칼세인-양성 세포의 % 로서 광-유도 독성을 측정하였다. 모든 Neurog2 mRNA 수준을 GFP 만을 발현하는 세포에 대해 측정하였다 (평균  $\pm$  s.e.m.; n = 3-4).

**도 43** 은 LITE-매개 유전자 발현에 대한 전사 활성화 도메인의 영향을 나타낸다. 상이한 전사 활성화 도메

인 (VP16, VP64 및 p65) 을 사용하는 LITE 에 의한 광의 존재 및 부재 하 Neurog2 상향조절. LITE 로 트랜스펙션한 Neuro-2a 세포를 24 시간 동안 466 nm 광으로 5 mW/cm<sup>2</sup> 의 강도 및 7% 의 사용률에서 (0.066 Hz 에서 1 초 펄스) 자극하였다. (평균 ± s.e.m.; n = 3-4)

**도 44A-C** 는 내인성 유전자 전사의 화학적 유도를 나타낸다. (a) 아브시스산 (ABA) 수용체 시스템을 기반으로 하는 화학 물질 유도 2 개 하이브리드 TALE 시스템의 설계를 나타내는 모식도. ABI 및 PYL 은 ABA 첨가 시 이량체화하며 ABA 철회시 해리된다. (b) ABA-의존적 Neurog2 상향조절의 시간 경과. 250 μM 의 ABA 를 TALE(Neurog2)-ABI 및 PYL-VP64 를 발현하는 HEK 293FT 세포에 첨가하였다. ABA 첨가 후 나타낸 시점에서 mRNA 증가 배수를 측정하였다. (c) ABA 자극 24 시간 후 Neurog2 mRNA 수준의 감소. 모든 Neurog2 mRNA 수준을 발현하는 GFP 대조군 세포에 대해 측정하였다 (평균 ± s.e.m.; n = 3-4).

**도 45A-C** 는 AAV 상정액 생성을 나타낸다. (a) 렌티바이러스 및 AAV 벡터 (GFP 운반) 를 사용하여 형질도입 효율을 시험하였다. (b) 1 차 배아 피질 뉴런을 동일 수의 AAV 또는 렌티바이러스-트랜스펙션된 293FT 세포로부터 유래한 300 및 250 μL 상정액으로 형질도입하였다. GFP 발현의 대표적 이미지를 7 d.p.i. 에서 수집하였다. 스케일 바 = 50 μm. (c) 나타낸 공정을 AAV 상정액의 생성 및 1 차 뉴런의 후속적 형질도입을 위해 개발하였다. 293FT 세포를 관심 유전자를 운반하는 AAV 벡터, AAV1 혈청형 패키징 벡터 (pAAV1), 및 PEL 을 사용하는 헬퍼 플라스미드 (pDF6) 로 트랜스펙션하였다. 48 시간 후, 상정액을 수확하고 0.45 μm PVDF 멤브레인을 통해 여과하였다. 1 차 뉴런을 이후 상정액으로 형질도입하고 남아 있는 분취액을 -80℃ 에서 보관하였다. 5-6 일 후 AAV 구축물 발현의 안정한 수준에 도달하였다. 이러한 공정에 따른 AAV 상정액 생성을 96-웰 포맷으로 96 개 이하의 상이한 바이러스 구축물의 생성에 사용할 수 있다 (도 37C 에서 나타낸 뉴런에서의 TALE 스크린에 이용).

**도 46** 은 DNaseI-민감성 크로마틴 부위에 의해 유도된 TALE 표적 부위의 선택을 나타낸다. ENCODE (<http://genome.ucsc.edu>) 로부터의 마우스 피질 조직 데이터를 기반으로 한 높은 DNaseI 민감성을 사용하여 개방 크로마틴 부위를 확인하였다. 전사 시작 부위의 2 kb 업스트림 부위 내의 최고 진폭을 갖는 피크를 표적화를 위해 선택하였다. TALE 결합 표적을 이후 피크의 중심에서 200 bp 부위 내에서 골랐다.

**도 47** 은 1 차 뉴런 건강에 대한 광 사용률의 영향을 나타낸다. 1 차 피질 뉴런 건강에 대한 광 자극의 효과를 사용률 7%, 0.8%, 및 무광 조건에 대해 비교하였다. 칼세인을 사용하여 뉴런 생존도를 평가하였다.

명시야 이미지를 캡처하여 형태 및 세포 무결성을 나타내었다. 466 nm 광의 5 mW/cm<sup>2</sup> 로 24 시간 동안 표적 사용률로 1 차 피질 뉴런을 자극하였다. 대표적 이미지, 스케일 바 = 50 μm. 펄스를 하기의 방식으로 수행하였다: 7% 사용률 = 0.067 Hz 에서 1 초 펄스, 0.8% 사용률 = 0.0167 Hz 에서 0.5 초 펄스.

**도 48** 은 광유전학적 자극 동안의 마우스 이미지를 나타낸다. 잠에서 깬, 자유로이 거동하는, LITE-주입 마우스를 뇌정위적으로 이식한 케틀라 및 광학 섬유로 사진찍었다.

**도 49** 는 마우스 변연계 아래 피질에서의 AAV1/2 에 의한 LITE 성분의 동시-형질도입 효율을 나타낸다. TALE(Grm2)-CIB1 단독, CRY2PHR-VP64 단독에 의해 형질도입되거나 동시-형질도입된 세포를, 모든 형질도입된 세포의 % 로서 계산하였다.

**도 50** 은 기준선 전사 조절에 대한 개별적인 LITE 성분의 기여도를 나타낸다. Grm2 mRNA 수준을 개별적인 LITE 성분으로 트랜스펙션된 1 차 뉴런에서 측정하였다. Grm2 TALE 1-CIB1 단독을 발현하는 1 차 뉴런은 전체 LITE 시스템을 발현하는 미자극 세포에서와 같이 Grm2 mRNA 수준에 있어서 유사한 증가를 초래하였다 (평균 ± s.e.m.; n = 3-4).

**도 51** 은 활성화, 배경 신호 및 감소 배수에 대한 LITE 성분 조작의 효과를 나타낸다. 단백질 변형을 이용하여, LITE 성분이 감소된 배경 전사 활성화를 초래하면서 광에 의한 유도 비율을 향상시킨다는 것을 발견하였다. 단백질 변형을 하기에서 상세히 토의한다. 간략하게, 내인성 핵 이출 신호에서의 돌연변이 및 핵 국부화 신호를 사용하여 CRY2PHR-VP64 성분의 핵 이입을 향상시켰다. 핵 국부화 또는 CIB1 전사 활성화를 감소시키는 것으로 의도된 CIB1 의 여러 변이를 추구하여 TALE-CIB1 성분의 배경 활성화에 대한 기여도를 감소시켰다. 시험한 CRY2PHR-VP64 및 TALE-CIB1 의 모든 조합의 결과를 상기에 나타낸다. 막대 그래프의 좌측에 있는 표는 각 조합에 사용한 도메인/돌연변이의 특정 조합을 표시한다. 표 및 막대 그래프의 각 행은 성분 세부사항, 광/무광 활성화, 및 특정 CRY2PHR/CIB1 조합에 대한 광에 의한 유도 비율을 포함한다. LITE 1.0 에 비해 감소된 배경 및 증가된 유도 배수를 모두 초래한 조합을 "+" 표시한 표 열에서 녹색으로 하이라이트 표시한다 (t-시험 p<0.05). **CRY2PHR-VP64 구축물:** 3 개의 새로운 구축물을 CRY2PHR-VP64 핵 이입을 향상시키

는 목적으로 설계하였다. 먼저, CRY2PHR 의 예측된 내인성 핵 이출 서열 내의 돌연변이 L70A 및 L74A 를 유도하여 단백질의 핵 이출을 제한하였다 (효과기 열에서 '\*' 로 나타냄). 두 번째로, α-입포틴 핵 국부화 서열을 CRY2PHR-VP64 의 N-말단에 융합시켰다 (효과기 열에서 'A' 로 나타냄). 세 번째로, SV40 핵 국부화 서열을 CRY2PHR-VP64 의 C-말단에 융합시켰다 (효과기 열에서 'P' 로 나타냄). **TALE-CIB1 링커:** LITE 1.0 에서 사용한 TALE 과 CIB1 사이의 SV40 NLS 링커를 TALE-CIB1 단백질의 핵 이출을 증가시키도록 설계된 여러 링커 중 하나로 대체하였다 (CIB1 링커 열에서 사용한 기호를 괄호 안에 나타냄): 유연성 글리신-세린 링커 (G), 아데노바이러스 유형 5 E1B 핵 이출 서열 (W), HIV 핵 이출 서열 (M), MAPK 핵 이출 서열 (K) 및 PTK2 핵 이출 서열 (P). **NLS\* 내인성 CIB1 핵 국부화 서열 돌연변이:** 핵 국부화 신호는 야생형 CIB1 서열 내에 존재한다.

이러한 신호는 NLS\* 구축물 내에서 K92A, R93A, K105A 및 K106A 에서 돌연변이화되어 TALE-CIB1 핵 국부화가 약화되었다 (NLS\* 열에서 'N' 으로 나타냄). **△CIB1 전사 인자 상동성 결실:** 가능한 기저 CIB1 전사 활성화를 제거하려는 노력으로, 결실 구축물을 고등 식물에서의 염기성 헬릭스-루프-헬릭스 전사 인자에 대한 높은 상동성 부위가 제거되게 설계하였다. 이들 결실된 부위는 Aaa230-256, Aaa276-307, Aaa308-334 로 이루어졌다 (ACIB1 열에서 '1', '2' 및 '3' 으로 나타냄). 각각의 경우, 결실된 부위를 3 개 잔기 GGS 링크로 대체하였다.

**CIB1 내로의 NES 삽입:** TALE- CIB1 의 광-의존적 핵 이입을 촉진하기 위한 하나의 계획은 CRY2PHR 과의 이량체화 접촉면에서 CIB1 에 NES 를 삽입하여, CRY2PHR 과 결합시 신호가 숨겨질 수 있는 것이다. 이를 위해, NES 를 공지된 CRY2 상호작용 도메인 CIBN 내 상이한 위치 (aa 1-170) 에서 삽입하였다. 위치는 하기와 같다 (NES 열에서 사용한 기호를 괄호 안에 나타냄): aa28 (1), aa52 (2), aa73 (3), aa120 (4), aa140 (5), aa160 (6). **\*bHLH 염기성 헬릭스-루프-헬릭스 돌연변이:** 직접적 CIB1-DNA 상호작용을 감소시키기 위해, CIB1 에서의 염기성 헬릭스-루프-헬릭스 부위의 여러 염기성 잔기를 돌연변이시켰다. 하기 돌연변이는 모든 \*bHLH 구축물 내에 존재한다 (도 51 의 \*bHLH 컬럼에서 'B' 로 나타냄): R175A, G176A, R187A 및 R189A.

**도 52** 는 TALE-CIB1 의 광 매개 동시-의존적 핵 이입의 설명을 나타낸다. (a) 광의 부재 하, TALE-CIB1 LITE 성분은 핵 국부화 신호, NLS 의 부재 (또는 약한 핵 이출 신호, NES 의 첨가) 로 인해 세포질에서 존재한다. 한편 NLS 를 함유하는 CRY2PHR-VP64 성분은 저절로 핵에 능동적으로 이입된다. (b) 청색광의 존재 하에, TALE-CIB1 은 CRY2PHR 에 결합한다. CRY2PHR-VP64 에 존재하는 강력한 NLS 는 이제 LITE 성분 모두의 복합체의 핵 이입을 매개하여, 이들이 표적화된 유전자좌에서 전사를 활성화시킬 수 있게 한다.

**도 53** 은 주목할만한 LITE 1.9 조합을 나타낸다. LITE 2.0 구축물에 추가로, 조작된 LITE 성분 스크린으로부터의 여러 CRY2PHR-VP64 :: TALE-CIB1 조합에 특히 주목하였다. 돌연변이화된 내인성 NLS 를 갖는 α-입포틴 NLS 효과기 구축물 및 △276-307 TALE-CIB1 구축물이 조합된 LITE 1.9.0 은 9 초과의 유도 비율 및 180 초과의 절대 광 활성화를 나타내었다. 미변형된 CRY2PHR-VP64 와 돌연변이화된 NLS, △318-334, AD5 NES TALE-CIB1 구축물이 조합된 LITE 1.9.1 은, 4 의 유도 비율을 1.06 의 배경 활성화로 얻었다. 2 미만의 배경 활성화 및 7 내지 12 범위의 유도 비율을 갖는 다른 LITE 1.9 조합의 선택을 또한 하이라이트 표시하였다.

**도 54A-D** 는 뉴런에서의 TALE SID4X 억제제 분석 및 적용을 나타낸다. a) 합성적 억제제를 4 개의 SID 도메인 (SID4X) 을 연쇄 연결시켜 구축하였다. 최적의 TALE -억제제 구조를 확인하기 위해, SID 또는 SID4X 를 마우스 p11 유전자를 표적화하도록 설계된 TALE 에 융합시켰다. (b) p11 mRNA 에서의 감소 배수를 qRT-PCR 을 사용하여 검정하였다. (c) AAV 에 패키징된 항시성 TALE 전사 억제제의 일반적 모식도. hSyn: 인간 시냅신 프로모터; 2A: 구제역-유래 2 A 펩티드; WPRE: 마뫼 간염 전사후 반응 요소; bGH pA: 소 성장 호르몬 폴리-A 신호. phiLOV2.1 (330bp) 을 효율적 AAV 패키징을 확실히 하기 위한 단길이 형광 마커로서 선택하였다. (d) 내인성 마우스 유전자좌 Grm5 및 Grm2 를 표적화하는 2 개 TALE 을 SID4X 에 융합시키고 1 차 뉴런에 바이러스 형질도입하였다. SID4X 를 통한 표적 유전자 하향 조절을 GFP 만을 발현하는 뉴런에서의 수준에 대하여 각각의 TALE 에 대해 나타낸다. (평균 ± s.e.m.; n = 3-4).

**도 55A-B** 는 다양한 집합의 epiTALE 이 뉴런 및 Neuro2a 세포에서의 전사 억제를 매개한다는 것을 나타낸다. a) 상이한 히스톤 효과기 도메인에 융합된 총 24 개 Grm2 표적화 TALE 을 AAV 를 사용하여 1 차 피질 마우스 뉴런에 형질도입하였다. Grm2 mRNA 수준을, GFP 단독으로 형질도입된 뉴런에 대해 RT-qPCR 을 사용하여 측정하였다. \* 는 p < 0.05 인 억제제를 나타낸다. b) 총 32 개 epiTALE 을 Neuro2A 세포에 트랜스펙션시켰다. 이들 중 20 개는 표적화된 Neurog2 유전자좌의 유의한 억제를 매개하였다 (\* = p<0.05).

**도 56A-D** 는 Neuro 2A 세포에서의 히스톤 개질과 함께 epiTALE 매개된 전사 억제를 나타낸다. (a) Neuro 2A 세포에서 쥐과 Neurog2 유전자좌를 표적화하는 SIRT3 및 히스톤 탈아세틸화 후생적 효과기 NcoR 에 융합된 TALE 을 Neurog2 전사체 수준에 대한 억제 활성화에 대해 검정하였다. (b) ChIP RT-qPCR 은 NcoR 및 SIRT3

epiTALE 에 대한 Neurog2 프로모터에서의 H3K.9 아세틸화에 있어서의 감소를 나타낸다. (c) 공지된 히스톤 메틸트랜스퍼라아제 결합 활성을 갖는 후생적 효과기 PHF19 를 Neurog2 mRNA 수준의 Neurog2 매개 억제를 표적화하는 TALE 에 융합시켰다. (d) ChIP RT-qPCR 은 PHF19 epiTALE 에 대한 Neurog2 프로모터에서의 H3K27me3 수준에 있어서 증가를 나타낸다.

**도 57A-G** 는 RNA-안내 DNA 결합 단백질 Cas9 가 전사 효과기 도메인을 특정 게놈 유전자좌에 표적화하는데 사용될 수 있다는 것을 나타낸다. (a) 유형 II 스트렙토코쿠스 피오게네스 CRISPR/Cas 시스템으로부터의 RNA-안내 뉴클레아제 Cas9 는 2 개의 알라닌 치환 (D10A 및 H840A) 을 도입함으로써 핵산분해적으로 불활성인 RNA-안내 DNA 결합 단백질 (Cas9\*\*) 로 전환될 수 있다. 모식도는 합성적 안내 RNA (sgRNA) 가 인간 게놈에서의 특정 유전자좌에 대한 Cas9\*\*-효과기 융합을 지시할 수 있다는 것을 나타낸다. sgRNA 는 표적 서열을 명기하는 5' 말단에서 20bp 안내 서열을 포함한다. 표적 게놈 DNA 상에서, 20bp 표적 부위는 그 뒤에 5'-NGG PAM 모티프가 뒤따를 수 있다. (b, c) 각각 인간 KLF4 및 SOX2 유전자좌에서의 sgRNA 표적 부위를 나타내난 모식도. 각각의 표적 부위를 청색 막대로 표시하고 상응하는 PAM 서열을 자홍색 막대로 표시한다. (d, e) Cas9\*\*-VP64 전사 활성화제 및 SID4X-Cas9\*\* 전사 억제제 구축물의 모식도. (f, g) 각각 Cas9\*\*-VP64 및 SID4X-Cas9\*\* 매개된 KLF4 의 활성화 및 SOX2 억제. 모든 mRNA 수준을 GFP 모의 트랜스펙션된 대조군 세포에 대해 측정하였다 (평균 ± s.e.m.; n = 3).

**도 58** 은 2 개 TALE 이 각각의 내인성 마우스 유전자좌 Grm5, Grin2a 및 Grm2 를 표적화하는, 설계된 6 개 TALE 을 나타낸다. TALE 을 전사 활성화제 도메인 VP64 또는 억제제 도메인 SID4X 에 융합시키고, 1 차 뉴런에 바이러스 형질도입하였다. VP64 를 통한 표적 유전자 상향조절 및 SID4X 를 통한 하향조절 모두를, GFP 만을 발현하는 뉴런에서의 수준에 대하여 각각의 TALE 에 대해 나타낸다.

**도 59A-B** 는 하기를 나타낸다: (A) SID4X 억제제 도메인을 강조하는 LITE 억제제 구축물. (B) Grm2 T1-LITE 및 Grm2 T2-LITE 를 사용하는 1 차 피질 뉴런에서의 내인성 Grm2 발현의 광-유도 억제. 하향조절 배수를 GFP 단독으로 형질도입된 뉴런에 대해 나타낸다 (평균 ± s.e.m.; 모든 서브패널에 대해 n = 3-4).

**도 60A-B** 는 CRY2PHR 및 CIB1 성분 교환을 나타낸다. (A) TALE-CIB1::CRY2PHR-VP64 는 TALE-CRY2PHR::CIB1-VP64 보다 더 높은 수준에서 Ngn2 를 활성화시킬 수 있었다. (B) Ngn2 LITE 의 활성화 비율 (광 대 무광) 배수는 두 설계 모두에 대해 유사한 효율을 나타낸다. 자극 매개변수는 도 36B 에서 사용한 것들과 동일하였다.

**도 61** 은 유도 Cas9 에 대한 Tet Cas9 벡터 설계를 나타낸다.

**도 62** 는 Cas9 및 EGFP 의 독시사이클린 유도 후 293FT 세포에서의 EGFP 발현 및 벡터를 나타낸다.

**발명을 실시하기 위한 구체적인 내용**

- [0051] 발명의 상세한 설명
- [0052] 용어 "핵산" 또는 "핵산 서열" 은 단일- 또는 이중-가닥 형태인 데옥시리보핵산 또는 리보핵산 올리고뉴클레오티드를 의미한다. 상기 용어는 핵산, 즉 자연적 뉴클레오티드의 공지된 유사체를 포함하는 올리고뉴클레오티드를 포함한다. 상기 용어는 또한 합성 백본을 갖는 핵산-유사 구조를 포함한다: 예를 들어 Eckstein, 1991; Baserga et al, 1992; Milligan, 1993; WO 97/03211; WO 96/39154; Mata, 1997; Strauss-Soukup, 1997; 및 Samstag, 1996 을 참고한다.
- [0053] 본원에서 사용하는 바와 같이, "재조합체" 는 합성되거나 시험관내 조작된 폴리뉴클레오티드 (예를 들어, "재조합체 폴리뉴클레오티드"), 세포 또는 기타 생물학적 시스템에서 유전자 산물을 생성시키기 위해 재조합체 폴리뉴클레오티드를 사용하는 방법, 또는 재조합체 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드 ("재조합체 단백질") 를 의미한다. "재조합체 수단" 은 상이한 근원으로부터의 다양한 코딩 부위 또는 도메인 또는 프로모터 서열을 갖는 핵산을 이의 발현을 위해, 예를 들어 본 발명의 벡터 내 폴리펩티드 코딩 서열의 유도 또는 항시성 발현을 위해 발현 카세트 또는 벡터에 라이게이션하는 것을 포함한다.
- [0054] 핵산에 대해 사용하는 경우 용어 "이중" 은 상기 핵산이 자연계에서 보통 발견되지 않는 세포 또는 바이러스 내에 존재하는 것을 나타내거나; 자연계에서 보통 발견되는 바처럼 서로 동일한 관계로 발견되지 않거나, 재조합적으로 조작되어 그의 발현 수준, 또는 세포 또는 구조에서의 다른 핵산 또는 다른 분자에 대한 물리적 관계가 자연계에서 보통 발견되지 않는, 둘 이상의 하위서열을 포함한다. 이러한 맥락에서 사용된 유사한 용어는 "외인성" 이다. 예를 들어, 이중 핵산은 통상 재조합적으로 생성되어, 자연계에서 발견되지 않는 방식으로

배열된 관련없는 유전자로부터의 둘 이상의 서열을 가지며; 그 예로, 본 발명의 아데노바이러스-기반 벡터 내로 삽입된 프로모터 서열에 작동가능하게 연결된 인간 유전자를 들 수 있다. 예를 들어, 관심 이중 핵산은 면역원성 유전자 생성물을 인코딩할 수 있으며, 이때 아데노바이러스를 담체 또는 약물-백신 조성물로서 치료적 또는 예방적으로 투여한다. 이중 서열은 다양한 조합의 프로모터 및 서열을 포함할 수 있으며, 이의 예를 본원에서 상세히 기재한다.

- [0055] "치료적 리간드" 는 치료 효과를 갖는 표적 세포의 수용체에 결합할 수 있는 물질일 수 있다.
- [0056] "치료 효과" 는 임의 종류의 의학적 치료의 결과일 수 있으며, 이의 결과는 당업자에 의해 바람직하고 이로인 것인지 판단된다. "치료 효과" 는 의학적 치료에 대한 반응으로서 발생하는 거동 또는 생리적 변화일 수 있다. 그 결과는 예측된, 미예측된, 또는 심지어 의도되지 않은 의학적 치료의 결과일 수 있다. "치료 효과" 는 예를 들어, 병원체에 의한 감염을 겪는 대상에서의 증상 감소를 포함할 수 있다.
- [0057] "표적 세포" 는 그의 활성화에 있어서의 변형이 원하는 결과 또는 반응을 유도할 수 있는 세포일 수 있다. 본원에서 사용하는 바와 같이, 세포는 시험관내 세포일 수 있다. 상기 세포는 전체 유기체로 발생할 수 없을 수 있는 단리된 세포일 수 있다.
- [0058] "리간드" 는 생물학적 목적으로 역할하도록 생분자에 결합하고 생분자와 복합체를 형성하는 임의의 물질일 수 있다. 본원에서 사용하는 바와 같이, "리간드" 는 또한 "항원" 또는 "면역원" 을 의미할 수 있다. 본원에서 사용하는 바와 같이 "항원" 및 "면역원" 은 상호교환가능하게 사용된다.
- [0059] 유전자 또는 핵산의 "발현" 은 세포적 유전자 발현 뿐 아니라 클로닝 시스템 및 임의의 다른 맥락에서의 핵산 (들) 의 전사 및 번역을 포함한다.
- [0060] 본원에서 사용하는 바와 같이, "벡터" 는 한 환경에서 다른 환경으로 개체를 이동시킬 수 있거나 이를 촉진시키는 도구이다. 예를 들어, 재조합 DNA 기술에서 사용한 일부 벡터는 개체, 예컨대 DNA 조각 (예컨대 이중 DNA 조각, 예컨대 이중 cDNA 조각) 이 표적 세포 내로 이동될 수 있게 한다. 본 발명은 바이러스 벡터, 박테리아 벡터, 원생동물 벡터, DNA 벡터 또는 이의 재조합체를 포함할 수 있는 재조합체 벡터를 포괄한다.
- [0061] 벡터에서의 발현을 위한 외인성 DNA (예를 들어, 관심 에피토프 및/또는 항원 및/또는 치료제를 인코딩함) 및 이러한 외인성 DNA 를 제공하는 문헌 뿐 아니라 핵산 분자의 발현 증강을 위한 전사 및/또는 번역 인자의 발현에 대해서, 그리고 "관심 에피토프", "치료적", "면역 반응", "면역학적 반응", "보호성 면역 반응", "면역학적 조성물", "면역원성 조성물" 및 "백신 조성물" 과 같은 용어에 관해서, 그 중에서도, 미국 특허 번호 5,990,091 (1999 년 11 월 23 일 발행), 및 WO 98/00166 및 WO 99/60164, 및 이에 인용된 문헌 및 특허 및 이들 PCT 출원의 공소제기에 있어서의 기록 문헌을 참조하고; 이들 모두는 본원에 참고로 포함된다. 따라서, 미국 특허 번호 5,990,091 및 WO 98/00166 및 WO 99/60164 및 이에 인용된 문헌 및 특허 및 이들 PCT 출원의 공소제기에 있어서의 기록 문헌, 및 이에 인용된 기타 문헌 또는 본원에 참고로 포함된 문헌은 본 발명의 실행에서 참조될 수 있고; 이에 인용된 모든 외인성 핵산 분자, 프로모터 및 벡터는 본 발명의 실행에서 사용될 수 있다. 이와 관련하여, 미국 특허 번호 6,706,693; 6,716,823; 6,348,450; 미국 특허 출원 일련 번호 10/424,409; 10/052,323; 10/116,963; 10/346,021; 및 WO 99/08713 (PCT/US98/16739 로부터, 1999 년 2 월 25 일 공개) 이 또한 언급된다.
- [0062] 본 발명의 양태는 전달 시스템을 통해 관심 유전자좌 또는 세포 또는 유기체에 전달되는 본 발명의 TALE 및 CRISPR-Cas 시스템을 포괄한다. 한 전달 수단은 벡터를 통한 것이며, 여기서 벡터는 바이러스 벡터, 예컨대 렌티- 또는 바칼로- 또는 바람직하게는 아데노-바이러스/아데노-관련 바이러스 벡터이나, 다른 전달 수단이 공지되어 있으며 (예컨대 효모 시스템, 미소낭포, 유전자 총/금 나노입자에 대한 부착 벡터) 제공된다. 일부 구현예에서, 하나 이상의 바이러스 또는 플라스미드 벡터는 나노입자, 엑소솜, 미소낭포 또는 유전자 총을 통해 전달될 수 있다.
- [0063] 본원에서 사용하는 바와 같이, 용어 "약물 조성물" 및 "약물", "백신 조성물", "백신", "백신 조성물", "치료적 조성물" 및 "치료적-면역학적 조성물" 은 항원 또는 병원체에 대한 보호를 유도하는 임의의 조성물을 포함한다. 일부 구현예에서, 보호는 병원체에 의한 감염의 억제 또는 예방으로 인한 것일 수 있다. 다른 구현예에서, 보호는 관심 항원(들) 에 대한 면역 반응에 의해 유도될 수 있거나, 항원에 대하여 효과적으로 보호하는 것일 수 있고; 예를 들어, 대상에 대한 투여 또는 주사 후, 표적화된 항원 또는 면역원에 대하여 보호성 면역 반응을 이끌어내거나 본 발명의 발명적인 아데노바이러스 벡터로부터 발현된 항원 또는 면역원에 대하여 효과적인 보호를 제공한다. 용어 "약학 조성물" 은 대상에게 전달되는 임의의 조성물을 의미한다.

일부 구현예에서, 상기 조성물은 병원체에 의한 감염을 억제하거나 방지하도록 전달될 수 있다.

- [0064] "치료적 유효량" 은 대상에게 투여시, 관심 유전자 생성물에 대해 치료적 반응 또는 면역 반응을 생성시키는, 관심 유전자를 인코딩하는 재조합 벡터의 양 또는 농도이다.
- [0065] 본원에서 사용하는 바와 같은 용어 "바이러스 벡터" 는 비제한적으로, 레트로바이러스, 아데노바이러스, 아데노-관련 바이러스, 알파바이러스 및 단순 포진 바이러스를 포함한다.
- [0066] 본 발명은 에너지 형태를 사용하는 내인성 또는 외인성 유전자 발현의 시공간적 제어를 포괄한다. 에너지 형태는 비제한적으로, 전자기 조사, 음파 에너지, 화학적 에너지 및 열 에너지를 포함할 수 있다. 본 발명의 바람직한 구현예에서, 에너지 형태는 전자기 조사, 바람직하게는 광 에너지이다. DNA 결합 아연 핑거 단백질에 연결된 전사 활성제와 같은 내인성 유전자의 발현 제어를 위한 이전의 접근방식은 시간적 또는 공간적 제어를 위한 메커니즘을 제공하지 않았다. 본원에 기재된 시스템의 광활성화를 위한 능력은 세포의 국부화 집단 내에서 정확한 시간에 시작하도록 유전자 발현 조절의 유도를 가능하게 한다.
- [0067] 상기 출원에서 상세히 설명한 바와 같은 제어의 양태는 적어도 하나 이상의 스위치(들) 에 관한 것이다. 본원에서 사용하는 바와 같은 용어 "스위치" 는 생물학적 기능의 활성화, 억제, 증강 또는 완료와 같은 상기 기능의 모든 양태를 포함하여, 변화에 영향을 주는 공동작용 방식으로 작용하는 성분의 집합 또는 시스템을 의미한다. 한 양태에서 용어 스위치는 유전자 조절 단백질의 기본 성분 및 이들 단백질이 인지하는 특이적 DNA 서열을 포함하는 유전적 스위치를 포함한다. 한 양태에서, 스위치는 유전자 조절에 사용되는 유도 및 억제 시스템에 관한 것이다. 일반적으로, 유전자 발현을 가능하게 하는 일부 분자 (유도물질로 지칭함) 가 존재하지 않으면 유도 시스템은 작동이 불가할 수 있다. 상기 분자는 "발현을 유도" 하는 것으로 일컬어진다. 이러한 일이 발생하는 방식은 제어 메커니즘 뿐 아니라 세포 유형에 있어서의 차이에 따라 좌우된다. 억제 시스템은 유전자 발현을 억제하는 일부 분자 (보조억제제로 지칭함) 가 존재하는 경우를 제외하고는 작동 가능하다. 상기 분자는 "발현을 억제" 하는 것으로 일컬어진다. 이러한 일이 발생하는 방식은 제어 메커니즘 뿐 아니라 세포 유형에 있어서의 차이에 따라 좌우된다. 본원에서 사용하는 바와 같은 용어 "유도" 는 수반된 분자 메커니즘에 관계없이 스위치의 모든 양태를 포함할 수 있다. 따라서 본 발명에 의해 이해되는 바와 같은 스위치는 비제한적으로 항생제 기반 유도 시스템, 전자기 에너지 기반 유도 시스템, 소분자 기반 유도 시스템, 핵 수용체 기반 유도 시스템 및 호르몬 기반 유도 시스템을 포함할 수 있다. 바람직한 구현예에서 스위치는 테트라사이클린 (Tet)/DOX 유도 시스템, 광 유도 시스템, 아브시스산 (ABA) 유도 시스템, 큐메이트 억제제/오퍼레이터 시스템, 4OHT/에스트로겐 유도 시스템, 액디손-기반 유도 시스템 또는 FKBP12/FRAP (FKBP12-라파마이신 복합체) 유도 시스템일 수 있다.
- [0068] 본 발명의 한 양태에서 하나 이상의 스위치는 TALE 또는 CRISPR-Cas 시스템과 연관될 수 있으며, 여기서 TALE 또는 CRISPR-Cas 시스템의 활성화는 스위치에 대한 하나 이상의 유도물질 에너지원과의 접촉에 의해 제어된다. 본 발명의 양태에 대해 본원에서 사용되는 바와 같은 용어 "접촉" 은 스위치와 유도물질 에너지원 사이의 임의의 관련적 관계를 의미하며, 이는 성분과의 물리적 상호작용 (함께 결합하는 분자 또는 단백질에서와 같이) 또는 에너지원에 의해 방사된 에너지의 경로에 있거나 상기 에너지에 의해 부딪치는 물리적 상호작용일 수 있다 (빛, 열 또는 소리의 흡수 또는 반사의 경우에서와 같이). 본 발명의 일부 양태에서 스위치와 유도물질 에너지원과의 접촉은 유도물질 에너지원의 적용에 의해 이루어진다. 본 발명은 또한 수동적 피드백 시스템을 통한 접촉을 포괄한다. 이는 TALE 또는 CRISPR-Cas 시스템 활성화, 이미 존재하며 따라서 적용할 필요가 없는 유도물질 에너지원과의 접촉에 의해 제어되는 임의의 수동적 조절 메커니즘을 비제한적으로 포함한다. 예를 들어 상기 에너지원은 세포 또는 세포 환경에 이미 존재하는 분자 또는 단백질일 수 있다. 수동적 접촉을 초래하는 상호작용은 비제한적으로 수용체/리간드 결합, 수용체/화학적 리간드 결합, 수용체/단백질 결합, 항체/단백질 결합, 단백질 이량체화, 단백질 이형이량체화, 단백질 다량체화, 핵 수용체/리간드 결합, 전사-후 변형 예컨대 인산화, 탈인산화, 유비퀴틴화 또는 탈유비퀴틴화를 포함할 수 있다.
- [0069] 2 개의 주요 분자 도구가 광반응성 전사 활성제-유사 (TAL) 효과기 시스템의 설계에 있어서 영향력을 행사하였다. 먼저, 조작된 TAL 효과기의 DNA 결합 특이성을 이용하여 게놈 내 특정 부위에 대해 복합체를 국부화시킨다. 두 번째로, 광-유도 단백질 이량체화를 사용하여 TAL 효과기에 의해 명시된 부위에 대해 활성화 또는 억제 도메인을 유도시켜, 다운스트림 유전자를 조절한다.
- [0070] 유도 효과기는 시간적 또는 공간적으로 특이적인 유전자 발현 제어가 필요한 시험관내 또는 생체내 적용에 고려된다. 시험관내 예: 발생 신호의 시기를 밝히기 위한 발생 유전자의 시간적으로 정확한 유도/억제, 세포-유형 패턴화 조직의 생성을 위한 세포 운명 재프로그래밍 인자의 공간적으로 제어된 유도. 시험관내 예: 특정

뇌 부위 내의 유전자 발현의 조합된 시간적 및 공간적 제어.

- [0071] 본 발명의 바람직한 구현예에서, 유도 효과기는 광 유도 전사 효과기 (LITE) 이다. LITE 시스템의 모듈성으로 인해, 임의의 수의 효과기 도메인이 전사 조절에 이용될 수 있다. 특히 유리한 구현예에서, 전사 활성화제 유사 효과기 (TALE) 및 활성화 도메인 VP64 가 본 발명에서 사용된다.
- [0072] LITE 는 시간적 및 공간적으로 정확한 방식으로 개별적 내인성 유전자의 발현을 조정하거나 변형시키도록 설계된다. 각각의 LITE 는 커스텀화된 DNA-결합 전사 활성화제 유사 효과기 (TALE) 단백질, 아라바돕시스 탈리아나 (*Arabidopsis thaliana*) 로부터의 광-반응성 크립토크롬 이형이량체, 및 전사 활성화/억제 도메인으로 이루어지는 2 성분 시스템을 포함할 수 있다. TALE 은 관심 유전자의 프로모터 서열에 결합하도록 설계된다. TALE 단백질은 크립토크롬 이형이량체의 절반 (크립토크롬-2 또는 CIB1) 에 융합되면서, 남아있는 크립토크롬 파트너가 전사 효과기 도메인에 융합된다. 효과기 도메인은 활성화제 예컨대 VP16, VP64 또는 p65, 또는 억제제 예컨대 KRAB, EnR 또는 SID 일 수 있다. LITE 의 미 자극 상태에서, TALE-크립토크롬2 단백질은 관심 유전자의 프로모터에 대해 국부화되나, CIB1-효과기 단백질에 결합하지는 않는다. 청색 스펙트럼 광으로 LITE 자극시, 크립토크롬-2 는 활성화되고, 형태적 변화를 거치며, 이의 결합 도메인이 밝혀진다. 다음으로 CIB1 은 크립토크롬-2 에 결합하여 관심 유전자의 프로모터 부위에 대해 효과기 도메인을 국부화시키고 유전자 과발현 또는 침묵화 (silencing) 를 개시시킨다.
- [0073] 활성화제 및 억제제 도메인은 종류, 강도, 메커니즘, 지속기간, 크기, 또는 임의의 수의 기타 매개변수를 기반으로 선택될 수 있다. 바람직한 효과기 도메인은 비제한적으로, 트랜스포사아제 도메인, 인테그라아제 도메인, 재조합효소 도메인, 리졸바아제 도메인, 인버타아제 도메인, 프로테아제 도메인, DNA 메틸트랜스퍼라아제 도메인, DNA 데메틸라아제 도메인, 히스톤 아세틸라아제 도메인, 히스톤 데아세틸라아제 도메인, 뉴클레아제 도메인, 억제제 도메인, 활성화제 도메인, 핵-국부화 신호 도메인, 전사-단백질 영입 도메인, 세포 흡수 활성화 관련 도메인, 핵산 결합 도메인 또는 항체 제시 도메인을 포함한다.
- [0074] LITE 또는 임의의 기타 유도 효과기에서의 유전자 표적화는 커스텀화된 TALE DNA 결합 단백질의 특이성을 통해 이루어질 수 있다. 관심 유전자의 프로모터 부위에서 표적 서열이 선택되며 상기 서열에 대해 커스텀화된 TALE 이 설계된다. TALE 의 중심 부분은 34 개 아미노산 길이의 탠덤 반복물로 이루어진다. 이들 반복물의 서열은 거의 동일하나, 각각의 반복물의 12 번째 및 13 번째 아미노산 (반복 가변성 2 잔기로 지칭함) 은 가변적이어서, 각각의 반복물의 뉴클레오티드-결합 특이성을 결정한다. 따라서, TALE 단량체 반복물의 적절한 배치를 갖는 구축물을 합성하여, 표적 프로모터 서열에 특이적인 DNA 결합 단백질을 제작한다.
- [0075] 본 발명의 유리한 구현예에서, 본원에 제공된 방법은 단리된, 비-자연 발생적인 재조합 또는 조작된 DNA 결합 단백질을 사용한다 (이는 향상된 효율 및 확장된 특이성을 갖는 핵산 서열의 표적화를 가능하게 하는 그의 조직 구조의 일부로서 TALE 단량체 또는 TALE 단량체 또는 반-단량체를 포함함).
- [0076] 자연 발생적 TALE 또는 "야생형 TALE" 은 수많은 종류의 프로테오박테리아에 의해 분비된 핵산 결합 단백질이다. TALE 폴리펩티드는 대부분 33, 34 또는 35 개 아미노산 길이며 서로 주로 아미노산 위치 12 및 13 에 있어서 상이한 고도로 보존된 단량체 폴리펩티드의 탠덤 반복물로 구성된 핵산 결합 도메인을 포함한다. 유리한 구현예에서 핵산은 DNA 이다. 본원에서 사용하는 바와 같이, 용어 "폴리펩티드 단량체", "TALE 단량체" 또는 "단량체" 는 TALE 핵산 결합 도메인 내의 고도로 보존된 반복적 폴리펩티드 서열을 의미하는데 사용되며, 용어 "반복 가변성 2-잔기" 또는 "RVD" 는 폴리펩티드 단량체의 위치 12 및 13 에서의 고도로 가변적인 아미노산을 의미하는데 사용된다. DNA 결합 도메인 내에 포함되는 TALE 단량체의 일반적 표시는  $X_{1-11}-(X_{12}X_{13})-X_{14-33}$  또는  $34$  또는  $35$  이며, 여기서 첨자는 아미노산 위치를 나타내고, X 는 임의의 아미노산을 나타낸다.  $X_{12}X_{13}$  은 RVD 를 나타낸다. 일부 폴리펩티드 단량체에서, 위치 13 에서의 가변적 아미노산은 결실되거나 부재하고, 이러한 단량체에서, RVD 는 단일 아미노산으로 이루어진다. 이러한 경우 RVD 는 대안적으로  $X^*$  로 표시될 수 있으며, 여기서 X 는  $X_{12}$  를 나타내고, (\*) 는  $X_{13}$  이 부재함을 나타낸다. DNA 결합 도메인은 TALE 단량체의 여러 반복물을 포함하며 이는  $(X_{1-11}-(X_{12}X_{13})-X_{14-33}$  또는  $34$  또는  $35$ )<sub>z</sub> 로서 표시되고, 이때 유리한 구현예에서, z 는 적어도 5 내지 40 이다. 추가로 유리한 구현예에서, z 는 적어도 10 내지 26 이다.
- [0077] TALE 단량체는 그의 RVD 에서의 아미노산 동일성에 의해 결정되는 뉴클레오티드 결합 친화성을 갖는다. 예를 들어, NI 의 RVD 를 갖는 폴리펩티드 단량체는 바람직하게는 아데닌 (A) 에 결합하고, NG 의 RVD 를 갖는 단량체는 바람직하게는 티민 (T) 에 결합하고, HD 의 RVD 를 갖는 단량체는 바람직하게는 시토신 (C) 에 결합하고, NN 의 RVD 를 갖는 단량체는 바람직하게는 아데닌 (A) 및 구아닌 (G) 둘 모두에 결합한다. 본 발

명의 또 다른 구현예에서, IG 의 RVD 를 갖는 단량체는 바람직하게는 T 에 결합한다. 따라서, TALE 의 핵산 결합 도메인에서의 폴리펩티드 단량체 반복물의 수 및 순서는 이의 핵산 표적 특이성을 결정한다. 본 발명의 추가 구현예에서, NS 의 RVD 를 갖는 단량체는 모든 4 개 염기 쌍을 인지하며 A, T, G 또는 C 에 결합할 수 있다. TALE 의 구조 및 기능은 예를 들어, Moscou et al, Science 326: 1501 (2009); Boch et al, Science 326:1509-1512 (2009); 및 Zhang et al, Nature Biotechnology 29: 149-153 (2011) 에 추가로 기재되어 있으며, 이들 각각은 그 전체가 본원에 참고로 포함된다.

[0078] 본 발명의 방법에서 사용한 폴리펩티드는 단리된, 비-자연 발생적인 재조합 또는 조작된 핵산 결합 단백질이며 이는 특정 핵산 서열을 표적화하도록 설계되는 폴리펩티드 단량체 반복물을 함유하는 핵산 또는 DNA 결합 부위를 갖는다.

[0079] 본원에서 기재한 바와 같이, HN 또는 NH 의 RVD 를 갖는 폴리펩티드 단량체는 바람직하게는 구아닌에 결합하며, 이로써 표적 핵산 서열을 함유하는 구아닌에 대한 높은 결합 특이성을 갖는 TALE 폴리펩티드가 생성되게 한다. 본 발명의 바람직한 구현예에서, RVD RN, NN, NK, SN, NH, KN, HN, NQ, HH, RG, KH, RH 및 SS 를 갖는 폴리펩티드 단량체는 바람직하게는 구아닌에 결합한다. 본 발명의 훨씬 더 유리한 구현예에서, RVD RN, NK, NQ, HH, KH, RH, SS 및 SN 을 갖는 폴리펩티드 단량체는 바람직하게는 구아닌에 결합하며 이로써 표적 핵산 서열을 함유하는 구아닌에 대한 높은 결합 특이성을 갖는 TALE 폴리펩티드가 생성되게 한다. 본 발명의 보다 더 유리한 구현예에서, RVD HH, KH, NH, NK, NQ, RH, RN 및 SS 를 갖는 폴리펩티드 단량체는 바람직하게는 구아닌에 결합하며 이로써 표적 핵산 서열을 함유하는 구아닌에 대한 높은 결합 특이성을 갖는 TALE 폴리펩티드가 생성되게 한다. 추가로 유리한 구현예에서, 구아닌에 대한 높은 결합 특이성을 갖는 RVD 는 RN, NH, RH 및 KH 이다. 또한, NV 의 RVD 를 갖는 폴리펩티드 단량체는 바람직하게는 아데닌 및 구아닌에 결합한다. 본 발명의 보다 바람직한 구현예에서, H\*, HA, KA, N\*, NA, NC, NS, RA 및 S\* 의 RVD 를 갖는 단량체는 비슷한 친화성으로 아데닌, 구아닌, 시토신 및 티민에 결합한다.

[0080] 본 발명의 보다 더 유리한 구현예에서 아데닌에 대한 특이성을 갖는 RVD 는 NI, RI, KI, HI 및 SI 이다. 본 발명의 보다 바람직한 구현예에서, 아데닌에 대한 특이성을 갖는 RVD 는 HN, SI 및 RI 이고, 가장 바람직하게는 아데닌 특이성에 대한 RVD 는 SI 이다. 본 발명의 보다 더 바람직한 구현예에서 티민에 대한 특이성을 갖는 RVD 는 NG, HG, RG 및 KG 이다. 본 발명의 추가로 유리한 구현예에서, 티민에 대한 특이성을 갖는 RVD 는 KG, HG 및 RG 이고, 가장 바람직하게는 티민 특이성에 대한 RVD 는 KG 또는 RG 이다. 본 발명의 보다 더 바람직한 구현예에서 시토신에 대한 특이성을 갖는 RVD 는 HD, ND, KD, RD, HH, YG 및 SD 이다. 본 발명의 추가로 유리한 구현예에서, 시토신에 대한 특이성을 갖는 RVD 는 SD 및 RD 이다. 본 발명의 가장 바람직한 구현예에 포함되는 대표적 RVD 및 이들이 표적화하는 뉴클레오티드에 대해서는 도 7B 를 참고한다. 추가로 유리한 구현예에서 변이체 TALE 단량체는 도 7A 에서 나타난 바와 같이 뉴클레오티드에 대한 특이성을 나타내는 임의의 RVD 를 포함할 수 있다. 모든 이러한 TALE 단량체는 관련되나 동일하지는 않은 표적 핵산 서열의 레퍼토리에 결합할 수 있는 퇴행적 TALE 폴리펩티드가 생성될 수 있게 한다. 본 발명의 추가 구현예에서, RVD NT 는 G 및 A 에 결합할 수 있다. 본 발명의 추가 구현예에서, RVD NP 는 A, T 및 C 에 결합할 수 있다. 본 발명의 보다 유리한 구현예에서, 하나 이상의 선택된 RVD 는 NI, HD, NG, NN, KN, RN, NH, NQ, SS, SN, NK, KH, RH, HH, KI, HI, RI, SI, KG, HG, RG, SD, ND, KD, RD, YG, HN, NV, NS, HA, S\*, N\*, KA, H\*, RA, NA 또는 NC 일 수 있다.

[0081] 핵산 또는 DNA 결합 도메인의 하나 이상의 폴리펩티드 단량체의 사전결정된 N-말단에서 C-말단 순서는 본 발명의 폴리펩티드가 결합할 상응하는 사전결정된 표적 핵산 서열을 결정한다. 본원에서 사용하는 바와 같이 단량체 및 적어도 하나 이상의 반-단량체는 관심 유전자 또는 게놈 유전자좌를 "표적화하도록 특이적으로 순서화된다". 식물 게놈에서, 자연적 TALE-결합 부위는 항상 티민 (T) 으로 시작하는데, 이는 TALE 폴리펩티드의 비-반복성 N-말단 내의 암호성 신호에 의해 명시될 수 있고; 일부 경우 상기 부위는 반복물 0 으로서 지칭될 수 있다. 동물 게놈에서, TALE 결합 부위는 반드시 티민 (T) 으로 시작할 필요는 없으며 본 발명의 폴리펩티드는 T, A, G 또는 C 로 시작하는 DNA 서열을 표적화할 수 있다. TALE 단량체의 탠덤 반복물은 항상, 반복성 전체 길이 TALE 단량체의 처음 20 개 아미노산과만 동일성을 공유할 수 있는 서열의 스트레치 또는 반-길이 반복물로 종료되며, 상기 반-반복물은 반-단량체로서 지칭될 수 있다 (도 8). 그러므로, 이는 표적화되는 핵산 또는 DNA 의 길이가 전체 단량체 수 + 2 와 동일하다는 것에 따른다.

[0082] 예를 들어 핵산 결합 도메인은, 사전결정된 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 개 뉴클레오티드 길이 핵산 서열에 결합시키기 위해 N-말단에서 C-말단 방향으로 배열된 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 개 이상의 폴리펩티드 단량체를

함유하도록 조작될 수 있다. 본 발명의 보다 유리한 구현예에서, 핵산 결합 도메인은 각각 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27 및 28 개 뉴클레오티드 길이의 핵산 서열을 표적화하기 위해 특이적으로 순서화되거나 배열된 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 개 이상의 전체 길이 폴리펩티드 단량체를 함유하도록 조작될 수 있다. 특정 구현예에서 폴리펩티드 단량체는 인접하다. 일부 구현예에서, 특히 TALE 폴리펩티드의 C-말단에서 존재하는 경우, 반-단량체가 하나 이상의 단량체 대신에 사용될 수 있다.

[0083] 폴리펩티드 단량체는 일반적으로 33, 34 또는 35 개 아미노산 길이이다. RVD 를 제외하고, 폴리펩티드 단량체의 아미노산 서열은 고도로 보존되거나 본원에 기재된 바와 같이, RVD 를 제외하고 폴리펩티드 단량체에서의 아미노산은 TALE 활성화에 영향을 주는 패턴을 나타내며, 이의 확인이 본 발명의 바람직한 구현예에서 사용될 수 있다. 단량체 서열 중의 아미노산의 대표적 조합 (RVD 제외) 은 TALE 활성화에 대한 효과를 갖는 것으로 출원인에 의해 나타내어진다 (도 10). 본 발명의 보다 바람직한 구현예에서, DNA 결합 도메인이  $(X_{1-11}-X_{12}X_{13}-X_{14-33}$  또는  $34$  또는  $35)_z$  를 포함하는 경우 (여기서  $X_{1-11}$  은 11 개의 인접한 아미노산의 사슬이고,  $X_{12}X_{13}$  은 반복 가변성 2 잔기 (RVD) 이고,  $X_{14-33}$  또는  $34$  또는  $35$  는 21, 22 또는 23 개의 인접한 아미노산의 사슬이고,  $z$  는 적어도 5 내지 26 임), 아미노산의 바람직한 조합은  $X_{1-4}$  에서의 [LTLD] 또는 [LTLA] 또는 [LTQV], 또는 위치  $X_{30-33}$  또는  $X_{31-34}$  또는  $X_{32-35}$  에서의 [EQHG] 또는 [RDHG] 이다. 또한, 단량체에서의 다른 관심 아미노산 조합은, 단량체가 34 개 아미노산 길이인 경우,  $X_{1-4}$  에서의 [LTPD] 및  $X_{16-20}$  에서의 [NQALE] 및  $X_{32-34}$  에서의 [DHG] 이다. 단량체가 33 또는 35 개 아미노산 길이인 경우, 상응하는 이동이 인접한 아미노산 [NQALE] 및 [DHG] 의 위치에서 발생하고; 바람직하게는, 본 발명의 구현예는  $X_{15-19}$  또는  $X_{17-21}$  에서 [NQALE] 및  $X_{31-33}$  또는  $X_{33-35}$  에서 [DHG] 를 가질 수 있다.

[0084] 본 발명의 추가 구현예에서 단량체에서의 관심 아미노산 조합은, 단량체가 34 개 아미노산 길이인 경우,  $X_{1-4}$  에서의 [LTPD] 및  $X_{16-20}$  에서의 [KRALE] 및  $X_{32-34}$  에서의 [AHG] 또는  $X_{1-4}$  에서의 [LTPE] 및  $X_{16-20}$  에서의 [KRALE] 및  $X_{32-34}$  에서의 [DHG] 이다. 단량체가 33 또는 35 개 아미노산 길이인 경우, 상응하는 이동이 인접한 아미노산 [KRALE], [AHG] 및 [DHG] 의 위치에서 발생한다. 바람직한 구현예에서, 인접한 아미노산의 위치는 ( $X_{1-4}$  에서의 [LTPD] 및  $X_{15-19}$  에서의 [KRALE] 및  $X_{31-33}$  에서의 [AHG]) 또는 ( $X_{1-4}$  에서의 [LTPE] 및  $X_{15-19}$  에서의 [KRALE] 및  $X_{31-33}$  에서의 [DHG]) 또는 ( $X_{1-4}$  에서의 [LTPD] 및  $X_{17-21}$  에서의 [KRALE] 및  $X_{33-35}$  에서의 [AHG]) 또는 ( $X_{1-4}$  에서의 [LTPE] 및  $X_{17-21}$  에서의 [KRALE] 및  $X_{33-35}$  에서의 [DHG]) 일 수 있다. 본 발명의 추가 구현예에서, 인접한 아미노산 [NGKQALE] 는 위치  $X_{14-20}$  또는  $X_{13-19}$  또는  $X_{15-21}$  에서 존재한다. 이들 대표적 위치는 본 발명의 다양한 구현예를 제안하며 본원에서 기재한 모든 TALE 단량체에서의 관심 아미노산의 조합 또는 추가적인 관심 아미노산을 확인하기 위한 안내를 제공한다 (도 9A-F 및 10).

[0085] 하기 제공된 것은 폴리펩티드 단량체의 보존된 부분의 예시적 아미노산 서열이다. 각 서열에서의 RVD 의 위치를 XX 또는 X\* 로 나타낸다 (여기서 (\*) 는 RVD 가 단일 아미노산이고 잔기 13 ( $X_{13}$ ) 이 부재하다는 것을 나타냄).

L T P A Q V V A I A S X X G G K Q A L E T V Q R L L P V L C Q D H G  
 L T P A Q V V A I A S X \* G G K Q A L E T V Q R L L P V L C Q D H G  
 L T P D Q V V A I A N X X G G K Q A L A T V Q R L L P V L C Q D H G  
 L T P D Q V V A I A N X X G G K Q A L E T L Q R L L P V L C Q D H G  
 L T P D Q V V A I A N X X G G K Q A L E T V Q R L L P V L C Q D H G  
 L T P D Q V V A I A S X X G G K Q A L A T V Q R L L P V L C Q D H G  
 L T P D Q V V A I A S X X G G K Q A L E T V Q R L L P V L C Q D H G  
 L T P D Q V V A I A S X X G G K Q A L E T V Q R V L P V L C Q D H G  
 L T P E Q V V A I A S X X G G K Q A L E T V Q R L L P V L C Q A H G  
 L T P Y Q V V A I A S X X G S K Q A L E T V Q R L L P V L C Q D H G  
 L T R E Q V V A I A S X X G G K Q A L E T V Q R L L P V L C Q D H G  
 L S T A Q V V A I A S X X G G K Q A L E G I G E Q L L K L R T A P Y G  
 L S T A Q V V A V A S X X G G K P A L E A V R A Q L L A L R A A P Y G

[0086]

[0087]

서열 (X<sub>1-11</sub>-X<sub>14-34</sub> 또는 X<sub>1-11</sub>-X<sub>14-35</sub>) (여기서 X 는 임의의 아미노산이고 첨자는 아미노산 위치임) 로 나타낼 수 있는 RVD 를 제외하고 TALE 단량체의 추가적인 목록을 도 9A-F 에 제공한다.

[0088]

[Zhang et al, Nature Biotechnology 29:149-153 (2011)] 에서 기재한 바와 같이, TALE 폴리펩티드 결합 효율은, 자연 발생적 TALE 의 DNA 결합 부위의 직접적 N-말단 또는 C-말단인 "캡핑 부위" 로부터의 아미노산 서열을 조작된 TALE DNA 결합 부위의 위치 N-말단 또는 C-말단에서의 조작된 TALE 에 포함시킴으로써 증가될 수 있다. 따라서 특정 구현예에서, 본원에서 기재된 TALE 폴리펩티드는 N-말단 캡핑 부위 및/또는 C-말단 캡핑 부위를 추가로 포함한다.

[0089]

N-말단 캡핑 부위의 예시적 아미노산 서열은 하기와 같다:

M D P I R S R T P S P A R E L L S G P Q P D G V Q P T A D R G V S P  
 P A G G P L D G L P A R R T M S R T R L P S P P A P S P A F S A D S  
 F S D L L R Q F D P S L F N T S L F D S L P P F G A H H T E A A T G  
 E W D E V Q S G L R A A D A P P P T M R V A V T A A R P P R A K P A  
 P R R R A A Q P S D A S P A A Q V D L R T L G Y S Q Q Q E K I K P  
 K V R S T V A Q H H E A L V G H G F T H A H I V A L S Q H P A A L G  
 T V A V K Y Q D M I A A L P E A T H E A I V G V G K Q W S G A R A L  
 E A L L T V A G E L R G P P L Q L D T G Q L L K I A K R G G V T A V  
 E A V H A W R N A L T G A P L N

[0090]

[0091]

[0092]

C-말단 캡핑 부위의 예시적 아미노산 서열은 하기와 같다:

R P A L E S I V A Q L S R P D P A L A A L T N D H L V A L A C L G  
 G R P A L D A V K K G L P H A P A L I K R T N R R I P E R T S H R  
 V A D H A Q V V R V L G F F Q C H S H P A Q A F D D A M T Q F G M  
 S R H G L L Q L F R R V G V T E L E A R S G T L P P A S Q R W D R  
 I L Q A S G M K R A K P S P T S T Q T P D Q A S L H A F A D S L E  
 R D L D A P S P M H E G D Q T R A S

[0093]

[0094]

본원에서 사용하는 바오 같이 사전결정된 N-말단 캡핑 부위, 반복 TALE 단량체를 포함하는 DNA 결합 도메인 및 C-말단 캡핑 부위의 "N-말단" 에서 "C-말단" 배향은 본 발명의 d-TALE 또는 폴리펩티드에서의 상이한 도메인의 구성을 위한 구조적 기초를 제공한다.

[0095]

전체 N-말단 및/또는 C-말단 캡핑 부위는 DNA 결합 부위의 결합 활성을 증강시키는데 필요하지 않다. 따라서 특정 구현예에서, N-말단 및/또는 C-말단 캡핑 부위의 단편은 본원에 기재된 TALE 폴리펩티드에 포함된다.

[0096]

특정 구현예에서, 본원에 기재된 TALE 폴리펩티드는 N-말단 캡핑 부위의 적어도 10, 20, 30, 40, 50, 54, 60, 70, 80, 87, 90, 94, 100, 102, 110, 117, 120, 130, 140, 147, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260 또는 270 개 아미노산을 포함하는 N-말단 캡핑 부위 단편을 함유한다. 특정 구현예에서, N-말단 캡핑 부위 단편 아미노산은 N-말단 캡핑 부위의 C-말단 (DNA-결합 부위 근부 말단) 의 것이

다. [Zhang et al., Nature Biotechnology 29:149-153 (2011)] 에서 기재한 바와 같이, C-말단 240 개 아미노산을 포함하는 N-말단 캡핑 부위 단편은 전체 길이 캡핑 부위와 동일하게 결합 활성을 증강시키는 한편, C-말단 147 개 아미노산을 포함하는 단편은 전체 길이 캡핑 부위 효능의 80% 초과를 유지하고, C-말단 117 개 아미노산을 포함하는 단편은 전체 길이 캡핑 부위 활성의 50% 초과를 유지한다.

[0097] 일부 구현예에서, 본원에 기재된 TALE 폴리펩티드는 C-말단 캡핑 부위의 적어도 6, 10, 20, 30, 37, 40, 50, 60, 68, 70, 80, 90, 100, 110, 120, 127, 130, 140, 150, 155, 160, 170, 180 개 아미노산을 포함하는 C-말단 캡핑 부위 단편을 함유한다. 특정 구현예에서, C-말단 캡핑 부위 단편 아미노산은 C-말단 캡핑 부위의 N-말단 (DNA-결합 부위 근부 말단) 의 것이다. [Zhang et al., Nature Biotechnology 29: 149-153 (2011)] 에서 기재한 바와 같이, C-말단 68 개 아미노산을 포함하는 C-말단 캡핑 부위 단편은 전체 길이 캡핑 부위와 동일하게 결합 활성을 증강시키는 한편, C-말단 20 개 아미노산을 포함하는 단편은 전체 길이 캡핑 부위 효능의 50% 초과를 유지한다.

[0098] 특정 구현예에서, 본원에 기재된 TALE 폴리펩티드의 캡핑 부위는 본원에 제공된 캡핑 부위 서열에 대해 동일한 서열을 가질 필요는 없다. 따라서 일부 구현예에서, 본원에 기재된 TALE 폴리펩티드의 캡핑 부위는 본원에 제공된 캡핑 부위 아미노산 서열에 대해 적어도 50%, 60%, 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 동일하거나 동일성을 공유하는 서열을 갖는다. 서열 동일성은 서열 상동성과 관련된다. 상동성 비교는 육안으로, 또는 통상 용이하게 이용가능한 서열 비교 프로그램을 보조로 사용하여 수행될 수 있다. 이들 시판 컴퓨터 프로그램은 둘 이상의 서열 사이의 상동성 퍼센트 (%) 를 계산할 수 있으며 또한 둘 이상의 아미노산 또는 핵산 서열에 의해 공유된 서열 동일성을 계산할 수 있다. 일부 바람직한 구현예에서, 본원에 기재된 TALE 폴리펩티드의 캡핑 부위는 본원에 제공된 캡핑 부위 아미노산 서열에 대해 95% 이상 동일하거나 동일성을 공유하는 서열을 갖는다.

[0099] 서열 상동성은, BLAST 또는 FASTA 를 비제한적으로 포함하는 당업계에 공지된 임의의 다수 컴퓨터 프로그램에 의해 생성될 수 있다. GCG Wisconsin Bestfit 패키지와 같은 정렬을 실행하기에 적합한 컴퓨터 프로그램을 또한 사용할 수 있다. 소프트웨어가 최적 정렬을 생성시키고 나면, 상동성 %, 바람직하게는 서열 동일성 % 를 계산할 수 있다. 소프트웨어는 통상 서열 비교의 일부로서 이를 수행하며 수치적 결과를 산출해낸다.

[0100] 본원에 기재된 유리한 구현예에서, 본 발명의 TALE 폴리펩티드는 하나 이상의 효과기 도메인에 연결된 핵산 결합 도메인을 포함한다. 용어 "효과기 도메인" 또는 "조절 또는 기능적 도메인" 은 핵산 결합 도메인에 의해 인지된 핵산 서열에 결합하는 것 외의 활성을 갖는 폴리펩티드 서열을 의미한다. 핵산 결합 도메인을 하나 이상의 효과기 도메인과 조합시킴으로써, 본 발명의 폴리펩티드를 사용하여, 핵산 결합 도메인이 특이적으로 결합하는 특정 표적 DNA 서열에 대해 효과기 도메인에 의해 매개된 하나 이상의 기능 또는 활성을 표적화할 수 있다. 용어 "효과기 도메인" 및 "기능적 도메인" 은 본 출원에서 상호교환가능하게 사용된다.

[0101] 본원에서 기재된 TALE 폴리펩티드의 일부 구현예에서, 효과기 도메인에 의해 매개된 활성은 생물학적 활성이다. 예를 들어, 일부 구현예에서 효과기 도메인은 전사 저해제 (즉, 억제제 도메인), 예컨대 mSin 상호작용 도메인 (SID), SID4X 도메인 또는 크루펠 (Kruppel)-연관 상자 (KRAB) 또는 KRAB 도메인의 단편이다. 일부 구현예에서 효과기 도메인은 전사의 증강제 (즉, 활성화 도메인), 예컨대 VP16, VP64 또는 p65 활성화 도메인이다. 이들 상이한 활성화 도메인이 Sox2 mRNA 수준에 대해 갖는 효과의 그래프적 비교를 도 11 에 제공한다.

[0102] 본원에서 사용하는 바와 같이, VP16 은 헤르페스바이러스 단백질이다. 이는 바이러스 전초기 유전자 발현을 특이적으로 활성화시키는 매우 강력한 전사 활성화제이다. VP16 활성화 도메인은 산성 잔기가 풍부하며 고전적 산성 활성화 도메인 (AAD) 으로서 간주된다. 본원에서 사용하는 바와 같이, VP64 활성화 도메인은 VP16 의 최소 활성화 도메인의 사랑체 반복물이다. 본원에서 사용하는 바와 같이, p65 는 NF-카파 B 전사 인자 복합체를 구성하는 2 개 단백질 중 하나이다. 다른 단백질은 p50 이다. p65 활성화 도메인은 p65 서브 유닛의 일부이며 이는 p50 의 부재 하에서도 강력한 전사 활성화제이다. 특정 구현예에서, 효과기 도메인은 포유동물 단백질 또는 이의 생물학적 활성 단편이다. 이러한 효과기 도메인은 "포유동물 효과기 도메인" 으로서 지칭한다.

[0103] 일부 구현예에서 핵산 결합은 예를 들어, 트랜스포사아제 도메인, 인테그라아제 도메인, 재조합효소 도메인, 리졸바아제 도메인, 인버타아제 도메인, 프로테아제 도메인, DNA 메틸트랜스퍼라아제 도메인, DNA 히드록실메틸라아제 도메인, DNA 데메틸라아제 도메인, 히스톤 아세틸라아제 도메인, 히스톤 데아세틸라아제 도메인, 뉴클레아제 도메인, 억제제 도메인, 활성화제 도메인, 핵-국부화 신호 도메인, 전사-조절 단백질 (또는 전사 복합체 영입)

도메인, 세포 흡수 활성 관련 도메인, 핵산 결합 도메인, 항체 제시 도메인, 히스톤 개질 효소, 히스톤 개질 효소의 영입제; 히스톤 개질 효소의 저해제, 히스톤 메틸트랜스퍼라아제, 히스톤 데메틸라아제, 히스톤 키나아제, 히스톤 포스파타아제, 히스톤 리보실라아제, 히스톤 데리보실라아제, 히스톤 유비퀴티나아제, 히스톤 데유비퀴티나아제, 히스톤 비오티나아제 및 히스톤 테일 프로테아제를 비제한적으로 포함하는 효과기 도메인 또는 기능적 도메인과 연결된다.

[0104] 일부 구현예에서, 효과기 도메인은 트랜스포사아제 활성, 인테그라아제 활성, 재조합효소 활성, 리졸바아제 활성, 인버타아제 활성, 프로테아제 활성, DNA 메틸트랜스퍼라아제 활성, DNA 데메틸라아제 활성, 히스톤 아세틸라아제 활성, 히스톤 데아세틸라아제 활성, 뉴클레아제 활성, 핵-국부화 신호전달 활성, 전사 억제제 활성, 전사 활성제 활성, 전사 인자 영입 활성 또는 세포 흡수 신호전달 활성을 비제한적으로 포함하는 활성을 나타내는 단백질 도메인이다. 본 발명의 다른 바람직한 구현예는 본원에 기재된 활성의 임의 조합을 포함할 수 있다.

[0105] [Zhang et al, Nature Biotechnology 29: 149-153 (2011)] 에서 기재된 바와 같이, 핵산 결합 도메인 및 효과기 도메인을 갖는 TALE 폴리펩티드를 사용하여, 핵산 결합 도메인에 의해 인지된 사전결정된 핵산 서열을 갖는 게놈 위치에 대해 효과가 도메인의 활성을 표적화할 수 있다. 본원에 기재된 본 발명의 일부 구현예에서, TALE 폴리펩티드는 유전자 조절 활성, 예컨대 조절, 코딩 및/또는 유전자간 부위에 대한 전사 또는 번역 변형제 활성, 예컨대 인핸서 및/또는 억제제 활성을 표적화하기 위해 설계되고 사용되며, 이는 코딩 부위의 전사 업스트림 및 다운스트림에 영향을 줄 수 있고 유전자 발현을 증강시키거나 억제하는데 사용될 수 있다. 예를 들어, TALE 폴리펩티드는 전사 인자로부터의 DNA-결합 도메인을 갖는 효과기 도메인, 전사 인자 (활성제, 억제제, 보조-활성제, 보조-억제제) 로부터의 효과기 도메인, 사일런서 (silencer), 핵 호르몬 수용체 및/또는 크로마틴 연관 단백질 및 그의 변형제 (예를 들어 메틸라아제, 키나아제, 포스파타아제, 아세틸라아제 및 데아세틸라아제) 를 포함할 수 있다. 바람직한 구현예에서, TALE 폴리펩티드는 뉴클레아제 도메인을 포함할 수 있다. 보다 바람직한 구현예에서 뉴클레아제 도메인은 비-특이적 FokI 엔도뉴클레아제 촉매 도메인이다.

[0106] 추가 구현예에서, 유전자 발현 조절을 위해 유용한 도메인은 또한 암유전자의 유전자 산물로부터 수득될 수 있다. 본 발명의 추가로 유리한 구현예에서, 인테그라아제 또는 트랜스포사아제 활성을 갖는 효과기 도메인은 외인성 핵산 서열의 특정 핵산 서열 부위 내로의 통합을 촉진시키고, 특정 내인성 핵산 서열을 제거 (녹-아웃 (knock-out)) 하고/하거나 예컨대 DNA 메틸트랜스퍼라아제, DNA 데메틸라아제, 히스톤 아세틸라아제 및 히스톤 데아세틸라아제 활성을 촉진시킴으로써 후생적 신호 및 결과적인 유전자 조절을 변형시키는데 사용될 수 있다. 다른 구현예에서, 뉴클레아제 활성을 갖는 효과기 도메인은 본 발명의 폴리펩티드가 특이적으로 결합하는 표적 서열을 잘라내거나 소화시켜 게놈 구조를 변형시키는데 사용될 수 있으며, 이들 부위에 외인성 유전자를 도입시킬 수 있다. 추가 구현예에서, 인버타아제 활성을 갖는 효과기 도메인은 DNA 단편의 배향을 바꾸어 게놈 구조를 변형시키는데 사용될 수 있다.

[0107] 특히 유리한 구현예에서, 본 발명의 방법에서 사용된 폴리펩티드는 전사 활성을 표적화하는데 사용될 수 있다. 본원에서 사용하는 바와 같이, 용어 "전사 인자" 는 전사 제어를 위해 관심 유전자 또는 게놈 유전자좌와 연관된 특정 DNA 서열에 결합하는 단백질 또는 폴리펩티드를 의미한다. 전사 인자는 RNA 중합효소의 관심 유전자로의 영입을 촉진 (활성제로서) 하거나 차단 (억제제로서) 할 수 있다. 전사 인자는 그의 기능을 단독으로 또는 더 큰 단백질 복합체의 일부로서 수행할 수 있다. 전사 인자에 의해 사용되는 유전자 조절의 메커니즘은 비제한적으로, a) RNA 중합효소 결합의 안정화 또는 탈안정화, b) 히스톤 단백질의 아세틸화 또는 탈아세틸화 및 c) 보조-활성제 또는 보조-억제제 단백질의 영입을 포함한다. 또한, 전사 인자는 기본 전사, 전사의 증강, 발생, 세포간 신호전달에 대한 반응, 환경적 신호에 대한 반응, 세포-주기 제어 및 발병과정을 비제한적으로 포함하는 생물학적 활성에 있어서 그 역할을 수행한다. 전사 인자에 대한 정보와 관련하여, [Latchman and DS (1997) Int. J. Biochem. Cell Biol. 29 (12): 1305-12; Lee TI, Young RA (2000) Annu. Rev. Genet. 34: 77-137 및 Mitchell PJ, Tjian R (1989) Science 245 (4916): 371-8] 가 언급되고, 이는 본원에 그 전체가 참고로 포함된다.

[0108] LITE 의 광 반응성은 크립토크롬-2 및 CIB1 의 활성화 및 결합을 통해 이루어진다. 상기 언급한 바와 같이, 청색광 자극은 크립토크롬-2 에 있어서의 활성화 형태적 변화를 유도하여, 이의 결합 파트너 CIB1 의 영입을 초래한다. 이러한 결합은 빠르고 가역적이며, 펄스화된 자극 후 <15 초 내에 포화가 이루어지고 자극 종료 후 <15 분에 기준선으로 복귀된다. 이러한 신속한 결합 속도는 유도제의 흡수 및 청소율보다는 전사/번역 및 전사체/단백질 분해의 속도에 의해서만 일시적으로 결합한 LITE 시스템을 초래한다. 크립토크롬-2 활성화는 또한 매우 민감하여, 낮은 광 강도 자극의 사용을 가능하게 하며 광독성의 위험성을 완화시킨다. 추가로, 미손상 포유동물 뇌와 같은 맥락에서, 가변적인 광 강도는 LITE 자극 부위의 크기를 제어하는데 사용될 수

있어, 벡터 단독 전달이 제공할 수 있는 것보다 더 큰 정확성을 부여한다.

- [0109] LITE 시스템의 모듈성은 임의 수의 효과기 도메인이 전사 조절에 사용될 수 있게 한다. 따라서, 활성화 및 억제제 도메인은 종류, 강도, 메커니즘, 지속기간, 크기, 또는 임의 수의 기타 매개변수를 기반으로 선택될 수 있다.
- [0110] 출원인은 그 다음으로 LITE 시스템의 2 가지 원형적 시현을 제시한다. 제 1 예는 마우스 유전자 NEUROG2의 전사를 활성화시키도록 설계된 LITE 이다. 마우스 NEUROG2의 업스트림 프로모터 부위에 위치한 서열 TGAATGATGATAATACGA를 표적으로서 선택하고, 상기 서열에 매치되도록 TALE을 설계하고 합성하였다. TALE 서열을 핵 국부화 신호(아미노산: SPKKKRKVEAS)를 통해 크립토크롬-2에 대한 서열에 연결시켜, 사이토솔로부터의 단백질의 핵 공간으로의 이동을 촉진시킨다. 동일한 핵 국부화 신호를 사용하여, 전사 활성화 도메인 VP64에 연결된 CIB1 도메인을 포함하는 제 2 벡터를 합성하였다. 이러한 제 2 벡터, 또한 GFP 서열을 2A 전사 스킵 신호에 의해 CIB1-VP64 융합 서열로부터 분리하였다. 각 구축물의 발현을 편재성, 항시성 프로모터(CMV 또는 EF1- $\alpha$ )에 의해 유도하였다. Neuro 2A 세포주로부터의 마우스 신경아세포종 세포를 2 개 벡터로 동시-트랜스펙션시켰다. 벡터 발현이 가능하도록 인큐베이션한 후, 샘플을 488 nm LED의 어레이로부터의 주기적 펄스의 청색광에 의해 자극하였다. 미자극된 동시-트랜스펙션 샘플 및 형광 리포터 YFP 로만 트랜스펙션된 샘플을 대조군으로서 사용하였다. 각 실험 완료시, mRNA를 qPCR을 통해 분석한 샘플로부터 정제하였다.
- [0111] 크립토크롬-2 및 CIB1의 절두 형태를 클로닝하고 전체 길이 형태의 크립토크롬-2 및 CIB1과 조합으로 시험하여, 각각의 이형이량체 쌍의 효과를 측정하였다. 크립토크롬-2 단백질의 보존된 광반응성 부위 및 전체 길이 형태 CIB1로 이루어지는 CRY2PHR 도메인의 조합은 Neurog2 mRNA 수준의 최고 상향조절을 초래하였다(YFP 샘플의 ~22 배 초과 및 미자극 동시-트랜스펙션 샘플의 ~7 배 초과). 전체 길이 크립토크롬-2(CRY2)와 전체 길이 CIB1의 조합은 낮은 절대 활성화 수준(YFP의 ~4.6 배 초과)뿐 아니라 낮은 기준선 활성화(미자극 동시-트랜스펙션 샘플에 대한 YFP의 ~1.6 배 초과)를 초래하였다. 이들 크립토크롬 단백질 페어링은 필요한 유도의 절대 수준 및 LITE 시스템의 기준선 "누수도"를 최소화하기 위한 필요성에 따라 특정 용도를 위해 선택될 수 있다.
- [0112] 가역성 및 활성화의 속도는 LITE 시스템에 대한 중요한 설계 매개변수이다. LITE 시스템의 속도를 분석하기 위해, Neurog2 TALE-CRY2PHR 및 CIB1-VP64 형태 시스템으로 이루어지는 구축물을 시험하여 이의 활성화 및 불활성화 속도를 측정하였다. 추출하기 전 짧게는 0.5 시간 내지 길게는 24 시간 동안 샘플을 자극하였다. Neurog2 발현의 상향조절은 최단 시점, 0.5 시간에서 관찰되었다(YFP 샘플에 대해 ~5 배). Neurog2 발현은 자극의 12 시간에서 최대치였다(YFP 샘플에 대해 ~19 배). 불활성화 속도를 6 시간 동안 동시-트랜스펙션 샘플을 자극하여 분석하여, 이때 자극을 중단하고, 샘플을 0 내지 12 시간 동안 배양물 중 유지시켜 mRNA가 분해되게 하였다. Neurog2 mRNA 수준은 자극 완료 후 0.5 시간에서 최대치였고(YFP 샘플에 대해 ~16 배), 그 후 상기 수준은 ~3 시간 반감기로 저하된 후 12 시간까지 기준선 수준 근처로 복구되었다.
- [0113] 제 2의 원형적 예는 인간 유전자 KLF4의 전사를 활성화시키도록 설계된 LITE 이다. 인간 KLF4의 업스트림 프로모터 부위에 위치한 서열 TTCTACTTATAAC를 표적으로서 선택하고, 상기 서열에 매치되도록 TALE을 설계하고 합성하였다. TALE 서열을 핵 국부화 신호(아미노산: SPKKKRKVEAS)를 통해 CRY2PHR에 대한 서열에 연결시켰다. 상기 기재된 동일한 CIB1-VP64 활성화 단백질을 또한 LITE 시스템의 이러한 시현에 사용하였다. HEK293FT 세포주로부터의 인간 배아 신장 세포를 2 개 벡터로 동시-트랜스펙션하였다. 벡터 발현이 가능하도록 인큐베이션한 후, 488 nm LED의 어레이로부터의 주기적 펄스 청색광에 의해 샘플을 자극하였다. 미자극 동시-트랜스펙션 샘플 및 형광 리포터 YFP 로만 트랜스펙션된 샘플을 대조군으로서 사용하였다. 각 실험의 종료시, mRNA를 qPCR을 통해 분석한 샘플로부터 정제하였다.
- [0114] LITE 시스템의 광-강도 반응을, 샘플을 증가한 광 동력(0-9 mW/cm<sup>2</sup>)으로 자극하여 시험하였다. KLF4 mRNA 수준의 상향조절이 0.2 mW/cm<sup>2</sup>의 낮은 자극에 대해 관찰되었다. KLF4 상향조절은 5 mW/cm<sup>2</sup>에서 포화되었다(YFP 샘플에 대해 2.3 배). 세포 생존도 시험을 또한 9 mW/cm<sup>2</sup> 이하의 동력에 대해 수행하였고, >98% 세포 생존도를 나타내었다. 유사하게, 자극 사용률 가변화에 대한 KLF4 LITE 반응을 시험하였다(1.6-100%). 상이한 사용률 간에 KLF4 활성화에 있어서의 차이가 관찰되지 않았는데, 이는 15 초 마다 0.25 초의 낮은 자극 패러다임이 최대 활성화를 초래한다는 것을 나타낸다.
- [0115] 본 발명은 에너지원 예컨대 전자기 조사, 음파 에너지 또는 열 에너지를 고려한다. 유리하게는, 전자기 조

사는 가시광의 성분이다. 바람직한 구현예에서, 광은 약 450 내지 약 495 nm 파장의 청색광이다. 특히 바람직한 구현예에서, 파장은 약 488 nm 이다. 또 다른 바람직한 구현예에서, 광 자극은 펄스를 통한 것이다. 광 동력은 약 0-9 mW/cm<sup>2</sup> 범위일 수 있다. 바람직한 구현예에서, 15 초 마다 0.25 초의 낮은 자극 패러다임은 최대 활성화를 초래한다.

- [0116] 본 발명은 특히, 세포에서의 관심 게놈 유전자좌의 발현을 변형시키거나 게놈 또는 비유전적 유전자좌를 교란하는 유도 방법에 관한 것이며, 이때 상기 게놈 또는 비유전적 유전자좌는 데옥시리보핵산 (DNA) 결합 폴리펩티드를 포함하는 비-자연 발생적 또는 조작된 조성물과 접촉될 수 있다.
- [0117] 본 발명의 세포는 원핵세포 또는 진핵세포, 유리하게는 동물 세포, 보다 유리하게는 포유동물 세포일 수 있다.
- [0118] 상기 폴리펩티드는 적어도 5 개 이상의 전사 활성제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체 (화학적 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도됨) 를 포함하는 DNA 결합 도메인을 포함할 수 있다. 화학적 또는 에너지 민감성 단백질 또는 이의 단편은 화학적 원천의 결합에 의해 유도시 형태적 변화를 거쳐 이를 상호작용 파트너에 결합시킬 수 있다. 폴리펩티드는 또한 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 변이체 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인을 포함할 수 있으며, 이때 화학적 또는 에너지 민감성 단백질 또는 이의 단편은 화학적 원천에 의한 유도시 상호작용 파트너에 결합할 수 있다. 상기 방법은 또한 화학적 원천을 적용하고 게놈 유전자좌의 발현이 변형되는 것을 측정하는 것을 포함할 수 있다.
- [0119] 여러 상이한 설계의 이러한 화학적 유도 시스템이 존재한다: 1. 아브시스산 (ABA) 에 의해 유도가능한 ABI-PYL 기반 시스템 (예를 들어, <http://stke.sciencemag.org/cgi/content/abstract/sigtrans;4/164/rs2> 참고), 2. 라파마이신 (또는 라파마이신 기반의 관련 화학 물질) 에 의해 유도가능한 FKBP-FRB 기반 시스템 (예를 들어, <http://www.nature.com/nmeth/journal/v2/n6/full/nmeth763.html> 참고), 3. 지베렐린 (GA) 에 의해 유도가능한 GID1-GAI 기반 시스템 (예를 들어, <http://www.nature.com/nchembio/journal/v8/n5/full/nchembio.922.htm> 참고).
- [0120] 본 발명에 의해 고려되는 또 다른 시스템은 하위-세포 국부화에 있어서의 변화를 기반으로 한 화학적 유도 시스템이다. 출원인은 또한, 폴리펩티드가 적어도 5 개 이상의 전사 활성제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체 (화학적 또는 에너지 민감성 단백질에 대한 추가의 링커인 적어도 하나 이상의 효과기 도메인에 연결된 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도됨) 를 포함하는 DNA 결합 도메인을 포함하는 시스템을 개발하였다. 상기 단백질은 화학적 또는 에너지 민감성 단백질에 대한 에너지 이동 또는 화학 물질 결합시 전체 폴리펩티드의 하위-세포 국부화 (즉, 세포질로부터의 전체 폴리펩티드의 세포 핵 내로의 이동) 에 있어서의 변화를 초래할 것이다. 하나의 하위-세포 부분 또는 세포소기관 (그의 활성이 효과기 도메인에 대한 기질의 결합으로 인해 격리됨) 으로부터의 전체 폴리펩티드의 또 다른 곳 (기질이 존재함) 으로의 이동은 전체 폴리펩티드가 이의 바람직한 기질 (즉 포유동물 핵 내의 게놈 DNA) 과 접촉될 수 있게 하며 표적 유전자 발현의 활성화 또는 억제를 초래한다.
- [0121] 이러한 유형의 시스템은 또한 효과기 도메인이 뉴클레아제인 경우 세포 내 관심 게놈 유전자좌의 절단을 유도하는데 사용될 수 있다.
- [0122] 이러한 화학적 유도 시스템에 대한 설계는 4-히드록시타목시펜 (4OHT) 에 의해 유도가능한 에스트로겐 수용체 (ER) 기반 시스템이다 (예를 들어, <http://www.pnas.org/content/104/3/1027.abstract> 참고). ERT2 로 지칭되는 에스트로겐 수용체의 돌연변이화된 리간드-결합 도메인은 4-히드록시타목시펜 결합시 세포 핵 내로 전좌된다. 2 개의 탠덤 ERT2 도메인을 유연성 펩티드 링커와 함께 연결시킨 후, 포유동물 게놈 내의 특정 서열을 표적화하는 TALE 단백질에 융합시키고 하나 이상의 효과기 도메인에 연결시켰다. 상기 폴리펩티드는 효과기 도메인에 연결된 TALE 단백질이 불활성이 되게 하는 4OHT 의 부재 하에 세포의 세포질 내에 존재할 것이다. 4OHT 의 존재 하에, 탠덤 ERT2 도메인에 대한 4OHT 의 결합은 전체 펩티드가 세포 핵 내로 이동하도록 유도하여, 효과기 도메인에 연결된 TALE 단백질이 활성이 되게 한다.
- [0123] 4-히드록시타목시펜 (4OHT) 에 의해 유도가능한 에스트로겐 수용체 (ER) 기반 시스템의 또 다른 구현예에서, 본 발명은 핵 이출 신호 (NES) 를 포함할 수 있다. 유리하게는, NES 는 LDLASLIL 의 서열을 가질 수 있다. 본 발명의 추가 구현예에서 임의의 핵 수용체, 갑상선 호르몬 수용체, 레티노산 수용체, 에스트로겐 수용체, 에스트로겐-관련 수용체, 글루코코르티코이드 수용체, 프로게스테론 수용체, 안드로겐 수용체의 임의의 자연 발

생적 또는 조작된 유도체는 ER 기반 유도 시스템과 유사한 유도 시스템에서 사용될 수 있다.

- [0124] 또 다른 유도 시스템은 에너지, 열 또는 전파에 의해 유도가능한 일시적 수용체 포텐셜 (TRP) 이온 채널을 사용하는 설계를 기반으로 한다 (예를 들어, <http://www.sciencemag.org/content/336/6081/604> 참고). 이들 TRP 패밀리는 단백질은 광 및 열을 포함하는 상이한 자극에 반응한다. 상기 단백질이 광 또는 열에 의해 활성화되는 경우, 이온 채널이 개방되며 칼슘과 같은 이온이 혈장 멤브레인 내로 유입되게 된다. 이러한 이온 유입물은 TALE 단백질 및 하나 이상의 효과기 도메인을 포함하는 폴리펩티드에 연결된 세포내 이온 상호작용 파트너에 결합할 것이며, 결합은 폴리펩티드의 하위세포 국부화의 변화를 유도하여, 전체 폴리펩티드가 세포 핵에 유입되게 한다. 핵 내로 들어오거나, 효과기 도메인에 연결된 TALE 단백질은 활성이 되며 세포 내 표적 유전자 발현을 조정한다.
- [0125] 이러한 유형의 시스템은 또한, 효과기 도메인이 뉴클레아제인 경우 세포 내 관심 게놈 유전자좌의 절단을 유도하는데 사용될 수 있다. 레이저 또는 기타 형태의 에너지원으로 광이 생성될 수 있다. 전파 형태로 전달된 에너지원으로부터의 에너지 흡수 후 열을 방출하는 나노-입자로부터 또는 에너지원으로부터의 온도 상승 결과로 열이 생성될 수 있다.
- [0126] 광 활성화가 유리한 구현예일 수 있으나, 때때로 이는 특히 광이 피부 또는 기타 기관을 관통할 수 없는 생체내 적용에 대해 불리하다. 이러한 경우, 에너지 활성화의 다른 방법, 특히, 유사한 효과를 갖는 전계 에너지 및/또는 초음파가 고려된다. 필요한 경우, 또 다른 에너지원에 의한 최대 효과를 위해 LITE 시스템의 단백질 페어링이 변경되고/되거나 변형될 수 있다.
- [0127] 전계 에너지는 생체내 조건 하에 약 1 볼트/cm 내지 약 10 k볼트/cm 의 하나 이상의 전기 펄스를 사용하여, 바람직하게는 실질적으로 당업계에서 기재된 바와 같이 투여된다. 펄스 대신 또는 이에 추가로, 전계는 연속 방식으로 전달될 수 있다. 전기 펄스는 1  $\mu$ s 내지 500 밀리초, 바람직하게는 1  $\mu$ s 내지 100 밀리초 동안 적용될 수 있다. 전계는 약 5 분 동안 펄스 방식으로 또는 연속적으로 적용될 수 있다.
- [0128] 본원에서 사용하는 바와 같이, '전계 에너지' 는 세포가 노출되는 전기 에너지이다. 바람직하게는 전계는 생체내 조건 하에 약 1 볼트/cm 내지 약 10 k볼트/cm 또는 그 이상의 강도를 갖는다 (WO97/49450 참고).
- [0129] 본원에서 사용하는 바와 같이, 용어 "전계" 는 가변적 커패시턴스 및 전압에서의 하나 이상의 폴스를 포함하며 지수형 및/또는 구형파 및/또는 변조파 및/또는 변조 구형파 형태를 포함한다. 전계 및 전기에 대한 지시대상은 세포 환경에서의 전기 전위 차이의 존재에 대한 지시대상을 포함하는 것으로 의도되어야 한다. 이러한 환경은 당업계에서 알려져 있는 바와 같이 정전기, 교류 (AC), 직류 (DC) 등에 의해 설정될 수 있다. 전계는 균일하고, 비-균일하거나, 시간 의존적 방식으로 강도 및/또는 방향에 있어서 가변적일 수 있다.
- [0130] 전계의 단일 또는 다수 적용 뿐 아니라 초음파의 단일 또는 다수 적용이 또한 임의의 순서 및 임의의 조합으로 가능하다. 초음파 및/또는 전계는 단일 또는 다수 연속 적용으로서 또는 펄스로서 (박동성 전달) 전달될 수 있다.
- [0131] 전기천공법은 외래 물질을 살아있는 세포 내로 도입하기 위해 시험관내 및 생체내 절차 모두에서 사용되어왔다. 시험관내 적용으로, 살아 있는 세포의 샘플을 먼저 관심 작용제와 혼합하고 평행 플레이트와 같은 전극 사이에 둔다. 이후, 전극이 전계를 세포/이식물 혼합물에 적용한다. 시험관내 전기천공법을 수행하는 시스템의 예는 Electro Cell Manipulator ECM600 제품 및 Electro Square Porator T820 (둘 모두 BTX Division of Genetronics, Inc 사제 (미국 특허 번호 No 5,869,326 참고)) 을 포함한다.
- [0132] 공지된 전기천공법 기술 (시험관내 및 생체내 모두) 은 짧은 고전압 펄스를 처리 부위 근처에 위치한 전극에 적용함으로써 작용한다. 전극 사이에서 생성된 전계는 세포 멤브레인은 일시적으로 다공성이 되게 하여, 관심 작용제가 세포에 유입되게 한다. 공지된 전기천공법 적용에서, 이러한 전계는 1000 V/cm 에 따라, 약 100 .mu.s 지속기간의 단일 구형파 펄스를 포함한다. 이러한 펄스는 예를 들어 Electro Square Porator T820 의 공지된 적용으로 생성될 수 있다.
- [0133] 바람직하게는, 전계는 시험관내 조건 하에 약 1 V/cm 내지 약 10 kV/cm 의 강도를 갖는다. 따라서, 전계는 1 V/cm, 2 V/cm, 3 V/cm, 4 V/cm, 5 V/cm, 6 V/cm, 7 V/cm, 8 V/cm, 9 V/cm, 10 V/cm, 20 V/cm, 50 V/cm, 100 V/cm, 200 V/cm, 300 V/cm, 400 V/cm, 500 V/cm, 600 V/cm, 700 V/cm, 800 V/cm, 900 V/cm, 1 kV/cm, 2 kV/cm, 5 kV/cm, 10 kV/cm, 20 kV/cm, 50 kV/cm 또는 그 이상의 강도를 가질 수 있다. 보다 바람직하게는 시험관내 조건 하에 약 0.5 kV/cm 내지 약 4.0 kV/cm 이다. 바람직하게는 전계는 생체내 조건 하에 약 1 V/cm 내지 약 10 kV/cm 의 r아도를 갖는다. 그러나, 전계 강도는 표적 부위에 전달된 펄스의 수가 증가될

때 감소될 수 있다. 따라서, 낮은 전계 강도에서의 전계의 박동성 전달이 예상된다.

- [0134] 바람직하게는 전계의 적용은 다수 펄스 예컨대 동일 강도 및 커패시턴스의 이중 펄스 또는 가변적 강도 및/또는 커패시턴스의 순차적 펄스의 형태로의 적용이다. 본원에서 사용하는 바와 같이, 용어 "펄스" 는 가변적 커패시턴스 및 전압에서의 하나 이상의 전기 펄스를 포함하며 지수형 및/또는 구형파 및/또는 변조파/구형파 형태를 포함한다.
- [0135] 바람직하게는 전기 펄스는 지수형파 형태, 구형파 형태, 변조파 형태 및 변조 구형파 형태에서 선택되는 파형으로서 전달된다.
- [0136] 바람직한 구현에는 낮은 전압에서 직류를 이용한다. 따라서, 출원인은 1V/cm 내지 20V/cm 의 전계 강도에서, 100 밀리초 이상, 바람직하게는 15 분 이상 동안 세포, 조직 또는 조직 덩어리에 적용되는 전계의 사용을 개시하고 있다.
- [0137] 초음파는 유리하게는 약 0.05 W/cm<sup>2</sup> 내지 약 100 W/cm<sup>2</sup> 의 동력 수준에서 투여된다. 진단용 또는 치료용 초음파가 사용될 수 있거나, 이의 조합물이 사용될 수 있다.
- [0138] 본원에서 사용하는 바와 같이, 용어 "초음파" 는 인간 청력의 범위를 넘는 높은 주파수의 기계적 진동으로 이루어지는 에너지 형태를 의미한다. 초음파 스펙트럼의 낮은 주파수 한계는 일반적으로 약 20 kHz 에서 선택될 수 있다. 초음파의 대부분의 진단용 적용은 1 및 15 MHz' 범위의 주파수를 이용한다 (Ultrasonics in Clinical Diagnosis, P. N. T. Wells, ed., 2nd. Edition, Publ. Churchill Livingstone [Edinburgh, London & NY, 1977]로부터).
- [0139] 초음파는 진단적 및 치료적 적용 모두에서 사용되어왔다. 진단 도구로서 사용되는 경우 ("진단용 초음파"), 750 mW/cm<sup>2</sup> 이하의 에너지 밀도가 사용되어왔으나, 초음파는 통상 약 100 mW/cm<sup>2</sup> 이하의 에너지 밀도 범위 (FDA 권고사항) 로 사용된다. 물리치료에서, 초음파는 통상 약 3 내지 4 W/cm<sup>2</sup> 까지의 범위 (WHO 권고사항) 로 에너지원으로서 사용된다. 다른 치료적 적용에서, 높은 강도의 초음파, 예를 들어 100 W/cm 에서 1 kW/cm<sup>2</sup> 이하 (또는 더 높음) 에서의 HIFU 를 단기간 동안 이용할 수 있다. 본 명세서에서 사용하는 바와 같은 용어 "초음파" 는 진단적, 치료적 및 초점성 초음파를 포함하는 것으로 의도된다.
- [0140] 초점성 초음파 (FUS) 는 열 에너지가 침습성 프로브 없이 전달되게 한다 (Morocz et al 1998 Journal of Magnetic Resonance Imaging Vol.8, No. 1, pp.136-142 참고). 또 다른 형태의 초점성 초음파는 고강도 초점성 초음파 (HIFU) 이며 이는 [Moussatov et al in Ultrasonics (1998) Vol.36, No.8, pp.893-900] 및 [TranHuuhue et al in Acustica (1997) Vol.83, No.6, pp.1103-1106] 에 의해 검토된다.
- [0141] 바람직하게는, 진단용 초음파 및 치료용 초음파의 조합이 이용된다. 이러한 조합은 제한되는 것으로 의도되지 않으나, 당업자는 임의의 다양한 조합의 초음파가 사용될 수 있다는 것을 이해할 것이다. 추가적으로, 에너지 밀도, 초음파 주파수 및 노출 기간은 가변적일 수 있다.
- [0142] 바람직하게는 초음파 에너지원에 대한 노출은 약 0.05 내지 약 100 Wcm<sup>-2</sup> 의 동력 밀도에서 이루어진다. 보다 더 바람직하게는, 초음파 에너지원에 대한 노출은 약 1 내지 약 15 Wcm<sup>-2</sup> 의 동력 밀도에서 이루어진다.
- [0143] 바람직하게는 초음파 에너지원에 대한 노출은 약 0.015 내지 약 10.0 MHz 의 주파수에서 이루어진다. 보다 바람직하게는 초음파 에너지원에 대한 노출은 약 0.02 내지 약 5.0 MHz 또는 약 6.0 MHz 의 주파수에서 이루어진다. 가장 바람직하게는, 초음파는 3 MHz 의 주파수에서 적용된다.
- [0144] 바람직하게는 노출은 약 10 밀리초 내지 약 60 분의 기간 동안이다. 바람직하게는 노출은 약 1 초 내지 약 5 분의 기간 동안이다. 보다 바람직하게는, 초음파는 약 2 분 동안 적용된다. 그러나 와해될 특정 표적 세포에 따라서, 노출은 장기간, 예를 들어 15 분 동안일 수 있다.
- [0145] 유리하게는, 표적 세포는 약 0.05 Wcm<sup>-2</sup> 내지 약 10 Wcm<sup>-2</sup> 의 음향 동력 밀도에서 약 0.015 내지 약 10 MHz 범위의 주파수로 초음파 에너지원에 노출된다 (WO 98/52609 참고). 그러나 예를 들어, 100 Wcm<sup>-2</sup> 초과의 음향 동력 에너지에서, 감소된 기간, 예를 들어 밀리초 범위 또는 그 이하 동안 1000 Wcm<sup>-2</sup> 의 음향 동력 에너지에서 초음파 에너지원에 대한 노출과 같은 대안물이 또한 가능하다.

- [0146] 바람직하게는 초음파의 적용은 다수 펄스의 형태이며; 따라서 연속파 및 펄스형파 (초음파의 박동성 전달) 모두 임의의 조합으로 이용될 수 있다. 예를 들어, 연속파 초음파가 적용된 후 펄스형파 초음파가 적용될 수 있으며, 그 반대도 가능하다. 이는 임의 횡수로, 임의 순서 및 조합으로 반복될 수 있다. 펄스형파 초음파는 연속파 초음파의 배경에 반하여 적용될 수 있으며, 임의 수의 펄스가 임의 수의 군에서 사용될 수 있다.
- [0147] 바람직하게는, 초음파는 펄스형파 초음파를 포함할 수 있다. 매우 바람직한 구현예에서, 초음파는  $0.7 \text{ Wcm}^{-2}$  또는  $1.25 \text{ Wcm}^{-2}$  의 동력 밀도에서 연속파로서 적용된다.
- [0148] 초음파의 사용은 광파 같이, 표적에 정확히 초점화될 수 있으므로 유리하다. 또한, 초음파는 광파 달리 조직 내로 보다 깊숙히 초점화될 수 있으므로 유리하다. 따라서, 이는 전체-조직 침투 (예컨대 간엽에 제한되지 않는) 또는 전체 기관 (예컨대 전체 간 또는 전체 근육, 예컨대 심장에 제한되지 않는) 치료요법에 보다 양호하게 적합하다. 또 다른 중요한 이점은 초음파가 광범위한 진단적 및 치료적 적용에서 사용되는 비-침습성 자극이라는 것이다. 예를 들어, 초음파는 의료 영상화 기술, 및 추가적으로, 정형외과 치료요법에 있어서 널리 공지된다. 추가로, 척추동물 대상에 대한 초음파 적용에 적합한 기구가 광범위하게 이용가능하며 그의 용도가 당업계에 널리 공지되어있다.
- [0149] LITE 의 신속한 전사 반응 및 내인성 표적화는 전사 역학의 연구에 이상적인 시스템에 기여한다. 예를 들어, LITE 는 표적 유전자의 유도된 발현시 mRNA 스플라이스 변이체 생성의 역학 연구에 사용될 수 있다. 전사 주기의 다른 한편, mRNA 분해 연구는 종종, 과잉 유전자에서의 발현 수준 변화를 일으키는 강력한 세포외 자극에 대한 반응으로 수행된다. LITE 는 내인성 표적의 전사를 가역적으로 유도하는데 이용될 수 있으며, 그 후 자극이 중단될 수 있고 고유 표적의 분해 속도가 추적될 수 있다.
- [0150] LITE 의 시간적 정확도는 실험적 개입과 맞물려 시간 유전적 조절에 대한 동력을 제공할 수 있다. 예를 들어, 장기간 강화작용 (LTP) 에서 예상된 관여를 갖는 표적은 기관형 또는 해리성 뉴런 배양물에 있어서, 그러나 LTP 를 유도하는 자극 동안에만 조정되어, 세포의 정상 발생의 간섭이 방지될 수 있다. 유사하게, 질환 표현형을 나타내는 세포 모델에서, 특정 치료요법의 유효성에 관여되는 것으로 예상되는 표적은 처리 동안에만 조정될 수 있다. 반대로, 유전적 표적은 병리적 자극 동안에만 조정될 수 있다. 외부 실험 자극에 대한 유전적 신호의 시기가 관련되는 임의 수의 실험은 LITE 조정의 효용으로부터 잠재적으로 유익할 수 있다.
- [0151] 생체내 맥락은 유전자 발현 제어를 위한 LITE 의 용도에 대해 동등하게 많은 기회를 제공한다. 상기 언급한 바와 같이, 광유도성은 이전에 성취불가능한 공간적 정확도에 대한 가능성을 제공한다. 광전극 기술의 발전을 이용하여, 자극 광 섬유 리드 (lead) 는 정확한 뇌 부위 내에 위치할 수 있다. 자극 부위 크기는 광 강도에 의해 조정될 수 있다. 이는 미국 가출원 번호 61/671,615 의 바이러스 벡터 또는 분자 슬레드 (sled) 를 통한 LITE 의 전달과 함께 수행될 수 있거나, 트랜스제닉 LITE 동물이 이용가능해진 경우, 정확한 뇌 부위에서의 유전자 발현을 여전히 조정가능하게 하면서 바이러스의 사용을 없앨 수 있다. LITE 는 투명한 유기체, 예컨대 고정 제브라피시에서 사용되어 매우 정확한 레이저 유도 국부 유전자 발현 변화를 가능하게 할 수 있다.
- [0152] 본 발명은 또한 CRISPR/Cas 시스템을 사용하는 멀티플렉스 게놈 조작을 고려한다. 정상적인 유전적 변이체 및 요소의 기능적 해명은 정확한 게놈 편집 기술을 필요로 한다. 유형 II 원핵세포 CRISPR (군집화된 규칙적 공간화 단길이 군집화 규칙적 공간화 단길이 회귀성 반복물) 적용 면역계가 RNA-안내 부위-특이적 DNA 절단을 촉진시키는 것으로 나타난 바 있다. 출원인은 2 개 상이한 유형 II CRISPR 시스템을 조작하였고 Cas9 뉴클레아제가 단길이 RNA 에 의해 유도되어 인간 및 마우스 세포에서 내인성 게놈 유전자좌에서의 정확한 절단을 유도할 수 있다는 것을 입증한다. Cas9 는 또한 최소 돌연변이유발 활성을 갖는 상동성-유도 복구가 촉진되도록 절단 효소로 전환될 수 있다. 마지막으로, 다수의 안내 서열은 단일 CRISPR 어레이 내로 인코딩되어 포유동물 게놈 내 여러 부위의 동시 편집이 가능하게 하는데, 이는 CRISPR 기술의 용이한 프로그래밍가능성 및 광범위한 적용가능성을 입증한다.
- [0153] 일반적으로, "CRISPR 시스템" 은 전사체 및 CRISPR-연관 ("Cas") 유전자의 활성을 유도하거나 이의 발현에 관여된 다른 요소, 예컨대 Cas 유전자 인코딩 서열, tracr (트랜스-활성화 CRISPR) 서열 (예를 들어 tracrRNA 또는 활성 부분 tracrRNA), tracr-mate 서열 ("직접 반복물" 및 내인성 CRISPR 시스템의 맥락에 있어서 tracrRNA-처리 부분 직접 반복물 포함), 안내 서열 (또한 내인성 CRISPR 시스템의 맥락에 있어서 "스페이스" 로서 지칭함), 또는 CRISPR 유전자좌로부터의 기타 서열 및 전사체를 집합적으로 의미한다. 일부 구현예에서, CRISPR 시스템의 하나 이상의 요소는 유형 I, 유형 II 또는 유형 III CRISPR 시스템에서 유래한다. 일부 구현예에서, CRISPR 시스템의 하나 이상의 요소는 내인성 CRISPR 시스템을 포함하는 특정 유기체, 예컨대 스트렙토코쿠스 피

오게네스에서 유래한다. 일반적으로, CRISPR 시스템은 표적 서열 부위에서의 CRISPR 복합체 형성을 촉진하는 요소를 특징으로 한다 (또한 내인성 CRISPR 시스템의 맥락에서 프로토스페이서로서 지칭함). CRISPR 복합체 형성의 맥락에서, "표적 서열" 은 안내 서열이 상보성을 갖도록 설계된 서열을 의미하며, 이때 표적 서열과 안내 서열 사이의 하이브리드화가 CRISPR 복합체의 형성을 촉진한다. 표적 서열은 임의의 폴리뉴클레오티드, 예컨대 DNA 또는 RNA 폴리뉴클레오티드를 포함할 수 있다. 일부 구현예에서, 표적 서열은 세포의 핵 또는 세포질 내에 위치한다.

[0154] 통상, 내인성 CRISPR 시스템의 맥락에서, CRISPR 복합체 (하나 이상의 Cas 단백질과 복합화되며 표적 서열에 하이브리드화된 안내 서열을 포함함) 의 형성은 표적 서열 내 또는 표적 서열 근처 (예를 들어 표적 서열로부터 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50 또는 그 이상의 염기쌍) 하나 또는 둘 모두의 가닥의 절단을 초래한다. 이론에 속박되지 않고, 모든 또는 일부의 tracr 서열은 또한 예컨대 안내 서열에 작동가능하게 연결되는 tracr mate 서열의 전부 또는 일부에 대한 하이브리드화에 의해 CRISPR 복합체의 일부를 형성할 수 있다. 일부 구현예에서, CRISPR 시스템의 하나 이상의 요소의 발현을 유도하는 하나 이상의 벡터가 숙주 세포 내로 도입되어, CRISPR 시스템의 요소의 발현이 하나 이상의 표적 부위에서 CRISPR 복합체 형성을 유도한다. 예를 들어, Cas 효소, tracr-mate 서열에 연결된 안내 서열, 및 tracr 서열은 각각 개별 벡터 상 개별 조절 요소에 작동가능하게 연결될 수 있다. 대안적으로, 동일하거나 상이한 조절 요소로부터 발현된 둘 이상의 요소는 단일 벡터 내에서 조합될 수 있으며, CRISPR 시스템의 임의 성분을 제공하는 하나 이상의 추가적 벡터는 제 1 벡터에 포함되지 않는다. 단일 벡터 내에서 조합되는 CRISPR 시스템 요소는 임의의 적합한 배향으로 배열될 수 있으며, 예를 들어 하나의 요소가 제 2 의 요소 (의 "업스트림") 에 대해 5' 또는 제 2 의 요소 (의 "다운스트림") 에 대해 3' 로 배열될 수 있다. 하나의 요소의 코딩 서열은 제 2 의 요소의 코딩 서열의 동일 또는 반대 가닥 상에 위치할 수 있고, 동일 또는 반대 방향으로 배향될 수 있다. 일부 구현예에서, 단일 프로모터는 안내 서열, tracr mate 서열 (임의로는 안내 서열에 작동가능하게 연결됨) 및 하나 이상의 인트론 서열 (예를 들어 각각 상이한 인트론, 하나 이상의 인트론 중 둘 이상 또는 모두 단일 인트론) 내에 포매된 tracr 서열 중 하나 이상 및 CRISPR 효소를 인코딩하는 전사체의 발현을 유도한다. 일부 구현예에서, CRISPR 효소, 안내 서열, tracr mate 서열 및 tracr 서열은 동일 프로모터에 작동가능하게 연결되며 이로부터 발현된다.

[0155] 일부 구현예에서, 벡터는 하나 이상의 삽입 부위, 예컨대 제한 엔도뉴클레아제 인지 서열 (또한 "클로닝 부위" 로서 지칭함) 을 포함한다. 일부 구현예에서, 하나 이상의 삽입 부위 (예를 들어 약 또는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 이상의 삽입 부위) 는 하나 이상의 벡터의 하나 이상의 서열 요소의 업스트림 및/또는 다운스트림에 위치한다. 일부 구현예에서, 벡터는 tracr mate 서열의 업스트림, 및 임의로는 tracr mate 서열에 작동가능하게 연결된 조절 요소의 다운스트림 삽입 부위를 포함하여, 안내 서열의 삽입 부위 내로의 삽입 후 발현시 안내 서열이 진핵세포에서의 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 유도한다. 일부 구현예에서, 벡터는 둘 이상의 삽입 부위를 포함하며, 각각의 삽입 부위는 2 개의 tracr mate 서열 사이에 위치하여 각 부위에서의 안내 서열 삽입을 가능하게 한다. 이러한 배열에서, 둘 이상의 안내 서열은 단일 안내 서열의 둘 이상의 카피, 둘 이상의 상이한 안내 서열, 또는 이들의 조합을 포함할 수 있다. 다수의 상이한 안내 서열을 사용하는 경우, 단일 발현 구축물을 사용하여 CRISPR 활성을 세포 내 다수의 상이한, 상응하는 표적 서열에 대해 표적화할 수 있다. 예를 들어, 단일 벡터는 약 또는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 이상의 안내 서열을 포함할 수 있다. 일부 구현예에서, 약 또는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 이상의 이러한 안내-서열-함유 벡터가 제공될 수 있으며, 임의로는 세포에 전달된다.

[0156] 일부 구현예에서, 벡터는 Cas 단백질과 같은 CRISPR 효소를 인코딩하는 효소-코딩 서열에 작동가능하게 연결된 조절 요소를 포함한다. Cas 단백질의 비제한적 예는 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (또한 Csn1 및 Csx12 으로 공지됨), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, 이의 동족체, 또는 이의 변형된 형태를 포함한다. 일부 구현예에서, 미변형된 CRISPR 효소는 DNA 절단 활성을 갖는다 (예컨대 Cas9). 일부 구현예에서, CRISPR 효소는 표적 서열의 위치에서, 예컨대 표적 서열 내 및/또는 표적 서열의 보체 내 하나 또는 둘 모두의 가닥의 절단을 유도한다. 일부 구현예에서, CRISPR 효소는 표적 서열의 첫 번째 또는 마지막 뉴클레오티드로부터 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500 이상의 염기쌍 내 하나 또는 둘 모두의 가닥의 절단을 유도한다. 일부 구현예에서, 벡터는 상응하는 야생형 효소에 대하여 돌연변이화되어 상기 돌연변이화된 CRISPR 효소가 표적 서열을 함유하는 표적 폴리뉴클레오티드의 하나 또는 둘 모두의 가닥을 절단하는 능력이 소실되는, CRISPR 효소를 인코딩한다. 예를 들어, S. 피오게네스로부터의 Cas9 의 RuvC

I 효소 도메인에서의 아스파르테이트-알라닌 치환 (D10A) 은 두 가닥 모두를 절단하는 뉴클레아제로부터의 Cas9 를 니카아제 (nickase) (단일 가닥을 절단함) 로 전환시킨다. Cas9 가 니카아제가 되도록 하는 돌연변이의 다른 예는 비제한적으로, H840A, N854A 및 N863A 를 포함한다. 추가예로서, Cas9 의 둘 이상의 촉매 도메인 (RuvC I, RuvC II 및 RuvC III) 은, 모든 DNA 절단 활성이 실질적으로 결핍되는 돌연변이화된 Cas9 를 생성시키도록 돌연변이화될 수 있다. 일부 구현예에서, D10A 돌연변이는 H840A, N854A 또는 N863A 돌연변이 중 하나 이상과 조합되어 모든 DNA 절단 활성이 실질적으로 결핍되는 Cas9 효소를 생성시킨다. 일부 구현예에서, CRISPR 효소는 돌연변이화된 효소의 DNA 절단 활성이 이의 비-돌연변이화 형태에 대하여 약 25%, 10%, 5%, 1%, 0.1%, 0.01% 이하인 경우 모든 DNA 절단 활성이 실질적으로 결핍되는 것으로 고려된다.

[0157] 일부 구현예에서, CRISPR 효소를 인코딩하는 효소 코딩 서열은 특정 세포, 예컨대 진핵세포에서의 발현을 위해 최적화된 코돈이다. 진핵세포는 특정 유기체, 예컨대 인간, 마우스, 랫트, 토끼, 개 또는 비-인간 영장류를 비제한적으로 포함하는 포유동물의 것들 또는 이로부터 유래한 것들일 수 있다. 일반적으로 코돈 최적화는, 선천적 서열의 하나 이상의 코돈 (예를 들어 약 또는 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 50 이상의 코돈) 을 선천적 아미노산 서열을 유지하면서 숙주 세포의 유전자에서 보다 자주 또는 가장 자주 사용되는 코돈으로 대체함으로써, 관심 숙주 세포에서의 증강된 발현을 위해 핵산 서열을 변형시키는 프로세스를 의미한다. 다양한 종류는 특정 아미노산의 특정 코돈에 대한 특정한 편향을 나타낸다. 코돈 편향 (유기체 사이의 코돈 사용빈도에 있어서의 차이) 은 종종 메신저 RNA (mRNA) 의 번역 효율과 상호관련되는데, 이는 다른 것들 중에서도, 번역되는 코돈의 특성 및 특정 트랜스퍼 RNA (tRNA) 분자의 이용가능성에 의존적인 것으로 여겨진다. 세포 내에서의 선택된 tRNA 의 우세는 일반적으로 펩티드 합성에서 가장 자주 사용된 코돈의 반영이다. 따라서, 유전자는 코돈 최적화를 기반으로 주어진 유기체에서의 최적 유전자 발현에 맞추어질 수 있다. 코돈 사용빈도 표는 예를 들어 [www.kazusa.or.jp/codon/](http://www.kazusa.or.jp/codon/) 에서 이용가능한 "코돈 사용빈도 데이터베이스" (2002 년 7 월 9 일 방문) 에서 용이하게 이용가능하며, 이들 표는 수많은 방식으로 적합화될 수 있다. [Nakamura, Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28:292 (2000)] 를 참고한다. 특정 숙주 세포에서의 발현을 위한 특정 서열을 최적화하는 코돈에 대한 컴퓨터 알고리즘이 또한 이용가능하며, 예컨대 Gene Forge (Aptagen; Jacobus, PA) 가 또한 이용가능하다. 일부 구현예에서, CRISPR 효소를 인코딩하는 서열에서의 하나 이상의 코돈 (예를 들어 1, 2, 3, 4, 5, 10, 15, 20, 25, 50 이상, 또는 모든 코돈) 은 특정 아미노산에 대해 가장 자주 사용된 코돈에 상응한다.

[0158] 일부 구현예에서, 백터는 하나 이상의 핵 국부화 서열 (NLS), 예컨대 약 또는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 이상의 NLS 를 포함하는 CRISPR 효소를 인코딩한다. 일부 구현예에서, CRISPR 효소는 아미노-말단에서 또는 근처에서 약 또는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 이상의 NLS 를, 카르복시-말단에서 또는 근처에서 약 또는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 이상의 NLS 를, 또는 이들의 조합을 포함한다 (예를 들어 아미노-말단에서 하나 이상의 NLS 및 카르복시-말단에서 하나 이상의 NLS). 하나 초과 NLS 가 존재하는 경우, 각각은 서로에게 독립적으로 선택될 수 있어, 단일 NLS 가 하나 초과 카피로 존재할 수 있고/있거나 하나 초과 다른 NLS 와 조합으로 하나 이상의 카피로 존재할 수 있다. 일부 구현예에서, NLS 의 가장 가까운 아미노산이 N- 또는 C-말단으로부터의 폴리펩티드 사슬을 따라 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 개 이상의 아미노산 내에 있는 경우, NLS 는 N- 또는 C-말단 근처에 있는 것으로 고려된다. NLS 의 비제한적 예는 하기에서 유래한 NLS 서열을 포함한다: 아미노산 서열 PKKKRKV 를 갖는 SV40 바이러스 거대 T-항원의 NLS; 뉴클레오플라스민으로부터의 NLS (예를 들어 서열 KRPAATKKAGQAKKKK 를 갖는 뉴클레오플라스민 이연 NLS); 아미노산 서열 PAAKRVKLD 또는 QRRLNELKRSP 를 갖는 c-myc NLS; 서열 NQSSNFGPMKGGNFGGRSSGPYGGGGYFAKPRNQGGY 를 갖는 hRNP1 M9 NLS; 임포틴-알파로부터의 IBB 도메인의 서열 RMRIZFKNGKDTAELRRRRRVEVSLERKAKKDEQILKRRNV; 자궁근종 T 단백질의 서열 VSRKRPRP 및 PPKKARED; 인간 p53 의 서열 POPKKKPL; 마우스 c-ab1 IV 의 서열 SALIKKKKMAP; 인플루엔자 바이러스 NS1 의 서열 DRLRR 및 PKQKKRK; 간염 바이러스 델타 항원의 서열 RKLKKIKKL; 마우스 Mx1 단백질의 서열 REKKKFLKRR; 인간 폴리(ADP-리보오스) 중합효소의 서열 KRKGDEVDGVEAKKSKK; 및 스테로이드 호르몬 수용체 (인간) 글루코코르티코이드의 서열 RKCLQAGMNEARKTKK.

[0159] 일반적으로, 하나 이상의 NLS 는 진핵세포의 핵에서 검출가능한 양으로 CRISPR 효소를 축적하는 것을 유도하기에 충분한 강도의 것이다. 일반적으로, 핵 국부화 활성의 강도는 CRISPR 효소에서의 NLS 의 수, 특히 사용한 NLS(들), 또는 이들 인자의 조합에서 유래할 수 있다. 핵 내 축적의 검출은 임의의 적합한 기술에 의해 수행될 수 있다. 예를 들어, 검출가능 마커는 CRISPR 효소에 융합될 수 있어, 핵의 위치를 검출하기 위한 수단과의 조합으로 (예를 들어 핵에 대한 특이적 염색 예컨대 DAPI) 세포 내 위치가 가시화될 수 있다. 세포 핵은 또한 세포로부터 단리될 수 있으며, 이의 내용물은 단백질 검출을 위한 임의의 적합한 프로세스, 예컨대 면역조직화학, 웨스턴 블롯 또는 효소 활성 검정에 의해 분석될 수 있다. 핵 내 축적은 또한 예컨대

CRISPR 효소 또는 복합체에 노출되지 않은 대조군, 또는 하나 이상의 NLS 가 결핍되는 CRISPR 효소에 노출된 대조군과 비교하여 CRISPR 복합체 형성 효과에 대한 검정 (예를 들어 DNA 절단 또는 표적 서열에서의 돌연변이에 대한 검정, 또는 CRISPR 복합체 형성 및/또는 CRISPR 효소 활성화에 의해 영향받은 변경된 유전자 발현 활성화에 대한 검정) 에 의해 간접적으로 검출될 수 있다.

[0160] 본 발명의 또 다른 구현예에서, 본 발명은 유도 Cas9 를 포함할 수 있는 유도 CRISPR 에 관한 것이다.

[0161] CRISPR 시스템은 하기: I. CRISPR/Cas 시스템 키메라 RNA (chiRNA) 폴리뉴클레오티드 서열에 작동가능하게 연결된 제 1 조절 요소 (이때, 상기 폴리뉴클레오티드 서열은 (a) 진핵세포에서의 표적 서열에 하이브리드화할 수 있는 안내 서열, (b) tracr mate 서열 및 (c) tracr 서열을 포함할 수 있음), 및 II. 하나 이상의 핵 국부화 서열을 포함할 수 있는 CRISPR 효소를 인코딩하는 효소-코딩 서열에 작동가능하게 연결된 제 2 조절 요소를 포함할 수 있는 하나 이상의 벡터를 포함할 수 있는 벡터 시스템 내에서 인코딩될 수 있으며, 여기서 (a), (b) 및 (c) 는 5' 에서 3' 배향으로 배열되고, 성분 I 및 II 는 시스템의 동일하거나 상이한 벡터 상에 위치하고, 전사되는 경우, tracr mate 서열은 tracr 서열에 하이브리드화되고 안내 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 유도하고, CRISPR 복합체는 (1) 표적 서열에 대해 하이브리드화되는 안내 서열, 및 (2) tracr 서열에 대해 하이브리드화되는 tracr mate 서열로 복합화되는 CRISPR 효소를 포함할 수 있고, CRISPR 효소를 인코딩하는 효소 코딩 서열은 추가로 이중 기능적 도메인을 인코딩한다.

[0162] 유리한 구현예에서, 유도 Cas9 를 렌티바이러스에서 제조할 수 있다. 예를 들어, 도 61 은 Tet Cas9 벡터 설계를 나타내고 도 62 는 293FT 세포에서의 벡터 및 EGFP 발현을 나타낸다. 특히, 유도 테트라사이클린 시스템은 유도 CRISPR 에 대해 고려된다. 벡터는 [Markusic et al, Nucleic Acids Research, 2005, Vol. 33, No. 6 e63] 에서 기재된 바와 같이 설계될 수 있다. 테트라사이클린-의존적 전사 조절 시스템은 테트라사이클린 억제제 단백질 (TetR) 및 특이적 DNA-결합 부위, 테트라사이클린 오퍼레이터 서열 (TetO) 로 이루어지는 대장균 Tn10 테트라사이클린 저항성 오퍼레이터를 기반으로 한다. 테트라사이클린의 부재 하, TetR 은 이량체화하며 TetO 에 결합한다. 테트라사이클린 또는 독시사이클린 (테트라사이클린 유도체) 은 TetR 에 결합하고 형태적 변화를 유도할 수 있어, TetO 로부터의 이의 해리를 초래한다. 유리한 구현예에서, 벡터는 CRISPR 복합체의 조절된 발현을 위해 자가조절된 rtTA 발현을 갖는 단일 Tet-On 렌티바이러스 벡터일 수 있다. 테트라사이클린 또는 독시사이클린은 유도 CRISPR 복합체 활성화를 위해 고려될 수 있다.

[0163] 또 다른 구현예에서, 큐메이트 유전자-스위치 시스템은 유도 CRISPR 에 대해 고려된다. 유사한 시스템은 [Mullick et al., BMC Biotechnology 2006, 6:43 doi: 10.1186/1472-6750-6-43] 에서 기재된 바와 같다. 유도 큐메이트 시스템은 3 가지 상이한 전략을 사용하여 포유동물 세포에서의 유전자 발현을 조절하기 위한 박테리아 오페론 (cmt 및 cym) 의 조절 메커니즘을 포함한다. 억제제 구성에서, 조절은 강력한 항시성 프로모터의 다운스트림에 위치한 오퍼레이터 부위 (CuO) 에 대한 억제제 (CymR) 의 결합에 의해 매개된다. 큐메이트, 소분자의 첨가로, 억제가 완화된다. 전사촉진제 구성에서, CymR 과 VP16 의 활성화 도메인과의 융합에 의해 형성된 키메라 전사촉진제 (cTA) 단백질은 CMV 최소 프로모터의 업스트림에 위치한 여러 카피의 CuO 에 결합하는 경우 전사를 활성화시킬 수 있다. 큐메이트 첨가는 DNA 결합을 없애며, 따라서 cTA 에 의한 전사촉진을 없앤다. 본 발명은 또한 역 큐메이트 활성화제 (rcTA) 를 고려하는데, 이는 큐메이트의 부재 하에서보다는 존재 하에 전사를 활성화시킨다. CymR 은 강력한 프로모터, 예컨대 CMV 로부터의 발현을 가역적으로 차단하는 억제제로서 사용될 수 있다. 큐메이트 억제제/오퍼레이터 시스템의 특정 양태는 미국 특허 번호 7745592 에서 추가로 기재된다.

[0164] 다른 유도 시스템은 예컨대 비제한적으로, 중금속 [Mayo KE et al, Cell 1982, 29:99-108; Searle PF et al, Mol Cell Biol 1985, 5: 1480-1489 and Brinster RL et al, Nature (London) 1982, 296:39-42], 스테로이드 호르몬 [Hynes NE et al, Proc Natl Acad Sci USA 1981, 78:2038-2042; Klock G et al, Nature (London) 1987, 329:734-736 and Lee F et al, Nature (London) 1981, 294:228-232.], 열 충격 [Nouer L: Heat Shock Response. Boca Raton, FL: CRC; 1991] 및 개발된 기타 시약 [Mullick A, Massie B: Transcription, translation and the control of gene expression. In Encyclopedia of Cell Technology Edited by: Speir RE. Wiley; 2000: 1140-1164 and Fussenegger M, . Biotechnol Prog 2001, 17: 1-51] 에 의한 조절을 고려한다. 그러나, 이러한 유도 포유동물 프로모터로는 "오프 (off)" 상태의 "누수성" 및 유도물질 (열 충격, 중금속, 글루코코르티코이드 등) 의 다면발현성 효과와 같은 한계가 존재한다. 포유동물 세포에서 세포적 프로세스로의 간섭을 감소시키기 위한 시도로, 곤충 호르몬 (엑디손) 의 사용이 제시된 바 있다 [No D et al., Proc Natl Acad Sci USA 1996, 93:3346-3351]. 또 다른 명쾌한 시스템은 유도물질로서 라파마이신을 사용하나 [Rivera VM et al, Nat Med 1996, 2: 1028-1032], 면역억제제로서의 라파마이신의 역할은 이의 생체내 용

도에 대한 주요 한계이며, 따라서 유전자 발현 제어를 위해 생물학적으로 불활성인 화합물을 발견할 필요가 있었다 [Saez E et al., Proc Natl Acad Sci USA 2000, 97: 14512-14517].

- [0165] 본 발명은 또한 본 발명의 폴리펩티드를 인코딩하는 핵산을 포함한다. 핵산은 프로모터, 유리하게는 인간 시냅신 I 프로모터 (hSyn) 를 포함할 수 있다. 특히 유리한 구현예에서, 핵산은 아데노 연관 바이러스 벡터 (AAV) 내로 패키징될 수 있다.
- [0166] 또한 본 발명에 의해 고려되는 것은 하나 초과인 아데노바이러스 혈청형으로부터의 바이러스 준입자 (subviral particle) 를 포함할 수 있는 재조합 아데노바이러스 및 재조합 벡터이다. 예를 들어, 아데노바이러스 벡터가 특정 조직 또는 세포 유형에 대해 변경된 향성 (tropism) 을 나타낼 수 있다는 것이 공지되어 있으며 (Havenga, M.J.E. et al., 2002), 따라서, 상이한 아데노바이러스 캡시드, 즉 섬유 또는 각종 아데노바이러스 혈청형으로부터의 펜톤 단백질의 혼합 및 매칭이 유리할 수 있다. 섬유 및 펜톤을 포함하는 아데노바이러스 캡시드의 변형은 미변형된 아데노바이러스와 상이한 향성을 갖는 아데노바이러스 벡터를 초래할 수 있다. 표적 세포를 감염시키는 능력에 있어서 최적화되고 변형되는 아데노바이러스 벡터는 치료적 및 예방적 용량이 유의하게 감소될 수 있게 하여, 감소된 국부적 및 전과 독성을 초래한다.
- [0167] 바이러스 벡터 유전자 전달 시스템은 유전자 전달 및 유전자 치료요법 적용에서 흔히 사용된다. 상이한 바이러스 벡터 시스템은 그만의 고유한 이점 및 불리한 점을 갖는다. 본 발명의 병원체-유래 리간드를 발현시키는 데 사용될 수 있는 바이러스 벡터는 비제한적으로, 하기에 보다 상세히 기재된 아데노바이러스 벡터, 아데노-연관 바이러스 벡터, 알파바이러스 벡터, 단순포진 바이러스 벡터 및 레트로바이러스 벡터를 포함한다.
- [0168] 아데노바이러스의 추가적인 일반적 특징은 아데노바이러스의 생물학이 상세히 분석된다는 것이며; 상기 아데노바이러스는 종종 인간 병리학과 연관되지 않고; 아데노바이러스는 이의 DNA 를 숙주 세포 내로 도입하는데 있어서 매우 효과적이고; 아데노바이러스는 광범위한 세포를 감염시킬 수 있으며 넓은 숙주 범위를 갖고; 아데노바이러스는 상대적으로 쉽게 다량으로 생성될 수 있고; 아데노바이러스는 바이러스 게놈의 초기 부위 1 ("E1") 에서의 결실에 의해 결손이 되고/되거나 비-복제형이 될 수 있다.
- [0169] 아데노바이러스는 비-외피형 DNA 바이러스이다. 아데노바이러스의 게놈은 각 가닥의 5'-말단에 공유 결합된 55-kDa 말단 단백질을 갖는 대략 36,000 개 염기쌍 ("bp") 의 선형 이중가닥 DNA 분자이다. 아데노바이러스 DNA 는 약 100 bp 의 동일한 반전 말단 반복물 ("ITR") 을 함유하는데, 정확한 길이는 혈청형에 따라 좌우된다. 복제의 바이러스 기원은 게놈 말단에서 정확히 ITR 내에 위치한다. DNA 합성은 2 단계로 발생한다. 먼저, 복제는 가닥 교체에 의해 진행되고, 딸 듀플렉스 분자 및 부모 대체된 가닥을 생성시킨다. 대체된 가닥은 단일 가닥이며 "팬핸들 (panhandle)" 중간체를 형성할 수 있는데, 이는 복제가 개시되게 하고 딸 듀플렉스 분자를 생성시킨다. 대안적으로, 복제는 게놈의 양쪽 말단으로부터 진행될 수 있으며 동시에, 팬핸들 구조 형성에 대한 필요사항이 제거된다.
- [0170] 생산적 감염 주기 동안, 바이러스 유전자는 2 개 상으로 발현된다: 바이러스 DNA 복제까지의 기간인 초기상, 및 바이러스 DNA 복제의 개시와 동시에 일어나는 후기상. 초기상 동안, 부위 E1, E2, E3 및 E4 에 의해 인코딩되는 초기 유전자 산물만이 발현되며, 이는 바이러스 구조 단백질의 합성을 위해 세포를 준비시키는 수많은 기능을 수행한다 (Berk, A.J., 1986). 후기상 동안, 후기 바이러스 유전자 산물이 초기 유전자 산물에 추가로 발현되며 숙주 세포 DNA 및 단백질 합성이 정지된다. 결과적으로, 세포는 바이러스 DNA 및 바이러스 구조 단백질의 생성에 역점을 두게 된다 (Tooze, J., 1981).
- [0171] 아데노바이러스의 E1 부위는 표적 세포의 감염 후 발현된 아데노바이러스의 첫 번째 부위이다. 상기 부위는 2 개 전사 단위, E1A 및 E1B 유전자로 이루어지며, 이들 둘 모두는 1 차 (배아) 설치류 배양물의 종양성 형질 전환에 필요하다. E1A 유전자 산물의 주요 기능은 휴지기 세포가 세포 주기에 들어가게끔 유도하고 세포 DNA 합성을 재개시키며, E1B 유전자 및 바이러스 게놈의 기타 초기 부위 (E2, E3 및 E4) 를 전사적으로 활성화시키는 것이다. E1A 유전자 단독을 갖는 1 차 세포의 트랜스펙션은 제한되지 않은 증식 (불멸화) 을 유도할 수 있으나, 완전한 형질전환을 초래하지는 않는다. 그러나 대부분의 경우 E1A 의 발현은 프로그래밍된 세포 사멸 (세포자멸사) 의 유도를 초래하며, 가끔만 불멸화가 수득된다 (Jochemsen et al., 1987). E1B 유전자의 동시발현은 세포자멸사의 유도를 방지하고 완전한 형태적 형질전환이 발생하도록 하기 위해 필요하다. 구축된 불멸 세포주에서, E1A 의 높은 수준의 발현은 E1B 의 부재 하 완전한 형질전환을 일으킬 수 있다 (Roberts, B.E. et al., 1985).
- [0172] E1B 인코딩 단백질은 바이러스 복제가 가능하도록 세포 기능을 적용하는데 있어서 E1A 를 보조한다. 본질적

으로 핵 내에서 국부화되는 복합체를 형성하는 E1B 55 kD 및 E4 33 kD 단백질은 숙주 단백질의 합성을 저해하고 바이러스 유전자 발현을 촉진하는데 있어서 그 기능을 수행한다. 이들의 주요 영향은 감염의 후기상의 발병에 부수적으로, 바이러스 mRNA 를 핵으로부터 세포질에 선택적으로 이동시키는 것을 확립하는 것이다. E1B 21 kD 단백질은 생산적 감염 주기의 올바른 시간적 제어를 위해 중요하며, 이로써 바이러스 생명 주기가 완료되기 전에 숙주 세포의 때 이른 사멸을 방지한다. E1B 21 kD 유전자 산물을 발현할 수 없는 돌연변이체 바이러스는 숙주 세포 염색체 DNA (deg-표현형) 의 과도한 분해에 의해, 그리고 증강된 세포변성 효과 (cyt-표현형; Telling et al, 1994) 에 있어서 동반되는 단축된 감염 주기를 나타낸다. deg 및 cyt 표현형은 추가로 E1A 유전자가 돌연변이화되는 경우 억제되는데, 이는 이러한 표현형이 E1A 와 상관관계에 있다는 것을 나타낸다 (White, E. et al., 1988). 추가로, E1B 21 kDa 단백질은 E1A 가 다른 바이러스 유전자를 스위치 온 (switch on) 하는 속도를 감소시킨다. E1B 21 kD 가 이러한 E1A 의존적 기능을 해소하는 메커니즘은 아직 알려져 있지 않다.

[0173] 예를 들어 레트로바이러스와 반대로, 아데노바이러스는 숙주 세포의 게놈에 효율적으로 통합되지 않고, 미분화 세포를 감염시킬 수 있으며, 재조합 유전자를 생체내 효율적으로 이동시킬 수 있다 (Brody et al., 1994). 이들 특징은 아데노바이러스로 하여금 예를 들어 관상 동맥 또는 면역원의, 이를 필요로 하는 세포, 조직 또는 대상 내로의 생체내 유전자 이동에 대한 효과적인 후보물로 만든다.

[0174] 다수 결실을 함유하는 아데노바이러스 벡터는 벡터의 운반 수용력을 증가시키고 복제 적격 아데노바이러스 (RCA) 를 생성시키는 재조합 가능성을 감소시킨다. 아데노바이러스가 다수의 결실을 함유하는 경우, 각각의 결실이, 단독으로 존재하는 경우, 복제 결함 및/또는 미-복제 아데노바이러스를 초래할 필요는 없다. 결실 중 하나가 아데노바이러스가 복제 결함이 되거나 미-복제가 되게 하는 한, 추가적인 결실이 다른 목적, 예를 들어 이중 뉴클레오티드 서열에 대한 아데노바이러스 게놈의 운반 수용력을 증가시키기 위해 포함될 수 있다. 바람직하게는, 결실 중 하나 초과는 기능적 단백질의 발현을 방지하고 아데노바이러스가 복제 결함 및/또는 미-복제 및/또는 약화되게 한다. 보다 바람직하게는, 모든 결실은 아데노바이러스가 복제-결함 및/또는 미-복제 및/또는 약화되게 할 수 있는 결실이다. 그러나, 본 발명은 또한 복제 적격 및/또는 야생형인 아데노바이러스 및 아데노바이러스 벡터를 포함하는데, 즉 대상에서의 감염 및 복제를 위해 필요한 모든 아데노바이러스 유전자를 포함한다.

[0175] 아데노바이러스 재조합체를 이용하는 본 발명의 구현에는 E1-결함 또는 결실, 또는 E3-결함 또는 결실, 또는 E4-결함 또는 결실, 또는 E1 및 E3, 또는 E1 및 E4, 또는 E3 및 E4, 또는 E1, E3 및 E4 결실을 포함하는 아데노바이러스 벡터, 또는 모든 바이러스 유전자가 결실되는 "헬퍼 의존적 (gutless)" 아데노바이러스 벡터를 포함할 수 있다. 아데노바이러스 벡터는 E1, E3 또는 E4 유전자에서의 돌연변이, 또는 이들 또는 모든 아데노바이러스 유전자에서의 결실을 포함할 수 있다. E1 돌연변이는, E1-결함 아데노바이러스 돌연변이체가 복제-결함 및/또는 비-허용 세포에서 미-복제성인 것으로 일컬어지며 적어도 매우 약화되기 때문에, 벡터의 안전역 (safety margin) 을 증가시킨다. E3 돌연변이는 메커니즘을 방해함으로써 항원의 면역원성을 증강시켜, 아데노바이러스가 MHC 클래스 I 분자를 하향 조절한다. E4 돌연변이는 후기 유전자 발현을 억제함으로써 아데노바이러스의 면역원성을 감소시켜, 그에 따라 동일한 벡터를 이용하는 반복된 백신 재접종을 가능하게 할 수 있다. 본 발명은 E1, 또는 E3, 또는 E4, 또는 E1 및 E3, 또는 E1 및 E4 에 있어서 결실되거나 돌연변이화되는 임의의 혈청형 또는 혈청군의 아데노바이러스 벡터를 포괄한다. 이들 아데노바이러스 유전자의 결실 또는 돌연변이는 이들 단백질 활성을 손상 또는 실질적으로 완전히 손실시킨다.

[0176] "헬퍼 의존적" 아데노바이러스 벡터는 아데노바이러스 벡터 패밀리에서의 또 다른 유형의 벡터이다. 이의 복제에는 헬퍼 바이러스 및 E1a 및 Cre 모두를 발현하는 특이적 인간 293 세포주, 자연적 환경에서 존재하지 않는 조건이 필요하며; 벡터는 모든 바이러스 유전자가 없고, 따라서 백신 운반체로서의 상기 벡터는 비-면역원성이고 백신 재접종을 위해 수회 접종될 수 있다. "헬퍼 의존적" 아데노바이러스 벡터는 또한 관상 동맥 또는 면역원(들) 을 수용하기 위해 36 kb 공간을 포함하며, 따라서 다수의 항원 또는 면역원을 세포 내로 동시 전달할 수 있다.

[0177] 아데노-연관 바이러스 (AAV) 는 인간 집단에 대해 내인성인 단일 가닥 DNA 파르보바이러스이다. 다양한 종류로부터의 세포에서 생산적 감염이 가능함에도 불구하고, AAV 는 데펜도바이러스 (dependovirus) 여서, 이의 고유 복제를 위해 아데노바이러스 또는 포진 바이러스로부터 헬퍼 기능을 필요로 한다. 이들 헬퍼 바이러스로부터의 헬퍼 기능 부재 하에서, AAV 는 세포를 감염시키고, 핵에서 탈피되고, 이의 게놈이 숙주 염색체 내로 통합되거나, 새로운 바이러스 입자를 복제하거나 생성시키지는 않을 것이다.

- [0178] AAV 의 게놈은 박테리아 플라스미드로 클로닝되며 잘 분석된다. 바이러스 게놈은 각각 145 개 염기의 2 개 말단 반복물을 포함하는 4682 개 염기로 이루어진다. 이들 말단 반복물은 바이러스에 대한 DNA 복제의 기원으로서 역할한다. 일부 조사자들이 또한 이들이 인헨서 기능을 갖는다는 것을 제안한 바 있다. 게놈의 나머지는 2 개의 기능적 도메인으로 나뉜다. 게놈의 좌측 부분은 바이러스 DNA 복제 및 생명유지 유전자 발현을 조절하는 rep 기능을 코딩한다. 생명유지 게놈의 우측 면은 구조적 캡시드 단백질 VP1, VP2 및 VP3 을 인코딩하는 cap 유전자를 함유한다. rep 및 cap 유전자 모두에 의해 인코딩되는 단백질은 생산적 AAV 복제 동안 trans 에서 기능한다.
- [0179] AAV 는 형질도입 벡터로서 사용하기 위한 이상적인 후보물로 고려되며, 이러한 방식으로 사용되고 있다. 이러한 AAV 형질도입 벡터는 trans 에서 주어진 아데노바이러스 또는 포진 바이러스 헬퍼 기능의 존재 하에 복제하기에 충분한 cis-작용 기능을 포함한다. 재조합 AAV (rAAV) 는 많은 실험실에서 구축되고 있으며 다양한 혈통의 세포 내로 외인성 유전자를 운반하는데 사용되고 있다. 이들 벡터에서, AAV cap 및/또는 rep 유전자는 바이러스 게놈으로부터 결실되고 선택 DNA 조각으로 대체된다. 현재 벡터는 4300 개 이하 염기의 삽입된 DNA 를 수용할 수 있다.
- [0180] rAAV 를 생성시키기 위해서, 원하는 바이러스 구축물을 함유하는 플라스미드를 아데노바이러스-감염된 세포 내로 트랜스펙션한다. 추가로, 제 2 헬퍼 플라스미드를 이들 세포 내로 동시트랜스펙션하여 재조합 바이러스 구축물의 복제 및 패키징에 의무적인 AAV rep 및 cap 유전자를 제공한다. 이러한 조건 하에, AAV 의 rep 및 cap 단백질은 trans 에서 rAAV 구축물의 복제 및 패키징을 자극하는 작용을 한다. 트랜스펙션 3 일 후, rAAV 를 아데노바이러스와 함께 세포로부터 수확한다. 오염 아데노바이러스를 이후 열 처리에 의해 불활성화시킨다.
- [0181] 단순 포진 바이러스 1 (HSV-1) 은 80 개 초과 유전자를 인코딩하는 153 kb 의 게놈을 갖는 외피형, 이중가닥 DNA 바이러스이다. 이의 광범위한 숙주 범위는 세포 멤브레인에서 발견된 세포외 헤파린 술페이트 분자에 대한 바이러스 외피 당단백질의 결합으로 인한 것이다 (WuDunn & Spear, 1989). 바이러스의 부화는 외피 당단백질 gD 및 섬유모세포 성장 인자 수용체를 필요로 한다 (Kaner, 1990). HSV 는 세포를 용해적으로 감염시킬 수 있거나 잠복성을 구축할 수 있다. HSV 벡터를 사용하여 광범위한 세포 유형을 감염시킨다 (Lowenstein, 1994; Huard, 1995; Miyanochara, 1992; Liu, 1996; Goya, 1998).
- [0182] 재조합 HSV 벡터 및 암플리콘 벡터로 지칭되는 2 개 유형의 HSV 벡터가 존재한다. 재조합 HSV 벡터는 상동 재조합 이벤트를 통해 전사 단위를 HSV 게놈 내로 직접 삽입하여 생성된다. 암플리콘 벡터는 선택 전사 단위, 복제 기원 및 패키징 신호를 갖는 플라스미드를 기반으로 한다.
- [0183] HSV 벡터는 외래 유전자의 삽입을 위한 큰 수용력, 뉴런에서의 잠복성 구축력, 넓은 숙주 범위 및 18 개월까지 CNS 에 이식유전자 발현을 부여하는 능력의 명백한 이점을 갖는다 (Carpenter & Stevens, 1996).
- [0184] 레트로바이러스는 유전자 이동 프로토콜에서 광범위하게 사용되는 외피형 단일가닥 RNA 바이러스이다. 레트로바이러스는 gag, pro, pol 및 env 로 지칭되는 4 개 유전자 부위로 이루어진 약 7-10 kb 의 이베체 게놈을 갖는다. 이들 유전자 부위는 구조적 캡시드 단백질, 바이러스 프로테아제, 인테그라아제 및 바이러스 역전사 효소 및 외피 당단백질을 각각 인코딩한다. 게놈은 또한 패키징 신호 및 cis-작용 서열 (장길이-말단 반복물 (LTR) 로 지칭) 을 각각의 말단에 갖는데, 이는 전사 제어 및 통합에 있어서 역할을 갖는다.
- [0185] 가장 흔하게 사용되는 레트로바이러스 벡터는 몰로니 (Moloney) 쥐과 백혈병 바이러스 (Mo-MLV) 를 기반으로 하며 외피 당단백질의 수용체 결합 표면 도메인에 따라 가변적 세포 향성을 갖는다.
- [0186] 재조합 레트로바이러스 벡터는 모든 레트로바이러스 유전자로부터 결실되는데, 이는 마커 또는 치료적 유전자, 또는 둘 모두로 대체된다. 재조합 레트로바이러스를 증식시키기 위해, 바이러스 유전자, gag, pol 및 env 를 trans 에서 제공하는 것이 필요하다.
- [0187] 렌티바이러스는 유사분열 및 유사분열 후 세포 모두를 감염시키고 그의 유전자를 발현시키는 능력을 갖는 복잡한 레트로바이러스이다. 가장 흔하게 공지된 렌티바이러스는 인간 면역결핍 바이러스 (HIV) 이며, 이는 다른 바이러스의 외피 당단백질을 사용하여 광범위한 세포 유형을 표적화한다.
- [0188] 원형 신드비스 바이러스 (Sindbis virus) (SIN), 쉘리키 삼림 바이러스 (Semliki Forest virus) (SFV) 및 베네수엘라 말 뇌염 바이러스 (Venezuelan equine encephalitis virus) (VEE) 를 포함하는 알파바이러스는 20 면체 캡시드 내에 + 가닥 RNA 게놈을 함유하는 외피형 바이러스의 군을 구성한다.

- [0189] 본 발명의 바이러스 벡터는 시험관내 및 생체내 모두에서 세포에 항원 또는 면역원 발현 핵산을 전달하는데 유용하다. 특히, 본 발명의 벡터는 핵산을 세포, 보다 바람직하게는 포유동물 세포에 전달하거나 이동시키는 데 유리하게 이용될 수 있다. 관심 핵산은 펩티드 및 단백질, 바람직하게는 치료적 (예를 들어, 의약품 또는 수의학용) 또는 면역원성 (예를 들어, 백신용) 펩티드 또는 단백질을 인코딩하는 핵산을 포함한다.
- [0190] 바람직하게는, 관심 항원 또는 면역원을 인코딩하는 코돈은 "최적화된" 코돈이며, 즉, 상기 코돈은 예를 들어, 인플루엔자 바이러스에 의해 종종 사용되는 코돈 대신, 대상의 종류에서 고도로 발현된 유전자에서 종종 나타나는 것들이다. 이러한 코돈 사용빈도는 동물 세포에서 항원 또는 면역원의 효율적 발현을 제공한다. 다른 구현예에서, 예를 들어 관심 항원 또는 면역원이 박테리아, 효모 또는 또 다른 발현 시스템에서 발현되는 경우, 코돈 사용빈도 패턴은 항원 또는 면역원이 발현되는 유기체에서의 고도로 발현된 유전자에 대한 코돈 편향을 나타내도록 변경된다. 코돈 사용빈도 패턴은 많은 종류의 고도로 발현된 유전자에 대한 문헌에서 공지되어 있다 (예를 들어, Nakamura et al, 1996; Wang et al, 1998; McEwan et al. 1998).
- [0191] 추가적인 대안물로서, 바이러스 벡터는 배양물에서 세포를 감염시켜 원하는 유전자 산물을 발현하는데, 예를 들어 관심 단백질 또는 펩티드를 생성시키는데 사용될 수 있다. 바람직하게는, 단백질 또는 펩티드는 배지로 분비되고 당업계에 공지된 일상적 기술을 사용하여 그로부터 정제될 수 있다. 단백질의 세포외 분비를 유도하는 신호 펩티드 서열은 당업계에 공지되어 있으며 이를 인코딩하는 뉴클레오티드 서열은 당업계에 공지된 일상적 기술에 의해 관심 펩티드 또는 단백질을 인코딩하는 뉴클레오티드 서열에 작동가능하게 연결될 수 있다. 대안적으로, 세포는 용해될 수 있으며 발현된 재조합 단백질은 세포 용해물로부터 정제될 수 있다. 바람직하게는 세포는 동물 세포이고, 보다 바람직하게는 포유동물 세포이다. 또한, 특정 관심 바이러스 벡터에 의한 형질도입에 적합한 세포가 바람직하다. 이러한 세포는 PER.C6 세포, 911 세포 및 HEK293 세포를 포함한다.
- [0192] 숙주 세포 배양을 위한 배양 배지는 조직 배양에 흔히 사용되는 배지, 그 중에서도 예컨대 M199-얼 (earle) 배이스, 이글 MEM (E-MEM), 돌베코 MEM (DMEM), SC-UCM102, UP-SFM (GIBCO BRL), EX-CELL302 (Nichirei), EX-CELL293-S (Nichirei), TFBM-01 (Nichirei), ASF 104 를 포함한다. 특정 세포 유형에 적합한 배양 배지는 미국 미생물 보존 센터 (American Type Culture Collection) (ATCC) 또는 유럽 세포 배양물 보존 센터 (European Collection of Cell Cultures) (ECACC) 에서 발견될 수 있다. 배양 배지에는 아미노산 예컨대 L-글루타민, 염, 항진균 또는 항박테리아제 예컨대 Fungizone®, 페니실린-스트렙토마이신, 동물 혈청 등이 보충될 수 있다. 세포 배양 배지는 임의로는 혈청-불포함일 수 있다.
- [0193] 본 발명은 또한 LITE 또는 보노 발명에서 유용한 하나 이상의 작용제를 발현 또는 과발현할 수 있는 세포주 또는 트랜스제닉 동물에 관한 것이다. 바람직하게는 세포주 또는 동물은 하나 이상의 LITE 를 발현하거나 과발현한다.
- [0194] 트랜스제닉 동물은 통상 척추동물, 보다 바람직하게는 설치류, 예컨대 랫트 또는 마우스일 뿐 아니라 기타 포유동물 예컨대 인간, 염소, 돼지 또는 소 등을 포함한다.
- [0195] 이러한 트랜스제닉 동물은 새로운 유용한 화합물에 대한 스크리닝 검정에서 및 질환의 동물 모델로서 유용하다. 상기 정의한 바와 같이 하나 이상의 폴리펩티드를 특이적으로 발현시켜, 질환 발전에 대한 이러한 폴리펩티드의 효과를 연구할 수 있다. 추가로, 유전자 치료요법 및 각종 약물을 포함하는 치료요법을 트랜스제닉 동물에서 시험할 수 있다. 트랜스제닉 동물의 생성 방법은 당업계에 공지되어 있다. 예를 들어, 유전자를 배아 내로 도입하기 위한 여러 가능한 경로가 존재한다. 이는 (i) 배아 줄기 세포의 직접 트랜스펙션 또는 레트로바이러스 감염, 이후 발생의 배반포 단계에서의 이들 세포의 배아 내로의 도입; (ii) 초기 배아의 레트로바이러스 감염; 및 (iii) DNA 의 접합자 또는 초기 배아 세포 내로의 직접 미세주입을 포함한다. 유전자 및 /또는 전이유전자는 또한 당업계에 공지된 유전적 조절 요소 및/또는 구조적 요소를 포함할 수 있다. 전이유전자 도입을 위한 표적 세포의 유형은 배아 줄기 세포 (ES) 이다. ES 세포는 시험관내 배양된 사전-이식 배아로부터 수득될 수 있으며 배아에 융합된다 (Evans et al, 1981, Nature 292: 154-156; Bradley et al, 1984, Nature 309:255-258; Gossler et al, 1986, Proc. Natl. Acad. Sci. USA 83:9065-9069; and Robertson et al, 1986 Nature 322:445-448). 이식유전자는 DNA 트랜스펙션, 미세주입, 또는 레트로바이러스-매개 형질도입과 같은 다양한 표준 기술에 의해 ES 세포 내로 효율적으로 도입될 수 있다. 생성된 형질전환 ES 세포는 이후 비-인간 동물로부터의 배반포와 조합될 수 있다. 도입된 ES 세포는 이후 배아를 균집화시키며 생성된 키메라 동물의 생식세포 계열에 기여한다 (Jaenisch, 1988, Science 240: 1468-1474).
- [0196] LITE 는 또한 가치있는 생체내 시간적 정확도를 제공할 수 있다. LITE 는, 예를 들어 C. 엘레간스 (C.

elegans) 성장의 특정 단계 동안에만 특정 세포자멸사 유전자를 억제하여, 발생의 특정 단계 동안 유전자 발현을 변형시키는데 사용될 수 있다. LITE 는 유전적 신호를 특정 실험 창 (window) 에 대해 시기 조정하는데 사용될 수 있다. 예를 들어, 학습에 연루된 유전자는 미손상 설치류 또는 영장류 뇌의 정확도 부위에서의 학습 자극 동안에만 과발현되거나 억제될 수 있다. 또한, LITE 는 질환 발전의 특정 단계 동안에만 유전자 발현 변화를 유도하는데 사용될 수 있다. 예를 들어, 암유전자는 종양이 특정 크기 또는 전이 단계에 도달하고 단 한번 과발현될 수 있다. 반대로, 알츠하이머병 발전에서 의심된 단백질은 특정 뇌 부위 내에서 및 동물 생 (life) 에서 규정된 시점에서만 녹다운될 수 있다. 이들 예가 LITE 시스템의 잠재적 적용을 철저히 열거하지는 않으나, 이들은 LITE 가 강력한 기술일 수 있는 일부 분야를 강조한다.

[0197] 본 발명의 치료적 및 진단적 조성물은 장애를 치료하거나 진단하기에 충분한 양으로 개인에게 투여된다. 유효량은 개인의 상태, 체중, 성별 및 연령과 같은 다양한 인자에 따라 가변적일 수 있다. 기타 인자는 투여 방식을 포함한다.

[0198] 약학 조성물은 다양한 경로 예컨대 피하, 국소, 경구 및 근육내에 의해 개인에게 제공될 수 있다.

[0199] 본원에 개시된 방법에 따라 확인된 화합물은 적절한 투약량에서 단독으로 사용될 수 있다. 대안적으로, 기타 작용제의 동시-투여 또는 순차적 투여가 바람직할 수 있다.

[0200] 본 발명은 또한 본 발명의 신규한 치료 방법에서 사용하기 위해 적합한 국소, 경구, 전신 및 비경구 약학 제형을 제공하는 목적을 갖는다. 활성 성분으로서 본 발명에 따라 확인된 화합물을 함유하는 조성물은 투여용의 종래 비히클 중 광범위한 치료 투약 형태로 투여될 수 있다. 예를 들어, 화합물은 정제, 캡슐 (각각 시간적 방출형 및 서방형 제형 포함), 알약, 분말, 과립, 엘릭시르, 팅크, 용액, 현탁액, 시럽 및 에멀전과 같은 경구 투약 형태로, 또는 주사에 의해 투여될 수 있다. 마찬가지로, 이는 또한 정맥내 (볼루스 및 주입물 모두), 복강내, 피하, 국소 (흡장의 존재 또는 부재 하), 또는 근육내 형태로 투여될 수 있다 (모든 사용 형태는 약학 업계의 당업자에게 널리 공지됨).

[0201] 유리하게는, 본 발명의 화합물은 단일 1 일 용량으로 투여될 수 있거나, 총 1 일 투약량이 1 일 2, 3 또는 4 회 의 분할 용량으로 투여될 수 있다. 추가로, 본 발명의 화합물은 적합한 비강내 비히클의 국소 사용을 통해, 또는 경피 경로를 통해 비강내 형태로 투여될 수 있으며, 경피 피부 패치의 이들 형태를 사용하는 것은 당업자에게 널리 공지되어있다. 경피 전달 시스템의 형태로 투여되기 위해서는, 투약량의 투여는 물론 투약 계획 전체에 걸쳐 간헐식 보다는 연속식일 것이다.

[0202] 하나 초과 의 활성제로의 병용 치료를 위해서, 활성제가 별개의 투약량 제형 중에 존재하는 경우, 상기 활성제는 동시에 투여될 수 있거나, 각각 별도로 시차 시간으로 투여될 수 있다.

[0203] 본 발명의 화합물을 이용하는 투약 계획은 환자의 유형, 인종, 연령, 체중, 성별 및 의학적 상태; 치료하는 병상의 중증도; 투여 경로; 한 환자의 신장, 간 및 심혈관 기능; 및 이용한 이의 특정 화합물을 포함하는 다양한 인자에 따라 선택된다. 통상의 지식을 갖는 의사는 병상의 진행을 방지하고, 대응하거나 막기 위해 필요한 약물의 유효량을 용이하게 결정하고 처방할 수 있다. 독성을 갖지 않는 효능을 산출해내는 범위 내의 약물 농도를 얻는데 있어서 최적의 정확도는 표적 부위에 대한 약물의 이용가능성 속도를 기반으로 한 계획을 필요로 한다. 이는 약물의 분포도, 평형 및 제거에 대한 고려를 포함한다.

[0204] 본 발명 및 이의 이점이 상세히 기재되어 있으나, 첨부된 특허청구범위에서 정의한 바와 같이 본 발명의 취지 및 범주에서 벗어나지 않고 다양한 변화, 대응 및 변형이 본원에서 이루어질 수 있다는 것이 이해되어야 한다.

[0205] 하기 실시예에서 본 발명을 추가로 설명하며, 이는 설명의 목적으로만 제공되며 어떠한 식으로도 본 발명을 제한하는 것으로 의도되지 않는다.

[0206] **실시예**

[0207] *실시예 1*

[0208] 내인성 포유동물 계놈으로부터의 유전자 발현을 직접적으로 조정하는 능력은 정상 유전자 기능 및 질환 메커니즘을 해명하기 위해 중요하다. 세포 집단 내에서 유전자 발현의 공간적 및 시간적 제어를 추가로 개선시키는 진전은 유전자 조정의 유용성을 확장시키는 가능성을 갖는다. 출원인은 이전에 잔타모나스 오리제 (Xanthomonas oryze) 로부터 전사 활성제-유사 효과기 (TALE) 를 개발하여, 부위-특이적 DNA 결합 단백질의 신속한 설계 및 구축을 가능하게 하였다. 출원인은 내인성 포유동물 계놈에서의 광-조절 유전자 발현을 가능하게 하는 분자 도구 세트를 개발하였다. 상기 시스템은 아라비도프시스 탈리아나 (Arabidopsis thaliana)

로부터의 광-민감성 이량체화 단백질 도메인에 연결된 조작된 인공 전사 인자로 이루어진다. 상기 시스템은 450nm -500nm 범위의 광에 반응하며 포유동물 세포에서 6.2 mW/cm<sup>2</sup> 강도로의 광 자극 후 전분화능 인자의 발현에 있어서 유의한 증가를 유도할 수 있다. 출원인은 광범위한 유전자의 표적화를 위한 도구를 개발하고 있다. 출원인은 유전자 발현의 광-매개 제어를 위한 도구상자가 현존하는 광유전학적 방법을 보완할 수 있으며 장차 뇌에서의 특정 유전자의 시기-, 세포 유형- 및 농도 의존적 역할을 해명하는 것을 도울 수 있다고 믿는다.

[0209] 내인성 포유동물 계놈으로부터의 유전자 발현을 직접적으로 조정하는 능력은 정상 유전자 기능 및 질환 메커니즘을 해명하기 위해 중요하다. 출원인은 내인성 포유동물 계놈에서의 광-조절 유전자 발현을 가능하게 하는 분자 도구 세트 개발을 제시한다. 이러한 시스템은 아라비도프시스 탈리아나로부터의 광-민감성 이량체화 단백질 도메인 크립토크롬 2 (CRY2) 및 CIB1 에 연결된 전사 활성제 유사 효과기 (TALE) 및 활성화 도메인 VP64 로 이루어진다. 출원인은 KLF4 및 Neurog2 의 프로모터 부위를 표적화하도록 설계된 이들 LITE 구축물로 트랜스펙션한 HEK293FT 및 Neuro-2a 세포의 청색광 자극이 표적 발현에 있어서 유의한 증가를 초래한다는 것을 나타내는데, 이는 TALE-기재 광학 유전자 발현 조정 기술의 기능성을 입증한다.

[0210] 도 1 은 공간적 및 시간적 정확도에 대한 필요성을 설명하는 모식도를 나타낸다.

[0211] 도 2 는 전사 활성제 유사 효과기 (TALE) 를 나타낸다. TALE 는 그의 서열의 코어에서 34 개 aa 반복물로 이루어진다. 각각의 반복물은 TALE 에 의해 결합되는 표적 DNA 에서의 염기에 상응한다. 반복물은 위치 12 및 13 에서 2 개의 가변적인 아미노산에 의해서만 상이하다. 이러한 대응의 코드가 설명된 바 있으며 (Boch, J et al. , Science, 2009 and Moscou, M et al. , Science, 2009) 이 도면에서 나타낸다. 출원인은 이러한 코드를 통합하며 계놈 내에 선택 서열을 결합시킬 수 있는 설계자 TALE 의 합성 방법을 개발하였다 (Zhang, F et al. , Nature Biotechnology, 2011).

[0212] 도 3 은 LITE: TALE/크립토크롬 전사 활성화의 설계를 나타낸다. 각각의 LITE 는 CRY2 에 융합된 TALE 및 VP64 (전사 인자) 에 융합된 크립토크롬 결합 파트너 CIB1 을 포함할 수 있는 2-성분 시스템이다. 불활성 상태에서, TALE 은 이의 융합된 CRY2 도메인을 관심 유전자의 프로모터 부위에 국부화시킨다. 이때, CIB1 은 CRY2 에 결합할 수 없고, CIB1-VP64 가 핵 공간 내에서 미결합된 채로 있다. 488 nm (청색) 광으로 자극 시, CRY2 는 형태적 변화를 거쳐, 이의 CIB1 결합 부위가 밝혀진다 (Liu, H et al. , Science, 2008). CIB1 의 신속한 결합은 융합된 VP64 도메인의 영입을 초래하여, 표적 유전자의 전사를 유도한다.

[0213] 도 4 는 LITE 활성화에 대한 크립토크롬 이량체 절두의 효과를 나타낸다. CRY2 및 CIB1 의 활성을 변경시키는 것으로 알려져 있는 절두를 전체 길이 단백질에 대해 비교하였다. Neurog2 의 프로모터에 대해 표적화된 LITE 를 도메인의 각 조합에 대해 Neuro-2a 세포에서 시험하였다. 488 nm 광으로 자극한 후, Neurog2 의 전사 수준을 자극된 샘플 및 미자극된 샘플에 대해 qPCR 을 사용하여 정량하였다.

[0214] 도 5 는 KLF4 LITE 의 광-강도 의존적 반응을 나타낸다.

[0215] 도 6 은 Neurog2 LITE 의 활성화 반응속도 및 Neurog2 LITE 의 불활성화 반응속도를 나타낸다.

[0216] 실시예 2

[0217] 정상 유전자 발현은 주의깊게 조직된 시간적 및 공간적 성분을 갖는 동적 프로세스이며, 이의 정확도는 정상 발생, 항상성 유지 및 유기체의 진보를 위해 필요하다. 그 다음, 유전자 또는 유전자 세트의 증가, 감소 또는 변경된 기능에 의한 필요한 유전자 발현 패턴의 조절장애는 광범위한 집합의 병리학과 연결된다. 시공간적 정확도 방식으로 유전자 발현을 조정할 수 있는 기술은 정상 생물학적 프로세스 및 질환 메커니즘을 책임지는 유전적 신호의 해명을 가능하게 할 것이다. 이러한 기술적 필요성을 다루기 위해, 출원인은 내인성 유전자 발현의 광-매개 제어를 제공하는 광-유도 전사 효과기 (LITE) 를 개발하였다.

[0218] 유도 유전자 발현 시스템은 통상 삽입된 개방 판독 프레임 또는 shRNA 서열의 화학적 유도 활성화가 가능하도록 설계되어, 각각 유전자 과발현 또는 억제를 초래한다. 과발현을 위한 개방 판독 프레임 사용의 불리한 점은 스플라이스 변이의 손실 및 유전자 크기의 제한을 포함한다. RNA 간섭을 통한 유전자 억제, 인간 생물학에 있어서의 이의 형질전환적 영향력에도 불구하고, 복잡한 오프-타겟 효과에 의해 저해될 수 있다. 에스트로젠, 액티손 및 FKBP12/FRAP 기반 시스템을 포함하는 특정 유도 시스템은 오프-타겟 내인성 유전자를 활성화시키는 것으로 공지되어있다. 장기간 항생제 치료의 잠재적 유해 효과는 테트라사이클린 전사촉진제 (TET) 기반 시스템의 사용을 복잡하게 만들 수 있다. 이러한 화학적 유도 시스템의 생체내, 시간적 정확도는 유도제

흡수 및 제거의 속도에 따라 좌우된다. 또한, 유도제가 일반적으로 전신 전달되므로, 이러한 시스템의 공간적 정확도는 외인성 벡터 전달의 정확도에 의해 제약된다.

[0219] 이러한 제한에 대응하여, LITE 는 시간적 및 공간적으로 정확한 방식으로 개별적 내인성 유전자의 발현을 조정하도록 설계된다. 각각의 LITE 는 커스텀화된 DNA-결합 전사 활성화제 유사 효과기 (TALE) 단백질, 아라바둑 시스 탈리아나로부터의 광-반응성 크립토크롬 이형이량체 및 전사 활성화/억제 도메인으로 이루어지는 2 성분 시스템이다. TALE 은 관심 유전자의 프로모터 서열에 결합하도록 설계된다. TALE 단백질은 크립토크롬 이형이량체의 절반 (크립토크롬-2 또는 CIB1) 에 융합되면서, 남아있는 크립토크롬 파트너가 전사 효과기 도메인에 융합된다. 효과기 도메인은 활성화제 예컨대 VP16, VP64 또는 p65, 또는 억제제 예컨대 KRAB, EnR 또는 SID 일 수 있다. LITE 의 미자극 상태에서, TALE-크립토크롬2 단백질은 관심 유전자의 프로모터에 대해 국부화되나, CIB1-효과기 단백질에 결합하지는 않는다. 청색 스펙트럼 광으로 LITE 자극시, 크립토크롬-2 는 활성화되고, 형태적 변화를 거치며, 이의 결합 도메인이 밝혀진다. 다음으로 CIB1 은 크립토크롬-2 에 결합하여 관심 유전자의 프로모터 부위에 대해 효과기 도메인을 국부화시키고 유전자 과발현 또는 침묵화를 개시시킨다.

[0220] LITE 에서의 유전자 표적화는 커스텀화된 TALE DNA 결합 단백질의 특이성을 통해 이루어질 수 있다. 관심 유전자의 프로모터 부위에서 표적 서열이 선택되며 상기 서열에 대해 커스텀화된 TALE 이 설계된다. TALE 의 중심 부분은 34 개 아미노산 길이의 탠덤 반복물로 이루어진다. 이들 반복물의 서열은 거의 동일하나, 각각의 반복물의 12 번째 및 13 번째 아미노산 (반복 가변성 2 잔기로 지칭함) 은 가변적이어서, 각각의 반복물의 뉴클레오티드-결합 특이성을 결정한다. 따라서, TALE 단량체 반복물의 적절한 배치를 갖는 구축물을 합성하여, 표적 프로모터 서열에 특이적인 DNA 결합 단백질을 제작한다.

[0221] LITE 의 광 반응성은 크립토크롬-2 및 CIB1 의 활성화 및 결합을 통해 이루어진다. 상기 언급한 바와 같이, 청색광 자극은 크립토크롬-2 에 있어서의 활성화 형태적 변화를 유도하여, 이의 결합 파트너 CIB1 의 영입을 초래한다. 이러한 결합은 빠르고 가역적이며, 펄스화된 자극 후 <15 초 내에 포화가 이루어지고 자극 종료 후 <15 분에 기준선으로 복귀된다. 이러한 신속한 결합 속도는 유도제의 흡수 및 청소율보다는 전사/번역 및 전사체/단백질 분해의 속도에 의해서만 일시적으로 결합한 LITE 시스템을 초래한다. 크립토크롬-2 활성화는 또한 매우 민감하여, 낮은 광 강도 자극의 사용을 가능하게 하며 광독성의 위험성을 완화시킨다. 추가로, 미손상 포유동물 뇌와 같은 맥락에서, 가변적인 광 강도는 LITE 자극 부위의 크기를 제어하는데 사용될 수 있어, 벡터 단독 전달이 제공할 수 있는 것보다 더 큰 정확성을 부여한다.

[0222] LITE 시스템의 모듈성은 임의 수의 효과기 도메인이 전사 조정에 사용될 수 있게 한다. 따라서, 활성화 및 억제제 도메인은 종류, 강도, 메커니즘, 지속시간, 크기, 또는 임의 수의 기타 매개변수를 기반으로 선택될 수 있다.

[0223] 출원인은 그 다음으로 LITE 시스템의 2 가지 원형적 시현을 제시한다. 제 1 예는 마우스 유전자 NEUROG2 의 전사를 활성화시키도록 설계된 LITE 이다. 마우스 NEUROG2 의 업스트림 프로모터 부위에 위치한 서열 TGAATGATGATAATACGA 를 표적으로서 선택하고, 상기 서열에 매치되도록 TALE 을 설계하고 합성하였다. TALE 서열을 핵 국부화 신호 (아미노산: SPKKKRKVEAS) 를 통해 크립토크롬-2 에 대한 서열에 연결시켜, 사이토솔로부터의 단백질의 핵 공간으로의 이동을 촉진시킨다. 동일한 핵 국부화 신호를 사용하여, 전사 활성화제 도메인 VP64 에 연결된 CIB1 도메인을 포함하는 제 2 벡터를 합성하였다. 이러한 제 2 벡터, 또한 GFP 서열을 2A 전사 스킵 신호에 의해 CIB1-VP64 융합 서열로부터 분리하였다. 각 구축물의 발현을 편재성, 항시성 프로모터 (CMV 또는 EF1- $\alpha$ ) 에 의해 유도하였다. Neuro 2A 세포주로부터의 마우스 신경아세포종 세포를 2 개 벡터로 동시-트랜스펙션시켰다. 벡터 발현이 가능하도록 인큐베이션한 후, 샘플을 488 nm LED 의 어레이로부터의 주기적 펄스의 청색광에 의해 자극하였다. 미자극된 동시-트랜스펙션 샘플 및 형광 리포터 YFP 로만 트랜스펙션된 샘플을 대조군으로서 사용하였다. 각 실험 완료시, mRNA 를 qPCR 을 통해 분석한 샘플로부터 정제하였다.

[0224] 크립토크롬-2 및 CIB1 의 절두 형태를 클로닝하고 전체 길이 형태의 크립토크롬-2 및 CIB1 과 조합으로 시험하여, 각각의 이형이량체 쌍의 효과를 측정하였다. 크립토크롬-2 단백질의 보존된 광반응성 부위 및 전체 길이 형태 CIB1 로 이루어지는 CRY2PHR 도메인의 조합은 Neurog2 mRNA 수준의 최고 상향조절을 초래하였다 (YFP 샘플의 ~22 배 초과 및 미자극 동시-트랜스펙션 샘플의 ~7 배 초과). 전체 길이 크립토크롬-2 (CRY2) 와 전체 길이 CIB1 의 조합은 낮은 절대 활성화 수준 (YFP 의 ~4.6 배 초과) 뿐 아니라 낮은 기준선 활성화 (미자극 동시-트랜스펙션 샘플에 대한 YFP 의 ~1.6 배 초과) 를 초래하였다. 이들 크립토크롬 단백질 페어링은 필요

한 유도의 절대 수준 및 LITE 시스템의 기준선 "누수도" 를 최소화하기 위한 필요성에 따라 특정 용도를 위해 선택될 수 있다.

[0225] 가역성 및 활성화의 속도는 LITE 시스템에 대한 중요한 설계 매개변수이다. LITE 시스템의 속도를 분석하기 위해, Neurog2 TALE-CRY2PHR 및 CIB1-VP64 형태 시스템으로 이루어지는 구축물을 시험하여 이의 활성화 및 불활성화 속도를 측정하였다. 추출하기 전 짧게는 0.5 시간 내지 길게는 24 시간 동안 샘플을 자극하였다. Neurog2 발현의 상향조절은 최단 시점, 0.5 시간에서 관찰되었다 (YFP 샘플에 대해 ~5 배). Neurog2 발현은 자극의 12 시간에서 최대치였다 (YFP 샘플에 대해 ~19 배). 불활성화 속도를 6 시간 동안 동시-트랜스펙션 샘플을 자극하여 분석하여, 이때 자극을 중단하고, 샘플을 0 내지 12 시간 동안 배양물 중 유지시켜 mRNA 가 분해되게 하였다. Neurog2 mRNA 수준은 자극 완료 후 0.5 시간에서 최대치였고 (YFP 샘플에 대해 ~16 배), 그 후 상기 수준은 ~3 시간 반감기로 저하된 후 12 시간까지 기준선 수준 근처로 복구되었다.

[0226] 제 2 의 원형적 예는 인간 유전자 KLF4 의 전사를 활성화시키도록 설계된 LITE 이다. 인간 KLF4 의 업스트림 프로모터 부위에 위치한 서열 TTCTACTTATAAC 를 표적으로서 선택하고, 상기 서열에 매치되도록 TALE 을 설계하고 합성하였다. TALE 서열을 핵 국부화 신호 (아미노산: SPKKKRKVEAS) 를 통해 CRY2PHR 에 대한 서열에 연결시켰다. 상기 기재된 동일한 CIB1-VP64 활성화제 단백질을 또한 LITE 시스템의 이러한 시험에 사용하였다. HEK293FT 세포주로부터의 인간 배아 신장 세포를 2 개 벡터로 동시-트랜스펙션하였다. 벡터 발현이 가능하도록 인큐베이션한 후, 488 nm LED 의 어레이로부터의 주기적 펄스 청색광에 의해 샘플을 자극하였다. 미자극 동시-트랜스펙션한 샘플 및 형광 리포터 YFP 로만 트랜스펙션한 샘플을 대조군으로서 사용하였다. 각 실험의 종료시, mRNA 를 qPCR 을 통해 분석한 샘플로부터 정제하였다.

[0227] LITE 시스템의 광-강도 반응을, 샘플을 증가한 광 동력 (0-9 mW/cm<sup>2</sup>) 으로 자극하여 시험하였다. KLF4 mRNA 수준의 상향조절이 0.2 mW/cm<sup>2</sup> 의 낮은 자극에 대해 관찰되었다. KLF4 상향조절은 5 mW/cm<sup>2</sup> 에서 포화되었다 (YFP 샘플에 대해 2.3 배). 세포 생존도 시험을 또한 9 mW/cm<sup>2</sup> 이하의 동력에 대해 수행하였고, >98% 세포 생존도를 나타내었다. 유사하게, 자극 사용률 가변화에 대한 KLF4 LITE 반응을 시험하였다 (1.6-100%). 상이한 사용률 간에 KLF4 활성화에 있어서의 차이가 관찰되지 않았는데, 이는 15 초 마다 0.25 초의 낮은 자극 패러다임이 최대 활성화를 초래한다는 것을 나타낸다.

[0228] LITE 가 유전자 발현 제어를 위한 유리한 선택을 나타내는 잠재적 적용이 존재한다. LITE 가 특히 유리한 수많은 시험관내 적용이 존재한다. 이러한 모든 경우, LITE 는 올바른 스플라이스 변이체 발현을 위한 가능성을 갖는 내인성 유전자 발현 유도의 이점을 갖는다.

[0229] LITE 활성화가 광유도성이기 때문에, 차폐 또는 래스터화 레이저 스캐닝을 통해 생성된 공간적으로 규정된 광 패턴을 사용하여 국한된 세포 하위집합에서의 발현 수준을 변형시킬 수 있다. 예를 들어, 공간적으로 제약된 세포 집합에서만 세포간 신호전달 분자를 과발현하거나 침묵화시켜, 자극 부위로부터 그의 거리에 관련하여 가까운 세포의 반응은 세포 비-자율적 프로세스의 공간적 특징을 해명하는 것을 도울 수 있다. 추가적으로, 세포 재프로그래밍 생물학에 있어서의 최근 진전은 전사 인자 집합의 과발현이 하나의 세포 유형, 예컨대 섬유모세포를 또 다른 세포 유형, 예컨대 뉴런 또는 심장근육세포로 형질전환하는데 이용될 수 있다는 것을 나타내고 있다. 추가로, 조직 내 세포 유형의 올바른 공간적 분포는 적절한 기관형 기능에 있어서 중요하다. LITE 를 사용하는 재프로그래밍 인자의 과발현은 조직 조각 적용을 위해 공간적으로 정확한 방식으로 다수의 세포 혈통을 재프로그래밍하는데 이용될 수 있다.

[0230] LITE 의 신속한 전사 반응 및 내인성 표적화는 전사 역학의 연구에 이상적인 시스템에 기여한다. 예를 들어, LITE 는 표적 유전자의 유도된 발현시 mRNA 스플라이스 변이체 생성의 역학 연구에 사용될 수 있다. 전사 주기의 다른 한편, mRNA 분해 연구는 종종, 과잉 유전자에서의 발현 수준 변화를 일으키는 강력한 세포의 자극에 대한 반응으로 수행된다. LITE 는 내인성 표적의 전사를 가역적으로 유도하는데 이용될 수 있으며, 그 후 자극이 중단될 수 있고 고유 표적의 분해 속도가 추적될 수 있다.

[0231] LITE 의 시간적 정확도는 실험적 개입과 맞물려 시간 유전적 조절에 대한 동력을 제공할 수 있다. 예를 들어, 장기간 강화작용 (LTP) 에서 예상된 관여를 갖는 표적은 기관형 또는 해리성 뉴런 배양물에 있어서, 그러나 LTP 를 유도하는 자극 동안에만 조정되어, 세포의 정상 발생의 간섭이 방지될 수 있다. 유사하게, 질환 표현형을 나타내는 세포 모델에서, 특정 치료요법의 유효성에 관여되는 것으로 예상되는 표적은 처리 동안에만 조정될 수 있다. 반대로, 유전적 표적은 병리적 자극 동안에만 조정될 수 있다. 외부 실험 자극에 대한

유전적 신호의 시기가 관련되는 임의의 수의 실험은 LITE 조정의 효용으로부터 잠재적으로 유익할 수 있다.

[0232] 생체내 맥락은 유전자 발현 제어를 위한 LITE 의 용도에 대해 동등하게 많은 기회를 제공한다. 상기 언급한 바와 같이, 광유도성은 이전에 성취불가능한 공간적 정확도에 대한 가능성을 제공한다. 광전극 기술의 발전을 이용하여, 자극 광 섬유 리드는 정확한 뇌 부위 내에 위치할 수 있다. 자극 부위 크기는 광 강도에 의해 조정될 수 있다. 이는 바이러스 벡터를 통한 LITE 의 전달과 함께 수행될 수 있거나, 트랜스제닉 LITE 동물이 이용가능해진 경우, 정확한 뇌 부위에서의 유전자 발현을 여전히 조정가능하게 하면서 바이러스의 사용을 없앨 수 있다. LITE 는 투명한 유기체, 예컨대 고정 제브라피시에서 사용되어 매우 정확한 레이저 유도 국부 유전자 발현 변화를 가능하게 할 수 있다.

[0233] LITE 는 또한 가치있는 생체내 시간적 정확도를 제공할 수 있다. LITE 는, 예를 들어 C. 엘레간스 성장의 특정 단계 동안에만 특정 세포자멸사 유전자를 억제하여, 발생의 특정 단계 동안 유전자 발현을 변형시키는데 사용될 수 있다. LITE 는 유전적 신호를 특정 실험 장에 대해 시기 조정하는데 사용될 수 있다. 예를 들어, 학습에 연루된 유전자는 미손상 설치류 또는 영장류 뇌의 정확도 부위에서의 학습 자극 동안에만 과발현되거나 억제될 수 있다. 또한, LITE 는 질환 발전의 특정 단계 동안에만 유전자 발현 변화를 유도하는데 사용될 수 있다. 예를 들어, 암유전자는 종양이 특정 크기 또는 전이 단계에 도달하고 단 한번 과발현될 수 있다. 반대로, 알츠하이머병 발전에서 의심된 단백질은 특정 뇌 부위 내에서 및 동물 생에서 규정된 시점에서만 녹다운될 수 있다. 이들 예가 LITE 시스템의 잠재적 적용을 철저히 열거하지는 않으나, 이들은 LITE 가 강력한 기술일 수 있는 일부 분야를 강조한다.

[0234] 실시예 3

[0235] 포유동물 TALE 전사 억제제의 개발

[0236] 출원인은 조사자가 내인성 유전자의 전사를 억제할 수 있도록 포유동물 TALE 억제제 구조를 개발하였다. TALE 억제제는 유전자 뿐 아니라 미-코딩 전사체 예컨대 마이크로RNA 를 억제하는 가능성을 가져, 이들이 특정 유전적 요소의 정상적 역할을 시험하기 위한 매우 바람직한 도구가 되게 한다. 포유동물 세포에서 TALE 과 함께 사용하기 위한 적합한 억제 도메인을 확인하기 위해, 인간 SOX2 유전자의 프로모터를 표적화하는 TALE 을 사용하여 억제 도메인 후보물 무리의 전사 억제 활성을 평가하였다 (도 12a). 진핵세포 숙주 종류 범위에 걸친 억제 도메인을 선택하여, 카에노르합디티스 엘레간스 (*Caenorhabditis elegans*) 로부터의 PIE-1 억제 도메인 (PIE-1) (Batchelder, C. et al. Transcriptional repression by the *Caenorhabditis elegans* germ-line protein PIE-1. *Genes Dev.* 13, 202-212 (1999)), 드로소필라 멜라노가스테르 (*Drosophila melanogaster*) 로부터의 Ubx 유전자 내 QA 도메인 (Ubx-QA) (Tour, E., Hittinger, C.T. & McGinnis, W. Evolutionarily conserved domains required for activation and repression functions of the *Drosophila* Hox protein Ultrabithorax. *Development* 132, 5271-5281 (2005)), 아라비도프시스 탈리아나로부터의 IAA28 억제 도메인 (IAA28-RD)(4), mSin 상호작용 도메인 (SID) (Ayer, D.E., Laherty, CD., Lawrence, Q.A., Armstrong, A.P. & Eisenman, R.N. Mad proteins contain a dominant transcription repression domain. *Mol. Cell. Biol.* 16, 5772-5781 (1996)), Tbx3 억제 도메인 (Tbx3-RD), 및 호모 사피엔스 (*Homo Sapiens*) 로부터의 크루펠-연관 박스 (KRAB) (Margolin, J.F. et al. Kruppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. U S A* 91, 4509-4513 (1994)) 억제 도메인을 포함하는 강력한 합성 억제제 발견 기회를 증가시켰다. KRAB 의 상이한 절두가 가변적 수준의 전사 억제를 나타내는 것으로 공지되어있기 때문에 (Margolin, J.F. et al. Kruppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. U S A* 91, 4509-4513 (1994)), KRAB 의 3 개 상이한 절두를 시험하였다 (도 12c). 이러한 TALE 억제제 후보물을 HEK 293FT 세포에서 발현시켰고, 이는 2 개의 광범위하게 사용되는 포유동물 전사 억제 도메인, SID (Ayer, D.E., Laherty, CD., Lawrence, Q.A., Armstrong, A.P. & Eisenman, R.N. Mad proteins contain a dominant transcription repression domain. *Mol. Cell. Biol.* 16, 5772-5781 (1996)) 및 KRAB (Margolin, J.F. et al. Kruppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. U S A* 91, 4509-4513 (1994)) 도메인을 운반하는 TALE 이 내인성 SOX2 발현을 억제할 수 있는 한편, 다른 도메인은 전사 활성에 대해 적은 효과를 갖는다는 것이 발견되었다 (도 12c). TALE 결합으로 인한 SOX2 전사의 잠재적 교란을 제어하기 위해서, 임의의 효과기 도메인을 갖지 않는 SOX2-표적화 TALE DNA 결합 도메인 단독은 SOX2 의 전사 활성에 대해 효과를 갖지 않았다 (GFP 의 발현 또는 모의물과 유사) (도 12c, Null 조건). SID 도메인이 KRAB 도메인보다 내인성 SOX2 유전자좌를 26% 더 전사 억제시킬 수 있었기 때문에 (도 12c), 후속 연구를 위해 SID 도메인을 사용하는 것으로 결정되었다.

- [0237] 내인성 전사의 하향 조절을 위한 SID 억제제 도메인의 유효성을 추가로 시험하기 위해서, SID 를 이전의 실험으로부터의 CACNA1C-표적 TALE 과 조합하였다 (도 12d). qRT-PCR 을 사용하여, CACNA1C-표적화 TALE 상의 VP64 도메인의 SID 로의 대체가 CACNA1C 전사를 억제할 수 있었다는 것이 발견되었다. NH-함유 TALE 억제제는 NN-함유 TALE 과 유사한 수준의 전사 억제를 달성할 수 있었던 한편 (~4 배 억제), NK 를 사용하는 TALE 억제제는 상당히 덜 활성이었다 (~2 배 억제) (도 12d). 이러한 데이터는 SID 가 실제로 적합한 억제 도메인이라는 것을 입증하면서, 또한 NK 보다 더 적합한 G-표적화 RVD 로서 NH 를 추가로 지지한다.
- [0238] TALE 은 내인성 게놈 상의 특정 서열을 인지하도록 용이하게 커스텀화될 수 있다. 본원에서, TALE 도구상자의 두 가지 중요한 한계를 다루기 위해 일련의 스크리닝을 수행하였다. 함께, 강건한 TALE 억제제 구조 뿐만 아니라 미절충된 활성 강도를 갖는 보다 엄중한 G-특이적 RVD 의 확인은 포유동물 전사 및 게놈 기능을 조사하기 위한 TALE 의 유용성을 추가로 확장시킨다.
- [0239] TALE 과 함께 사용할 강건한 신규 억제제 도메인으로서 SID (mSin 상호작용 도메인) 를 확인한 후, 포유동물 세포에서 TALE 과 함께 사용하기 위한 SID 도메인을 기반으로 하는 보다 활성인 억제제 도메인 구조를 또한 설계하고 확인하였다. 상기 도메인은 SID4X 로 지칭되며, 이는 짧은 랩티드 링커에 의해 연결된 4 개 SID 도메인의 탠덤 반복물이다. 상이한 TALE 억제제 구조를 시험하기 위해, 마우스 (*Mus musculus*) p11 (s100a10) 유전자의 프로모터를 표적화하는 TALE 을 사용하여, 일련의 후보 TALE 억제제 구조의 전사 억제 활성을 평가하였다 (도 13a). TALE 의 상이한 절두가 가변적 수준의 전사 활성화 활동을 나타내는 것으로 공지되어 있기 때문에, SID 또는 SID4X 도메인에 융합된 TALE 의 2 개 상이한 절두를 시험하였는데, 한 형태는 DNA 결합 탠덤 반복물에 측면 위치하는 N- 및 C-말단에서 136 및 183 개 아미노산을 갖고, 또 다른 형태는 N- 및 C-말단에서 240 및 183 개 아미노산을 보유한다 (도 13b, c). 후보 TALE 억제제를 마우스 Neuro2A 세포에서 발현시켰고, SID 및 SID4X 도메인 모두를 운반하는 TALE 이 내인성 p11 발현을 4.8 배까지 억제할 수 있는 한편, GFP-인코딩 음성 대조군 구축물은 표적 유전자의 전사에 효과를 갖지 않는다는 것이 발견되었다 (도 13b, c). TALE 결합으로 인한 p11 전사의 잠재적 교란을 제어하기 위해, 임의의 효과기 도메인을 갖지 않는 p11-표적화 TALE DNA 결합 도메인 (시험한 구축물과 동일한 N- 및 C-말단 절두를 가짐) 은 내인성 p11 의 전사 활성화에 대해 효과를 갖지 않았다 (도 13b, c, null 구축물).
- [0240] SID4X 도메인을 포함하는 구축물이 TALE DNA 결합 도메인의 절두에 따라 SID 도메인보다 내인성 p11 유전자좌를 167% 및 66% 더 전사 억제할 수 있었기 때문에 (도 13c), SID4X 도메인에 융합된, 각각 N- 및 C-말단에서 136 및 183 개 아미노산을 포함하는 절두된 TALE DNA 결합 도메인이 표적 유전자 발현의 하향 조절을 가능하게 하는 강력한 TALE 억제제 구조이며 SID 도메인을 이용하는 이전 설계보다 더 활성이라는 결론이 내려졌다.
- [0241] mSin 상호작용 도메인 (SID) 및 SID4X 도메인은 포유동물 발현에 대해 코돈 최적화되었으며 측면위치 NheI 및 XbaI 제한 부위 (Genscript) 로 합성되었다. TALE DNA 결합 도메인의 절두 변이체를 PCR 증폭하고 NheI 및 XbaI 제한 부위를 사용하여 SID 또는 SID4X 도메인에 융합시킨다. TALE 결합으로 생긴 전사에 대한 임의의 효과를 제어하기 위해서, PCR 클로닝을 사용하여 TALE DNA 결합 도메인 단독만을 운반하는 발현 벡터를 구축하였다. 모든 구축물의 코딩 부위를 Sanger 서열분석을 사용하여 완전히 확인하였다. 2 개 상이한 유형의 TALE 구조의 비교를 도 14 에 나타낸다.
- [0242] 실시예 4
- [0243] 포유동물 TALE 전사 활성제 및 뉴클레아제의 개발
- [0244] 커스텀화된 TALE 은 전사 조정 및 게놈 편집을 포함하는 광범위한 게놈 조작 적용에 사용될 수 있다. 본원에서, 출원인은 계층적 라이게이션 절차를 사용하는 커스텀 TALE 전사 인자 (TALE-TF) 및 뉴클레아제 (TALEN) 의 신속한 구축을 위한 도구상자를 기재하고있다. 상기 도구상자는 1 주일 내 커스텀 TALE-TF 및 TALEN 의 알맞고 신속한 구축을 촉진시키며, 동시에 다수의 표적에 대한 TALE 을 구축하기 위해 용이하게 규모 확대될 수 있다. 출원인은 또한 각각 정량 역전사 PCR 및 Surveyor 뉴클레아제를 사용하여 커스텀 TALE-TF 및 TALEN 의 포유동물 세포에서의 활성을 시험하기 위한 세부사항을 제공한다. TALE 도구상자는 광범위한 생물학적 적용을 가능하게 할 것이다.
- [0245] TALE 은 숙주 식물에서 유전자 전사를 조정하여 박테리아 군집화를 촉진시키기 위해 잔토모나스 종에 의해 사용된 자연적 박테리아 효과기 단백질이다 (Boch, J. & Bonas, U. *Xanthomonas AvrBs3 family-type III effectors: discovery and function*. *Annu. Rev. Phytopathol.* 48, 419-436 (2010) and Bogdanove, A. J., Schornack, S. & Lahaye, T. *TAL effectors: finding plant genes for disease and defense*. *Curr. Opin.*

Plant Biol. 13, 394-401 (2010)). 단백질의 중심 부위는 34-aa 서열의 탠덤 반복물을 함유하며 (단량체로 지칭) 이는 DNA 인지 및 결합에 필요하다 (Romer, P. et al. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. Science 318, 645-648 (2007); Kay, S., Hahn, S., Marois, E., Hause, G. & Bonas, U. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. Science 318, 648-651 (2007); Kay, S., Hahn, S., Marois, E., Wieduwild, R. & Bonas, U. Detailed analysis of the DNA recognition motifs of the Xanthomonas type III effectors AvrBs3 and AvrBs3Deltarepl6. Plant J. 59, 859-871 (2009) and Romer, P. et al. Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper Bs3 alleles. Plant Physiol. 150, 1697-1712 (2009).) (도 8). 자연 발생적 TALE 은 1.5 내지 33.5 개 범위의 가변적 수의 단량체를 갖는 것으로 발견되었다 (Boch, J. & Bonas, U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annu. Rev. Phytopathol. 48, 419-436 (2010)). 각각의 단량체의 서열이 고도로 보존됨에도 불구하고, 이들은 반복 가변성 2 잔기로 지칭되는 2 개 위치에 있어서 주로 상이하다 (RVD, 12 번째 및 13 번째 위치). 최근 보고서는 이들 2 개 잔기의 동일성이 각각의 TALE 반복물의 뉴클레오티드-결합 특이성을 결정하며 간단한 암호가 각각의 RVD 의 표적 염기를 명시한다는 것을 발견한 바 있다 (NI = A, HD = C, NG = T, NN = G 또는 A) (Boch, J. et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326, 1509-1512 (2009) and Moscou, M.J. & Bogdanove, A.J. A simple cipher governs DNA recognition by TAL effectors. Science 326, 1501 (2009)). 따라서, 각각의 단량체는 하나의 뉴클레오티드를 표적화하고, TALE 에서의 단량체 선형 서열은 5' 에서 3' 배향으로 표적 DNA 서열을 명시한다. 식물 게놈 내의 자연적 TALE-결합 부위는 항상 티민으로 시작하는데 (Boch, J. et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326, 1509-1512 (2009) and Moscou, M.J. & Bogdanove, A.J. A simple cipher governs DNA recognition by TAL effectors. Science 326, 1501 (2009)), 이는 아마도 TALE 의 비반복적 N 말단 내 암호성 신호에 의해 명시된다. 탠덤 반복물 DNA-결합 도메인은 항상 반-길이 반복물로 종료된다 (0.5 반복물, 도 8). 따라서, 표적화되는 DNA 서열의 길이는 전체 반복 단량체의 수 + 2 와 동일하다.

[0246] 식물에서, 병원체는 종종 숙주-특이적이다. 예를 들어, 푸사리움 옥시스포룸 (*Fusarium oxysporum*) f. sp. 리코페르시시 (*lycopersici*) 는 토마토 시들병을 일으키나 오직 토마토만을 공격하고, F. 옥시스포룸 f. 디안티이 (*dianthii*) 푸시니아 그라미니스 (*Puccinia graminis*) f. sp. 트리티시 (*tritici*) 는 오직 밀만을 공격한다. 식물은 대부분의 병원체에 저항하기 위한 현존하며 유도된 방어 기작을 갖는다. 식물 세대에 걸친 돌연변이 및 재조합 이벤트는, 특히 병원체가 식물보다 더 많은 빈도로 복제하는 바, 민감성이 생기게 하는 유전적 변이성을 일으킨다. 식물에서는 비-숙주 저항성이 존재할 수 있는데, 예를 들어 숙주 및 병원체는 양립불가하다. 또한 수평 저항성, 예를 들어, 통상 많은 유전자에 의해 제어되는 모든 종류의 병원체에 대한 부분적 저항성, 및 수직 저항성, 예를 들어, 통상 소수 유전자에 의해 제어되는 일부 종류의 병원체에 대한 전체 저항성 (다른 종류에 대해서는 아님) 이 존재할 수 있다. 유전자 대 유전자 (Gene-for-Gene) 수준에 있어서, 식물 및 병원체는 함께 진화하며, 한 나머지에서의 유전적 변화는 다른 것에서 변화한다. 따라서, 자연적 가변성을 사용하여, 육종자는 수율, 품질, 균일성, 강인성, 저항성에 대해 가장 유용한 유전자를 조합한다. 저항성 유전자의 근원은 선천적 또는 외래 품종, 조상 전래 (Heirloom) 품종, 야생 식물 동류 및 유도된 돌연변이 (예를 들어 돌연변이유발제로 식물 물질 처리) 를 포함한다. 본 발명을 사용하여, 식물 육종자에게 돌연변이를 유도하는 새로운 도구가 제공된다. 따라서, 당업자는 저항성 유전자 근원의 게놈을 분석할 수 있으며, 원하는 특징 또는 특성을 갖는 품종에서 본 발명을 이용하여 이전의 돌연변이유발제보다 더 많은 정확도로 저항성 유전자의 발생을 유도하여, 이에 따라 식물 육종 프로그램을 가속화하고 향상시킨다.

[0247] 출원인은, 잘못된 라이게이션이 최소화되도록 하는 라이게이션 어댑터의 비유사성 최대화 및 별개의 소화 및 라이게이션 단계의 단일 골든 게이트 (Golden Gate) (Engler, C, Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. PLoS ONE 3, e3647 (2008); Engler, C, Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIS restriction enzymes. PLoS ONE 4, e5553 (2009) and Weber, E., Engler, C, Gruetzner, R., Werner, S. & Marillonnet, S. A modular cloning system for standardized assembly of multigene constructs. PLoS ONE 6, e16765 (2011)) 반응으로의 조합을 비롯한 소수의 최적화를 갖는 TALE 어셈블리 시스템을 추가로 향상시킨다. 간단하게, 각각의 뉴클레오티드-특이적 단량체 서열은 TALE 탠덤 반복물 내의 단량체 위치를 고유하게 명시하는 라이게이션 어댑터로 증폭된다. 이러한 단량체 라이브러리가 생성되고나면, 이는 많은 TALE 의 어셈블리를 위해 편리하게 재사용될 수 있다. 필요한 각각의 TALE 에 대해서, 적절한 단

량체가 먼저 육량체에 라이게이션되고, 그 다음 PCR 을 통해 증폭된다. 이후, 적절한 TALE 클로닝 백본을 갖는 두 번째 끝의 게이트 소화-라이게이션 (도 8) 으로, 완전히 어셈블리된, 서열-특이적 TALE 이 산출된다.

백본은 TALE N 및 C 말단에 의해 측면 위치한 ccdB 음성 선택 카세트를 함유하는데, 이는 TALE 이 성공적으로 구축되는 경우 탠덤 반복물 DNA-결합 도메인에 의해 대체된다. ccdB 는 빈 백본으로 형질전환된 세포에 비하여 선택하여, 이로써 탠덤 반복물이 삽입된 클론이 산출된다 (Cermak, T. et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 39, e82 (2011)).

[0248] 단량체 DNA-결합 도메인의 어셈블리는 적절한 TALE-TF 또는 TALEN 클로닝 백본 내로 삽입되어 커스텀화 TALE-TF 및 TALEN 이 구축될 수 있다. TALE-TF 는 TALE C 말단 내 자연적 활성화 도메인을 합성 전사 활성화 도메인 VP64 로 대체함으로써 구축된다 (Zhang, F. et al. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat. Biotechnol. 29, 149-153 (2011); 도 8). 전사 시작 부위의 결합 부위 업스트림을 표적화함으로써, TALE-TF 는 부위-특이적 방식으로 전사 복합체를 영입하고 유전자 전사를 개시한다. TALEN 은 TALE DNA-결합 도메인의 C-말단 절두 (+63 aa) (Miller, J.C. et al. A TALE nuclease architecture for efficient genome editing. Nat. Biotechnol. 29, 143-148 (2011)) 를 비롯하여 FokI 엔도뉴클레아제 촉매 도메인과 융합하여 구축된다 (도 14). +63-aa C-말단 절두는 또한 전사 조절에 충분한 최소 C 말단으로서 기능하는 것으로 나타난다 (Zhang, F. et al. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat. Biotechnol. 29, 149-153 (2011)). TALEN 은 ~17 개 염기에 의해 분리되는 2 개 표적 서열에 대한 결합을 통해 이량체를 형성한다.

결합 부위쌍 사이에서, FokI 촉매 도메인은 이량체화되며 이중 가닥 파단물 (DSB; 도 8) 을 도입함으로써 분자 가위로서 기능한다. 보통, DSB 는 비상동 말단-연결 (Huertas, P. DNA resection in eukaryotes: deciding how to fix the break. Nat. Struct. Mol. Biol. 17, 11-16 (2010)) 경로 (NHEJ) 에 의해 보수되어, 소 결실 및 기능적 유전자 녹아웃을 초래한다. 대안적으로, TALEN-매개 DSB 는 상동 재조합을 자극할 수 있어, 외인성 공여체 DNA 주형의 부위-특이적 삽입이 가능하게 한다 (Miller, J.C. et al. A TALE nuclease architecture for efficient genome editing. Nat. Biotechnol. 29, 143-148 (2011) and Hockemeyer, D. et al. Genetic engineering of human pluripotent cells using TALE nucleases. Nat. Biotechnol. 29, 731-734 (2011)).

[0249] VP64 활성화 도메인으로 구축되는 TALE-TF 와 함께, 본 발명의 다른 구현에는 VP16 및 p65 활성화 도메인으로 구축되는 TALE 폴리펩티드에 관한 것이다. 이러한 상이한 활성화 도메인의 Sox2 mRNA 수준에 대한 효과의 그래프적 비교를 도 11 에서 제공한다.

[0250] 실시예 5

[0251] 도 17 은 LITE 기능성에 대한 크립토크롬2 이형이량체 배향의 효과를 나타낸다. 뉴로제닌 (Neurogenin) 2 (Neurog2) LITE 의 2 개 형태를 합성하여 크립토크롬 2 포토리아제 상동성 부위 (CRY2PHR)/칼슘 및 인테그린-결합 단백질 1 (CIB1) 이량체 배향의 효과를 조사하였다. 한 형태에서, CIB1 도메인을 TALE (Neurog2) 도메인의 C-말단에 융합시킨 한편, CRY2PHR 도메인을 VP64 도메인의 N-말단에 융합시켰다. 반대 형태에서, CRY2PHR 도메인을 TALE (Neurog2) 도메인의 C-말단에 융합시킨 한편, CIB1 도메인을 VP64 도메인의 N-말단에 융합시켰다. 각각의 플라스미드 집합을 Neuro2a 세포에서 트랜스펙션시키고 자극한 후 (466 nm, 5 mW/cm<sup>2</sup>, 15 초 당 1 초 펄스, 12 시간) qPCR 분석을 위해 수확하였다. 자극 LITE 및 미자극 LITE Neurog2 발현 수준을 자극 GFP 대조군 샘플로부터의 Neurog2 수준으로 정규화시켰다. TALE-CRY2PHR/CIB1-VP64 LITE 는 상승된 기저 활성 및 높은 광 유도 Neurog2 발현을 나타내었는데, 이는 더 높은 절대 활성도가 필요한 상황에 대한 이의 적합성을 제시하였다. TALE-CIB1/CRY2PHR-VP64 LITE 의 상대적 광 유도 활성이 그의 대응물보다 낮았으나, 낮은 기저 활성은 최소 기준선 활성화를 필요로 하는 적용에서의 이의 유용성을 제시하였다. 또한, TALE-CIB1 구축물은 TALE-CRY2PHR 구축물에 비해 크기가 더 작았고, 바이러스 패키징과 같은 적용에 대해 잠재적 이점을 갖는다.

[0252] 도 18 은 마우스 피질 뉴런 배양물에서의 대사자극성 글루타메이트 수용체 2 (mGlu2) LITE 활성을 나타낸다. mGlu2 표적화 LITE 를 플라스미드 pAAV-간간 시냅신 I 프로모터 (hSyn)-HA-TALE(mGlu2)-CIB1 및 pAAV-hSyn-CRY2PHR-VP64-2A-GFP 를 통해 구축하였다. 이러한 융합 구축물을 이후 아데노 연관 바이러스 벡터 (AAV) 내로 패키징하였다. 추가적으로, hSyn-TALE-VP64-2A-GFP 및 GFP 만을 운반하는 AAV 가 생성되었다. 배아 마우스 (E16) 피질 배양물을 폴리-L-리신 코팅된 24 웰 플레이트 상에 플레이팅하였다. 5 일 후 시

시험관내 뉴런 배양물을 TALE(mGluR2)-CIB1 및 CRY2PHR-VP64 AAV 스톡 (stock) 의 혼합물로 동시 형질도입하였다. 대조군 샘플을 TALE(mGluR2)-VP64 AAV 또는 GFP AAV 로 형질도입하였다. AAV 형질도입 6 일 후, 실험 샘플을 2 가지 광 펄스 패러다임 중 어느 하나를 사용하여 자극하였다: 분 당 0.5 초 및 30 초 당 0.25 초. 뉴런을 24 시간 동안 자극하고 qPCR 분석을 위해 수확하였다. 모든 mGluR2 발현 수준을 각각의 자극 GFP 대조군에 대해 정규화하였다. 데이터는 LITE 시스템이 시험관내 1 차 뉴런 배양물 중 표적 유전자의 광-의존적 활성화를 유도하는데 사용될 수 있다는 것을 제시하였다.

[0253] 도 19 는 LITE AAV 벡터로의 1 차 마우스 뉴런의 형질도입을 나타낸다. 1 차 마우스 피질 뉴런 배양물을 LITE 시스템의 2 개 성분인 hSyn-CRY2PHR-VP64-2A-GFP 및 hSyn-HA-TALE-CIB1 을 인코딩하는 AAV 벡터로 5 일에 시험관내 동시-형질도입하였다. 좌측 패널: 형질도입 6 일 후, 뉴런 배양물은 hSyn-CRY2PHR-VP64-2A-GFP 벡터로부터 GFP 의 높은 발현을 나타내었다. 우측 패널: 동시-형질도입된 뉴런 배양물을 고정시키고 hSyn-HA-TALE-CIB1 에서의 TALE 도메인의 N-말단 상 HA 에피토프에 특이적인 항체로 염색하였다. 적색 신호는 HA 발현을 나타내었는데, 특히 핵 신호가 강력하였다 (청색 채널에서의 DAPI 에 의해 염색된 DNA). 이러한 영상은 함께, 각 LITE 성분의 발현이 1 차 마우스 뉴런 배양물에서 이루어질 수 있다는 것을 제시하였다 (스케일 바 =50 um).

[0254] 도 20 은 생체내 LITE 성분의 발현을 나타낸다. hSyn-CRY2PHR-VP64 를 운반하는 혈청형 1/2 의 AAV 벡터를 HEK293FT 세포의 트랜스펙션을 통해 생성시키고, 헤파린 컬립 결합을 통해 정제하였다. 미손상 마우스 뇌 내로의 주입을 위해 벡터를 농축하였다. 1 uL 의 정제된 AAV 스톡을, 정위 수술 및 주사에 의해 8 주령 수컷 C57BL/6 마우스의 해마 및 변연계 아래 피질에 주사하였다. 생체내 형질도입 7 일 후, 마우스를 안락사시키고 뇌 조직을 파라포름알데히드 관류에 의해 고정하였다. 뇌 박편을 마이브라톱 상에서 제조하고, 영상화를 위해 장착하였다. 해마 및 변연계 아래 피질에서의 강력하고 광범위한 GFP 신호는 LITE 성분 CRY2PHR-VP64 의 효율적 형질도입 및 높은 발현을 제시하였다.

[0255] 실시예 6

[0256] NES 요소 사용에 의한 향상된 설계

[0257] 에스트로젠 수용체 T2 (ERT2) 는 누수 쟁점을 갖는다. ERT2 도메인은 4-히드록시테스토스테론 (4OHT) 부재 하에서도 핵에 유입될 수 있어, TAL 에 의한 표적 유전자 활성화의 배경 수준을 초래한다. NES (핵 이출 신호) 는 살아있는 세포의 세포질에 대해 단백질을 표적화하는 펩티드 신호이다. NES 를 현존 구축물에 첨가함으로써, 출원인은 4OHT 의 부재 하 ERT2-TAL 단백질이 핵 내로의 유입을 방지하여, ERT2 도메인의 "누수" 로 인한 배경 활성화 수준을 저하시키는 것을 목적으로 한다.

[0258] 도 21 은 사용한 특정 NES 펩티드 서열이 LDLASLIL 인 구축물의 향상된 설계를 나타낸다.

[0259] 도 22 는 40H 타목시펜의 부재 및 존재 하에서의 Sox2 mRNA 수준을 나타낸다. Y-축은 qRT-PCR 에 의해 측정된 바와 같은 Sox2 mRNA 수준이다. X-축은 상부에서 기재한 상이한 구축물 설계의 패널이다. + 및 - 표시는 0.5uM 4OHT 의 존재 또는 부재를 나타낸다.

[0260] 실시예 7

[0261] CRISPR/Cas 시스템을 사용하는 멀티플렉스 게놈 조작

[0262] 임상적 유전자 변이체 및 요소의 기능적 해명은 정확한 게놈 편집 기술을 필요로 한다. 유형 II 원핵생물 CRISPR (군집화 규칙적 공간화 단길이 회귀성 반복물) 조정 면역계는 RNA-안내 부위-특이적 DNA 절단을 촉진시키는 것으로 나타난 바 있다. 출원인은 2 개 상이한 유형 II CRISPR 시스템을 조작하였고, Cas9 뉴클레아제가 단길이 RNA 에 의해 유도되어 인간 및 마우스 세포에서 내인성 게놈 유전자좌에서의 정확한 절단을 유도할 수 있다는 것을 입증한다. Cas9 는 또한 절단 효소로 전환되어 최소의 돌연변이유발 활성으로 상동성-유도 수복을 촉진시킬 수 있다. 마지막으로, 다수의 안내 서열은 단일 CRISPR 어레이에 인코딩되어 포유동물 게놈 내의 여러 부위의 동시 편집을 가능하게 하는데, 이는 CRISPR 기술의 용이한 프로그램성 및 광범위한 적용성을 입증한다.

[0263] 원핵생물 CRISPR 적응 면역계는 재구성되고 조작되어 포유동물 세포에서의 멀티플렉스 게놈 편집을 매개할 수 있다.

[0264] 정확하고 효율적인 게놈 표적화 기술이 개별적 유전 요소의 선택적 교란을 가능하게 함으로써 임상적 유전적 변이의 조직적 역조작을 가능하게 하기 위해서 필요하다. 게놈-편집 기술 예컨대 설계자 아연 핑거 (ZF) (M.

H. Porteus, D. Baltimore, Chimeric nucleases stimulate gene targeting in human cells. *Science* 300, 763 (May 2, 2003); J. C. Miller et al., An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* 25, 778 (Jul, 2007); J. D. Sander et al, Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods* 8, 67 (Jan, 2011) and A. J. Wood et al, Targeted genome editing across species using ZFNs and TALENs. *Science* 333, 307 (Jul 15, 2011), 전사 활성제-유사 효과기 (TALE) (A. J. Wood et al, Targeted genome editing across species using ZFNs and TALENs. *Science* 333, 307 (Jul 15, 2011); M. Christian et al., Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757 (Oct, 2010); F. Zhang et al, Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* 29, 149 (Feb, 2011); J. C. Miller et al., A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* 29, 143 (Feb, 2011); D. Reyon et al, FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol* 30, 460 (May, 2012); J. Boch et al., Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509 (Dec 11, 2009) and M. J. Moscou, A. J. Bogdanove, A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501 (Dec 11, 2009)), 및 호밍 (homing) 메가뉴클레아제 (B. L. Stoddard, Homing endonuclease structure and function. *Quarterly reviews of biophysics* 38, 49 (Feb, 2005)) 가 표적화 게놈 변형이 가능하도록 시작함에도 불구하고, 확장가능하고, 알맞고, 조작하기 용이한 새로운 기술에 대한 필요성이 존재한다. 여기서, 출원인은 유형 II 원핵생물 CRISPR 적응 면역계 (H. Deveau, J. E. Garneau, S. Moineau, CRISPR/Cas system and its role in phage-bacteria interactions. *Annual review of microbiology* 64, 475 (2010); P. Horvath, R. Barrangou, CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327, 167 (Jan 8, 2010); K. S. Makarova et al., Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 9, 467 (Jun, 2011) and D. Bhaya, M. Davison, R. Barrangou, CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet* 45, 273 (2011)) 로부터의 RNA-유도 Cas9 뉴클레아제 기반의 정확도 게놈 조작 도구의 새로운 클래스의 개발을 보고하고있다 (M. Jinek et al, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816 (Aug 17, 2012); G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci USA* 109, E2579 (Sep 25, 2012) and J. E. Garneau et al, The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67 (Nov 4, 2010)).

[0265]

스트렙토코쿠스 피오게네스 SF370 유형 II CRISPR 유전자좌는 Cas9 뉴클레아제 뿐 아니라 2 개 비-코딩 RNA: tracrRNA 및 예비-crRNA 어레이 (동일한 직접적 반복물 (DR) 에 의해 공간이 남겨진 뉴클레아제 안내 서열 (스페이서) 을 함유함) 를 포함하는 4 개 유전자로 이루어진다 (도 27) (E. Deltcheva et al, CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602 (Mar 31, 2011)). 출원인은 핵심 성분의 이중 발현을 통해 포유동물 염색체에 표적화 이중 가닥 파단물 (DSB) 을 도입하기 위해 이러한 원핵생물 RNA-프로그램가능 뉴클레아제 시스템을 이용하는 것을 탐색하였다. tracrRNA, 예비-crRNA, 숙주 인자 RNase III 및 Cas9 뉴클레아제의 발현이 시험관내 DNA 절단 (M. Jinek et al., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816 (Aug 17, 2012) and G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci US A* 109, E2579 (Sep 25, 2012)) 및 원핵세포 (R. Sapranauskas et al, The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* 39, 9275 (Nov, 2011) and A. H. Magadan, M. E. Dupuis, M. Villion, S. Moineau, Cleavage of phage DNA by the *Streptococcus thermophilus* CRISPR3-Cas system. *PLoS One* 7, e40913 (2012)) 에서 필요하며 그에 충분하다는 것이 이전에 나타난 바 있다. 출원인은 S. 피오게네스 Cas9 (SpCas9) 및 RNase III (SpRNase III) 를 코돈 최적화하고 핵 국부화 신호 (NLS) 를 부착하여 포유동물 세포에서의 핵 구획화를 확실히 하였다. 인간 293FT 세포에서의 이러한 구축물의 발현이 SpCas9 를 핵에 대해 표적화하는데 필요하다는 것이 밝혀졌다 (도 23A). CRISPR 의 비-코딩 RNA 성분을 재구성하기 위해, 출원인은 RNA 중합효소 III U6 프로모터 하 (도 23B) 에 89-뉴클레오티드 (nt) tracrRNA (도 28) 를 발현시켰다. 유사하게, 출원인은 DR 에 측면 위치한 단일 안내 스페이서를 포함하는 예비-crRNA 어레이의 발현을 유도하기 위해 U6 프로모터를 사용하였다 (도 23B). 출원인은 NGG 에 선행하는 인간 EMX1 유전자좌에서 30-염기쌍 (bp) 부위를 표적화하기 위해 필요한 프로토스페이서 인접 모티프 (PAM) 인 초기 스페이서 (프로

토스페이서) 를 설계하였다 (도 23C 및 도 27) (H. Deveau et al., Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol* 190, 1390 (Feb, 2008) and F. J. Mojica, C. Diez-Villasenor, J. Garcia-Martinez, C. Almendros, Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733 (Mar, 2009)).

[0266] CRISPR 시스템 (SpCas9, SpRNase III, tracrRNA 및 예비-crRNA) 의 이중 발현이 포유동물 염색체의 표적화 절단을 달성할 수 있는지를 시험하기 위해, 출원인은 CRISPR 성분의 상이한 조합으로 293FT 세포를 트랜스펙션시켰다. 포유동물 DNA 에서의 DSB 가 삽입결실-형성 비-상동 말단 연결 (NHEJ) 경로에 의해 부분적으로 수복되기 때문에, 출원인은 내인성 표적 절단이 검출되도록 SURVEYOR 검정 (도 29) 을 사용하였다 (도 23D 및 도 28B). 모든 4 개의 필요한 CRISPR 성분의 동시-트랜스펙션은 프로토스페이서의 효율적 절단을 초래하였으며 (도 23D 및 도 28B), 이는 Sanger 서열분석에 의해 이후에 확인되었다 (도 23E). 흥미롭게도, SpRNase III 은 프로토스페이서의 절단에 필요하지 않았으며 (도 23D), 89-nt tracrRNA 는 이의 부재 하에 진행된다 (도 28C). 유사하게, 예비-crRNA 의 성숙은 RNase III (도 23D 및 도 30) 을 필요로 하지 않으며, 이는 예비-crRNA 성숙을 보조하는 내인성 포유동물 RNase가 존재할 수 있다는 것을 제시한다 (M. Jinek, J. A. Doudna, A three-dimensional view of the molecular machinery of RNA interference. *Nature* 457, 405 (Jan 22, 2009); C. D. Malone, G. J. Hannon, Small RNAs as guardians of the genome. *Cell* 136, 656 (Feb 20, 2009) and G. Meister, T. Tuschl, Mechanisms of gene silencing by double-stranded RNA. *Nature* 431, 343 (Sep 16, 2004)). 임의의 남아 있는 RNA 또는 Cas9 성분을 제거하는 것은 CRISPR 시스템의 게놈 절단 활성을 없앴다 (도 23D). 이러한 결과는 포유동물 세포에서 효율적 CRISPR-매개 게놈 변형을 위한 최소의 3-성분 시스템을 규정한다.

[0267] 다음으로, 출원인은 EMX1 유전자좌 내의 추가적인 프로토스페이서를 표적화함으로써 진핵세포에서의 CRISPR-매개 절단의 일반화가능성을 탐색하였다 (도 24A). 동시-전달을 향상시키기 위해, 예비-crRNA 및 SpCas9 모두를 유도하도록 발현 벡터를 설계하였다 (도 31). 동시에, 출원인은 최근 시험관내 인증된 키메라 crRNA-tracrRNA 하이브리드 (도 24B, 상부) 설계를 조정하였는데 (M. Jinek et al., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816 (Aug 17, 2012)), 이때 자연적 crRNA:tracrRNA 듀플렉스를 모방하기 위해 성숙 crRNA 가 합성 줄기-루프를 통해 부분적 tracrRNA 에 융합된다 (도 24B, 하부). 출원인은 SpCas9 가 예비-crRNA (DR-스페이서-DR) 및 tracrRNA 와 동시-발현되는 경우 모든 프로토스페이서 표적의 절단을 관찰하였다. 그러나, 모든 키메라 RNA 설계가 그의 게놈 표적의 절단을 촉진시킬 수 있는 것은 아니다 (도 24C, 표 1). 출원인은 이후 인간 PVALB 및 마우스 Th 유전자좌를 표적화하는 예비-crRNA 및 키메라 RNA 를 설계함으로써 인간 및 마우스 세포 모두에서 추가적인 게놈 유전자좌의 표적화를 시험하였다 (도 32). 출원인은 crRNA:tracrRNA 설계를 사용하여 모든 3 개의 마우스 Th 및 1 개의 PVALB 표적에서 효율적 변형을 달성하였으며, 따라서 다수 유기체에 걸쳐 상이한 유전자좌를 변형시키는데 있어서 CRISPR 시스템의 광범위한 적용가능성을 입증한다 (표 1). 동일한 프로토스페이서 표적에 대해서, 키메라 RNA 의 절단 효율은 crRNA:tracrRNA 듀플렉스의 효율보다 더 낮거나 검출가능하지 않았다. 이는 RNA 의 발현 및 안정성에 있어서의 차이, 내인성 RNAi 장비에 의한 분해, 또는 비효율적 Cas9 로딩 또는 표적 인지를 초래하는 2 차 구조로 인한 것일 수 있다.

[0268] 효율적 게놈 편집은 높은 정확도 및 효율 모두를 갖는 뉴클레아제 표적 특이적 게놈 유전자좌를 필요로 한다. CRISPR-매개 절단의 특이성을 조사하기 위해서, 출원인은 스페이서 및 이의 포유동물 프로토스페이서 표적 사이의 단일-뉴클레오티드 미스매치를 분석하였다 (도 25A). 출원인은 PAM 의 5' 12-bp 이하의 단일-염기 미스매치가 SpCas9 에 의한 게놈 절단을 완전히 없앤 반면, 더 먼 업스트림 돌연변이를 갖는 스페이서가 프로토스페이서 표적에 대한 활성을 보유하고 있다는 것을 관찰하였다 (도 25B). 이는 Cas9 특이성의 이전의 박테리아 및 시험관내 연구와 일치한다 (M. Jinek et al., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816 (Aug 17, 2012) and R. Sapranaukas et al., The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* 39, 9275 (Nov, 2011)). 추가로, CRISPR 은 동일한 EMX1 프로토스페이서를 표적화하는 TALE 뉴클레아제 (TALEN) 쌍만큼 효율적으로 게놈 절단을 매개할 수 있다 (도 25, C 및 D).

[0269] 게놈의 표적화된 변형은 이상적으로 실수-유발 NHEJ 메커니즘에서 발생하는 돌연변이를 방지한다. 야생형 SpCas9 는 부위-특이적 DSB 를 매개할 수 있는데, 이는 NHEJ 또는 상동성-유도 수복 (HDR) 을 통해 수복될 수 있다. 출원인은 SpCas9 의 RuvC I 도메인에서의 아스파르테이트-알라닌 치환 (D10A) 을 조작하여 뉴클레아제를 DNA 니카아제로 전환시켰는데 (SpCas9n, 도 26A) (M. Jinek et al., A programmable dual-RNA-guided DNA

endonuclease in adaptive bacterial immunity. *Science* 337, 816 (Aug 17, 2012); G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* 109, E2579 (Sep 25, 2012) and R. Sapranaukas et al, The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* 39, 9275 (Nov, 2011)), 이는 니킹된 게놈 DNA 가 균일하게 또는 고-적합도 HDR 을 통해 통상 수복되기 때문이다. SURVEYOR (도 26B) 및 327 개 암플리콘의 서열분석은 SpCas9n 에 의해 유도된 임의의 삽입-결실을 검출하지 못했다. 그러나, 니킹된 DNA 가 드문 경우 DSB 중간체를 통해 진행될 수 있으며 NHEJ 이벤트를 초래한다는 것은 주목할만 하다 (M. T. Certo et al., Tracking genome engineering outcome at individual DNA breakpoints. *Nat Methods* 8, 671 (Aug, 2011)). 출원인은 프로토스페이서 근처의 제한 부위 쌍을 도입하기 위해 상동성 수복 주형을 갖는 동일한 EMX1 유전자좌에서의 Cas9-매개 HDR 을 시험하였다 (도 26C). SpCas9 및 SpCas9n 은 유사한 수준에서의 EMX1 유전자좌 내로의 수복 주형의 통합을 촉진하였는데 (도 26D), 출원인은 Sanger 서열분석을 통해 추가로 확인하였다 (도 26E). 이러한 결과는 표적화된 게놈 삽입을 촉진시키기 위한 CRISPR 의 유용성을 입증한다. 야생형 SpCas9 의 14-bp (시드 (seed) 서열로부터 12-bp 및 PAM 으로부터 2-bp) 표적 특이성 (도 25B) 을 고려하여, 니카아제의 사용은 오프-타겟 돌연변이를 감소시킬 수 있다.

[0270] 마지막으로, 어레이된 스페이서를 갖는 CRISPR 유전자좌의 자연적 구조 (도 27) 는 멀티플렉스화 게놈 조작의 가능성을 제시한다. EMX1- 및 PVALB-표적화 스페이서 쌍을 인코딩하는 단일 CRISPR 어레이를 사용하여, 출원인은 유전자좌 모두에서 효율적 절단을 검출하였다 (도 26F). 출원인은 119-bp 간격의 EMX1 내 2 개 표적에 대한 스페이서를 사용하여 동시발생 DSB 를 통해 더 큰 게놈 부위의 표적화된 결실을 추가로 시험하였고, 1.6% 결실 효능을 관찰하였으며 (182 개 암플리콘 중 3 개; 도 26G), 그에 따라 CRISPR 시스템이 단일 게놈 내에 멀티플렉스화 편집을 매개할 수 있다는 것을 입증하였다.

[0271] 서열-특이적 DNA 절단을 프로그래밍하도록 RNA 를 사용하는 능력은 새로운 클래스의 게놈 조작 도구를 규정한다. 여기서, 출원인은 효율적 게놈 편집이 촉진되도록 *S. 피오게네스* CRISPR 시스템이 포유동물 세포에서 이종으로 재구성될 수 있다는 것을 나타내며; 동반 연구는 여러 인간 세포주에서 고효율 CRISPR-매개 게놈 표적화를 독립적으로 확인하였다 (Mali et al.). 그러나, CRISPR 시스템의 여러 양태는 이의 효율 및 다재다능성이 증가되도록 추가로 향상될 수 있다. NGG PAM 에 대한 필요조건은 크로마틴 및 DNA 메틸화 상태로 인한 crRNA 2 차 구조 또는 게놈 접근성에 의해 제기된 잠재적 제약을 해명하지 않고 인간 게놈에서 평균 8-bp 마다로 *S. 피오게네스* CRISPR 표적 공간을 제한한다 (도 33). 이러한 제한 중 일부는 미생물적 다양성에 걸친 (K. S. Makarova et al., Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 9, 467 (Jun, 2011)) 이의 상이한 PAM 필요조건 및 Cas9 효소 패밀리를 이용함으로써 극복될 수 있다 (H. Deveau et al., Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol* 190, 1390 (Feb, 2008) and F. J. Mojica, C. Diez-Villasenor, J. Garcia-Martinez, C. Almendros, Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733 (Mar, 2009)). 실제로, 다른 CRISPR 유전자좌는 포유동물 세포 내로 이식될 가능성이 있으며; 예를 들어, 스트렙토코쿠스 써모필루스 LMD-9 CRISPR1 은 또한 포유동물 게놈 절단을 매개할 수 있다 (도 34). 마지막으로, 포유동물 세포에서의 멀티플렉스 게놈 편집을 실행하는 능력은 기초 과학, 생물공학 및 의학에 걸쳐 강력한 적용을 가능하게 한다 (P. A. Carr, G. M. Church, Genome engineering. *Nat Biotechnol* 27, 1151 (Dec, 2009)).

[0272] 실시예 8

[0273] CRISPR/Cas 시스템을 사용하는 멀티플렉스 게놈 조작: 보충 물질

[0274] 세포 배양 및 트랜스펙션. 인간 배아 신장 (HEK) 세포주 293FT (Life Technologies) 를 10% 소 태아 혈청 (HyClone), 2mM GlutaMAX (Life Technologies), 100U/mL 페니실린 및 100µg/mL 스트렙토마이신이 보충된 돌베코 개질 이글 배지 (DMEM) 에서 37°C 에서 5% CO2 인큐베이션하면서 유지시켰다. 마우스 neuro2A (N2A) 세포주 (ATCC) 를 37°C 에서 5% CO2 와 함께, 5% 소 태아 혈청 (HyClone), 2mM GlutaMAX (Life Technologies), 100U/mL 페니실린 및 100µg/mL 스트렙토마이신이 보충된 DMEM 으로 유지시켰다.

[0275] 웰 당 200,000 개 세포의 밀도로 트랜스펙션 1 일 전 293FT 또는 N2A 세포를 24-웰 플레이트 (Corning) 에 시딩하였다. 세포를 제조사의 권고 프로토콜에 따라 리포펙타민 2000 (Life Technologies) 을 사용하여 트랜스펙션하였다. 24-웰 플레이트의 각 웰에 대해 총 800ng 플라스미드를 사용하였다.

- [0276] 게놈 변형에 대한 Surveyor 검정 및 서열분석. 293FT 또는 N2A 세포를 상기 기재한 바와 같이 플라스미드 DNA 로 트랜스펙션하였다. 게놈 DNA 추출 전 트랜스펙션후 72 시간 동안 37°C 에서 세포를 인큐베이션하였다. 게놈 DNA 를 제조사의 프로토콜에 따라 QuickExtract DNA 추출 키트 (Epicentre) 를 사용하여 추출하였다. 간략하게, 세포를 QuickExtract 용액에 재현탁하고 65°C 에서 15 분 동안 및 98°C 에서 10 분 동안 인큐베이션하였다.
- [0277] 각 유전자에 대한 CRISPR 표적 부위를 둘러싸고 있는 게놈 부위를 PCR 증폭하고, 생성물을 제조사의 프로토콜에 따라 QiaQuick Spin Column (Qiagen) 을 사용하여 정제하였다. 총 400ng 의 정제된 PCR 생성물을 2 $\mu$ l 10X Taq 중합효소 PCR 완충액 (Enzymatics) 및 초순수와 총 부피 20 $\mu$ l 로 혼합하고, 재-어닐링 프로세스를 거치게 하여 이형이중체가 형성되게 하였다: 95°C 에서 10 분, 95°C 에서 85°C (-2°C/초로 램핑 (ramping)), 85°C 에서 25°C (-0.25°C/초), 및 25°C 에서 1 분 고정. 재어닐링 후, 생성물을 제조사의 권고 프로토콜에 따라 SURVEYOR 뉴클레아제 및 SURVEYOR 인헨서 S (Transgenomics) 로 처리하고, 4-20 Novex TBE 폴리-아크릴아미드 겔 (Life Technologies) 에서 분석하였다. 겔을 SYBR Gold DNA 염색제 (Life Technologies) 로 30 분 동안 염색하고 Gel Doc 겔 영상화 시스템 (Biorad) 으로 영상화하였다. 정량은 상대적 밴드 강도를 기반으로 하였다.
- [0278] 상동 재조합의 검출을 위한 제한 단편 길이 다형성 검정. HEK 293FT 및 N2A 세포를 플라스미드 DNA 로 트랜스펙션하고, 37°C 에서 72 시간 동안 인큐베이션한 후 상기 기재한 바와 같이 게놈 DNA 를 추출하였다. 표적 게놈 부위를 상동 재조합 (HR) 주형의 상동성 암 (arm) 외부의 프라이머를 사용하여 PCR 증폭하였다. PCR 생성물을 1% 아가로오스 겔 상에서 분리하고 MinElute GelExtraction 키트 (Qiagen) 를 사용하여 추출하였다. 정제된 생성물을 HindIII (Fermentas) 로 소화시키고 6% Novex TBE 폴리-아크릴아미드 겔 (Life Technologies) 에서 분석하였다.
- [0279] RNA 추출 및 정제. HEK 293FT 세포를 이전에 언급한 바와 같이 유지시키고 트랜스펙션하였다. 세포를 트립신화에 의해 수확한 후 인산 완충 식염수 (PBS) 로 세척하였다. 총 세포 RNA 를 제조사의 프로토콜에 따라 TRI 시약 (Sigma) 으로 추출하였다. 추출한 총 RNA 를 Naonodrop (Thermo Scientific) 을 사용하여 정량하고 동일 농도로 정규화하였다.
- [0280] 포유동물 세포에서의 crRNA 및 tracrRNA 발현의 노던 블롯 분석. RNA 를 동일 부피의 2X 로딩 완충액 (Ambion) 과 혼합하고, 5 분 동안 95°C 로 가열하고, 1 분 동안 얼음 상에서 냉각시킨 후, 8% 변성 폴리아크릴아미드 겔 (SequaGel, National Diagnostics) 상에, 30 분 이상 동안 겔을 사전-런닝시킨 후 로딩하였다. 샘플을 40 W 한도에서 1.5 시간 동안 전기영동하였다. 그 후, RNA 를 실온에서 1.5 시간 동안 반-건조 이동 장치 (Bio-rad) 내에서 300 mA 에서 Hybond N+ 멤브레인 (GE Healthcare) 에 옮겼다. RNA 를 Stratagene UV Crosslinker the Stratalinker (Stratagene) 상의 자동가교 버튼을 사용하여 멤브레인에 가교시켰다. 멤브레인을 ULTRAhyb-Oligo Hybridization 완충액 (Ambion) 중에서 30 분 동안, 42°C 에서 회전시키면서 사전-하이브리드화한 후, 프로브를 첨가하고 밤새 하이브리드화하였다. 프로브를 IDT 에서 주문하고 T4 폴리뉴클레오티드 키나아제 (New England Biolabs) 와 함께 [감마-32P] ATP (Perkin Elmer) 로 표지하였다. 멤브레인은 사전가온된 (42°C) 2xSSC, 0.5% SDS 로 1 분 동안 1 회 세척한 후 42°C 에서 2 회 30 분 세척하였다. 멤브레인은 1 시간 동안 또는 밤새 실온에서 인광체 스크린에 노출시킨 후 phosphorimager (Typhoon) 로 스캔하였다.
- [0281] 표 1. 포유동물 게놈 표적의 프로토스페이서 서열 및 변형 효율. 프로토스페이서 표적을 스트렙토코쿠스 피오게네스 유형 II CRISPR 및 스트렙토코쿠스 썬모필루스 CRISPR1 유전자좌 (인간 및 마우스 게놈에서 3 개의 상이한 유전자에 대하여 그의 필요한 PAM 을 가짐) 를 기반으로 하여 설계하였다. 세포를 Cas9 및 precrRNA/tracrRNA 또는 키메라 RNA 로 트랜스펙션하였다. 세포를 트랜스펙션 72 시간 후 분석하였다. 삽입-결실 % 를 나타낸 세포주로부터의 SURVEYOR 검정 결과를 기반으로 계산하고, 모든 스페이서 표적에 대해 N = 3 이고, 오류는 S.E.M. 이고, N.D. 는 SURVEYOR 검정을 사용하여 검출불가능이고; N.T. 는 이 연구에서 시험하지 않은 것이다.

[0282]

Gene#	표적 종류	유전자	프로토스페이스 ID	프로토스페이스 서열 (5'에서 3')	PAM	가닥	시원한 세포주	심인-결실 % (에비-crRNA + tracrRNA)	심인-결실 % (키메라 RNA)
S. 피오케비스 SF370 유형 II CHISPR	호모 사피엔스	EMM1	1	GGAAAGGCTGAGTCCGAGCAGAGAGAAA	GGG	+	293FT	29 ± 1.8	8.7 ± 0.62
		EMM3	2	CATTGGAGGTGACATCGATGTCCTCCCAT	TGG	-	293FT	2.1 ± 0.38	N.D.
		EMM4	3	GGACATCGATGTCCTCCATGACATGAGGG	TGG	+	293FT	14 ± 1.1	N.D.
		EMM5	4	CATGATGTCCTCCATGAGCTGGCTTCG	TGG	-	293FT	11 ± 1.7	N.D.
	EMM6	5	TTCGTGGCAATGAGCGTACGGTTGATGTA	TGG	-	293FT	4.3 ± 0.48	2.1 ± 0.31	
	EMM7	6	TGCTGGCAATGAGCGTACGGTTGATGTA	GGA	-	293FT	4.3 ± 0.88	0.41 ± 0.25	
	EMM8	7	TCCAGCTTCCTGCGTTTATGTTGCTTC	GGA	-	293FT	1.5 ± 0.13	N.D.	
	EMM9	8	GGAGGGAGGGGACAGATGGAGACTCAGG	AGG	-	293FT	7.8 ± 0.68	3.3 ± 1.2	
호모 사피엔스	PVAL3	9	AGGGGCTCGAGTTGGGTTTLAGGGAGAG	AGG	+	293FT	31 ± 2.0	8.5 ± 0.33	
	PVAL8	10	ATGCAAGGATGGGTGGCGAGGGGGCCGAT	TGG	+	293FT	N.D.	N.D.	
	PVAL9	11	GGTGGCAGAGGGGCGAGTTGGGTTTC	AGG	+	293FT	N.D.	N.D.	
마우스	7a	12	CAAGCAC TGAGTGCATTAGCTAATGCAAT	AGG	-	Neuro2A	27 ± 4.3	4.1 ± 2.2	
	7b	13	AATGCAATAGGTTACACCCAGGTGCTAG	GGG	-	Neuro2A	4.8 ± 1.2	N.D.	
	7c	14	ACACACATGGGAGAGCTCTGGGCTAGGAA	AGG	+	Neuro2A	11.3 ± 1.9	N.D.	
	EMM1	15	GGAGGGGTAGTATACAGAACACAGAGAA	GTAGGAT	-	293FT	14 ± 0.88	N.D.	
S. 세모펠루스 LMD9 CRISPR1 사피엔스	EMM1	16	AGCAATGATAGAGGAGTACAGAACTTCAACA	CTAGGAA	-	293FT	7.8 ± 0.77	N.D.	

[0283] 표 2. SURVEYOR 검증, RFLP 검증, 게놈 서열분석 및 노던 블롯에 사용한 프라이머 및 프로브에 대한 서열.

프라이머 명칭	검정	게놈 표적	프라이머 서열
Sp-EMX1-F	SURVEYOR 검정, 서열분석	<i>EMX1</i>	AAAACCACCCCTTCTCTCTGGC
Sp-EMX1-R	SURVEYOR 검정, 서열분석	<i>EMX1</i>	GGAGATTGGAGACACGGAGAG
Sp-PVALB-F	SURVEYOR 검정, 서열분석	<i>PVALB</i>	CTGGAAAGCCAATGCCTGAC
Sp-PVALB-R	SURVEYOR 검정, 서열분석	<i>PVALB</i>	GGCAGCAAACCTCTTGTCTCT
Sp-Th-F	SURVEYOR 검정, 서열분석	<i>Th</i>	GTGCTTTGCAGAGGCCTACC
Sp-Th-R	SURVEYOR 검정, 서열분석	<i>Th</i>	CCTGGAGCGCATGCAGTAGT
St-EMX1-F	SURVEYOR 검정, 서열분석	<i>EMX1</i>	ACCTTCTGTGTTCCACCATTTC
St-EMX1-R	SURVEYOR 검정, 서열분석	<i>EMX1</i>	TTGGGGAGTGCACAGACTTC
Sp-EMX1-RFLP-F	RFLP, 서열분석	<i>EMX1</i>	GGCTCCCTGGGTTCAAAGTA
Sp-EMX1-RFLP-R	RFLP, 서열분석	<i>EMX1</i>	AGAGGGGTCTGGATGTCGTAA
Pb_EMX1_sp1	노던 블롯 프로브	해당 없음	TAGCTCTAAAACCTCTTCTTCTGCTCGGAC
Pb_tracrRNA	노던 블롯 프로브	해당 없음	CTAGCCTTATTTTAACCTTGCTATGCTGTTT

[0284]

[0285] 보충적 서열

[0286] > U6-단길이 tracrRNA (스트렙토코쿠스 피오케네스 SF370)

```
GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCT
GTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAA
TACGTGACGTAGAAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTT
AAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTTCGATTTCTTGGCTTTAT
ATATCTTGTGGAAAGGACGAAACACCGGAACATTCAAAAACAGCATAGCAAGTTAA
AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT
```

[0287]

[0288] > U6-장길이 tracrRNA (스트렙토코쿠스 피오케네스 SF370)

```
GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCT
GTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAA
TACGTGACGTAGAAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTT
AAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTTCGATTTCTTGGCTTTAT
ATATCTTGTGGAAAGGACGAAACACCGGTAGTATTAAGTATTGTTTTATGGCTGATA
AATTTCTTTGAATTTCTCCTTGATTATTTGTTATAAAAAGTTATAAAAATAATCTTGTTG
GAACCATTCAAAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGA
AAAAGTGGCACCGAGTCGGTGCTTTTTTTT
```

[0289]

[0290] > U6-DR-BbsI 백본-DR (스트렙토코쿠스 피오케네스 SF370)

GAGGGCCTATTTCCCATGATTTCCTTCATATTTGCATATACGATACAAGGCT  
 GTTAGAGAGATAAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAA  
 TACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTT  
 AAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTGATTTCTTGGCTTTAT  
 ATATCTTGTGGAAAAGGACGAAACACCGGGTTTTAGAGCTATGCTGTTTTGAATGGTC  
 CAAAACGGGTCTTCGAGAAGACGTTTTAGAGCTATGCTGTTTTGAATGGTCCAAA  
 AC

[0291]

[0292] > U6-키메라 RNA-BbsI 백본 (스트렙토코쿠스 피오케네스 SF370)

GAGGGCCTATTTCCCATGATTTCCTTCATATTTGCATATACGATACAAGGCT  
 GTTAGAGAGATAAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAA  
 TACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTT  
 AAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTGATTTCTTGGCTTTAT  
 ATATCTTGTGGAAAAGGACGAAACACCGGGTCTTCGAGAAGACCTGTTTTAGAGCTA  
 GAAATAGCAAGTAAAATAAGGCTAGTCCG

[0293]

[0294] > 3xFLAG-NLS-SpCas9-NLS

ATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTA  
 CAAAGACGATGACGATAAGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCAC  
 GGAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTC  
 TGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCAGCAAGAAATCAAGG  
 TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTG  
 TTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAA  
 GATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAG  
 ATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGA  
 AGAGGATAAGAAGCAGGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGG  
 CCTACCACGAGAAGTACCCACCATCTACCACCTGAGAAAGAACTGGTGGACAGC  
 ACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCACATGATCAAGTTC  
 CGGGGCCACTTCTGATCGAGGGCGACCTGAACCCGACAACAGCGACGTGGACAA  
 GCTGTTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTTCGAGGAAAACCCATCAA  
 CGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGAC  
 GGCTGGAAAATCTGATCGCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCCGGC  
 AACCTGATTGCCCTGAGCCTGGGCCTGACCCCAACTTCAAGAGCAACTTCGACCTG  
 GCCGAGGATGCCAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAA  
 CCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCAAGAACCT  
 GTCCGACGCCATCTGCTGAGCGACATCTGAGAGTGAACACCGAGATCACCAAGG  
 CCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACAGGACCTGACC  
 CTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTC  
 GACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAG  
 AGTTCTACAAGTTCATCAAGCCATCTGGAAAAAGATGGACGGCACCGAGGAAGT  
 CTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGG  
 CAGCATCCCCACAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGG  
 AAGATTTTACCATTCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACC  
 TTCCGCATCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCTGG  
 ATGACCAGAAAGAGCGAGGAAACCATCACCCCTGGAACCTTCGAGGAAGTGGTGGAA  
 CAAGGGCGCTTCCGCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACC  
 TGCCCAACGAGAAGGTGCTGCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGT  
 ATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTC

[0295]

CTGAGCGGCGAGCAGAAAAAGGCCATCGTGACCTGCTGTTCAAGACCAACCGGAA  
 AGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACT  
 CCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCAC  
 GATCTGCTGAAAATTATCAAGGACAAGGACTTCTTGACAATGAGGAAAACGAGGA  
 CATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGA  
 GGAACGGCTGAAAACCTATGCCACCTGTTTCGACGACAAAGTGATGAAGCAGCTGA  
 AGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATC  
 CGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCTGAAAGTCCGACGGCTTCGC  
 CAACAGAAACTTATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACA  
 TCCAGAAAAGCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAAT  
 CTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGA  
 CGAGCTCGTGAAAGTGATGGGCCCGCACAAGCCCAGAACATCGTGATCGAAATGG  
 CCAGAGAGAACCAGACCACCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAA  
 GCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCC  
 GTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGG  
 GCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATG  
 TGGACCATATCGTGCCCTCAGAGCTTTCTGAAAGGACGACTCCATCGACAACAAGGTGC  
 TGACCAGAAGCGACAAGAACCAGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGT  
 CGTGAAGAAGATGAAGAATACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCC  
 AGAGAAAATTGACAATCTGACCAAGGCCGAGAGAGGGCGGCTGAGCGAACTGGAT  
 AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGT  
 GGCACAGATCCTGGACTCCCGATGAACACTAAGTACGACGAGAATGACAAGCTGA  
 TCCGGGAAGTGAAAGTGATCACCTGAAAGTCCAAGCTGGTGTCCGATTTCCGGAAG  
 GATTTCCAGTTTTACAAAGTGCGCGAGATCAACAATACTACCACCACGCCACGACGCC  
 TACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAG  
 CGAGTTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGA  
 GCGAGCAGGAAATCGGCAAGGCTACCGCAAGTACTTCTTACAGCAACATCATG  
 AACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCT  
 GATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTG  
 CCACCGTGCAGAAAGTGCTGAGCATGCCCAAGTGAATATCGTGAAAAAGACCGAG  
 GTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAA

[0296]

GCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCC  
 CCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAGGGCAAGTCCAAG  
 AAAGTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAGAAAGCAGCTT  
 CGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGG  
 ACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAG  
 AGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTC  
 CAAATATGTGAACTTCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCC  
 CGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAAGCACTACCTGGACG  
 AGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAAT  
 CTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGATAAGCCCATCAGAGAGCA  
 GGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTT  
 CAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGC  
 TGGACGCCACCTGATCCACCAGAGCATCACCGGCTGTACGAGACACGGATCGAC  
 CTGTCTCAGCTGGGAGGCGACAAGCGTCTGCTGCTACTAAGAAAGCTGGTCAAGCT  
 AAGAAAAAGAAA

[0297]

[0298] > SpRNase3-mCherry-NLS

ATGAAGCAGCTGGAGGAGTTACTTTCTACCTCTTTGACATCCAGTTTAAT  
 GACCTGACCCTGCTGGAAACCGCCTTCACTCACACCTCCTACGCGAATGAGCACC  
 CTAATGATGTGAGCCACAACGAGCGCCTGGAGTTTCTGGGGGATGCTGTCTTACAG  
 CTGATCATCTCTGAATATCTGTTTGCCAAATACCCTAAGAAAACCGAAGGGGACATG  
 TCAAAGCTGCGCTCCATGATAGTCAGGGAAGAGAGCCTGGCGGGCTTTAGTCGTTTT  
 TGCTCATTGACGCTTATATCAAGCTGGGAAAAGGCGAAGAGAAAGTCCGGCGGCAG  
 GAGGCGGATACAATTCTGGGCGATCTCTTTGAAGCGTTTCTGGGCGCACTTCTACT  
 GGACAAAAGGGATCGACGCAGTCCGCCCTTTCTGAAACAAGTGATGATCCCTCAGG  
 TCGAAAAGGGAAACTTCGAGAGAGTGAAGGACTATAAAACATGTTTGAGGAATTT  
 CTCCAGACCAAGGGAGATGTAGCAATAGATTATCAGGTAATAAGTGAGAAAGGACC  
 AGCTCACGCCAAACAATTGAAAGTTAGCATCGTTGTTAATGGCGCAGTGTGTCGAA  
 GGGCTTGGGTAAATCAAAAAACTGGCCGAGCAGGACGCTGCTAAAAACGCCCTCG  
 CTCAGCTCAGCGAGGTAGGATCCGTGAGCAAGGGCGAGGAGGATAACATGGCCATC  
 ATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGA  
 GTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCACCCAGACCGCCA

[0299] AGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTC  
 AGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCGACTACT  
 TGAAGCTGTCTTCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGAC  
 GGCGGCGTGGTGACCGTGACCCAGGACTCCTCCTGCAGGACGGCGAGTTCATCTA  
 CAAGGTGAAGCTGCGCGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGA  
 AGACCATGGGCTGGGAGGGCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTG  
 AAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTG  
 AGGTCAAGACCACCTACAAGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAAC  
 GTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACA  
 GTACGAACGCGCCGAGGGCCGCCACTCCACCGCGGCATGGACGAGCTGTACAAGA  
 AGCGTCTGTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAAAGAAA

[0300]

[0301] > 3xFLAG-NLS-SpCas9n-NLS (D10A 니카아제 돌연변이에 밀줄 그음)

ATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTA  
 CAAAGACGATGACGATAAGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCAC  
 GGAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTC  
 TGTGGGCTGGGCGGTGATCACCGACGAGTACAAGGTGCCAGCAAGAAATTC AAGG  
 TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTG  
 TTCGACAGCGGCAACAGCCGAGGCCACCCGGCTGAAGAGAACC GCCAGAAGAA  
 GATACACCAGACGGAAGAACC GGATCTGCTATCTGCAAGAGATCTTCAGCAACGAG  
 ATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCTGTTGGA  
 AGAGGATAAGAAGCAGGAGCGGCACCCATCTTCGGCAACATCGTGGACGAGGTGG  
 CCTACCACGAGAAGTACCCACCATCTACCACCTGAGAAAGAACTGGTGGACAGC  
 ACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCACATGATCAAGTTC  
 CGGGGCCACTTCTGATCGAGGGCGACCTGAACCCGACAACAGCGACGTGGACAA  
 GCTGTTTATCCAGCTGGTGCAGACCTACAACCAGCTGTTTCGAGGAAAAACCCATCAA  
 CGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGAC  
 GGCTGGAAAATCTGATCGCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCCGGC  
 AACCTGATTGCCCTGAGCCTGGGCTGACCCCAACTTCAAGAGCAACTTCGACCTG  
 GCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAA  
 CCTGCTGGCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCAAGAACCT  
 GTCCGACGCCATCTGCTGAGCGACATCTGAGAGTGAACACCGAGATCACCAAGG

[0302]

CCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACC  
 CTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTC  
 GACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAG  
 AGTTCTACAAGTTCATCAAGCCATCTTGAAAAGATGGACGGCACCGAGGAACTG  
 CTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGG  
 CAGCATCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGG  
 AAGATTTTTACCCATTCTGAAGGACAACCGGAAAAGATCGAGAAGATCCTGACC  
 TTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAAACAGCAGATTCGCCTGG  
 ATGACCAGAAAAGAGCGAGGAAACCATCACCCCTGGAACCTCGAGGAAGTGGTGGGA  
 CAAGGGCGCTTCCGCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAAGAACC  
 TGCCCAACGAGAAGGTGTGCCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGT  
 ATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTC  
 CTGAGCGGGCAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAA  
 AGTGACCGTGAAGCAGCTGAAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACT  
 CCGTGGAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCAC  
 GATCTGCTGAAAATTATCAAGGACAAGGACTTCTGGACAATGAGGAAAACGAGGA  
 CATTCTGGAAGATATCGTGTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGA  
 GGAACGGCTGAAAACCTATGCCACCTGTTTCGACGACAAAAGTGATGAAGCAGCTGA  
 AGCGGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATC  
 CGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCTGAAGTCCGACGGCTTCGC  
 CAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACA  
 TCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAAT  
 CTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGGA  
 CGAGCTCGTGAAGTGATGGGCCGACACAAGCCCGAGAACATCGTGATCGAAATGG  
 CCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAA  
 GCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCC  
 GTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGG  
 GCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATG  
 TGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGC  
 TGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGT  
 CGTGAAGAAGATGAAGAATACTGCGCGCAGCTGCTGAACGCCAAGCTGATTACCC

[0303]

AGAGAAAGTTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGAT  
 AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGT  
 GGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGA  
 TCCGGGAAGTGAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAG  
 GATTTCCAGTTTTACAAAGTGC GCGAGATCAACA ACTACCACCACGCCACGACGCC  
 TACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAG  
 CGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGA  
 GCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATG  
 AACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCT  
 GATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTG  
 CCACCGTGGGAAAGTGCTGAGCATGCCCAAGTGAATATCGTGAAAAAGACCGAG  
 GTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAA  
 GCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCC  
 CCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAG  
 AAAGTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAGAAGCAGCTT  
 CGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGG  
 ACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAG  
 AGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTC  
 CAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGTCCCC  
 CGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACG  
 AGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAAT  
 CTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCA  
 GGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCTGCCGCCTT  
 CAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGC  
 TGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGAC  
 CTGTCTCAGCTGGGAGGCGACAAGCGTCTGCTGCTACTAAGAAAGCTGGTCAAGCT  
 AAGAAAAAGAAA

[0304]

[0305] > hEMX1-HRTemplate-HindIII-NheI

GAATGCTGCCCTCAGACCCGCTTCCTCCCTGTCCTTGCTGTCCAAGGAGA  
 ATGAGGTCTCACTGGTGGATTTTCGACTACCCTGAGGAGCTGGCACCTGAGGGACA  
 AGGCCCCCCACCTGCCAGCTCCAGCCTCTGATGAGGGGTGGGAGAGAGCTACATG

[0306]

AGGTTGCTAAGAAAGCCTCCCCTGAAGGAGACCACACAGTGTGTGAGGTTGGAGTC  
 TCTAGCAGCGGGTTCTGTGCCCCCAGGGATAGTCTGGCTGTCCAGGCACTGCTCTTG  
 ATATAAACACCACCTCCTAGTTATGAAACCATGCCATTCTGCCTCTCTGTATGGAA  
 AAGAGCATGGGGCTGGCCCGTGGGGTGGTGTCCACTTTAGGCCCTGTGGGAGATCA  
 TGGGAACCCACGCAGTGGGTATAGGCTCTCTCATTTACTACTCACATCCACTCTGT  
 GAAGAAGCGATTATGATCTCTCCTCTAGAAACTCGTAGAGTCCCATGTCTGCCGGCT  
 TCCAGAGCCTGCACTCCTCCACCTTGGCTTGGCTTTGCTGGGGCTAGAGGAGCTAGG  
 ATGCACAGCAGCTCTGTGACCCTTTGTTTGAGAGGAACAGGAAAACCACCCTTCTCT  
 CTGGCCCACTGTGTCTCTCCTGCCCTGCCATCCCCTTCTGTGAATGTTAGACCCAT  
 GGGAGCAGCTGGTCAGAGGGGACCCCGCCTGGGGCCCCCTAACCTATGTAGCCTC  
 AGTCTTCCCATCAGGCTCTCAGCTCAGCCTGAGTGTGAGGGCCCCAGTGGCTGCTCT  
 GGGGGCCTCCTGAGTTTCTCATCTGTGCCCTCCCTCCCTGGCCAGGTGAAGGTGT  
 GGTTCAGAACCGGAGGACAAAGTACAAACGGCAGAAGCTGGAGGAGGAAGGGCC  
 TGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGGCGCATTGCCA  
 CGAAGCAGGCCAATGGGGAGGACATCGATGTCACCTCCAATGACaagttgctagcGGTGG  
 GCAACCACAAACCCACGAGGGCAGAGTGCTGCTTGCTGCTGGCCAGGCCCTGCGT  
 GGGCCCAAGCTGGACTCTGGCCACTCCCTGGCCAGGCTTTGGGGAGGCCTGGAGTC  
 ATGGCCCCACAGGGCTTGAAGCCCAGGGCCGCCATTGACAGAGGGACAAGCAATGG  
 GCTGGCTGAGGCCTGGGACCACTTGGCCTTCTCCTCGGAGAGCCTGCCTGCCTGGGC  
 GGGCCCGCCCGCCACCGCAGCCTCCAGCTGCTCTCCGTGTCTCCAATCTCCCTTTTG  
 TTTTGATGCATTTCTGTTTAATTTATTTTCCAGGCACCACTGTAGTTTGTAGTATCCC  
 AGTGTCCCCCTTCCCTATGGGAATAATAAAAAGTCTCTCTTAAATGACACGGGCATC  
 CAGCTCCAGCCCCAGAGCCTGGGGTGGTAGATTCCGGCTCTGAGGGCCAGTGGGGG  
 CTGGTAGAGCAAACGCGTTCAGGGCCTGGGAGCCTGGGGTGGGGTACTGGTGGAGG  
 GGGTCAAGGGTAATTCATTAACCTCTCTTTTGTGGGGGACCCTGGTCTCTACCTC  
 CAGCTCCACAGCAGGAGAAACAGGCTAGACATAGGGAAGGGCCATCCTGTATCTTG  
 AGGGAGGACAGGCCAGGTCTTTCTTAACGTATTGAGAGGTGGGAATCAGGCCAG  
 GTAGTTCAATGGGAGAGGGAGAGTGCTTCCCTCTGCCTAGAGACTCTGGTGGCTTCT  
 CCAGTTGAGGAGAAACCAGAGGAAAGGGGAGGATTGGGGTCTGGGGGAGGGAACA  
 CCATTCACAAAGGCTGACGGTTCAGTCCGAAGTCGTGGGCCACCAGGATGCTCA  
 CCTGTCTTGGAGAACCGCTGGGCAGGTTGAGACTGCAGAGACAGGGCTTAAGGCT

[0307]

GAGCCTGCAACCAGTCCCCAGTGAAGTCAAGGGCCTCCTCAGCCCAAGAAAGAGCAAC  
 GTGCCAGGGCCCGCTGAGCTCTTGTGTTCACCTG

[0308]

> NLS-StCsn1-NLS

[0309]

ATGAAAAGGCCGGCCGACGAAAAAGGCCGGCCAGGCAAAAAAGAAA  
 AAGTCCGACCTGGTACTTGGACTGGATATTGGTATCGTTCCGGTGGGAGTCGGAATC  
 CTCACAAGGTCACGGGGGAGATCATTACAAGAAGCTCGCGGATCTTCCCCGAGC  
 TCAGGCTGAGAACAACCTGGTGCGGAGAACG

[0310]

AATAGGCAGGGCAGGCGACTGGCGAGGAGGAAGAAACACAGGAGAGTC  
 CGATTGAACCGGCTGTTTCGAGGAGTCCGGTTTGATCACCGACTTTACGAAAATCTCG  
 ATTAACCTTAATCCCTATCAGCTTCGGGTGAAAGGCCTGACAGACGAACTTTCGAAT  
 GAGGAACTTTTCATCGCGCTGAAAAACATGGTCAAGCACAGAGGGATTTCTACCTC  
 GATGACGCCTCGGATGACGGAAATTCCTCAGTAGGAGATTATGCACAGATCGTGAA  
 AGAGAACTCAAAGCAACTGGAAACAAAGACACCGGGCAGATCCAACCTGAAAAGA  
 TACCAGACATACGGACAGCTCAGAGGAGATTTTACGGTGGAGAAGGACGGTAAAAA  
 GCACAGACTCATTAACGTATTTCCACGTCCGGCTACAGATCCGAAGCGCTCCGCAT  
 CCTTCAGACTCAACAGGAGTTCAACCCGCAAATTAAGTATGAGTTCATCAACCGCTA  
 TTTGGAAATCTTGACCGGAAAGCGCAAGTATTATCATGGGCCGGGTAATGAGAAAT  
 CCAGAACAGATTACGGCCGATACAGAACTTCGGGGGAAACCTTGATAACATCTTT  
 GGTATTTTGATTGGAAAGTGCACCTTTTACCCGGACGAGTTTCGAGCGGCCAAGGCG  
 TCATACACAGCACAAAGAGTTAATCTCTTGAATGATTTGAACAACCTTGACGGTCCCC  
 ACGGAGACAAAAGAAGCTCTCAAAGAGCAAAAAGAACCAATCATCAACTACGTCA  
 AGAACGAGAAGGCTATGGGGCCAGCGAAGCTGTTCAAGTATATCGCTAAACTTCTC  
 AGCTGTGATGTGGCGGACATCAAAGGGTACCGAATCGACAAGTCGGGAAAAGCGGA  
 AATTCACACGTTTGAAGCATATCGAAAGATGAAAACGTTGGAAACACTGGACATTG  
 AGCAGATGGACCGGAAACGCTCGACAAAAGTGGCATACGTGCTCACGTTGAATACT  
 GAACGAGAGGGAATCCAAGAGGCCCTTGAACATGAGTTCGCCGATGGATCGTTCAG  
 CCAGAAGCAGGTCGACGAACTTGTGCAATTCGCAAGGCGAATAGCTCCATCTTCG  
 GGAAGGGATGGCACAACCTTTTCGGTCAAACCTCATGATGGAGTTGATCCAGAACCTT  
 ATGAGACTTCGGAGGAGCAAATGACGATCTTGACGCGCTTGGGGAAACAGAAAACG  
 ACAAGCTCATCGAACAAAACCTAAGTACATTGATGAGAAATTGCTGACGGAAGAAAT  
 CTATAATCCGGTAGTAGCGAAATCGGTAAGACAAGCGATCAAAAATCGTGAACGCGG

[0311]

CGATCAAGGAATATGGTGACTTTGATAACATCGTAATTGAAATGGCTAGAGAGACG  
AACGAAGATGACGAGAAAAAGGCAATCCAGAAGATCCAGAAGGCCAACAAGGATG  
AAAAAGATGCAGCGATGCTTAAAGCGGCCAACCAATACAATGGAAAGGCGGAGCT  
GCCCCATTAGTGTTCACGGTCATAAACAGTTGGCGACCAAGATCCGACTCTGGCA  
TCAGCAGGGTGAGCGGTGTCTCTACACCGGAAAGACTATCTCCATCCATGACTTGAT  
TAACAATTCGAACCAGTTTGAAGTGGATCATAATTCTGCCCTGTCAATCACCTTTGA  
CGACTCGCTTGCGAACAAGGTGCTCGTGTACGCAACGGCAAATCAGGAGAAAAGGCC  
AGCGGACTCCGTATCAGGCGCTCGACTCAATGGACGATGCGTGGTCATTCGGGGAG  
CTGAAGGGCGTTCGTACGCGAGAGCAAGACTGAGCAACAAAAAGAAAGAGTATCT  
GCTGACAGAGGAGGACATCTCGAAATTCGATGTCAGGAAGAAGTTCATCGAGCGGA  
ATCTTGTGACACTCGCTACGCTCCAGAGTAGTACTGAACGCGCTCCAGGAACACT  
TTAGAGCGCACAAAATTGACACGAAGGTGTCAGTGGTGAGAGGGCAGTTCACATCC  
CAACTCCGCCGACATTGGGGCATCGAAAAGACGCGGGACACATATCACCATCATGC  
GGTGGACGCGCTGATTATTGCCGCTTCGTCCCAGTTGAATCTCTGGAAAAAGCAGAA  
GAACACGCTGGTGTGCTATTTCGGAGGATCAGCTTTTGGACATCGAAACCGGGGAGC  
TGATTTCCGACGATGAATACAAAGAATCGGTGTTTAAAGCACCATATCAGCATTTCG  
TGGACACGCTGAAGAGCAAAGAGTTTGAGGACAGCATCCTCTTTTCGTACCAAGTG  
GACTCGAAGTTTAAATCGCAAGATTTAGACGCCACAATCTACGCGACGAGGCAGGC  
GAAAGTGGGCAAAGATAAAGCAGATGAAACCTACGTCCTTGGTAAAAATCAAGGACA  
TCTACACTCAGGACGGGTACGATGCGTTCATGAAAATCTACAAGAAGGATAAGTCG  
AAGTTTCTCATGTACCGCCACGATCCACAGACTTTTCGAAAAAGTCATTGAGCCTATT  
TTGGAGAACTACCTAAACAAGCAAATCAACGAGAAAGGGAAAGAAGTCCCGTGCAA  
CCCCTTTCTGAAGTACAAGGAAGAGCACGGTTATATCCGCAAATACTCGAAGAAAG  
GAAATGGGCCTGAGATTAAGTCGCTTAAGTATTACGACTCAAAGTTGGGTAACCAC  
ATCGACATTACCCGAAAGACTCCAACAACAAAGTCGTGTTGCAGTCCGTCTCGCCC  
TGGCGAGCAGATGTGATTTTAATAAGACGACCGGCAAATATGAGATCCTTGGACTC  
AAATACGCAGACCTTCAATTCGAAAAGGGGACGGGCACTTATAAGATTTACAAGA  
GAAGTACAACGACATCAAGAAAAAGGAAGGGTTCGATTCAGATTCGGAGTTCAAAT  
TCACCTCTACAAAAACGACCTCCTGCTTGTGAAGGACACAGAAACGAAGGAGCAG  
CAGCTCTTTCGGTTCCTCTCACGCACGATGCCCAAACAAAAACATTACGTCGAACTT  
AAACCTTACGATAAGCAAAGTTTGAAGGGGGAGAGGCACTGATCAAAGTATTGGG

[0312]

TAACGTAGCCAATAGCGGACAGTGTAAGAAAGGGCTGGGAAAGTCCAATATCTCGA  
TCTATAAAGTACGAACAGATGTATTGGGAAACCAGCATATCATCAAAAATGAGGGG  
GATAAACCCAAACTCGATTTCAAGCGTCTGCTGCTACTAAGAAAGCTGGTCAAGCT  
AAGAAAAAGAAATAA

[0313]

[0314] > U6-St\_tracrRNA(7-97)

GAGGGCCTATTTCCATGATTCCTTCATATTTGCATATACGATACAAGGCT  
GTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAA  
TACGTGACGTAGAAAGTAATAATTTCTGGGTAGTTTGCAGTTTAAAATTATGTTTT  
AAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTGATTTCTTGGCTTTAT  
ATATCTTGTGGAAAAGGACGAAACACCGTTACTTAAATCTTGCAGAAGCTACAAAAGA  
TAAGGCTTCATGCCGAAATCAACACCTGTCATTTTATGGCAGGGTGTTCGTTATT  
TAA

[0315]

[0316] > EMX1\_TALEN\_좌측

ATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTA  
 CAAAGACGATGACGATAAAGTGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCAC  
 GGAGTCCCAGCAGCCGTAGATTTGAGAACTTTGGGATATTCACAGCAGCAGCAGGA  
 AAAGATCAAGCCCAAAGTGAAGTTCGACAGTCGCGCAGCATCACGAAGCGCTGGTGG  
 GTCATGGGTTTACACATGCCACATCGTAGCCTTGTCGCAGCACCCCTGCAGCCCTTG  
 GCACGGTCGCCGTCAAGTACCAGGACATGATTGCGGGCTTGCCGGAAGCCACACAT  
 GAGGCGATCGTCGGTGTGGGGAAACAGTGGAGCGGAGCCCGAGCGCTTGAGGCCCT  
 GTTGACGGTCGCGGGAGAGCTGAGAGGGCCTCCCTTCAGCTGGACACGGGCCAGT  
 TGCTGAAGATCGCGAAGCGGGGAGGAGTACGGCGGTGAGGCGGTGCACGCGTGG  
 CGCAATGCGCTACGGGAGCACCCCTCAACCTGACCCAGAGCAGGTCGTGGCAAT  
 TGCGAGCAACCACGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTG  
 TGCTGTGCCAAGCGCACGGACTTACGCCAGAGCAGGTCGTGGCAATTGCGAGCAAC  
 CACGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCA  
 AGCGCACGGACTAACCCAGAGCAGGTCGTGGCAATTGCGAGCAACATCGGGGGAA  
 AGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGG  
 TTGACCCAGAGCAGGTCGTGGCAATTGCGAGCAACCACGGGGAAAGCAGGCACT  
 CGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGCCTGACCCAG  
 AGCAGGTCGTGGCAATTGCGAGCAACCACGGGGAAAGCAGGCACTCGAAACCGTC

[0317] CAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGACTGACACCAGAGCAGGTCGT  
 GGCAATTGCGAGCAACATCGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGC  
 TGCCTGTGCTGTGCCAAGCGCACGGACTTACACCCGAACAAGTCGTGGCAATTGCG  
 AGCAACCACGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCT  
 GTGCCAAGCGCACGGACTTACGCCAGAGCAGGTCGTGGCAATTGCGAGCAACCACG  
 GGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCG  
 CACGGACTAACCCAGAGCAGGTCGTGGCAATTGCGAGCAACATCGGGGGAAAGCA  
 GGCCTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGGTTGA  
 CCCCAGAGCAGGTCGTGGCAATTGCGAGCAACATCGGGGGAAAGCAGGCACTCGAA  
 ACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGCCTGACCCAGAGCA  
 GGTTCGTGGCAATTGCGAGCAACCACGGGGAAAGCAGGCACTCGAAACCGTCCAGA  
 GGTTGCTGCCTGTGCTGTGCCAAGCGCACGGACTGACACCAGAGCAGGTCGTGGCA  
 ATTGCGAGCAACCACGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCC  
 TGTGCTGTGCCAAGCGCACGGACTCACGCCTGAGCAGGTAGTGGCTATTGCATCCAA  
 CAACGGGGGAGACCCGCACTGGAGTCAATCGTGGCCAGCTTTCGAGGCCGGACC  
 CCGCGCTGGCCGCACTCACTAATGATCATCTTGTAGCGCTGGCCTGCCTCGGCGGAC  
 GACCCGCCTTGATGCGGTGAAGAAGGGCTCCCGCACGCGCCTGCATTGATTAAG  
 CGGACCAACAGAAGGATCCCGAGAGGACATCACATCGAGTGGCAGGTTCCCAACT  
 CGTGAAGAGTGAACCTGAGGAGAAAAAGTCGGAGCTGCGGCACAAATTGAAATACG  
 TACCGCATGAATACATCGAACTTATCGAAATTGCTAGGAACTCGACTCAAGACAGA  
 ATCCTTGAGATGAAGGTAATGGAGTTCCTTATGAAGGTTTATGGATACCGAGGGAAG  
 CATCTCGGTGGATCACGAAAACCCGACGGAGCAATCTATACGGTGGGAGCCCGAT  
 TGATTACGGAGTGATCGTCGACACGAAAGCCTACAGCGGTGGGTACAATCTTCCAT  
 CGGGCAGGCAAGTGAATGCAACGTTATGTCGAAGAAAAATCAGACCAGGAACAAA  
 CACATCAATCCAAATGAGTGGTGGAAAGTGTATCCTTCATCAGTGACCGAGTTTAA  
 TTTTGTGTTGCTCTGGCATTTCAAAGGCAACTATAAGGCCAGCTCACACGGTTG  
 AATCACATTACGAACTGCAATGGTGGGTTTTGTCCGTAGAGGAACTGCTCATTGGT  
 GGAGAAATGATCAAAGCGGAACTCTGACACTGGAAGAAGTCAGACGCAAGTTTAA  
 CAATGGCGAGATCAATTTCCGCTCA

[0318]

[0319] >EMX1\_TALEN\_우측

ATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTA  
 CAAAGACGATGACGATAAGATGGCCCCAAAGAAGAAGCGGAAAGTTCGGTATCCAC  
 GGAGTCCCAGCAGCCGTAGATTTGAGAACTTTGGGATATTCACAGCAGCAGCAGGA  
 AAAGATCAAGCCCCAAAGTGAGGTCGACAGTCGCGCAGCATCACGAAGCGCTGGTGG  
 GTCATGGGTTTACACATGCCACATCGTAGCCTTGTCGCAGCACCTGCAGCCCTG  
 GCACGGTCGCCGTCAAGTACCAGGACATGATTGCGGGCTTGCCGGAAGCCACACAT  
 GAGGCGATCGTCGGTGTGGGGAAACAGTGGAGCGGAGCCCCGAGCGCTTGAGGCCCT  
 GTTGACGGTCGCGGGAGAGCTGAGAGGGCCTCCCCTTACGCTGGACACGGGCCAGT  
 TGCTGAAGATCGCGAAGCGGGGAGGAGTCACGGCGGTGCGAGGCGGTGCACGCGTGG  
 CGCAATGCGCTCACGGGAGCACCCCTCAACCTGACCCCAGAGCAGGTCGTGGCAAT  
 TGGCAGCAACCACGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTG  
 TGCTGTGCCAAGCGCACGGACTTACGCCAGAGCAGGTCGTGGCAATTGCGAGCAAC  
 CACGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCA  
 AGCGCACGGACTAACCCAGAGCAGGTCGTGGCAATTGCGAGCAACCACGGGGGA  
 AAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGG  
 GTTGACCCCAGAGCAGGTCGTGGCAATTGCGAGCAACATCGGGGGAAAGCAGGCAC  
 TCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGCCTGACCCAG  
 AGCAGGTCGTGGCAATTGCGAGCAACCACGGGGAAAGCAGGCACTCGAAACCGTC  
 CAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGACTGACACCAGAGCAGGTCGT  
 GGCAATTGCGAGCCATGACGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGC  
 TGCTGTGCTGTGCCAAGCGCACGGACTTACACCCGAAACAAGTCGTGGCAATTGCG  
 AGCCATGACGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCT  
 GTGCCAAGCGCACGGACTTACGCCAGAGCAGGTCGTGGCAATTGCGAGCCATGACG  
 GGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCG  
 CACGGACTAACCCAGAGCAGGTCGTGGCAATTGCGAGCAACGGAGGGGGAAAGC  
 AGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGGTTG  
 ACCCCAGAGCAGGTCGTGGCAATTGCGAGCAACGGAGGGGGAAAGCAGGCACTCG  
 AAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGCCTGACCCAGAG  
 CAGGTCGTGGCAATTGCGAGCCATGACGGGGAAAGCAGGCACTCGAAACCGTCCA  
 GAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGACTGACACCAGAGCAGGTCGTGG  
 CAATTGCGAGCAACGGAGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTG

[0320] CCTGTGCTGTGCCAAGCGCACGGACTCACGCCTGAGCAGGTAGTGGCTATTGCATCC  
 AACGGAGGGGGCAGACCCGCACTGGAGTCAATCGTGGCCAGCTTTCGAGGCCGGA  
 CCCCAGCGCTGGCCGCACTCACTAATGATCATCTTGTAGCGCTGGCCTGCCTCGGCGG  
 ACGACCCGCTTGGATGCGGTGAAGAAGGGGCTCCCGCACGCGCCTGCATTGATTA  
 AGCGGACCAACAGAAGGATTCCCGAGAGGAATCACATCGAGTGGCAGGTTCCCAAC  
 TCGTGAAGAGTGAAGTGAAGGAGAAAAAGTCGGAGCTGCGGCACAAATTGAAATAC  
 GTACCGCATGAATACATCGAACTTATCGAAATTGCTAGGAACTCGACTCAAGACAG  
 AATCCTTGAGATGAAGTAATGGAGTTCTTTATGAAGGTTTATGGATACCGAGGGAA  
 GCATCTCGGTGGATCACGAAAACCCGACGGAGCAATCTATACGGTGGGGAGCCCGA  
 TTGATTACGGAGTGATCGTCGACACGAAAGCCTACAGCGGTGGGTACAATCTTCCCA  
 TCGGGCAGGCAGATGAGATGCAACGTTATGTGCAAGAAAATCAGACCAGGAACAAA  
 CACATCAATCCAAATGAGTGGTGGAAAGTGTATCCTTCATCAGTGACCGAGTTTAAG  
 TTTTGTGTTGCTCTGGGCATTTCAAAGGCAACTATAAGGCCAGCTCACACGGTTG  
 AATCACATTACGAACTGCAATGGTGCGGTTTTGTCCGTAGAGGAACTGCTCATTGGT  
 GGAGAAATGATCAAAGCGGGAACTCTGACACTGGAAGAAGTCAGACGCAAGTTTAA  
 CAATGGCGAGATCAATTTCCGCTCA

[0321]

- [0322] 실시예 9
- [0323] AAV 구축물의 클로닝 (구축)
- [0324] **AAV-프로모터-TALE-효과기 백본의 구축.** AAV-프로모터-TALE-효과기의 구축을 위해 백본을 표준 서브클로닝 방법에 의해 클로닝하였다. 특히, 백터는 항생제 저항성 유전자, 예컨대 암피실린 저항성 및 프로모터-TALE-효과기 삽입물에 측면 위치한 2 개 AAV 역전 말단 반복물 (itr's) (서열, 하기 참조) 을 함유하였다. 프로모터 (hSyn), 효과기 도메인 (이 실시예에서 VP64, SID4X 또는 CIB1)/2 개 유형 IIS 제한 부위 (이 경우 BsaI) 를 갖는 스페이서를 함유하는 TALE 유전자의 N- 및 C-말단 부분을 상기 백터 내로 서브클로닝하였다. 서브클로닝을 수행하기 위해, 각각의 DNA 성분을 중합효소-연쇄 반응을 사용하여 증폭한 후 특정 제한 효소로 소화시켜 매치되는 DNA 점성 말단을 생성시켰다. 백터를 DNA 제한 효소로 유사하게 소화시켰다. 모든 DNA 단편을 이후 매치되는 말단에서 어닐링시키고 리가아제 효소를 사용하여 함께 융합하였다.
- [0325] **개별적 TALE 의 AAV-프로모터-TALE-효과기 백본 내로의 어셈블리.**
- [0326] 상이한 TALE 단량체 서열을 상기 기재한 AAV-프로모터-TALE-효과기 백본 내로 통합하기 위한 계획은 유형 IIS 제한 효소로의 개별 단량체의 제한 및 그의 고유한 오버행 (overhang) 의 라이게이션으로 12 내지 16 개 단량체의 어셈블리를 형성시켜 최종 TALE 을 형성시키고 이를 N- 및 C-말단 사이의 스페이서에 존재하는 유형 IIS 부위를 사용하여 AAV-프로모터-TALE-효과기 백본 내로 이를 라이게이션하는 것을 기반으로 하였다 (골든 게이트 어셈블리로 지칭함). TALE 단량체 어셈블리의 이러한 방법은 이전에 기재된 바 있다 (NE Sanjana, L Cong, Y Zhou, M M Cunniff, G Feng & F Zhang A transcription activator-like effector toolbox for genome engineering Nature Protocols 7, 171-192 (2012) doi: 10.1038/nprot.2011.431).
- [0327] 상기 요약한 일반적인 클로닝 계획을 사용하여, 상이한 프로모터, 효과기 도메인 및 TALE 단량체를 함유하는 AAV 백터를 용이하게 구축할 수 있다.
- [0328] 뉴클레오티드 서열:
- [0329] **좌측 AAV ITR**  

```
cctgcaggcagctgcgcgctcgtcgtcactgagccgccccgggcaaaagccggcgctcggcgacaccttggctgcc
cggcctcagtgagcagcgagcgcgcagagagggagtgccaactccatcactaggggttct
```
- [0330]
- [0331] **우측 AAV ITR**  

```
Aggaaccctagtgatggagtgccactccctctctcgcgctcgtcgtcactgagccggcgaccaaaggtcgc
ccgacgccgggcttggccggcgccctcagtgagcgcgagcgcgcagctgcctgcagg
```
- [0332]
- [0333] **hSyn 프로모터**  

```
gtgtctagactgcagaggccctcgtatgagtgcaagtgggttttaggaccaggatgagcgggggtgggggtcctac
ctgacgaccgaccccgaccactggacaagcaccacccccattcccaattgcgcatcccctatcagagagggggagggaaca
ggatgcggcgagggcgtcgcactccagcttcagcaccgcggacagtgctctcggcccgctggcggcgcgccaccgccgc
tcagcactgaaggcgcgctgactcactcgggtccccgcaactccccctccggccacttgctcgtccgcggccggccggc
ccagccggaccgcaccacgcgagggcgcgagatagggggcagggcgccacctctcgtcgtcggcgccggcgactcagcgtgc
ctcagctcgggtggcagcgggagtgctgtcgtcctgagagcagtcgagaa
```
- [0334]
- [0335] **TALE N-말단 (+136 AA 절두)**  

```
GTAGATTTGAGAACTTTGGGATATTCACAGCAGCAGCAGGAAAAGATCAA
GCCCAAAGTGAGGTCGACAGTCGCGCAGCATCACGAAGCGTGGTGGGTCATGGGT
TTACACATGCCACATCGTAGCCTTGTCGAGCACCCTGCAGCCCTTGGCACGGTGC
CCGTCAAGTACCAGGACATGATTGCGGCGTTGCCGGAAGCCACACATGAGGCGATC
GTCGGTGTGGGAAACAGTGGAGCGGAGCCGAGCGCTTAGGCCCTGTTGACGGT
CGCGGGAGAGCTGAGAGGGCCTCCCTTCAGCTGGACACGGGCCAGTTGCTGAAGA
TCGCGAAGCGGGGAGGATCACGGCGGTGAGGCGGTGCACGCGTGGCGCAATGC
GCTCACGGGAGCACCCCTCAAC
```
- [0336]
- [0337]

[0338] TALE C-말단 (+63 AA 절두)  
 CGGACCCCGCGCTGGCCGCACTCACTAATGATCATCTTGTAGCGCTGGCC  
 TGCCTCGGCGGACGACCCGCCTTGGATGCGGTGAAGAAGGGCTCCCGCACGCGCC  
 TGCATTGATTAAGCGACCAACAGAAGGATTCCCGAGAGGACATCACATCGAGTGG

[0339] CA

[0340] 암피실린 저항성 유전자

atgagtattcaacattccgtgctgcccttattccctttttgcggcattttgccttctgttttgcaccagaacgctggga  
 aagtaaaagatgctgaagatcagttgggtgacagtggttacatcgaactggatctcaacagcggtgaagatccttgagagtttccgccc  
 gaagaacgtttccaatgatgagcacttttaaagtctgctatgtggcgcggtattatcccgtattgacccgggcaagagcaactcggctgcc  
 gcatacactattctcagaatgacttggtgagtactaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgagtg  
 ctgccataacctgatgataaactcggccaacttactctgacaacgatcggaggaccgaaggagtaaccgctttttgcacaacatg  
 ggggatcatgtaactgccttgatcgttgggaaccggagctgaatgaagccataccaacacgagcgtgacaccacgatgcctgtagca  
 atggcaacaacgttgcaaacattactgccaacttactcttagcttcccgcaacaattaagactggatggagcgggataaagtt  
 gcaggaccactctgctcggccctccggctggctgtttattgctgataaatctggagccgggtgagcgtgggtctcgggtatcattgca  
 gcactggggccagatgtaagccctccgctatcgtatgtatctacacgacggggagtcaggcaactatgatgaacgaaatagacagatc  
 gctgagataggtgcctcactgattaagcattgg

[0341]

[0342] 실시예 10

[0343] 내인성 포유동물 전사의 광학적 제어

[0344] 내인성 포유동물 계놈의 전사를 직접적으로 조정하는 능력은 정상 유전자 기능 및 질환 메커니즘을 해명하기 위  
 해 중요하다. 여기서, 출원인은 광-유도 전사 효과기 (LITE), 커스터마이징 가능한 TALE DNA-결합 도메인을 아라  
 비도프시스 탈리아나로부터의 광-민감성 크립토크롬 2 단백질 및 이의 상호작용 파트너 CIB1 과 통합시키는 2-  
 하이브리드 시스템의 개발을 기재하고 있다. LITE 는 곧 활성화될 수 있어, 내인성 포유동물 유전자 발현  
 뿐 아니라 표적화 후생적 크로마틴 변형의 가역성 양방향 조절을 매개한다. 출원인은 이러한 시스템을 1 차  
 마우스 뉴런 뿐 아니라 각성 상태의, 거동 중인 마우스의 뇌 내에 생체내 적용하였다. LITE 시스템은 내인  
 성 세포 프로세스의 광유전학적 제어의 신규한 방식을 확립하며, 유전적 및 후생적 조절의 일상적 역할의 직접  
 적인 시험을 가능하게 한다.

[0345] 유전자 발현의 동적 성질은 생존계에서의 세포 프로그래밍, 항상성유지 및 환경적 적응을 가능하게 한다.  
 세포 및 유기체 기능에 대한 유전자의 기여도를 분석하는 것은 따라서 유전자 발현의 공간적 및 시간적 제어 조  
 정을 가능하게 하는 접근 방식을 필요로 한다. 미생물 및 식물-유래 광-민감성 단백질은 광유전학적 활성화제  
 로서 조작되어 높은 시공간적 분해능을 제공하는 광의 사용이 많은 세포 기능을 제어할 수 있게 한다  
 (Deisseroth, K. Optogenetics. Nature methods 8, 26-29, doi: 10.1038/nmeth.f.324 (2011); Zhang, F. et  
 al. The microbial opsin family of optogenetic tools. Cell 147, 1446-1457, doi: 10.1016/j  
 .cell.2011.12.004 (2011); Levskaya, A., Weiner, O. D., Lim, W. A. & Voigt, C. A. Spatiotemporal  
 control of cell signalling using a light-switchable protein interaction. Nature 461, 997-1001, doi:  
 10.1038/nature08446 (2009); Yazawa, M., Sadaghiani, A. M., Hsueh, B. & Dolmetsch, R. E. Induction of  
 protein-protein interactions in live cells using light. Nature biotechnology 27, 941-945, doi:  
 10.1038/nbt. 1569 (2009); Strickland, D. et al. TULIPs: tunable, light-controlled interacting protein  
 tags for cell biology. Nature methods 9, 379-384, doi: 10.1038/nmeth.1904 (2012); Kennedy, M. J. et  
 al. Rapid blue-light-mediated induction of protein interactions in living cells. Nature methods 7,  
 973-975, doi: 10.1038/nmeth. 1524 (2010); Shimizu-Sato, S., Huq, E., Tepperman, J. M. & Quail, P. H. A  
 light-switchable gene promoter system. Nature biotechnology 20, 1041-1044, doi: 10.1038/nbt734 (2002);  
 Ye, H., Daoud-El Baba, M., Peng, R. W. & Fussenegger, M. A synthetic optogenetic transcription device  
 enhances blood-glucose homeostasis in mice. Science 332, 1565-1568, doi:10.1126/science.1203535  
 (2011); Polstein, L. R. & Gersbach, C. A. Light-inducible spatiotemporal control of gene activation by  
 customizable zinc finger transcription factors. Journal of the American Chemical Society 134, 16480-  
 16483, doi: 10.1021/ja3065667 (2012); Bugaj, L. J., Choksi, A. T., Mesuda, C. K., Kane, R. S. &

Schaffer, D. V. Optogenetic protein clustering and signaling activation in mammalian cells. *Nature methods* (2013) and Zhang, F. et al. Multimodal fast optical interrogation of neural circuitry. *Nature* 446, 633-639, doi: 10.1038/nature05744 (2007)). 그러나, 광을 사용하여 내인성 전사 조절을 직접적으로 조정하기 위한 다목적이고 강건한 기술은 달성하기 어려운 채로 남아 있다.

[0346] 여기서, 출원인은 광-유도 전사 효과기 (LITE), 포유동물 세포에서의 내인성 유전적 및 후생적 프로세스의 시공간적으로 정확한 제어를 가능하게 하는 모듈식 광유전학적 시스템의 개발을 보고하고 있다. LITE 는 잔토모나스 종으로부터의 전사 활성제-유사 효과기 (TALE) 의 프로그래밍가능한 DNA-결합 도메인 (Boch, J. et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509-1512, doi: 10.1126/science.1178811 (2009) and Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501, doi: 10.1126/science.1178817 (2009)) 을 아라비도프시스 탈리아나로부터의 광-유도 이형이량체 단백질 크립토크롬 2 (CRY2) 및 CIB1 (Kennedy, M. J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. *Nature methods* 7, 973-975, doi: 10.1038/nmeth.1524 (2010) and Liu, H. et al. Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in Arabidopsis. *Science* 322, 1535-1539, doi: 10.1126/science.1163927 (2008)) 과 조합한다. 이들은 이중 유전자 요소의 도입을 필요로 하지 않고, 외인성 화학적 보조 인자에 의존적이지 않으며, 빠르고 가역적인 이량체화 반응속도를 나타낸다 (Levskaia, A., Weiner, O. D., Lim, W. A. & Voigt, C. A. Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* 461, 997-1001, doi: 10.1038/nature08446 (2009); Yazawa, M., Sadaghiani, A. M., Hsueh, B. & Dolmetsch, R. E. Induction of protein-protein interactions in live cells using light. *Nature biotechnology* 27, 941-945, doi: 10.1038/nbt.1569 (2009), Kennedy, M. J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. *Nature methods* 7, 973-975, doi: 10.1038/nmeth.1524 (2010); Shimizu-Sato, S., Huq, E., Tepperman, J. M. & Quail, P. H. A light-switchable gene promoter system. *Nature biotechnology* 20, 1041-1044, doi: 10.1038/nbt734 (2002) and Liu, H. et al. Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in Arabidopsis. *Science* 322, 1535-1539, doi: 10.1126/science.1163927 (2008)). 다른 광유전학적 도구와 같이, LITE 는 바이러스 벡터 내로 패키징될 수 있으며 특정 세포 집단을 탐색하도록 유전적으로 표적화될 수 있다. 출원인은 1 차 뉴런 뿐 아니라 마우스 뇌 (생체내) 에서의 이러한 시스템의 적용을 입증한다.

[0347] LITE 시스템은 2 개의 독립적 성분을 함유한다 (도 36A): 제 1 성분은 계놈 앵커 (anchor) 이며 광-민감성 CRY2 단백질에 융합된 커스텀화 TALE DNA-결합 도메인 (TALE-CRY2) 으로 이루어진다. 제 2 성분은 원하는 전사 효과기 도메인에 융합된 CIB1 으로 이루어진다 (CIB1-효과기). 효율적 핵 표적화를 확실히 하기 위해, 출원인은 핵 국부화 신호 (NLS) 를 양 모듈 모두에 부착시켰다. 광의 부재 하 (불활성 상태), TALE-CRY2 는 표적 유전자의 프로모터 부위에 결합하는 한편 CIB1-효과기는 핵 구획 내에 자유로이 남아있다. 청색광으로의 조명 (파장 ~450 nm) 은 CRY2 에서 형태적 변화를 촉발시키고 이후 CIB1-효과기 (도 36A 에서 나타낸 VP64) 를 표적 유전자좌에 영입하여 전사 조절을 매개한다. 이러한 모듈식 설계는 각각의 LITE 성분이 독립적일 수 있게 한다. 예를 들어, 동일한 계놈 앵커는 활성화 또는 억제 효과기와 조합되어 (Beerli, R. R., Segal, D. J., Dreier, B. & Barbas, C. F., 3rd. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14628-14633 (1998) and Cong, L., Zhou, R., Kuo, Y.-c, Cunniff, M. & Zhang, F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat Commun* 3, 968) 동일한 내인성 계놈 유전자좌에 걸쳐 양성적 및 음성적 전사 제어를 가할 수 있다.

[0348] 가장 효율적인 LITE 구조를 확인하기 위해서, 출원인은 TALE 및 전사 활성제 VP64 (Beerli, R. R., Segal, D. J., Dreier, B. & Barbas, C. F., 3rd. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14628-14633 (1998); Zhang, F. et al. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* 29, 149-153, doi: 10.1038/nbt.1775 (2011); Miller, J. C. et al. A TALE nuclease architecture for efficient genome editing. *Nature biotechnology* 29, 143-148, doi: 10.1038/nbt.1755 (2011) and Hsu, P. D. & Zhang, F. Dissecting neural function using targeted genome engineering technologies. *ACS chemical neuroscience* 3, 603-610, doi: 10.1021/cn300089k

(2012).) 를 각각 CRY2 및 CIB1 의 상이한 절두 (Kennedy, M. J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. Nature methods 7, 973-975, doi: 10.1038/nmeth.1524 (2010)) 에 융합시키고, 신경 계통-명시 전사 인자 뉴로제닌 2 (Neurog2) 의 청색광 조명 유도된 전사 변화를 측정하여 각 설계의 효능을 평가하였다 (도 36B). 출원인은 전체 길이 CRY2 뿐 아니라 포토리아제 상동성 부위 단독으로 이루어지는 절두 (CRY2PHR, 아미노산 1-498) 를 평가하였다 (Kennedy, M. J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. Nature methods 7, 973-975, doi: 10.1038/nmeth.1524 (2010)). CIB1 에 대해서, 출원인은 전체 길이 단백질 뿐 아니라 N-말단 도메인-단독 단편을 시험하였다 (CIBN, 아미노산 1-170) (Kennedy, M. J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. Nature methods 7, 973-975, doi:10.1038/nmeth.1524 (2010)). 4 개 초기 LITE 페어링 중 3 개는 Neuro 2a 세포에서 유의한 광-유도 Neurog2 mRNA 상향조절을 생성시켰다 ( $p < 0.001$ , 도 36B). 이들 중에서, TALE-CRY2PHR::CIB1-VP64 는 GFP-단독 대조군 또는 미자극 LITE 샘플로 정규화되는 경우 가장 높은 완전 광-매개 mRNA 증가를 산출하였으며 (도 36B), 따라서 후속 실험에 적용되었다.

[0349] 효율적 LITE 구조를 확립하여, 출원인은 파장 (도 40), 사용률 (도 41) 및 광 강도 (도 42 및 실시예 11) 를 포함하는 광 자극 매개변수를 체계적으로 최적화하였다 (Banerjee, R. et al. The signaling state of Arabidopsis cryptochrome 2 contains flavin semiquinone. The Journal of biological chemistry 282, 14916-14922, doi: 10.1074/jbc.M700616200 (2007)). 출원인은 또한 활성화 도메인 VP16 및 p65 를 VP64 에 추가로 비교하여 LITE CIB1-효과기 성분의 모듈성을 시험하였다. 모든 3 개 도메인은 유의한 광-의존적 Neurog2 mRNA 상향조절을 생성시켰다 ( $p < 0.001$ , 도 43). 출원인은 광-자극의 부재 하 그의 낮은 기저 활성으로 인해, 후속 실험을 위해 VP64 를 선택하였다.

[0350] 발현 속도가 조절 요소, mRNA 처리 및 전사체 안정성을 포함하는 많은 인자에 의존하는 바, 내인성 유전자 발현의 조작은 다양한 도전과제를 제시한다 (Moore, M. J. & Proudfoot, N. J. Pre-mRNA processing reaches back to transcription and ahead to translation. Cell 136, 688-700, doi: 10.1016/j.cell.2009.02.001 (2009) and Proudfoot, N. J., Furger, A. & Dye, M. J. Integrating mRNA processing with transcription. Cell 108, 501-512 (2002)). CRY2 와 CIB1 사이의 상호작용이 초단위 이하의 기간으로 발생함에도 불구하고 (Kennedy, M. J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. Nature methods 7, 973-975, doi: 10.1038/nmeth.1524 (2010)), LITE-매개 활성화는 전사의 내재 반응속도에 의해 제한될 가능성이 있다. 출원인은 30 분 내지 24 시간의 광 자극 시간 경과 동안 mRNA 수준을 측정함으로써 LITE-매개 Neurog2 발현의 온-키네틱스 (on-kinetics) 를 조사하였다 (도 36C). Neurog2 mRNA 의 상대적 수준은 광 자극 개시 30 분 후 만큼 상당히 일찍 증가하였으며 GFP-트랜스펙션된 음성 대조군에 비해 대략 20 배 상향조절되어 12 시간에서 포화될 때까지 꾸준히 상승하였다. 유사하게, 출원인은 6 시간 동안 세포를 자극하고 조명 중단 후 여러 시점에서의 Neurog2 전사체의 수준을 측정함으로써 시스템의 오프-키네틱스 (off-kinetics) 를 평가하였다 (도 36D). Neurog2 mRNA 수준은 자극후 30 분까지 짧게 증가하였으며, 효과는 잔류 CRY2PHR-CIB1 이량체화 또는 이전에 영입된 RNA 중합효소로 인한 것일 수 있다. 그 후, Neurog2 발현은 ~3 시간의 반감기로 감소하였는데, 이는 전사체가 광 자극의 부재 하 자연적 수준으로 복구하는 것을 입증한다. 반대로, 식물 호르몬 아브시스산 수용체를 기반으로 한 소분자 유도 TALE 시스템 (Liang, F.-S., Ho, W. Q. & Crabtree, G. R. Engineering the ABA Plant Stress Pathway for Regulation of Induced Proximity. Sci. Signal. 4, rs2-, doi: 10.1126/scisignal.2001449 (2011)) 은 약물 확산, 대사 또는 청소 (clearance) 에 의해 잠재적으로 제한되는 더 느린 온- 및 오프-키네틱스를 나타내었다 (도 44).

[0351] 출원인은 다음으로 바이러스 형질도입을 통한 뉴런 적용을 위한 LITE 의 유용성을 탐색하였다. 출원인은 TALE 유전자의 전달을 위한 아데노-연관 바이러스 (AAV)-기반 벡터 및 AAV 생성을 위한 단순화된 프로세스를 개발하였다 (도 37A 및 B, 도 45 및 실시예 11). AAV 의 ssDNA-기반 계놈은 재조합에 덜 취약하여, 렌티바이러스 벡터에 비하여 이점을 제공한다 (Holkers, M. et al. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. Nucleic acids research 41, e63, doi: 10.1093/nar/gks1446 (2013)).

[0352] 1 차 마우스 피질 뉴런에서의 전사를 조정하기 위한 AAV-매개 TALE 전달을 특정화하기 위해, 출원인은 신경전달 또는 뉴런 분화에 관여된 유전자, 이온 채널 서브유닛, 및 신경학적 질환에 연루된 유전자를 포함하는 총 28 개 쥐과 유전자좌를 표적화하는 TALE-VP64 전사 활성화제 패널을 구축하였다. 각 표적 유전자의 프로모터에서의 DNase I-민감성 부위는 TALE 결합 서열 선택에 대한 안내를 제공하였다 (도 46). 출원인은 TALE 활성이 출원인의 AAV-TALE 생성 프로세스를 사용하여 효율적으로 스크리닝될 수 있다는 것을 확인하였으며 (도 45), 이러

한 방식으로 선택되고 AAV 벡터를 사용하여 1 차 뉴런에 전달된 TALE 이 다양한 어레이의 유전자 표적을 가변적 정도로 활성화시켰다는 것을 발견하였다 (도 37C). 또한, AAV-TALE 의 입체공간적 전달은 마우스 전전두엽 피질에서의 강력한 생체내 발현을 매개하였다 (도 37D, E). 마우스 변연계 아래 피질 (ILC) 에서의 TALE(Grm2)-VP64 의 발현은 GFP-주입 대조군에 비해 Grm2 mRNA 수준에 있어서 2.5 배 증가를 유도하였다 (도 37F).

[0353] TALE 활성제를 배양된 1 차 뉴런에 전달하여, 출원인은 다음으로 LITE 성분의 전달을 위한 벡터로서 AAV 를 사용하는 것을 탐색하였다. 이를 위해, 출원인은 LITE 전이유전자가 포함된 각각의 재조합 AAV 의 총 바이러스 게놈 크기가 4.8 kb 의 패키징 한계를 초과하지 않았다는 것을 확실히 할 필요가 있었다 (Wu, Z., Yang, H. & Colosi, P. Effect of Genome Size on AAV Vector Packaging. Mol Ther 18, 80-86 (2009)). 출원인은 TALE N- 및 C-말단을 단축시키고 (N-말단에서 136 aa 를, C-말단에서 63 aa 를 유지), CRY2PHR (1.5kb) 및 CIB1 (1kb) 도메인을 교환하였다 (TALE-CIB1 및 CRY2PHR-VP64; 도 38A). 이러한 LITE 를 2 개 AAV 벡터의 조합에 의해 동시-형질도입을 통해 1 차 피질 뉴런에 전달시켰다 (도 38B; 개별 성분에 대해 83-92% 의 전달 효율, >80% 동시-형질도입 효율). 출원인은 0.8% 의 감소된 사용률로 2 광 펄스 주파수에서 Grm2-표적화 LITE 를 시험하여 뉴런 건강성을 확인하였다 (도 47). 두 자극 조건 모두 Grm2 mRNA 수준에 있어서 ~7 배 광 의존적 증가를 달성하였다 (도 38C). 추가의 연구는 실질적 표적 유전자 발현 증가가 신속히 이루어질 수 있다는 것을 확인하였다 (4 시간 내 4 배 상향조절; 도 38D). 또한, 출원인은 자극 후 mGluR2 단백질의 유의한 상향조절을 관찰하였는데, 이는 mRNA 수준에서 LITE 에 의해 영향받은 변화가 단백질 수준으로 옮겨지는 것을 입증한다 (GFP 수준에 대해  $p < 0.01$ , 무광 조건에 대해  $p < 0.05$ ; 도 38E).

[0354] LITE 시스템을 생체내 적용하기 위해서, 출원인은 CRY2PHR-VP64 LITE 성분 및 Grm2-표적화 TALE-CIB1 을 운반하는 고농도 AAV 벡터 ( $10^{12}$  DNaseI 저항성 입자/mL) 의 1:1 혼합물을 야생형 C57BL/6 마우스의 ILC 에 입체공간적으로 전달하였다. 생체내 LITE-발현 뉴런의 광학적 자극을 제공하기 위해서, 출원인은 주입 부위에서 광섬유 캐놀라를 이식하였다 (도 38F 및 도 48) (Zhang, F. et al. Optogenetic interrogation of neural circuits: technology for probing mammalian brain structures. Nat Protoc 5, 439-456, doi: 10.1038/nprot.2009.226 (2010)). 주입 부위에서의 뉴런은 TALE(Grm2)-CIB1 및 CRY2PHR-VP64 모두를 발현하는 >80% 의 형질도입된 세포를 갖는 바이러스 모두에 의해 효율적으로 동시-형질도입되었다 (도 38G 및 도 49). 수술 후 8 일에, 출원인은 고체-상태 473 nm 레이저를 이식된 섬유 캐놀라에 연결시켜 거동 중인 마우스의 ILC 를 자극하였다. 12 시간 자극 기간 후 (5 mW, 0.8% 사용률, 0.0167 Hz 에서 0.5 초 광 펄스 사용), Grm2 mRNA 에서의 변화에 대해 광 섬유 캐놀라 이식 부위로부터의 뇌 조직을 분석하였다 (도 38H). 출원인은 미자극 ILC ( $p \leq 0.01$ ) 에 비해 광 자극 후 Grm2 mRNA 에 있어서 유의한 증가를 관찰하였다. 종합하여, 이러한 결과로, LITE 가 배양된 뉴런에서 및 생체내에서 내인성 유전자 발현의 광학적 제어를 가능하게 한다는 것이 확인된다.

[0355] 생체내 LITE 활성제의 무광 조건에서 관찰된 기저 상향조절의 지속성으로 인해, 출원인은 배경의 근원을 확인하고 약화시키며 광-매개 유전자 유도의 효율 (유전자 발현의 광/무광 비) 을 향상시키는 것을 목표로 하는 또 다른 라운드의 최적화를 수행하였다. LITE 표적화 성분 TALE-CIB1 만을 발현하는 뉴런은 LITE 성분 모두 (모두 GFP 대조군에 대해  $p < 0.001$ ) 를 발현하는 미자극 뉴런에서 발견된 바와 유사한 Grm2 mRNA 증가를 생성시킨 한편, 효과기 성분 CRY2PHR-VP64 단독은 전사에 유의하게 영향을 미치지 않았는데 ( $p > 0.05$ , 도 50), 이는 LITE 에 의한 배경 전사 활성화가 오직 DNA 표적화 성분으로부터만 발생할 수 있다는 것을 시사한다.

[0356] 따라서, 출원인은 TALE-CIB1 에 의한 기저 표적 상향-조절을 감소시키기 위해 포괄적 스크리닝을 실행하였다 (도 51). 최적화는 다음의 2 가지 계획에 초점을 맞추었다: 첫 번째로, CIB1 은 식물 전사 인자이며 포유동물 세포에서도 고유 조절 효과를 가질 수 있다 (Liu, H. et al. Photoexcited CRY2 Interacts with CIB1 to Regulate Transcription and Floral Initiation III Arabidopsis. Science 322, 1535-1539, doi: 10.1126/science.1163927 (2008)). 출원인은 고등 식물의 염기성 헬릭스-루프 헬릭스 전사 인자 중 보존된 3 개 CIB1 부위를 결실시켜 이러한 효과를 없애는 것을 모색하였다 (도 51). 두 번째로, 출원인은 TALE-CIB1 이 광의 부재 하 표적 유전자좌에 결합하는 것을 방지하는 것을 목표로 하였다. 이를 달성하기 위해, 출원인은 LS-함유 CRY2PHR-VP64 로의 광-유도 이량체화때까지 세포질에서 국부화되도록 TALE-CIB1 을 조작하였다 (도 52). 2 가지 계획 모두를 시험하기 위해서, 출원인은 73 개의 별개 LITE 구조를 평가하고 12 개의 효과기-표적화 도메인 쌍을 확인하였는데 (도 51 및 도 53 에서 "+" 열로 나타냄) 이는 향상된 광-유도 효율 및 감소된 전체 기초선 모두를 가졌다 (본래 LITE1.0 과 비교한 무광 조건에서의 mRNA 증가 배수;  $p < 0.05$ ). LITE2.0 로 지정된, 2 가지 계획 모두를 통합하는 하나의 구조는 최고의 광 유도를 입증하였으며 (광/무광 =

20.4) 본래 구조와 비교하여 6 배 초과로 감소된 배경 활성화가 초래되었다 (도 38I). 또 다른 - LITE1.9.1 - 은 4 배 광 유도를 유지하면서 최소의 배경 mRNA 증가 (1.06) 를 생성시켰다 (도 53).

[0357] 출원인은 TALE 및 LITE 조정에 의해 접근가능한 프로세스의 범위를 더 확장시키는 것을 모색하였다. 내인성 전사 억제제는 종종 히스톤 메틸 메틸트랜스퍼라아제 (HMT) 및 데아세틸라아제 (HDAC) 와 같은 크로마틴 변형 효소에 의해 매개된다. 출원인은 이전에 mSin3 상호작용 도메인 (SID), mSin3-HDAC 복합체의 부분이 293FT 세포에서 표적 유전자를 하향조절하기 위하여 TALE 과 융합될 수 있다는 것을 나타낸 바 있다 (Beerli, R. R., Segal, D. J., Dreier, B. & Barbas, C. F., 3rd. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14628-14633 (1998) and Cong, L., Zhou, R., Kuo, Y.-c., Cunniff, M. & Zhang, F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat Commun* 3, 968, doi:[http://www.nature.com/ncomms/journal/v3/n7/supinfo/ncomms1962\\_S1.html](http://www.nature.com/ncomms/journal/v3/n7/supinfo/ncomms1962_S1.html) (2012)).

이러한 TALE 억제제를 추가로 향상시키기 위해, 출원인은 VP64 의 쿼드러플 (quadruple) VP16 텐덤 반복물 구조와 유사한 SID 의 4 개 반복물 (Beerli, R. R., Segal, D. J., Dreier, B. & Barbas, C. F., 3rd. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14628-14633 (1998)) 이 유전자 전사를 억제하는 이의 효력을 증가시킬 수 있다고 추론하였다. 실제로, TALE-SID4X 구축물은 293FT 세포에서 TALE-SID 만큼 2 배 효율적이었으며 (도 54A 및 54B) 또한 뉴런에서의 효율적 유전자 억제를 매개하였다 (도 54C 및 54D).

[0358] 출원인은 내인성 유전자좌에 대한 히스톤 효과기의 TALE-매개 표적화가 특이적 후생적 변형을 유도할 수 있어, 후생 뿐 아니라 전사 역학의 탐색을 가능하게 한다는 가설을 세웠다 (도 39A). 출원인은 CRY2PHR-SID4X 구축물을 생성시켰으며, 표적화된 Grm2 프로모터에서의 H3K9 아세틸화에 있어서 ~2 배 감소에 부수적으로 (도 39D), 뉴런에서의 Grm2 의 광-매개 전사 억제를 입증하였다 (도 39B 및 도 39C). 유전자좌 특이적 히스톤 개질에 대한 히스톤 잔기 표적의 다양성을 확장시키려는 노력으로, 출원인은 문헌으로부터 일련의 억제성 히스톤 효과기 도메인을 유도하였다 (표 6). 광범위한 계통발생적 스펙트럼에 걸쳐 끌어내어, 도메인은 HDAC, 히스톤 메틸트랜스퍼라아제 (HMT) 및 히스톤 아세틸트랜스퍼라아제 (HAT) 저해제 뿐 아니라 HDAC 및 HMT 영입 단백질을 포함하였다. 효율적 AAV 패키징이 촉진되도록 작은 크기의 기능적 절두 및 단백질이 바람직하다.

생성된 후생적-변형 TALE-히스톤 효과기 융합 구축물 (epiTALE) 을 1 차 뉴런 및 Neuro 2a 세포에서 각각 Grm2 및 Neurog2 전사를 억제하는 그의 능력에 대해 시험하였다 (도 39E, 도 39F 및 도 55). 1 차 뉴런에서, 24 개 epiTALE 중 23 개가  $p < 0.05$  의 통계적 기준을 사용하여 grm2 의 전사를 성공적으로 억제하였다.

유사하게, Neuro 2a 세포에서의 epiTALE 발현은 시험한 32 개 히스톤 효과기 도메인 중 20 개에 대해 Neurog2 발현 감소를 초래하였다 ( $p < 0.05$ ). 유망한 epiTALE 의 하위집합을 1 차 뉴런 및 Neuro 2a 세포에서 발현시키고, 표적화된 내인성 프로모터에서의 상대적 히스톤 잔기 마크 (mark) 수준을 ChIP-RT-qPCR 에 의해 정량하였다 (도 39G, 도 39H 및 도 56). 1 차 뉴런 또는 Neuro 2a 세포에서, H3K9me1, H4K20me3, H3K27me3, H3K9ac 및 H4K8ac 의 수준은 각각, KYP (A. 탈리아나), TgSET8 (T. 곤디), NUE 및 PHF19 (C. 트라초마티스 (trachomatis) 및 H. 사피엔스), Sin3a, Sirt3 및 NcoR (모두 H. 사피엔스) 및 hdac8, RPD3 및 Sir2a (X. 라에비스 (laevis), S. 세레비지에, P. 팔시파룸 (falciparum)) 에서 유래한 epiTALE 에 의해 변형되었다. 이러한 도메인은 준비된 후생적 효과기의 근원을 제공하여 LITE 에 의한 전사 및 후생적 제어 범위를 확장시킨다.

[0359] 뇌와 같은 이종 조직에서의 시공간적으로 정확한 생체내 유전자 조절을 달성하는 능력은 조사자로 하여금 발생, 학습, 기억 및 질환 진행과 같이 다양한 프로세스에서의 동적 유전자 조절의 역할에 대해 의문을 갖게끔 한다.

LITE 는 시간적으로 정확하고, 공간적으로 표적화된, 세포주, 1 차 뉴런 및 마우스 뇌에서의 생체내 내인성 유전자 발현의 바이모달 (bimodal) 제어를 가능하게 하는데 사용될 수 있다. LITE 의 TALE DNA 결합 성분은 광범위한 계통 유전자좌를 표적화하도록 커스터마이징될 수 있으며, 다른 DNA 결합 도메인 예컨대 RNA-안내 Cas9 효소 (Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819823 (2013)) 를 TALE 대신 사용하여 다재다능한 유전자좌-특이적 표적화를 가능하게 할 수 있다 (도 57). 신규한 방식의 LITE 조정은 또한 효과기 모듈을 후생적 변형 효소와 같은 새로운 관능기로 대체함으로써 이루어질 수 있다 (de Groote, M. L., Verschure, P. J. & Rots, M. G. Epigenetic Editing: targeted rewriting of epigenetic marks to modulate expression of selected target genes. *Nucleic acids research* 40, 10596-10613, doi:10.1093/nar/gks863 (2012)). 따라서, LITE 시스템은 현존하는 광유전학적 도구상자에 대한 새로운 집합의 능력이 생길 수 있게 하며, 광을 사용하는 내인성 유전자 조절을 변형시키기 위한 매우 일반화가능하고 다

재다능한 플랫폼을 구축한다.

- [0360] **방법 요약.** LITE 구축물을 GenJet 를 사용하여 Neuro 2A 세포에 트랜스펙션하였다. TALE 또는 LITE 구축물을 운반하는 AAV 벡터를 사용하여 마우스 1 차 배아 피질 뉴런 뿐 아니라 마우스 뇌 (생체내) 를 형질도입하였다. RNA 를 추출하고 역전사하고 TaqMan-기반 RT-qPCR 을 사용하여 mRNA 수준을 측정하였다. 광 방사 다이오드 또는 고체 상태 레이저를 각각 조직 배양물 및 생체내 광 전달에 사용하였다.
- [0361] **LITE 의 설계 및 구축.** 모든 LITE 구축물 서열은 실시예 11 에서 발견될 수 있다.
- [0362] **Neuro 2a 배양 및 실험.** Neuro 2a 세포 (Sigma-Aldrich) 를 5% HyClone 열-불활성화 FBS (Thermo Scientific), 1% 페니실린/스트렙토마이신 (Life Technologies) 이 보충된, 1:1 비의 OptiMEM (Life Technologies) 대 GlutaMax 및 나트륨 피루베이트를 갖는 고-글루코오스 DMEM (Life Technologies) 을 함유하는 배지에서 성장시키고, 2 일마다 1:5 에서 계대하였다. 120,000 개 세포를 24-웰 플레이트의 각 웰에 트랜스펙션 18-20 시간 전에 플레이팅하였다. 트랜스펙션 1 시간 전에, 배지를 5% HyClone 열-불활성화 FBS 및 1% 페니실린/스트렙토마이신이 보충된 DMEM 으로 바꾸었다. 세포를, 제조사의 지시사항에 따라 1.5  $\mu$ l 의 GenJet (SigmaGen Laboratories) 트랜스펙션 시약을 사용하여 웰 당 1.0  $\mu$ g 의 총 구축물 DNA (등몰비) 로 트랜스펙션하였다. 배지를 트랜스펙션 후 24 시간 및 44 시간에 교환하고, 광 자극을 48 시간에 시작하였다. 자극 매개변수는 하기와 같았다: 5 mW/cm<sup>2</sup>, 466 nm, 7% 사용률 (1 초 광 펄스 0.067 Hz), 도면 범례에서 다르게 나타내지 않는 한 24 시간 동안. RNA 를 제조사의 지시사항에 따라 RNeasy 키트 (Qiagen) 를 사용하여 추출하고, 샘플 당 1  $\mu$ g 의 RNA 를 qScript (Quanta Biosystems) 를 사용하여 역전사하였다. 상대적 mRNA 수준을, 표적화 유전자 뿐 아니라 내인성 대조군으로서 GAPDH 에 대해 특이적인 TaqMan 프로브 (Life Technologies, Taqman 프로브 ID 에 대해서는 표 3 참조) 를 사용하여 정량 실시간 PCR (qRT-PCR) 에 의해 측정하였다.  $\Delta\Delta$ Ct 분석을 사용하여, GFP 단독으로 형질도입되고 광 자극 처리된 음성 대조군에 대한 변화 배수를 수득하였다. LIVE/DEAD 검정 키트 (Life Technologies) 를 지시사항에 따라 사용하여 독성 실험을 수행하였다.
- [0363] **AAV 벡터 생성.** 293FT 세포 (Life Technologies) 를 항생제 불포함 D10 배지 (GlutaMax 및 나트륨 피루베이트를 갖는 DMEM 고 글루코오스, 10% 열-불활성화된 Hyclone FBS 및 1% 1M HEPES) 에서 성장시키고, 매일 1:2-2.5 에서 계대하였다. 계대 총 수를 10 이하로 유지시키고, 세포가 절대 85% 밀집도 (confluence) 를 넘지 않도록 성장시켰다. 트랜스펙션 전 날에, 21.5 mL 의 D10 배지 중  $1 \times 10^6$  세포를 15 cm 디쉬에 플레이팅하고 18-22 시간 동안 또는 ~80% 밀집도가 될 때까지 인큐베이션하였다. 트랜스펙션 시약으로서 사용하기 위해, 1 mg/mL 의 PEI "Max" (Polysciences) 를 물에 용해하고 용액의 pH 를 7.1 로 조정하였다. AAV 생성을 위해서, 10.4  $\mu$ g 의 pDF6 헬퍼 플라스미드, 8.7  $\mu$ g 의 pAAV1 혈청형 패키징 벡터, 및 5.2  $\mu$ g 의 pAAV 벡터 (관심 유전자 운반) 를 434  $\mu$ l 의 혈청-불포함 DMEM 에 첨가하고, 130  $\mu$ l 의 PEI "Max" 용액을 DMEM-희석 DNA 혼합물에 첨가하였다. DNA/DMEM/PEI 각테일을 볼텍싱하고 실온에서 15 분 동안 인큐베이션하였다. 인큐베이션한 후, 트랜스펙션 혼합물을 22 mL 의 완전 배지에 첨가하고, 짧게 볼텍싱하고, 293FT 세포의 15 cm 디쉬에 대해 배지를 교체하는데 사용하였다. 상청액 생성을 위해, 트랜스펙션 산청액을 48 시간에 수확하고, 0.45  $\mu$ m PVDF 필터 (Millipore) 를 통해 여과하고, 분취액으로 분배하고, 저장을 위해 -80°C 에서 동결시켰다.
- [0364] **1 차 피질 뉴런 배양.** 해리된 피질 뉴런을 E16 (Charles River Labs) 상에서 C57BL/6N 마우스 배아로부터 제조하였다. 피질 조직을 빙냉 HBSS - (50 mL 10x HBSS, 435 mL dH<sub>2</sub>O, 0.3 M HEPES pH 7.3, 및 1% 페니실린/스트렙토마이신) 중에서 해부하였다. 피질 조직을 20 mL 의 빙냉 HBSS 로 3 회 세척한 후 37°C 에서 20 분 동안, 8 mL 의 HBSS 중에서 240  $\mu$ l 의 2.5% 트립신 (Life Technologies) 과 함께 소화시켰다. 이후 피질을 1 mL FBS 를 함유하는 20 mL 의 가온된 HBSS 로 3 회 세척하였다. 피질을 2 mL 의 HBSS 중에서 분말화하고, 폴리-D-리신 코팅된 24-웰 플레이트 (BD Biosciences) 에서 150,000 개 세포/웰로 플레이팅하였다. 뉴런을 1X B27 (Life Technologies), GlutaMax (Life Technologies) 및 1% 페니실린/스트렙토마이신이 보충된 Neurobasal 배지 (Life Technologies) 중에서 유지시켰다.
- [0365] **1 차 뉴런 형질도입 및 광 자극 실험.** 1 차 피질 뉴런을 DIV 5 상에서 250  $\mu$ l 의 AAVI 상청액으로 형질도입하였다. 배지 및 상청액을 그 다음날 보통의 완전 neurobasal 로 대체하였다. Neurobasal 을 AAV 형질도입 6 일 후 1X B27, GlutaMax (Life Technologies) 및 1% 페니실린/스트렙토마이신을 함유하는 최소 필수 배지 (Life Technologies) 로 교환하여, 광 자극 동안 Neurobasal 에 함유된 HEPES 및 리보플라빈으로부터의 광독성 생성물의 형성을 방지하였다.

- [0366] 광 자극을  $5 \text{ mW/cm}^2$  의 강도, 0.8% 의 사용률 (0.033Hz 에서의 250 ms 펄스 또는 0.016Hz 에서의 500 ms 펄스), 24 시간 동안 466 nm 청색광 (도면 범례에서 다르게 나타내지 않는 한) 으로, AAV 형질도입 6 일 후 시작하였다 (DIV 11). Cells-to-Ct 키트를 제조사 지시사항에 따라 사용하여 (Life Technologies) RNA 추출 및 역전사를 수행하였다. Neuro 2a 세포에 대해 상기 기재한 바와 같은 TaqMan 프로브를 사용하여 정량 실시간 PCR (qRT-PCR) 에 의해 상대적 mRNA 수준을 측정하였다.
- [0367] **1 차 뉴런의 면역조직화학.** 1 차 뉴런의 면역조직화학을 위해, 세포를 수확 후 폴리-D-리신/라미닌 코팅된 커버슬립 (BD Biosciences) 상에 플레이팅하였다. AAV1-형질도입을 상기 기재한 바와 같이 수행하였다. 뉴런을 4% 파라포름알데히드 (Sigma Aldrich) 로 15 분 동안 실온에서 형질도입 후 7 일 고정시켰다. 블로킹 및 침투화를 DPBS (Life Technologies) 중 0.5% Triton-X100 (Sigma-Aldrich) 및 10% 정상 염소 혈청 (Life Technologies) 으로 1 시간 동안 실온에서 수행하였다. 뉴런을 1 차 항체와 함께 밤새 4°C 에서 인큐베이션 하고, DPBS 로 3 회 세척하고 2 차 항체와 함께 90 분 동안 실온에서 인큐베이션하였다. 항체 제공처 및 사용 농도에 대해서는 표 4 를 참조한다. 커버슬립을 마지막으로, DAPI 와 함께 Prolong Gold Antifade 시약 (Life Technologies) 사용하여 고정하고 X-Cite 120Q 광원 (Lumen Dynamics) 을 갖는 Axio Scope A.1 (Zeiss) 상에서 영상화하였다. AxioCam MRm 카메라 및 AxioVision 4.8.2 를 사용하여 영상을 획득하였다.
- [0368] **웨스턴 블롯.** 총 단백질을 용해물의 제조를 위해, 빙냉 용해 완충액 (RIPA, 세포 신호전달; 0.1% SDS, Sigma-Aldrich; 및 complete ultra protease inhibitor mix, Roche Applied Science) 중에서 1 차 피질 뉴런을 광 자극 (상기 참조) 후 수확하였다. 세포 용해물을 Bioruptor 초음파발생장치 (sonicator) (Diagenode) 에서 'M' 설정으로 5 분 동안 초음파처리하고, 21,000 x g 에서 10 분 동안 4°C 에서 원심분리하였다. RC DC 단백질 검정 (Bio-Rad) 을 사용하여 단백질 농도를 측정하였다. 레인 당 30-40  $\mu\text{g}$  의 총 단백질을, Precision Plus Protein Dual Color Standard (Bio-Rad) 와 함께 4-15% Tris-HCl 겔 (Bio-Rad) 상에서 비환원 조건 하에 분리하였다. 폴리비닐리덴 디플루오라이드 멤브레인 (Millipore) 으로의 습식 전기이동 후, Tris 완충 식염수 (TBS, Bio-Rad) 중 5% BLOT-QuickBlocker (Millipore) 중에서 45 분 동안 멤브레인을 블로킹하고, 웨스턴 블롯을 항-mGluR2 (Abcam, 1:1,000) 및 항- $\alpha$ -튜블린 (Sigma-Aldrich 1:20,000) 으로 밤새 4°C 에서 프로브 처리한 후, 세척하고 항-마우스-IgG HRP 항체 인큐베이션하였다 (Sigma-Aldrich, 1:5,000 - 1:10,000). 추가의 항체 세부사항에 대해서는 표 4 를 참조한다. ECL 웨스턴 블롯 기질 (SuperSignal West Femto 키트, Thermo Scientific) 을 통해 검출을 수행하였다. AlphaImager (Innotech) 시스템으로 블롯을 영상화하고, ImageJ 소프트웨어 1.46r 을 사용하여 정량하였다.
- [0369] **농축되고 정제된 AAV1/2 벡터의 생성.** 생체내 입체공간적 주입을 위한 농축되고 정제된 AAV 의 생성을, AAV1 상청액 생성에 대해 상기 요약한 동일한 초기 단계를 사용하여 수행하였다. 그러나 트랜스펙션에 대해서는, 동일 비의 AAV1 및 AAV2 혈청형 플라스미드를 AAV1 단독 대신 사용하였다. 구축물 ekdd 5 개 플레이트를 트랜스펙션하고 세포를 트랜스펙션 후 48 시간에 세포-스크래퍼로 수확하였다. HiTrap 헤파린 친화성 컬럼 (GE Healthcare) 을 사용하여 AAV1/2 입자의 정제를 수행하였다 (McClure, c., Cole, K. L., Wulff, P., Klugmann, M. & Murray, A. J. Production and titrating of recombinant adeno-associated viral vectors. J Vis Exp, e3348, doi:10.3791/3348 (2011)). 출원인은 500  $\mu\text{l}$  농축 컬럼 (100 kDa 절삭값, Millipore) 을 사용하여 구축물 당 100  $\mu\text{l}$  의 최종 부피에 이르기까지 제 2 농축 단계를 추가하여, 더 높은 바이러스 역가를 획득하였다. WPRE 에 대한 커스텀 Taqman 프로브를 사용하여 (Life Technologies) qRT-PCR 에 의해 AAV 적정을 수행하였다. qRT-PCR 전에, 농축된 AAV 를 DNaseI (New England Biolabs) 으로 처리하여 DNaseI-저항성 입자만에 대한 측정값을 획득하였다. DNaseI 열-불활성화 후, 바이러스 외피를 프로테이나아제 K 소해 (New England Biolabs) 에 의해 분해시켰다. 바이러스 역가를, 공지된 WPRE 카피수를 갖는 표준 곡선을 기반으로 계산하였다.
- [0370] **AAV1/2 의 입체공간적 주입 및 광학적 이식.** 모든 동물 절차는 동물 관리에 대한 MIT 위원회에 의해 승인받았다. 성체 (10-14 주령) 수컷 C57BL/6N 마우스를 케타민/자일라진 (100 mg/kg 케타민 및 10 mg/kg 자일라진) 을 복강내 (i.p.) 주입하여 마취하였고, 선제성 무통증이 주어졌다 (Buprenex, 1 mg/kg, i.p.). 승인된 절차에 따라 개두술을 수행하고 1  $\mu\text{l}$  의 AAV1/2 을 0.35/1.94/-2.94 (브레그마 (bregma) 에 대해 mm 으로 측면, 전면 및 하면 좌표) 에서 ILC 에 주입하였다. 동일한 수술적 절차 동안, 섬유를 갖는 광학적 캐논라 (Doric Lenses) 를, 브레그마에 대하여 0.35/1.94/-2.64 에 위치한 광학 섬유의 말단에 단독으로 ILC 에 이식하였다. 상기 캐논라를 Metabond 치아용 시멘트 (Parkell Inc) 및 Jet 의치 보수물 (Lang dental) 을 사용하여 두개골에 부착시켜, 이를 둘러싼 안정한 원뿔형을 구축하였다. 절개부를 봉합하고 적절한 수술-후 진

통제를 수술 후 3 일 동안 투여하였다.

- [0371] **ILC 뇌 섹션에 대한 면역조직화학.** 마우스에 케타민/자일라진 마취제의 치사량을 주입하고 PBS 및 4% 파라포름알데하이드 (PFA) 로 경심 관류하였다. 뇌를 4% PFA 중에서 4℃ 에서 밤새 추가적으로 고정시킨 후, 동결 보호를 위해 밤새 실온에서 30% 수크로오스에 이동시켰다. 뇌를 이후 Tissue-Tek 최적 절단 온도 (Optimal Cutting Temperature) (OCT) 화합물 (Sakura Finetek) 에 옮기고 -80℃ 에서 동결시켰다. 18 μm 섹션을 cryostat (Leica Biosystems) 상에서 절단하고 Superfrost Plus 유리 슬라이드 (Thermo Fischer) 에 탑재시켰다. 섹션을 15 분 동안 4% PFA 로 후-고정하고, 1 차 뉴런에 대해 상기 기재한 바와 같이 면역조직화학을 수행하였다.
- [0372] **ILC 에서의 광 자극 및 mRNA 수준 분석.** 수술 후 8 일에, 각성 상태이며 자유로이 움직이는 마우스를 섬유 패치 케이블 및 회전식 연결부를 통해 광학 이식물에 연결된 473 nm 레이저원 (OEM Laser Systems) 을 사용하여 자극하였다. 자극 매개변수는 1 차 뉴런에 사용한 바와 동일하였다: 5 mW (총 출력), 0.8% 사용률 (0.016 Hz 에서 500 ms 광 펄스), 총 12 시간 동안. 형질도입된 구축물 및 광 자극을 포함하는 실험 조건을 표 5 에 열거한다.
- [0373] 광 자극 종료 후, 마우스를 CO<sub>2</sub> 를 사용하여 안락사시키고, 전액골 피질 (PFC) 을 얼음 상에서 신속히 해부하고 RNA later (Qiagen) 중에서 4℃ 에서 밤새 인큐베이션하였다. 200 μm 섹션을 바이브라툼 (vibratome) (Leica Biosystems) 상에서 4℃ 에서 RNA later 중에서 절단하였다. 이후 섹션을 드라이 아이스 상 유리 커버슬라이드에서 동결시키고, 바이러스 형질도입된 ILC 를 형광 입체현미경 (Leica M165 FC) 하에 확인하였다. 광학 섬유 트랙 종료부에 직접적으로 배 쪽에 위치한 ILC 의 0.35 mm 직경 펀치를 추출하였다 (Harris unicorn, Ted Pella). 뇌 펀치 샘플을 이후 50 μl Cells-to-Ct RNA 용해 완충액 중에서 RNase-불포함 펠렛-막자 그라인더 (Kimble Chase) 를 사용하여 균질화하고, RNA 추출, 역전사 및 qRT-PCR 을 1 차 뉴런 샘플에 대해 기재한 바와 같이 수행하였다.
- [0374] **크로마틴 면역침전.** 뉴런 또는 Neuro2a 세포를 상기 기재한 바와 같이 배양하고 형질도입하거나 트랜스펙션 하였다. ChIP 샘플을, 세포 수 및 세포 유형에 대해 소수 조정하여, 이전에 기재된 바와 같이 제조하였다 (Blecher-Gonen, R. et al. High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. Nature protocols 8, 539-554 (2013)). 세포를 24-웰 포맷으로 수확하고, 96-웰 포맷으로 세척하고, 용해를 위해 미세원심분리 튜브에 옮겼다. 샘플을 이후, 30 초 온/오프 주기를 사용하여 21 분 동안 Biorupter 초음파처리 장치 (Diagenode) 로 수조 초음파처리에 의해 직접적으로 용해하였다. qPCR 을 사용하여 표적화된 유전자좌에서의 히스톤 마크의 축적을 평가하였다.
- [0375] **통계적 분석.** 모든 실험을 최소 2 개의 독립적인 생물학적 복제품로 수행하였다. 2 개 조건 비교시 스튜던트 양방향 t-검정, 여러 샘플을 서로 비교시 터키 사후검정 분석과 함께 ANOVA, 및 여러 샘플을 음성 대조군에 대해 비교시 던넷 사후검정 분석과 함께 ANOVA 를 사용하여, Prism (GraphPad) 으로 통계적 분석을 수행하였다.
- [0376] **실시예 11**
- [0377] **실시예 10 에 대한 보충적 정보: 내인성 포유동물 전사의 광학적 제어**
- [0378] **광자극 하드웨어 - 시험관내.** 커스텀 빌트 LED 광자극 장치를 사용하여 시험관내 광 자극 실험을 수행하였다. 모든 전자 부품을 커스텀 프린트 회로 보드 (ExpressPCB) 상에 탑재하였다. 피크가 466 nm 인 청색 LED (모델 #: YSL-R542B5C-A11, China Young Sun LED Technology; 'LED - Super Bright Blue' COM-00529 로서 SparkFun Electronics 에 의해 배포됨) 를 Corning 24-웰 플레이트의 웰과 함께 정렬된 3 개 군에서 어레이화 하였다. LED 전류 흐름을 25 mA Dyna0hm 드라이버 (LEDDynamics #4006-025) 에 의해 조절하였다. LED 어레이의 열을 Arduino UNO 마이크로컨트롤러 보드를 통해 TTL 컨트롤 (Fairchild Semiconductor PN2222BU-ND) 에 의해 어드레스하였다. Thorlabs PM100D 파워 미터 및 S120VC 광다이오드 검출기를 이용하여 광 출력량을 어레이 위 80 mm 의 거리로부터 측정하였다. 환기 공간을 제공하고 광 펄드 균일성을 최대화하기 위해, 80 mm 길이 환기 스페이서를 LED 어레이와 24-웰 샘플 플레이트 사이에 두었다. 팬 (Evercool EC5015M12CA) 을 스페이서 유닛의 한 웰에 따라 탑재하면서, 반대쪽 웰을 기류 증가가 가능하도록 간격을 두어 제작하였다.
- [0379] **ImageJ 소프트웨어를 사용하는 LIVE/DEAD® 검정의 정량.** LIVE/DEAD (Life Technologies) 염색 세포의 영상을 형광 현미경에 의해 캡처하고 하기와 같이 처리하였다: 배경을 제한다 (Process → Subtract Background).

형광 부위를 기반으로 한 역치를 세포 상태의 정확한 확인이 가능하도록 설정한다 (Image → Adjust → Threshold). 분할 분석을 수행하여 개별 세포의 자동화된 계수가 가능하게 한다 (Process → Binary → Watershed). 마지막으로, 잔해 신호를 필터링하고 세포를 계수한다 (Analyze → Analyze Particles). 사멸 세포의 % 로서 독성을 측정하였다.

[0380] **화학적 유도 TALE.** Neuro2A 세포를 5% HyClone 열-불활성화 FBS (Thermo Scientific), 1% 페니실린/스트렙토마이신 (Life Technologies) 및 25mM HEPES (Sigma Aldrich) 가 보충된, 1:1 비의 OptiMEM (Life Technologies) 대 GlutaMax 및 나트륨 피루베이트를 갖는 고-글루코오스 DMEM (Life Technologies) 을 함유하는 배지에서 성장시켰다. 150,000 개 세포를 24-웰 플레이트의 각 웰에 트랜스펙션 18-20 시간 전에 플레이팅하였다. 세포를, 웰 당 1  $\mu$ g 의 총 구축물 DNA (등몰비) 및 2  $\mu$ l 의 리포펙타민 2000 (Life Technologies) 으로, 제조사의 권고 프로토콜에 따라 트랜스펙션하였다. 배지를 트랜스펙션 후 12 시간에 교환하였다. 반응속도 시험을 위해서, 화학적 유도를 트랜스펙션 후 24 시간에 시작하였고, 이때 아브시스산 (ABA, Sigma Aldrich) 을 250  $\mu$ M 의 최종 농도로 새 배지에 첨가하였다. RNeasy 키트 (Qiagen) 를 제조사의 지시사항에 따라 사용하여 RNA 를 추출하고, 샘플 당 1  $\mu$ g 의 RNA 를 qScript (Quanta Biosystems) 를 사용하여 역전사하였다. 상대적 mRNA 수준을, 표적화 유전자 뿐 아니라 내인성 대조군으로서 GAPDH 에 대해 특이적인 TaqMan 프로브 (Life Technologies, Taqman 프로브 ID 에 대해서는 보충적 표 2 참조) 를 사용하여 정량 실시간 PCR (qRT-PCR) 에 의해 측정하였다.  $\Delta\Delta$ Ct 분석을 사용하여, GFP 로 모의 트랜스펙션한 세포 인 음성 대조군에 대한 변화 배수를 획득하였다.

[0381] **Cas9 전사 효과기.** HEK 293FT 세포를, 24 웰 디쉬에 시딩하고 24 시간 후 리포펙타민 2000 (Life Technologies) 을 사용하여 돌연변이체 Cas9 융합 단백질 및 합성 안내 RNA (sgRNA) 와 함께 동시-트랜스펙션하였다. 트랜스펙션 후 72 시간에, 총 RNA 를 정제하였다 (RNeasy Plus, Qiagen). 1 $\mu$ g 의 RNA 를 cDNA 로 역전사하였다 (qScript, Quanta BioSciences). 정량 실시간 PCR 을 제조사의 지시사항 (Life Technologies) 에 따라 수행하고, hKlf4 (Hs00358836\_m1), hSox2 (Hs01053049\_s1) 및 내인성 대조군 GAPDH (Hs02758991\_g1) 에 대해 TaqMan 검정을 사용하여 3 반복물로 수행하였다.

[0382] hSpCas9 활성제 플라스미드를 hEF1a 프로모터의 발현 하에 렌티바이러스 벡터 내로 클로닝하였다 (pLenti-EF1a-Cas9-NLS-VP64). hSpCas9 억제제 플라스미드를 동일 벡터 내로 클로닝하였다 (pLenti-EF1a-SID4x-NLS-Cas9-NLS). KLF4 유전자좌에 대해 표적화된 안내 서열 (20bp) 은 GCGCGCTCCACACAACCTCAC, GCAAAATAGACAATCAGCA, GAAGGATCTCGCCAATTTG 이다. SOX2 유전자좌에 대해 표적화된 안내 RNA 에 대한 스페이서 서열은 GCTGCCGGTTTTCATGAA, CCGGGCCCGCAGCAAACCTTC,GGGCTGTTCAGGAATAAAT 이다.

[0383] **광유전학적 작동기:** 미생물 및 식물-유래 광 민감 단백질을 광유전학적 작동기로서 조작하여, 멤브레인 전위 (520/Deisseroth, K. Optogenetics. Nature methods 8, 26-29, doi:10.1038/nmeth.f.324 (2011); Zhang, F. et al. The microbial opsin family of optogenetic tools. Cell 147, 1446-1457, doi:10.1016/j.cell.2011.12.004 (2011) and Yizhar, O., Fenno, L. E., Davidson, T. J., Mogri, M. & Deisseroth, K. Optogenetics in neural systems. Neuron 71, 9-34, doi: 10.1016/j.neuron.2011.06.004 (2011)), 세포내 생화학적 신호전달 (Airan, R. D., Thompson, K. R., Fenno, L. E., Bernstein, H. & Deisseroth, K. Temporally precise in vivo control of intracellular signalling. Nature 458, 1025-1029, doi:10.1038/nature07926 (2009)), 단백질 상호작용 (Levskaya, A., Weiner, O. D., Lim, W. A. & Voigt, C. A. Spatiotemporal control of cell signalling using a light-switchable protein interaction. Nature 461, 997-1001, doi:10.1038/nature08446 (2009); Yazawa, M., Sadaghiani, A. M., Hsueh, B. & Dolmetsch, R. E. Induction of protein-protein interactions in live cells using light. Nat Biotechnol 27, 941-945, doi:10.1038/nbt.1569 (2009); Strickland, D. et al. TULIPs: tunable, light-controlled interacting protein tags for cell biology. Nature methods 9, 379-384, doi:10.1038/nmeth.1904 (2012) and Kennedy, M. J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. Nature methods 7, 973-975, doi: 10. 1038/nmeth. 1524 (2010)) 및 이중 유전자 발현 (Yazawa, M., Sadaghiani, A. M., Hsueh, B. & Dolmetsch, R. E. Induction of protein-protein interactions in live cells using light. Nat Biotechnol 27,941-945, doi:10.1038/nbt.1569 (2009); Kennedy, M. J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. Nature methods 7, 973-975, doi:10.1038/nmeth.1524 (2010); Shimizu-Sato, S., Huq, E., Tepperman, J. M. & Quail, P. H. A light-switchable gene promoter system. Nat Biotechnol 20, 1041-1044, doi:10.1038/nbt734 (2002); Ye, H., Daoud-El Baba, M., Peng, R. W. & Fussenegger, M. A synthetic optogenetic transcription device enhances

blood-glucose homeostasis in mice. *Science* 332, 1565-1568, doi:10.1126/science.1203535 (2011); Wang, X., Chen, X. & Yang, Y. Spatiotemporal control of gene expression by a light-switchable transgene system. *Nature methods* 9, 266-269, doi:10.1038/nmeth.1892 (2012) and Polstein, L. R. & Gersbach, C. A. Light-inducible spatiotemporal control of gene activation by customizable zinc finger transcription factors. *J Am Chem Soc* 134, 16480-16483, doi:10.1021/ja3065667 (2012)) 을 포함하는 세포 기능의 광학적 제어를 가능하게 하였다.

[0384] **주변 광 노출:** 모든 세포를 자극 동안을 제외하고 모든 시기에 낮은 광 수준 (<0.01 mW/cm<sup>2</sup>) 에서 배양하였다. 실내에서의 주변 광 (0.1-0.2 mW/cm<sup>2</sup>) 으로서 선택한 이러한 예방책은 LITE 시스템을 상당히 활성화시키는 것으로 발견되었다 (도 36D). 생체내 실험 동안 광으로부터 동물을 보호하기 위해서는 특별한 예방책이 선택되지 않았다 - 이식된 광학 섬유 내 이상적 증식을 가정하더라도, 주변 광으로 인한 섬유 말단에서의 광 전달의 추정치는 <0.01 mW 였다 (200 μm 섬유 코어 직경 및 0.22 개구수를 기준으로 함).

[0385] **Neuro2A 세포에서의 광 자극 매개변수의 최적화:** 근UV 유도 세포독성을 최소화하기 위해서, 출원인은 TALE-CRY2 가 활성화되도록 466 nm 청색 LED 를 선택하였고, 파장은 450 nm 의 CRY2 흡수 최대치로부터 약간 적색-이동되었으나 여전히 80% 초과활성을 유지하였다 (Banerjee, R. et al. The signaling state of Arabidopsis cryptochrome 2 contains flavin semiquinone. *J Biol Chem* 282, 14916-14922, doi:10.1074/jbc.M700616200 (2007)) (도 42). 광 노출을 최소화하기 위해서, 출원인은 강하지 않은 자극 프로토콜을 선택하였다 (0.067 Hz 에서 1 초 광 펄스, ~7% 사용률). 이는 광 사용률이 광범위한 사용률 매개변수 (1.7% 내지 100% 사용률, 도 41) 에 대한 LITE-매개 전사 활성화에 대해 유의한 효과를 갖지 않았다는 출원인의 발견을 기반으로 하였다. 0 내지 10 mW/cm<sup>2</sup> 의 광 강도 범위를 갖는 조명으로, Ngn2 mRNA 수준이 5 mW/cm<sup>2</sup> 이하의 강도의 함수로서 증가하였다는 것이 밝혀졌다. 그러나, Ngn2 mRNA 수준의 증가는 10 mW/cm<sup>2</sup> 에서 감소하였는데 (도 36C), 이는 고강도 광이 LITE 기능 또는 세포 생리학에 대해 유해한 효과를 가질 수 있다는 것을 제시한다. 이러한 관찰을 더 양호하게 특징화하기 위해, 출원인은 살아있는 세포에 대해 칼세인 대조염색하는 에티디움 동종이량체-1 세포독성 검정을 수행하였고, 10 mW/cm<sup>2</sup> 의 높은 자극 강도에서 상당히 더 높은 % 의 에티디움-양성 세포를 발견하였다. 반대로, 5 mW/cm<sup>2</sup> 자극으로부터의 에티디움-양성 세포 계수는 비자극 대조군과 구별가능하지 않았다. 따라서 5 mW/cm<sup>2</sup> 는 낮은 세포독성을 유지하면서 강력한 LITE 활성화를 획득하기에 최적인 것으로 나타났다.

[0386] **1 차 뉴런에서의 광-유도 독성의 감소:** 뉴런에서 LITE 의 초기 적용으로, 배양된 뉴런이 Neuro 2a 세포보다 청색광에 훨씬 더 민감하다는 것이 밝혀졌다. 출원인이 Neuro 2a 세포에 대해 이전에 최적화한 자극 매개변수 (466 nm, 5 mW/cm<sup>2</sup> 강도, 7% 사용률, 0.067 Hz 에서 1 초 광 펄스, 총 24 시간 동안) 는 1 차 뉴런에서 >50% 독성을 초래하였다. 따라서 출원인은, 출원인이 광범위한 사용률이 LITE-매개 전사 활성화에 대해 약간의 효과를 갖는다는 것을 이전에 관찰했던 바 (도 41), 더 낮은 사용률로 생존을 시험하였다. 동일한 광 강도 (5 mW/cm<sup>2</sup>) 에서의 0.8% 의 감소된 사용률 (0.0167 Hz 에서 0.5 초 광 펄스) 은 미자극 배양물과 구별가능한 높은 생존률을 유지하기에 충분하였다 (도 47).

[0387] **생체내 실험에서의 광 전파 및 독성:** 이전의 연구는 뇌 조직에서의 상이한 파장의 광의 전파 효율을 조사하였다. 473 nm 광 (이 연구에서 사용한 파장) 에 대해, 0.35 mm 의 조직을 통과한 후 >90% 감쇠가 존재하였다 (Witten, Hana B. et al. Recombinase-Driver Rat Lines: Tools, Techniques, and Optogenetic Application to Dopamine-Mediated Reinforcement. *Neuron* 72, 721-733, doi:http://dx.doi.org/10.1016/j.neuron.2011.10.028 (2011)). 추정된 5 mW/cm<sup>2</sup> 광력 밀도를, 조직의 0.35 mm 조직 깊이 (이 연구에서 사용한 뇌 편치의 직경) 및 5 mW 의 총 전력 출력을 기준으로 추정하였다. 생체내 사용한 광 자극 사용률은 1 차 뉴런에 사용한 바 (도 47) 와 동일하였다 (0.8%, 0.0167 Hz 에서 0.5 초).

[0388] **CRY2 흡수 스펙트럼:** CRY2 의 흡수 스펙트럼 조명을 도 42 에 나타내었다. 상기 스펙트럼은 480 nm 초과활성 흡수에서 첨예한 감소를 나타내었다 (Banerjee, R. et al. The Signaling State of Arabidopsis Cryptochrome 2 Contains Flavon Semiquinone. *Journal of Biological Chemistry* 282, 14916-14922, doi:10.1074/jbc.M700616200 (2007)). 500 nm 초과활성 파장은 사실상 흡수되지 않았는데, 이는 황색 또는

적색광 민감성 단백질로의 추후 멀티모달 광학적 제어에 유용할 수 있다.

[0389]

**AAV1 상청액 프로세스의 발전:** 종래의 AAV 입자 생성은 번거로운 생성 및정제 프로세스를 필요로 하며, 많은 구축물을 동시에 시험하는 것이 비현실적이 되게 하였다 (Grieger, J. C., Choi, V. W. & Samulski, R. J. Production and characterization of adeno-associated viral vectors. Nat Protoc 1, 1412-1428, doi:10.1038/nprot.2006.207 (2006)). 이러한 연구에서, 간단하지만 상당히 효율적인 AAV 생성 프로세스는 트랜스펙션된 293FT 세포로부터의 여과된 상청액을 사용하였다 (도 43). 최근 보고는 293FT 세포에서 생성된 AAV 입자가 세포질에서 뿐만 아니라 배양 배지에서도 상당량으로 발견될 수 있었다는 것을 나타낸다 (Lock M, A. M., Vandenberghe LH, Samanta A, Toelen J, Debyser Z, Wilson JM. Rapid, Simple, and Versatile Manufacturing of Recombinant Adeno-Associated Viral Vectors at Scale. Human Gene Therapy 21, 1259-1271, doi:10.1089/hum.2010.055 (2010)). 숙주 세포의 사이토졸과 상청액 사이의 바이러스 입자 비는 AAV 혈청형에 따라 가변적이며, 폴리에틸렌이민 (PEI) 을 바이러스 패키징 플라스미드를 트랜스펙션하는데 사용한 경우 분비가 증진되었다 (Lock M, A. M., Vandenberghe LH, Samanta A, Toelen J, Debyser Z, Wilson JM. Rapid, Simple, and Versatile Manufacturing of Recombinant Adeno-Associated Viral Vectors at Scale. Human Gene Therapy 21, 1259-1271, doi:10.1089/hum.2010.055 (2010)). 현 연구에서, TALE 을 운반하는 AAV 벡터 (도 38A) 로 트랜스펙션되고 AAV1 혈청형을 사용하여 패키징된  $2 \times 10^5$  개 293FT 세포가 1 mL 당  $5.6 \pm 0.24 \times 10^{10}$  DNaseI 저항성 게놈 카피 (gc) 의 농도에서  $250 \mu\text{l}$  의 AAV1 을 생성시킬 수 있었다는 것이 발견되었다.  $250 \mu\text{l}$  의 여과된 상청액은 80-90% 의 효율로 150,000 개 1 차 피질 뉴런을 형질도입시킬 수 있었다 (도 38B 및 도 43). 이러한 프로세스는 또한 96-웰 포맷에 성공적으로 적합화되어, 동시에 96 개 이하의 상이한 구축물로부터 125 ul AAV1 상청액의 생성을 가능하게 하였다. 35 ul 의 상청액을 이후 96-웰 포맷으로 배양된 1 차 뉴런의 1 개 웰을 형질도입하는데 사용할 수 있어, 단일 웰로부터의 생물학적 3 반복물에서의 형질도입을 가능하게 하였다.

[0390] 표 3. 모든 Taqman 프로브에 대한 제품 정보 (Life Technologies)

표적	종	프로브 #
<i>Ngn2</i>	마우스	Mm00437603_g1
<i>Grm5 (mGluR5)</i>	마우스	Mm00690332_m1
<i>Grm2 (mGluR2)</i>	마우스	Mm01235831_m1
<i>Grin2a (NMDAR2A)</i>	마우스	Mm00433802_m1
<i>GAPD (GAPDH)</i>	마우스	4352932E
<i>KLF4</i>	인간	Hs00358836_m1
<i>GAPD (GAPDH)</i>	인간	4352934E
<i>WPRE</i>	커스텀	
<i>5-HT1A</i>	마우스	Mm00434106_s1
<i>5-HT1B</i>	마우스	Mm00439377_s1
<i>5-HTT</i>	마우스	Mm00439391_m1
<i>Arc</i>	마우스	Mm00479619_g1
<i>BDNF</i>	마우스	Mm04230607_s1
<i>c-Fos</i>	마우스	Mm00487425_m1
<i>CBP/P300</i>	마우스	Mm01342452_m1
<i>CREB</i>	마우스	Mm00501607_m1
<i>CRHR1</i>	마우스	Mm00432670_m1
<i>DNMT1</i>	마우스	Mm01151063_m1
<i>DNMT3a</i>	마우스	Mm00432881_m1
<i>DNMT3b</i>	마우스	Mm01240113_m1
<i>egr-1 (zif-268)</i>	마우스	Mm00656724_m1
<i>Gad65</i>	마우스	Mm00484623_m1
<i>Gad67</i>	마우스	Mm00725661_s1
<i>GR (GCR, NR3C1)</i>	마우스	Mm00433832_m1
<i>HAT1</i>	마우스	Mm00509140_m1
<i>HCRTR1</i>	마우스	Mm01185776_m1
<i>HCRTR2</i>	마우스	Mm01179312_m1
<i>HDAC1</i>	마우스	Mm02391771_g1
<i>HDAC2</i>	마우스	Mm00515108_m1
<i>HDAC4</i>	마우스	Mm01299557_m1
<i>JMJD2A</i>	마우스	Mm00805000_m1
<i>M1 (CHRM1)</i>	마우스	Mm00432509_s1
<i>MCH-R1</i>	마우스	Mm00653044_m1
<i>NET (SLC6A2)</i>	마우스	Mm00436661_m1
<i>NR2B</i> 서브유닛	마우스	Mm00433820_m1
<i>OXTR</i>	마우스	Mm01182684_m1
<i>Scn1a</i>	마우스	Mm00450580_m1
<i>SIRT1</i>	마우스	Mm00490758_m1
<i>Tet1</i>	마우스	Mm01169087_m1
<i>Tet2</i>	마우스	Mm00524395_m1
<i>Tet3</i>	마우스	Mm00805756_m1

[0391]

[0392] 표 4. 이 연구에서 사용한 항체에 대한 클론, 제품 번호 및 농도

표적	숙주	클론 #	제조사	제품 #	동형	농도
mGluR2	마우스	mG2Na-s	Abcam	Ab15672	IgG	1:1000
α-튜블린	마우스	B-5-1-2	Sigma-Aldrich	T5168	IgG1	1:20000
NeuN	마우스	A60	Millipore	MAB377	IgG1	1:200
HA (Alexa Fluor 594)	마우스	6E2	Cell Signaling	3444	IgG1	1:100
GFP	닭	다클론	Aves Labs	GFP-1020	IgY	1:500

표적	숙주	접합체	제조사	제품 #	농도
마우스 IgG	염소	HRP	Sigma-Aldrich	A9917	1:5000-10000
마우스 IgG	염소	Alexa Fluor 594	Life Technologies	A11005	1:1000
닭 IgG	염소	Alexa Fluor 488	Life Technologies	A11039	1:1000

표적	숙주	에피토프	제조사	제품 #	동형	농도
H3K9me1	마우스	1-18	Millipore	17-680	IgG	2 μl/IP
H3K9me2	마우스	1-18	Millipore	17-681	IgG	4 μl/IP
H3K9Ac	토끼	다클론	Millipore	17-658	IgG	3 μg/IP
H4K20me1	토끼	15-24	Millipore	17-651	IgG	4 μg/IP
H4K8Ac	토끼	다클론	Millipore	17-10099	IgG	1.5 μl/IP
H4K20me3	토끼	18-22	Millipore	17-671	IgG	7 μl/IP
H3K27me3	토끼	다클론	Millipore	17-622	IgG	4 μg/IP

[0393]

[0394] 표 5. CHIP-qPCR 에 사용한 qPCR 프라이머

표적	프라이머
<i>Grm2</i> 프로모터	정방향 : CTGTGCTGAAGGATCTGGGG 역방향 : ATGCTGCAGGCATAGGACAA
<i>Neurog2</i> 프로모터	정방향 : GAGGGGGAGAGGGACTAAAGA 역방향 : GCTCTCCCTCCCCAGCTTA
<i>Myt-1</i> 프로모터 대조군	Cell Signaling Technologies SimpleChIP® 마우스 MYT-1 프로모터 프라이머 #8985
<i>RPL30</i> 인트론 2 대조군	Cell Signaling Technologies SimpleChIP® 마우스 RPL30 인트론 2 프라이머 #7015

[0395]

[0396] 표 6. TALE 에 의해 표적화된 게놈 서열

<i>5-HT1B</i>	TATCTGAACTCTCC
<i>5-HTT</i>	TGTCTGTCTTGCAT
<i>Arc</i>	TGGCTGTTGCCAGG
<i>BDNF</i>	TACCTGGAGCTAGC
<i>c-Fos</i>	TACACAGGATGTCC
<i>DNMT3a</i>	TTGGCCCTGTGCAG
<i>DNMT3b</i>	TAGCGCAGCGATCG
<i>gad65</i>	TATTGCCAAGAGAG
<i>gad67</i>	TGACTGGAACATAC
<i>GR (GCR, NR3C1)</i>	TGATGGACTGTAT
<i>HAT1</i>	TGGACCTTCTCCCT
<i>HCRTR1</i>	TAGGTCTCCTGGAG
<i>HCRTR2</i>	TGGCTCAGGAACTT
<i>HDAC1</i>	TTCTCTAAGCTGCC
<i>HDAC2</i>	TGAGCCCTGGAGGA
<i>HDAC4</i>	TGCCTAAGATGGAG
<i>JMJD2A</i>	TGTAGTGAGTGTC
<i>MCH-R1</i>	TGTCTAGGTGATGT
<i>NET</i>	TCTCTGCTAGAAGG
<i>Scn1a</i>	TCTAGGTCAAGTGT
<i>SIRT1</i>	TCCTCTGCTCCGCT
<i>tet1</i>	TCTAGGAGTGTAGC
<i>tet3</i>	TGCCTGGCTGCTGG
<i>5-HT1B</i>	TATCTGAACTCTCC
<i>Grm2</i>	TCAGAGCTGTCTC
<i>Grm5</i>	TGCAAGAGTAGGAG
<i>5-HT2A</i>	TAGTGACTGATTCC
<i>Grin2a</i>	TTGGAGGAGCACCA
<i>Neurog2</i>	TGAATGATGATAATAC

[0397]

[0398] 표 7. 마우스 변연계 아래 피질 (ILC) 에서의 Grm2 의 생체내 LITE-매개 활성화에 대한 바이러스 형질도입 및 광 자극 매개변수. 동측성 LITE-발현 반구에서의 Grm2 mRNA 수준을, 도 39J 에서 나타낸 모든 3 가지 실험 조건에 대해 대측성 mCherry-발현 대조군 반구와 비교한다.

실험 조건	ILC 반구체 (동측성)	ILC 반구체 (대측성)
	AAV 벡터	AAV 벡터
	광	
	자극	
GFP	GFP	mCherry
LITE / 무광	TALE-CIB1::CRY2PHR-VP64	mCherry
LITE / + 광	TALE-CIB1::CRY2PHR-VP64	mCherry

[0399]

[0400] 표 8. HDAC 영입 효과기 도메인

하위유형/복합체	명칭	기질 (공지된 경우)	변형 (공지된 경우)	유기체	전체 크기 (aa)	선택된 절두 (aa)	최종 크기 (aa)	촉매적 도메인
Sin3a	MeCP2	-	-	짐취 (R. norvegicus)	492	207-492 (Nan)	286	-
Sin3a	MBD2b	-	-	H. 사피엔스	262	45-262 (Boeke)	218	-
Sin3a	Sin3a	-	-	H. 사피엔스	1273	524-851 (Laherty)	328	627-829; HDAC1 상호작용
NcoR	NcoR	-	-	H. 사피엔스	2440	420-488 (Zhang)	69	-
NuRD	SALL1	-	-	마우스 (M. musculus)	1322	1-93 (Lauberth)	93	-
CoREST	RCOR1	-	-	H. 사피엔스	482	81-300 (Gu, Ouyang)	220	-

[0401]

[0402] Nan, X. et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393, 386-389 (1998).

[0403] Boeke, J., Ammerpohl, O., Kegel, S., Moehren, U. & Renkawitz, R. The minimal repression domain of MBD2b overlaps with the methyl-CpG-binding domain and binds directly to Sin3A. *Journal of Biological Chemistry* 275, 34963-34967 (2000).

[0404] Laherty, C. D. et al. Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. *Cell* 89, 349-356 (1997).

[0405] Zhang, J., Kalkum, M., Chait, B. T. & Roeder, R. G. The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Molecular cell* 9, 611-623 (2002).

[0406] Lauberth, S. M. & Rauchman, M. A conserved 12-amino acid motif in Sall1 recruits the nucleosome remodeling and deacetylase corepressor complex. *Journal of Biological Chemistry* 281, 23922-23931 (2006).

[0407] Gu, H. & Roizman, B. Herpes simplex virus-infected cell protein 0 blocks the silencing of viral DNA by dissociating histone deacetylases from the CoREST, AiREST complex.

[0408] Ouyang, J., Shi, Y, Valin, A., Xuan, Y & Gill, G. Direct binding of CoREST1 to SUMO-2/3 contributes to gene-specific repression by the LSD1/CoREST1/HDAC complex. *Molecular cell* 34, 145-154 (2009)

[0409] 표 9. HDAC 효과기 도메인

하위유형/ 복합체	명칭	기질 (공지된 경우)	변형 (공지된 경우)	유기체	전체 크기 (aa)	선택된 절두 (aa)	최종 크기 (aa)	촉매적 도메인
HDAC I	<b>HDAC 8</b>	-	-	아프리카 발톱 개구리 ( <i>X. laevis</i> )	325	1-325	325	1-272: HDAC
HDAC I	<b>RPD3</b>	-	-	S. 세레비지에	433	19-340	322 (Vannier)	19-331: HDAC
HDAC IV	<b>MesoL04</b>	-	-	M. 로티	300	1-300 (Gregorette)	300	-
HDAC IV	<b>HDAC 11</b>	-	-	H. 사피엔스	347	1-347 (Gao)	347	14-326: HDAC
HD2	<b>HDT1</b>	-	-	A. 탈리아나	245	1-211 (Wu)	211	-
SIRT I	<b>SIRT3</b>	H3K9Ac H4K16Ac H3K56Ac	-	H. 사피엔스	399	143-399 (Scher)	257	126-382: SIRT
SIRT I	<b>HST2</b>	-	-	C. 알비칸스	331	1-331 (Hnisz)	331	-
SIRT I	<b>CobB</b>	-	-	대장균 ( <i>K12</i> )	242	1-242 (Landry)	242	-
SIRT I	<b>HST2</b>	-	-	S. 세레비지에	357	8-298 (Wilson)	291	-
SIRT III	<b>SIRT5</b>	H4K8Ac H4K16Ac	-	H. 사피엔스	310	37-310 (Gertz)	274	41-309: SIRT
SIRT III	<b>Sir2A</b>	-	-	P. 팔시파룸	273	1-273 (Zhu)	273	19-273: SIRT
SIRT IV	<b>SIRT6</b>	H3K9Ac H3K56Ac	-	H. 사피엔스	355	1-289 (Tennen)	289	35-274: SIRT

[0410]

[0411] Vannier, D., Balderes, D. & Shore, D. Evidence that the transcriptional regulators SIN3 and RPD3, and a novel gene (SDS3) with similar functions, are involved in transcriptional silencing in *S. cerevisiae*. *Genetics* 144, 1343-1353 (1996).

[0412] Gregorette, I., Lee, Y.-M. & Goodson, H. V. Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. *Journal of molecular biology* 338,17-31 (2004).

[0413] Gao, L., Cueto, M. A., Asselbergs, F. & Atadja, P. Cloning and functional characterization of HDAC11, a novel member of the human histone deacetylase family. *Journal of Biological Chemistry* 277,25748-25755 (2002).

[0414] Wu, K., Tian, L., Malik, K., Brown, D. & Miki, B. Functional analysis of HD2 histone deacetylase homologues in *Arabidopsis thaliana*. *The Plant Journal* 22, 19-27 (2000).

[0415] Scher, M. B., Vaquero, A. & Reinberg, D. SirT3 is a nuclear NAD<sup>+</sup>-dependent histone deacetylase that translocates to the mitochondria upon cellular stress. *Genes & development* 21,920-928 (2007).

[0416] Hnisz, D., Schwarzmuller, T. & Kuchler, K. Transcriptional loops meet chromatin: a dual, Aelayer network controls white,Aiopaque switching in *Candida albicans*. *Molecular microbiology* 74, 1-15 (2009).

- [0417] Landry, J. et al. The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases. *Proceedings of the National Academy of Sciences* 97, 5807-5811 (2000).
- [0418] Wilson, J. M., Le, V. Q., Zimmerman, c., Marmorstein, R. & Pillus, L. Nuclear export modulates the cytoplasmic Sir2 homologue Hst2. *EMBO reports* 7,1247-1251 (2006).
- [0419] Gertz, M. & Steegborn, C. Function and regulation of the mitochondrial Sirtuin isoform Sirt5 in Mammalia. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1804, 1658-1665 (2010).
- [0420] Zhu, A. Y. et al. Plasmodium falciparum Sir2A preferentially hydrolyzes medium and long chain fatty acyl lysine. *ACS chemical biology* 7, 155-159 (2011).
- [0421] Tennen, R. I., Berber, E. & Chua, K. F. Functional dissection of SIRT6: identification of domains that regulate histone deacetylase activity and chromatin localization. *Mechanisms of ageing and development* 131, 185-192 (2010).

[0422] 표 10. 히스톤 메틸트랜스퍼라아제 (HMT) 효과기 도메인

하위유형/복합체	명칭	기질 (공지된 경우)	변형 (공지된 경우)	유기체	전체 크기(aa)	선택된 절두 (aa)	최종 크기 (aa)	촉매적 도메인
SET	NUE	H2B, H3, H4	-	C. 트라초마티스	219	1-219 (Pennini)	219	-
SET	vSET	-	H3K27me3	P. 부르사리아 클로벨라 바이러스	119	1-119 (Mujtaba)	119	4-112: SET2
SUV39 패밀리	EHMT 2/G9A	H1.4K2, H3K9, H3K27	H3K9me1/2, H1K25me1	마우스 (M. musculus)	1263	969-1263 (Tachibana)	295	1025-1233: preSET, SET, postSET
SUV39	SUV39 H1	-	H3K9me2/3	H. 사피엔스	412	79-412 (Snowden)	334	172-412: preSET, SET, postSET
Suvar3-9	dim-5	-	H3K9me3	N. 크라사	331	1-331 (Rathert)	331	77-331: preSET, SET, postSET
Suvar3-9 (SUVH 서브패밀리)	KYP	-	H3K9me1/2	A. 탈리아나	624	335-601	267 (Jackson)	-
Suvar3-9 (SUVR 서브패밀리)	SUVR4	H3K9me1	H3K9me2/3	A. 탈리아나	492	180-492	313 (Thorstensen)	192-462: preSET, SET, postSET
Suvar4-20	SET4	-	H4K20me3	C. 엘레간스	288	1-288 (Vielle)	288	-
SET8	SET1	-	H4K20me1	C. 엘레간스	242	1-242 (Vielle)	242	-
SET8	SETD8	-	H4K20me1	H. 사피엔스	393	185-393	209 (Couture)	256-382: SET
SET8	TgSET 8	-	H4K20me1/2/3	T. 곤디	1893	1590-1893 (Sautel)	304	1749-1884: SET

- [0423]
- [0424] Pennini, M. E., Perrinet, S. p., Dautry-Varsat, A. & Subtil, A. Histone methylation by NUE, a novel nuclear effector of the intracellular pathogen Chlamydia trachomatis. *PLoS pathogens* 6, e1000995 (2010).
- [0425] Mujtaba, S. et al. Epigenetic transcriptional repression of cellular genes by a viral SET protein. *Nature cell biology* 10, 1114-1122 (2008).
- [0426] Tachibana, M., Matsumura, Y, Fukuda, M., Kimura, H. & Shinkai, Y G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *The EMBO journal* 27, 2681-2690 (2008).
- [0427] Snowden, A. W., Gregory, P. D., Case, C. C. & Pabo, C. O. Gene-specific targeting of H3K9 methylation is sufficient for initiating repression in vivo. *Current biology* 12,2159-2166 (2002).
- [0428] Rathert, P., Zhang, X., Freund, C., Cheng, X. & Jeltsch, A. Analysis of the substrate specificity of the Dim-5 histone lysine 메틸transferase using peptide arrays. *Chemistry & biology* 15,5-11 (2008).
- [0429] Jackson, J. P. et al. Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana. *Chromosoma* 112,308-315 (2004).
- [0430] Thorstensen, T. et al. The Arabidopsis SUVR4 protein is a nucleolar histone methyltransferase with preference for monomethylated H3K9. *Nucleic acids research* 34, 5461-5470 (2006).

[0431] Vielle, A. et al. H4K20me1 Contributes to Downregulation of X-Linked Genes for C. elegans Dosage Compensation. PLoS Genetics 8, e1002933 (2012).

[0432] Couture, J.-F., Collazo, E., Brunzelle, J. S. & Trievel, R. C. Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. Genes & development 19, 1455-1465 (2005).

[0433] Sautel, C. I. F. et al. SET8-mediated methylations of histone H4 lysine 20 mark silent heterochromatic domains in apicomplexan genomes. Molecular and cellular biology 27, 5711-5724 (2007).

[0434] 표 11. 히스톤 메틸트랜스퍼라아제 (HMT) 영입 효과기 도메인

하위유형/복합체	명칭	기질 (공지된 경우)	변형 (공지된 경우)	유기체	전체 크기(aa)	선택된 절두 (aa)	최종 크기 (aa)	촉매적 도메인
-	<b>Hp1a</b>	-	H3K9me3	마우스 (M. musculus)	191	73-191	119 (Hathaway)	121-179: chromoshadow
-	<b>PHF19</b>	-	H3K27me3	H. 사피엔스	580	(1-250) + GGSG 링커 + (500-580)	335 (Ballaré)	163-250: PHD2
-	<b>NIPPI</b>	-	H3K27me3	H. 사피엔스	351	1-329 (Jin)	329	310-329: EED

[0435]

[0436] Hathaway, N. A. et al. Dynamics and memory of heterochromatin in living cells. Cell (2012).

[0437] Ballaré, C. et al. Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. Nature structural & molecular biology 19, 1257-1265 (2012).

[0438] Jin, Q. et al. The protein phosphatase-1 (PP1) regulator, nuclear inhibitor of PP1 (NIPPI), interacts with the polycomb group protein, embryonic ectoderm development (EED), and functions as a transcriptional repressor. Journal of Biological Chemistry 278, 30677-30685 (2003).

[0439] 표 12. 히스톤 아세틸트랜스퍼라아제 저해제 효과기 도메인

하위유형/복합체	명칭	기질 (공지된 경우)	변형 (공지된 경우)	유기체	전체 크기(aa)	선택된 절두 (aa)	최종 크기 (aa)	촉매적 도메인
-	<b>SET/TAF-1β</b>	-	-	마우스 (M. musculus)	289	1-289 (Cervoni)	289	-

[0440]

[0441] Cervoni, N., Detich, N., Seo, S.-B., Chakravarti, D. & Szyf, M. The oncoprotein Set/TAF-1beta, an inhibitor of histone acetyltransferase, inhibits active demethylation of DNA, integrating DNA methylation and transcriptional silencing. Journal of Biological Chemistry 277,25026-25031 (2002)

[0442] 보충적 서열

[0443] > TALE(Ngn2)-NLS-CRY2

MSRTRLSPPPAPSPAFSADSFSDLRQFDPSLFNTSLFDSLPPFGAHHTEAATG  
 EWDEVQSGLRADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGY  
 SQQQKEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALP  
 EATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARKGGVTAVEAVH

[0444]

AWRNALTGAPLNLTPQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASN  
 GGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLT  
 PEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQA  
 LETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAI  
 ASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQR  
 LLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGK  
 QALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQV  
 VAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVAQLSRPDPA  
 LAALTNHDLVALACLGRRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVADHAQVVR  
 VLGFFQCHSHAQAFDDAMTQFGMSRHLLQLFRRVGVTELEARSGLTPPASQRWDRI  
 LQASGMKRAKPSPTSTQTPDQASLHAFADSLERDL DAPSPMHEGDQTRAS<sup>ASPKKKKR</sup>  
 VEASKMDKKTIVWFRRLRIEDNPALAAAAHEGSVFPVFIWCPEEEGQFYPPGRASRWW  
 MKQSLAHLSQLKALGSDTLTIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDHTV  
 KEKLVERGISVQSYNGDLLYEPWEIYCEKGPFTSFNSYWKKCLDMSIESVMLPPPWRL  
 MPITAAAEAIWACSI EELGLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDYAK  
 NSKKVVGNSTSLLSPYLHFG EISVRHV FQCARMKQIIWARDKNSEGEESADLFLRGIGLR  
 EYSRYICFNFPFTHEQSLLSHLRFFPWDADVDKFKAWRQGRGTGYPLVDAGMRELWATG  
 WMHNRIRVIVSSFAVKFLLLPWKWGMKYFWDTL DADLECDILGWQYISGSIPDGHEL  
 DRLDNPALQGAKYDPEGEYIRQWLP ELARLPTEWIIHPWDAPLTVLKASGVELGTNYA  
 KPIVDIDTARELLAKAISRTREAQIMIGAAPDEIVADSFEALGANTIKEPGLCPSVSSNDQ  
 QVPSAVRYNGSKRVKPEEEEEERDMKKS RGFDERELFSTAESSSSSVFFVSQCSLASEG  
 KNLEGIQDSSDQITTS LGKNG

[0445]

[0446]

>TALE(Ngn2)-NLS-CRY2PHR

MSRTRLPSPPAPSPAFSADSFSDLRQFDP SLFNTSLFDSLPPFGAHTTEAATG  
 EWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGY  
 SQQQKEKIKPVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTAVVKYQDMIAALP  
 EATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARKGGVTAVEAVH  
 AWRNALTGAPLNLTPQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASN  
 GGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLT

[0447]

PEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 LETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAI  
 ASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQR  
 LLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGK  
 QALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQV  
 VAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVAQLSRPDPA  
 LAALTNHDLVALACLGGPALDAVKKGLPHAPALIKRTNRRIPERTSHRVADHAQVVR  
 VLGFFQCHSHAQAFDDAMTQFGMSRHGLLQLFRRVGVTELEARSGLTPPASQRWDRI  
 LQASGMKRAKPSPTSTQTPDQASLHAFADSLERDLDPSPMHEGDQTRASASPKKKRK  
 VEASKMDKKTIVWFRRLDRIEDNPALAAAAHEGSVFPVFIWCPEEGQFYPGRASRWW  
 MKQSLAHLSQLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDHTV  
 KEKLVVERGISVQSYNGDLLYEPWEIYCEKPKPFTSFNSYWKCLDMSIESVMLPPPWRL  
 MPITAAAEAIWACSEELGLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDYAK  
 NSKKVVGNSTSLLSPYLHFGEISVRHVFCARMKQIHWARKNSEGEESADLFLRGIGLR  
 EYSRYICFNFPFTHEQSLLSHLRFFPDADVDKFKAWRQGRGTGYPLVDAGMRELWATG  
 WMHNRIRVIVSSFAVKFLLLPWKWGMKYFWDTLDDADLECDILGWQYISGSIPDGHEL  
 DRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIIHPWDAPLTVLKASGVELGTNYA  
 KPJVDDIDTARELLAKAISRTREAQIMIGAAP

[0448]

[0449] > CIB1-NLS-VP64\_2A\_GFP

MNGAIGGDLLLNFPMDSVLERQRAHLKYLNPFDSPLAGFFADSSMITGGEM  
 DSYLSTAGLNLPMMYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKMKMT  
 MNRDDLVEEGEEESKITEQNNGSTKSIKMKHKAKKEENNFSNDSSKVTKELEKTDYI  
 HVRARRGQATDSHSIAERVRREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQ  
 RQIEFLSMKLAIVNPRPDFDMDDIFAKEVASTPMTVVPSPPEMVLGYSHEMVHSGYSSE  
 MVNSGYLHVNPMQVNTSSDPLSCFNNGEAPSMWDSHVQNLVGNLGVASPKKKRKY  
 EASGSGRADALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLDM  
 LINSRGSGEGRGSLTTCGDVEENPGPVSKGEELFTGVVPILVELDGDVNGHKFVS  
 GEGEDATYGKLTCLKFICTTGKLPVPWPTLVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGY  
 VQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKIDFKEDGNILGHKLEYNYNSHNVYI

[0450]

MADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS  
 KDPNEKRDHMLLEFVTAAGITLGMDELYK

[0451]

[0452] > CIBN-NLS-VP64\_2A\_GFP

MNGAIGGDLLLNFPMDSVLERQRAHLKYLNPFDSPLAGFFADSSMITGGEM  
 DSYLSTAGLNLPMMYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKMKMT  
 MNRDDLVEEGEEESKITEQNNGSTKSIKMKHKAKKEENNFSNDSSKVTKELEKTDYI  
 ASPKKRKYVEASGSGRADALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSD  
 ALDDFDLMLINSRGSGEGRGSLTTCGDVEENPGPVSKGEELFTGVVPILVELDGDVNG  
 HKFVS  
 GEGEDATYGKLTCLKFICTTGKLPVPWPTLVTTLTYGVCFSRYPDHMKQHDF  
 FKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKIDFKEDGNILGHKLE  
 YNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH  
 YLSTQSALS  
 KDPNEKRDHMLLEFVTAAGITLGMDELYK

[0453]

[0454] > CIB1-NLS-VP16\_2A\_GFP

MNGAIGGDLLLLNFPDMSVLERQRAHLKYLNPFDSPLAGFFADSSMITGGEM  
 DSYLSTAGLNLPMMYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKKKKMT  
 MNRDDLVEEGEEEEKSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYI  
 HVRARRGQATDSHSIAERVRREKISERMKFLQDLVPGCDKITGKAGMLDEHINYVQSLQ  
 RQIEFLSMKLAIVNPRPDFDMDDIFAKEVASTPMTVVPSPEMVLVSGYSHEMVHSGYSSE  
 MVNSGYLHVNPMQVNTSSDPLSCFNNGEAPSMWDSHVQNLVGNLGVASPKKKRKKV  
 EASAPPTDVSLGDELHLDGEDVAMAHADALDDFDLMLGDGDSPPGPGFTPHDSAPYG  
 ALDMADFEFEQMFTDALGIDEYGGFEPGIRRSRGSGEGRSLLTCGDVEENPGPVSKGE  
 ELFTGVVPIVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTLT  
 YGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFECDTLVNRI  
 ELKGIDFKEDGNILGHKLEYNYNHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHY  
 QQNTPIGDGPVLLPDNHYLSTQSAISKDPNEKRDMVLLLEFVTAAGITLGMDELYK

[0455]

[0456] > CIB1-NLS-p65\_2A\_GFP

MNGAIGGDLLLLNFPDMSVLERQRAHLKYLNPFDSPLAGFFADSSMITGGEM  
 DSYLSTAGLNLPMMYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKKKKMT  
 MNRDDLVEEGEEEEKSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYI  
 HVRARRGQATDSHSIAERVRREKISERMKFLQDLVPGCDKITGKAGMLDEHINYVQSLQ  
 RQIEFLSMKLAIVNPRPDFDMDDIFAKEVASTPMTVVPSPEMVLVSGYSHEMVHSGYSSE

[0457]

MVNSGYLHVNPMQVNTSSDPLSCFNNGEAPSMWDSHVQNLVGNLGVASPKKKRKKV  
 EASPSGQISNQALALAPSSAPVLAQTMVPSSAMVPLAQPPAPAPVLTGPPQSLAPVVK  
 STQAGEGTLSEALLHLQFDAQEDL GALLGNSTDPGVFTDLASVDNSEFQQLNQGVSMS  
 HSTAEPMLMEYPAITRLVTGSQRPPDPAPTPLGTSGLPNGLSGDEDFSSIADMDFSALLS  
 QISSGQSRGSGEGRSLLTCGDVEENPGPVSKGEELFTGVVPIVELDGDVNGHKFSVS  
 GEGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTLT YGVQCFSRYPDHMKQHDFFKSAM  
 PEGYVQERTIFFKDDGNYKTRAEVKFECDTLVNRIELKGIDFKEDGNILGHKLEYNYNH  
 NVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSA  
 LSKDPNEKRDMVLLLEFVTAAGITLGMDELYK

[0458]

[0459] > HA-TALE(12mer)-NLS-VP64\_2A\_GFP

MYPYDVDPDYAVDLRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGFTHAHI  
 VALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGP  
 PLQDGTGQLLKIARRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASXXGGKQALET  
 VQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASX  
 XGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLT  
 PEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLP  
 VLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVV  
 AIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQ  
 AHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGRPALES  
 I VAQLSRPDPALAALNDHLVALACLGRPALDAVKKGLPHAPALIKRTNRRIPERTSHR  
 VAASPKKKRKKVEASGSGRADALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLG  
 SDALDDFDLMLINSRGSGEGRSLLTCGDVEENPGPVSKGEELFTGVVPIVELDGDV  
 NGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTLT YGVQCFSRYPDHMKQ  
 HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFECDTLVNRIELKGIDFKEDGNILGH  
 KLEYNYNHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDN  
 HYLSTQSAISKDPNEKRDMVLLLEFVTAAGITLGMDELYK

[0460]

[0461] > HA-TALE(12mer)-NLS-SID4X\_2A\_phiLOV2.1

MYPYDVPDYAVDLRRTLGYSSQQQEKIKPKVVRSTVAQHHEALVGHGFTHAHI  
VALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGP  
PLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPQVVAIASXXGGKQALET

[0462] VQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASX  
XGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLT  
PEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLP  
VLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQ  
ALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQV  
AIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQ  
AHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALET  
VAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHR  
VAASPKKKRKRKVEASPKKKRKRKVEASGSGMNIQMLLEAADYLERREREAEHGYASMLPG  
SGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHG  
YASMLPGSGMNIQMLLEAADYLERREREAEHGYASMLPSRSGSGEGRGSLLTCGDVE  
ENPGPIEKSFVITDPRLPDYPIIFASDGFLELTEYSREEIMGRNARFLQGPETDQATVQKIR  
DAIRDQRETTVQLINITYKSGKKFWNLHLQPVDRDRKGGGLQYFIGVQLVGSDDHV

[0463]

[0464] > HA-TALE(12mer)-NLS-CIB1

MYPYDVPDYAVDLRRTLGYSSQQQEKIKPKVVRSTVAQHHEALVGHGFTHAHI  
VALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGP  
PLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPQVVAIASXXGGKQALET  
VQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASX  
XGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLT  
PEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLP  
VLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQ  
ALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQV  
AIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQ  
AHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALET  
VAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHR  
VAASPKKKRKRKVEASNGAIGGDLNLPDMSVLERQRAHLKYLNPFDSPLAGFFADSSM  
ITGGEMDSYLSLAGLNLPMMYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNE  
KKKKMTMNRDDLVEEGEEKSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSKVTKE  
LEKTDYIHVRARRGQATDSSHIAERVREKISERMKFLQDLVPGCDKITGKAGMLDEIIN  
YVQSLQRQIEFLSMKLAIVNRPDFDMDIFAKEVASTPMTVVPSPPEMVLVSGYSEMVH  
SGYSEMVNISGYLHVNPMPQQVNTSSDPLSCFNNGEAPSMWDSHVQNLVGNLGV

[0465]

[0466] > CRY2PHR-NLS-VP64\_2A\_GFP

MKMDKKTIVWFRDLRIEDNPALAAAHEGVSFVPVFIWCPEEEGQFYPRGRAS  
 RWWMKQSLAHLSSQLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVR  
 DHTVKEKLVVERGISVQSYNGDLLYEPWEIYCEKGPFTSFNSYWKKCLDMSIESVMLPP  
 PWRLMPITAAAEAIWACSIIEELGLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLI  
 DYAKNSKKVVGNSTSLSPYLHFGEISVRHVVFQCARMKQIIWARDKNSEGEESADLFLR  
 GIGLREYSRYICFNFPFTHQSLLSHLRFFPWDADVDKFKAWRQGRGTGYPLVDAGMREL  
 WATGWMHNRIRVIVSSFAVKFLLLPWKWGMKYFWDTLDDADLECDILGWQYISGSIPD  
 GHELDRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIIHPWDAPLTVLKASGVELG  
 TNYAKPIVDIDTARELLAKAISRTREAQIMIGAAPASPKKKRKRVEASGSGRADALDDFDL  
 DMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLINSRGSGEGRGSLTTC  
 GDVEENPGPVSKGEELFTGVVPIVVELDGDVNGHKFVSVEGEGDATYGKLTCLKFICTT  
 GKLPVPWPTLVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT  
 RAEVKFEGDTLVNRIELKIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI  
 RHNIEDGSVQLADHYQNTPIGDGPVLLPDNHVYLSQSALS KDPNEKRDMVLLLEFVTA  
 AGITLGMDELYKV

[0467]

[0468] > CRY2PHR-NLS-SID4X\_2A\_phiLOV2.1

MKMDKKTIVWFRDLRIEDNPALAAAHEGVSFVPVFIWCPEEEGQFYPRGRAS  
 RWWMKQSLAHLSSQLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVR  
 DHTVKEKLVVERGISVQSYNGDLLYEPWEIYCEKGPFTSFNSYWKKCLDMSIESVMLPP  
 PWRLMPITAAAEAIWACSIIEELGLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLI  
 DYAKNSKKVVGNSTSLSPYLHFGEISVRHVVFQCARMKQIIWARDKNSEGEESADLFLR  
 GIGLREYSRYICFNFPFTHQSLLSHLRFFPWDADVDKFKAWRQGRGTGYPLVDAGMREL  
 WATGWMHNRIRVIVSSFAVKFLLLPWKWGMKYFWDTLDDADLECDILGWQYISGSIPD  
 GHELDRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIIHPWDAPLTVLKASGVELG  
 TNYAKPIVDIDTARELLAKAISRTREAQIMIGAAPASPKKKRKRVEASGSGMNIQMLLEAA  
 DYLERREREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQ  
 MLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASMLP  
 SRSRGSGEGRGSLTTCGDVEENPGPIEKSFVITDPRLPDYPIHFASDGFLELTEYSREEIMG  
 RNARFLQGPETDQATVQKIRDAIRDQRETTVQLINYTKSGKKFWNLLHLQPVDRKGGGL  
 QYFIGVQLVGS DHV

[0469]

[0470] > TALE(KLF4)-NLS-CRY2PHR

MSRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATG  
 EWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGY  
 SQQQQEKIKPKVRSVVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALP  
 EATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVH  
 AWRNALTGAPLNLTPQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASH  
 DGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLT  
 PEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLP  
 VLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQV  
 AIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQA  
 HGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQ  
 RLLPVLCQAHGLTPEQVVAIASHDGGRPALLESIVAQLSRPDPALAAALTNDHLVALACLG  
 GRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVADHAQVVRVLGFFQCHSHPAQAFDD  
 AMTQFGMSRHGLLQLFRRVGVTELEARSGLTLPASQRWDRILQASGMKRAKPSPTSTQ  
 TPDQASLHAFADSLERDLDAVSPMHEGDQTRASASPKKKRKEASKMDKKTIVWFRRD  
 LRIEDNPALAAAAHEGSVFPVFIWCPEEEGQFYPGRASRWWWKQSLAHLSSQLKALGS  
 DLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDHTVKEKLVVERGISVQSYNGD  
 LLYEPWEIYCEKGPFTSFNSYWKKCLDMSIESVMLPPPWRMPITAAAEAIWACSIEEL  
 GLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDYAKNSKKVVGNSTSLSPYL  
 HFGEISVRHVFCARMKQIHWARKNSEGEESADLFLRGIGLREYSRYICFNFPFTHEQSL  
 LSHLRFFPWDADVDKFKAWRQGRGTGYPLVDAGMRELWATGWMHNRIRVIVSSFAVKF  
 LLLPWKWMKMYFDWTLDDADLECDILGWQYISGSIPDGHELDRLDNPALQGAKYDPEG  
 EYIRQWLPELARLPTEWIIHPWDAPLTVLKASGVELGTNYAKPIVDIDTARELLAKAISR  
 TREAQIMIGAAP

[0471]

[0472] > HA-NLS-TALE(p11, N136)-SID

MYPYDVPDYASPKKKRKEASVLDLRTLGYSSQQQEKIKPKVRSVVAQHHEA  
 LVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEA  
 LLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPQVVAIA  
 SNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG  
 LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQR

[0473]

LLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGG  
 KQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQ  
 VVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVL  
 CQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHG  
 LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQR  
 LLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGG  
 KQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALLESIVAQLSRPDPALAAALTNDH  
 LVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVADHAQVVRVLGFFQCHS  
 HPAQAFDDAMTQFGMSRHGLLQLFRRVGVTELEARSGLTLPASQRWDRILQASGMKRA  
 KPSPTSTQTPDQASLHAFADSLERDLDAVSPMHEGDQTRASASGSGMNIQMLLEAADYL  
 ERREREAEHGYASMLP.

[0474]

[0475]

> HA-NLS-TALE(p11, N136)-SID4X

MYPYDVPDYASPKKKRKEASVLDLRTLGYSSQQQEKIKPKVRSSTVAQHHEA  
 LVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEA  
 LLTVAGELRGPPLQLDTGQLLKIARKGGVTAVEAVHAWRNALTGAPLNLTPQVVAIA  
 SNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG  
 LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQR  
 LLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGG  
 KQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQ  
 VVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHG  
 LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQR  
 LLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGG  
 KQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAAALTNDH  
 LVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVADHAQVVRVLGFFQCHS  
 HPAQAFDDAMTQFGMSRHGLLQLFRRVGVTELEARSGLPPASQRWDRILQASGMKRA  
 KPSPTSTQTPDQASLHAFADSLERDLAPSPMHEGDQTRASASGSGMNIQMLLEAADYL

[0476]

ERREREAHEGYASMLPGSGMNIQMLLEAADYLERREREAHEGYASMLPGSGMNIQML  
 LEAADYLERREREAHEGYASMLPGSGMNIQMLLEAADYLERREREAHEGYASMLPSR

[0477]

[0478]

>HA-TALE(ng2, C63)-GS-cib1-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSLRAADAPPPPTMRVAVTAARPPRAKPAARRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARKG  
 GVTAVEAVHAWRNALTGAPLNLTPQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASGGGGSGGGGSNGAIGDLLLLNFPDMSVLERQR  
 AHLKYLNPFTDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRLSISPE  
 TTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEESKITEQNGSTKSIKK  
 MKHKAKKEENFNSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRRREKISERMKF  
 LQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVAS  
 TPMTVVPSPEMVLGSGYSHEMVHSGYSSEMVSNGYLHVNPMQQVNTSSDPLSCFNNGEA  
 PSMWDSHVQONLYGNLGV

[0479]

[0480] > HA-TALE(ng2, C63)-wNES-cib1-mutNLS  
 YPYDVPDYASRTRLSPSPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRADAPPPTMRVAVTAARPPRAKAPRRRAAQPSDASPA  
 QVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV

[0481] LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASLYPERLRRILTNGAIGDLLLLNFPDMSVLERQRA  
 HLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRLSISPET  
 TLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEKSKITEQNNGSTKSIKKM  
 KHKAKKEENNFSNDSKVTKELEKTDYIHVRARRGQATDSHSAERVREKISERMKFL  
 QDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFMDDIFAKEVAST  
 PMTVVPSPEMVLSGYSHEMVHSGYSSEMVN<sup>SGYL</sup>HVNPMQQVNTSSDPLSCFNNGEAP  
 SMWDSHVQNL<sup>YGNL</sup>GV

[0482]  
 [0483] > HA-TALE(ng2, C63)-mNES-cib1-mutNLS

YPYDVPDYASRTRLSPSPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRADAPPPTMRVAVTAARPPRAKAPRRRAAQPSDASPA  
 QVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASLQLPPLERLTLNGAIGDLLLLNFPDMSVLERQR

[0484] AHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRLSISPE  
 TTLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEKSKITEQNNGSTKSIKK  
 MKHKAKKEENNFSNDSKVTKELEKTDYIHVRARRGQATDSHSAERVREKISERMKF  
 LQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFMDDIFAKEVAS  
 TPMTVVPSPEMVL<sup>SGYS</sup>HEMVHSGYSSEMVN<sup>SGYL</sup>HVNPMQQVNTSSDPLSCFNNGEA  
 PSMWDSHVQNL<sup>YGNL</sup>GV

[0485]

[0486] > HA-TALE(ng2, C63)-ptk2NES-cib1-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASLDLASLJLNGAIGDLLLLNFPDMSVLERQRAHL  
 KYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRSLISPETTL  
 GTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEKSKITEQNNGSTKSIKMKH  
 KAKKEENFNDSKSVTKELEKTDYIHVRARRGQATDSHSIAERVREKISERMKFLQD  
 LVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVASTPM  
 TVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSM  
 WDSHVQONLYGNLGV

[0487]

[0488] > HA-TALE(ng2, C63)-mapkNES-cib1-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAA

[0489]

QVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASLQKKLEEELELNGAIGDLLLLNFPDMSVLERQRA  
 HLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRSLISPET  
 TLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEKSKITEQNNGSTKSIKMK  
 KHKAKKEENFNDSKSVTKELEKTDYIHVRARRGQATDSHSIAERVREKISERMKFL  
 QDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVAST  
 PMTVVPSPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGEAP  
 SMWDSHVQONLYGNLGV

[0490]

[0491] > HA-TALE(ng2, C63)-GS-cib1Δ3-mutNLS

YPYDVPDYASRTRLSPSPAPSPAFSADSFSDLLRQFDPSTLNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRADAPPPTMRVAVTAARPPRAKPAARRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ

[0492] ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASGGGGSGGGGSNGAIGDLLLLNFPDMSVLERQR  
 AHLKYLNPFDSPLAGFFADSSMITGGEMDSYLSLAGLNLPMMYGETTVEGDSRLSISPE  
 TTLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEKSKITEQ<sup>NN</sup>NGSTKSIKK  
 MKHKAKKEENNFSNDSSKVTKLEKTDYIHVRARRGQATDSHSAERVRREKISERMKF  
 LQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVAS  
 TPMTVVPSPEMVLSGYSHEMVHSGYSSEM<sup>VNS</sup>GYLHVNP<sup>MQQ</sup>VNTSS

[0493]  
 [0494] > HA-TALE(ng2, C63)-wNLS-cib1Δ3-mutNLS

YPYDVPDYASRTRLSPSPAPSPAFSADSFSDLLRQFDPSTLNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRADAPPPTMRVAVTAARPPRAKPAARRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASLYPERLRRIL<sup>TNG</sup>AIGDLLLLNFPDMSVLERQRA  
 HLKYLNPFDSPLAGFFADSSMITGGEMDSYLSLAGLNLPMMYGETTVEGDSRLSISPET  
 TLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEKSKITEQ<sup>NN</sup>NGSTKSIKKM  
 KHKAKKEENNFSNDSSKVTKLEKTDYIHVRARRGQATDSHSAERVRREKISERMKFL  
 QDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVAST  
 PMTVVPSPEMVLSGYSHEMVHSGYSSEM<sup>VNS</sup>GYLHVNP<sup>MQQ</sup>VNTSS

[0495]

[0496] > HA-TALE(ng2, C63)-mNLS-cib1△3-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPPPTMRVAVTAARPPRAKPAARRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARKG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASSPKKKRKRVEASNGAIGGDLLLNFPMDSVLERQR  
 AHLKYLNPFTDSPLAGFFADSSMITGGEMDSYLSLAGLNLPMMYGETTVEGDSRSLISPE  
 TTLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEEEKSKITEQNNGSTKSIKK  
 MKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKISERMKF  
 LQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVAS  
 TPMTVVPSPPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSS

[0497]

[0498] > HA-TALE(ng2, C63)-GS-cib1-mutNLS-mutbHLH

YPYDVPDYASRTRLPSPPAPSPAFAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPPPTMRVAVTAARPPRAKPAARRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARKG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL

[0499]

PVLCQAHGLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASGGGGSGGGGSNGAIGGDLLLNFPMDSVLERQR  
 AHLKYLNPFTDSPLAGFFADSSMITGGEMDSYLSLAGLNLPMMYGETTVEGDSRSLISPE  
 TTLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEEEKSKITEQNNGSTKSIKK  
 MKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARA<sup>AA</sup>QATDSHSIAE<sup>AV</sup>AREKISERMK  
 FLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVA  
 STPMTVVPSPPEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQQVNTSSDPLSCFNNGE  
 APSMWDSHVQONLYGNLGV

[0500]

[0501] > HA-TALE(ng2, C63)-wNES-cib1-mutNLS-mutbHLH

SRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGE  
WDEVQSGLRAADAPPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGY  
QQQEQEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPE  
ATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHA  
WRNALTGAPLNLTPSEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNN  
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPE  
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPV  
CQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALE  
TVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS  
NNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGL  
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQ  
ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
IASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALI  
KRTNRRIPERTSHRVAASLYPERLRRILNGAIGGDLLNFPDMSVLERQRAHLKYLNPT  
FDSPLAGFFADSSMITGGEMDSYLSLAGLNLPMMYGETTVEGDSRLSISPETTLGTGNFK  
AAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEEKSKITEQNGSTKSIKKMKHKAKKE  
ENNFSNDSKVTKELEKTDYIHVRARAAQATDSHSIAEAVA REKISERMKFLQDLVPGC  
DKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFMDDDIFAKEVASTPMTVVPS

[0502] PEMVLSGYSHEMVHSGYSSEMVNSGYLHVNPMQVNTSSDPLSCFNNGEAPSMWDSH  
VQNLYGNLGV

[0503]

[0504] > HA-TALE(ng2, C63)-GS-cib1Δ1-mutNLS

YPYDVDPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
HHTEAATGEWDEVQSGLRAADAPPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAA  
QVDLRTLGYQQQEQEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRG  
GVTAVEAVHAWRNALTGAPLNLTPSEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE  
EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
GLPHAPALIKRTNRRIPERTSHRVAASGGGGSGGGGSNGAIGGDLLNFPDMSVLERQR  
AHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSLAGLNLPMMYGETTVEGDSRLSISPE  
TTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEEKSKITEQNGSTKSIKK  
MKHKAKKEENNFSNDSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKISERMKF  
LQDLVPGCDKITGKAGMLDEIINYVQSLQRGGSVASTPMTVVVPSPEMVLSGYSHEMVH  
SGYSSEMVNSGYLHVNPMQVNTSSDPLSCFNNGEAPSMWDSHVQNLYGNLGV

[0505]

[0506]

> HA-TALE(ng2, C63)-wNLS-cib1 $\Delta$ 1-mutNLS

YPYDVDPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPTMRVAVTAARPPRAKAPRRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSVTAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL

[0507]

ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASLYPERLRRILTNGAIGDLLLLNFPDMSVLERQRA  
 HLKYLNPFTDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRLSISPET  
 TLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEESKITEQNNGSTKSIKKM  
 KHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKISERMKFL  
 QDLVPGCDKITGKAGMLDEIINYVQSLQRGGSGEEESKITEQNNGSTKSIKKMKHKAK  
 KEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKISERMKFLQDLVP  
 GCDKITGKAGMLDEIINYVQSLQRGGSVASTPMTVVPPEMVLVSGYSHEMVHSGYSSE  
 MVNSGYLHVNPMQVNTSSDPLSCFNNGEAPSMWDSHVQNLVGNLGV

[0508]

[0509]

> HA-TALE(ng2, C63)-GS-cib1 $\Delta$ 2-mutNLS

YPYDVDPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPTMRVAVTAARPPRAKAPRRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSVTAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASGGGGSGGGGNGAIGDLLLLNFPDMSVLERQR

[0510]

AHLKYLNPFTDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRLSISPE  
 TTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEESKITEQNNGSTKSIKK  
 MKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVRREKISERMKF  
 LQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFMDDDIFAKEVAS  
 TPMTVVPPEMVLVSGYGGSPSCFNNGEAPSMWDSHVQNLVGNLGV

[0511]

[0512] > HA-TALE(ng2, C63)-wNLS-cib1△2-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSTLNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKAPRRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARG  
 GVTAVEAVHAWRNALTGAPLNLTPQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASLYPERLRRILTNGAIGDLLLLNFPDMSVLERQRA  
 HLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRSLISPET  
 TLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEEEKSKITEQNNGSTKSIKKM  
 KHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVREKISERMKFL  
 QDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVAST  
 PMTVVPSPEMVLVSGYGGSPSCFNNGEAPSMWDSHVQNLVGNLGV

[0513]

[0514] > HA-TALE(ng2, C63)-NLS-cib1-mutNLS-mutbHLH)

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSTLNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKAPRRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARG

[0515]

GVTAVEAVHAWRNALTGAPLNLTPQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASSPKKKR<sup>K</sup>VEASNGAIGDLLLLNFPDMSVLERQR  
 AHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRSLISPE  
 TTLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEEEKSKITEQNNGSTKSIKK  
 MKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARA<sup>AA</sup>QATDSHSIAE<sup>AV</sup>AREKISERMK  
 FLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDMDDIFAKEVA  
 STPMTVVPSPEMVLVSGYSEM<sup>V</sup>HSGYSSEM<sup>V</sup>NSGYLHVNP<sup>MQ</sup>VNTSSDPLSCFNNGE  
 APSMWDSHVQNLVGNLGV

[0516]

[0517] > HA-TALE(ng2, C63)-NLS-cib1 $\Delta$ 1-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDDLRRQFDPSTLNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPTMRVAVTAARPPRAKAPRRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARKG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGGKQALETVQRLLPVLCQAH

[0518] GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASSPKKKRKEASNGAIGGDLNFPDMSVLERQR  
 AHLKYLNPFTDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRSLISPE  
 TTLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEEKSKITEQNNGSTKSIKK  
 MKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHIAERVRREKISERMKF  
 LQDLVPGCDKITGKAGMLDEIINYVQSLQRGGSVASTPMTVVPSPPEMVLSGYSHEMVHS  
 GYSSEMVNSGYLHVNPMQVNTSSDPLSCFNNGEAPSMWDSHVQNLVGNLGV

[0519]

[0520] > HA-TALE(ng2, C63)-NLS-cib1 $\Delta$ 2-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDDLRRQFDPSTLNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPTMRVAVTAARPPRAKAPRRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARKG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASLYPERLRRILTNGAIGGDLNFPDMSVLERQRA  
 HLKYLNPFTDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRSLISPET  
 TLGTGNFKAAKFDTETKDCNEAAKKMTMNRDDLVEEGEEEEKSKITEQNNGSTKSIKKM  
 KHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHIAERVRREKISERMKFL  
 QDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFMDDIFAKEVAST  
 PMTVVPSPPEMVLSGYGGSPSCFNNGEAPSMWDSHVQNLVGNLGV

[0521]

[0522] > HA-TALE(ng2, C63)-GS-iNES1-cib1-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDDLRRQFDPSTLNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPTMRVAVTAARPPRAKAPRRRAAQPSDASPAA

[0523]

QVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARKG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNHDLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASGGGGSGGGGNGAIGDLLLLNFPDMSVLERQR  
 AHLKYLPERLRRILTNPTFDSPLAGFFADSSMITGGEMDSYLSLAGLNLPMMYGETT  
 VEGDSRLSISPETTLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEEEKSKITE  
 QNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERV  
 RREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFD  
 MDDIFAKEVASTPMTVVPSPPEMVLSGYSHEMVHSGYSSEMVN<sup>SGYL</sup>HVNPMQQVNTS  
 SDPLSCFNNGEAPSMWDSHVQNL<sup>YGNL</sup>GV

[0524]

[0525] > HA-TALE(ng2, C63)-GS-iNES2-cib1-mutNLS

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDP<sup>SLFNTSL</sup>FDSLPPFGA  
 HHTEAATGEWDEVQSLRAADAPPPTMRVAVTAARPPRAKPA<sup>RRRRAA</sup>QPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARKG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ

[0526]

ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNHDLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASGGGGSGGGGNGAIGDLLLLNFPDMSVLERQR  
 AHLKYL<sup>NPTFDS</sup>PLAGFFADSSMITGGEMDLYPERLRRIL<sup>TSYL</sup>STAGLNLPMMYGETT  
 VEGDSRLSISPETTLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEEEKSKITE  
 QNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERV  
 RREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFD  
 MDDIFAKEVASTPMTVVPSPPEMVLSGYSHEMVHSGYSSEMVN<sup>SGYL</sup>HVNPMQQVNTS  
 SDPLSCFNNGEAPSMWDSHVQNL<sup>YGNL</sup>GV

[0527]

[0528] > HA-TALE(ng2, C63)-GS-iNES3-cib1-mutNLS

YPYDVDPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASGGGGSGGGGSNGAIGDLLLLNFPDMSVLERQR  
 AHLKYLNPFTDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGLYPERLRR  
 ILTDSRLSISPETTLGTGNFKA<sup>AA</sup>KFDTETKDCNEA<sup>AA</sup>KKMTMNRDDLVEEGEEEEKSKITEQ  
 NNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVR  
 REKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFDM

[0529] DDIFAKEVASTPMTVVPSPPEMVLGYSHEMVHSGYSSEMVNNGYLHVNPMQQVNTSSD  
 PLSCFNNGEAPSMWDSHVQNLVGNLGV

[0530]

[0531] > HA-TALE(ng2, C63)-GS-iNES4-cib1-mutNLS

YPYDVDPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASGGGGSGGGGSNGAIGDLLLLNFPDMSVLERQR  
 AHLKYLNPFTDSPLAGFFADSSMITGGEMDSYLSTAGLNLPMMYGETTVEGDSRLSISPE  
 TTLGTGNFKA<sup>AA</sup>KFDTETKDCNEA<sup>AA</sup>A<sup>AA</sup>KKMTMNRDDLVEEGLYPERLRRILTEEEKSKIT  
 EQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAER  
 VRREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDF  
 DMDDIFAKEVASTPMTVVPSPPEMVLGYSHEMVHSGYSSEMVNNGYLHVNPMQQVNT  
 SSDPLSCFNNGEAPSMWDSHVQNLVGNLGV

[0532]

[0533] > HA-TALE(ng2, C63)-GS-iNES5-cib1-mutNLS

YPYDVDPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSVTAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV

[0534] LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASGGGGSGGGGSNGAIGDLLLLNFPDMSVLERQR  
 AHLKYLNPFDSPLAGFFADSSMITGGEMDSYLSLAGLNLPMMYGETTVEGDSRLSISPE  
 TTLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEEEKSKITEQNGSTKSIKK  
 LYPERRRILTMKHKAKKEENNFSNDSSKVTKLEKTDYIHVRARRGQATDSHSIAERV  
 RREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFD  
 MDDIFAKEVASTPMTVVPSPPEMVLVSGYSHEMVHSGYSSEMVNNGYLHVNPMQQVNTS  
 SDPLSCFNNGEAPSMWDSHVQNLVGNLGV

[0535] > HA-TALE(ng2, C63)-GS-iNES6-cib1-mutNLS

[0536] YPYDVDPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAA  
 QVDLRTLGYSSQQQEKIKPKVRSVTAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK  
 YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARG  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASGGGGSGGGGSNGAIGDLLLLNFPDMSVLERQR

[0537] AHLKYLNPFDSPLAGFFADSSMITGGEMDSYLSLAGLNLPMMYGETTVEGDSRLSISPE  
 TTLGTGNFKA<sup>AA</sup>KFDTETKDCNE<sup>AA</sup>KKMTMNRDDLVEEGEEEEKSKITEQNGSTKSIKK  
 MKHKAKKEENNFSNDSSKVTLPERLRRILTKLEKTDYIHVRARRGQATDSHSIAERV  
 RREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNPRPDFD  
 MDDIFAKEVASTPMTVVPSPPEMVLVSGYSHEMVHSGYSSEMVNNGYLHVNPMQQVNTS  
 SDPLSCFNNGEAPSMWDSHVQNLVGNLGV

[0538] > HA-TALE(ng2, C63)-GS-iNES6-cib1-mutNLS

[0539] > HA-TALE(ng2, C63)-NLS-cib1△1

MYPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLNFNTSLFDSLPPFG  
 AHHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPA  
 AQVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAV  
 KYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAGR  
 GGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVV  
 AIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQA  
 HGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQ  
 RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIG  
 GKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQ  
 VVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLC  
 QAHGLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDA  
 VKKGLPHAPALIKRTNRRIPERTSHRVAASSPKKKRKRVEASNGAIGGDLLLNFPMDSVLE  
 RQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSLAGLNLPMMYGETTVEGDSRLS  
 ISPETTLGTGNFKKRKFDTETKDCNEKMKMTMNRDDLVEEGEEKSKITEQNNGSTKS  
 IKMKHAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVREKISER  
 MKFLQDLVPGCDKITGKAGMLDEIINYVQSLQRGGSVASTPMTVVPSPPEMVLVSGYSHE  
 MVHSGYSSEMVNSGYLHVNPQQVNTSSDPLSCFNNGEAPSMWDSHVQNLYGNLGV

[0540]

[0541] > HA-TALE(ng2, C63)-NLS-cib1△2

YPYDVPDYASRTRLPSPPAPSPAFSADSFSDLLRQFDPSLNFNTSLFDSLPPFGA  
 HHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPA  
 AQVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK

[0542]

YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAGR  
 GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTP  
 EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL  
 PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQ  
 ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVA  
 IASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH  
 GLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKK  
 GLPHAPALIKRTNRRIPERTSHRVAASSPKKKRKRVEASNGAIGGDLLLNFPMDSVLERQR  
 AHLKYLNPTFDSPLAGFFADSSMITGGEMDSYLSLAGLNLPMMYGETTVEGDSRLSISPE  
 TTLGTGNFKKRKFDTETKDCNEKMKMTMNRDDLVEEGEEKSKITEQNNGSTKSIKK  
 MKHAKKEENNFSNDSSKVTKELEKTDYIHVRARRGQATDSHSIAERVREKISERMKF  
 LQDLVPGCDKITGKAGMLDEIINYVQSLQRQIEFLSMKLAIVNRPDFDMDIFAKEVAS  
 TPMTVVPSPPEMVLVSGYGGSPSCFNNGEAPSMWDSHVQNLYGNLGV

[0543]

[0544] > 알파-입포틴-NLS-CRY2PHR-NLS-VP64\_2A\_GFP

MKRPAAATKKAGQAKKKKKMDKKTIVWFRDLRIEDNPALAAAHEGSVFP  
 VFIWCPEEEGQFYPGRASRWWMKQSLAHLSQLKALGSDLTLIKTHNTISAILDCIRVTG  
 ATKVVFNHLYPDVSIVRDHTVKEKLVVERGISVQSYNGDLLYEPWEIYCEKGKPFSTFNS  
 YWKKCLDMSIESVMLPPPWRMPITAAAEAIWACSIEELGLENEAEKPSNALLTRAWSP  
 GWSNADKLLNEFIEKQLIDYAKNSKKVVGNSTSLLSPYLHFGEISVRHVQFCARMKQII  
 WARDKNSEGEESADFLRGIGLREYSRYICFNPFPTHEQSLLSHLRFFPWDADVDFKA  
 WRQGRGTGYPLVDAGMRELWATGWMHNRIRVIVSSFAVKFLLLPWKWGMKYFWDTLL  
 DADLECDILGWQYISGSIPDGHELDRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIH  
 HPWDAPLTVLKASGVELGTNYAKPIVDIDTARELLAKAISRTREAQIMIGAAPASPKKKR  
 KVEASGSGRADALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDL  
 DMLNSRGS GEGRGSLLTCGDVEENPGPVSKGEELFTGVVPILVELDGDVNGHKFSVSG  
 EGEGDATY GKLTLKFICTTGKLPVPWPTLVTTLYGVQCFSRYPDHMKQHDFFKSAMP  
 EGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNH

[0545]  
 NVYIMADKQKNGIKVNFKIRHNIEDGVSQADHYQQNTPIGDGPVLLPDNHYLSTQSAL  
 SKDPNEKRDHMLLEFVTAAGITLGMDELYKV

[0546]

[0547] > mutNES-CRY2PHR-NLS-VP64\_2A\_GFP

MEQKLISEEDLKMDKKTIVWFRDLRIEDNPALAAAHEGSVFPVFIWCPEEE  
 GQFYPGRASRWWMKQSLAHLSQLKAAGSDATLIKTHNTISAILDCIRVTGATKVVFNH  
 LYDPVSLVRDHTVKEKLVVERGISVQSYNGDLLYEPWEIYCEKGKPFSTFNSYWKKCLD  
 MSIESVMLPPPWRMPITAAAEAIWACSIEELGLENEAEKPSNALLTRAWSPGWSNADK  
 LLNEFIEKQLIDYAKNSKKVVGNSTSLLSPYLHFGEISVRHVQFCARMKQIIWARDKNSE  
 GEESADFLRGIGLREYSRYICFNPFPTHEQSLLSHLRFFPWDADVDFKAWRQGRGTGY  
 PLVDAGMRELWATGWMHNRIRVIVSSFAVKFLLLPWKWGMKYFWDTLLDADLECDILG  
 WQYISGSIPDGHELDRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIHHPWDAPLTV  
 LKASGVELGTNYAKPIVDIDTARELLAKAISRTREAQIMIGAAPASPKKKRKVEASGSGR  
 ADALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLNSRGS  
 GEGRGSLLTCGDVEENPGPVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATY  
 GKLTLKFICTTGKLPVPWPTLVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIF  
 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNHNVYIMADKQ  
 KNGIKVNFKIRHNIEDGVSQADHYQQNTPIGDGPVLLPDNHYLSTQSALS  
 SKDPNEKRDHMLLEFVTAAGITLGMDELYKV

[0548]

[0549] > CRY2PHR-NLS-VP64-NLS\_2A\_GFP

MKMDKKTIVWFRDLRIEDNPALAAAHEGSVFPVFIWCPEEEGQFYPGRAS  
 RWWMKQSLAHLSQLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYPDVSIVR  
 DHTVKEKLVVERGISVQSYNGDLLYEPWEIYCEKGKPFSTFNSYWKKCLDMSIESVMLPP  
 PWRLMPITAAAEAIWACSIEELGLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLI  
 DYAKNSKKVVGNSTSLLSPYLHFGEISVRHVQFCARMKQIIWARDKNSEGEESADFLR  
 GIGLREYSRYICFNPFPTHEQSLLSHLRFFPWDADVDFKAWRQGRGTGYPLVDAGMREL  
 WATGWMHNRIRVIVSSFAVKFLLLPWKWGMKYFWDTLLDADLECDILGWQYISGSIPD  
 GHELDRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIHHPWDAPLTVLKASGVELG  
 TNYAKPIVDIDTARELLAKAISRTREAQIMIGAAPASPKKKRKVEASGSGRADALDDFDL  
 DMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLNSPKKKRKEASSR  
 GSGEGRGSLLTCGDVEENPGPVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDAT  
 YGKLTTLKFICTTGKLPVPWPTLVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQER

[0550]

TIFFKDDGNYKTRAEVKFEGDITLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMAD  
 KQKNGIKVNFKIRHNIEDGSVQLADHYQNTPIGDGPVLLPDNHYLSTQSALS KDPNEK  
 RDHMLLEFVTAAGITLGMDELKVK

[0551]

[0552]

> Neurog2-TALE(N240,C63)-PYL

MSRTRLPSPPAPSPAFSADSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATG  
 EWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSPDASPAAQVDLRTLGY  
 SQQQQEKIKPKVRSVTAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALP  
 EATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQDGTGQLLKIARGGVTAVEAVH  
 AWRNALTGAPLNLTPQVVAIASHNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNI  
 GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPE  
 QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHNGGKQALETVQRLLPV  
 LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQAL  
 ETVQRLLPVLCQAHGLTPEQVVAIASHNGGKQALETVQRLLPVLCQAHGLTPEQVVAIA  
 SNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGL  
 TPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLP  
 VLCQAHGLTPEQVVAIASNIGGGRPALESIVAQLSRPDPALAALNDHLVALACLGGRPA  
 LDAVKKGLPHAPALIKRTNRRIPERTSHRVAASMANSESSSPVNEEENSQRISTLHHQT  
 MPSDLTQDEFTQLSQSIAEFHTYQLGNRCSSLLAQRIHAPPETVWSVVRFRDRPQIYKH  
 FIKSCNVSEDFEMRVGCTRDVNVISGLPANTSRERLDLLDDRRVTGFSITGGEHRLRNY  
 KSVTTVHRFEKEEEEEERIWTVVLESYVVDVPEGNSEEDTRLFADTVIRLNLQKLASITEA  
 MNRNNNNNNSSQVR

[0553]

[0554]

>ABI-NLS-VP64

MVPLYGFTSICGRRPEMEEAVSTIPRFLQSSSGSMLDGRFDPQSAAHFFGVYD  
 GHGGSQVANYCRERMHLALAEIIEAKEKPMLCDGDTWLEKWKKALFNSFLRVDSEIESV  
 APETVGSTSVVAVVFPESHIFVANCSDRAVLCRGKTALPLSVDHKPDREDEAARIEAAG  
 GKVIQWNGARVFGVLAMRSIGDRYLKPSIIPDPEVTAVKRVKEDDCLILASDGVWDV  
 MTDEEACEMARKRILLWHKKNVAGDASLLADERRKEGKDPAAAMSAEYLSKLAIQR  
 GSKDNISVVVDLKPRLKSKPLNASPKKKRKEASGSGRADALDDFDLDMLGSDAL  
 DDFDLMLGSDALDDFDLMLGSDALDDFDLMLIN

[0555]

[0556] > hSpCas9(D10A, H840A)-링커-NLS-VP64

MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLF  
 DSGETAEAATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFTHRLEESFLVEEDKK  
 HERHPIFGNIVDEVAYHIEKYPTTYHLRKKLVDSSTDKADLRLIYLALAHMIKFRGHFLIEG  
 DLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGE  
 KKNGLFGNLIASLGLTPNFKSNFDLAEDAQLQLSKDQTYDDDLNLLAQIGDQYADLFL  
 AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDQLTKALVRRQQLPEKYKEIFF  
 DQSKNGYAGYIDGCASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRQRTFDNGSIPH  
 QIHLGELHAILKRQEDFYFLKDNREKIEKILTRIPYYVYVGLARGNSRFAWMTRKSEETI  
 TPWNFEVVDKGGASQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE  
 GMRKPAFLSGEQKKAIVDLDFKTNRKVTYKQLKEDYFKKIECFDSVEISGVEDRPNASL  
 GTYHDLKHKDKDFLDNEENEDILEDIVLTLTLFEDREMIERLKYAHLFDDKYMQL  
 KRRRYTGWRSLRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDDSLTFKEDIQKA  
 QVSGQGDLSLHEHIANLAGSPAIKKGIQTVKVVDELVKVMGRHKKPENIVIEMARENQTT  
 QKGGKNSRERMKRIIEGKELGSQLKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL  
 DINRLSDYDVAIVPQSFLKDDSIDNKVLRSDKNRKGSDNVPSEEVVKKMKNYWRQL  
 LNAKLITQRKFDNLTKAERGGSELKAGPIKRQLVETROITKHVAQILDSRMNTKYDE  
 NDKLIREVKVITLKSCLVSDFFKDFQFYKVRINNYHHHAHDAYLNAVVGTAIKKYPKL  
 ESEFVYGDYKVVYDVRKMIKSEQEIHKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET  
 NGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTFGGFSKESILPKRNSDKLIARCK  
 DWDPKKYGGFDSPTVAYSVLVAKVEKGGKSKLKSVKELGITIMERSSEKFNPIDFLE  
 AKGYKEVKKDLIHLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHY  
 EKLGSPEDNEQKQLFVEQHKHYLDEHIEQISEFSKRVILADANLTKVLSAYNKHHRDKPI  
 REQAENIIHLFTLNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQL  
 LGGDSAGGGGSGGGGSGGGGSGPKKKRKRVAAGSGRADALDDFDLMLGSDALDDF  
 DLDMLGSDALDDFDLMLGSDALDDFDLMLIN

[0557]

[0558] >SID4X-NLS-FLAG-링커-hSpCas9(D10A, H840A)-NLS

MGSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLE  
 RREREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLL  
 EAADYLERREREAEHGYASMLPSRSAGPKKKRKRVAADYKDDDDKSRLEPGEKPYKC  
 PECGKSFSQSGALTRHQRTHTRDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNT  
 DRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSF

[0559]

HRLEESFLVEEDKHERHPFIGNIVDEVAYHEKYPTTYHLRKKLVDSTDKADLRRLIYLAL  
AHMIKFRGHFLIEGDLNPDNSDVVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK  
SRRLLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAQLQSKDTFYDDDLN  
LAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDUTLLKAL  
VRQQLPEKYKEIFPDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDL  
LRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNS  
RFAWMTRKSEETITPWNFEEVVDKASASQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT  
VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS  
VEISGVEDRFNASLGTYHDLKIKKDKDFLDNEENEDILEDIVLTLTLFEDREMIERLKT  
YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFML  
IHDDSLTFKEDIQKAQVSGQDLSLHEHIANLAGSPAIKKGILOTVKVVDELVKVMGRHK  
PENIVEMARENQTTQKGGKNSRERMKRIEKGKELGSQLKEHPVENTQLQNEKLYLYY  
LQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLRSDKNRGRKSDNVPSEE  
VVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGSELDKAGFIKRQLVETRQITKHVA  
QILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNA  
VVGTALIKKYPKLESEFVYGDYKVVYDVRKMIAKSEQEIGKATAKYFFYSNIMNPFKTEIT  
LANGEIRKRPLIETNETGEIVWDKGRDFATVRKVLSPQVNVVKKTEVQTGGFSKESIL  
PKRNSDKLIARKKDWDPKKGFGDSPTVAYSVLVVAKVEKGGKSKKLSVKELLGITM  
ERSSFENPIDFLEAKGYKEVKKDLIKLPKYSLFELENGRKRMLASAGELQKGNELALP  
SKYVNFYLYASHYKLGSPEDNEQKQLFVEQHKHYLDEHEQISEFSKRVLADANLTK  
VLSAYNKHHRDKPIREQAENIHLFTLNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ  
SITGLYETRIDLSQLGGDKRPAATKKAGQAKKKK

[0560]

[0561] 후생적 효과기 도메인 서열

[0562] >hs\_NCoR

ASSPKKKRKVEASMGMLMEDPMKVYKDRQFMNVWTDHEKEIFKDKFIQHP  
KNFGLIASYLERKSVPCVLYYYLTKKNENYKEF

[0563]

[0564] >pf\_Sir2A

ASSPKKKRKVEASMGMLMISFLKKTQSITLLELAKIHKCKHVVALTGSST  
AESNIPFRGSSNSIWSKYDPRYGTIWFWKYPEKIWEVIRDISSDYEIEINNGHVALSTL  
ESLGYLKSVVQTQNVVDGLHEASGNTKVISLHGNVFEAVCCTCNKIVKLNKIMLQKTSHEM

[0565]

HQLPPECPCGGIFKPNILFGEVVSSDLLKEAEEIACDLLLLVIGTSSTVSTATNLCHFAC  
KKKKKIVEINISKTYITNKMSDYHVCAKFSELTKVANILKGSSEKNKIMEF

[0566]

[0567] >nc\_DIM5

ASSPKKKRKVEASMEKAFRPHFFNHGKPDANPKEKKNCHWCQIRSFATHAQ  
LPISIVNREDDAFLNPNFRFIDHSIIGKNVPVADQSFVRVGCSCASDEECMYSTCQCLDEMA  
PDSDEEADPYTRKKRFAYYSQGAKKGLLRDRVLQSQEPIYEQHCACSKDCPNRVVE  
RGRTVPLQIFRTKDRGWGVKCPVNIKRGQFVDRYLGEIITSEEADRRRAESTIARRKDVY  
LFALDKFSDPDSLPLLAGQPLEVDGEYMSGPTRFINHSCDPNMAIFARVGDHADKHHI  
DLALFAIKDIPKGTELTFDYVNGLTGLESDAHDPSKISEMTKCLCGTAKCRGYLWEF

[0568]

- [0569] >sc\_HST2  
 ASSPKKKRKVEASTEMSVRKIAAHMKSNNPAKVIFMVGAGISTSCGIPDFRSP  
 GTGLYHNLARLKLPEAVFDVDFQSDPLPFYTLAKELYPGNFRPSKFHYLLKLFQDK  
 DVLKRVTQNIIDTLERQAGVKDDLIIEAHGSFAHCHCIGCGKVYPPQVFKSKLAEHPIKD  
 FVKCDVCGELVKPAIVFFGEDLPDSFSETWLNSEWLREKITTSGKHPQQPLVIVVGTSL  
 AVYPFASLPPEEIPRKVKRVLCNLETVGDFAKANKRPTDLIVHQYSDEFQAEQLVEELGWQE  
 DFEKILTAQGGMGEF
- [0570]
- [0571] >hs\_SIRT3  
 ASSPKKKRKVEASMVGAGISTPSGIPDFRSPGSLYSLNQYDLPYPEAIFELP  
 FFFHNPKPFFTLAKELYPGNYKPNVTHYFLRLLHDKGLLRLTYQNIIDGLERVSGIPASK  
 LVEAHGTFASATCTVCQRPFPGEDIRADVMADRVPKPCVCTGVVVKPDIVFFGEPLPQRFL  
 LHVVDVPMADLLILGTSLEVEPFASLTEAVRSSVPRLLINRDLVGPLAWHPRSRDVAQL  
 GDVVHGVESLVELLWTEEMRDLVQRETGKLDGPDKEF
- [0572]
- [0573] >hs\_NIPP1  
 ASSPKKKRKVEASMAAAAANSGLPLFDCPTWAGKPPPGLHLDVVKGDKLIE  
 KLIIDEKYYLFGRNPDLCDFDIDHQSCSRVHAALVYHKHLKRVFLIDLNSTHGTFLGHI  
 RLEPHKQQIPIDSTVSGASTRAYTLREKQTLPSAVKGDEKMGGEDELKLLGLPEE  
 ETELDNLTEFNTAHNKRISTLTIEGNLDIQRPKRKRKNSRVTFSEDEIINPEDVDPSVGR  
 FRNMVQTAVVVPVKKRVEGPGSLGLEESGSRMQNFASFGLYGGPLPPTHSEAGSQPH  
 GIHGTLIGGLPMPYPNLAPDVLTPVVPSAVNMNPAVNPVYNPEAVNEEF
- [0574]
- [0575] >ct\_NUE  
 ASSPKKKRKVEASMTNSTQDTLYLSLHGGIDSAIPYVRRVEQLLQFSFLPE  
 LQFQNAAVKQRIQRLCYREEKRLAVSSLAKWLGQLHKQRLRAPKNPPVAICWINSYVG  
 YGVFARESIPAWSYIGEYTGILRRRQALWLDENDYCFRYVPVPRYSFRYFTIDSGMQNV  
 TRFINHSDNPNLEAIGAFENGIFHIIIRAIKDILPGEELCYHYGPLYWKHRKKREEFVPQEE  
 EF
- [0576]
- [0577] >hs\_MBD2b  
 ASSPKKKRKVEASARYLGNTVDLSSDFRGTGMMPSKLQKNKQRLRNDPLN  
 QNKGKPDNLNTLPIRQTASIFKQPVTKVTNHPNSNKVSDPQRMNEQPRQLFWEKRLQGL  
 SASDVTEQIIKTMELPKGLQGVGPGSNDLTLASAVASALHTSSAPITGQVSAAVEKNPAV  
 WLNTSQPLCKAFIVTDEDIRKQEERVQVVRKILEDALMADILSRAADTEEMDIEMDSGD  
 EAEF
- [0578]
- [0579] >ca\_HST2  
 ASSPKKKRKVEASMPSLDDILKPVAEAVKNGKVTFFNGAGISTGAGIPDFRS  
 PDTGLYANLAKLNLPEAEAVFDIDFFKEDPKPFYTLAEELYPGNFAPTKFHFIKLLQDQ  
 GSLKRVTQNIIDTLERLAGVEDKYIVEAHGSFASNHCVDCHKEMTTETLKTYMKDKKI  
 PSCQHCEGYVKPDIVFFGEGPLVVKFFDLWEDDCEDVEVAIVAGTSLTVFPFASLPGEVVK  
 KCLRVLVNKEKVGTFKHEPRKSDIHALHDCDIVAERLCTLLGLDDKLENYEKEKIKYSK  
 AETKEIKMHEIEDKLKEEHLKEDKHTTKVDRKKEKQNDANDKELEQLIDKAKAEF
- [0580]

- [0581] >hs\_PHF19  
 ASSPKKKRKVEASMENRALDPGTRDSYGATSHLPNKGALAKVKNNFKDLMS  
 KLTEGQYVLCRWTDGLYYLGKIKRVSSSKQSCLVTFEDNSKYVVLWKDIQHAGVPGEE  
 PKCNICLGKTSGPLNEILICGKCGLYHQCHIPIAGSADQPLLTPWFCRRCIFALAVRKG  
 GALKKGAIARTLQAVKMLVLSYQPEELEWDSPHRTNQQCYCYCGGPGEWYLRMLQCY  
 RCRQWFHEACTQCLNEPMMFGDRFYLFCSVCNQGPGGSGSDSSAEGASVPERPDEGID  
 SHTFESISEDSSLSHLKSSITNYFGAAGRLACGEKYQVLARRVTPEGKVQYLVEWEGTT  
 PYEF
- [0582]
- [0583] >hs\_HDAC11  
 ASSPKKKRKVEASMLHTTQLYQHVPETRWPIVYSPRYNITFMGLEKLHPFDA  
 GKWGVKVINFLKEEKLLSDSMLVEAREASEEDLLVVHTRRYLNELKWSFAVATITEIPPVI  
 FLPNFLVQRKVLRLRPTQTGGTIMAGKLAVERGWAINVGGGFHHCSSDRGGGFCAYAD  
 ITLAIKFLFERVEGISRATIIDLDAHQNGHERDFMDDKRVYIMDVYNRHIYPGDRFAKQ
- [0584]  
 AIRRKVELEWGTEDDEYLDKVERNIKSLQEHLPDVVVYNAGTDILEGDRLGGLSISPA  
 GIVKRDELVFRMVRGRRVPILMVTSGGYQKRTARIADSILNLFGLGLIGPESPSVSAQNS  
 DTPLPPAVPEF
- [0585]
- [0586] >ml\_MesoLo4  
 ASSPKKKRKVEASMPLQIVHHPDYDAGFATNHRFPMSKYPLLMEALRARGL  
 ASPDALNTTEPAPASWLKLAHAADYVDQVISCVPEKIEREIGFPVGPVSLRAQLATGG  
 TILAAARLALRHGIACNTAGGSHHARRAQGAGFCTFNDVAVASLVLLDEGAAQNILVVD  
 LDVHQGDGTADILSDEPGVFTFSMHGERNYPVRKIASDLIALPDGTGDAAYLRRLATIL  
 PELSARARWDIVFYNAGVDVHAEDRLGRLALSNGGLRARDEMVIHFALGIPVCGVI  
 GGGYSTDVPALASRHAILFEVASTYAEF
- [0587]
- [0588] >pbv1\_vSET  
 ASSPKKKRKVEASMFNDRVIVKKSPLGGYGVFARKSFEEKGELVEECLCIVRH  
 NDDWGTALEDYLSRKNMSAMALGFGAIFNHKDPNARHELTAGLKRMRIFTIKPIAIG  
 EEITISYGDDYWLSRPLTQNEF
- [0589]
- [0590] >at\_KYP  
 ASSPKKKRKVEASDISGGLFEGKIPATNRVDDSPVSPTSFTYIKSLIIEPNVIIP  
 KSSTGCNCRGSDTSSKACAKLNGGNFPYVDLNDGRLIESRDVVFECGPHCGCGPKC  
 VNRTSQKRLRFNLEVFSAKKGWAVRSWEYIPAGSPVCEYIGVVRRTADVDTISDNEYI  
 FEIDCQQTMQGLGGRQRRLDVAVPMNNGVSQSSDENAPEFCIDAGSTGNFARFINHS  
 CEPNLFVQCVLSSHQDIRLARVVLFAADNISPMQELTYDYGALDSVHEF
- [0591]
- [0592] >tg\_TgSET8  
 ASSPKKKRKVEASASRRTGEFLRDAQAPSRWLKRSKTGQDDGAFCLLETWLA  
 GAGDDAAGGERGRDREGAADKAKQREERRQKELEERFEEMKVEFEKAQRMIAARRAA  
 LTGEIYSDGKGSKKPRVPSLPENDDALIEIIDPEQGILKWPLSVMSIRQRTVIYQECLRR  
 DLTACIHLTKVPGKGRAVFAADTILKDDFVVEYKGEELCSEREAREQRYNRSKVPMGS  
 FMFYFKNGSRMMAIDATDEKQDFGPARLINHSRRNPNMTPRAITLGFNSEPRLIFVARR  
 NIEKGEELLVDYGERDPDIKEHPWLNSEF
- [0593]

- [0594] >hs\_SIRT6  
 ASSPKKKRKVEASMSVNYAAGLSPYADKGGKGLPEIFDPPEELERKVVWELAR  
 LVWQSSSVVFHTGAGISTASGIPDFRPHGVWTMEERGLAPKFDTTFESARPTQTHMAL  
 VQLERVGLLRFLVSQNVDDLHVRSGFPRDKLAELHGNMFVEECAKCKTQYVRDVTVVG
- [0595] TMGLKATGRLECTVAKARGLRACRGELRDITLDWEDSLPDRDLALADEASRNADLSITL  
 GTSLQIRPSGNLPLATKRRGGRLVIVNLQPTKHDRHADLRIHGYVDEVMTRLMKHLGLE
- [0596] IPAWDGPRVLERALPPLEF
- [0597] >ce\_Set1  
 ASSPKKKRKVEASMKVAAKKLATSRRMRKDRAAAASPSSDIENSENPSSLASH  
 SSSSGRMTPSKNTRSRRKGVSVKDVSNHKITEFFQVRRSNRKTQSKQISDEAKHALRDTVL  
 KGTNERLLEVYKDVVKGRGIRTKVNFEEKGDFVVEYRGVMMMEYSEAKVIEEQYSNDEEI  
 GSYMYFFEHNKKWCIDATKESPWKGRLINHSVLRPNLKTQVVEIDGSHHLILVARRQI  
 AQGEELLYDYGDRSAETIAKNPWLVTNTEF
- [0598]
- [0599] >mm\_G9a  
 ASSPKKKRKVEASVRTEKIICRDVARGYENVPIPCVNGVDGEPEDYKYISE  
 NCETSTMNIDRNITHLQHCTCVDDCSSNCLCGQLSIRCWYDKDGRLLQEFNKIEPLIFE  
 CNQACSCWRSCNRRVVQSGIKVRLQLYRTAKMGWGVRAALQTIPQGTVICEYVVELISD  
 AEADVREDDSYLFDLNDKDGVEYCIDARYYGNISRFINHLCDPNIPVRVFMLHQDLRFP  
 RIAFFSSRDIRTGEELGFDYGDRFWDIKSKYFTCCQCGSEKCKHSAEIALEQSRLARLDP  
 HPELLPDLSSLPPINTEF
- [0600]
- [0601] >hs\_SIRT5  
 ASSPKKKRKVEASSSMADFRKFFAKAKHIVIIISGAGVSAESGVPTFRGAGGY  
 WRKWQAQDLATPLAFAHNPSRVWFYHYRREVMGSKEPNAGHRAIAECETRLGKQGR  
 RVVVITQNIDELHRKAGTKNLLEIHGSLFKTRCTSCGVVAENYKSPICPALSCKGAPEPG  
 TQDASIPVEKLPRCEEAGCGGLLRPHVVWFGENLDPAILVEVDRELAHCDLCLVVGTS  
 VVYPAAMFAPQVAARGVPVAEFNTETTPATNRFHFQGPCGTTLPEALACHENETVSE  
 F
- [0602]
- [0603] >x1\_HDAC8  
 ASSPKKKRKVEASMSRVVVKPVASMEEMAAFHTDAYLQHLHKVSEEGDND  
 DPETLEYGLGYDCPITEGIYDYAAAVGGATLTAAEQIIEGKTRIAVNWPGGWHHAKKD  
 EASGFCYLNDVAVLGIKLREKFDRLVLYVMDLHHGDGVEDAFSFTSKVMTVSLHKFSP  
 GFFPGTGDVSDIGLGKGRYYSINVPLQDGIQDDKYYQICEGVLKEVFTTFNPEAVVLQLG  
 ADTIAGDPMCSFNMTPEGIGKCLKYVLQWQLPTLILGGGGYHLPNTARCWTYLTALIVG  
 RTLSEIPDHEFFTEYGPDYVLEITPSCRPRDRNDTQKVQEILQSIKGNLKRVEVF
- [0604]
- [0605] >mm\_HP1a  
 ASSPKKKRKVEASMKEGENNKPREKSEGKSRKSSFSNSADDIKSKKKREQSN  
 DIARGFERGLEPEKIIIGATDSCGDLMLMKWKDTDEADLVLAKEANVKCPQIVIAFYEE  
 RLTHWYHPEDAENKEKESAKSEF
- [0606]
- [0607] >at\_HDT1  
 ASSPKKKRKVEASMEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKG  
 EFVPLHVKVGQNQLVLTSTENIPQLFCDLVFDKEFELSHTWKGKGSVYFVGYKTPNIEP  
 QGYSEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDESDDGMEDESDSDE  
 DSEEEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTPQKTDEKKGKGAANQSEF
- [0608]

- [0609] >mm\_SA11  
 ASSPKKKRKVEASMSRRKQAKPQHFQSDPEVASLPRRDGDTEKGQPSRPTKS  
 KDAHVCGRCCAEFFELSDLLLHKKSCTKNQLVLIVNESPASPAKTFPPGPSLNDEF
- [0610]
- [0611] >hs\_SETD8  
 ASSPKKKRKVEASSCDSTNAAIKQALKKPIKQAPRKAQGKTQQNRKL  
 TDFYPVRRSSRSKAEQSEERKRIDELIESGKEEGMKIDLIDGKGRGVIAATKQFSRGDFV  
 VEYHGLDIEITDAKKREALYAQDPSTGCMYFYFQYLSKTYCVDATRETNRLGRLINHSK  
 CGNCQTKLHDIDGVPHLILIASRDIAAGEELLYDYGDRSKASIEAFPWLKHEF
- [0612]
- [0613] >sc\_RPD3  
 ASSPKKKRKVEASRRVAYFYDADVGNAYGAGHPMKPHRIRMAHSLIMNY  
 GLYKMEIYRAKPAKQEMCQFHTDEYIDFLSRVTPDNLEMFKRESVKFNVGDDCPVF  
 DGLYEYCSISGGGMEGAARLNRGKCDVAVNYAGGLHHAKKSEASGFCYLNDIVLGIIE  
 LLRYHPRVLYIDIDVHHGDGVEEAFYTTDRVMTCSEFHKYGEFFPGTGELRDIGVGAGKN  
 YAVNVPLRDGIDDATYRSVFEPVIKKIMEWYQPSAVVLQCGGDSLSGDRLGCFNLSME  
 GHANCVNYVKSFGIPMMVVGGGGYTMRNVARTWCFETGLLNNVVLDKDLPYEF
- [0614]
- [0615] >ec\_CobB  
 ASSPKKKRKVEASMEKPRVLVLTGAGISAESGIRTFRAADGLWEEHRVEDVA  
 TPEGFDRDPELVQAFYNARRRQLQQPEIQPNAHLALAKLQDALGDRFLLVTQNIDNLH  
 ERAGNTNVIHMHGELLKVRCSQSQVLDWTGDVTPEDKCHCCQFPAPLRPHVWVWFG  
 MPLGMDEIYMALSMADIFIAIGTSGHVYPAAGFVHEAKLHGAHTVELNLEPSQVGNFA  
 EKYYGPASQVPEFVEKLLKGLKAGSIAEF
- [0616]
- [0617] >hs\_SUV39H1  
 ASSPKKKRKVEASNLKCVRIKQFHKDLERELRRHRSKTPRHLDPSLANY  
 LVQKAKQRRALRRWEQELNAKRSHLGRITVENEVDLDGPPRAFVYINEYRVGEGITLNQ  
 VAVGCECQDCLWAPGTGCCPGASLHKFAYNDQGVRLRAGLPIYECNSRCRCGYDCP  
 NRVVQKGIYDLCIFRTDDGRGWVVRTLEKIRKNSFVMEYVGEIITSEEARRGQIYDRQ  
 GATYLFDLDYVEDVYTVDAAYYGNISHFVNHSKDPNLQVYVNFIDNLDERLPRIAFFAT  
 RTIRAGEELTFDYNMQVDPVDMESTRMDSNFGLAGLPGSPKKRVRIECKCGTESCRKYL  
 FEF
- [0618]
- [0619] >hs\_RCOR1  
 ASSPKKKRKVEASSNSWEEGSSGSSSDEEHGGGGMRVGPQYQAVVPDFDPA  
 KLARRSQERDNLGMLVWSPNQNLSEAKLDEYIAIAKEKHGYNMEQALGMLFWHKHNI  
 EKSLADLPNFTPPDEWTVEDKVLFEQAFSFGKTFHRIQQMLPDKSIASLVKFYYSWK  
 KTRTKTSVMDRHRARKQKRERESEDELEEANGNNPIDIEVDQNKESKKEVPPTETVPQV  
 KKEKHSTEF
- [0620]
- [0621] >hs\_s in3a  
 ASSPKKKRKVEASYKESVHLETYPKERATEGIAMEIDYASCKRLGSSYRALP  
 KSYQQPKCTGRTPCKEVLNDTWVSPSWSEDSTFVSSKKTQYEEHIYRCEDERFELDV  
 VLETNLATIRVLEAIQKLSRLSAEEQAKFRLDNTLGGTSEVIHRKALQRIYADKAADIID  
 GLRKNPSIAVPIVLKRLKMKEEEWREAQRGFNKVWREQNEKYLYKSLDHQGINFKQND  
 TKVLRKSKLLNEIESIYDERQEATEENAGVPVGPPLSLAYEDKQILEDAAALIIHHVKR  
 QTGIQKEDKYKIKQIMHHPIDLLFAQRGDLSDVEEEEEEMDVDEATGAVEF
- [0622]

[0623] >at\_SUVR4

ASSPKKKRKVEASQAYLHVSLARISDEDCCANCKGNCLSAADFPTCARETS  
 GEYAYTKEGLLKEKFLDTCLMKKEPDSFPKVYCKDCPLERDHDKGTYGKCDGHLIRK  
 FIKECWRKCGCDMQCGNRVVQRGIRCQLQVYFTQEGKGWGLRTLQDLPKGTFIGEYIG  
 EILTNTELYDRNVRSSSERHTYPVTLADADWGSEKDLKDEEALCLDATICGNVARFINHR  
 CEDANMIDIPIEIEIPDRHYHIAFFTLRDVKAMDELTDWYMFNDKSHPVKAFRCCC  
 GSESCRDRKIKGSQKSIERRKIVSAKKQQGSKEVSKKRKEF

[0624] >rn\_MeCP2\_NLS

ASSPKKKRKVEASVQVKRVLEKSPGKLLVKMPFQASPGGKGGGATTSAQ  
 VMVIKRPGRKRKAADPQAIPKGRKPGSVVAAAAAEAKKAVKESSIRSVQETVLP  
 KKRKTRETVSIEVKEVVKPLLVLSTLGEKSGKGLKTKCKSPGRKSKESSPKGRSSASSPPK

[0625] KEHHHHHHHAESPKAPMPLPPPPPEPQSSSEDPIPPPEQDLSSSICKEEKMPRAGSLESD  
 GCPKEPAKTQPMVAAAAATTTTTTTTTVAEKYKHRGEGEKDIVSSMPPRNREEPVDSR  
 TPVTERVSEF

[0626] >mm\_SET-TAF1B

ASSPKKKRKVEASMAPKRQSAILPQPKPRPAAAPKLEDKSASPLPKGEKE  
 QQEAIEHIDEVQNEIDRLNEQASEEILKVEQKYNKLRQPFQKRSELIKIPNFWVTTFFVN  
 HPQVSALLGEEDEEALHYLTRVEVTEFEDIKSGYRIDFYFDENPYFENKVLKSEFHLNES  
 GDPSSKSTEIKWKSGLDLTKRSSQTQNKASRKRQHEEPESFFTWFTHSDAGADELGEV  
 IKDDIWPNPLQYYLVPDMDEEGEAEDDDDDDEEEGLEDEEGDEDEGEDDDEDE  
 GEEGEEDEGEDDEF

[0627] >ce\_Set4

ASSPKKKRKVEASMLHEQIANISVTFNDIPRSDHSMTPTELCYFDDFATTLV  
 VDSVLNFTTHKMSKKRRYLYQDEYRTARTVMKTFREQRDWTNAIYGLLTLRSVSHFLS  
 KLPPNKLFEFRDHIVRFLNMFILDSGYTIQECKRYSQEGHQAKLVSTGVWSRGDKIERL  
 SGVVCLLSEDEDSEILAQEGSDFSMYSTRKRCSTLWLGPGAYINHDCRPTCEFVSHGST  
 AHIRVLRDMVPGDEITCFYGSEFFGPNNDCECCTCEKNMNGAFSYLRGNENAEPHISEK  
 KTKYELRSRSEF

[0628] 광자극 하드웨어 제어 스트립트

[0629] 하기의 Arduino 스트립트를 광 자극된 24-웰 플레이트의 각 4-웰 열의 개별 제어가 가능하도록 사용하였다.

```
[00573] //Basic control code for LITE LED array using Arduino UNO
[00574] //LED column address initialization to PWM-ready Arduino outputs
[00575] int led1_pin = 3;
[00576] int led2_pin = 5;
[00577] int led3_pin = 6;
[00578] int led4_pin = 9;
[00579] int led5_pin = 10;
[00580] int led6_pin = 11;
[00581] //Maximum setting for Arduino PWM
[00582] int uniform_brightness = 255;
[00583] //PWM settings for individual LED columns
[00584] int led1_brightness = uniform_brightness/2;
[00585] int led2_brightness = uniform_brightness/2;
[00586] int led3_brightness = uniform_brightness/2;
[00587] int led4_brightness = uniform_brightness/2;
```

[0634]

```

[00588]     int led5_brightness = uniform_brightness/2;
[00589]     int led6_brightness = uniform_brightness/2;
[00590]     // 'on' time in msec
[00591]     unsigned long uniform_stim_time = 1000; //
[00592]     // individual 'on' time settings for LED columns
[00593]     unsigned long led1_stim_time = uniform_stim_time;
[00594]     unsigned long led2_stim_time = uniform_stim_time;
[00595]     unsigned long led3_stim_time = uniform_stim_time;
[00596]     unsigned long led4_stim_time = uniform_stim_time;
[00597]     unsigned long led5_stim_time = uniform_stim_time;
[00598]     unsigned long led6_stim_time = uniform_stim_time;
[00599]     // 'off' time in msec
[00600]     unsigned long uniform_off_time = 14000;
[00601]     // individual 'off' time settings for LED columns
[00602]     unsigned long led1_off_time = uniform_off_time;
[00603]     unsigned long led2_off_time = uniform_off_time;
[00604]     unsigned long led3_off_time = uniform_off_time;
[00605]     unsigned long led4_off_time = uniform_off_time;
[00606]     unsigned long led5_off_time = uniform_off_time;
[00607]     unsigned long led6_off_time = uniform_off_time;
[00608]     unsigned long currentMillis = 0;
[00609]     // initialize timing and state variables
[00610]     unsigned long led1_last_change = 0;
[00611]     unsigned long led2_last_change = 0;
[00612]     unsigned long led3_last_change = 0;
[00613]     unsigned long led4_last_change = 0;
[00614]     unsigned long led5_last_change = 0;
[00615]     unsigned long led6_last_change = 0;
[00616]     int led1_state = HIGH;
[00617]     int led2_state = HIGH;
[00618]     int led3_state = HIGH;
[00619]     int led4_state = HIGH;
[00620]     int led5_state = HIGH;
[00621]     int led6_state = HIGH;
[00622]     unsigned long led1_timer = 0;
[00623]     unsigned long led2_timer = 0;
[00624]     unsigned long led3_timer = 0;
[00625]     unsigned long led4_timer = 0;
[00626]     unsigned long led5_timer = 0;
[00627]     unsigned long led6_timer = 0;
[00628]     void setup() {

```

[0635]

```

[00629] // setup PWM pins for output
[00630] pinMode(led1_pin, OUTPUT);
[00631] pinMode(led2_pin, OUTPUT);
[00632] pinMode(led3_pin, OUTPUT);
[00633] pinMode(led4_pin, OUTPUT);
[00634] pinMode(led5_pin, OUTPUT);
[00635] pinMode(led6_pin, OUTPUT);
[00636] //LED starting state
[00637] analogWrite(led1_pin, led1_brightness);
[00638] analogWrite(led2_pin, led2_brightness);
[00639] analogWrite(led3_pin, led3_brightness);
[00640] analogWrite(led4_pin, led4_brightness);
[00641] analogWrite(led5_pin, led5_brightness);
[00642] analogWrite(led6_pin, led6_brightness);
[00643] }
[00644] void loop() {
[00645]   currentMillis = millis();
[00646]   //identical timing loops for the 6 PWM output pins
[00647]   led1_timer = currentMillis - led1_last_change;
[00648]   if (led1_state == HIGH) { //led state is on
[00649]     if (led1_timer >= led1_stim_time) { //TRUE if stim time is complete
[00650]       analogWrite(led1_pin, 0); //turn LED off
[00651]       led1_state = LOW; //change LED state variable
[00652]       led1_last_change = currentMillis; //mark time of most recent change
[00653]     }
[00654]   }
[00655]   else { //led1 state is off
[00656]     if (led1_timer >= led1_off_time) { //TRUE if off time is complete
[00657]       analogWrite(led1_pin, led1_brightness); //turn LED on
[00658]       led1_state = HIGH; //change LED state variable
[00659]       led1_last_change = currentMillis; //mark time of most recent change
[00660]     }
[00661]   }
[00662]   led2_timer = currentMillis - led2_last_change;
[00663]   if (led2_state == HIGH) {
[00664]     if (led2_timer >= led2_stim_time) {
[00665]       analogWrite(led2_pin, 0);
[00666]       led2_state = LOW;
[00667]       led2_last_change = currentMillis;
[00668]     }
[00669]   }

```

[0636]

```

[00670]     else[ //led2 state is off
[00671]         if (led2_timer >= led2_off_time) {
[00672]             analogWrite(led2_pin, led2_brightness);
[00673]             led2_state = HIGH;
[00674]             led2_last_change = currentMillis;
[00675]         }
[00676]     }
[00677]     led3_timer = currentMillis - led3_last_change;
[00678]     if (led3_state == HIGH) {
[00679]         if (led3_timer >= led3_stim_time) {
[00680]             analogWrite(led3_pin, 0);
[00681]             led3_state = LOW;
[00682]             led3_last_change = currentMillis;
[00683]         }
[00684]     }
[00685]     else[ //led3 state is off
[00686]         if (led3_timer >= led3_off_time) {
[00687]             analogWrite(led3_pin, led3_brightness);
[00688]             led3_state = HIGH;
[00689]             led3_last_change = currentMillis;
[00690]         }
[00691]     }
[00692]     led4_timer = currentMillis - led4_last_change;
[00693]     if (led4_state == HIGH) {
[00694]         if (led4_timer >= led4_stim_time) {
[00695]             analogWrite(led4_pin, 0);
[00696]             led4_state = LOW;
[00697]             led4_last_change = currentMillis;
[00698]         }
[00699]     }
[00700]     else[ //led4 state is off
[00701]         if (led4_timer >= led4_off_time) {
[00702]             analogWrite(led4_pin, led4_brightness);
[00703]             led4_state = HIGH;
[00704]             led4_last_change = currentMillis;
[00705]         }
[00706]     }
[00707]     led5_timer = currentMillis - led5_last_change;
[00708]     if (led5_state == HIGH) {
[00709]         if (led5_timer >= led5_stim_time) {
[00710]             analogWrite(led5_pin, 0);

```

[0637]

```

[00711]         led5_state = LOW;
[00712]         led5_last_change = currentMillis;
[00713]     }
[00714] }
[00715] else[ //led5 state is off
[00716]     if (led5_timer >= led5_off_time){
[00717]         analogWrite(led5_pin, led5_brightness);
[00718]         led5_state = HIGH;
[00719]         led5_last_change = currentMillis;
[00720]     }
[00721] }
[00722] led6_timer = currentMillis - led6_last_change;
[00723] if (led6_state == HIGH){
[00724]     if (led6_timer >= led6_stim_time){
[00725]         analogWrite(led6_pin, 0);
[00726]         led6_state = LOW;
[00727]         led6_last_change = currentMillis;
[00728]     }
[00729] }
[00730] else[ //led6 state is off
[00731]     if (led6_timer >= led6_off_time){
[00732]         analogWrite(led6_pin, led6_brightness);
[00733]         led6_state = HIGH;
[00734]         led6_last_change = currentMillis;
[00735]     }
[00736] }
[00737] }

```

[0638]

[0639] *실시예 12*

[0640] *내인성 포유동물 전사의 광학적 제어*

[0641]

1 차 마우스 피질 뉴런에서의 전사를 조정하기 위한 AAV-매개 TALE 전달의 효능을 시험하기 위해서, 출원인은 각각 mGluR5, NMDA 서브유닛 2A 및 mGluR2 를 인코딩하는 3 개 마우스 신경전달물질 수용체: Grm5, Grin2a 및 Grm2 의 유전자좌를 표적화하는 6 개 TALE-DNA 결합 도메인을 구축하였다 (도 58). 표적 부위 접근 가능성을 증가시키기 위해, 출원인은 UCSC 게놈 브라우저로부터 마우스 피질 DNase I 민감성 데이터를 사용하여 추정 상 개방형 크로마틴 부위를 확인하였다. 각각의 표적 유전자의 프로모터 내 DNase I 민감성 부위는 TALE 결합 서열의 선택을 위한 안내를 제공하였다 (도 46). 각각의 TALE 에 대해서, 출원인은 mSin3 상호작용 도메인 (SID) 의 퀴드러플 탠덤 반복물 또는 전사 활성제로서 VP64 를 이용하였다 (Beerli, R.R., Segal, D.J., Dreier, B. & Barbas, C.F., 3rd Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proc Natl Acad Sci USA 95, 14628-14633 (1998) and Ayer, D.E., Laherty, C.D., Lawrence, Q.A., Armstrong, A.P. & Eisenman, R.N. Mad proteins contain a dominant transcription repression domain. Molecular and Cellular Biology 16, 5772-5781 (1996)) as a repressor. Applicants have previously shown that a single SID fused to TALE downregulated a target gene effectively in 293FT cells (Cong, L., Zhou, R., Kuo, Y.-c., Cunniff, M. & Zhang, F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. Nat Commun 3, 968 (2012)). 이러한 TALE 억제제를 추가로 향상시키기 위해, 출원인은 VP64 의 성공적 퀴드러플 VP16 반복물 구조와 유사한 SID 의 4 개 반복물 (Beerli, R.R., Segal, D.J., Dreier, B. & Barbas, C.F., 3rd Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proc Natl Acad Sci USA 95, 14628-14633 (1998)) 이 이의 억제 활성을 증가시킬 수 있다고 추론하였다. 이는 실제로 TALE-SID4X 구축물이 293FT 세포에서 TALE-SID 에 대해 억제를 ~2 배 증진시킨 바와 같은 경우이다 (도 54).

[0642]

출원인은 6 개 TALE-VP64 구축물 중 4 개 (T1, T2, T5 및 T6) 가 AAV-형질도입된 1 차 뉴런에서의 그의 표적 유전자 Grm5 및 Grm2 를, 각각 2- 및 8-배까지 효율적으로 활성화시켰다는 것을 발견하였다 (도 58). 유사

하게, 6 개 TALE-SID4X 억제제 중 4 개 (T9, T10, T11, T12) 가 그의 내인성 표적 Grin2a 및 Grm2 의 발현을 각각 2- 및 8-배까지 감소시켰다 (도 58). 함께, 이러한 결과는 구성적 TALE 이 뉴런에서의 내인성 표적 유전자 발현을 양상적 또는 음성적으로 조절할 수 있다는 것을 나타낸다. 특히, 주어진 TALE 에 의한 효율적인 활성화 또는 억제는 반대편 방향에서의 전사 조절에 있어서 그의 효율을 예측하지 않았다. 따라서, 다수의 TALE 은 특정 유전자좌에 대한 가장 효율적 TALE 을 확인하기 위해 스크리닝될 필요가 있을 수 있다.

[0643] LITE 의 뉴런 적용을 위해, 출원인은 출원인의 6 개 구성적 TALE 활성화제 비교를 기반으로 (도 58), 1 차 뉴런에서 가장 강력한 수준의 표적 상향조절을 나타낸 Grm2 TALE (T6) 을 선택하였다. 출원인은 0.8% 의 동일한 사용물로 2 광 펄스 주파수를 사용하여 이의 기능을 조사하였다. 자극 조건 모두 Grm2 mRNA 수준에 있어서 ~7 배 광-의존적 증가를 획득하였다 (도 38C). 추가 연구로, 유의한 표적 유전자 발현 증가가 빠르게 이루어질 수 있다는 것이 확인되었다 (4 시간 내에 4-배 상향조절; 도 38D). 또한, 출원인은 자극 후 mGluR2 단백질의 유의한 상향조절을 관찰하였는데, 이는 mRNA 수준에서 LITE 에 의해 영향받은 변화가 단백질 도메인으로 옮겨지는 것을 입증한다 (도 38E). 종합하여, 이러한 결과로, LITE 가 뉴런에서의 내인성 유전자 발현의 시간적으로 정확한 광학적 제어를 가능하게 한다는 것이 확인된다.

[0644] 출원인의 이전에 실행한 LITE 활성화제에 대한 치하로서, 출원인은 다음으로 TALE-SID4X 구축물을 기반으로 LITE 억제제를 조작하였다. 구성적 Grm2 TALE (T11 및 T12, 도 59A) 은 최고 수준의 전사 억제를 매개하였으며, LITE 억제제로서 선택되었다 (도 59A, B). 광-유도 억제제 모두가 각각 T11 및 T12 에 대해 1.95 배 및 1.75 배 감소로 Grm2 발현의 유의한 하향조절을 매개하였는데, 이는 뉴런에서의 광학적으로 제어된 억제제의 실행 가능성을 입증한다 (도 38G).

[0645] AAV 를 사용하여 뉴런 내에 LITE 를 전달하기 위해, 출원인은 LITE 전이유전자가 포함된 총 바이러스 게놈 크기가 4.8 kb 를 초과하지 않았다는 것을 확실히 해야만 했다 (Wu, Z., Yang, H. & Colosi, P. Effect of Genome Size on AAV Vector Packaging. Mol Ther 18, 80-86 (2009) and Dong JY, F.P., Frizzell RA Quantitative analysis of the packaging capacity of recombinant adeno-associated virus. Human Gene Therapy 7, 2101-2112 (1996)). 이를 위해, 출원인은 TALE N- 및 C-말단을 단축시키고 (N-말단에서 136 aa 를, C-말단에서 63 aa 를 유지), CRY2PHR 및 CIB1 도메인을 교환하였다 (TALE-CIB1 및 CRY2PHR-VP64; 도 38A). 이러한 스위치는 LITE 의 각 성분이 AAV 벡터 내로 들어맞게 하였으며 광-매개 전사 조절의 효능을 감소시키지 않았다 (도 60). 이러한 LITE 는 2 개 AAV 벡터의 조합에 의한 동시-형질도입을 통해 1 차 피질 뉴런 내로 효율적으로 전달될 수 있다 (도 38B; >80% 동시-형질도입 효율을 갖는 2개별 성분에 대해 83-92% 의 전달 효율).

[0646] 실시예 13

[0647] 유도성 렌티바이러스 Cas9

[0648] 렌티바이러스 제조. pCasES10 (렌티바이러스 이동 플라스미드 백본을 함유함) 을 클로닝한 후, 10% 소 태아 혈청을 포함하고 항생제를 포함하지 않는 DMEM 중에서 트랜스펙션하기 전날에 낮은 계대 (p=5) 에서의 HEK293FT 를 T-75 플라스크에 50% 밀집도로 시딩하였다. 20 시간 후, 배지를 OptiMEM (혈청 불포함) 배지로 교환하고 4 시간 후 트랜스펙션을 수행하였다. 세포를 10 ug 의 렌티바이러스 이동 플라스미드 (pCasES10) 및 다음의 패키징 플라스미드로 트랜스펙션하였다: 5 ug 의 pMD2.G (VSV-g 유형) 및 7.5 ug 의 psPAX2 (gag/pol/rev/tat). 양이온 지질 전달제 (50uL 리포펙타민 2000 및 100ul Plus 시약) 와 함께 4 mL OptiMEM 중에서 트랜스펙션을 수행하였다. 6 시간 후, 배지를 10% 소 태아 혈청을 갖는 항생제-불포함 DMEM 으로 교환하였다.

[0649] 렌티바이러스 정제. 바이러스 상청액을 48 시간 후 수확하였다. 상청액을 먼저 잔해가 없도록 하고 0.45um 저 단백질 결합 (PVDF) 필터를 통해 여과하였다. 이를 2 시간 동안 24,000 rpm 에서 초원심분리로 회전시켰다. 바이러스 펠렛을 50ul 의 DMEM 중에서 밤새 4°C 에서 재현탁하였다. 이를 분취하고 즉시 -80°C 에서 동결시켰다.

[0650] FACS 를 사용하는 클론 분리. HEK293FT 및 HUES64 인간 배아 줄기 세포의 클론 단리를 위해, 세포를 현탁액 중 1ul 또는 5ul 의 정제된 바이러스로 감염시켰다. 감염 후 24 시간에, 1uM 독시사이클린을 세포 배양 배지에 첨가하였다. 24 시간 또는 48 시간 이상 후, 세포를 BD FACSAria IIu 기기 상에서 형광-보조 세포 분류 (FACS) 처리하여, 독시사이클린 처리 후 EGFP (및 따라서 Cas9) 를 강건하게 발현한 단일 세포를 단리하였다. 추가 사용을 위해 세포를 벌크로 또는 개별 웰에 플레이팅하여 통합된 유도 Cas9 를 갖는 클론 집단이 선택되게 하였다. 분류 효율은 항상 >95% 였고, EGFP 형광을 확인하기 위해 플레이팅한 직후 세포가

가시화되었다.

[0651] 도 61 은 Tet Cas9 벡터 설계를 나타낸다.

[0652] 도 62 는 293FT 세포에서의 벡터 및 EGFP 발현을 나타낸다.

[0653] pCasES020 유도 Cas9 의 서열:

```

caacttggatagaaaagttggctccgaattgcacctcagggtccgaggttctagacgagtttactccctatcagtgatagaga
acgatgtcgagttactccctatcagtgatagagaacgtatgtcgagttactccctatcagtgatagagaacgtatgtcgagttactccctatc
agtgatagagaacgtatgtcgagttactccctatcagtgatagagaacgtatgtcgagttactccctatcagtgatagagaacgtatgtcgag
gtaggcgtgtcaggtggaggcctatataagcagagctcgtttagtgaaaccgctcagatcgcaaaggcgcaattcgaccecaagttgtacag
ccaccATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTACAA
AGACGATGACGATAAGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGA
GTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGT
GGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCAGCAAGAAATCAAGGTGC
TGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTC
GACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGAT
ACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATG
GCCAAGGTGGACGACAGCTTCTCCACAGACTGGAAGAGTCCTTCCTGGTGAAGA

```

[0654]

```

GGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCT
ACCACGAGAAGTACCCACCATCTACCACCTGAGAAAGAACTGGTGGACAGCACC
GACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCACATGATCAAGTCCGG
GGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCT
GTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTTCGAGGAAAACCCCATCAACG
CCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGG
CTGGAATACTGATCGCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTTCGGCAA
CCTGATTGCCCTGAGCCTGGCCTGACCCCAACTTCAAGAGCAACTTCGACCTGGC
CGAGGATGCCAAACTGCAGCTGAGCAAGGACCTACGACGACGACCTGGACAACC
TGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGT
CCGACGCCATCCTGTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCC
CCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCT
GCTGAAAAGCTCTCGTGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGA
CCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAG
TTCTACAAGTTCATCAAGCCCATCCTGGAAGAAGATGGACGGCACCGAGGAAGTCT
CGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCA
GCATCCCCACAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAA
GATTTTACCATTCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTC
CGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAAACAGCAGATTCGCCTGGAT
GACCAGAAAAGAGCGAGGAAAACCATCACCCCTGGAACCTTCGAGGAAGTGGTGGAC
AAGGGCGCTTCCGCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCT
GCCCAACGAGAAGGTGCTGCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGT
ATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAAGCCCGCCTTC
CTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAA
AGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGTTCGACT
CCGTGGAAAATCTCCGGCGTGAAGATCGGTTCAACGCCTCCCTGGGCACATACCAC
GATCTGCTGAAAATTATCAAGGACAAGGACTTCTGGACAATGAGGAAAACGAGGA
CATTCTGGAAGATATCGTGTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGA
GGAACGGCTGAAAACCTATGCCACCTGTTTCGACGACAAAGTGATGAAGCAGCTGA
AGCGGCGGAGATACCCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATC
CGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCTGAAGTCCGACGGCTTCGC

```

[0655]

CAACAGAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACA  
TCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAAT  
CTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGA  
CGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCCGAGAACATCGTGATCGAAATGG  
CCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAA  
GCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCC  
GTGAAAAACCCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGG  
GCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATG  
TGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGC  
TGACCAGAAGCGACAAGAACCAGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGT  
CGTGAAGAAGATGAAGAATACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCC  
AGAGAAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCTGAGCGAACTGGAT  
AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGT  
GGCACAGATCCTGGACTCCCGGATGAACTAAGTACGACGAGAATGACAAGCTGA  
TCCGGGAAGTGAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAG  
GATTTCCAGTTTTACAAAGTGCGCGAGATCAACAATACTACCACCACGCCACGACGCC  
TACCTGAACGCCGTCTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAG  
CGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGA  
GCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATG  
AACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCT  
GATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTG  
CCACCGTGGGAAAGTGCTGAGCATGCCCAAGTGAATATCGTGAAAAAGACCGAG  
GTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCAAGAGGAACAGCGATAA  
GCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCC  
CCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAAGTGAAAAAGGGCAAGTCCAAG  
AAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTT  
CGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGG  
ACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAG  
AGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTC  
CAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCC  
CGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAAGCACTACCTGGACG

[0656]

AGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAAT  
 CTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCA  
 GGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTT  
 CAAGTACTTTGACACCACCATCGACCAGGAAAGAGGTACACCAGCACCAAAGAGGTGC  
 TGGACGCCACCCTGATCCACCAGAGCATCACCGCCTGTACGAGACACGGATCGAC  
 CTGTCTCAGCTGGGAGGCGACAAAAGGCCGGCGGCCACGAAAAGGCCGGCCAGG  
 CAAAAAGAAAAAGgaattctctagaGGCAGTGGAGAGGGCAGAGGAAGTCTGCTAACAT  
 GCGGTGACGTCGAGGAGAATCCTGGCCCAGTGAGCAAGGGCGAGGAGCTGTTTACC  
 GGGGTGGTGCCCATCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAG  
 CGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCA  
 TCTGCACCACGGCAAGCTGCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCT  
 ACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCA  
 AGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGAC  
 GGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG  
 CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGC  
 TGGAGTACAACACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAAC  
 GGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT  
 CGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCG  
 ACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGC  
 GATCACATGGTCTCTGCTGGAGTTCTGTGACCGCCCGGGATCACTCTCGGCATGGAC  
 GAGCTGTACAAGCTCGAGGGAAGCGGAGCTACTAAGTTCAGCCTGCTGAAGCAGGC  
 TGGCGACGTGGAGGAGAACCCTGGACCTatgtctaggctggacaagagcaagtataaacggagctctggaa  
 ttactcaatggtgtcggatcgaaggcctgacgacaaggaaactgctcaaaagctgggagttgagcagcctaccctgtactggcacgtga  
 agaacaagcgggcccctgctcgtatgccctgccaatcgagatgctggacagcctacataccacttctgcccctggaaggcagtcattggc  
 aagacttctgcggaacaacgccaagtcataccgctgtgtctctcctcacaatcgcgacggggtaaaagtcatctcggcaccgcccac  
 agagaacagctacgaaacctggaatacagctcgcgttctgtgtcagcaaggcttctcctggagaacgactgtacgctctgtccg  
 gtggccactttactggtgctgtaftggagggaacagagcatcaagtagcaaaagaggaaaagagacacctaccaccgattctatg  
 cccccacttctgagacaagcaattgagctgttcgaccggcaggagccgaacctgcttcttctggcctggaactaatcatatgtgcctg  
 gagaacagctaaagtgcgaaagcggcggccgaccgagcccctgacgatttgacttagacatgctcccagccgatcccctgacgatt  
 ttgacctgacatgctccccgggtaatgtacaagtggtgaattccgcaattcgatatcaagcttatcgataatcaacctctggattacaaat  
 tgtgaaagattgactggtattcttaactatgtgtccttttacgctatgtggatacctgcttfaatgcctttgtatcatgctattgctcccgtatgg

[0657]

ctttcattttctcctctgtataaaactggtgctgtctctttatgaggagttgtggcccggtcaggaacgtggcgtggtgactgtgtttg  
 ctgacgcaacccccactggttggggcattgccaccacctgtcagctccttccgggactttcgcttccccctccctattgccacggcgaac  
 tcatcgcgcctgcttcccgtctggacaggggctcggtgttgggactgacaattccggtgtgtcgggaaatcatctcttcc  
 ctggctgctcgcctgtgttggccacctggattctgcgcgggacgtcctctctctacgtccctcggccctcaatccagcggacctctctccg  
 cggcctgctccggctctcggcctcttccgctcttgcctcgcctcagacgagtcggatctcctttggccgctccccgcatcgata  
 ccgtcgacctgagacctgaaaaacatggagcaatcacaagtagcaatacagcagctaccaatgctgattgtgcttgctagaagcaca  
 agaggaggaggaggtgggtttccagtcacacctcaggfaccttfaagcaatgactacaaggcagctgtagatcttagccattttaa  
 agaaaaagggggactggaaaggctaattcactccaacgaagacaagatattctgactgtggatctaccacacaaggctacttccct  
 gattggcagaactacacaccaggcggcaggatcagatattcactgaccttggatgggtactacaagtagtaccagttgagcaagagaag  
 gtagaagaagccaatgaaggagagaacaccgcttggtaacctgtgagcctgcatgggatggatgacccggagagagaagtattaga  
 gtggagggttgcagccgctgacttcatcatatgcccggagagctgcatccggactgtactgggtctctctgttagaccagatctgagc  
 ctgggagctctctggtaactggaaccactgcttgaacctcaataaagcttgcctgagtgctcaagtagtgtgtgcccgtctgtgtgtg  
 actctgtaactagagatcctcagaccctttatgctagtggtgaaaatctctagcagggcccgttfaaccctgctgatcagcctcactgtgc  
 cttctagttgccacctgtgttggctccccctgcttcccttaccctggaaagggtgccactcccactgtccttctataaaatgagg  
 aattgcatcgattgtctgtagtggtgtcattctattctgggggtgggggtggggcaggacagcaagggggaggattgggaagacaatag  
 caggcatctgggggatgggtggctctatgcttctgaggggaaagaaccagctggggctctaggggatccccacgcgcctgta  
 gggcgcatataagcgcgggggtgtgggtgtacgcgagcgtgaccgctacacttgcagcgcctagcggccctccttctccttctc  
 ccttcttctcgcacgtcggcgttccccgtcaagctcaaatcggggctcctttagggttccgatttagtctttacggcacctcgac  
 ccaaaaaactgattagggtgatgttccagtagtggccatcgcctgatagacgggttttcccttggacttggagtcacagttctta  
 agtggactcttgttcaaaactggaacaactcaacctatctcgtctattctttgattataagggttttggcatttccgctattggtaaa  
 aatgagctgatttaacaaaaaattaacgcaatfaattctgtggaatgtgtgtcagttagggtgtgaaagtccccaggctcccagcaggc  
 agaagtatgcaaaatgcatctcaattagtcagcaaccaggtgtgaaagtccccaggctccccagcagagaagtatcaaaatgcatg  
 catctcaattagtcagcaaccatagctcccccctaacctccatcccccctaacctccgcccagttccgcccattctcccccattggc  
 tgaactaattttttttatgacagggccgagccctcctcctctgagctattccagaagtgtgaggaggctttttggaggcctaggctttt  
 gcaaaaaactccccggagctgtatcattttcggatctgatcagcagctgttgaacaattaatcagcagatagatagcagatagataat  
 acgacaagggtgaggaaactaaacctggccaagttgaccagtccttccgggtctcaccgcgcgagctcgggagcgggtcgagttc  
 tggaccgaccggctcgggttctccgggactctggaggacgacttccgggtgtgtccgggacgagctgacctgttcatcagcgcg  
 gtccaggaccaggtgttccggacaacacctggcctgggtgtgggtgctcggcctggagagctgtacggaggtgtcggagggtcg  
 tgtccacgaactccggagcctccggcggccatgaccgagatcggcagcagcgggtggggcgggagttcgcctcgcgcgacc  
 cggccggcaactcgtgactctgtgcccagagcaggactgacagctgctacgagattcattccaccgcccctctatgaaagggtt  
 gggctcggaaatcgttttccgggacgccggctgagatctcagcgcgggatctcatgtggagttcttcccacccaactgtttatt

[0658]

gcagcttataatgggttacaataaagaatagcatcacaattcacaataaagcatttttctactgattctagttggtgttgcctcaactcat  
 caatgatcttatcatgtctgtataacctgacgaccttagctagagcttgccgtaatcatggtcatagctgtttcctgtgtgaaattgtatccgctca  
 caattccacacaacatagcagccggaagcataaagftaaagcctgggggcttaatgagtgagtaactcacattaattgctgtgctca  
 ctgcccgctttccagtcgggaaacctgtctgaccagctgcattaatgaatcggccaacgcgaggagagggctgttctgtattggcgct  
 cttccgctctcctcactgactcgtcgtcgtcgttctggctgctggcgagcgggtatcagctcactcaaaaggcgttaatacgggtatcca  
 cagaatcaggggataacgcaggaaagaacatgtgagcaaaaaggccagcaaaaaggccaggaaccgtaaaaaggccgctgtctggcgtt  
 ttccataggtccgccccctgacgagcatcacaataatcgacgctcaagtcaaggtggcgaacccgacaggactataaagatacca  
 ggctgttccccctggaagctcctctgctcgtctcctgttccgacctgccgttaccggatacctgtccgctttctccctcgggaagcgtg  
 gcgctttctcatagctcagctgtaggtatctcagttcgggtgtagctgtctcctcaagctgggctgtgtcacgaacccccgttccagccc  
 gaccgctgcgcttaccgtaactatctgtctgagtcacccggtaagacacgacttatgccactggcagcagccactgtaacaggat  
 tagcagagcggaggtatgtaggcgggtctacagagttctgaagtgggtgcttaactcggctactactagaagaacagattttgtatctgcg  
 ctctgtgaagccagttacctcggaaaaagagttgtagctcttgatccggcaaaaaccaccgctgtagcgtggtttttgtttgcaa  
 gcagcagattaccgagaaaaaaggatctcaagaagatctttgatctttctacgggctgacgctcagtggaacgaaaaactacgtt  
 aagggattttgctatgagattcaaaaaggatctcacctagatctttaaattaaaatgaagtttaatacaactaaagtatatatgagtaa  
 actggtctgacagttaccaatgcttaacagtgaggcacctatctcagcagatctgtctatcttctcattccatagttgctgactccccgtcgtg  
 agataactacgatacgggagggcttaccatctgccccagtgctgcaatgataaccgagaccacgctcaccggctccagattatcagc  
 aataaacgaccgagcgggagggcggcagcagaagtggctcctcaactttaccgctccatcagctatattaattgtgccgggaagcta  
 gagtaagtagttccagtttaatagttgcaacgttggcttctgctacagcagcagctggtgtcacgctcgtcgtttgtatggcttca  
 gctccggtcccaacgatcaaggcaggtatcatgacccccatgttgcaaaaaagcgggttagctcctcctggctcccgatcgtgtcagaa  
 gtaagttggcccaggttatcactcatggttatgacagcactgcataatctctactgcatgccatccgtaagatctttctgtgactggtga  
 gfactcaaccaagtcattctgagaatagtgatgacggcgaccgagttgctcttccccggcgtcaatcgggataataccgcccacatagca  
 gaactttaaaggtctcatcattgaaaaagcttctggggcgaaaaactcfaaggatctaccgctgttgagatcagttcagtaaccaccat  
 cgtgcacccaactgatctcagcatctttactttaccagcgttctgggtgagcaaaaacaggaaggcaaaatcggcaaaaaagggaat  
 aagggcgacacggaatgtgaaactacatactctcttttcaatattgaagcattatcaggggttattgtctcatgagcggatacatattg  
 aatgtatttagaaaaatacaaataggggttccgcccacatccccgaaaagtgccacctgacgtcagcggatcgggagatctcccgatc  
 ccctatgggtgactctcagtaacaatctgctctgafccgcatagtaagccagtatctctccctgctgtgttggaaggtcgtgagtagtgc  
 gcgagcaaaatfaagctacaacaaggcgaagcgtgaccgacaattgcatgaaatctgcttagggtaggcgttttgcgctctcgcgat  
 gtacgggcccagatatacgcgttgacattgattatgactagtattaaatagtaatacaattacggggtcattatgctatgcccataatgagttcc  
 gcgttacataactacgtaaatggcccgcctggctgaccgccaacgacccccgccattgacgtcaataatgacgtatgtcccatagta  
 acgccaataggactttcattgacgtcaatgggtggagattttacggtaaacctcccacttggcagtatcaagtgatcatatgccaagta  
 cggccccctattgacgtcaatgacggtaaatggccccctggcattatgcccagtatcatgacctatgggacttctacttggcagtatcatcta

[0659]

cgattagtcacgctattaccatgggtgatcgcgffttggcagtcacatcaatggcgctggatagcggttgactcacgggatttccaagtctcc  
 accccattgacgtcaatgggagttgtttggcacaaaatcaacgggactttccaaaatgctgtaacaactccgccccattgacgcaaatgg  
 gcggtagggctgtacgggtgggaggtctatataaagcagcgctgttgcctgtactgggtctctctggttagaccagatctgagcctgggagctc  
 tctggctaactagggaaacctgcttaagcctcaataaagcttgccttgagtgtctcaagtagtgtgtgccctctgtgtgtgactctggtaa  
 ctagatccctcagacccttttagctagtgtaaactctctagcagtgccgcccgaacagggacttgaagcgaagggaaccagag  
 gagctctctcagcagggactcggctgtgaaagcgcgcacggcaagagcgcgagggcgcgactggtgagtagcgaataattttgac  
 tagcggaggctagaaggagagatgggtgagagcgtcagtagtaagcggggagaatagatcgcgatgggaaaaattcgggttaa  
 ggccaggggggaaagaaaaataataaataatagatgtggcaagcagggagctagaacgattcgaattaatcctggctgttaga  
 aacatcagaaggctgtagacaaactgggacgctacaacctccctcagacagatcagaagaacttagatcattatataatcacgtag  
 caacctctattgtgtcatcaaaagtagagataaaagacaccaaggaagctttagacaagatagaggaaagcgaatacaaaagtaaga  
 ccaccgcagcagcagggcctgctgcttgcacacctggaggagagatagagggacaattggagaagtgaattatataataaagta  
 gtaaaaattgaaccattaggtagcaccaccaaggcaaaagagaagtggtgagcagagagaaaaagagcagtggaataggagcttt  
 gttcctgggtcttgggagcagcaggaagcactatggcgagcgtcaatgacgctgacggtagcagccagacaattattgtctgtatag  
 tgcagcagcagaacaattgtgtagggctattgagcgcacacagatctgtgcaactcacagcttggggcatcaagcagctccaggcaa  
 gaatcctgctgtggaagatacctaaagatcaacagctcctgggatttggggtgctctgaaaactatttgcaccactgctgtgcctt  
 ggaatgctagtgtgagtaataatctctggaacagatttggaaatcacagcctggatggagtgaggacagagaataacaattacaagc  
 ttaatacactccttaattgaagaatcgcaaaaccgaagaaagaatgaacaagaatttgaatagataaatggcaagttgtggaatt  
 ggttaacatacaaaattggctgtgtatataaaattattcataatgatagtaggagcttggtaggttaagaatagttttgctgtacttctatag  
 tgaatagagttaggcagggatattaccattatcgtttcagaccacctccaacccgaggggacccagagcccgaaggaatagaag  
 aagaaggtggagagagacagagacagatcattcagattagtgaaacggatcggcactgctgccaattctcagacaaatggcagat  
 attcatcccaattttaaagaaaagggggattgggggtacagtcaggggaaagaatagtagacataatgcaacagacatacaaaact  
 aaagaattacaacaaatacaaaattcaaaatttccgggttattacagggacagcagagatccagtttgggttaattaa

[0660]

[0661]

[0662]

[0663]

[0664]

[0665]

[0666]

[0667]

[0668]

[0669]

[0670]

[0671]

1. 하기를 포함하는, 세포에서의 관심 게놈 유전자좌의 발현을 변형하는 유도 방법:

(a) (i) 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 개 이상의 전사 활성제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 데옥시리보핵산 (DNA) 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너와 결합할 수 있음), 및 /또는

(ii) 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 상호작용 파트너에 결합함)

을 포함하는 DNA 결합 폴리펩티드를 포함하는 비-자연발생적 또는 조작된 조성물과 게놈 유전자좌를 접촉시키고;

(b) 에너지원을 적용시키고;

(c) 게놈 유전자좌의 발현이 변형된 것을 측정함.

2. 단락 1 에 있어서, 적어도 하나 이상의 효과기 도메인이 트랜스포사아제 도메인, 인테그라아제 도메인, 재조합효소 도메인, 리졸바아제 도메인, 인버타아제 도메인, 프로테아제 도메인, DNA 메틸트랜스퍼라아제 도메인, DNA 데메틸라아제 도메인, 히스톤 아세틸라아제 도메인, 히스톤 데아세틸라아제 도메인, 뉴클레아제 도메인, 억제제 도메인, 활성제 도메인, 핵-국부화 신호 도메인, 전사-단백질 영입 도메인, 세포 흡수 활성 관련 도메인, 핵산 결합 도메인 또는 항체 제시 도메인으로 이루어지는 군에서 선택되는 방법.

3. 단락 2 에 있어서, 적어도 하나 이상의 효과기 도메인이 뉴클레아제 도메인 또는 재조합효소 도메인인 방법.

4. 단락 3 에 있어서, 뉴클레아제 도메인이 비-특이적 FokI 엔도뉴클레아제 측매 도메인인 방법.

5. 단락 1-4 중 어느 하나에 있어서, 에너지 민감성 단백질이 크립토크롬-2 (CRY2) 인 방법.

6. 단락 1-5 중 어느 하나에 있어서, 상호작용 파트너가 크립토크롬-상호작용 염기성 헬릭스-루프-헬릭스 (CIB1) 인 방법.

- [0672] 7. 단락 1-6 중 어느 하나에 있어서, 에너지원이 전자기 조사, 음파 에너지 또는 열 에너지로 이루어지는 군에서 선택되는 방법.
- [0673] 8. 단락 7 에 있어서, 전자기 조사가 가시광의 성분인 방법.
- [0674] 9. 단락 8 에 있어서, 가시광 성분이 450nm-500nm 범위의 파장을 갖는 방법.
- [0675] 10. 단락 8 에 있어서, 가시광 성분이 청색광인 방법.
- [0676] 11. 단락 1 에 있어서, 에너지를 적용하는 것이  $6.2 \text{ mW/cm}^2$  이상 강도로의 청색광 자극을 포함하는 방법.
- [0677] 12. 단락 1-11 중 어느 하나에 있어서, DNA 결합 도메인이 ( $X_{1-11}-X_{12}X_{13}-X_{14-33}$  또는  $^{34}$  또는  $^{35}$ ) $z$  를 포함하는 방법으로 서, 여기서  $X_{1-11}$  이 11 개의 인접한 아미노산 사슬이고,  $X_{12}X_{13}$  이 반복 가변 이중잔기 (RVD) 이고,  $X_{14-33}$  또는  $^{34}$  또는  $^{35}$  가 21, 22 또는 23 개의 인접한 아미노산 사슬이고,  $z$  가 적어도 5 내지 40 이고, 하나 이상의 RVD 가 NI, HD, NG, NN, KN, RN, NH, NQ, SS, SN, NK, KH, RH, HH, HI, KI, RI, SI, KG, HG, RG, SD, ND, KD, RD, YG, HN, NV, NS, HA, S\*, N\*, KA, H\*, RA, NA 및 NC 로 이루어지는 군에서 선택되고, 이때 (\*) 가  $X_{13}$  에서의 아미노산이 부재함을 의미하는 방법.
- [0678] 13. 단락 12 에 있어서,  $z$  가 적어도 10 내지 26 인 방법.
- [0679] 14. 단락 12 에 있어서:
- [0680]  $X_{1-11}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}-X_{14-34}$  또는  $X_{1-11}-X_{14-35}$ ) 에서의 아미노산 1-11 로 나타내는 12 개의 인접한 아미노산의 서열이거나,
- [0681]  $X_{14-34}$  또는  $X_{14-35}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}-X_{14-34}$  또는  $X_{1-11}-X_{14-35}$ ) 에서의 아미노산 12-32 또는 12-33 으로 나타내는 21 또는 22 개의 인접한 아미노산의 서열인 방법.
- [0682] 15. 단락 12 에 있어서, 하나 이상의 RVD 가 (a) 구아닌 (G) 의 인지를 위한 HH, KH, NH, NK, NQ, RH, RN, SS; (b) 아데닌 (A) 의 인지를 위한 SI; (c) 티민 (T) 의 인지를 위한 HG, KG, RG; (d) 시토신 (C) 의 인지를 위한 RD, SD; (e) A 또는 G 의 인지를 위한 NV; 및 (f) A 또는 T 또는 G 또는 C 의 인지를 위한 H\*, HA, KA, N\*, NA, NC, NS, RA, S\* (여기서 (\*) 는  $X_{13}$  에서의 아미노산이 부재함을 의미함) 로 이루어지는 군에서 선택되는 방법.
- [0683] 16. 단락 15 에 있어서:
- [0684] G 의 인지를 위한 RVD 가 RN, NH, RH 또는 KH 이거나;
- [0685] A 의 인지를 위한 RVD 가 SI 이거나;
- [0686] T 의 인지를 위한 RVD 가 KG 또는 RG 이고;
- [0687] C 의 인지를 위한 RVD 가 SD 또는 RD 인 방법.
- [0688] 17. 단락 12 에 있어서, 하기 중 하나 이상이 존재하는 방법:
- [0689]  $X_{1-4}$  에서의 [LTLD] 또는 [LTLA] 또는 [LTQV], 또는
- [0690] 위치  $X_{30-33}$  또는  $X_{31-34}$  또는  $X_{32-35}$  에서의 [EQHG] 또는 [RDHG].
- [0691] 18. 단락 1-17 중 어느 하나에 있어서:
- [0692] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 147 개의 인접한 아미노산을 포함하거나,
- [0693] C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 68 개의 인접한 아미노산을 포함하거나,
- [0694] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 136 개의 인접한 아미노산을 포함하고 C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 183 개의 인접한 아미노산을 포함하는 방법.
- [0695] 19. 단락 1-18 중 어느 하나에 있어서, 관심 게놈 유전자좌가 분화 인자, 전사 인자, 신경전달물질 이동체, 신경전달물질 신타아제, 시냅스 단백질, 가소성 단백질, 시냅스전 활성 구역 단백질, 시냅스후 밀도 단백질, 신경

전달물질 수용체, 후생적 변형제, 신경 운명 선별화 인자, 액손 안내 분자, 이온 채널, CpG 결합 단백질, 유비퀴틴화 단백질, 호르몬, 호메오박스 단백질, 성장 인자, 발암유전자 또는 원발암유전자를 인코딩하는 유전자와 관련된되는 방법.

- [0696] 20. 하기를 포함하는, 세포에서의 관심 게놈 유전자좌의 발현을 억제하는 유도 방법:
- [0697] (a) (i) 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 개 이상의 전사 활성제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너와 결합할 수 있음), 및/또는
- [0698] (ii) 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 상호작용 파트너에 결합함)
- [0699] 을 포함하는 DNA 결합 폴리펩티드를 포함하는 비-자연발생적 또는 조작된 조성물과 게놈 유전자좌를 접촉시키고;
- [0700] (b) 에너지원을 적용시키고;
- [0701] (c) 게놈 유전자좌의 발현이 억제된 것을 측정함.
- [0702] 21. 단락 20 에 있어서, 폴리펩티드가 하나 이상의 SID 억제제 도메인을 포함하는 방법.
- [0703] 22. 단락 21 에 있어서, 폴리펩티드가 4 개 이상의 SID 억제제 도메인을 포함하는 방법.
- [0704] 23. 단락 21 에 있어서, 폴리펩티드가 SID4X 억제제 도메인을 포함하는 방법.
- [0705] 24. 단락 20 에 있어서, 폴리펩티드가 KRAB 억제제 도메인을 포함하는 방법.
- [0706] 25. 단락 20-24 중 어느 하나에 있어서, 에너지 민감성 단백질이 크립토크롬-2 (CRY2) 인 방법.
- [0707] 26. 단락 20-25 중 어느 하나에 있어서, 상호작용 파트너가 크립토크롬-상호작용 염기성 헬릭스-루프-헬릭스 (CIB1) 인 방법.
- [0708] 27. 단락 20-26 중 어느 하나에 있어서, 에너지원이 전자기 조사, 음파 에너지 또는 열 에너지로 이루어지는 군에서 선택되는 방법.
- [0709] 28. 단락 20 에 있어서, 전자기 조사가 가시광의 성분인 방법.
- [0710] 29. 단락 28 에 있어서, 가시광 성분이 450nm-500nm 범위의 파장을 갖는 방법.
- [0711] 30. 단락 28 에 있어서, 가시광 성분이 청색광인 방법.
- [0712] 31. 단락 20 에 있어서, 에너지원을 적용하는 것이  $6.2 \text{ mW/cm}^2$  이상 강도로의 청색광 자극을 포함하는 방법.
- [0713] 32. 단락 20-31 중 어느 하나에 있어서, DNA 결합 도메인이  $(X_{1-11}-X_{12}X_{13}-X_{14-33}$  또는  $34$  또는  $35$ ) $Z$  를 포함하는 방법으로서, 여기서  $X_{1-11}$  이 11 개의 인접한 아미노산 사슬이고,  $X_{12}X_{13}$  이 반복 가변 이중잔기 (RVD) 이고,  $X_{14-33}$  또는  $34$  또는  $35$  가 21, 22 또는 23 개의 인접한 아미노산 사슬이고,  $Z$  가 적어도 5 내지 40 이고, 하나 이상의 RVD 가 NI, HD, NG, NN, KN, RN, NH, NQ, SS, SN, NK, KH, RH, HH, HI, KI, RI, SI, KG, HG, RG, SD, ND, KD, RD, YG, HN, NV, NS, HA, S\*, N\*, KA, H\*, RA, NA 및 NC 로 이루어지는 군에서 선택되고, 이때 (\*) 가  $X_{13}$  에서의 아미노산이 부재함을 의미하는 방법.
- [0714] 33. 단락 32 에 있어서,  $Z$  가 적어도 10 내지 26 인 방법.
- [0715] 34. 단락 32 에 있어서:
- [0716]  $X_{1-11}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}-X_{14-34}$  또는  $X_{1-11}-X_{14-35}$ ) 에서의 아미노산 1-11 로 나타내는 11 개의 인접한 아미노산의 서열이거나,
- [0717]  $X_{14-34}$  또는  $X_{14-35}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}-X_{14-34}$  또는  $X_{1-11}-X_{14-35}$ ) 에서의 아미노산 12-32 또는 12-33

으로 나타내는 21 또는 22 개의 인접한 아미노산의 서열인 방법.

- [0718] 35. 단락 20-34 중 어느 하나에 있어서:
- [0719] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 147 개의 인접한 아미노산을 포함하거나,
- [0720] C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 68 개의 인접한 아미노산을 포함하거나,
- [0721] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 136 개의 인접한 아미노산을 포함하고 C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 183 개의 인접한 아미노산을 포함하는 방법.
- [0722] 36. 단락 20-35 중 어느 하나에 있어서, 관심 게놈 유전자좌가 분화 인자 또는 이온 채널 성분을 인코딩하는 유전자와 관련되는 방법.
- [0723] 37. 단락 36 에 있어서, 분화 인자가 SRY-박스-2 (SOX2) 이며 유전자 SOX2 에 의해 인코딩되는 방법.
- [0724] 38. 단락 36 에 있어서, 분화 인자가 p11 이며 유전자 p11 에 의해 인코딩되는 방법.
- [0725] 39. 단락 36 에 있어서, 이온 채널 성분이 CACNA1C 이며 유전자 CACNA1C 에 의해 인코딩되는 방법.
- [0726] 40. 하기를 포함하는, 세포에서의 관심 게놈 유전자좌의 발현을 활성화하는 유도 방법:
- [0727] (a) (i) 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 개 이상의 TALE 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너와 결합할 수 있음), 및
- [0728] (ii) 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 상호작용 파트너에 결합함)
- [0729] 을 포함하는 DNA 결합 폴리펩티드를 포함하는 비-자연발생적 또는 조작된 조성물과 게놈 유전자좌를 접촉시키고;
- [0730] (b) 에너지원을 적용시키고;
- [0731] (c) 게놈 유전자좌의 발현이 활성화된 것을 측정함.
- [0732] 41. 단락 40 에 있어서, 폴리펩티드가 하나 이상의 VP16 또는 VP64 활성화제 도메인을 포함하는 방법.
- [0733] 42. 단락 40 에 있어서, 폴리펩티드가 하나 이상의 p65 활성화제 도메인을 포함하는 방법.
- [0734] 43. 단락 40-42 중 어느 하나에 있어서, 에너지 민감성 단백질이 CRY2 인 방법.
- [0735] 44. 단락 40-43 중 어느 하나에 있어서, 상호작용 파트너가 CIB1 인 방법.
- [0736] 45. 단락 40 에 있어서, 에너지원이 전자기 조사, 음파 에너지 또는 열 에너지로 이루어지는 군에서 선택되는 방법.
- [0737] 46. 단락 45 에 있어서, 전자기 조사가 가시광의 성분인 방법.
- [0738] 47. 단락 46 에 있어서, 가시광 성분이 450nm-500nm 범위의 파장을 갖는 방법.
- [0739] 48. 단락 46 에 있어서, 가시광 성분이 청색광인 방법.
- [0740] 49. 단락 40 에 있어서, 에너지원을 적용하는 것이  $6.2 \text{ mW/cm}^2$  이상 강도로의 청색광 자극을 포함하는 방법.
- [0741] 50. 단락 40-49 중 어느 하나에 있어서, DNA 결합 도메인이  $(X_{1-11}-X_{12}X_{13}-X_{14-33}$  또는  $34$  또는  $35$ ) $_z$  를 포함하는 방법으로서, 여기서  $X_{1-11}$  이 11 개의 인접한 아미노산 사슬이고,  $X_{12}X_{13}$  이 반복 가변 이중잔기 (RVD) 이고,  $X_{14-33}$  또는  $34$  또는  $35$  가 21, 22 또는 23 개의 인접한 아미노산 사슬이고,  $z$  가 적어도 5 내지 40 이고, 하나 이상의 RVD 가 NI, HD, NG, NN, KN, RN, NH, NQ, SS, SN, NK, KH, RH, HH, HI, KI, RI, SI, KG, HG, RG, SD, ND, KD, RD, YG, HN, NV, NS, HA, S\*, N\*, KA, H\*, RA, NA 및 NC 로 이루어지는 군에서 선택되고, 이때 (\*) 가  $X_{13}$  에서의 아미노산이 부재함을 의미하는 방법.

- [0742] 51. 단락 50 에 있어서,  $z$  가 적어도 10 내지 26 인 방법.
- [0743] 52. 단락 50 에 있어서:
- [0744]  $X_{1-11}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}$ - $X_{14-34}$  또는  $X_{1-11}$ - $X_{14-35}$ ) 에서의 아미노산 1-11 로 나타내는 11 개의 인접한 아미노산의 서열이거나,
- [0745]  $X_{14-34}$  또는  $X_{14-35}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}$ - $X_{14-34}$  또는  $X_{1-11}$ - $X_{14-35}$ ) 에서의 아미노산 12-32 또는 12-33 으로 나타내는 21 또는 22 개의 인접한 아미노산의 서열인 방법.
- [0746] 53. 단락 40-52 중 어느 하나에 있어서:
- [0747] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 147 개의 인접한 아미노산을 포함하거나,
- [0748] C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 68 개의 인접한 아미노산을 포함하거나,
- [0749] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 136 개의 인접한 아미노산을 포함하고 C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 183 개의 인접한 아미노산을 포함하는 방법.
- [0750] 54. 단락 40-53 중 어느 하나에 있어서, 관심 게놈 유전자좌가 분화 인자, 후생적 변형체 또는 이온 채널 성분을 인코딩하는 유전자와 관련되는 방법.
- [0751] 55. 단락 54 에 있어서, 분화 인자가 뉴로제닌-2 이며 유전자 NEUROG2 에 의해 인코딩되는 방법.
- [0752] 56. 단락 54 에 있어서, 분화 인자가 크루펠 유사 인자 4 이며 유전자 KLF-4 에 의해 인코딩되는 방법.
- [0753] 57. 단락 54 에 있어서, 후생적 변형체가 Tet 메틸시토신 디옥시메틸아제 1 이며 유전자 tet-1 에 의해 인코딩되는 방법.
- [0754] 58. 단락 54 에 있어서, 이온 채널 성분이 CACNA1C 이며 유전자 CACNA1C 에 의해 인코딩되는 방법.
- [0755] 59. 세포에서의 게놈 유전자좌의 발현을 유도가능하게 변형하기 위한 비-자연발생적 또는 조작된 조성물로서, 상기 조성물이 하기를 포함하는 DNA 결합 폴리펩티드를 포함하고:
- [0756] (i) 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너와 결합할 수 있음), 및/또는
- [0757] (ii) 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 상호작용 파트너에 결합함);
- [0758] 상기 폴리펩티드가 코돈 최적화된 핵산 분자에 의해 인코딩되고 이로부터 번역되어, 폴리펩티드가 바람직하게는 게놈 유전자좌의 DNA 에 결합하고,
- [0759] 상기 폴리펩티드가 에너지원의 적용시 게놈 유전자좌의 발현을 변형시키는 조성물.
- [0760] 60. 단락 59 에 있어서, 적어도 하나 이상의 효과기 도메인이 트랜스포사아제 도메인, 인테그라아제 도메인, 재조합효소 도메인, 리졸바아제 도메인, 인버타아제 도메인, 프로테아제 도메인, DNA 메틸트랜스퍼라아제 도메인, DNA 데메틸라아제 도메인, 히스톤 아세틸라아제 도메인, 히스톤 데아세틸라아제 도메인, 뉴클레아제 도메인, 억제제 도메인, 활성화제 도메인, 핵-국부화 신호 도메인, 전사-단백질 영입 도메인, 세포 흡수 활성화 관련 도메인, 핵산 결합 도메인 또는 항체 제시 도메인으로 이루어지는 군에서 선택되는 조성물.
- [0761] 61. 단락 59 에 있어서, 적어도 하나 이상의 효과기 도메인이 뉴클레아제 도메인 또는 재조합효소 도메인인 조성물.
- [0762] 62. 단락 61 에 있어서, 뉴클레아제 도메인이 비-특이적 FokI 엔도뉴클레아제 촉매 도메인인 조성물.
- [0763] 63. 단락 59-62 중 어느 하나에 있어서, 에너지 민감성 단백질이 크립토크롬-2 (CRY2) 인 조성물.
- [0764] 64. 단락 59-63 중 어느 하나에 있어서, 상호작용 파트너가 크립토크롬-상호작용 염기성 헬릭스-루프-헬릭스 (CIB1) 인 조성물.

- [0765] 65. 단락 59-64 중 어느 하나에 있어서, 에너지원이 전자기 조사, 음파 에너지 또는 열 에너지로 이루어지는 군에서 선택되는 조성물.
- [0766] 66. 단락 65 에 있어서, 전자기 조사가 가시광의 성분인 조성물.
- [0767] 67. 단락 66 에 있어서, 가시광 성분이 450nm-500nm 범위의 파장을 갖는 조성물.
- [0768] 68. 단락 66 에 있어서, 가시광 성분이 청색광인 조성물.
- [0769] 69. 단락 59 에 있어서, 에너지원을 적용하는 것이  $6.2 \text{ mW/cm}^2$  이상 강도로의 청색광 자극을 포함하는 조성물.
- [0770] 70. 단락 59-69 중 어느 하나에 있어서, DNA 결합 도메인이 ( $X_{1-11}$ - $X_{12}X_{13}$ - $X_{14-33}$  또는  $^{34}$  또는  $^{35}$ )z 를 포함하는 조성물로서, 여기서  $X_{1-11}$  이 11 개의 인접한 아미노산 사슬이고,  $X_{12}X_{13}$  이 반복 가변 이중잔기 (RVD) 이고,  $X_{14-33}$  또는  $^{34}$  또는  $^{35}$  가 21, 22 또는 23 개의 인접한 아미노산 사슬이고, z 가 적어도 5 내지 40 이고, 하나 이상의 RVD 가 NI, HD, NG, NN, KN, RN, NH, NQ, SS, SN, NK, KH, RH, HH, HI, KI, RI, SI, KG, HG, RG, SD, ND, KD, RD, YG, HN, NV, NS, HA, S\*, N\*, KA, H\*, RA, NA 및 NC 로 이루어지는 군에서 선택되고, 이때 (\*) 가  $X_{13}$  에서의 아미노산이 부재함을 의미하는 조성물.
- [0771] 71. 단락 70 에 있어서, z 가 적어도 10 내지 26 인 조성물.
- [0772] 72. 단락 70 에 있어서:
- [0773]  $X_{1-11}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}$ - $X_{14-34}$  또는  $X_{1-11}$ - $X_{14-35}$ ) 에서의 아미노산 1-11 로 나타내는 12 개의 인접한 아미노산의 서열이거나,
- [0774]  $X_{14-34}$  또는  $X_{14-35}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}$ - $X_{14-34}$  또는  $X_{1-11}$ - $X_{14-35}$ ) 에서의 아미노산 12-32 또는 12-33 으로 나타내는 21 또는 22 개의 인접한 아미노산의 서열인 조성물.
- [0775] 73. 단락 70 에 있어서, 하나 이상의 RVD 가 (a) 구아닌 (G) 의 인지를 위한 HH, KH, NH, NK, NQ, RH, RN, SS; (b) 아데닌 (A) 의 인지를 위한 SI; (c) 티민 (T) 의 인지를 위한 HG, KG, RG; (d) 시토신 (C) 의 인지를 위한 RD, SD; (e) A 또는 G 의 인지를 위한 NV; 및 (f) A 또는 T 또는 G 또는 C 의 인지를 위한 H\*, HA, KA, N\*, NA, NC, NS, RA, S\* (여기서 (\*) 는  $X_{13}$  에서의 아미노산이 부재함을 의미함) 로 이루어지는 군에서 선택되는 조성물.
- [0776] 74. 단락 73 에 있어서:
- [0777] G 의 인지를 위한 RVD 가 RN, NH, RH 또는 KH 이거나;
- [0778] A 의 인지를 위한 RVD 가 SI 이거나;
- [0779] T 의 인지를 위한 RVD 가 KG 또는 RG 이고;
- [0780] C 의 인지를 위한 RVD 가 SD 또는 RD 인 조성물.
- [0781] 75. 단락 70 에 있어서, 하기 중 하나 이상이 존재하는 조성물:
- [0782]  $X_{1-4}$  에서의 [LTLD] 또는 [LTLA] 또는 [LTQV], 또는
- [0783] 위치  $X_{30-33}$  또는  $X_{31-34}$  또는  $X_{32-35}$  에서의 [EQHG] 또는 [RDHG].
- [0784] 76. 단락 59-75 중 어느 하나에 있어서:
- [0785] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 147 개의 인접한 아미노산을 포함하거나,
- [0786] C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 68 개의 인접한 아미노산을 포함하거나,
- [0787] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 136 개의 인접한 아미노산을 포함하고 C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 183 개의 인접한 아미노산을 포함하는 조성물.
- [0788] 77. 단락 59-76 중 어느 하나에 있어서, 관심 계층 유전자좌가 분화 인자, 전사 인자, 신경전달물질 이동체, 신경전달물질 신타아제, 시냅스 단백질, 가소성 단백질, 시냅스전 활성 구역 단백질, 시냅스후 밀도 단백질, 신경

전달물질 수용체, 후생적 변형제, 신경 운명 선별화 인자, 액손 안내 분자, 이온 채널, CpG 결합 단백질, 유비퀴틴화 단백질, 호르몬, 호메오박스 단백질, 성장 인자, 발암유전자 또는 원발암유전자를 인코딩하는 유전자와 관련된 조성물.

- [0789] 78. 세포에서의 게놈 유전자좌의 발현을 유도가능하게 억제하기 위한 비-자연발생적 또는 조작된 조성물로서, 상기 조성물이 하기를 포함하는 DNA 결합 폴리펩티드를 포함하고:
- [0790] (i) 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너와 결합할 수 있음), 및/또는
- [0791] (ii) 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 상호작용 파트너에 결합함);
- [0792] 상기 폴리펩티드가 코돈 최적화된 핵산 분자에 의해 인코딩되고 이로부터 발현되어, 폴리펩티드가 바람직하게는 게놈 유전자좌의 DNA 에 결합하고,
- [0793] 상기 폴리펩티드가 에너지원의 적용시 게놈 유전자좌의 발현을 억제하는 조성물.
- [0794] 79. 단락 78 에 있어서, 폴리펩티드가 1 개 이상의 SID 억제제 도메인을 포함하는 조성물.
- [0795] 80. 단락 79 에 있어서, 폴리펩티드가 4 개 이상의 SID 억제제 도메인을 포함하는 조성물.
- [0796] 81. 단락 78 에 있어서, 폴리펩티드가 SID4X 억제제 도메인을 포함하는 조성물.
- [0797] 82. 단락 78 에 있어서, 폴리펩티드가 KRAB 억제제 도메인을 포함하는 조성물.
- [0798] 83. 단락 78-82 중 어느 하나에 있어서, 에너지 민감성 단백질이 크립토크롬-2 (CRY2) 인 조성물.
- [0799] 84. 단락 78-83 중 어느 하나에 있어서, 상호작용 파트너가 크립토크롬-상호작용 염기성 헬릭스-루프-헬릭스 (CIB1) 인 조성물.
- [0800] 85. 단락 78-84 중 어느 하나에 있어서, 에너지원이 전자기 조사, 음파 에너지 또는 열 에너지로 이루어지는 군에서 선택되는 조성물.
- [0801] 86. 단락 78 에 있어서, 전자기 조사가 가시광의 성분인 조성물.
- [0802] 87. 단락 86 에 있어서, 가시광 성분이 450nm-500nm 범위의 파장을 갖는 조성물.
- [0803] 88. 단락 86 에 있어서, 가시광 성분이 청색광인 조성물.
- [0804] 89. 단락 78 에 있어서, 에너지원을 적용하는 것이  $6.2 \text{ mW/cm}^2$  이상 강도로의 청색광 자극을 포함하는 조성물.
- [0805] 90. 단락 78-89 중 어느 하나에 있어서, DNA 결합 도메인이  $(X_{1-11}-X_{12}X_{13}-X_{14-33}$  또는  $34$  또는  $35$ )/z 를 포함하는 방법으로서, 여기서  $X_{1-11}$  이 11 개의 인접한 아미노산 사슬이고,  $X_{12}X_{13}$  이 반복 가변 이중잔기 (RVD) 이고,  $X_{14-33}$  또는  $34$  또는  $35$  가 21, 22 또는 23 개의 인접한 아미노산 사슬이고, z 가 적어도 5 내지 40 이고, 하나 이상의 RVD 가 NI, HD, NG, NN, KN, RN, NH, NQ, SS, SN, NK, KH, RH, HH, HI, KI, RI, SI, KG, HG, RG, SD, ND, KD, RD, YG, HN, NV, NS, HA, S\*, N\*, KA, H\*, RA, NA 및 NC 로 이루어지는 군에서 선택되고, 이때 (\*) 가  $X_{13}$  에서의 아미노산이 부재함을 의미하는 조성물.
- [0806] 91. 단락 90 에 있어서, z 가 적어도 10 내지 26 인 조성물.
- [0807] 92. 단락 90 에 있어서:
- [0808]  $X_{1-11}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}-X_{14-34}$  또는  $X_{1-11}-X_{14-35}$ ) 에서의 아미노산 1-11 로 나타내는 11 개의 인접한 아미노산의 서열이거나,
- [0809]  $X_{14-34}$  또는  $X_{14-35}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}-X_{14-34}$  또는  $X_{1-11}-X_{14-35}$ ) 에서의 아미노산 12-32 또는 12-33 으로 나타내는 21 또는 22 개의 인접한 아미노산의 서열인 조성물.

- [0810] 93. 단락 78-92 중 어느 하나에 있어서:
- [0811] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 147 개의 인접한 아미노산을 포함하거나,
- [0812] C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 68 개의 인접한 아미노산을 포함하거나,
- [0813] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 136 개의 인접한 아미노산을 포함하고 C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 183 개의 인접한 아미노산을 포함하는 조성물.
- [0814] 94. 단락 78-93 중 어느 하나에 있어서, 관심 게놈 유전자좌가 분화 인자 또는 이온 채널 성분을 인코딩하는 유전자와 관련되는 조성물.
- [0815] 95. 단락 94 에 있어서, 분화 인자가 SRY-박스-2 (SOX2) 이며 유전자 SOX2 에 의해 인코딩되는 조성물.
- [0816] 96. 단락 94 에 있어서, 분화 인자가 p11 이며 유전자 p11 에 의해 인코딩되는 조성물.
- [0817] 97. 단락 95 에 있어서, 이온 채널 성분이 CACNA1C 이며 유전자 CACNA1C 에 의해 인코딩되는 조성물.
- [0818] 98. 세포에서의 게놈 유전자좌의 발현을 유도가능하게 활성화하기 위한 비-자연발생적 또는 조작된 조성물로서, 상기 조성물이 하기를 포함하는 DNA 결합 폴리펩티드를 포함하고:
- [0819] (i) 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 개 이상의 전사 활성제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너와 결합할 수 있음), 및/또는
- [0820] (ii) 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 상호작용 파트너에 결합함);
- [0821] 상기 폴리펩티드가 코돈 최적화된 핵산 분자에 의해 인코딩되고 이로부터 발현되어, 폴리펩티드가 바람직하게는 게놈 유전자좌의 DNA 에 결합하고,
- [0822] 상기 폴리펩티드가 에너지원의 적용시 게놈 유전자좌의 발현을 활성화하는 조성물.
- [0823] 99. 단락 98 에 있어서, 폴리펩티드가 하나 이상의 VP16 또는 VP64 활성제 도메인을 포함하는 조성물.
- [0824] 100. 단락 98 에 있어서, 폴리펩티드가 하나 이상의 p65 활성제 도메인을 포함하는 조성물.
- [0825] 101. 단락 98-100 중 어느 하나에 있어서, 에너지 민감성 단백질이 CRY2 인 조성물.
- [0826] 102. 단락 98-101 중 어느 하나에 있어서, 상호작용 파트너가 CIB1 인 조성물.
- [0827] 103. 단락 98 에 있어서, 에너지원이 전자기 조사, 음파 에너지 또는 열 에너지로 이루어지는 군에서 선택되는 조성물.
- [0828] 104. 단락 103 에 있어서, 전자기 조사가 가시광의 성분인 조성물.
- [0829] 105. 단락 104 에 있어서, 가시광 성분이 450nm-500nm 범위의 파장을 갖는 조성물.
- [0830] 106. 단락 104 에 있어서, 가시광 성분이 청색광인 조성물.
- [0831] 107. 단락 98 에 있어서, 에너지원을 적용하는 것이  $6.2 \text{ mW/cm}^2$  이상 강도로의 청색광 자극을 포함하는 조성물.
- [0832] 108. 단락 98-107 중 어느 하나에 있어서, DNA 결합 도메인이  $(X_{1-11}-X_{12}X_{13}-X_{14-33}$  또는  $^{34}$  또는  $^{35})_z$  를 포함하는 방법으로서, 여기서  $X_{1-11}$  이 11 개의 인접한 아미노산 사슬이고,  $X_{12}X_{13}$  이 반복 가변 이중잔기 (RVD) 이고,  $X_{14-33}$  또는  $^{34}$  또는  $^{35}$  가 21, 22 또는 23 개의 인접한 아미노산 사슬이고,  $z$  가 적어도 5 내지 40 이고, 하나 이상의 RVD 가 NI, HD, NG, NN, KN, RN, NH, NQ, SS, SN, NK, KH, RH, HH, HI, KI, RI, SI, KG, HG, RG, SD, ND, KD, RD, YG, HN, NV, NS, HA, S\*, N\*, KA, H\*, RA, NA 및 NC 로 이루어지는 군에서 선택되고, 이때 (\*) 가  $X_{13}$  에서의 아미노산이 부재함을 의미하는 조성물.
- [0833] 109. 단락 108 에 있어서,  $z$  가 적어도 10 내지 26 인 조성물.

- [0834] 110. 단락 108 에 있어서:
- [0835]  $X_{1-11}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}$ - $X_{14-34}$  또는  $X_{1-11}$ - $X_{14-35}$ ) 에서의 아미노산 1-11 로 나타내는 11 개의 인접한 아미노산의 서열이거나,
- [0836]  $X_{14-34}$  또는  $X_{14-35}$  중 하나 이상이 도 9 의 서열 ( $X_{1-11}$ - $X_{14-34}$  또는  $X_{1-11}$ - $X_{14-35}$ ) 에서의 아미노산 12-32 또는 12-33 으로 나타내는 21 또는 22 개의 인접한 아미노산의 서열인 조성물.
- [0837] 111. 단락 98-110 중 어느 하나에 있어서:
- [0838] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 147 개의 인접한 아미노산을 포함하거나,
- [0839] C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 68 개의 인접한 아미노산을 포함하거나,
- [0840] N-말단 캡핑 부위 또는 이의 단편이 야생형 N-말단 캡핑 부위의 136 개의 인접한 아미노산을 포함하고 C-말단 캡핑 부위 또는 이의 단편이 야생형 C-말단 캡핑 부위의 183 개의 인접한 아미노산을 포함하는 조성물.
- [0841] 112. 단락 98-111 중 어느 하나에 있어서, 관심 게놈 유전자좌가 분화 인자, 후생적 변형제, 이온 채널 성분 또는 수용체를 인코딩하는 유전자와 관련되는 조성물.
- [0842] 113. 단락 112 에 있어서, 분화 인자가 뉴로제닌-2 이며 유전자 NEUROG2 에 의해 인코딩되는 조성물.
- [0843] 114. 단락 112 에 있어서, 분화 인자가 크루펠 유사 인자 4 이며 유전자 KLF-4 에 의해 인코딩되는 조성물.
- [0844] 115. 단락 112 에 있어서, 후생적 변형제가 Tet 메틸시토신 디옥시메틸아제 1 이며 유전자 tet-1 에 의해 인코딩되는 조성물.
- [0845] 116. 단락 112 에 있어서, 이온 채널 성분이 CACNA1C 이며 유전자 CACNA1C 에 의해 인코딩되는 조성물.
- [0846] 117. 단락 112 에 있어서, 수용체가 대사자극성 글루타메이트 수용체이며 유전자 mGlu2 에 의해 인코딩되는 조성물.
- [0847] 118. 단락 98-117 중 어느 하나에 있어서, 발현이 화학적으로 유도가능한 조성물.
- [0848] 119. 단락 118 에 있어서, 화학적으로 유도가능한 발현 시스템이 4-히드록시타목시펜 (4OHT) 에 의해 유도가능한 에스트로겐 수용체 (ER) 기반 시스템인 조성물.
- [0849] 120. 단락 119 에 있어서, 핵 이출 신호 (NES) 를 추가로 포함하는 조성물.
- [0850] 121. 단락 120 에 있어서, NES 가 LDLASLIL 의 서열을 갖는 조성물.
- [0851] 122. 단락 98-121 중 어느 하나에 따른 조성물을 인코딩하는 핵산.
- [0852] 123. 단락 122 에 있어서, 프로모터를 포함하는 핵산.
- [0853] 124. 단락 123 에 있어서, 프로모터가 인간 시냅신 I 프로모터 (hSyn) 인 핵산.
- [0854] 125. 단락 122-124 중 어느 하나에 있어서, 아데노 연관 바이러스 벡터 (AAV) 내로 패키징되는 핵산.
- [0855] 126. 하기를 포함하는, 관심 게놈 유전자좌의 발현을 변형하는 유도 방법:
- [0856] (a) (i) 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 개 이상의 전사 활성화제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너와 결합할 수 있음), 및/또는
- [0857] (ii) 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 상호작용 파트너에 결합함)
- [0858] 을 포함하는 DNA 결합 폴리펩티드를 포함하는 비-자연발생적 또는 조작된 조성물과 게놈 유전자좌를 접촉시키고;
- [0859] (b) 에너지원을 적용시키고;

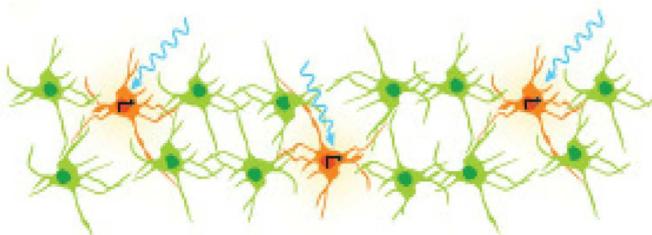
- [0860] (c) 게놈 유전자좌의 발현이 변형된 것을 측정함.
- [0861] 127. 게놈 유전자좌의 발현을 유도가능하게 변형하기 위한 비-자연발생적 또는 조작된 조성물로서, 상기 조성물이 하기를 포함하는 DNA 결합 폴리펩티드를 포함하고:
- [0862] (i) 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 개 이상의 전사 활성화제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너와 결합할 수 있음), 및/또는
- [0863] (ii) 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 상호작용 파트너에 결합함);
- [0864] 상기 폴리펩티드가 바람직하게는 게놈 유전자좌의 DNA 에 결합하고,
- [0865] 상기 폴리펩티드가 에너지원의 적용시 게놈 유전자좌의 발현을 변형하는 조성물.
- [0866] 128. 하기를 포함하는, 세포에서의 관심 게놈 유전자좌의 발현을 교란하기 위한 유도 방법:
- [0867] (a) 데옥시리보핵산 (DNA) 결합 폴리펩티드를 포함하는 비-자연발생적 또는 조작된 조성물과 게놈 유전자좌를 접촉시키고;
- [0868] (b) 유도물질 근원을 적용하고;
- [0869] (c) 게놈 유전자좌의 발현 교란이 발생하는 것을 측정함.
- [0870] 129. 단락 128 에 있어서, 발현을 교란하는 것이 발현을 변형하고 (상향 또는 하향), 발현 결과를 변형하거나 (예컨대 뉴클레아제를 사용) 발현 이동을 제거하고, 예를 들어 의존적 선택사항에 대하여 발현을 변형하는 것인 방법.
- [0871] 130. 단락 128 또는 129 에 있어서, 유도물질 근원이 에너지원 (예컨대 파동 또는 열) 또는 소분자인 방법.
- [0872] 131. 단락 126-130 중 어느 하나에 있어서, DNA 결합 폴리펩티드가 하기를 포함하는 방법:
- [0873] (i) 에너지 민감성 단백질 또는 이의 단편에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 5 개 이상의 전사 활성화제-유사 효과기 (TALE) 단량체 및 적어도 하나 이상의 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 형태적 변화를 거쳐, 상호작용 파트너와 결합할 수 있음), 및/또는
- [0874] (ii) 상호작용 파트너에 연결된 적어도 하나 이상의 효과기 도메인 또는 관심 게놈 유전자좌를 표적화하도록 특이적으로 유도된 적어도 하나 이상의 TALE 단량체 또는 반-단량체를 포함하는 DNA 결합 도메인 (이때 상기 에너지 민감성 단백질 또는 이의 단편은 유도물질 에너지원에 의한 유도시 상호작용 파트너에 결합함).
- [0875] 132. 하기를 포함하는, 세포에서의 관심 게놈 유전자좌의 발현을 교란하기 위한 유도 방법:
- [0876] (a) I. CRISPR/Cas 시스템 키메라 RNA (chiRNA) 폴리뉴클레오티드 서열에 작동가능하게 연결된 제 1 조절 요소 (이때, 상기 폴리뉴클레오티드 서열은 (a) 진핵세포에서의 표적 서열에 하이브리드화할 수 있는 안내 서열, (b) tracr mate 서열 및 (c) tracr 서열을 포함함), 및
- [0877] II. 적어도 하나 이상의 핵 국부화 서열을 포함하는 CRISPR 효소를 인코딩하는 효소-코딩 서열에 작동가능하게 연결된 제 2 조절 유도 요소를 포함하는 하나 이상의 벡터를 포함하는 벡터 시스템을 게놈 유전자좌와 접촉시키고
- [0878] {여기서 (a), (b) 및 (c) 는 5' 에서 3' 배향으로 배열되고,
- [0879] 성분 I 및 II 는 시스템의 동일하거나 상이한 벡터 상에 위치하고,
- [0880] 전사되는 경우, tracr mate 서열은 tracr 서열에 하이브리드화되고 안내 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 유도하고,
- [0881] CRISPR 복합체는 (1) 표적 서열에 대해 하이브리드화되는 안내 서열, 및 (2) tracr 서열에 대해 하이브리드화되는 tracr mate 서열로 복합화되는 CRISPR 효소를 포함하고,

- [0882] CRISPR 효소를 인코딩하는 효소 코딩 서열은 추가로 이중 기능적 도메인을 인코딩함};
- [0883] (b) 유도물질 근원을 적용하고;
- [0884] (c) 게놈 유전자좌의 발현 교란이 발생하는 것을 측정함.
- [0885] 133. 단락 132 에 있어서, 발현을 교란하는 것이 발현을 변형하고 (상향 또는 하향), 발현 결과를 변형하거나 (예컨대 뉴클레아제를 사용) 발현 이동을 제거하고, 예를 들어 의존적 선택사항에 대하여 발현을 변형하는 것인 방법.
- [0886] 134. 단락 132 또는 133 에 있어서, 유도물질 근원이 화학 물질인 방법.
- [0887] 135. 단락 132 내지 134 중 어느 하나에 있어서, 벡터가 렌티바이러스인 방법.
- [0888] 136. 단락 132 내지 135 중 어느 하나에 있어서, 제 2 조절 유도 요소가 테트라사이클린-의존적 조절 시스템을 포함하는 방법.
- [0889] 137. 단락 132 내지 135 중 어느 하나에 있어서, 제 2 조절 유도 요소가 큐메이트 유전자 스위치 시스템을 포함하는 방법.
- [0890] 138. 세포가 원핵세포 또는 진핵세포인, 단락 1-137 중 어느 하나에 따른 조성물, 핵산 또는 방법.
- [0891] 139. 단락 138 에 있어서, 진핵세포가 동물 세포인 조성물, 핵산 또는 방법.
- [0892] 140. 단락 139 에 있어서, 동물 세포가 포유동물 세포인 조성물, 핵산 또는 방법.
- [0893] \* \* \*
- [0894] 본 발명의 바람직한 구현예를 이에 따라 상세히 기재하였으며, 본 발명의 취지 또는 범주를 벗어나지 않고 그의 많은 명백한 변형이 가능하므로, 상기 단락에 의해 정의된 본 발명이 상기 상세한 설명에서 나타낸 특정 세부사항에 제한되지 않는다는 것이 이해될 것이다.

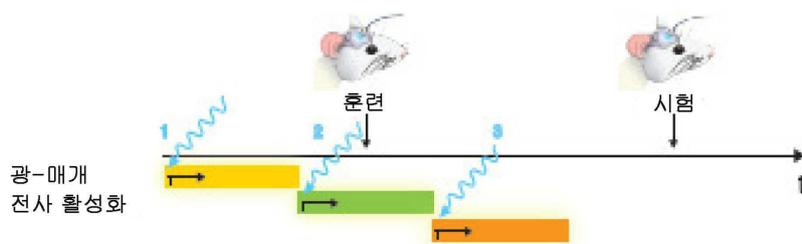
**도면**

**도면1**

**공간적 제어**



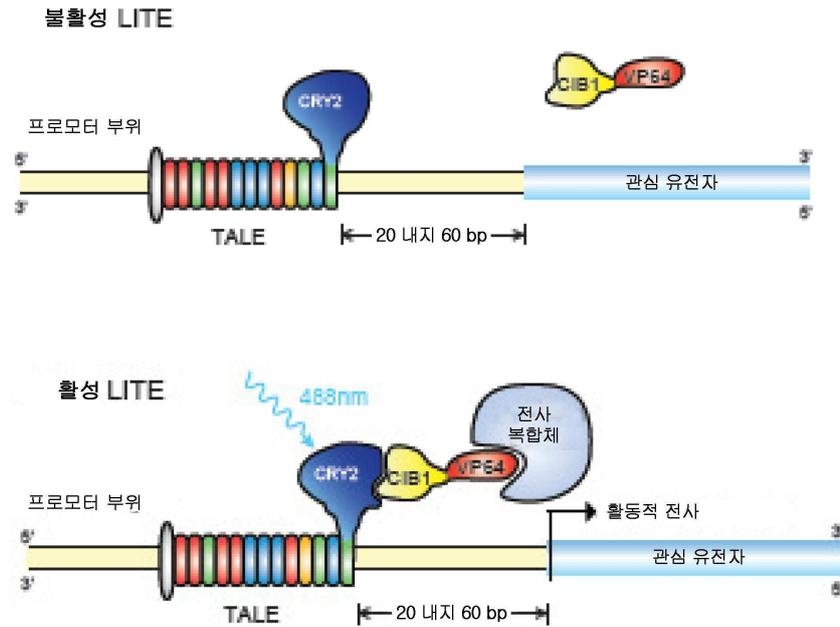
**시간적 제어**



도면2

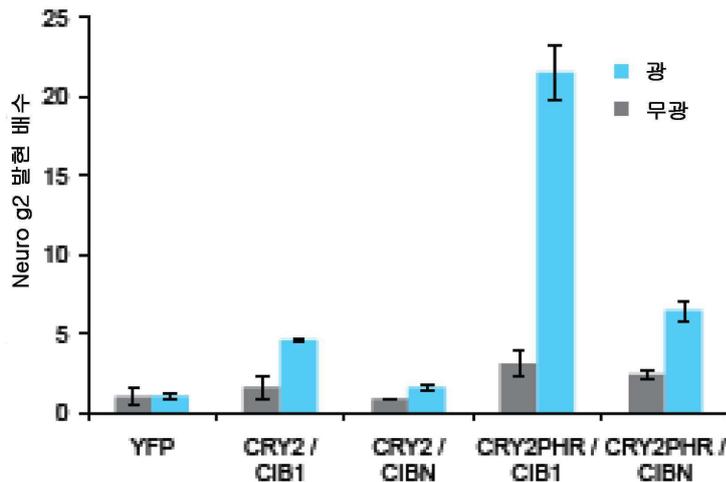


도면3



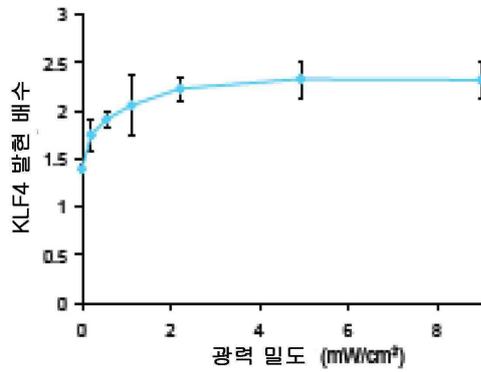
도면4

LITE 활성화에 대한 크립토크롬 이량체 절두의 효과



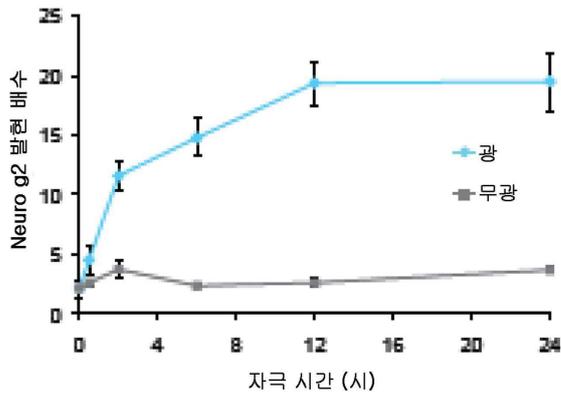
도면5

KLF4 LITE 의 광-강도 의존적 반응

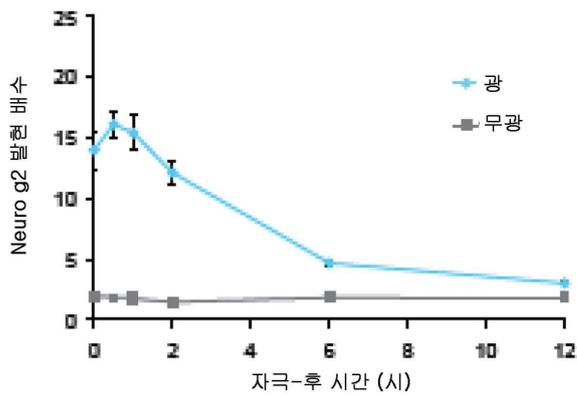


도면6

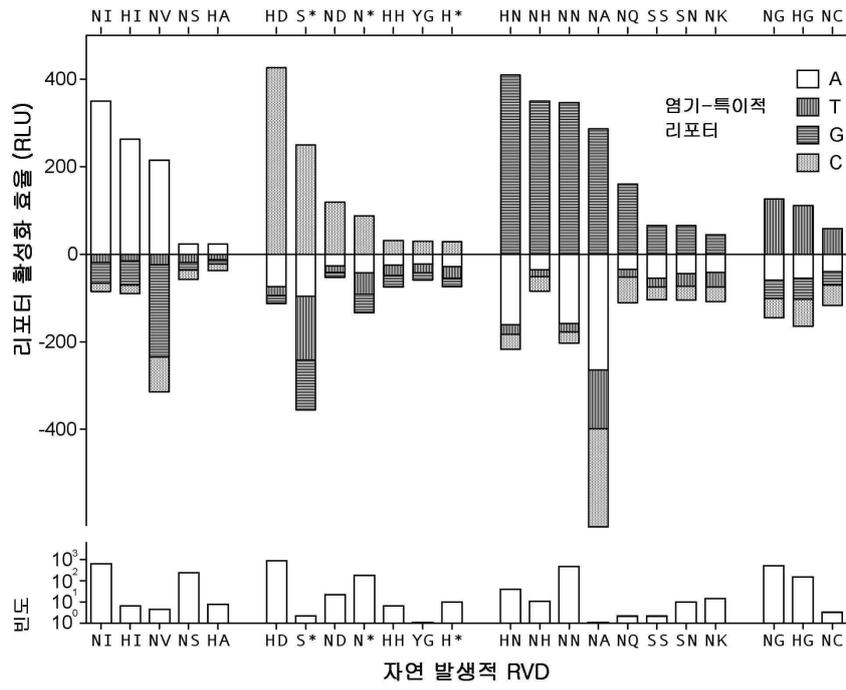
Neurog2 LITE 의 활성화 반응속도



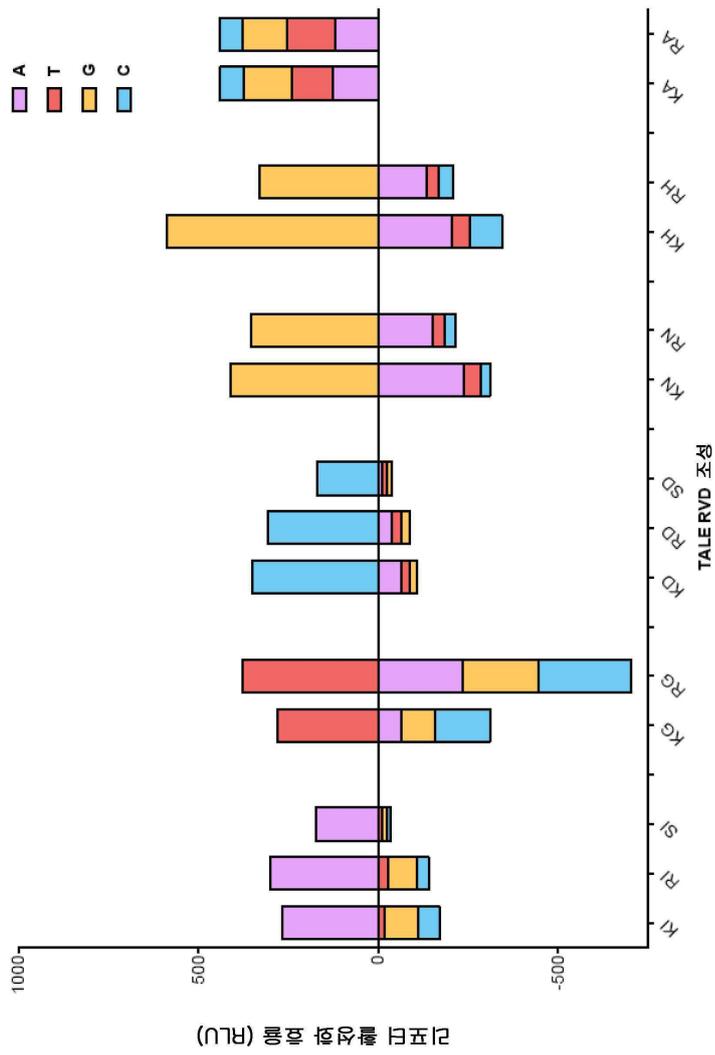
Neurog2 LITE 의 불활성화 반응속도



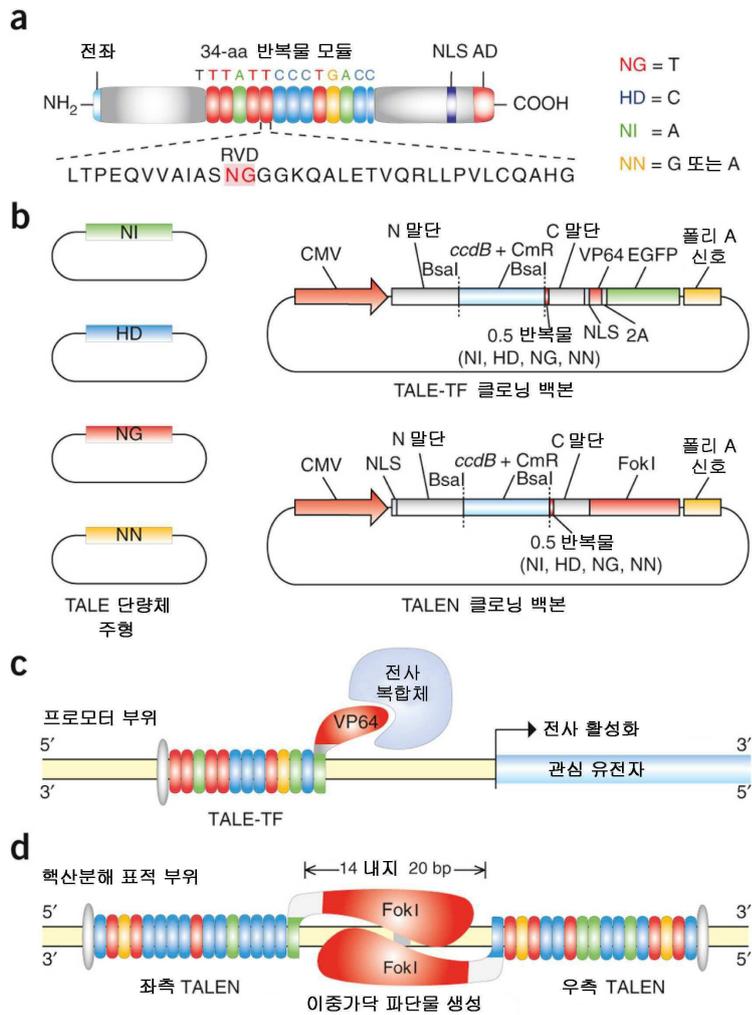
도면7a



도면7b



도면8



도면9a

SEQ ID NO.	단량체 (RVD 제거)	빈도
	LTPDQVVVAIASGGKQALETQRLLPVLCQDHG	754
	LTPEQVVVAIASGGKQALETQRLLPVLCQAHG	278
	LTPDQVVVAIASGGKQALETQRLLPVLCQAHG	254
	LTPAQVVVAIASGGKQALETQRLLPVLCQAHG	147
	LTPAQVVVAIASGGKQALETQRLLPVLCQDHG	143
	LTPDQVVVAIASGGKQALETQRLLPVLCQAHG	107
	LTPDQVVVAIANGGKQALETQRLLPVLCQDHG	72
	LTLDQVVVAIASGGKQALETQRLLPVLCQDHG	47
	LTPQQVVVAIASGGKQALETQRLLPVLCQAHG	47
	LTPDQVVVAIANGGKQALETQRLLPVLCQAHG	40
	LTPNQVVVAIASGGKQALETQRLLPVLCQDHG	35
	LTLDQVVVAIASGGKQALETQRLLPVLCQAHG	27
	LTPAQVVVAIANGGKQALETQRLLPVLCQDHG	20
	LTPEQVVVAIASGGKQALETQALLPVLCQAHG	19
	LTLDQVVVAIASGSKQALETQRLLPVLCQDHG	19
	LTQDQVVVAIASGGKQALETQRLLPVLCQDHG	18
	LSPDQVVVAIASGGKQALETQRLLPVLCQDHG	13
	LTPDQVVVAIANGGKQALETQRLLPVLCQDHG	13
	LTPDQVVVAIASGGKQALETQRLLPVLCQDHG	11
	LTPDQVVVAIASGGKQALETQRLLPVLRQAHG	11
	LTPDQVVVAIASGGNQALETQRLLPVLCQAHG	11
	LTPDQVVVAIASGGKQALATVQRLLPVLCQAHG	10
	LTPAQVVVAIANGGKQALETQRLLPVLCQAHG	9
	LTLAQVVVAIASGGKQALETQRLLPVLCQAHG	9
	LTPEQVVVAIACGGKQALETQRLLPVLCQAHG	9
	LTPAQVVVAIASGGKQALETVQQLLPVLCQAHG	9
	LTPQQVVVAIASGGRPALETQRLLPVLCQAHG	9
	LTPDQVVVAIASGSKQALETQRLLPVLCQDHG	8
	LTPNQVVVAIASGGKQALETQRLLPVLCQAHG	8
	LTPDQVVVAIASGGKQALGTQRLLPVLCQDHG	8
	LTLAQVVVAIASGGKQALETQRLLPVLCQDHG	8
	LTPAQAVVAIASGGKQALETQRLLPVLCQDHG	7
	LTPAQVVVAIASGGNQALETQRLLPVLCQDHG	7
	LTPDQVVVAIASGGKQALETQRLLPVLCQAHG	7
	LTPDQVVVAIANGGKQALETQRLLPVLCQAHG	7
	LTPDQVVVTIASGGKQALETQRLLPVLCQDHG	7
	LTPAQVVVAIANGGKQALETVRLLPVLCQDHG	7
	LTPDQVVVAIASGGNQALETQRLLPVLCQDHG	6
	LTPDQVVVAIASGGKQALETQRLLPVLCQTHG	6

도면9b

SEQ ID NO.	단량체 (RVD 제거)	빈도
	LPPDQVVVAIASGGKQALETVQRLLPVLCQDHG	6
	LTSdqVVVAIASGGKQALETVQRLLPVLCQDHG	6
	LTPAQVVVAIASGGKQALETVQRLLPVLCQDHG	6
	LIPAQVVVAIASGGKQALETVQRLLPVLCQDHG	5
	LTPAQVVVAIASGGKQALETMQRLLPVLCQAHG	5
	LTRDQVVVAIASGGKQALETVQRLLPVLCQDHG	5
	LTPDQVVVAIASGGKQALETVQRLLPVLCQDHG	5
	LIPDQVVVAIANGGKQALETVQRLLPVLCQAHG	5
	LTPDQVVVAIASGGKQALETVQRLLPVLCQNHG	5
	LTLdqVVVAIASGGKQALETVQRLLPVLCQDHG	4
	LTPDQLVAIANGGKQALETVQRLLPVLCQDHG	4
	LTPDQVVVAIASGGKQALETVQRLLPVLCQGHG	4
	LTPDQVVVAIASGGKQALETVQRLLPVLCQEHG	4
	LTLdkVVVAIASGGKQALETVQRLLPVLCQDHG	4
	LTPAQVVVAIASGSKQALETVQRLLPVLCQAHG	4
	LTPDKVVVAIASGGKQALETVQRLLPVLCQAHG	4
	LTQdqVVVAIASGGKQALETVQRLLPVLYQDHG	4
	LTPAQVVVAIVSGGKQALETVQRLLPVLCQAHG	4
	LTPDKVVVAIANGGKQALETVQRLLPVLCQDHG	4
	LTQdqVVVAIASGGKQALETVQRLLPVLCQAHG	4
	LTPDQVMAIANGGKQALETVQRLLPVLCQDHG	4
	LTTdqVVVAIASGGKQALETVQRLLPVLCQAHG	4
	LTPDQVVVAIASGSKQALETVQRLLPVLCQAHG	3
	LTPDQVVVAIANGGKQALETVQRLLPVLCQAHG	3
	LTQEQVVVAIASGGKQALETVQRLLPVLCQAHG	3
	LTPDQVVVTIANGGKQALETVQRLLPVLCQAHG	3
	LSPAQVVVAIASGGKQALETVQRLLPVLCQDHG	3
	LTPDQVVVAIASGGKQALEMVQRLLPVLCQAHG	3
	LIPDQVVVAIASGGKQALETVQRLLPVLCQDHG	3
	LTPVQVVVAIASGGKQALETVQRLLPVLCQDHG	3
	LTPDQVVVAIASGGKQALKTVQRLLPVLCQDHG	3
	LTPDQVVVAIASGGKQALETMQRLLPVLCQAHG	3
	LTPAQVVVAIASGGKQALETVQRLLFPVLCQDHG	3
	LTPAQVVVAIASGGKQALETVQQLLPVLCQAHG	3
	LTPAQVVALASGGKQALETVQRLLPVLCQDHG	3
	LTPDQVVVAIASGGKQALETVQRLLPVLCQDHG	3
	LTPDQVVVAIASGGKQALATVQRLLPVLCQDHG	3
	LTQVQVVVAIASGGKQALETVQRLLPVLCQAHG	3
	LTPDQVVVAIARGGKQALETVQRLLPVLCQAHG	3

도면9c

SEQ ID NO.	단량체 (RVD 제거)	빈도
	LPPDQVVVAIASGGKQALETVQRLLPVLCQAHG	3
	LTLDQVVVAIASGGKQALETVQRLLPVLCQAHG	3
	LSPDQVVVAIANGGKQALETQRLLPVLCQTHA	3
	LNPDQVVVAIASGGKQALETVQRLLPVLCQDHG	3
	LTPDQVMAIASGGKQALETVQRLLPVLCQDHG	3
	LTPAQVVVAIASGGKQALETVRRLLPVLCQAHG	3
	LTPDQVVVAIASGGKQTLETVQRLLPVLCQDHG	3
	LTPDQVMTIASGGKQALETVQRLLPVLCQDHG	3
	LTPAQVVVTIASGGKQALETVQRLLPVLCQDHG	3
	LTPAQVVVAIASGGKQALETVQRLLPVLCRAHG	3
	LSPDQVVVAIASGGKQALETVQRLLPVLCQAHG	3
	LTPDQVVVGIASGGKQALETVQRLLPVLCQDHG	2
	LTPDQVVVAIASGGKQALETVQRLLPVLCQANG	2
	LTPAQVVVAIASGGKQALETVQRLLPVLCQTHG	2
	LTPDQVVVAIASGGKQALEMVQRLLPVLCQDHG	2
	LTPDQVVVAIASGGKQALETMRLLPVLCQDHG	2
	LTPDQVVVAIANGGKQALATVQRLLPVLCQDHG	2
	LTPDQVVVTIASGGKQALETVQRLLPVLCQAHG	2
	LTPDQVVVAIASGGKQALETVQRLLTVLCQDHG	2
	MTPDQVVVAIASGGKQALETVQRLLPVLCQDHG	2
	LAPDQVVAVASGGKQALETVQRLLPVLCQDHG	2
	LTPAQVVVAIASGGKQALKTVQQLLPVLCQDHG	2
	LTPDQVVVAIARGGKQALETVQRLLPVLCQDHG	2
	LTPDQVVVAIASGGKQALETVQQLLPVLCQAHG	2
	LTPDQVLAIASGGKQALETQRLLPVLCQDHG	2
	LTPEQVVVAIARGGKQALETVQRLLPVLCQAHG	2
	LTPAQVVVAIASGGKQALETMRLLPVLCRAHG	2
	LTPDQVVVAIANGGKQALEMVQRLLPVLCQDHG	2
	LTTDQVVVTIASGGKQALETVQRLLPVLCQDHG	2
	LTPTQVMAIANGGKQALETVQRLLPVLCQDHG	2
	LTPQQVVVAIASGGKQALETVQALLPVLCQAHG	2
	LTPDQVVVAIASGGKQALETVQRLLPMLCQDHG	2
	L TSAQVVVAIANGGKQALETVQRLLPVLCQDHG	2
	LTPDQVVVAIASGGKQALETVQQLLPVLCQDHG	2
	LTPDQVVVAIANGGKQALATVQRLLPVLCQAHG	2
	LTPAQVVVAIASGGKQALETVQRLLPMLCQAHG	2
	LTLDQVVVAIASGGKQALETVQRLLPVLCQARG	2
	LTPAQVVVAIASGGKQALETQRLLPVLCQDHG	2
	LTPDQVVVAIANGGKQALETVQRLLPVLCQNHG	2

도면9d

SEQ ID NO.	단량체 (RVD 제거)	빈도
	LTPDQVVTIASGGKQALEMVQRLLPVLCQDHG	2
	LTPDQVVAIASGGKQALERVQRLLPVLCQEHG	1
	LTPEQVVAIACGGKQALETVQALLPVLRQAHG	1
	LTPDQVVAIASGGKQALETVQRLLPVLCRDHG	1
	LTPEQVVAIASGGKQALETVQRLLPMLCQAHG	1
	LTPEQVVAIACGGKQALETVQRLLPVLRHAHG	1
	LTPDQVVAIASGGKQALETVQRLLPVLCQHHG	1
	LIPDQVVAIASGGKQALETVQRLLPVLCQHHG	1
	LTRAQVVAIASGGKQALETVQRLLPVLCQEHG	1
	LTPDQVVAIANGGKQAVGTVQRLLPVLCQAHG	1
	LTLQVVAIASGGKQALETVQRLLPVLCQEHG	1
	LTPAQVVAIASGGKQALETVQRLLPMLCQDHG	1
	LTPDQVVAIASGSKQALETMQRLLPVLCQDHG	1
	LTPDQVVAIASGGKQALETVQRLLPVLCQKHG	1
	LTLQVVAIASGGKQALETVQRLLPVLCQTHG	1
	LTPDQVVAIASGGKQALEAVQRLLPVLCQDHG	1
	LTPAQVVTIASGGKQALETVQRLLPVLCQEHG	1
	LTPAQVMAIASGGKQALETVQRLLPVLCQDHG	1
	LTREQVVAIASGGKQALETVQRLLPVLRQAHG	1
	LTLAQVVAIANGGKQALETVQRLLPVLCQAHG	1
	LTLQVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTPQQVVAIASGGKQALETVQRLLPVLCQEHG	1
	LSPDQVVAIANGGKQALETVQRLLPVLCQDHG	1
	LTPDQVVAIANGGKQALETVQRLLPVLCQHHG	1
	LTPEQVVAIASGGKQALETVQALLPVLRQAHG	1
	LSQDQVVAIASGGKQALETVQRLLPVLCQDHG	1
	LPPEQVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTPDQVVAIASGGKQALEAVQRLLPVLCQAHG	1
	LTPDQVVAIANGGKQALETVQRLLPVLCQEHG	1
	LTLQVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTPDQVVAIASGGKQALETVQRVLPVLCQDHG	1
	LIPAQVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTPAQVVAIASGGKQALETVQRLLPVLRQAHG	1
	LTPAQVVAIASGSKQALETVQRLLPVLCQTHG	1
	LTPQQVVAIASGGKQALETVQRLLPVLCQDHG	1
	LTPDQVVAIANGGKQAVETVQRLLPVLCQAHG	1
	LSPDQVVTIASGGKQALETLQRLLPVLCQDHG	1
	LTPVQVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTLQVVAIASGSKQALETVQRLLPVLCQTHG	1

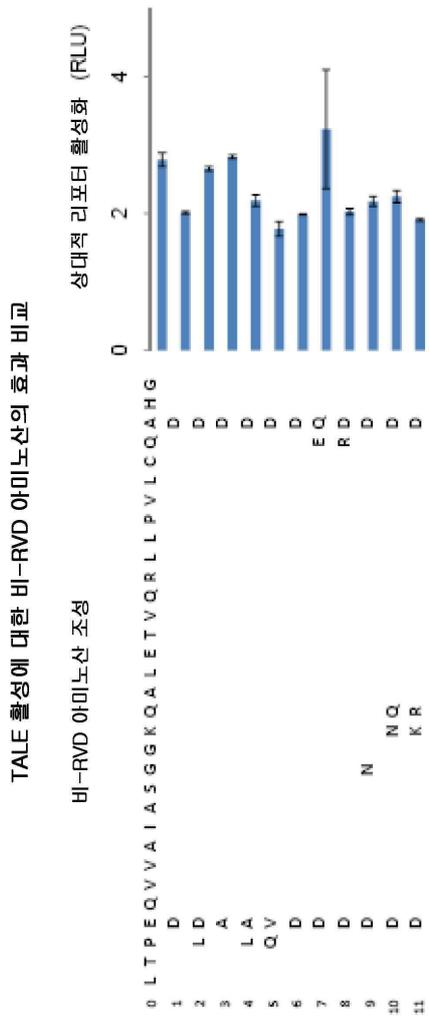
도면9e

SEQ ID NO.	단량체 (RVD 제거)	빈도
	LTPAQVVAIACGGKQALETVRRLLPVLCQAHG	1
	LTPAQVVAIASGSKQALETVQRLLPVLCQAHG	1
	LPPAQVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTPDQVVAIASGGKQALETVQRLLPVLFQEHG	1
	LTPAKVVAIASGGKQALETVQRLLPVLCQDHG	1
	LTPVQVVAIASGGKQALETVQRLLPVLCQDHG	1
	LTPDQVVAIASGGKQALETVQRLLPGLCQDHG	1
	LTLAQVVAIANGGKQALETVQRLLPVLCQDHG	1
	LTPAQVVAIASGGKQALETVQRLLTVLCQDHG	1
	LPPAQVVAIASGGKQALETVQRLLPVLCQDHG	1
	LTPAQAVAIASGGKQALETVQRLLPVLCQAHG	1
	LTPAQVVAIVSGGKQALETVQRLLPVLCQTHG	1
	LTPDQVVAVAGGGKQALETVQRLLPVLCQDHG	1
	LTPDQVVAIASGGKQALETVQRLLPVLCQAHG	1
	LPPAQVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTTDQVVAIASGGKQALETVQRLLPVLCQDHG	1
	LTPDQVVAIANGGKQALETVQRLLVPVLCQDHG	1
	LTPDQVVAIASGGKQALETVQRLLPVLCQTHA	1
	LTLAQVVAIASGGKQALETVQRLLPVLCQTHG	1
	LTPNQLVAIANGGKQALETVQRLLPVLCQDHG	1
	LSPAQVVAIASGSKQALETVQRLLPVLCQDHG	1
	LTPDQVVAIASGGKQALETVQRVLPVLCQAHG	1
	LTPDQVMAIANGGKQALETVQRLLPVLCQAHG	1
	LTPEQVVAIASGGRQALETVQRLLPVLCQAHG	1
	LTPAQVVAIASGGKQALETVQWLLPVLCQAHG	1
	LTPDKVVAIASGGKQALETVQRLLPVLCQDHG	1
	LTPAQVMAIANGGKQALETVQRLLPVLCQDHG	1
	LTQDQVVAIASGGKQALETVQRLLPVLCQANG	1
	LTPAQVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTPDQVVAIASGGKQALETMQRLLPVLCQDHG	1
	LTPDQVVAIASGSKQALETVQRLLPVLQRDHG	1
	LTPYQVVAIASGSKQALETVQRLLPVLCQDHG	1
	LTPYQVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTLDQVVAIASGGKQALETVQRLLPVLCQEHG	1
	LTLEQVVAIASGGKQALETVQRLLVLCQAHG	1
	LTPDQVVAIASGGKQALETVRRLLQVLCQDHG	1
	LTPDQVVAIASGGKQALETVQRLLPVLQRDHG	1
	LTPDQVVSANGGKQALETVQRLLPVLCQAHG	1
	LTPDQVVAIANGGKQALETVQRLLPVLCQTHG	1

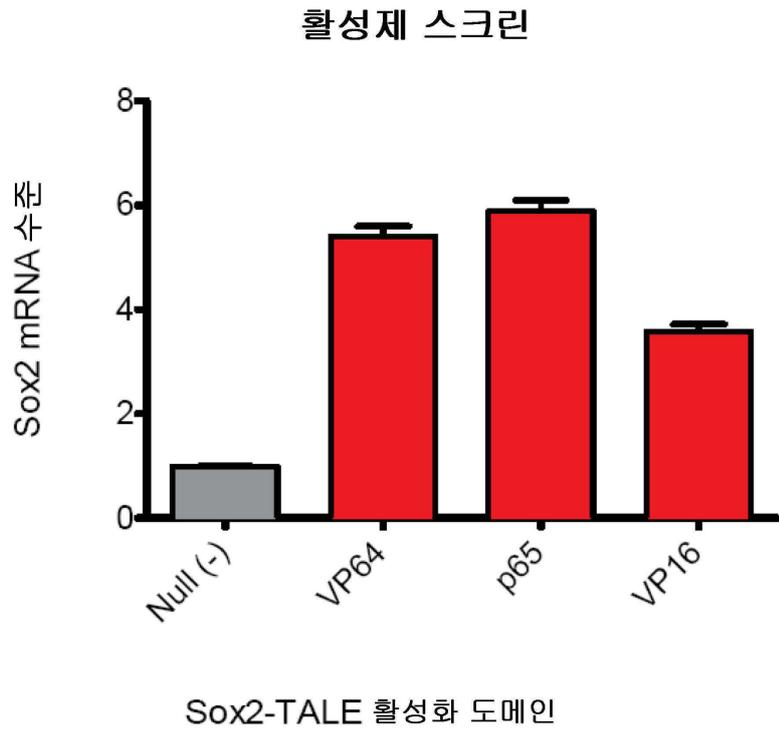
도면9f

SEQ ID NO.	단량체 (RVD 제거)	빈도
	LTPDQVVVAIASGGKQALETVKRLLPVLCQAHG	1
	LTTDQVVVAIANGGKQALETVQRLLPVLCQDHG	1
	LIPQQVVVAIASGGKQALETVQRLLPVLCQDHG	1
	LTLTQVVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTPTQVVVAIASGGKQALETVQRLLPVLCQDHG	1
	LTPTQVMAIANGGKQALETVQRLLPVLCQAHG	1
	LTPDQVVAVASGGKQALETVQRLLPVLCQAHG	1
	LTPAQVVVAIASGSKQALETVQRLLPVLCQDHG	1
	LTPGQVVVAIASGGKRALETVQRLLPVLCQDHG	1
	LTPDQVVVIAISGGKQALETVQRLLPVLCQAHG	1
	LPPDQVVVAIASGSKQALETVQRLLPVLCQDHG	1
	LTPDQVVVTIANGSKQALETVQRLLPVLCQAHG	1
	LTPAQVVVAIASGGKQALETVQRLLQVLCQDHG	1
	LTPDHVVVAIASGGKQALETVQRLLPVLCQDHG	1
	LTPDQVVVAIASGGKQALETVQRLLQVLCQDHG	1
	LTPDQVVVAIASGGRQALETVQRLLPVLCQAHG	1
	LHPGQVVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTLDQVVVSIASGGKQALETVQRLLPVLCQDHG	1
	LTPDQVVVAIASGGKQALETVQRLLPALCQDHG	1
	LTPDQVVVAIASGGKPALETVQRLLPVLCQAHG	1
	LTPAQVVVAIASGGKQALKTVQRLLPVLCQAHG	1
	LTPDQVVVAIASGGKRALETVQRLLPVLCQAHG	1
	LNPDQVVVAIASGGKQALETVQRLLPVLCQAHG	1
	LTPDQVVVAIASGGKQALETVKRLLPVLCQDHG	1
	LTLDQVVVAIANGGKQALETVQRLLPVLCQAHG	1
	LTPAQVVVAIASGGKQALETVQRLLPVLCRDHG	1
	LTPAQVLAIASGGKQALETVQRLLTVLCQDHG	1
	LTPAQVVVAIASGGKQALETMQRLLPVLCQDHG	1
	LTPDQVVVAIASGGKQALETVQRLLPGLCQAHG	1
	LTREQVVVAIASGGKQALETVQALLPVLQAHG	1
	LTPAQVVVAIASGGKQALETVQRLLPVLCQVHG	1
	LTPNQVVVAIASGGKQALETVQRLLLVLCQDHG	1
	LTPDQVMAIASGGKQALETVQRLLPVLCQAHG	1
	LTREQVVVAIASGGKQALETVQRLLPVLCQDHG	1
	LSTAQVVVAIASGGKQALEGIGEQLLKLRTPYQ	1
	LSTAQVVAVASGGKPALEAVRAQLLALRAAPYQ	1

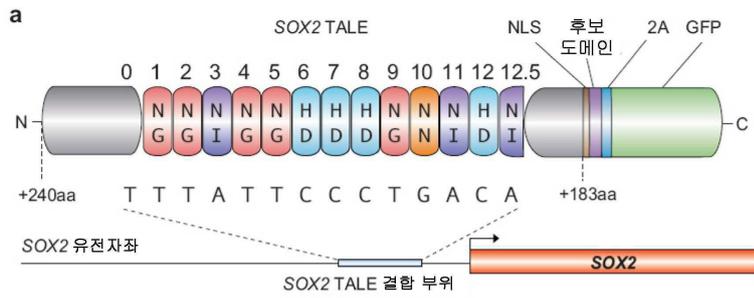
도면10



도면11

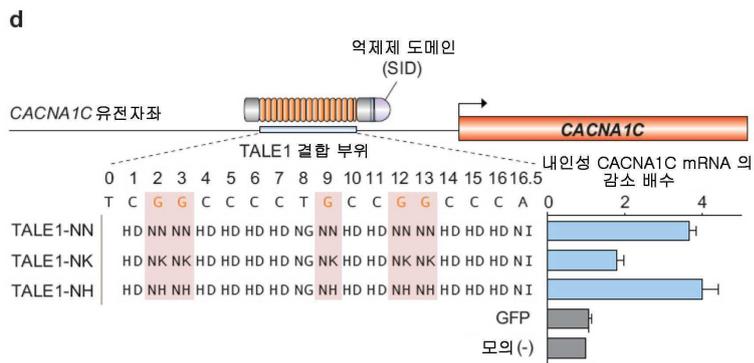
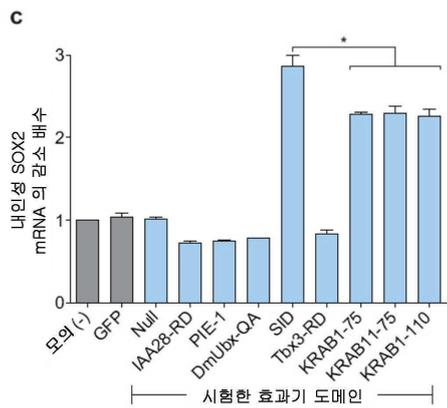


도면12

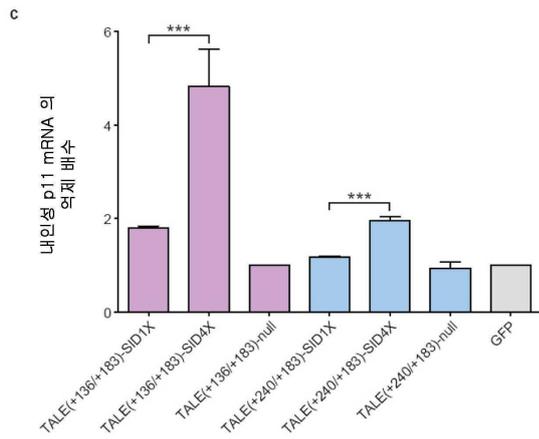
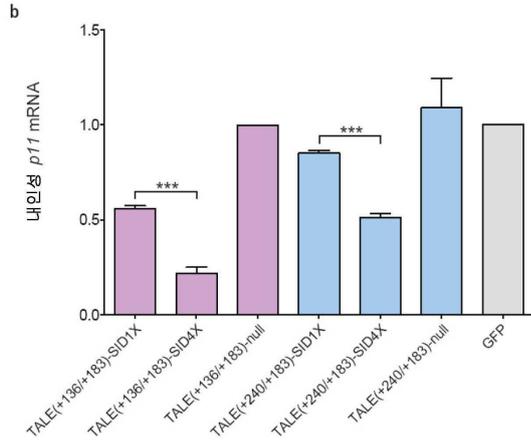
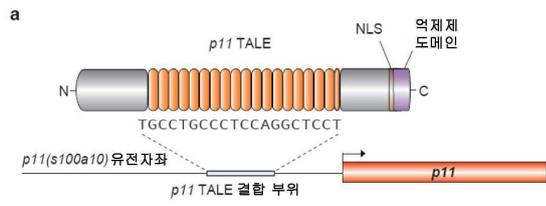


**b**

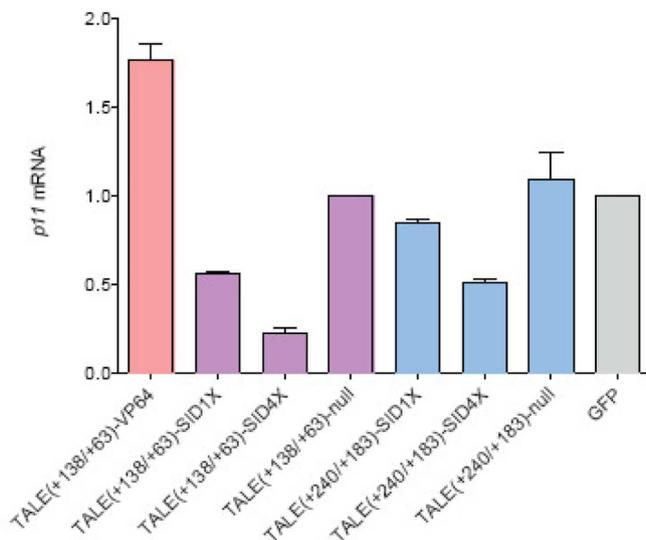
후보 도메인	유기체 기원
PIE-1	C. 엘레간스
Ubx-QA	D. 멜라노가스테르
IAA28-RD	A. 탈리아나
SID	H. 사피엔스
Tbx3-RD	H. 사피엔스
KRAB	H. 사피엔스



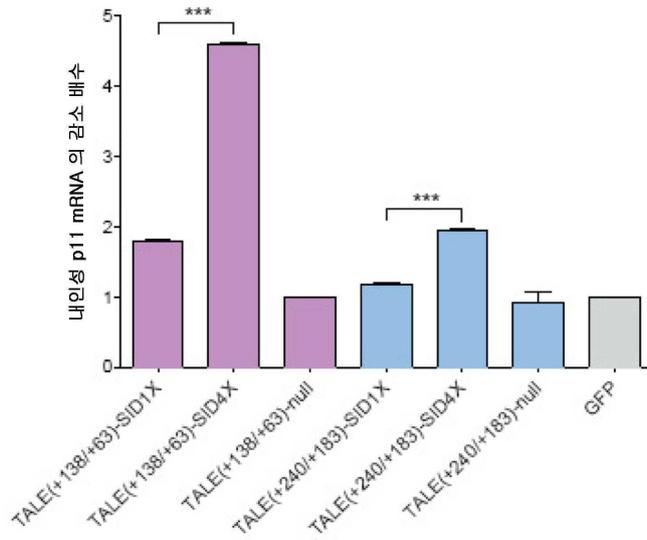
도면13



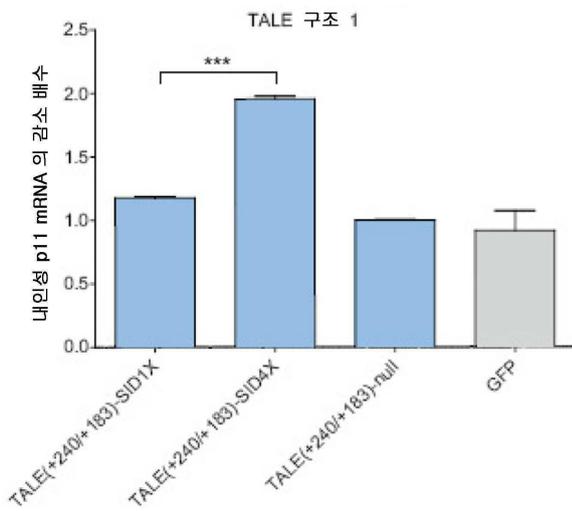
도면14a



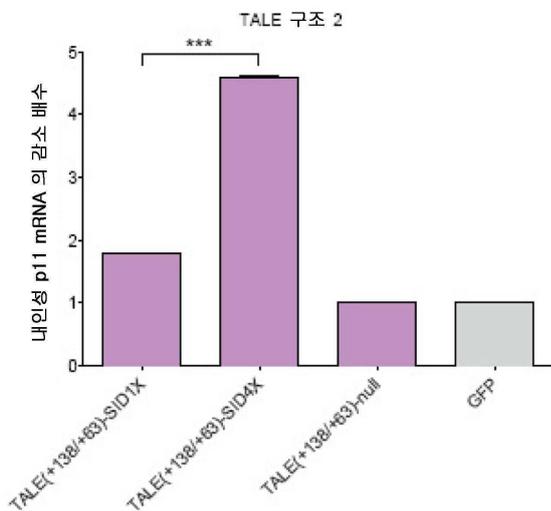
도면14b



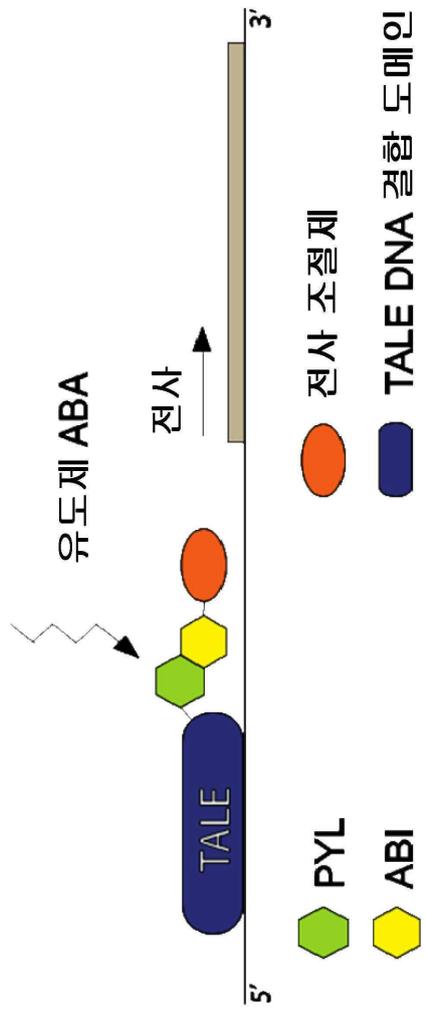
도면14c



도면14d

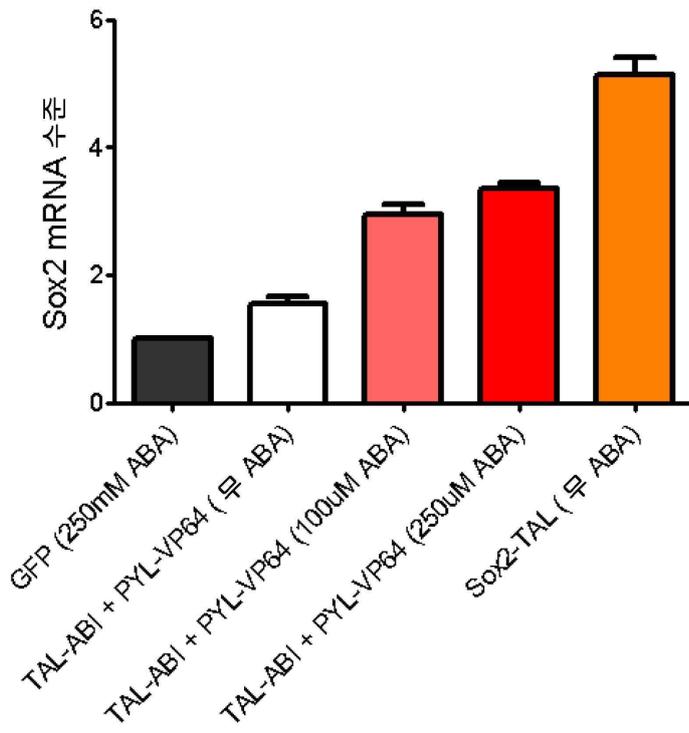


도면 15a



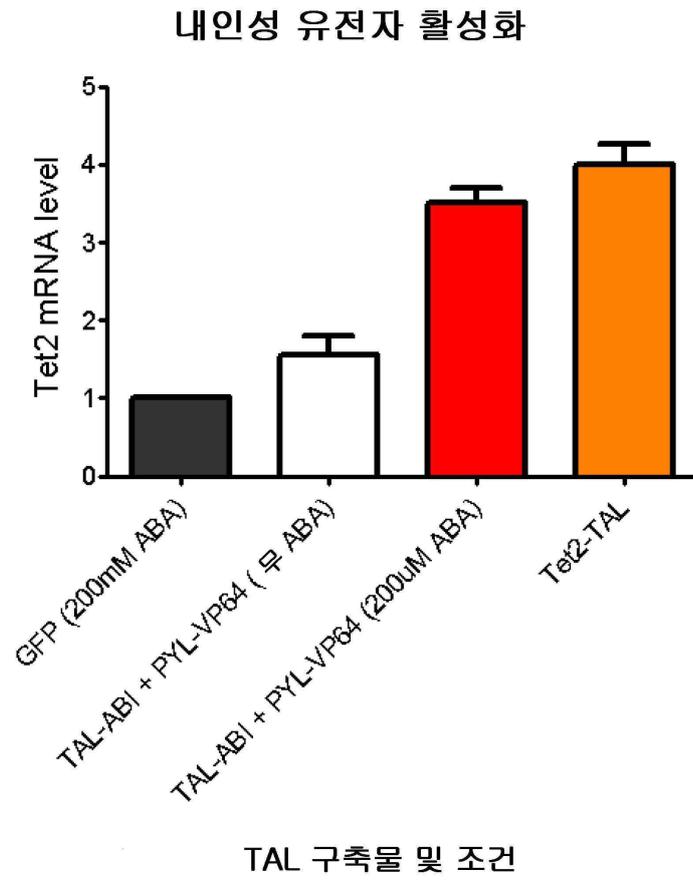
도면15b

내인성 유전자 활성화

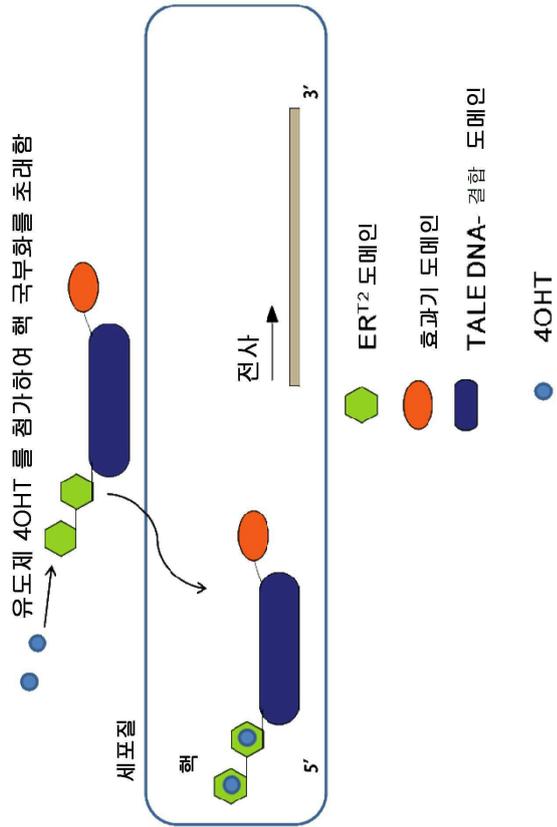


TAL 구조물 및 조건

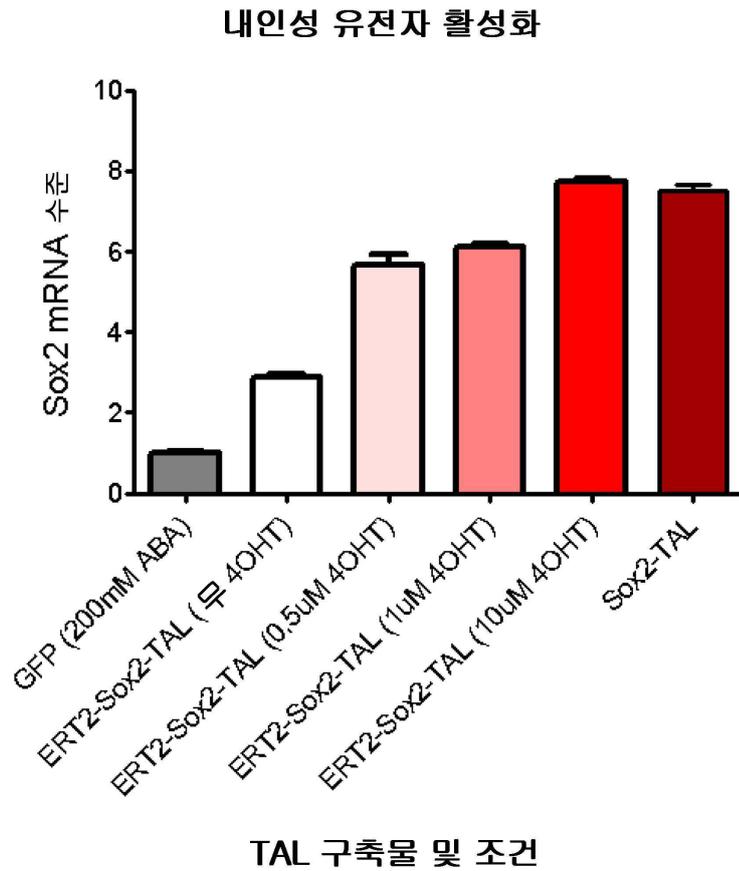
도면15c



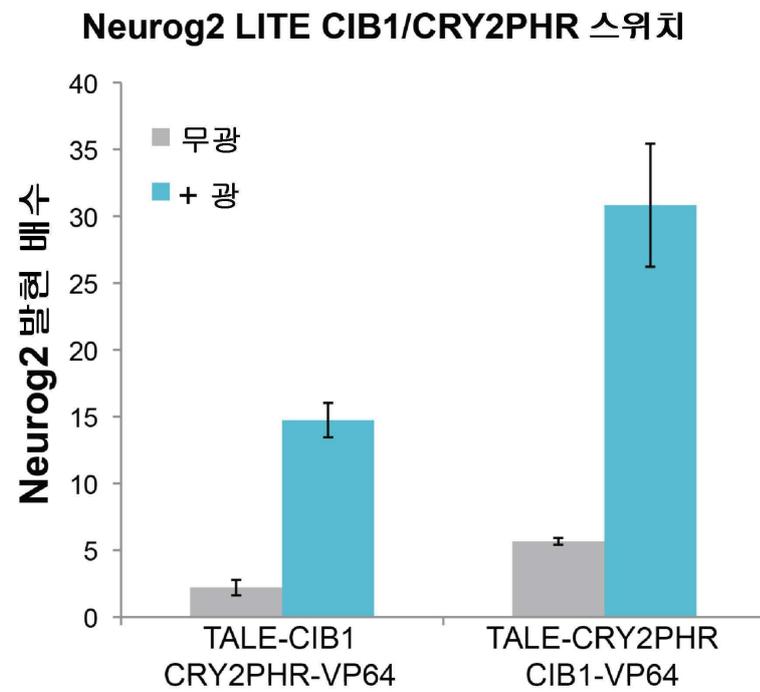
도면16a



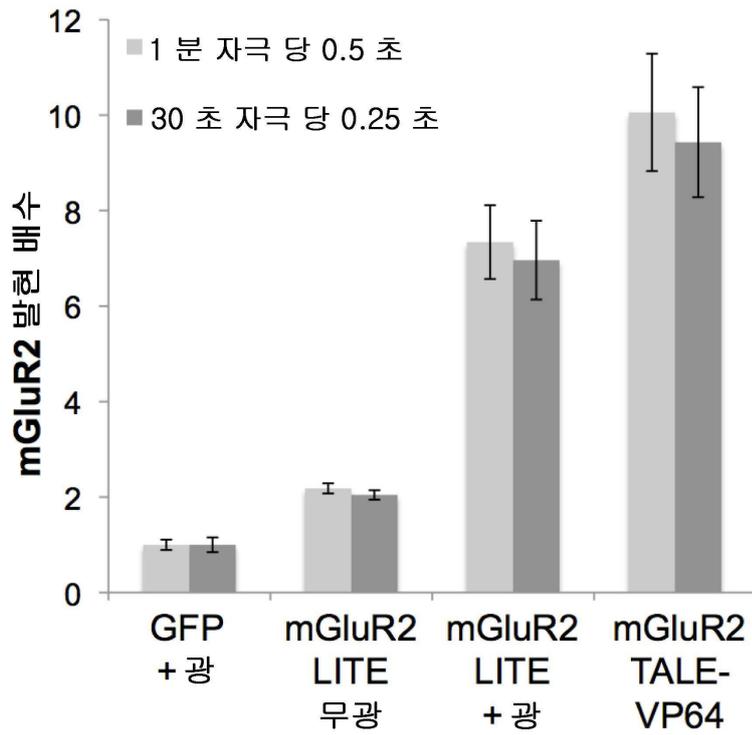
도면16b



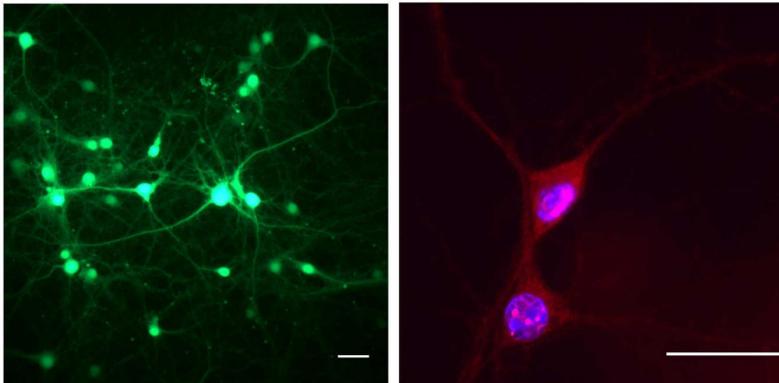
도면17



도면18



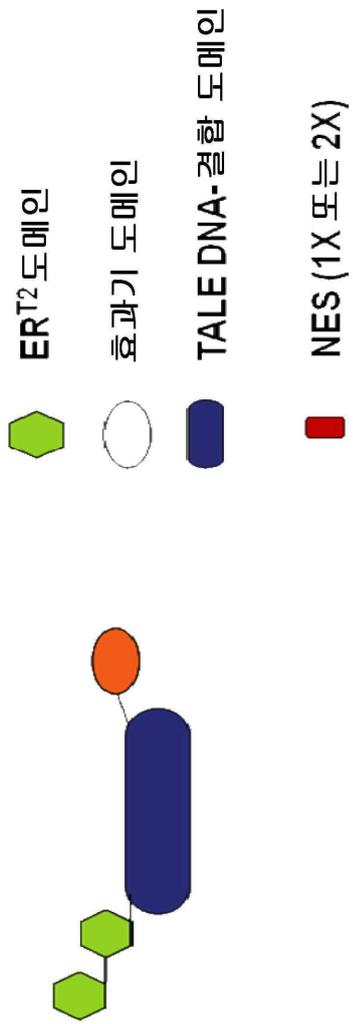
도면19



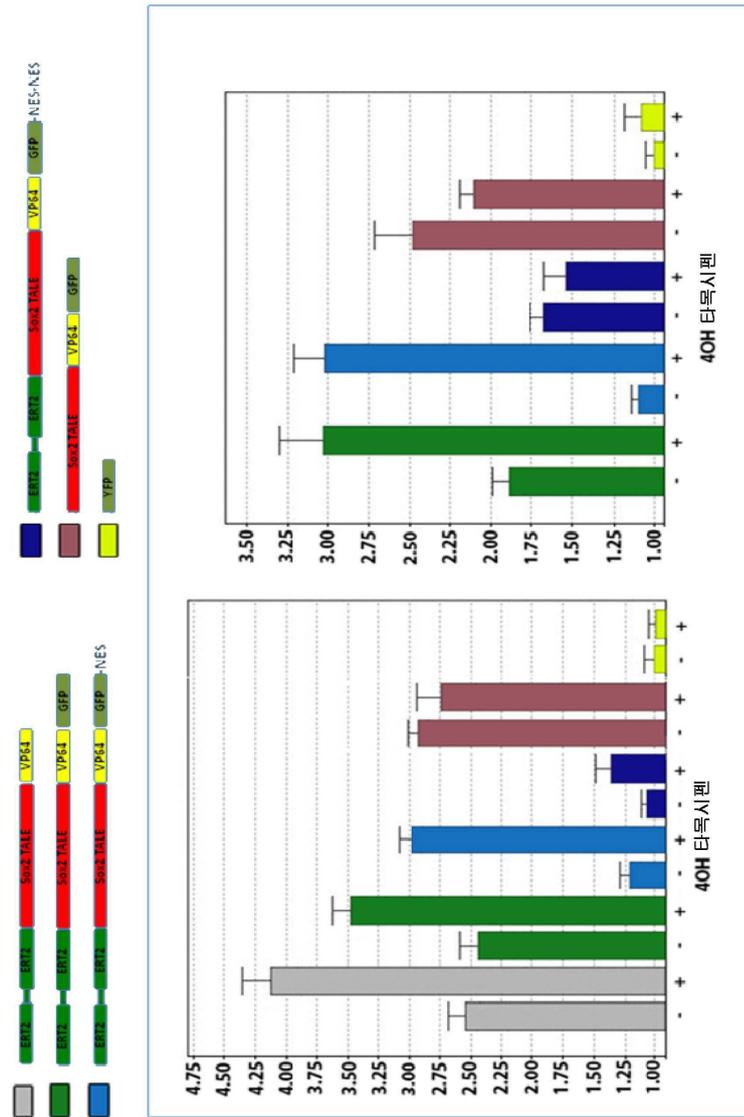
도면20



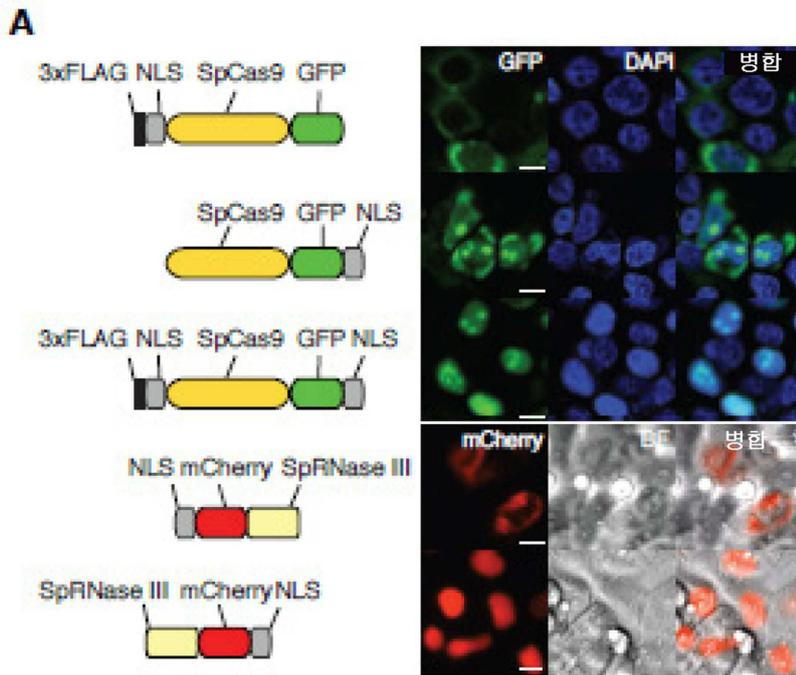
도면21



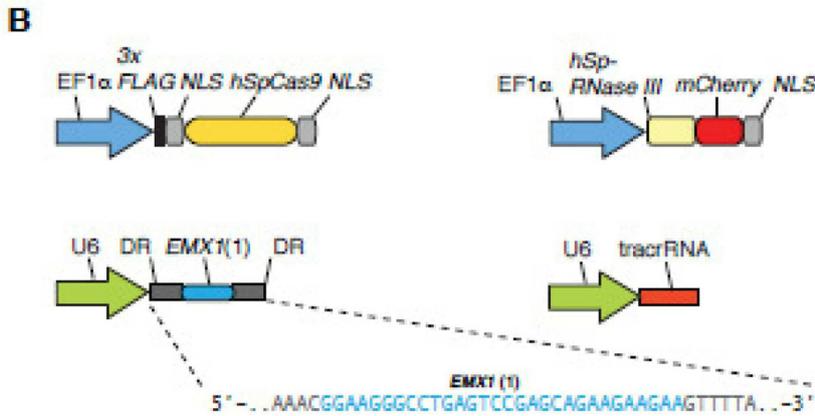
도면22



도면23a



도면23b



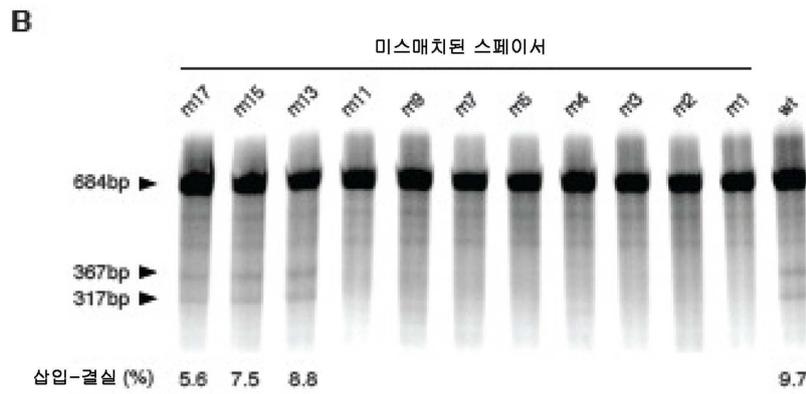
도면23c







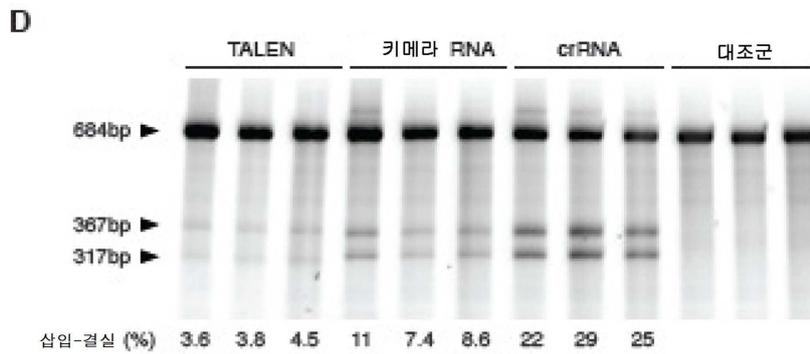
도면25b



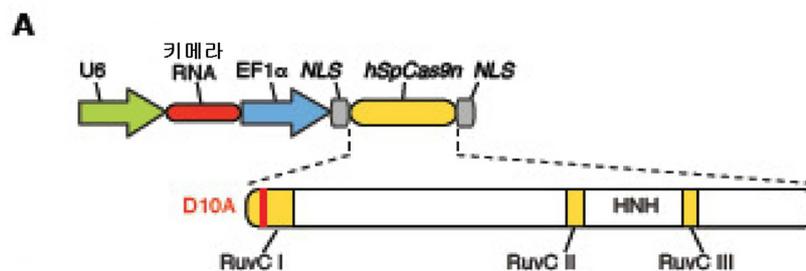
도면25c



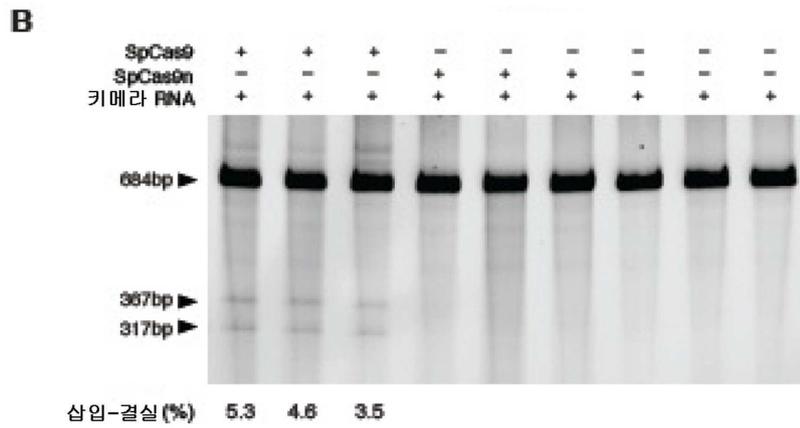
도면25d



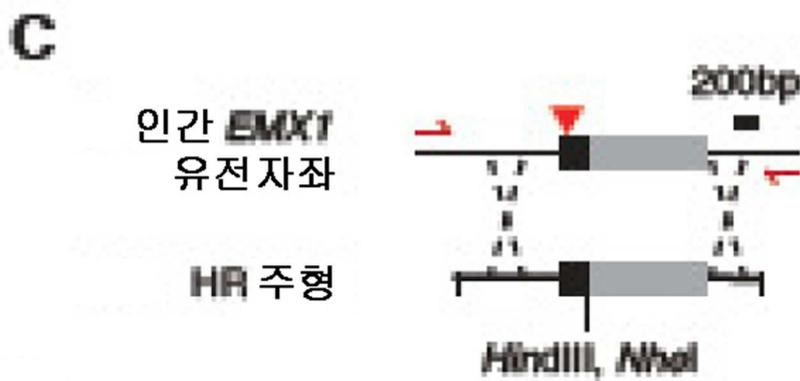
도면26a



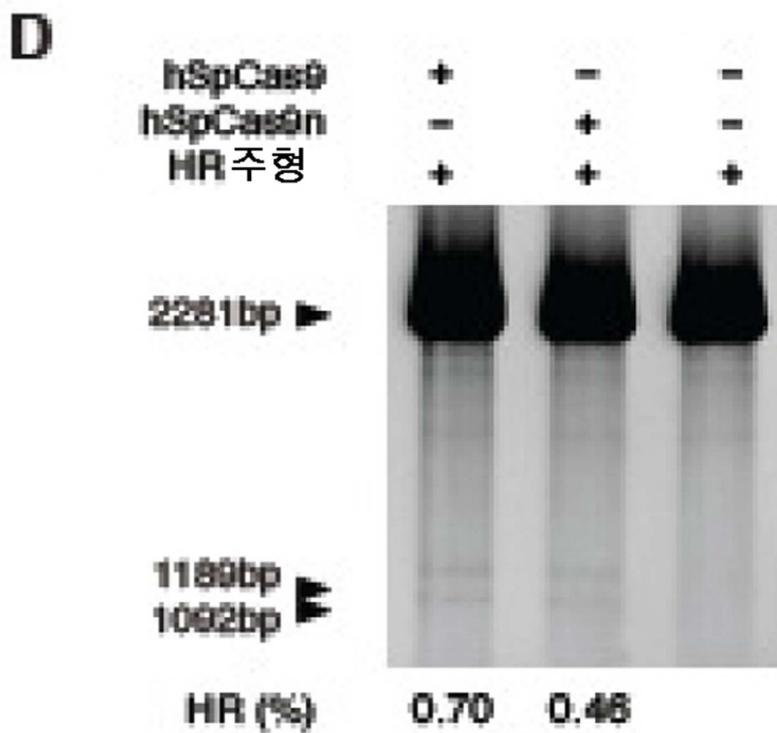
도면26b



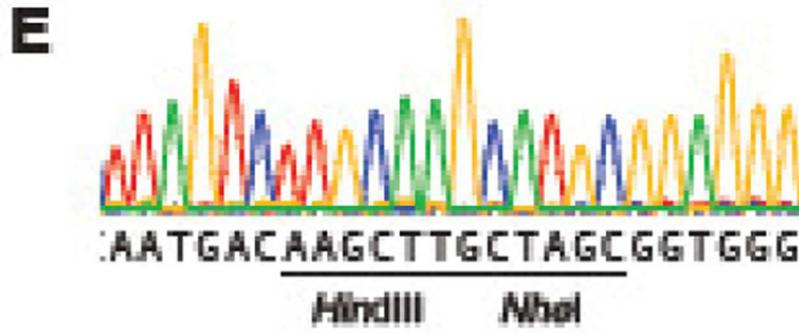
도면26c



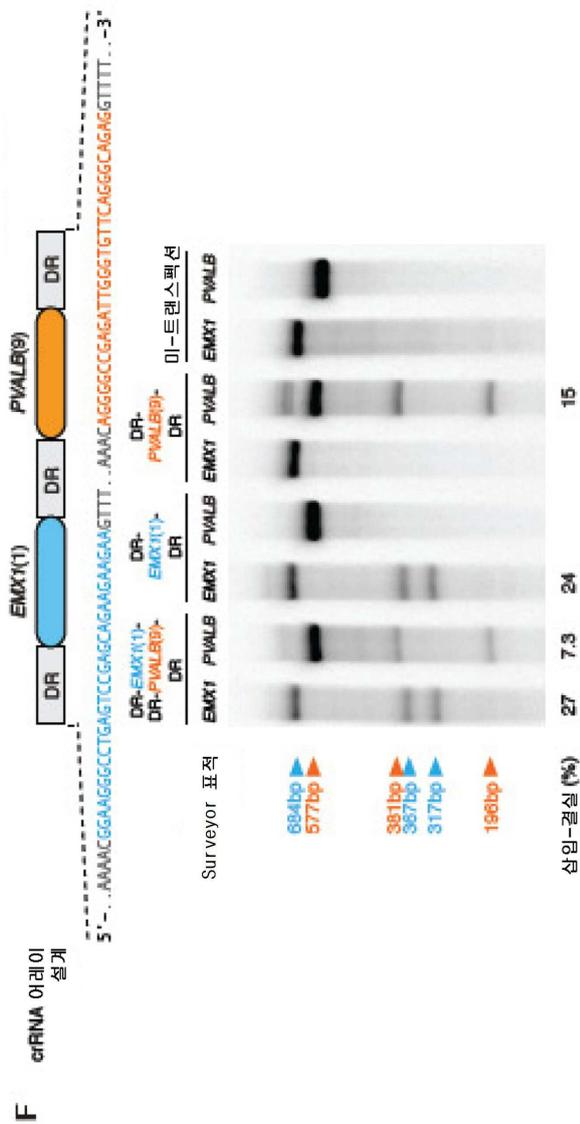
도면26d



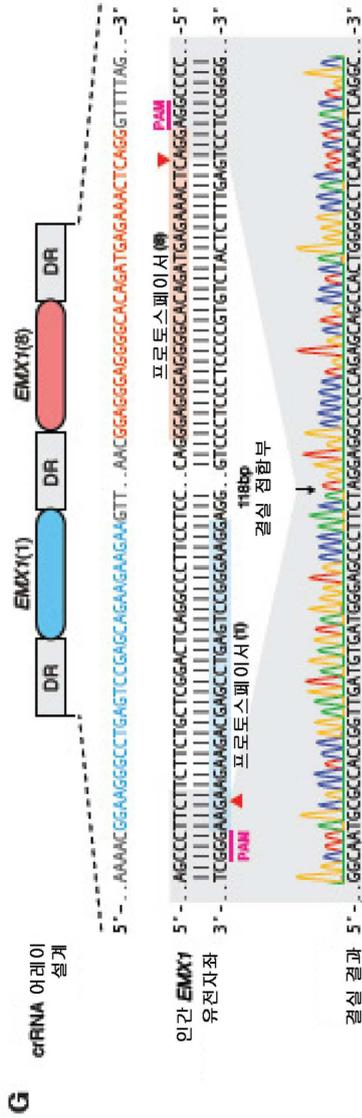
도면26e



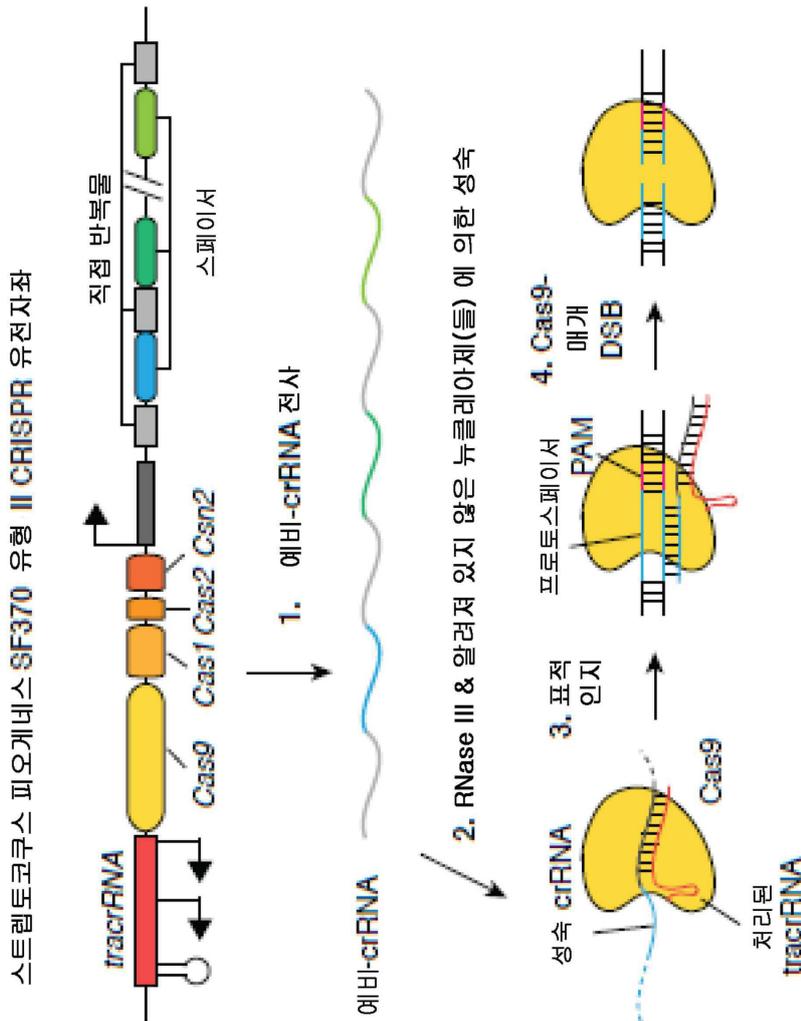
도면26f



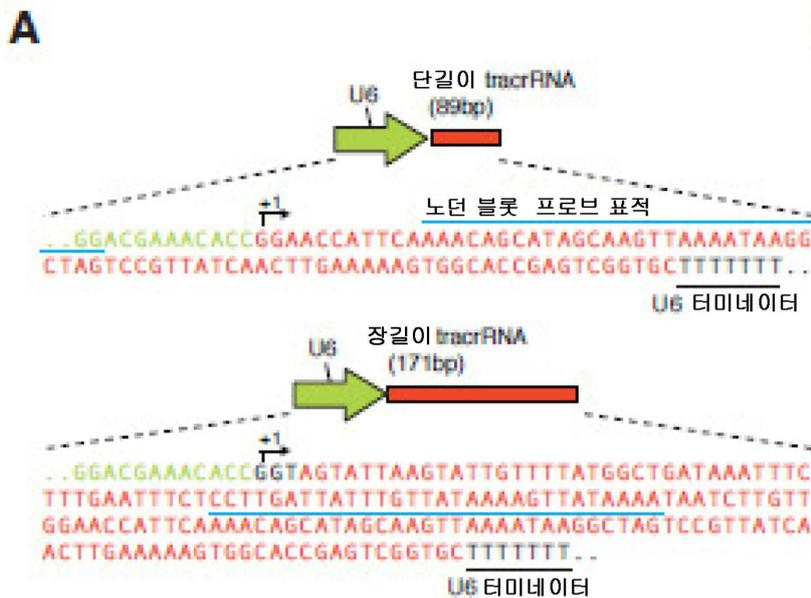
도면 26g



도면27

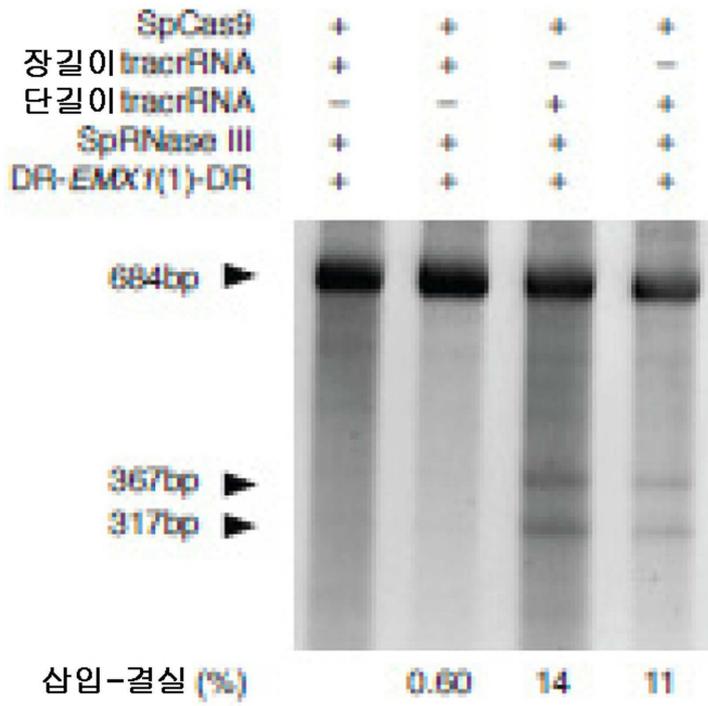


도면28a



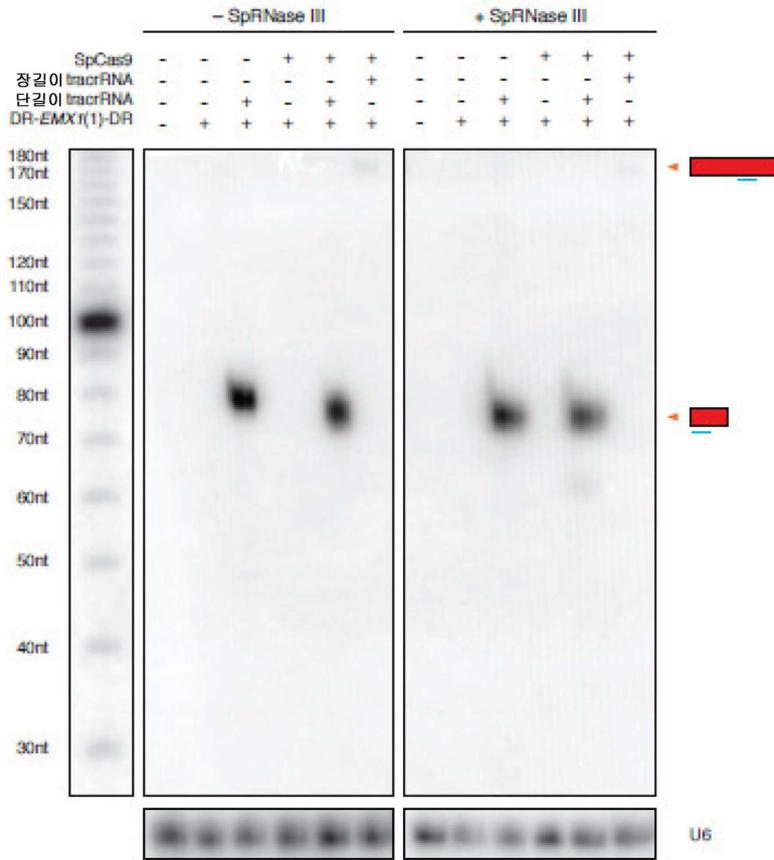
도면28b

**B**

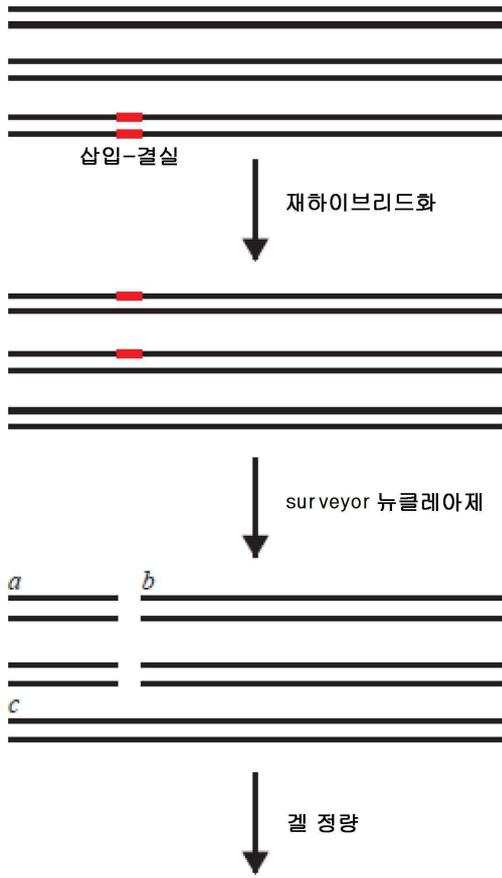


도면28c

C

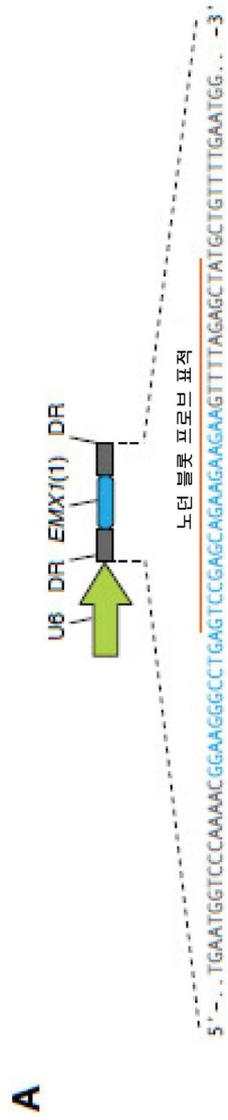


도면29



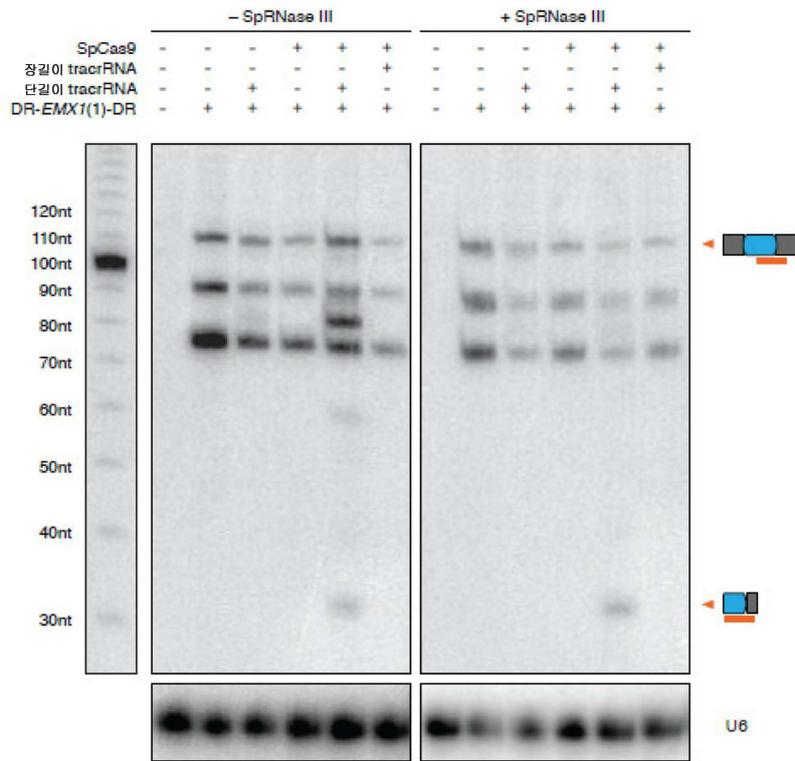
$$\text{삽입-결실 \%} = \left( 1 - \sqrt{1 - \frac{a+b}{a+b+c}} \right) * 100$$

도면30a

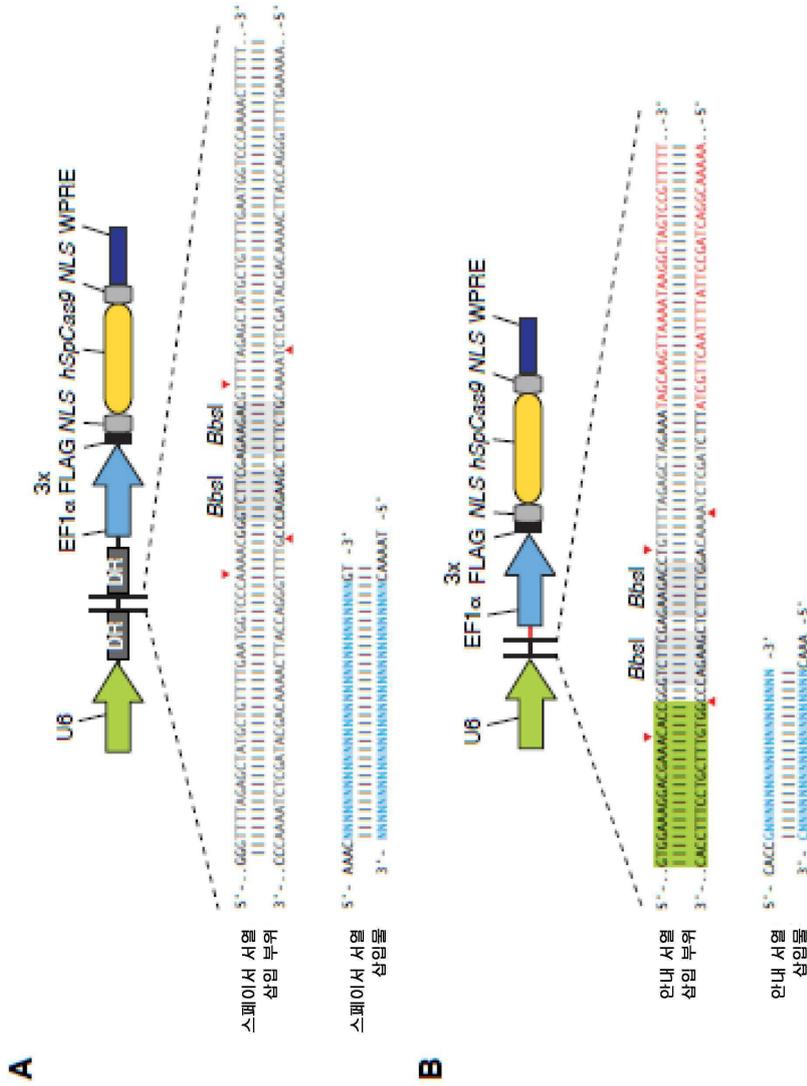


도면30b

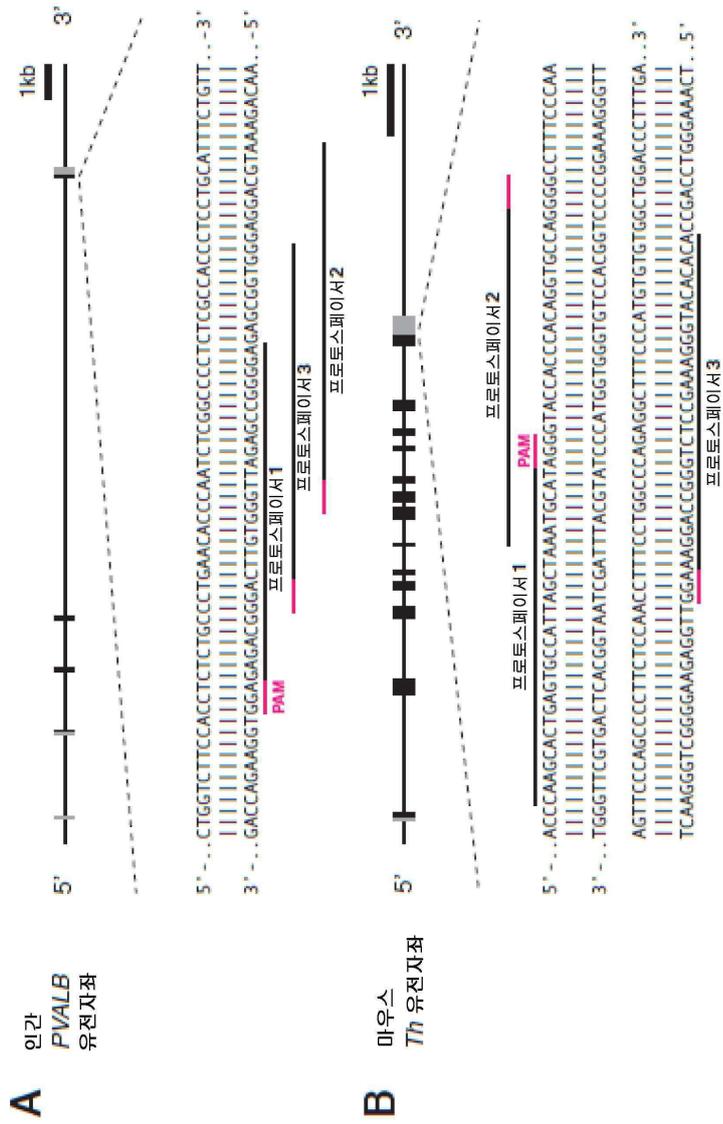
**B**



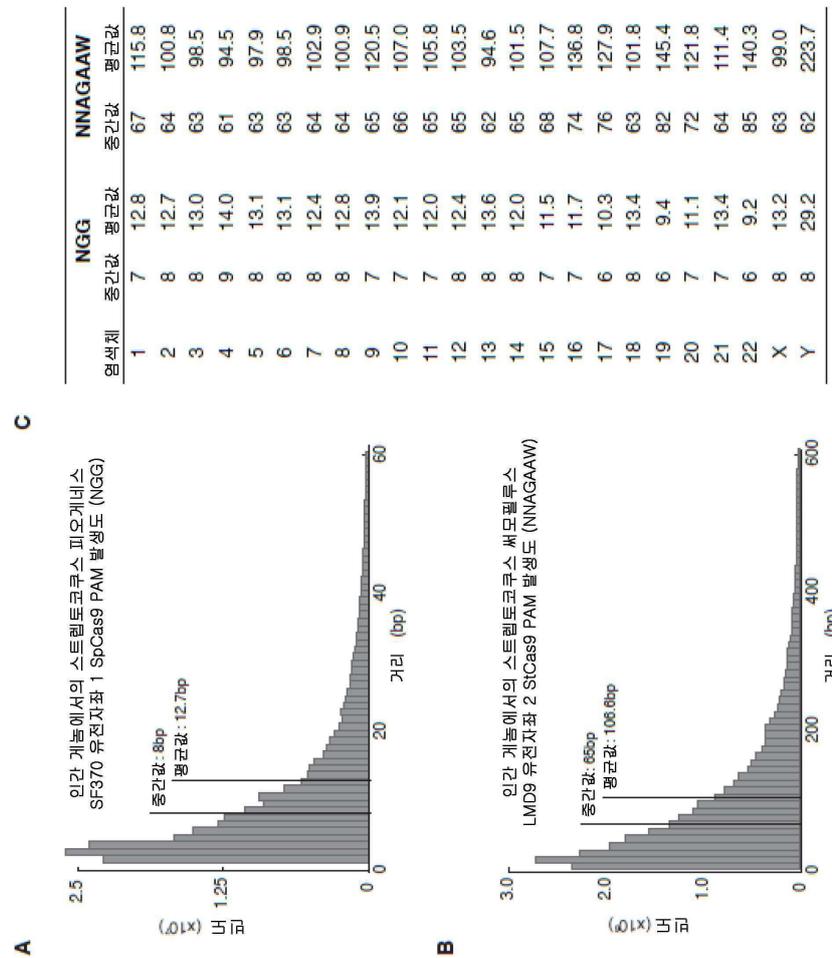
도면31



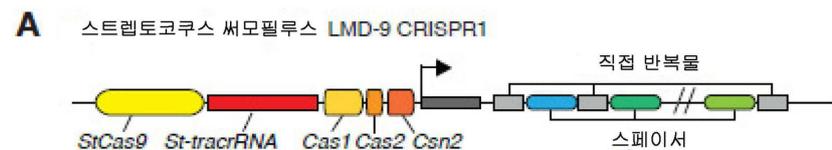
도면32



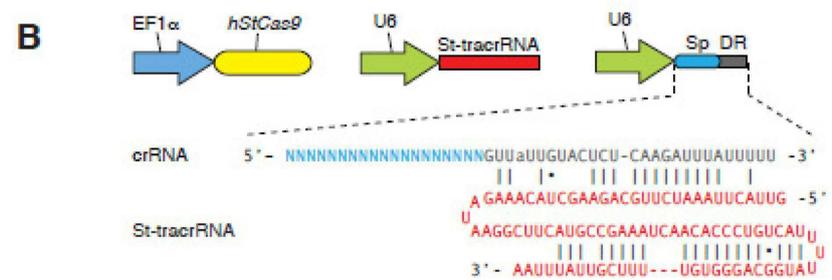
도면33



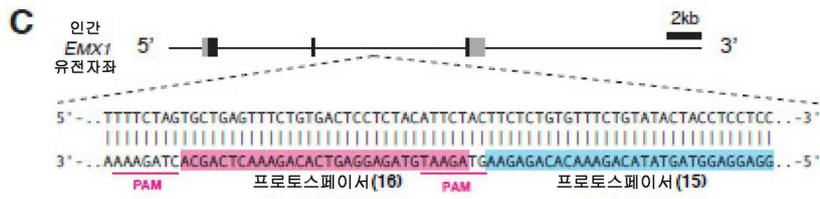
도면34a



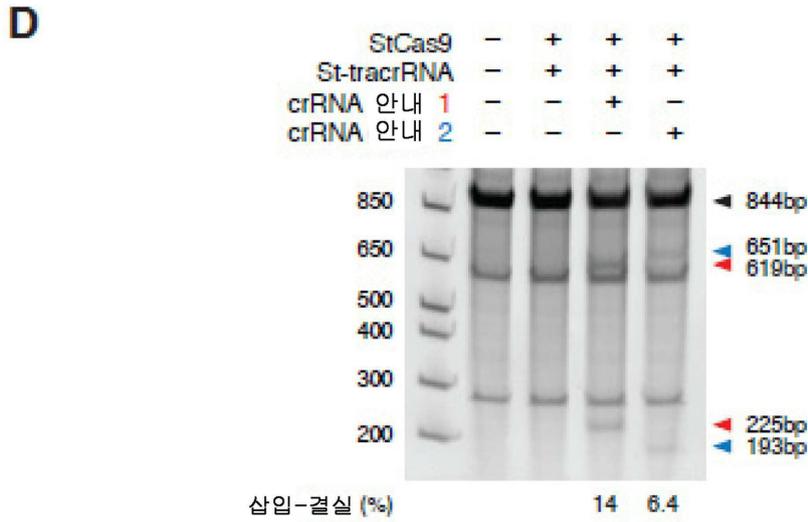
도면34b



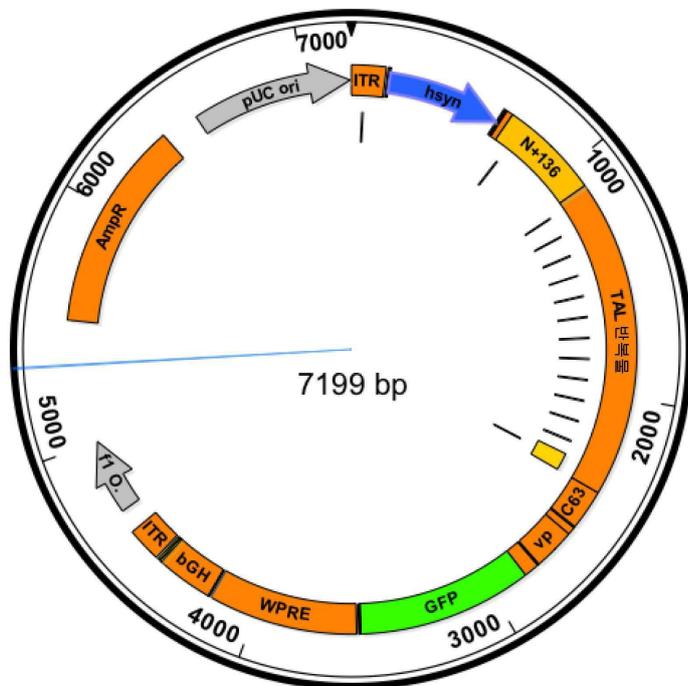
도면34c



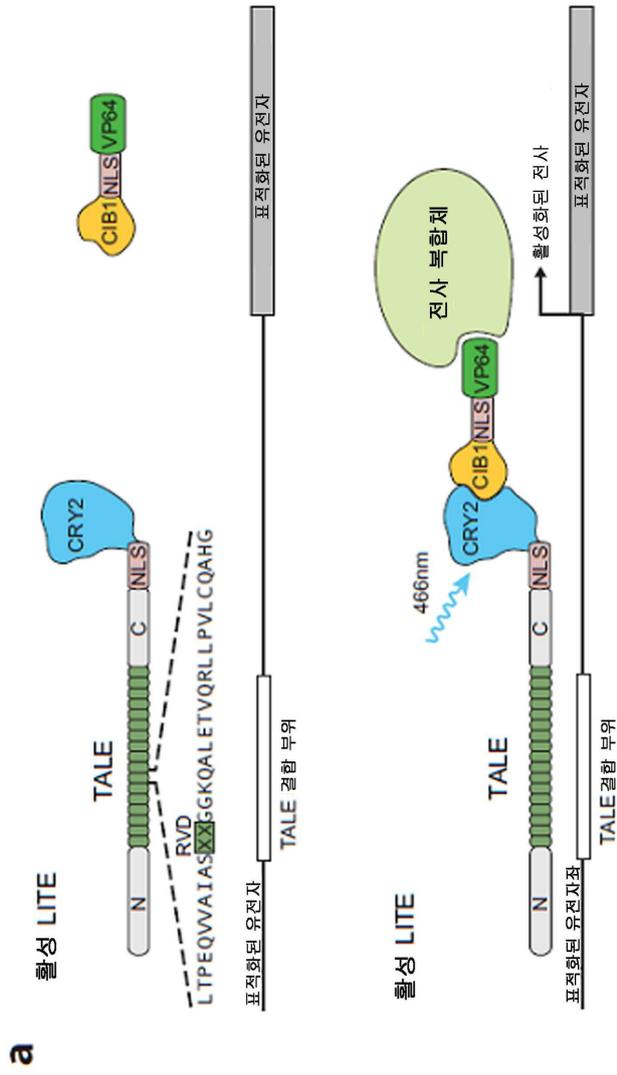
도면34d



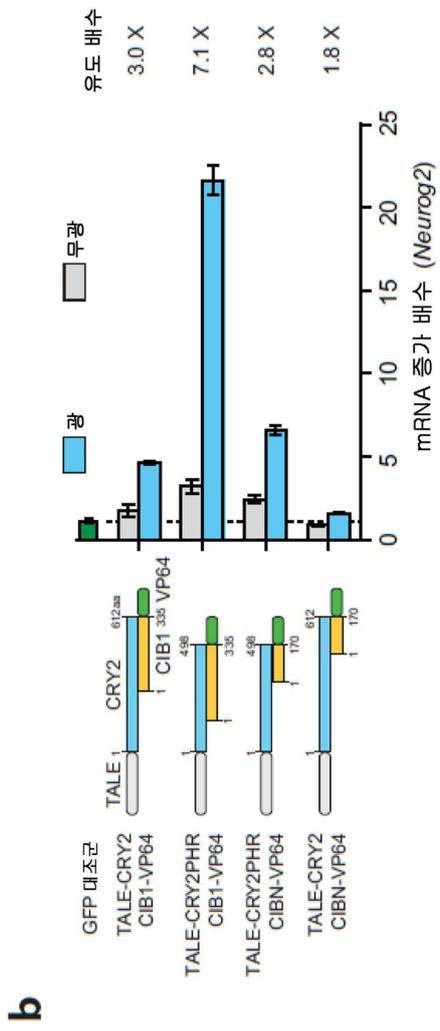
도면35



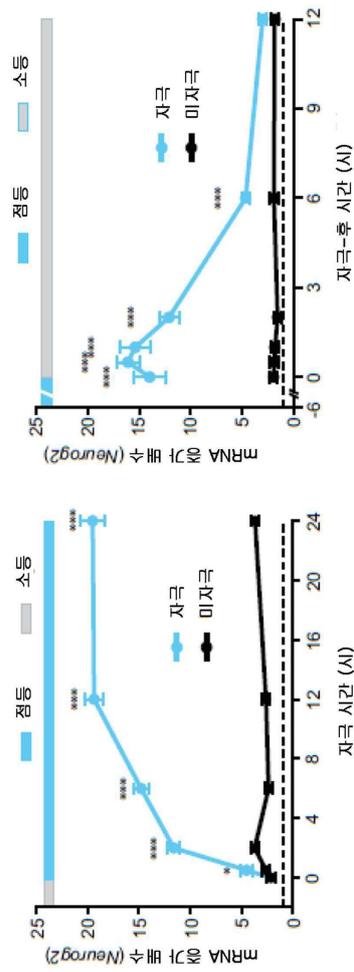
도면 36a



도면36b

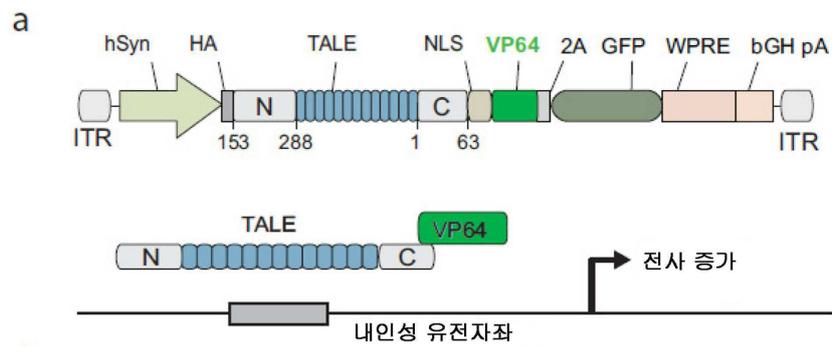


도면36c

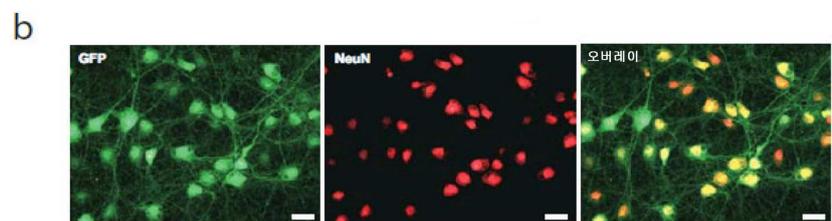


c

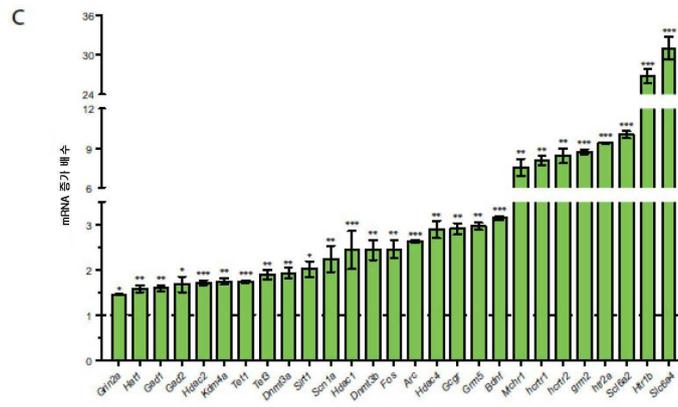
도면37a



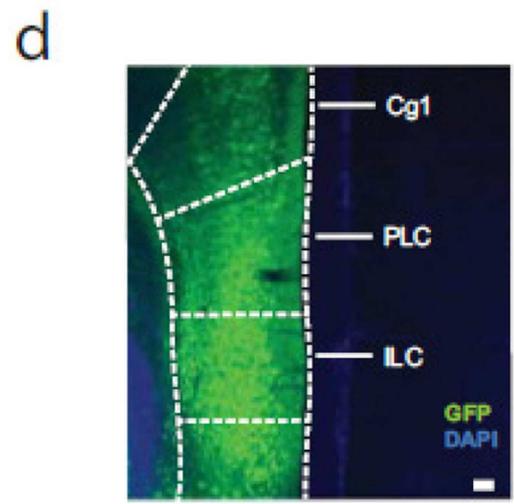
도면37b



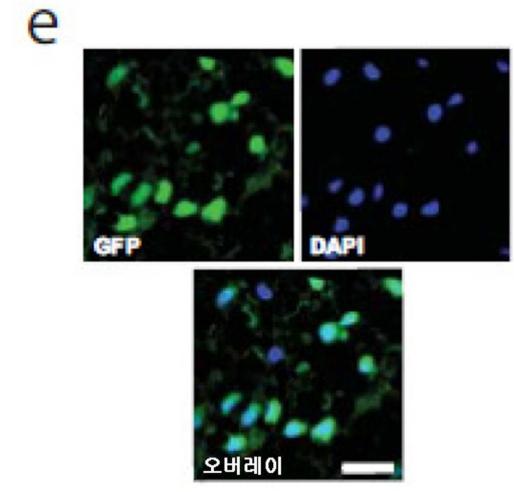
도면37c



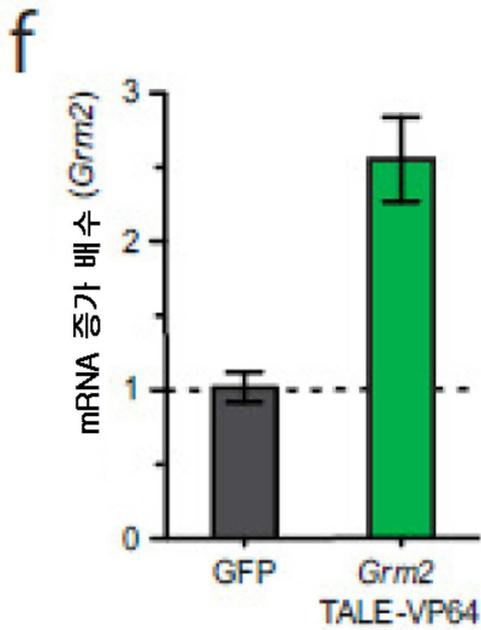
도면37d



도면37e

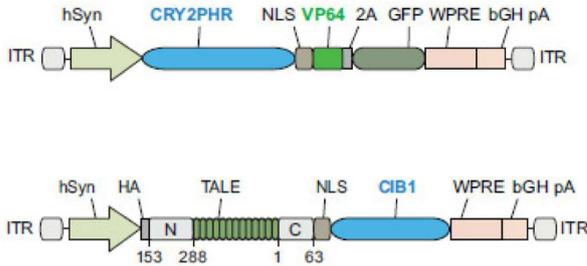


도면37f



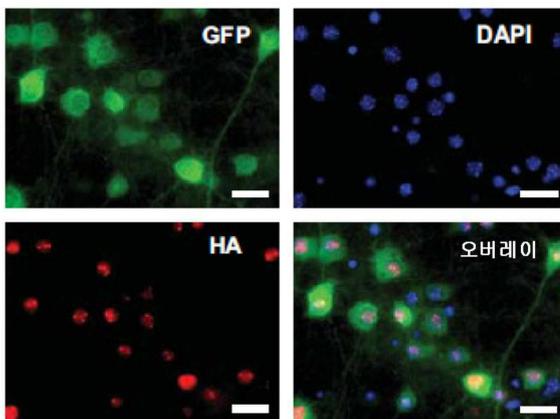
도면38a

**a**

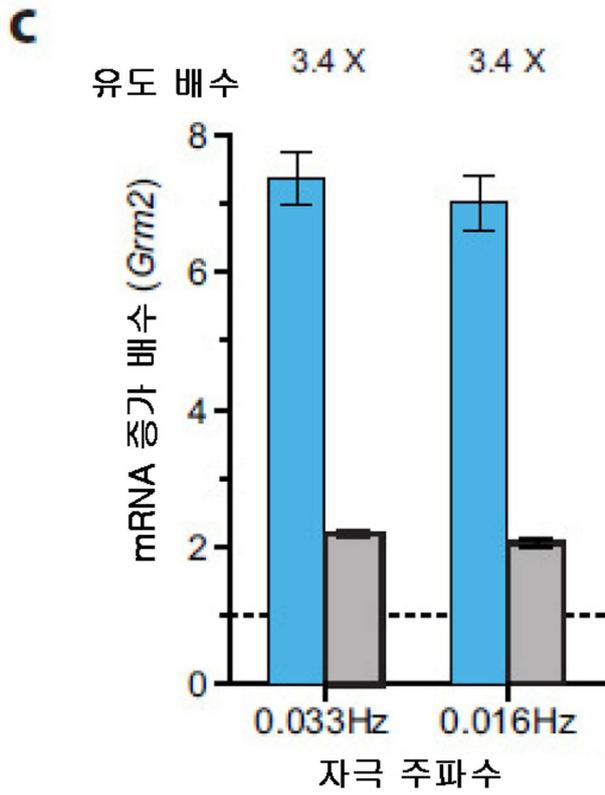


도면38b

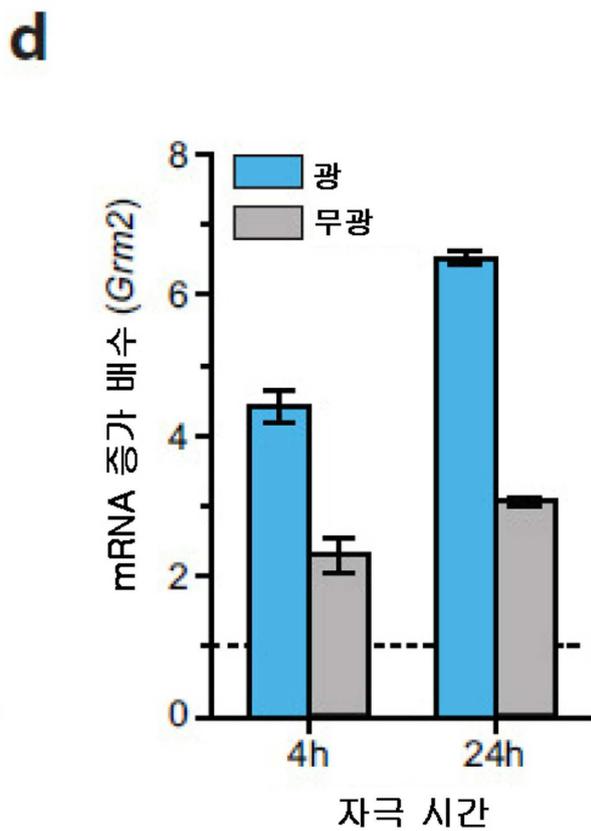
**b**



도면38c

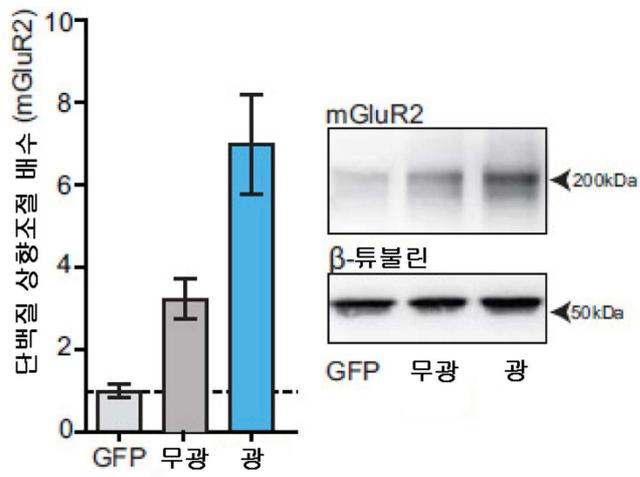


도면38d



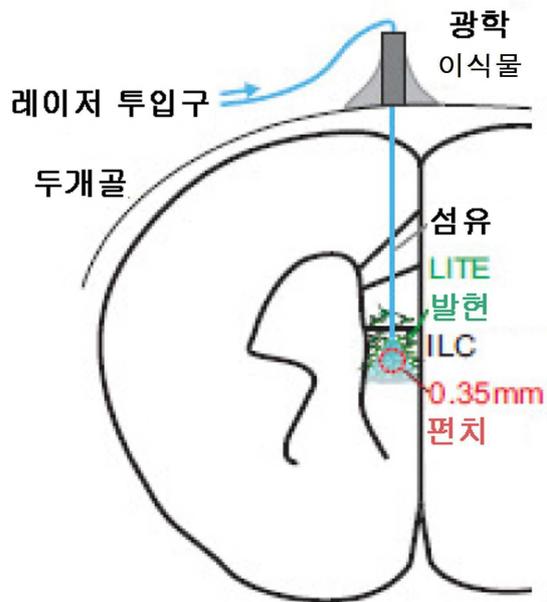
도면38e

e



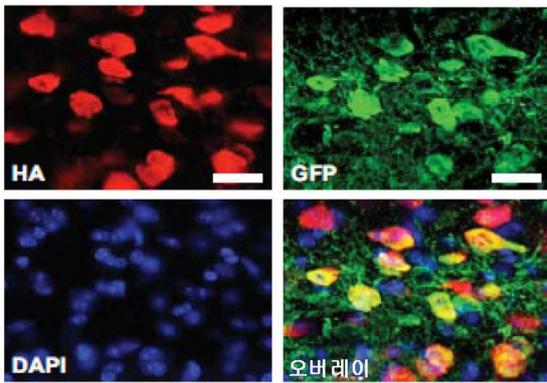
도면38f

f



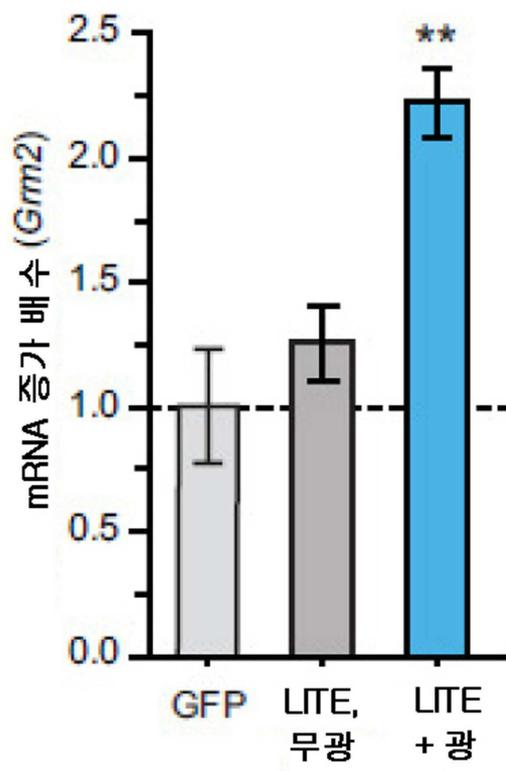
도면38g

g

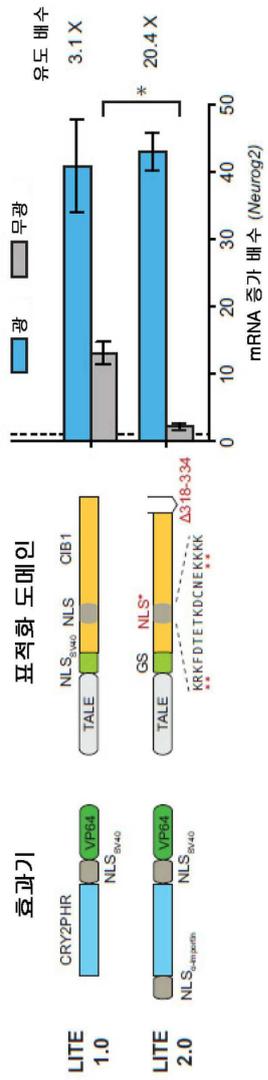


도면38h

h



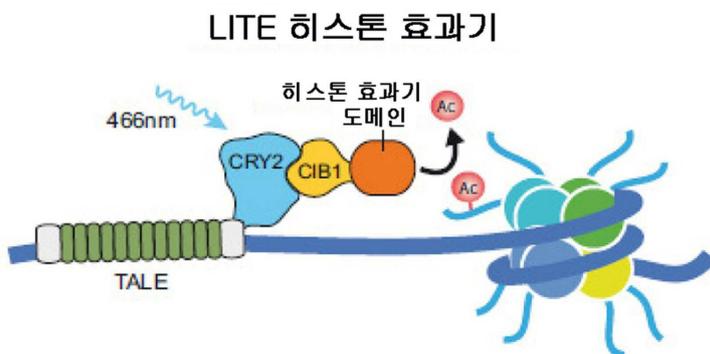
도면38i



i

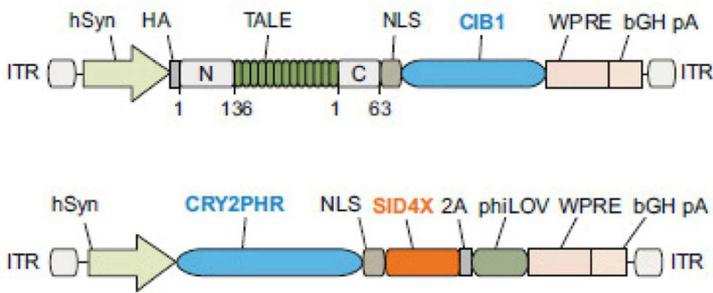
도면39a

a



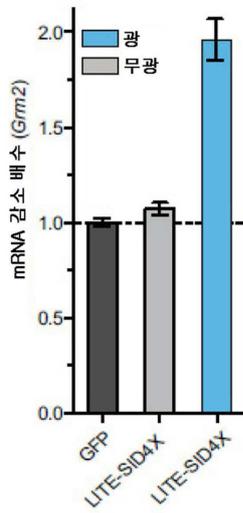
도면39b

b



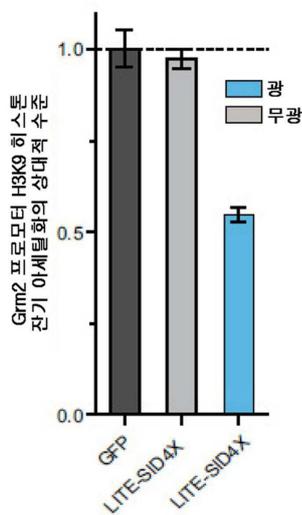
도면39c

c

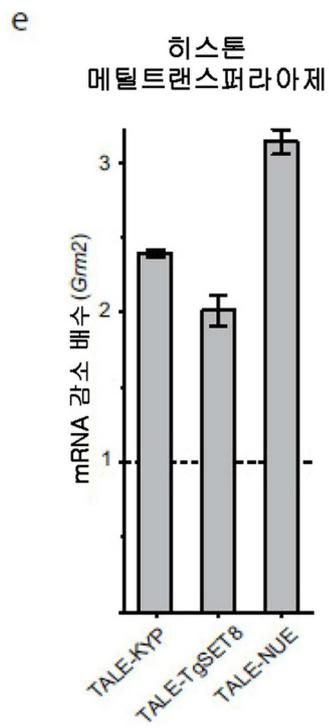


도면39d

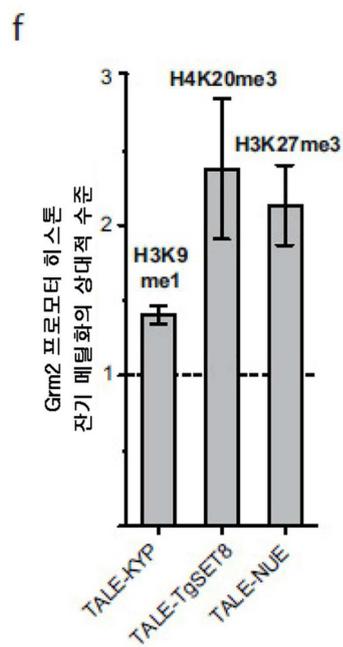
d



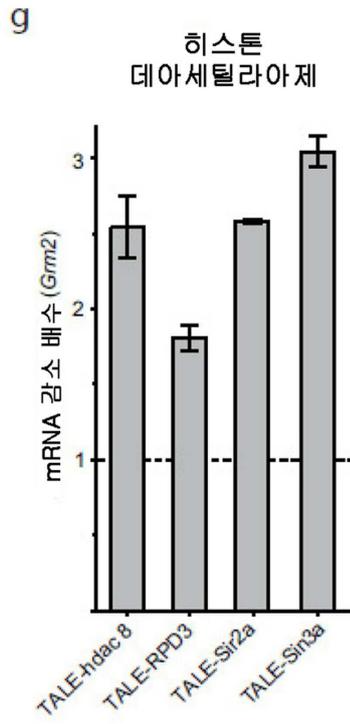
도면39e



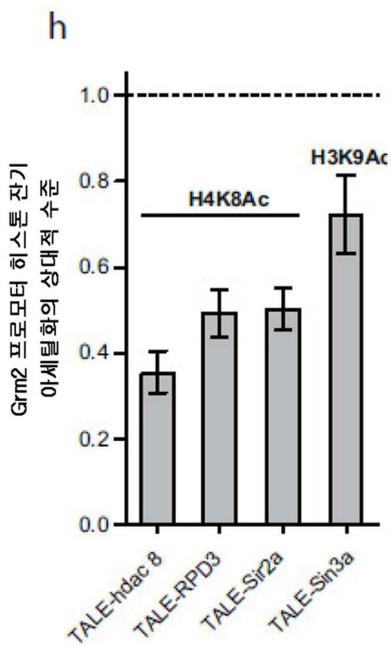
도면39f



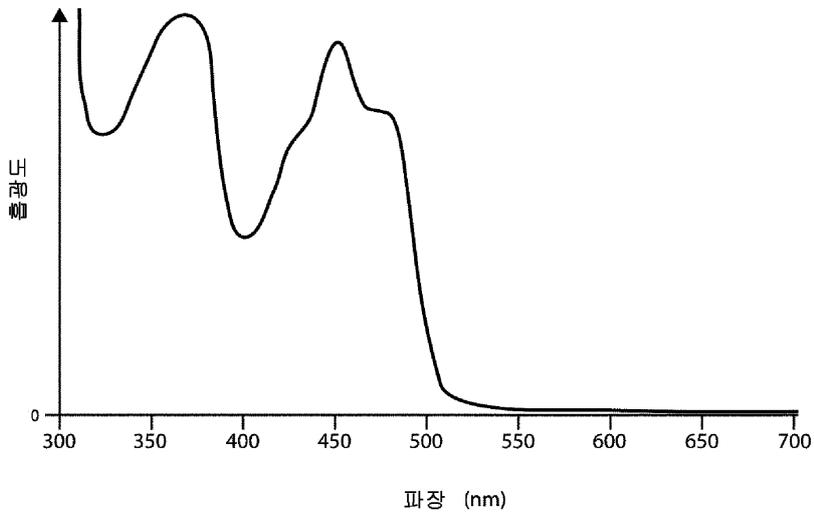
도면39g



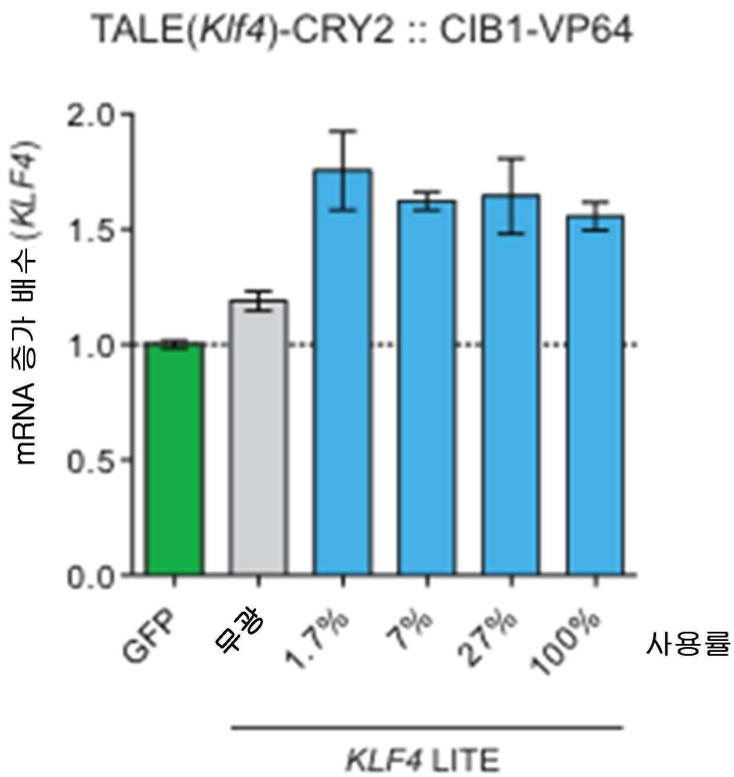
도면39h



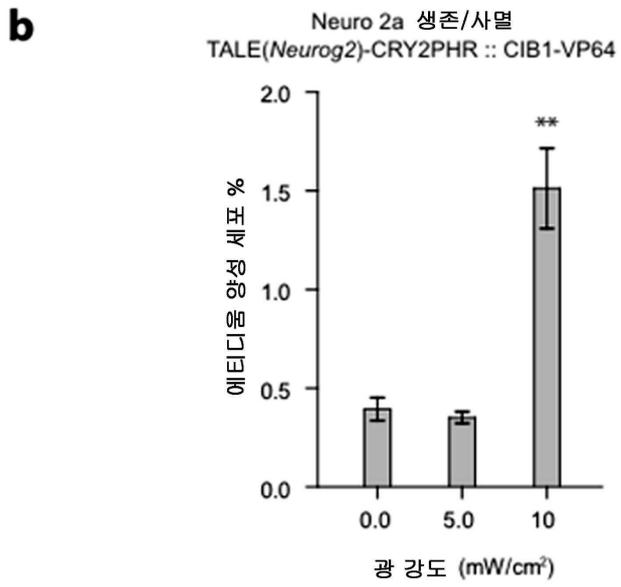
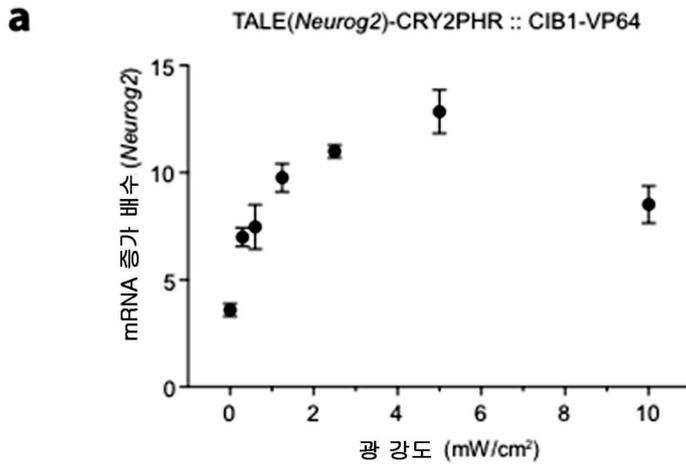
도면40



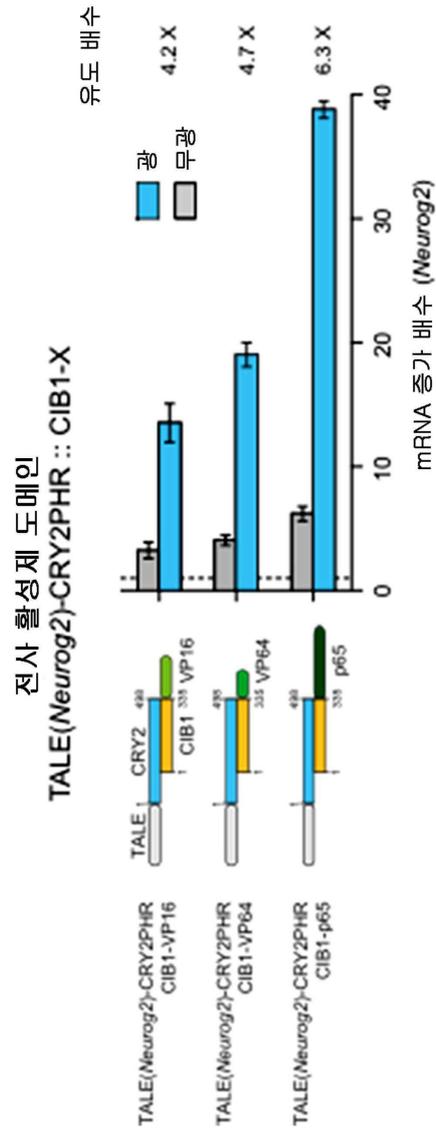
도면41



도면42

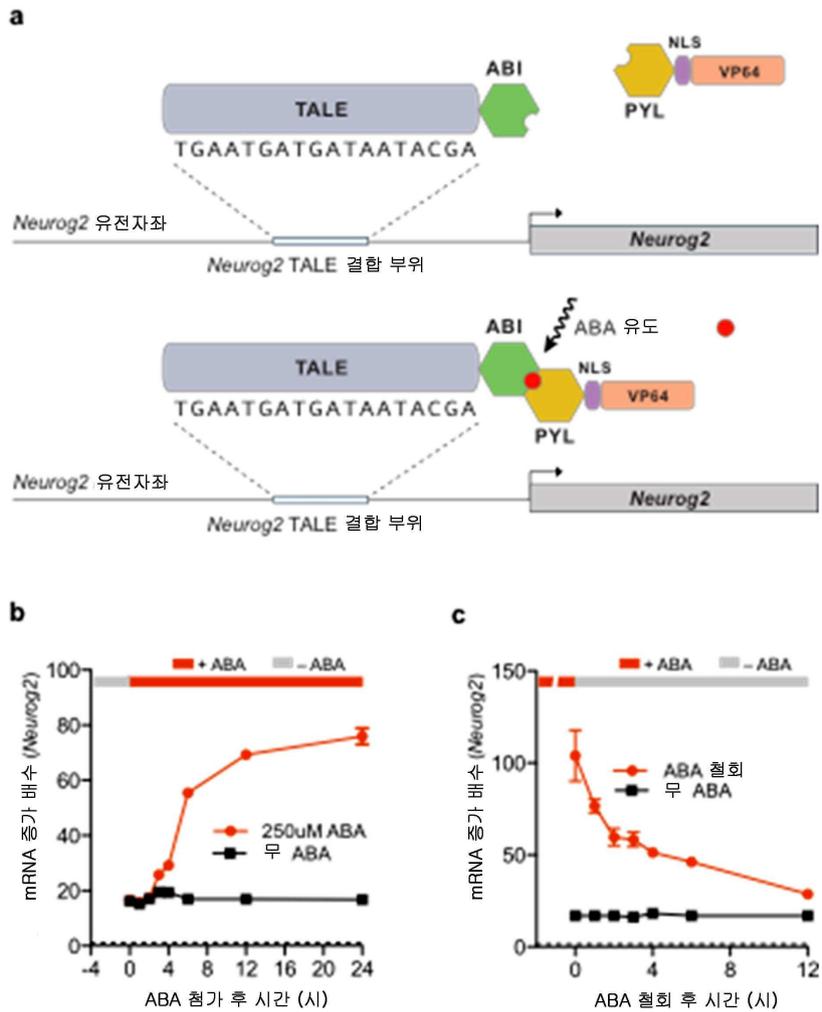


도면43

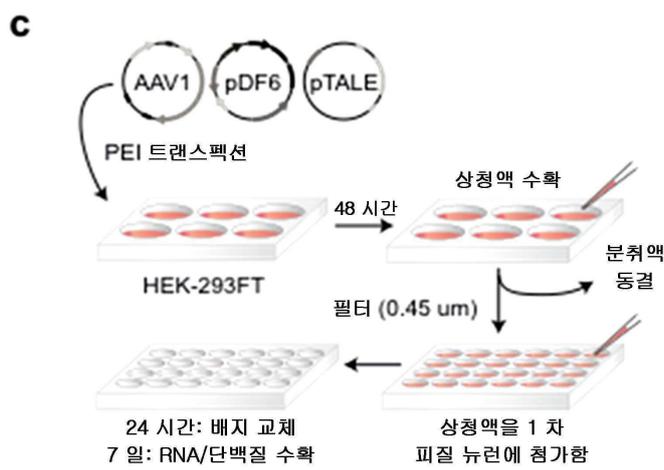
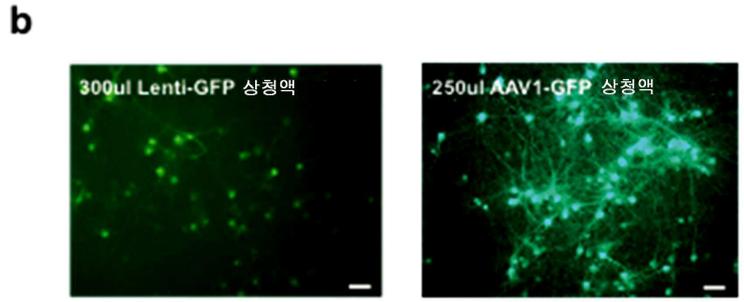
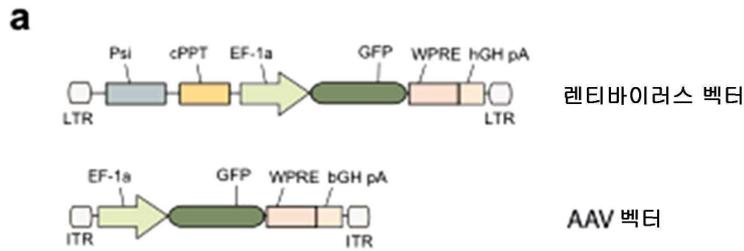


도면44

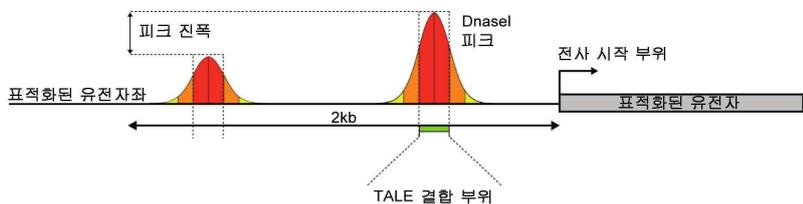
화학적 유도  
TALE(*Neurog2*)-PYL :: ABI-VP64



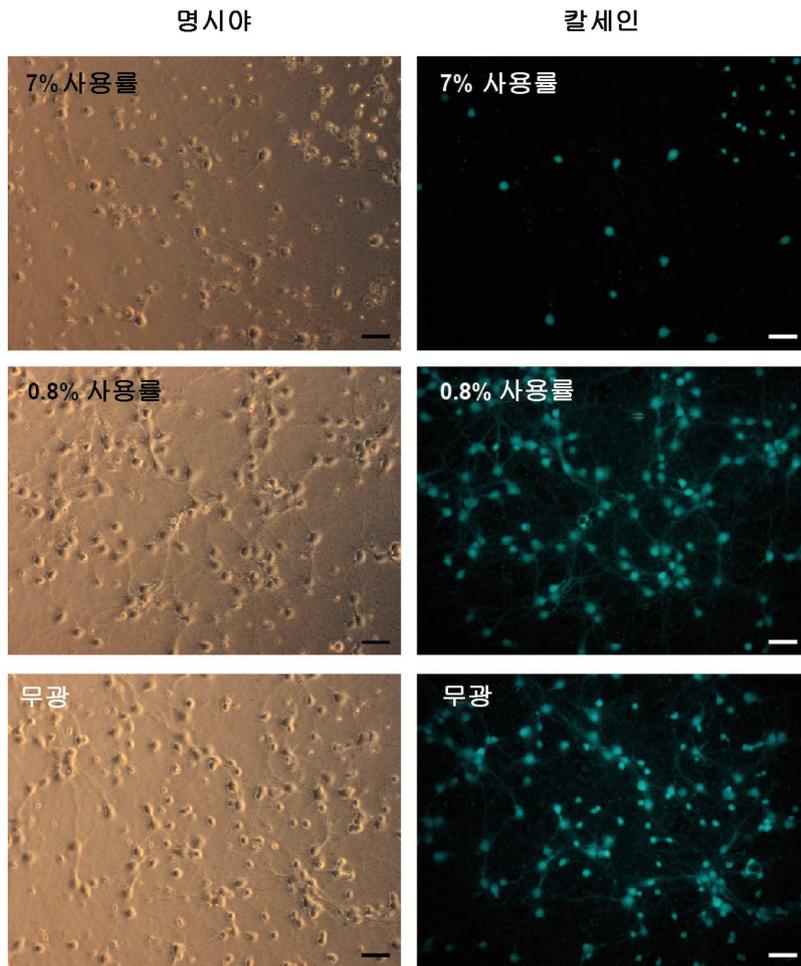
도면45



도면46



도면47



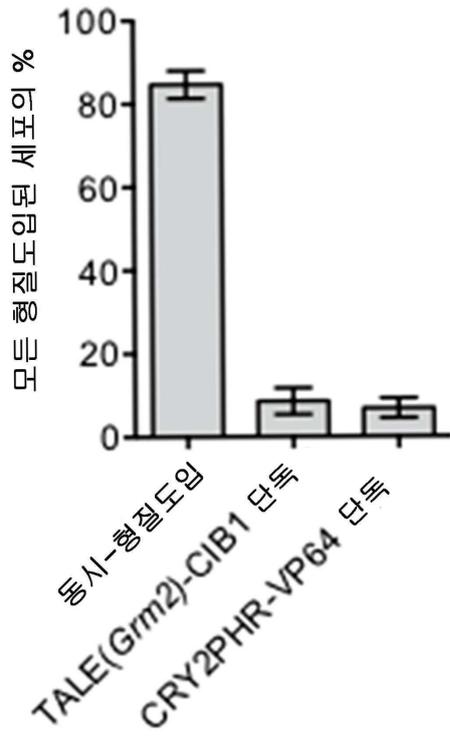
도면48

### 광유전학적 자극 동안의 마우스 이미지



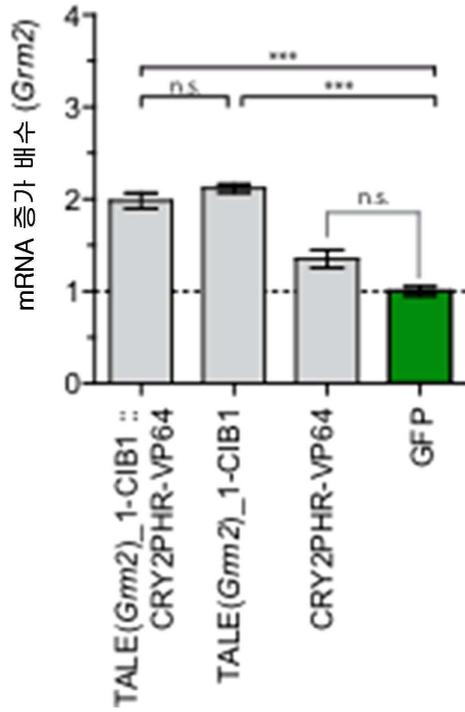
도면49

TALE(*Grm2*)-CIB1 :: CRY2PHR-VP64  
 생체내 동시-형질도입 효율

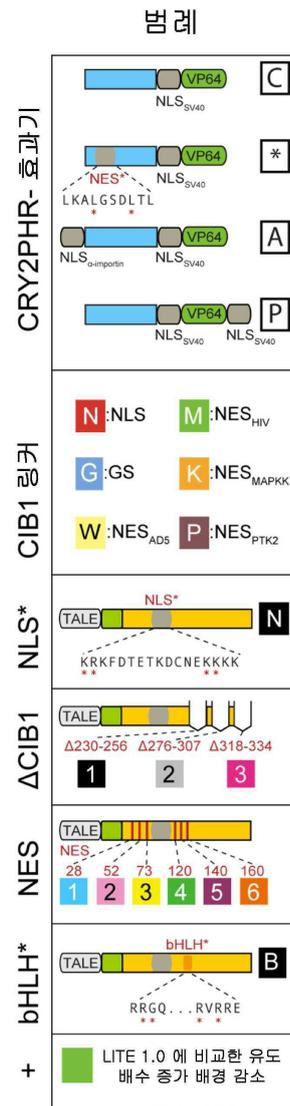


도면50

개별적  
TALE(*Grm2*)-CIB1 :: CRY2PHR-VP64  
성분의 배경 기여도

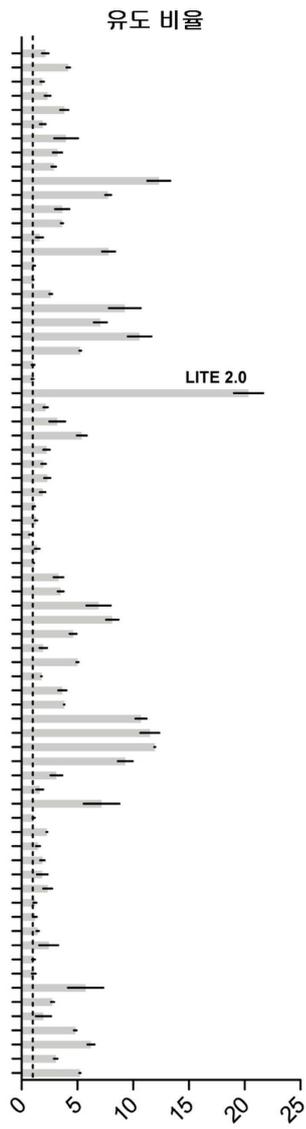


도면51a

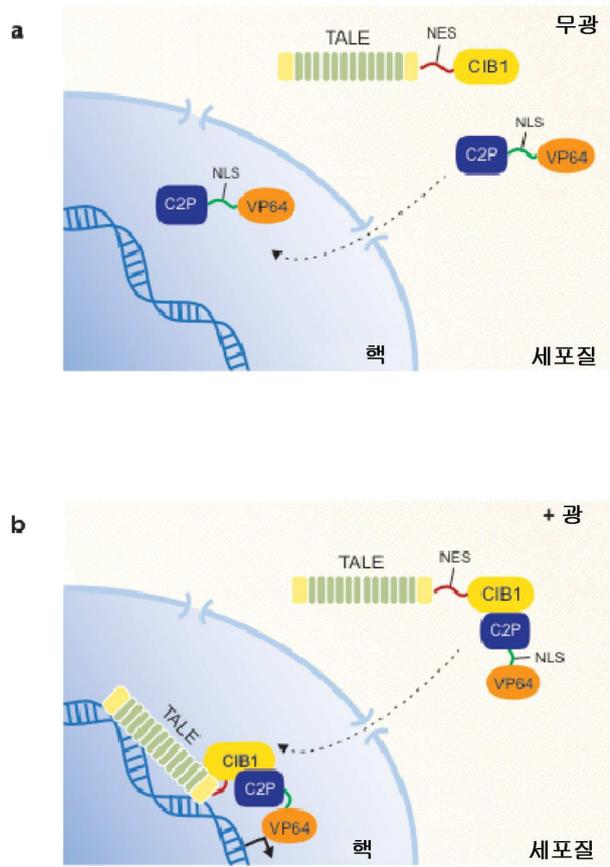




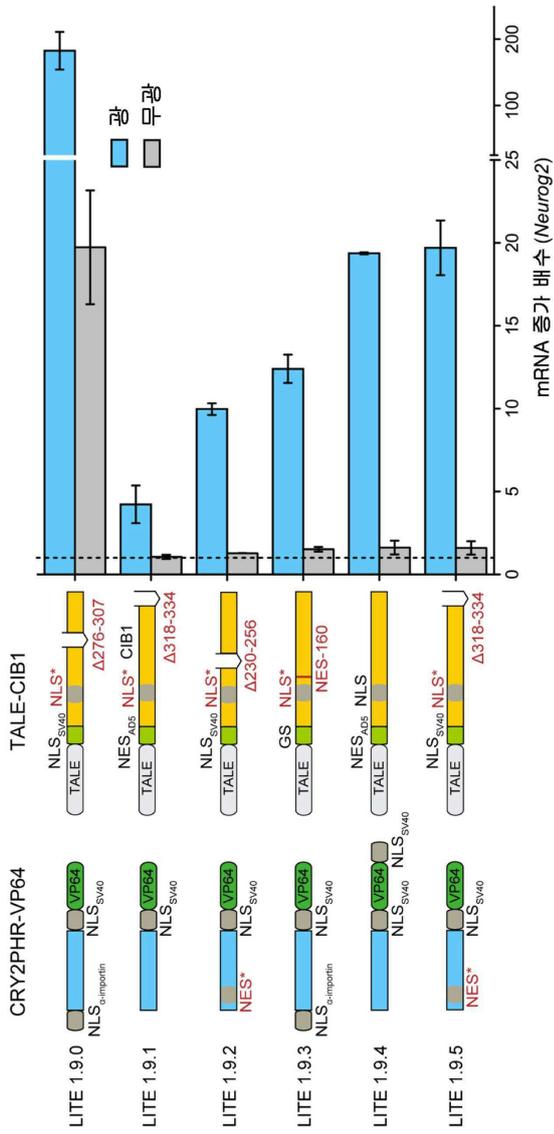
도면51c



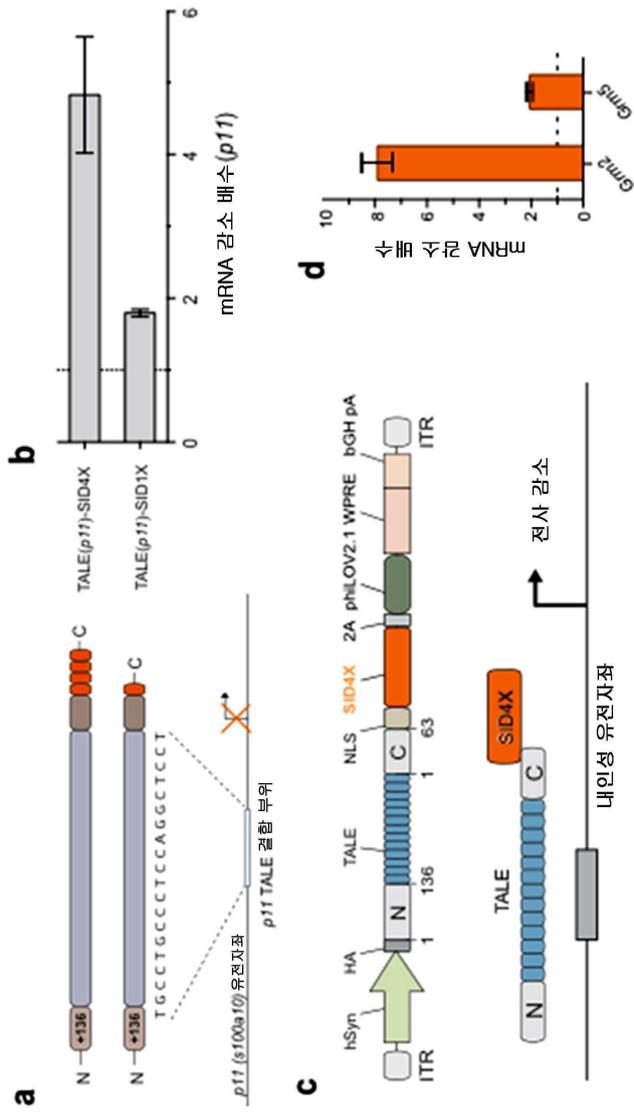
도면52



도면53

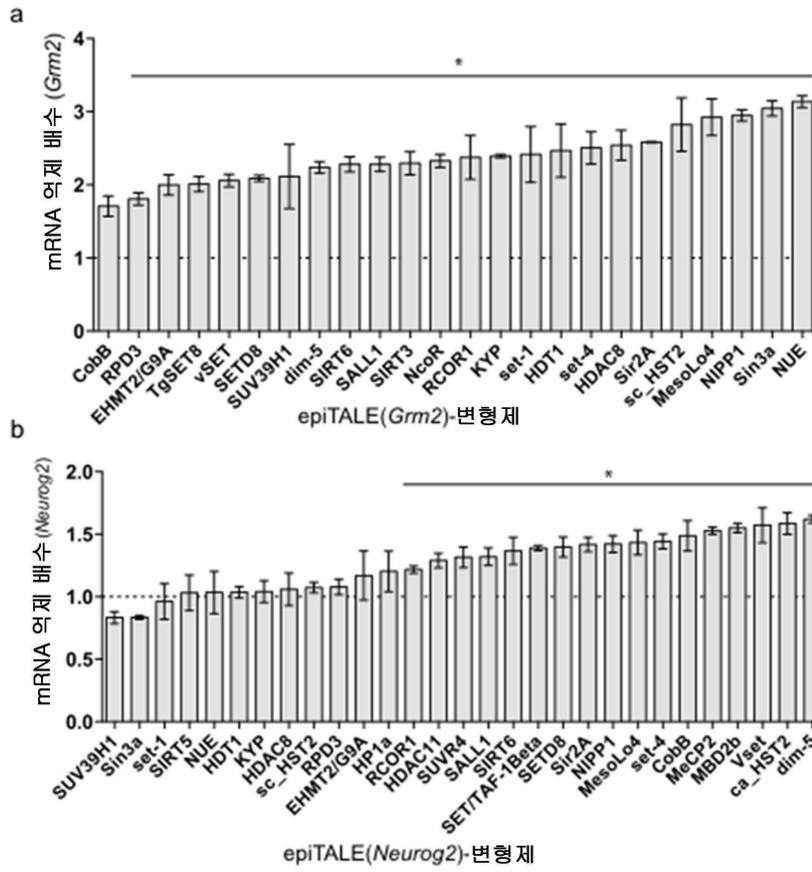


도면54



도면55

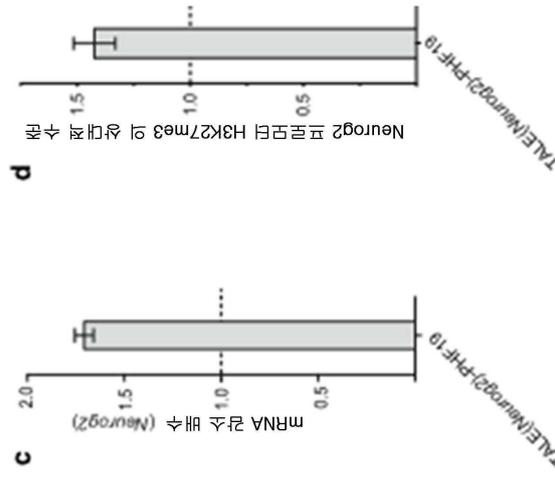
뉴런 및 Neuro 2a 세포에서의 EpiTALE 매개 억제



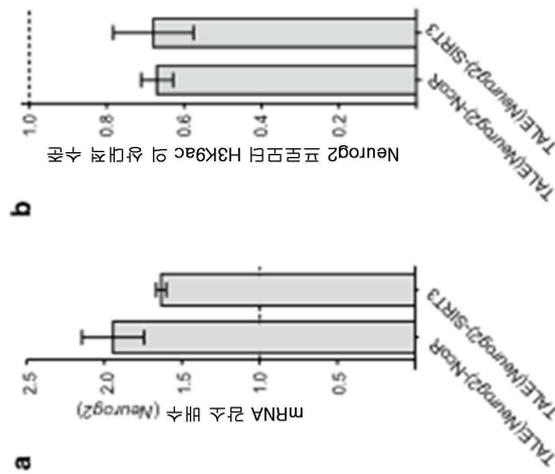
도면56

EpiTALE (Neurog2)  
매개 억제 및 히스톤 개질

히스톤 메틸화

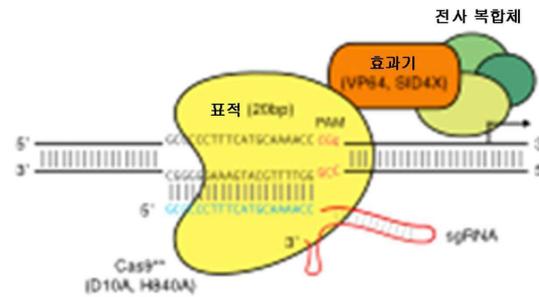


히스톤 탈아세틸화



도면57a

a



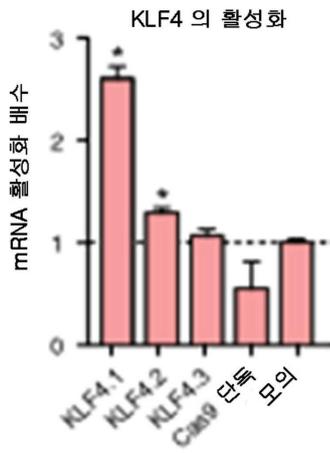
도면57b





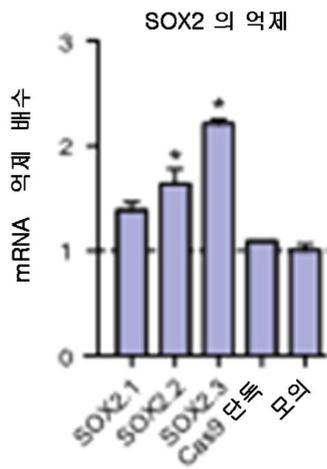
도면57f

f

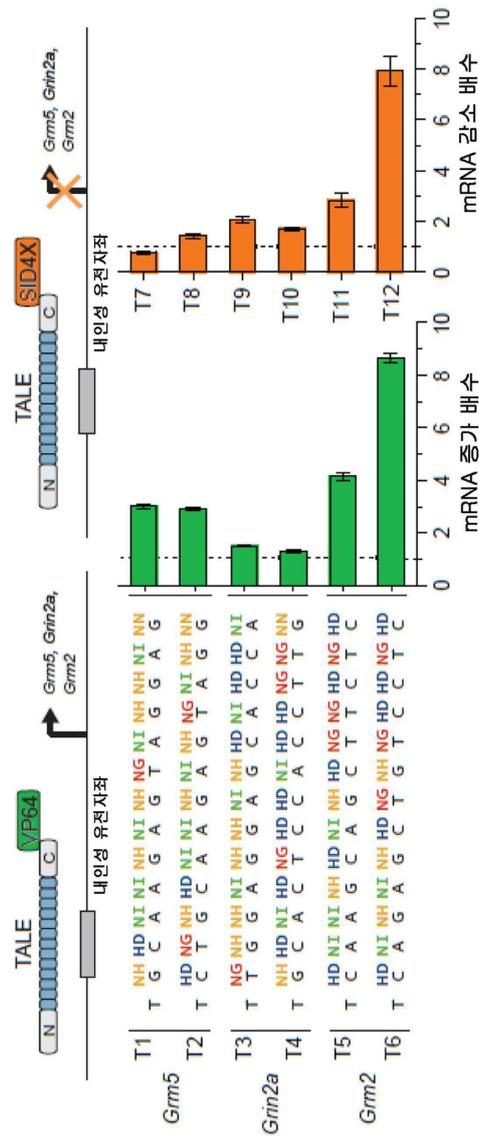


도면57g

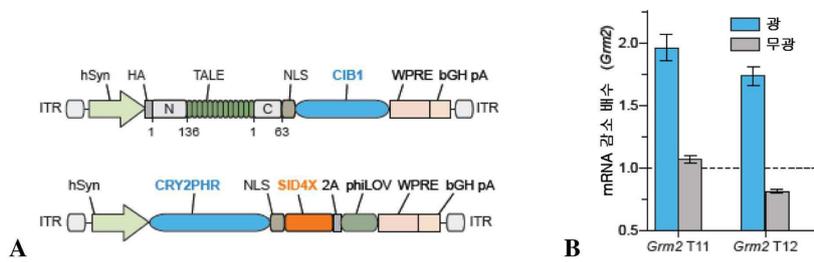
g



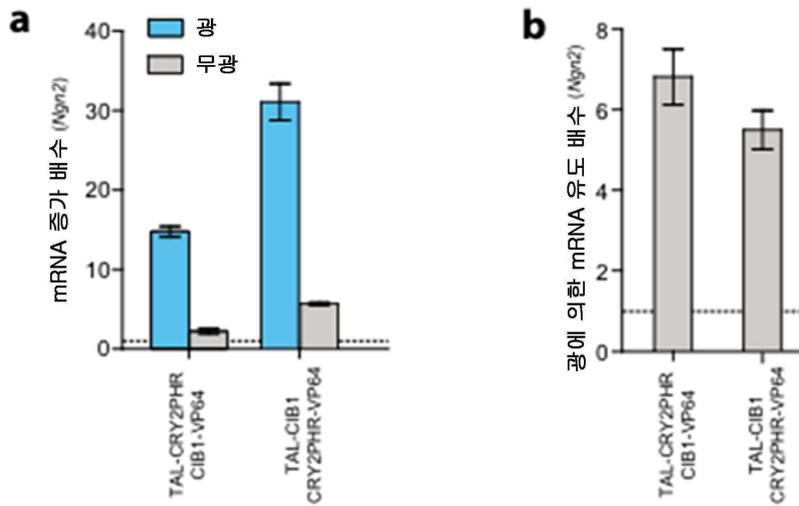
도면58



도면59



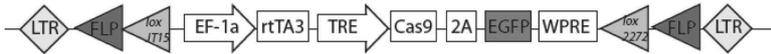
도면60



도면61

Cas9 ES 트랜스제닉 렌티바이러스 카세트

1) 바이시스트로닉 rtTA3 및 Cas9



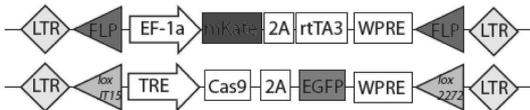
2) 자기조절 rtTA3 피드백



3) 단방향 폴리A를 갖는 역전 Cas9-EGFP 카세트

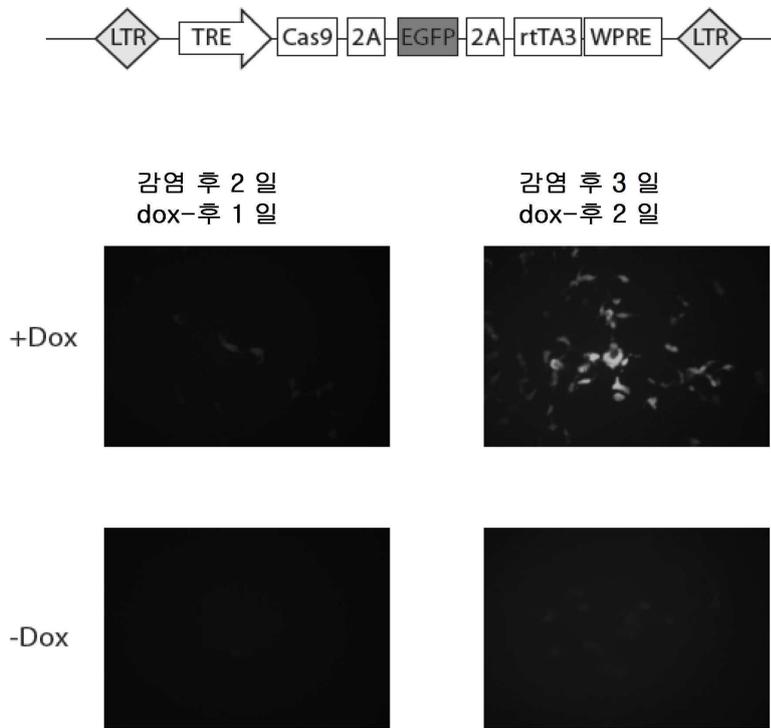


4) rtTA3 및 Cas9에 대한 별개의 바이러스



도면62

정제된 렌티바이러스를 통한 HEK293FT 에서의 Dox-민감성 자동유도제 Cas9



서열목록

SEQUENCE LISTING

<110> THE BROAD INSTITUTE, INC.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

<120> INDUCIBLE DNA BINDING PROTEINS AND GENOME PERTURBATION TOOLS AND APPLICATIONS THEREOF

<130> 44790.99.2005

<140> PCT/US2013/051418

<141> 2013-07-21

<150> 61/835,973

<151> 2013-06-17

<150> 61/794,458

<151> 2013-03-15

<150> 61/736,465

<151> 2012-12-12

<150> 61/721,283

<151> 2012-11-01

<150> 61/675,778

<151> 2012-07-25

<160> 513

<170> PatentIn version 3.5

<210> 1

<211> 4

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
peptide"

<400> 1

Leu Thr Leu Asp

1

<210> 2

<211> 4

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
peptide"

<400> 2

Leu Thr Leu Ala

1

<210> 3

<211> 4

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
  
peptide"

<400> 3

Leu Thr Gln Val

1  
 <210> 4  
 <211> 4  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 peptide"  
 <400> 4  
 Glu Gln His Gly  
 1  
 <210> 5  
 <211> 4  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 peptide"  
 <400> 5  
 Arg Asp His Gly  
 1  
 <210> 6  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence  
  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 peptide"  
 <400> 6  
 Leu Asp Leu Ala Ser Leu Ile Leu  
 1                    5  
 <210> 7  
 <211> 4  
 <212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 7

Leu Thr Pro Asp

1

<210> 8

<211> 5

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 8

Asn Gln Ala Leu Glu

1                    5

<210> 9

<211> 5

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 9

Lys Arg Ala Leu Glu

1                    5

<210> 10

<211> 4

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 10

Leu Thr Pro Glu

1

<210> 11

<211> 7

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 11

Asn Gly Lys Gln Ala Leu Glu

1                    5

<210> 12

<211> 34

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<220><221> MOD\_RES

<222> (12)..(13)

<223> Any amino acid

<400> 12

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys

1                    5                    10                    15

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp

                  20                    25                    30

His Gly

<210> 13

<211> 33

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<220><221> MOD\_RES

<222> (12)..(12)

<223> Any amino acid

<400> 13

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Xaa Gly Gly Lys Gln

1	5	10	15
Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His			
	20	25	30
Gly			

<210> 14

<211> 34

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<220><221> MOD\_RES

<222> (12)..(13)

<223> Any amino acid

<400> 14

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Xaa Xaa Gly Gly Lys

1	5	10	15
Gln Ala Leu Ala Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp			
	20	25	30
His Gly			

<210> 15

<211> 34

<212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <220><221> MOD\_RES  
 <222> (12)..(13)  
 <223> Any amino acid  
 <400> 15  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Xaa Xaa Gly Gly Lys  
  
 1                    5                    10                    15  
 Gln Ala Leu Glu Thr Leu Gln Arg Leu Leu Pro Val Leu Cys Gln Asp  
  
                   20                    25                    30  
 His Gly

<210> 16  
 <211> 34  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <220><221> MOD\_RES  
 <222> (12)..(13)  
 <223> Any amino acid  
 <400> 16  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Xaa Xaa Gly Gly Lys  
  
 1                    5                    10                    15  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp  
  
                   20                    25                    30  
 His Gly

<210> 17

<211> 34  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <220><221> MOD\_RES  
 <222> (12)..(13)  
 <223> Any amino acid  
 <400> 17  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys

1	5	10	15
Gln Ala Leu Ala Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp			
	20	25	30
His Gly			

<210> 18  
 <211> 34  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <220><221> MOD\_RES  
 <222> (12)..(13)  
 <223> Any amino acid  
 <400> 18  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys

1	5	10	15
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp			
	20	25	30
His Gly			

<210> 19  
 <211> 34  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <220><221> MOD\_RES  
 <222> (12)..(13)  
 <223> Any amino acid  
 <400> 19

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys

1	5	10	15
Gln Ala Leu Glu Thr Val Gln Arg Val Leu Pro Val Leu Cys Gln Asp			
	20	25	30
His Gly			

<210> 20  
 <211> 34  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <220><221> MOD\_RES  
 <222> (12)..(13)  
 <223> Any amino acid  
 <400> 20

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys

1	5	10	15
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala			
	20	25	30
His Gly			

<210> 21  
 <211> 34  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <220><221> MOD\_RES  
 <222> (12)..(13)  
 <223> Any amino acid  
 <400> 21  
 Leu Thr Pro Tyr Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Ser Lys  
  
 1                    5                    10                    15  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp  
  
                   20                    25                    30  
 His Gly

<210> 22  
 <211> 34  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <220><221> MOD\_RES  
 <222> (12)..(13)  
 <223> Any amino acid  
 <400> 22  
 Leu Thr Arg Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys  
  
 1                    5                    10                    15  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp  
  
                   20                    25                    30

His Gly

<210> 23

<211> 35

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<220><221> MOD\_RES

<222> (12)..(13)

<223> Any amino acid

<400> 23

Leu Ser Thr Ala Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys

1	5	10	15
Gln	Ala	Leu	Glu
Gly	Ile	Gly	Glu
Gln	Leu	Leu	Lys
Leu	Arg	Thr	Ala
20	25	30	
Pro	Tyr	Gly	
35			

<210> 24

<211> 35

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<220><221> MOD\_RES

<222> (12)..(13)

<223> Any amino acid

<400> 24

Leu Ser Thr Ala Gln Val Val Ala Val Ala Ser Xaa Xaa Gly Gly Lys

1	5	10	15
Pro	Ala	Leu	Glu
Ala	Val	Arg	Ala
Gln	Leu	Leu	Ala
Leu	Arg	Ala	Ala

20 25 30

Pro Tyr Gly  
35

<210> 25  
<211> 288  
<212> PRT  
<213> Artificial Sequence  
<220><221> source  
<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"  
<400> 25

Met Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu Leu  
1 5 10 15

Ser Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly Val  
20 25 30

Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg Thr  
35 40 45

Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala Phe  
50 55 60

Ser Ala Asp Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser Leu  
65 70 75 80

Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His His  
85 90 95

Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu Arg  
100 105 110

Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala Ala  
115 120 125

Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln Pro  
130 135 140

Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr  
145 150 155 160

Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val

165 170 175  
 Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His  
 180 185 190  
 Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val  
 195 200 205

Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala  
 210 215 220  
 Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala  
 225 230 235 240  
 Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp  
 245 250 255  
 Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val  
 260 265 270

Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn  
 275 280 285

<210> 26  
 <211> 183  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 26

Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro  
 1 5 10 15  
 Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu  
 20 25 30

Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala  
 35 40 45  
 Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser  
 50 55 60  
 His Arg Val Ala Asp His Ala Gln Val Val Arg Val Leu Gly Phe Phe

65                    70                    75                    80  
Gln Cys His Ser His Pro Ala Gln Ala Phe Asp Asp Ala Met Thr Gln  
                          85                    90                    95  
  
Phe Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe Arg Arg Val Gly  
                          100                    105                    110  
Val Thr Glu Leu Glu Ala Arg Ser Gly Thr Leu Pro Pro Ala Ser Gln  
                          115                    120                    125  
Arg Trp Asp Arg Ile Leu Gln Ala Ser Gly Met Lys Arg Ala Lys Pro  
                          130                    135                    140  
Ser Pro Thr Ser Thr Gln Thr Pro Asp Gln Ala Ser Leu His Ala Phe  
145                    150                    155                    160

Ala Asp Ser Leu Glu Arg Asp Leu Asp Ala Pro Ser Pro Met His Glu  
                          165                    170                    175  
Gly Asp Gln Thr Arg Ala Ser  
                          180

<210> 27

<211> 18

<212> DNA

<213> Mus sp.

<400> 27

tgaatgatga taatacga

18

<210> 28

<211> 11

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 28

Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser

1                    5                    10

<210> 29

<211> 14

<212> DNA

<213> Homo sapiens

<400> 29

ttcttactta taac

14

<210> 30

<211> 7

<212> PRT

<213> Simian virus 40

<400> 30

Pro Lys Lys Lys Arg Lys Val

1                    5

<210> 31

<211> 16

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown: Nucleoplasmin bipartite  
NLS peptide"

<400> 31

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys

1                    5                    10                    15

<210> 32

<211> 9

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown: C-myc NLS peptide"

<400> 32

Pro Ala Ala Lys Arg Val Lys Leu Asp

1                    5

<210> 33

<211> 11

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown: C-myc NLS peptide"

<400> 33

Arg Gln Arg Arg Asn Glu Leu Lys Arg Ser Pro

1                    5                    10

<210> 34

<211> 38

<212> PRT

<213> Homo sapiens

<400> 34

Asn Gln Ser Ser Asn Phe Gly Pro Met Lys Gly Gly Asn Phe Gly Gly

1                    5                    10                    15

Arg Ser Ser Gly Pro Tyr Gly Gly Gly Gly Gln Tyr Phe Ala Lys Pro

                  20                    25                    30

Arg Asn Gln Gly Gly Tyr

                  35

<210> 35

<211> 42

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown: IBB domain from importin-alpha

NLS polypeptide"

<400> 35

Arg Met Arg Ile Glx Phe Lys Asn Lys Gly Lys Asp Thr Ala Glu Leu

1                    5                    10                    15

Arg Arg Arg Arg Val Glu Val Ser Val Glu Leu Arg Lys Ala Lys Lys

                  20                    25                    30

Asp Glu Gln Ile Leu Lys Arg Arg Asn Val

                  35                    40

<210> 36

<211> 8

<212> PRT  
 <213> Unknown  
 <220><221> source  
 <223> /note="Description of Unknown: Myoma T NLS peptide"

<400> 36  
 Val Ser Arg Lys Arg Pro Arg Pro  
 1                    5

<210> 37  
 <211> 8  
 <212> PRT  
 <213> Unknown  
 <220><221> source  
 <223> /note="Description of Unknown: Myoma T NLS peptide"

<400> 37  
 Pro Pro Lys Lys Ala Arg Glu Asp  
 1                    5

<210> 38  
 <211> 7  
 <212> PRT  
 <213> Homo sapiens  
 <400> 38  
 Pro Gln Pro Lys Lys Lys Pro

1                    5  
 <210> 39  
 <211> 12  
 <212> PRT  
 <213> Mus sp.  
 <400> 39

Ser Ala Leu Ile Lys Lys Lys Lys Lys Met Ala Pro  
 1                    5                    10

<210> 40  
 <211> 5  
 <212> PRT

<213> Influenza virus

<400> 40

Asp Arg Leu Arg Arg

1                    5

<210> 41

<211> 7

<212> PRT

<213> Influenza virus

<400> 41

Pro Lys Gln Lys Lys Arg Lys

1                    5

<210> 42

<211> 10

<212> PRT

<213> Hepatitis virus

<400> 42

Arg Lys Leu Lys Lys Lys Ile Lys Lys Leu

1                    5                    10

<210> 43

<211> 10

<212> PRT

<213> Mus sp.

<400> 43

Arg Glu Lys Lys Lys Phe Leu Lys Arg Arg

1                    5                    10

<210> 44

<211> 20

<212> PRT

<213> Homo sapiens

<400> 44

Lys Arg Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala Lys Lys

1                    5                    10                    15

Lys Ser Lys Lys

20

<210> 45

<211> 17

<212> PRT

<213> Homo sapiens

<400> 45

Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala Arg Lys Thr Lys

1                    5                    10                    15

Lys

<210> 46

<211> 33

<212> DNA

<213> Homo sapiens

<400> 46

ggaagggcct gagtccgagc agaagaagaa ggg 33

<210> 47

<211> 33

<212> DNA

<213> Homo sapiens

<400> 47

cattggaggt gacatcgatg tctcctccat tgg 33

<210> 48

<211> 33

<212> DNA

<213> Homo sapiens

<400> 48

ggacatcgat gtcacctcca atgactaggg tgg 33

<210> 49

<211> 33

<212> DNA

<213> Homo sapiens

<400> 49

catcgatgtc ctccccattg gcctgcttcg tgg 33

<210> 50  
 <211> 33  
 <212> DNA  
 <213> Homo sapiens  
 <400> 50  
 ttcgtggcaa tgcgccaccg gttgatgtga tgg 33  
 <210> 51  
 <211> 33  
 <212> DNA  
 <213> Homo sapiens  
 <400> 51  
 tcgtggcaat gcgccaccg ttgatgtgat ggg 33  
 <210> 52  
 <211> 33  
 <212> DNA  
 <213> Homo sapiens  
 <400> 52  
 tccagcttct gccgtttgta ctttgcctc egg 33  
 <210> 53  
 <211> 33  
 <212> DNA  
 <213> Homo sapiens  
 <400> 53  
 ggaggaggagg gcacagatga gaaactcagg agg 33  
  
 <210> 54  
 <211> 33  
 <212> DNA  
 <213> Homo sapiens  
 <400> 54  
 aggggccgag attgggtgtt cagggcagag agg 33  
 <210> 55  
 <211> 33  
 <212> DNA  
 <213> Homo sapiens

<400> 55  
atgcaggagg gtggcgagag gggccgagat tgg 33  
<210> 56  
<211> 33  
<212> DNA  
<213> Homo sapiens  
<400> 56  
ggtggcgaga ggggccgaga ttgggtgttc agg 33  
<210> 57  
<211> 33  
<212> DNA  
<213> Mus musculus  
<400> 57  
caagcactga gtgccattag ctaaatgcat agg 33  
<210> 58  
<211> 33  
<212> DNA  
<213> Mus musculus  
<400> 58  
aatgcatagg gtaccaccca caggtgccag ggg 33  
<210> 59  
<211> 33  
<212> DNA  
<213> Mus musculus  
<400> 59  
acacacatgg gaaagcctct gggccaggaa agg 33  
<210> 60  
<211> 37  
<212> DNA  
<213> Homo sapiens  
<400> 60  
ggaggaggta gtatacagaa acacagagaa gtagaat 37  
<210> 61  
<211> 37

<212> DNA  
 <213> Homo sapiens  
 <400> 61  
 agaatgtaga ggagtcacag aaactcagca ctagaaa 37

<210> 62  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 primer"  
 <400> 62  
 aaaaccaccc ttctctctgg c 21

<210> 63  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 primer"  
 <400> 63  
 ggagattgga gacacggaga g 21

<210> 64  
 <211> 20  
  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 primer"  
 <400> 64  
 ctggaaagcc aatgcctgac 20  
 <210> 65  
 <211> 20

<212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
         primer"  
 <400> 65  
 ggcagcaaac tccttgctct 20  
 <210> 66  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
         primer"  
 <400> 66  
 gtgctttgca gaggcctacc 20  
 <210> 67  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
         primer"  
 <400> 67  
 cctggagcgc atgcagtagt 20  
 <210> 68  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
  
 <223> /note="Description of Artificial Sequence: Synthetic  
         primer"  
 <400> 68

accttctgtg tttccacat tc 22

<210> 69

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 69

ttggggagtg cacagacttc 20

<210> 70

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 70

ggctccctgg gttcaaagta 20

<210> 71

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 71

agaggggtct ggatgtcgta a 21

<210> 72

<211> 30

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

```

    probe"
<400
> 72
tagctctaaa acttcttctt ctgctcggac          30
<210> 73
<211> 30
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
    probe"
<400> 73
ctagccttat ttaacttgc tatgctgttt          30
<210> 74
<211> 335
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
    polynucleotide"
<400> 74
gagggcctat ttccatgat tccttcatat ttgcatatac gatacaaggc tgtagagag          60

ataattggaa ttaatttgc tgtaaacaca aagatattag tacaaaatac gtgacgtaga          120
aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaat ggactatcat          180
atgcttaccg taacttgaaa gtatttcgat ttcttggctt tataatctt gtggaagga          240
cgaaacaccg gaaccattca aaacagcata gcaagttaaa ataaggctag tccgttatca          300
acttgaaaaa gtggcaccga gtcggtgctt ttttt          335
<210> 75
<211> 423
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic

```

polynucleotide"  
 <400> 75  
 gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaagc tgtagagag 60  
 ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120  
 aagtaataat ttcttgggta gtttcagtt taaaattat gttttaaataat ggactatcat 180  
 atgcttaccg taacttgaat gtatttcgat ttcttggctt tatatatctt gtggaaagga 240  
 cgaacaccg gtagtattaa gtattgtttt atggctgata aatttctttg aatttctcct 300  
 tgattatttg ttataaaagt tataaaataa tcttgttggga accattcaaa acagcatagc 360  
 aagttaaaat aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cgggtccttt 420

ttt 423

<210> 76

<211> 339

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 76

gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaagc tgtagagag 60  
 ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120  
 aagtaataat ttcttgggta gtttcagtt taaaattat gttttaaataat ggactatcat 180  
 atgcttaccg taacttgaat gtatttcgat ttcttggctt tatatatctt gtggaaagga 240  
  
 cgaacaccg ggttttagag ctatgctgtt ttgaatggc ccaaacggg tcttcgagaa 300  
 gacgttttag agctatgctg ttttgaatgg tcccaaac 339

<210> 77

<211> 309

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 77

gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaagc tgtagagag 60

ataattgga ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120  
 aagtaataat ttcttgggta gtttgcagt ttaaaattat gttttaaata ggactatcat 180  
  
 atgcttaccg taacttgaag gtatttcgat ttcttggcctt tatatatctt gtggaaagga 240  
 cgaaacaccg ggtcttcgag aagacctgtt ttagagctag aaatagcaag ttaaaataag 300  
 gctagtccg 309  
 <210> 78  
 <211> 4269  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
           polynucleotide"  
 <400> 78  
 atggactata aggaccacga cggagactac aaggatcatg atattgatta caaagacgat 60  
 gacgataaga tggcccaaaa gaagaagcgg aaggtcggta tccacggagt cccagcagcc 120  
  
 gacaagaagt acagcatcgg cctggacatc ggcaccaact ctgtgggctg ggccgtgatc 180  
 accgacgagt acaaggtgcc cagcaagaaa ttcaaggtgc tgggcaaac cgaccggcac 240  
 agcatcaaga agaacctgat cggagccctg ctgttcgaca gcgcgaaac agccgaggcc 300  
 acccgctga agagaaccgc cagaagaaga tacaccagac ggaagaaccg gatctgctat 360  
 ctgcaagaga tcttcagcaa cgagatggcc aaggtggacg acagcttctt ccacagactg 420  
 gaagatcct tcctgggtga agaggataag aagcacgagc ggcaccccat cttcggcaac 480  
 atcgtggacg aggtggccta ccacgagaag taccaccca tctaccacct gagaaagaaa 540  
  
 ctggtggaca gcaccgacaa ggccgacctg cggctgatct atctggccct ggccacatg 600  
 atcaagtcc ggggccactt cctgatcgag ggcgacctga accccgacaa cagcgacgtg 660  
 gacaagctgt teatccagct ggtgcagacc tacaaccagc tgttcgagga aaaccctatc 720  
 aacgccagcg gcgtggacgc caaggccatc ctgtctgcca gactgagcaa gagcagacgg 780  
 ctggaatac tgatcgccca gctgcccggc gagaagaaga atggcctgtt cggaacctg 840  
 attgcctga gcctgggect gaccccaac ttcaagagca acttcgacct ggccgaggat 900  
 gccaaactgc agctgagcaa ggacacctac gacgacgacc tggacaacct gctggcccag 960  
  
 atcggcgacc agtacgccga cctgtttctg gccgcaaga acctgtccga cgccatcctg 1020  
 ctgagcgaca tctgagagt gaacaccgag atcaccaagg ccccctgag gcctctatg 1080

atcaagagat acgacgagca ccaccaggac ctgaccctgc tgaagctct cgtgcggcag 1140  
 cagctgcctg agaagtacaa agagattttc ttcgaccaga gcaagaacgg ctacgccggc 1200  
 tacattgacg gcggagccag ccaggaagag ttctacaagt tcatcaagcc catcctggaa 1260  
 aagatggacg gcaccgagga actgctcgtg aagctgaaca gagaggacct gctgcggaag 1320  
 cagcggacct tcgacaacgg cagcatcccc caccagatcc acctgggaga gctgcacgcc 1380  
  
 attctgcggc ggcaggaaga tttttacceca ttctgaagg acaaccggga aaagatcgag 1440  
 aagatcctga ccttccgcat ccctactac gtgggccctc tggccagggg aaacagcaga 1500  
 ttgcctgga tgaccagaaa gagcagggaa accatcacc cctggaactt cgaggaagtg 1560  
 gtggacaagg gcgcttccgc ccagagcttc atcgagcggg tgaccaactt cgataagaac 1620  
 ctgccaacg agaaggtgct gcccaagcac agcctgctgt acgagtactt caccgtgat 1680  
 aacgagctga ccaaagtgaa atacgtgacc gagggaatga gaaagcccgc cttcctgagc 1740  
 ggcgagcaga aaaaggccat cgtggacctg ctgttcaaga ccaaccggaa agtgaccctg 1800  
  
 aagcagctga aagaggacta cttcaagaaa atcgagtgtc tcgactccgt ggaaatctcc 1860  
 ggcgtggaag atcggttcaa gcctccctg ggcacatacc acgatctgct gaaaattatc 1920  
 aaggacaagg acttctgga caatgaggaa aacgaggaca ttctggaaga tatcgtgctg 1980  
 acctgacac tgtttgagga cagagagatg atcgaggaac ggctgaaaac ctatgcccac 2040  
 ctgttcgacg acaaagtgat gaagcagctg aagcggcggg gatacaccgg ctggggcagg 2100  
 ctgagccgga agctgatcaa cggcatccgg gacaagcagt cggcaagac aatcctggat 2160  
 ttctgaagt ccgacggctt cgccaacaga aacttcatgc agctgatcca cgacgacagc 2220  
  
 ctgacctta aagaggacat ccagaaagcc caggtgtccg gccagggcga tagcctgcac 2280  
 gagcacattg ccaatctggc gcgcagcccc gccattaaga agggcatcct gcagacagtg 2340  
 aaggtggtgg acgagctcgt gaaagtgatg ggccggcaca agcccagaaa catcgtgatc 2400  
 gaaatggcca gagagaacca gaccaccag aaggacaga agaacagccg cgagagaatg 2460  
 aagcggatcg aagagggat caaagagctg ggcagccaga tcctgaaaga acacccctg 2520  
 gaaaacacce agctgcagaa cgagaagctg tacctgtact acctgcagaa tgggcgggat 2580  
 atgtactgg accaggaact ggacatcaac cgctgtccg actacgatgt ggacatc 2640  
  
 gtgcctcaga gctttctgaa ggacgactcc atcgacaaca aggtgctgac cagaagcgac 2700  
 aagaaccggg gcaagagcga caactgccc tccgaagagg tcgtgaagaa gatgaagaac 2760  
 tactggcggc agctgctgaa cgccaagctg attaccaga gaaagtctga caatctgacc 2820  
 aaggccgaga gaggcggcct gagcgaactg gataaggccg gcttcatcaa gagacagctg 2880  
 gtggaaaccc ggcagatcac aaagcacgtg gcacagatcc tggactcccg gatgaacact 2940

aagtacgacg agaatgacaa gctgatccgg gaagtgaaag tgatcacct gaagtccaag 3000  
ctggtgtccg atttccggaa ggatttccag ttttacaag tgcgcgagat caacaactac 3060

caccacgccc acgacgccta cctgaacgcc gtcgtgggaa ccgcctgat caaaaagtac 3120  
cctaagctgg aaagcgagtt cgtgtacggc gactacaagg tgtacgacgt gcggaagatg 3180  
atcgccaaga gcgagcagga aatcggaag gctaccgcca agtacttctt ctacagcaac 3240  
atcatgaact ttttcaagac cgagattacc ctggccaacg gcgagatccg gaageggcct 3300  
ctgatcgaga caaacggcga aaccggggag atcgtgtggg ataagggccg ggattttgcc 3360  
accgtcgga aagtgtgag catgccccaa gtgaatatcg tgaaaaagac cgaggtgcag 3420  
acaggcggct tcagcaaaga gtctatcctg cccaagagga acagcgataa gctgatcgcc 3480

agaaagaagg actgggaccc taagaagtac ggcggcttcg acagcccccac cgtggcctat 3540  
tctgtgctgg tgggtggccaa agtggaaaag ggcaagtcca agaaactgaa gagtgtgaaa 3600  
gagctgctgg ggatcacctat catgaaaaga agcagcttcg agaagaatcc catcgacttt 3660  
ctggaagcca agggctacaa agaagtgaaa aaggacctga tcatcaagct gcctaagtac 3720  
tcctgtttcg agctggaaaa cggcccgaag agaatgctgg cctctgccgg cgaactgcag 3780  
aagggaaacg aactggcctt gccctccaaa tatgtgaact tcctgtacct ggccagccac 3840  
tatgagaagc tgaagggtc ccccaggat aatgagcaga aacagctgtt tgtggaacag 3900

cacaagcact acctggacga gatcatcgag cagatcagcg agttctccaa gagagtgatc 3960  
ctggccgacg ctaatctgga caaagtgtg tccgcctaca acaagcaccg ggataagccc 4020  
atcagagagc aggccgagaa tateatccac ctgtttaccc tgaccaatct gggagcccct 4080  
gccgccttca agtactttga caccaccatc gaccggaaga ggtacaccag caccaaagag 4140  
gtgctggacg ccacctgat ccaccagagc atcaccggcc tgtacgagac acggatcgac 4200  
ctgtctcagc tgggagcga caagcgtcct gctgctacta agaaagctgg tcaagetaag 4260  
aaaaagaaa 4269

<210> 79

<211> 1449

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 79

atgaagcagc tggaggagtt actttctacc tctttcgaca tccagtttaa tgacctgacc 60  
 ctgctggaaa ccgccttcac tcacacctcc tacgcgaatg agcaccgcct actgaatgtg 120  
 agccacaacg agcgcctgga gtttctgggg gatgctgtct tacagctgat catctctgaa 180  
 tatctgtttg ccaaataccc taagaaaacc gaaggggaca tgtcaaagct gcgctccatg 240  
 atagtcaggg aagagagcct ggcgggcttt agtcgttttt gctcattcga cgcttatatc 300

aagctgggaa aaggcgaaga gaagtccggc ggcaggaggc gcgatacaat tetgggcgat 360  
 ctctttgaag cgtttctggg cgcacttcta ctggacaaag ggatcgacgc agtccgccgc 420  
 tttctgaaac aagtgatgat ccttcaggtc gaaaaggga acttcgagag agtgaaggac 480  
 tataaaacat gtttgcagga atttctccag accaaggag atgtagcaat agattatcag 540  
 gtaataagtg agaaaggacc agctcacgcc aaacaattcg aagttagcat cgttgtaat 600  
 ggcgcagtgt tgtcgaaggc cttgggtaaa tcaaaaaaac tggccgagca ggacgctgct 660  
 aaaaacgccc tcgctcagct cagcagggia ggatccgtga gcaaggcgca ggaggataac 720

atggccatca tcaaggagtt catgcgcttc aaggtgcaca tggagggctc cgtgaacggc 780  
 cagcagttcg agatcgaggg cgagggcgag ggccgccctt acgagggcac ccagaccgcc 840  
 aagctgaagg tgaccaaggg tggccccctg cccttcgctt gggacatcct gtcccctcag 900  
 ttcatgtacg gctccaaggc ctactgaag caccccgcg acatccccga ctacttgaag 960  
 ctgtccttcc ccgaggctt caagtgggag cgcgtgatga acttcgagga cggcggcgtg 1020  
 gtgacctga cccaggactc ctccctgcag gacggcgagt tcatctaaa ggtgaagctg 1080  
 cgcgccacca acttcccctc cgacggcccc gtaatgcaga agaagaccat gggctgggag 1140

gcctcctccg agcggatgta ccccaggac ggcgccctga agggcgagat caagcagagg 1200  
 ctgaagctga aggacggcgg ccaactacgac gctgaggtca agaccaccta caaggccaag 1260  
 aagcccgtgc agctgcccgg cgcctacaac gtcaacatca agttggacat cacctcccac 1320  
 aacgaggact acaccatcgt ggaacagtac gaacgcgccg agggccgcca ctccaccggc 1380  
 ggcatggacg agctgtacaa gaagcgtcct gctgctacta agaaagctgg tcaagctaag 1440  
 aaaaagaaa 1449

<210> 80

<211> 4269

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 80

atggactata aggaccacga cggagactac aaggatcatg atattgatta caaagacgat	60
gacgataaga tggcccaaaa gaagaagcgg aaggtcggta tccacggagt cccagcagcc	120
gacaagaagt acagcatcgg cctggccatc ggcaccaact ctgtgggctg gccctgatc	180
accgacgagt acaaggtgcc cagcaagaaa ttcaaggtgc tgggcaaac cgaccggcac	240
agcatcaaga agaacctgat cggagccctg ctgttcgaca gcggcgaaac agccgaggcc	300
acccggctga agagaaccgc cagaagaaga tacaccagac ggaagaaccg gatctgctat	360
ctgcaagaga tcttcagcaa cgagatggcc aaggtggacg acagcttctt ccacagactg	420
gaagagtctt tcctggtgga agaggataag aagcacgagc ggcaccccat cttcggaac	480
atcgtggacg aggtggccta ccacgagaag taccaccca tctaccacct gagaaagaaa	540
ctggtggaca gcaccgaaa ggccgacctg cggtgatct atctggcct ggcccatg	600
atcaagtcc ggggccactt cctgatcgag ggcgacctga accccgaaa cagcgactg	660
gacaactgt tcatccagct ggtgcagacc tacaaccagc tgttcgagga aaacccatc	720
aacgccagcg gcgtggacgc caaggccatc ctgtctgcca gactgagcaa gagcagacgg	780
ctggaaaatc tgatgcccc gctgcccggc gagaagaaga atggcctgtt cggaacctg	840
attgccctga gectgggctt gaccccaac ttcaagagca acttcgacct ggccgaggat	900
gccaactgc agctgagcaa ggacacctac gacgacgacc tggacaacct gctggcccag	960
atcggcgacc agtacccga cctgtttctg gccgccaaga acctgtccga cgccatctg	1020
ctgagcgaca tctgagagt gaacaccgag atcaccaagg ccccctgag gcctctatg	1080
atcaagat acgacgagca ccaccaggac ctgacctgc tgaagctct cgtgcggcag	1140
cagctgcctg agaagtaca agagattttc ttcgaccaga gcaagaacgg ctacccggc	1200
tacattgacg gcggagccag ccaggaagag ttctacaagt tcatcaagcc catcctgga	1260
aagatggacg gcaccgagga actgctctg aagctgaaca gagaggacct gctgcggaag	1320
cagcggacct tegacaacgg cagcatcccc caccagatcc acctgggaga getgcacgcc	1380
attctgcggc ggcaggaaga tttttacca ttctgaagg acaaccggga aaagatcgag	1440
aagatctga ccttcgcgat ccctactac gtggccctc tggccagggg aaacagcaga	1500
ttcgctgga tgaccagaaa gagcgaggaa accatcccc cctggaactt cgaggaagtg	1560
gtggacaagg gcgcttcgc ccagagcttc atcgagcgga tgaccaactt cgataagaac	1620
ctgccaacg agaaggtgct gcccaagcac agcctgctgt acgagtactt caccgtgtat	1680
aacgagctga ccaaagtga atactgacc gagggaatga gaaagccgc cttcctgagc	1740

ggcgagcaga aaaaggccat cgtggacctg ctgttcaaga ccaaccggaa agtgaccgtg 1800  
 aagcagctga aagaggacta cttcaagaaa atcgagtgtc tcgactccgt ggaatctcc 1860  
 ggcggtggaag atcggttcaa cgctccctg ggcacatacc acgatctgct gaaaattatc 1920  
 aaggacaagg acttcttga caatgaggaa aacgaggaca ttctggaaga tatcgtgctg 1980  
 acctgacac tgtttgagga cagagagatg atcgaggaac ggctgaaaac ctatgcccac 2040  
  
 ctgttcgacg acaaagtgat gaagcagctg aagcggcgga gatacaccgg ctggggcagg 2100  
 ctgagccgga agctgatcaa cggcatccgg gacaagcagt ccggcaagac aatcctggat 2160  
 ttctgaagt ccgacggctt cgccaacaga aacttcatgc agctgatcca cgacgacagc 2220  
 ctgaccttta aagaggacat ccagaaagcc caggtgtccg gccagggcga tagcctgcac 2280  
 gagcacattg ccaatctggc cggcagcccc gccattaaga agggcatcct gcagacagtg 2340  
 aaggtggtgg acgagctcgt gaaagtgatg ggccggcaca agcccagaaa catcgtgatc 2400  
 gaaatggcca gagagaacca gaccaccag aaggacaga agaacagccg cgagagaatg 2460  
  
 aagcggatcg aagaggcat caaagagctg ggcagccaga tcctgaaaga acacccctg 2520  
 gaaaacacce agctgcagaa cgagaagctg tacctgtact acctgcagaa tgggggggat 2580  
 atgtactgg accaggaact ggacatcaac cggctgtccg actacgatgt ggacatatac 2640  
 gtgcctcaga gctttctgaa ggacgactcc atcgacaaca aggtgctgac cagaagcgac 2700  
 aagaaccggg gcaagagcga caacgtgcc tccgaagagg tcgtgaagaa gatgaagaac 2760  
 tactggcggc agctgctgaa cgccaagctg attaccaga gaaagttcga caatctgacc 2820  
 aaggccgaga gagggcgct gagcgaactg gataaggccg gcttcatcaa gagacagctg 2880  
  
 gtgaaacc ggcagatcac aaagcacgtg gcacagatcc tggactcccg gatgaacact 2940  
 aagtacgac agaatgaaa gctgatccgg gaagtgaag tgatcacct gaagtccaag 3000  
 ctggtgtccg atttccgaa ggatttccag ttttacaag tgcgcgagat caacaactac 3060  
 caccagccc acgacgcta cctgaacgcc gtctgtggaa ccgacctgat caaaaagtac 3120  
 cctaagctgg aaagcgagtt cgtgtacggc gactacaagg tgtacgacgt gcggaagatg 3180  
 atcgccaaga gcgagcagga aatcggaag gctaccgcca agtacttctt ctacagcaac 3240  
 atcatgaact tttcaagac cgagattacc ctggccaacg gcgagatccg gaagcgcct 3300  
  
 ctgatcgaga caaacggcga aaccggggag atcgtgtggg ataagggcgg ggattttgcc 3360  
 accgtcggga aagtgtgag catgcccaca gtgaatatcg tgaaaaagac cgaggtgcag 3420  
 acagcggct tcagcaaga gtctatcctg cccaagagga acagcgaata gctgatcgcc 3480  
 agaaagaagg actgggacc taagaagtac ggcggcttcg acagccccac cgtggcctat 3540  
 tctgtgctgg tgggtggcaa agtgaaaaag ggcaagtcca agaaactgaa gagtgtgaaa 3600

gagctgctgg ggatcaccat catggaaaga agcagcttcg agaagaatcc catcgacttt 3660  
ctggaagcca agggctacaa agaagtgaag aaggacctga tcatcaagct gcctaagtac 3720

tccctgttcg agctggaaaa cggccggaag agaatgetgg cctctgccgg cgaactgcag 3780  
aagggaacg aactggcctt gccctccaaa tatgtgaact tcctgtacct ggccagccac 3840  
tatgagaagc tgaagggtc ccccgaggat aatgagcaga aacagctgtt tgtggaacag 3900  
cacaagcact acctggacga gatcatcgag cagatcagcg agttctccaa gagagtgatc 3960  
ctggccgacg ctaatctgga caaagtgtg tccgcctaca acaagcaccg ggataagccc 4020  
atcagagagc aggccgagaa tatcatccac ctgtttacc tgaccaatct gggagcccct 4080  
gccgccttca agtactttga caccaccatc gaccggaaga ggtacaccag caccaaagag 4140

gtgctggacg ccacctgat ccaccagagc atcaccggcc tgtacgagac acggatcgac 4200  
ctgtctcage tgggaggcga caagcgtcct gctgtacta agaaagctgg tcaagctaag 4260  
aaaaagaaa 4269

<210> 81  
<211> 2012  
<212> DNA  
<213> Artificial Sequence  
<220><221> source  
<223> /note="Description of Artificial Sequence: Synthetic  
          polynucleotide"  
<400> 81

gaatgctgcc ctgacaccg cttctcctt gtcttctgt gtccaaggag aatgaggtct 60  
cactgggtgga tttcggacta cctgaggag ctggcacctg agggacaagg cccccacct 120

gccagctcc agcctctgat gaggggtggg agagagctac atgaggttgc taagaaagcc 180  
tccctgaag gagaccacac agtgtgtgag gttggagtct ctgacagcgg gttctgtgcc 240  
cccagggata gtctggctgt ccaggcactg ctcttgatat aaacaccacc tectagtatt 300  
gaaaccatgc ccattctgcc tctctgtatg gaaaagagca tggggctggc ccgtggggtg 360  
gtgtccactt taggccctgt gggagatcat gggaaccac gcagtgggtc ataggctctc 420  
tcatttacta ctcacatcca ctctgtgaag aagcgattat gatctctct ctagaaactc 480  
gtagagtccc atgtctgccg gcttcagag cctgcactcc tccaccttgg cttggctttg 540

ctggggctag aggagctagg atgcacagca gctctgtgac cctttgtttg agaggaacag 600  
gaaaaccacc cttctctctg gccactgtg tcctcttct gccctgcat cccttctgt 660

gaatgttaga cccatgggag cagctggtca gaggggaccc cggcctgggg ccctaacc 720  
 tatgtagcct cagtcttccc atcaggctct cagctcagcc tgagtgttga ggccccagtg 780  
 gctgctctgg gggcctcctg agtttctcat ctgtgccctt ccctccctgg cccagtgaa 840  
 ggtgtggttc cagaaccgga ggacaaagta caaacggcag aagctggagg aggaagggcc 900  
 tgagtccgag cagaagaaga agggctcca tcacatcaac cggtggcgca ttgccacgaa 960

gcaggccaat ggggaggaca tcgatgtcac ctccaatgac aagcttgcta gcggtgggca 1020  
 accacaaacc cacgagggca gagtgtgctt tgctgctggc caggcccctg cgtgggcca 1080  
 agctggactc tggcactcc ctggccaggc tttggggagg cctggagtca tggccccaca 1140  
 gggcttgaag cccggggccg ccattgacag agggacaagc aatgggctgg ctgaggcctg 1200  
 ggaccacttg gccttctct cggagagcct gcctgcctgg gcgggcccgc ccgccaccgc 1260  
 agcctcccag ctgctctccg tgtctccaat ctcccttttg ttttgatgca tttctgtttt 1320  
 aatttatttt ccaggcacca ctgtagttta gtgatcccca gtgtcccctt tcctatggg 1380

aataataaaa gtctctctct taatgacacg ggcatccagc tccagcccca gacctgggg 1440  
 tggtagattc cgctctgag gggcagtggg ggttggtaga gcaaacgcgt tcagggcctg 1500  
 ggagcctggg gtggggtact ggtggagggg gtcaagggtt attcattaac tcctctcttt 1560  
 tgttggggga cctggtctc tacctccagc tccacagcag gagaacagg ctagacatag 1620  
 ggaaggcca tctgtatct tgaggaggga caggcccagg tcttcttaa cgtattgaga 1680  
 ggtgggaatc agggccaggt agttcaatgg gagagggaga gtgcttcctt ctgcctagag 1740  
 actctggtgg ctctccagt tgaggagaaa ccagaggaaa ggggaggatt ggggtctggg 1800

ggaggaaca ccattcaca aggctgacgg ttccagtcg aagtcgtggg cccaccagga 1860  
 tgctcacctg tcttggaga accgctgggc aggttgagac tgacagaca gggcttaagg 1920  
 ctgagcctgc aaccagtccc cagtgactca gggcctctc agccaagaa agagcaact 1980  
 gccaggccc gctgagctct tgtgttcacc tg 2012

<210> 82

<211> 3462

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 82

atgaaaaggc cggcggccac gaaaaaggcc ggccaggcaa aaaagaaaa gtccgacctg 60

gtacttggac tggatattgg tateggttcg gtgggagtcg gaatcctcaa caaggtcacg 120

ggggagatca ttcacaagaa ctccgggatc ttccccgcag ctccagctga gaacaacttg 180

gtcgggagaa cgaataggca gggcaggcga ctggcgagga ggaagaaaca caggagagtc 240

cgattgaacc ggctgttcga ggagtcgggt ttgatcaccg actttacgaa aatctcgatt 300

aaccttaate cctatcagct tcgggtgaaa ggcctgacag acgaacttc gaatgaggaa 360

cttttcatcg cgctgaaaaa catggtcaag cacagagga tttctacct cgatgacgcc 420

tcggatgacg gaaattcctc agtaggagat tatgcacaga tcgtgaaaga gaactcaaag 480

caactggaac caaagacacc ggggcagatc caacttgaac gataccagac ataccgacag 540

ctcagaggag attttacggt ggagaaggac ggtaaaaagc acagactcat taacgtattt 600

cccacgtcgg cgtacagatc cgaagcgtc cgcaccttc agactcaaca ggagttcaac 660

ccgcaaatca ctgatgagtt catcaaccgc ttttggaaa tcttgaccgg aaagcgcaag 720

tattatcatg ggccgggtaa tgagaaatcc agaacagatt acggccgata cagaacttcg 780

ggggaaacct tggataacat ctttggattt ttgattggaa agtgacactt ttaccggac 840

gagtttcgag cggccaaggc gtcatacaca gcacaagagt ttaatctctt gaatgatttg 900

aacaacttga cggccccac ggagacaaag aagctctcca aagagcaaaa gaaccaaatc 960

atcaactacg tcaagaacga gaaggctatg gggccagcga agctgttcaa gtatatcgct 1020

aaacttctca gctgtgatg ggcggacatc aaagggtacc gaatcgacaa gtcgggaaaa 1080

gcggaaatc acacgtttga agcatatcga aagatgaaaa cgttggaaac actggacatt 1140

gagcagatgg accgggaaac gctcgacaaa ctggcatacg tgctcacgtt gaatactgaa 1200

cgagagggaa tccaagaggc ccttgaacat gatttcgccg atggatcgtt cagccagaag 1260

caggtcgacg aacttgtgca attccgcaag gcgaatagct ccatcttcgg gaagggatgg 1320

cacaactttt cggtcaaac catgatggag ttgatcccag aactttatga gacttcggag 1380

gagcaaatga cgatcttgac gcgcttgggg aaacagaaaa cgacaagctc atcgaacaaa 1440

actaagtaca ttgatgagaa attgctgacg gaagaaatct ataatccggt agtagcgaac 1500

tcggtgaagc aagcgatcaa aatcgtgaac gcggcgatca aggaatatgg tgactttgat 1560

aacatcgtaa ttgaaatggc tagagagacg aacgaagatg acgagaaaaa ggcaatccag 1620

aagatccaga aggccaacaa ggatgaaaaa gatgcagcga tgcttaaacg ggccaaccaa 1680

tacaatggaag aggcggagct gcccattca gtgtttcacg gtcataaaca gttggcgacc 1740

aagatccgac tctggcatca gcagggtgag cgggtgtctct acaccgaaa gactatctcc 1800  
atccatgact tgattaacaa ttcgaaccag tttgaagtgg atcatattct gccctgtca 1860  
atcacctttg acgactcgct tgcgaacaag gtgctcgtgt acgcaacggc aaatcaggag 1920  
aaaggccagc ggactccgta tcaggcgctc gactcaatgg acgatgcgtg gtcattccgg 1980  
gagctgaagg cgttcgtacg cgagagcaag aactgagca acaaaaagaa agagtatctg 2040  
ctgacagagg aggacatctc gaaattcgat gtcaggaaga agttcatcga gcggaatctt 2100  
gtcgacactc gctacgcttc cagagtagta ctgaacgcgc tccaggaaca ctttagagcg 2160  
  
cacaaaattg acacgaaggt gtcagtgggtg agagggcagt tcacatcca actccgccga 2220  
cattggggca tcgaaaagac gcgggacaca tatcaccatc atgcggtgga cgcgctgatt 2280  
attgccgctt cgtcccagtt gaatctctgg aaaaagcaga agaacacgct ggtgtcgtat 2340  
tcggaggatc agcttttggg catcgaacc ggggagctga tttccgacga tgaatacaaa 2400  
gaatcgggtg ttaaggcacc atatcagcat ttcgtggaca cgctgaagag caaagagttt 2460  
gaggacagca tcctcttttc gtaccaagtg gactcgaagt ttaatcgcaa gatttcagac 2520  
gccacaatct acgcgacgag gcaggcgaag gtgggcaaag ataaagcaga tgaaacctac 2580  
  
gtccttggta aaatcaagga catctacact caggacgggt acgatgcgtt catgaaaatc 2640  
tacaagaagg ataagtcgaa gtttctcatg taccgccacg atccacagac tttcgaaaaa 2700  
gtcattgagc ctattttggg gaactacct aacaagcaaa tcaacgagaa agggaaagaa 2760  
gtcccgtgca acccctttct gaagtacaag gaagagcacg gttatatccg caaatactcg 2820  
aagaaaggaa atgggcctga gattaagtgc cttaagtatt acgactcaa gttaggtaac 2880  
cacatcgaca ttaccccgaa agactccaac acaaaagtgc tgttcgagtc cgtctcgccc 2940  
tggcgagcag atgtgtattt taataagacg accggcaat atgagatcct tggactcaa 3000  
  
tacgcagacc ttcaattcga aaaggggacg ggcacttata agatttcaca agagaagtac 3060  
aacgacatca agaaaaagga aggggtcgat tcagattcgg agttcaaatt caccctctac 3120  
aaaaacgacc tcctgcttgt gaaggacaca gaaacgaagg agcagcagct ctttcggttc 3180  
ctctcagca cgatgccaa acaaaaacat tacgtcgaac ttaaacctta cgataagcaa 3240  
aagtttgaag ggggagaggc actgatcaa gtattgggta acgtagcaa tagcggacag 3300  
tgtaagaaag ggctgggaaa gtccaatc tcgatctata aagtacgaac agatgtattg 3360  
ggaaaccagc atatcatcaa aaatgagggg gataaaccca aactcgattt caagctcct 3420  
  
gctgctacta agaaagctgg tcaagctaag aaaaagaat aa 3462

<210> 83

<211> 340

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polynucleotide"

<400> 83

gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgtagagag 60  
ataattggaa ttaatttgac tgtaaacaca aagatattag taaaaatac gtgacgtaga 120  
aagtaataat ttcttgggta gtttgcagtt taaaattat gttttaaataat ggactatcat 180  
atgcttaccg taacttgaaa gtatttcgat ttcttggcctt tatatatctt gtggaagga 240

cgaaacaccg ttacttaaat cttgcagaag ctacaaagat aaggcttcat gccgaaatca 300  
acacctgtc attttatggc aggggtgtttt cgttatntaa 340

<210> 84

<211> 2601

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polynucleotide"

<400> 84

atggactata aggaccacga cggagactac aaggatcatg atattgatta caaagacgat 60  
gacgataaga tggcccaaaa gaagaagcgg aaggtcggta tccacggagt cccagcagcc 120  
gtagatttga gaactttggg atattcacag cagcagcagg aaaagatcaa gcccaaagtg 180

aggtcgacag tcgcgcagca tcacgaagcg ctggtgggtc atgggtttac acatgcccac 240

atcgtagcct tgtcgcagca ccctgcagcc cttggcacgg tcgccgtcaa gtaccaggac 300

atgattgcgg cgttgccgga agccacacat gaggcgatcg tcggtgtggg gaaacagtgg 360

agcggagccc gagcgttga ggccctgttg acggtcgcgg gagagctgag agggcctccc 420

cttcagctgg acacgggcca gttgctgaag atcgcaagc ggggaggagt cacggcggtc 480

gaggcgggtgc acgctggcg caatgcgctc acgggagcac ccctcaacct gacccagag 540

caggtcgtgg caattgcgag caaccacggg ggaaagcagg cactcgaaac cgtccagagg 600

ttgctgcctg tgctgtgcca agcgcacgga cttacgccag agcaggtcgt ggcaattgcg 660

agcaaccacg ggggaaagca ggcactcga accgtccaga ggttgctgcc tgtgctgtgc 720

caagcgacg gactaacccc agagcaggtc gtggcaattg cgagcaacat cgggggaaag 780  
 caggcactcg aaaccgtcca gaggttgctg cctgtgctgt gccaaagcga cgggttgacc 840  
 ccagagcagg tcgtggcaat tgcgagcaac cacgggggaa agcaggcact cgaaaccgtc 900  
 cagaggttgc tgctgtgct gtgccaagcg cacggcctga cccagagca ggtcgtggca 960  
 attgcgagca accacggggg aaagcaggca ctcgaaaccg tccagaggtt gctgcctgtg 1020  
  
 ctgtgccaag cgcacggact gacaccagag caggctgtgg caattgcgag caacatcggg 1080  
 ggaaagcagg cactcgaaac cgtccagagg ttgctgctg tgctgtgcca agcgcacgga 1140  
 cttacacccc aacaagtctg ggcaattgctg agcaaccacg ggggaaagca ggcaactcga 1200  
 accgtccaga ggttgctgcc tgtgctgtgc caagcgacg gacttacgcc agagcaggtc 1260  
 gtggcaattg cgagcaacca cgggggaaag caggcactcg aaaccgtcca gaggttgctg 1320  
 cctgtgctgt gccaaagcga cggactaacc ccagagcagg tcgtggcaat tgcgagcaac 1380  
 atcgggggaa agcaggcact cgaaaccgtc cagaggttgc tgctgtgct gtgccaagcg 1440  
  
 cacgggttga cccagagca ggtcgtggca attgcgagca acatcggggg aaagcaggca 1500  
 ctgaaaccg tccagaggtt gctgctgtg ctgtgccaag cgcacggcct gaccccagag 1560  
 caggctgtgg caattgcgag caaccacggg ggaaagcagg cactcgaaac cgtccagagg 1620  
 ttgctgctg tgctgtgcca agcgcacgga ctgacaccag agcaggctgt ggcaattgctg 1680  
 agcaaccacg ggggaaagca ggcaactcga accgtccaga ggttgctgcc tgtgctgtgc 1740  
 caagcgacg gactcacgcc tgagcaggta gtggctattg catccaaca cgggggcaga 1800  
 cccgcaactg agtcaatctg ggcccagctt tcgagccggg accccgcgt ggccgcaactc 1860  
  
 actaatgac atcttgiagc gctggcctgc ctggcggac gaccgcctt ggatgcggtg 1920  
 aagaaggggc tcccgcacg gcctgcattg attaagcggg ccaacagaag gattcccag 1980  
 aggacatcac atcgagtggc aggttcccaa ctgctgaaga gtgaacttga ggagaaaaag 2040  
 tcggagctgc ggcacaaatt gaaatcgtc ccgatgaat acatcgaact tatcgaatt 2100  
 gctaggaact cgactcaaga cagaatcctt gagatgaagg taatggagt ctttatgaag 2160  
 gtttatggat accgagggaa gcatctcggg ggatcacgaa aaccgcaggg agcaatctat 2220  
 acggtgggga gcccgattga ttacggagt atcgtcgaca cgaaagccta cagcgtggg 2280  
  
 tacaatctc ccatcgggca ggcagatgag atgcaacgtt atgtcgaaga aatcagacc 2340  
 aggaacaaac acatcaatcc aatgagtgg tggaaagtgt atccttcac agtgaccgag 2400  
 ttttaagttt tgttgtctc tgggcattc aaaggcaact ataaggcca gctcacacgg 2460  
 ttgaatcaca ttacgaactg caatggtgct gttttgtccg tagaggaact gctcattggt 2520  
 ggagaaatga tcaaagcggg aactctgaca ctggaagaag tcagacgcaa gtttaacaat 2580

ggcgagatca atttccgctc a 2601

<210> 85

<211> 2600

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polynucleotide"

<400> 85

atggactata aggaccacga cggagactac aaggatcatg atattgatta caaagacgat 60

gacgataaga tggcccaaaa gaagaagcgg aaggtcggta tccacggagt cccagcagcc 120

gtagatttga gaactttggg atattcacag cagcagcagg aaaagatcaa gcccaaagtg 180

aggtcgacag tcgcgacga tcacgaagcg ctggtgggtc atgggtttac acatgcccac 240

atcgtagcct tgtcgacga ccctgcagcc cttggcacgg tcgccgtcaa gtaccaggac 300

atgattgcgg cgttgccgga agccacacat gaggcgatcg tcggtgtggg gaaacagtgg 360

agcggagccc gagcgcttga ggcctgttg acggtcggg gagagctgag agggcctccc 420

cttcagctgg acacgggcca gttgctgaag atcggaagc ggggaggagt cacggcggtc 480

gaggcggcgc acgcgtggcg caatgcgctc acgggagcac ccctcaacct gaccccagag 540

caggtcgtgg caattgcgag caaccacggg ggaaagcagg cactcgaaac cgtccagagg 600

ttgctgcctg tgctgtgcca agcgcacgga cttacgccag agcaggtcgt ggcaattgctg 660

agcaaccacg ggggaaagca ggcactcga accgtccaga ggttgctgcc tgtgctgtgc 720

caagcgacg gactaacccc agagcaggtc gtggcaattg cgagcaacca cgggggaaag 780

caggcactcg aaaccgtcca gaggttgctg cctgtgctgt gccaagcgca cgggttgacc 840

ccagagcagg tcgtggcaat tgcgagcaac atcgggggaa agcaggcact cgaaacctgc 900

cagaggttgc tgctgtgct gtgccaagcg cacggcctga ccccagagca ggtcgtggca 960

attgcgagca accacggggg aaagcaggca ctcgaaaccg tccagaggtt gctgcctgtg 1020

ctgtgccaag cgcacggact gacaccagag caggtcgtgg caattgcgag ccatgacggg 1080

ggaaagcagg cactcgaaac cgtccagagg ttgctgctg tgctgtgcca agcgcacgga 1140

cttacaccg aacaagtctg ggcaattgctg agccatgacg ggggaaagca ggcactcga 1200

accgtccaga ggttgctgcc tgtgctgtgc caagcgacg gacttacgcc agagcaggtc 1260

gtggcaattg cgagccatga cgggggaaag caggcactcg aaaccgtcca gaggttgctg 1320

cctgtgctgt gccaaagcga cggactaacc ccagagcagg tcgtggcaat tgcgagcaac 1380  
 ggagggggaa agcaggcact cgaaaccgtc cagaggttgc tgcctgtgct gtgccaagcg 1440  
 cacgggttga ccccagagca ggtcgtggca attgcgagca acggaggggg aaagcaggca 1500  
 ctcgaaaccg tccagaggtt gctgcctgtg ctgtgccaag cgcacggcct gaccccagag 1560  
 caggtcgtgg caattgcgag ccatgacggg ggaaagcagg cactcgaaac cgtccagagg 1620

ttgctgcctg tgetgtgcca agcgcacgga ctgacaccag agcaggtcgt ggcaattgcg 1680  
 agcaacggag ggggaaagca ggcaactgaa accgtccaga ggttgctgcc tgtgctgtgc 1740  
 caagcgcacg gactcacgcc tgagcaggta gtggctattg catccaacgg agggggcaga 1800  
 cccgcactgg agtcaatcgt ggcccagctt tcgaggccgg accccgcgct ggcccactc 1860  
 actaatgac atctttagc gctggcctgc ctggcggac gacccgcctt ggatgcggtg 1920  
 aagaaggggc tcccgcacg ccctgcattg attaagcgga ccaacagaag gattcccag 1980  
 aggaatcaca tcgagtggca ggttccaac tcgtgaagag tgaacttgag gagaaaaagt 2040

cggagctgcg gcacaaattg aaatacgtac cgcataaata catcgaactt atcgaaattg 2100  
 ctaggaacte gactcaagac agaatccttg agatgaaggt aatggagttc tttatgaagg 2160  
 tttatggata ccgaggggaa catctcgggtg gatcacgaaa acccgacgga gcaatctata 2220  
 cgggtggggag cccgattgat tacggagtga tcgtcgacac gaaagcctac agcggtggtt 2280  
 acaatcttc catcgggcag gcagatgaga tgcaacgtta tgtcgaagaa aatcagacca 2340  
 ggaacaaaca catcaatcca aatgagtggg ggaaagtgt tccctcatca gtgaccgagt 2400  
 ttaagtttt gtttgtctct gggcatttca aaggcaacta taaggcccag ctcacacggt 2460

tgaatcacat tacgaactgc aatggtgctg ttttgtccgt agaggaactg ctcatgtgtg 2520  
 gagaaatgat caaagcggga actctgacac tggagaagt cagacgcaag tttacaatg 2580  
 gcgagatcaa tttccgctca 2600

<210> 86  
 <211> 141  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polynucleotide"  
 <400> 86  
 cctgcaggca gctgctgctc cgctcgctca ctgaggccgc ccgggcaaag cccgggcgctc 60

ggcgacctt tggtcgccc gcctcagtga gcgagcgagc ggcagagag ggagtggcca 120

actccatcac taggggttcc t 141

<210> 87

<211> 141

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polynucleotide"

<400> 87

aggaaccct agtgatggag ttggcactc cctctctgcg cgctcgctcg ctcaactgagg 60

ccggcgacc aaaggtgcc cgacggccgg gctttgccc ggcgccctca gtgagcgagc 120

gagcgcgag ctgcctgcag g 141

<210> 88

<211> 485

<212

> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polynucleotide"

<400> 88

gtgtctagac tgcagagggc cctgcgtatg agtgaagtg ggttttagga ccaggatgag 60

gcggggtggg ggtgcctacc tgacgaccga ccccgaccca ctggacaagc acccaacccc 120

cattcccaa attgcgcatc ccctatcaga gaggggagg gaaacagga tgcggcgagg 180

cgcgtgcga ctgccagctt cagcaccgag gacagtgcct tcgccccgc ctggcgggcg 240

gcgccaccgc cgctcagca ctgaaggcgc gctgacgtca ctcgccggtc ccccgcaaac 300

tccccttccc ggccaccttg gtcggtccg cgccgcccgg ggcccagccg gaccgacca 360

cgcgaggcgc gagatagggg ggcacgggag cgaccatctg cgctcgggcg ccggcgactc 420

agcgtgcct cagtctcggg tgggcagcgg aggagtctg tcgtgcctga gagcgagtc 480

gagaa 485

<210> 89

<211> 408

<212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"  
 <400> 89  
 gtagatttga gaactttggg atattcacag cagcagcagg aaaagatcaa gcccaaagtg 60  
  
 aggtcgacag tcgcgcagca tcacgaagcg ctggtgggtc atgggtttac acatgcccac 120  
 atcgtagcct tgcgcagca ccctgcagcc cttggcacgg tcgccgtcaa gtaccaggac 180  
 atgattgcgg cgttgccgga agccacacat gaggcgatcg tcggtgtggg gaaacagtgg 240  
 agcggagccc gagcgttga ggccctgttg acggtcgcgg gagagctgag agggcctccc 300  
 cttcagctgg acacgggcca gttgctgaag atcgcgaaac ggggaggagt cacggcggtc 360  
 gaggcgggtc acgctggcg caatgcgctc acgggagcac ccctcaac 408  
 <210> 90  
 <211> 164  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"  
 <400> 90  
 cggaccccgc gctggccgca ctcaactatg atcatcttgt agcgctggcc tgcctcggcg 60  
 gacgaccccgc cttggatcgc gtgaagaagg ggctcccgca cgcgctgca ttgattaagc 120  
 ggaccaacag aaggattccc gagaggacat cacatcgagt ggca 164  
 <210> 91  
 <211> 858  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"  
 <400> 91

atgagtattc aacatttccg tgtcgcctt attccctttt ttgcggcatt ttgccttcct 60  
 gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca 120  
 cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc 180  
 gaagaacgtt ttccaatgat gagcactttt aaagtctgc tatgtggcgc ggtattatcc 240  
 cgtattgacg ccgggcaaga gcaactcggc cgcgcatac actattctca gaatgacttg 300  
 gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta 360  
 tgcagtgtcg ccataacat gaggataac actgcggcca acttacttct gacaacgatc 420

ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt 480  
 gatcgttggg aaccggagct gaatgaagcc ataccaaagc acgagcgtga caccacgatg 540  
 cctgtagcaa tggcaacaac gttgcgcaa ctattaactg gcgaactact tactctagct 600  
 tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc acttctgcgc 660  
 tcggccttc cggtctgctg gtttattgct gataaatctg gagccggtga gcgtgggtct 720  
 gcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac 780  
 acgacgggga gtcaggcaac tatggatgaa cgaatatagac agatcgctga gataggtgcc 840

tcaactgatta agcattgg 858

<210> 92

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"

<400> 92

gcgcgtcca cacaactcac 20

<210> 93

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"

<400> 93

gcaaaaatag acaatcagca 20

<210> 94  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <400> 94  
 gaaggatctc ggccaatttg 20

<210> 95  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <400> 95  
 gctgccgggt ttgcatgaa 20

<  
 <210> 96  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <400> 96  
 ccgggccgc agcaaacttc 20

<210> 97  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
oligonucleotide"  
<400> 97  
ggggctgtca gggaataaat 20  
<210>  
> 98  
<211> 20  
<212> DNA  
<213> Artificial Sequence  
<220><221> source  
<223> /note="Description of Artificial Sequence: Synthetic  
primer"  
<400> 98  
ctgtgctgaa ggatctgggg 20  
<210> 99  
<211> 20  
<212> DNA  
<213> Artificial Sequence  
<220><221> source  
<223> /note="Description of Artificial Sequence: Synthetic  
primer"  
<400> 99  
atgctgcagg cataggacaa 20  
<210> 100  
<211> 21  
<212> DNA  
<213> Artificial Sequence  
<220><221> source  
<223> /note="Description of Artificial Sequence: Synthetic  
primer"  
<400> 100  
gagggggaga gggactaaag a 21  
<210> 101  
<211> 19

<212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic primer"  
 <400> 101  
 gctctccctc cccagctta 19  
 <210> 102  
 <211> 14  
 <212> DNA  
 <213>  
 > Mus sp.  
 <400> 102  
 tatctgaact ctcc 14  
 <210> 103  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 103  
 tgtctgtctt gcat 14  
 <210> 104  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 104  
 tggctgttgc cagg 14  
 <210> 105  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 105  
 tacctggagc tagc 14  
 <210> 106  
 <211> 14

<212> DNA  
 <213> Mus sp.  
 <400> 106  
 tacacaggat gtcc 14  
 <210> 107  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 107  
 ttggccctgt gcag 14  
 <210> 108  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 108  
 tagcgcagcg atcg 14  
 <210> 109  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 109  
 tattgccaag agag 14  
  
 <210> 110  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 110  
 tgactggaac atac 14  
 <210> 111  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 111  
 tgatggactt gtat 14

<210> 112  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 112  
 tggacettct ccct 14  
 <210> 113  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 113  
 taggtctcct ggag 14  
  
 <210> 114  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 114  
 tggctcagga actt 14  
 <210> 115  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 115  
 ttctctaagc tgcc 14  
 <210> 116  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 116  
 tgagccctgg agga 14  
 <210> 117  
 <211> 14  
 <212> DNA  
 <213> Mus sp.

<400> 117	
tcgctaagat ggag	14
<210> 118	
<211> 14	
<212> DNA	
<213> Mus sp.	
<400> 118	
tgtagtgagt gttc	14
<210> 119	
<211> 14	
<212> DNA	
<213> Mus sp.	
<400> 119	
tgtctaggtg atgt	14
<210> 120	
<211> 14	
<212> DNA	
<213> Mus sp.	
<400> 120	
tctctgctag aagg	14
<210> 121	
<211> 14	
<212> DNA	
<213> Mus sp.	
<400> 121	
tctaggtcaa gtgt	14
<210> 122	
<211> 14	
<212> DNA	
<213> Mus sp.	
<400> 122	
tcctctgctc cgct	14
<210> 123	

<211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 123  
 tctaggagtg tagc 14  
 <210> 124  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 124  
 tgcttgctg ctgg 14  
 <210> 125  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 125  
 tatctgaact ctcc 14  
  
 <210> 126  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 126  
 tcagagctgt cctc 14  
 <210> 127  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 127  
 tgcaagagta ggag 14  
 <210> 128  
 <211> 14  
 <212> DNA  
 <213> Mus sp.  
 <400> 128

tagtgactga ttcc 14

<210> 129

<211> 14

<212> DNA

<213> Mus sp.

<400> 129

ttggaggagc acca 14

<210> 130

<211> 16

<212> DNA

<213> Mus sp.

<400> 130

tgaatgatga taatac 16

<210> 131

<211> 4

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 131

Gly Gly Ser Gly

1

<210> 132

<211> 1603

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 132

Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala Phe

1                    5                    10                    15

Ser Ala Asp Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser Leu  
 20 25 30  
 Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His His  
 35 40 45  
 Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu Arg  
 50 55 60  
  
 Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala Ala  
 65 70 75 80  
 Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln Pro  
 85 90 95  
 Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr  
 100 105 110  
 Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val  
 115 120 125  
  
 Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His  
 130 135 140  
 Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val  
 145 150 155 160  
 Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala  
 165 170 175  
 Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala  
 180 185 190  
  
 Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp  
 195 200 205  
 Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val  
 210 215 220  
 Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn  
 225 230 235 240  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys  
 245 250 255  
  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala

260	265	270	
His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly			
275	280	285	
Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys			
290	295	300	
Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn			
305	310	315	320
Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val			
325	330	335	
Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala			
340	345	350	
Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu			
355	360	365	
Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala			
370	375	380	
Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg			
385	390	395	400
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val			
405	410	415	
Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val			
420	425	430	
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu			
435	440	445	
Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu			
450	455	460	
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr			
465	470	475	480
Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala			
485	490	495	
Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly			
500	505	510	

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 515 520 525  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 530 535 540  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
 545 550 555 560  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 565 570 575  
  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 580 585 590  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 595 600 605  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 610 615 620  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 625 630 635 640  
  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 645 650 655  
 Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 660 665 670  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 675 680 685  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 690 695 700  
  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 705 710 715 720  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 725 730 735  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 740 745 750  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala

755	760	765	
Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly			
770	775	780	
Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Arg			
785	790	795	800
Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala			
	805	810	815
Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly			
	820	825	830
Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro			
835	840	845	
Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser His			
850	855	860	
Arg Val Ala Asp His Ala Gln Val Val Arg Val Leu Gly Phe Phe Gln			
865	870	875	880
Cys His Ser His Pro Ala Gln Ala Phe Asp Asp Ala Met Thr Gln Phe			
	885	890	895
Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe Arg Arg Val Gly Val			
900	905	910	
Thr Glu Leu Glu Ala Arg Ser Gly Thr Leu Pro Pro Ala Ser Gln Arg			
915	920	925	
Trp Asp Arg Ile Leu Gln Ala Ser Gly Met Lys Arg Ala Lys Pro Ser			
930	935	940	
Pro Thr Ser Thr Gln Thr Pro Asp Gln Ala Ser Leu His Ala Phe Ala			
945	950	955	960
Asp Ser Leu Glu Arg Asp Leu Asp Ala Pro Ser Pro Met His Glu Gly			
	965	970	975
Asp Gln Thr Arg Ala Ser Ala Ser Pro Lys Lys Lys Arg Lys Val Glu			
	980	985	990
Ala Ser Lys Met Asp Lys Lys Thr Ile Val Trp Phe Arg Arg Asp Leu			
995	1000	1005	

Arg Ile Glu Asp Asn Pro Ala Leu Ala Ala Ala Ala His Glu Gly  
 1010 1015 1020

Ser Val Phe Pro Val Phe Ile Trp Cys Pro Glu Glu Glu Gly Gln  
 1025 1030 1035

Phe Tyr Pro Gly Arg Ala Ser Arg Trp Trp Met Lys Gln Ser Leu  
 1040 1045 1050

Ala His Leu Ser Gln Ser Leu Lys Ala Leu Gly Ser Asp Leu Thr  
 1055 1060 1065

Leu Ile Lys Thr His Asn Thr Ile Ser Ala Ile Leu Asp Cys Ile  
 1070 1075 1080

Arg Val Thr Gly Ala Thr Lys Val Val Phe Asn His Leu Tyr Asp  
 1085 1090 1095

Pro Val Ser Leu Val Arg Asp His Thr Val Lys Glu Lys Leu Val  
 1100 1105 1110

Glu Arg Gly Ile Ser Val Gln Ser Tyr Asn Gly Asp Leu Leu Tyr  
 1115 1120 1125

Glu Pro Trp Glu Ile Tyr Cys Glu Lys Gly Lys Pro Phe Thr Ser  
 1130 1135 1140

Phe Asn Ser Tyr Trp Lys Lys Cys Leu Asp Met Ser Ile Glu Ser  
 1145 1150 1155

Val Met Leu Pro Pro Pro Trp Arg Leu Met Pro Ile Thr Ala Ala  
 1160 1165 1170

Ala Glu Ala Ile Trp Ala Cys Ser Ile Glu Glu Leu Gly Leu Glu  
 1175 1180 1185

Asn Glu Ala Glu Lys Pro Ser Asn Ala Leu Leu Thr Arg Ala Trp  
 1190 1195 1200

Ser Pro Gly Trp Ser Asn Ala Asp Lys Leu Leu Asn Glu Phe Ile  
 1205 1210 1215

Glu Lys Gln Leu Ile Asp Tyr Ala Lys Asn Ser Lys Lys Val Val  
 1220 1225 1230

Gly Asn Ser Thr Ser Leu Leu Ser Pro Tyr Leu His Phe Gly Glu

1235	1240	1245
Ile Ser Val Arg His Val Phe	Gln Cys Ala Arg Met	Lys Gln Ile
1250	1255	1260
Ile Trp Ala Arg Asp Lys Asn	Ser Glu Gly Glu Glu	Ser Ala Asp
1265	1270	1275
Leu Phe Leu Arg Gly Ile Gly	Leu Arg Glu Tyr Ser	Arg Tyr Ile
1280	1285	1290
Cys Phe Asn Phe Pro Phe Thr	His Glu Gln Ser Leu	Leu Ser His
1295	1300	1305
Leu Arg Phe Phe Pro Trp Asp	Ala Asp Val Asp Lys	Phe Lys Ala
1310	1315	1320
Trp Arg Gln Gly Arg Thr Gly	Tyr Pro Leu Val Asp	Ala Gly Met
1325	1330	1335
Arg Glu Leu Trp Ala Thr Gly	Trp Met His Asn Arg	Ile Arg Val
1340	1345	1350
Ile Val Ser Ser Phe Ala Val	Lys Phe Leu Leu Leu	Pro Trp Lys
1355	1360	1365
Trp Gly Met Lys Tyr Phe Trp	Asp Thr Leu Leu Asp	Ala Asp Leu
1370	1375	1380
Glu Cys Asp Ile Leu Gly Trp	Gln Tyr Ile Ser Gly	Ser Ile Pro
1385	1390	1395
Asp Gly His Glu Leu Asp Arg	Leu Asp Asn Pro Ala	Leu Gln Gly
1400	1405	1410
Ala Lys Tyr Asp Pro Glu Gly	Glu Tyr Ile Arg Gln	Trp Leu Pro
1415	1420	1425
Glu Leu Ala Arg Leu Pro Thr	Glu Trp Ile His His	Pro Trp Asp
1430	1435	1440
Ala Pro Leu Thr Val Leu Lys	Ala Ser Gly Val Glu	Leu Gly Thr
1445	1450	1455
Asn Tyr Ala Lys Pro Ile Val	Asp Ile Asp Thr Ala	Arg Glu Leu
1460	1465	1470

Leu Ala Lys Ala Ile Ser Arg Thr Arg Glu Ala Gln Ile Met Ile  
 1475 1480 1485

Gly Ala Ala Pro Asp Glu Ile Val Ala Asp Ser Phe Glu Ala Leu  
 1490 1495 1500

Gly Ala Asn Thr Ile Lys Glu Pro Gly Leu Cys Pro Ser Val Ser  
 1505 1510 1515

Ser Asn Asp Gln Gln Val Pro Ser Ala Val Arg Tyr Asn Gly Ser  
 1520 1525 1530

Lys Arg Val Lys Pro Glu Glu Glu Glu Glu Arg Asp Met Lys Lys  
 1535 1540 1545

Ser Arg Gly Phe Asp Glu Arg Glu Leu Phe Ser Thr Ala Glu Ser  
 1550 1555 1560

Ser Ser Ser Ser Ser Val Phe Phe Val Ser Gln Ser Cys Ser Leu  
 1565 1570 1575

Ala Ser Glu Gly Lys Asn Leu Glu Gly Ile Gln Asp Ser Ser Asp  
 1580 1585 1590

Gln Ile Thr Thr Ser Leu Gly Lys Asn Gly  
 1595 1600

<210> 133

<211> 1492

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 133

Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala Phe  
 1 5 10 15  
 Ser Ala Asp Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser Leu  
 20 25 30  
 Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His His  
 35 40 45

Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu Arg  
 50 55 60  
  
 Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala Ala  
 65 70 75 80  
 Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln Pro  
 85 90 95  
 Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr  
 100 105 110  
 Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val  
 115 120 125  
  
 Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His  
 130 135 140  
 Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val  
 145 150 155 160  
 Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala  
 165 170 175  
 Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala  
 180 185 190  
  
 Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp  
 195 200 205  
 Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val  
 210 215 220  
 Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn  
 225 230 235 240  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys  
 245 250 255  
  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 260 265 270  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 275 280 285  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys

290 295 300  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 305 310 315 320  
  
 Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 325 330 335  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 340 345 350  
 Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 355 360 365  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 370 375 380  
  
 Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 385 390 395 400  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 405 410 415  
 Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val  
 420 425 430  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 435 440 445  
  
 Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu  
 450 455 460  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 465 470 475 480  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 485 490 495  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 500 505 510  
  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 515 520 525  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 530 535 540

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
 545                    550                    555                    560  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                          565                    570                    575  
  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
                          580                    585                    590  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
                          595                    600                    605  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                          610                    615                    620  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 625                    630                    635                    640  
  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
                          645                    650                    655  
 Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
                          660                    665                    670  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
                          675                    680                    685  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
                          690                    695                    700  
  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 705                    710                    715                    720  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
                          725                    730                    735  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
                          740                    745                    750  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
                          755                    760                    765  
  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                          770                    775                    780  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Arg

785                    790                    795                    800  
 Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala  
                                  805                    810                    815  
 Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly  
                                  820                    825                    830  
  
 Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro  
                                  835                    840                    845  
 Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser His  
                                  850                    855                    860  
 Arg Val Ala Asp His Ala Gln Val Val Arg Val Leu Gly Phe Phe Gln  
 865                    870                    875                    880  
 Cys His Ser His Pro Ala Gln Ala Phe Asp Asp Ala Met Thr Gln Phe  
                                  885                    890                    895  
  
 Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe Arg Arg Val Gly Val  
                                  900                    905                    910  
 Thr Glu Leu Glu Ala Arg Ser Gly Thr Leu Pro Pro Ala Ser Gln Arg  
                                  915                    920                    925  
 Trp Asp Arg Ile Leu Gln Ala Ser Gly Met Lys Arg Ala Lys Pro Ser  
                                  930                    935                    940  
 Pro Thr Ser Thr Gln Thr Pro Asp Gln Ala Ser Leu His Ala Phe Ala  
 945                    950                    955                    960  
  
 Asp Ser Leu Glu Arg Asp Leu Asp Ala Pro Ser Pro Met His Glu Gly  
                                  965                    970                    975  
 Asp Gln Thr Arg Ala Ser Ala Ser Pro Lys Lys Lys Arg Lys Val Glu  
                                  980                    985                    990  
 Ala Ser Lys Met Asp Lys Lys Thr Ile Val Trp Phe Arg Arg Asp Leu  
                                  995                    1000                    1005  
 Arg Ile Glu Asp Asn Pro Ala Leu Ala Ala Ala Ala His Glu Gly  
                                  1010                    1015                    1020  
  
 Ser Val Phe Pro Val Phe Ile Trp Cys Pro Glu Glu Glu Gly Gln  
                                  1025                    1030                    1035

Phe Tyr Pro Gly Arg Ala Ser Arg Trp Trp Met Lys Gln Ser Leu  
 1040 1045 1050  
 Ala His Leu Ser Gln Ser Leu Lys Ala Leu Gly Ser Asp Leu Thr  
 1055 1060 1065  
 Leu Ile Lys Thr His Asn Thr Ile Ser Ala Ile Leu Asp Cys Ile  
 1070 1075 1080  
  
 Arg Val Thr Gly Ala Thr Lys Val Val Phe Asn His Leu Tyr Asp  
 1085 1090 1095  
 Pro Val Ser Leu Val Arg Asp His Thr Val Lys Glu Lys Leu Val  
 1100 1105 1110  
 Glu Arg Gly Ile Ser Val Gln Ser Tyr Asn Gly Asp Leu Leu Tyr  
 1115 1120 1125  
 Glu Pro Trp Glu Ile Tyr Cys Glu Lys Gly Lys Pro Phe Thr Ser  
 1130 1135 1140  
  
 Phe Asn Ser Tyr Trp Lys Lys Cys Leu Asp Met Ser Ile Glu Ser  
 1145 1150 1155  
 Val Met Leu Pro Pro Pro Trp Arg Leu Met Pro Ile Thr Ala Ala  
 1160 1165 1170  
 Ala Glu Ala Ile Trp Ala Cys Ser Ile Glu Glu Leu Gly Leu Glu  
 1175 1180 1185  
 Asn Glu Ala Glu Lys Pro Ser Asn Ala Leu Leu Thr Arg Ala Trp  
 1190 1195 1200  
  
 Ser Pro Gly Trp Ser Asn Ala Asp Lys Leu Leu Asn Glu Phe Ile  
 1205 1210 1215  
 Glu Lys Gln Leu Ile Asp Tyr Ala Lys Asn Ser Lys Lys Val Val  
 1220 1225 1230  
 Gly Asn Ser Thr Ser Leu Leu Ser Pro Tyr Leu His Phe Gly Glu  
 1235 1240 1245  
 Ile Ser Val Arg His Val Phe Gln Cys Ala Arg Met Lys Gln Ile  
 1250 1255 1260  
  
 Ile Trp Ala Arg Asp Lys Asn Ser Glu Gly Glu Glu Ser Ala Asp

1265	1270	1275
Leu Phe Leu Arg Gly Ile Gly	Leu Arg Glu Tyr Ser	Arg Tyr Ile
1280	1285	1290
Cys Phe Asn Phe Pro Phe Thr	His Glu Gln Ser Leu	Leu Ser His
1295	1300	1305
Leu Arg Phe Phe Pro Trp Asp	Ala Asp Val Asp Lys	Phe Lys Ala
1310	1315	1320
Trp Arg Gln Gly Arg Thr Gly	Tyr Pro Leu Val Asp	Ala Gly Met
1325	1330	1335
Arg Glu Leu Trp Ala Thr Gly	Trp Met His Asn Arg	Ile Arg Val
1340	1345	1350
Ile Val Ser Ser Phe Ala Val	Lys Phe Leu Leu Leu	Pro Trp Lys
1355	1360	1365
Trp Gly Met Lys Tyr Phe Trp	Asp Thr Leu Leu Asp	Ala Asp Leu
1370	1375	1380
Glu Cys Asp Ile Leu Gly Trp	Gln Tyr Ile Ser Gly	Ser Ile Pro
1385	1390	1395
Asp Gly His Glu Leu Asp Arg	Leu Asp Asn Pro Ala	Leu Gln Gly
1400	1405	1410
Ala Lys Tyr Asp Pro Glu Gly	Glu Tyr Ile Arg Gln	Trp Leu Pro
1415	1420	1425
Glu Leu Ala Arg Leu Pro Thr	Glu Trp Ile His His	Pro Trp Asp
1430	1435	1440
Ala Pro Leu Thr Val Leu Lys	Ala Ser Gly Val Glu	Leu Gly Thr
1445	1450	1455
Asn Tyr Ala Lys Pro Ile Val	Asp Ile Asp Thr Ala	Arg Glu Leu
1460	1465	1470
Leu Ala Lys Ala Ile Ser Arg	Thr Arg Glu Ala Gln	Ile Met Ile
1475	1480	1485
Gly Ala Ala Pro		
1490		

<210> 134

<211> 665

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 134

Met Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn Phe Pro Asp Met  
 1                    5                    10                    15  
 Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr Leu Asn Pro Thr  
                   20                    25                    30  
 Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser Ser Met Ile Thr  
                   35                    40                    45  
 Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly Leu Asn Leu Pro  
  
                   50                    55                    60  
 Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser Arg Leu Ser Ile  
 65                    70                    75                    80  
 Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys Lys Arg Lys Phe  
                   85                    90                    95  
 Asp Thr Glu Thr Lys Asp Cys Asn Glu Lys Lys Lys Lys Met Thr Met  
                   100                    105                    110  
 Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu Lys Ser Lys Ile  
  
                   115                    120                    125  
 Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys Lys Met Lys His  
                   130                    135                    140  
 Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser Lys Val  
 145                    150                    155                    160  
 Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His Val Arg Ala Arg Arg  
                   165                    170                    175  
 Gly Gln Ala Thr Asp Ser His Ser Ile Ala Glu Arg Val Arg Arg Glu  
  
                   180                    185                    190

Lys Ile Ser Glu Arg Met Lys Phe Leu Gln Asp Leu Val Pro Gly Cys  
 195 200 205  
 Asp Lys Ile Thr Gly Lys Ala Gly Met Leu Asp Glu Ile Ile Asn Tyr  
 210 215 220  
 Val Gln Ser Leu Gln Arg Gln Ile Glu Phe Leu Ser Met Lys Leu Ala  
 225 230 235 240  
 Ile Val Asn Pro Arg Pro Asp Phe Asp Met Asp Asp Ile Phe Ala Lys  
 245 250 255  
 Glu Val Ala Ser Thr Pro Met Thr Val Val Pro Ser Pro Glu Met Val  
 260 265 270  
 Leu Ser Gly Tyr Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu  
 275 280 285  
 Met Val Asn Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn  
 290 295 300  
 Thr Ser Ser Asp Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser  
 305 310 315 320  
 Met Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val Ala  
 325 330 335  
 Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Gly Ser Gly Arg Ala  
 340 345 350  
 Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu  
 355 360 365  
 Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe  
 370 375 380  
 Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp  
 385 390 395 400  
 Met Leu Ile Asn Ser Arg Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu  
 405 410 415  
 Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro Val Ser Lys Gly Glu  
 420 425 430  
 Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp

435                      440                      445  
 Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala  
 450                      455                      460  
 Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu  
 465                      470                      475                      480  
 Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln  
 485                      490                      495  
 Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys  
  
 500                      505                      510  
 Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys  
 515                      520                      525  
 Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp  
 530                      535                      540  
 Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp  
 545                      550                      555                      560  
 Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn  
  
 565                      570                      575  
 Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe  
 580                      585                      590  
 Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His  
 595                      600                      605  
 Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp  
 610                      615                      620  
 Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu  
  
 625                      630                      635                      640  
 Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile  
 645                      650                      655  
 Thr Leu Gly Met Asp Glu Leu Tyr Lys  
 660                      665

<210> 135

<211> 500

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 135

Met Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn Phe Pro Asp Met

1                    5                    10                    15

Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr Leu Asn Pro Thr

20                    25                    30

Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser Ser Met Ile Thr

35                    40                    45

Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly Leu Asn Leu Pro

50                    55                    60

Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser Arg Leu Ser Ile

65                    70                    75                    80

Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys Lys Arg Lys Phe

85                    90                    95

Asp Thr Glu Thr Lys Asp Cys Asn Glu Lys Lys Lys Lys Met Thr Met

100                    105                    110

Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu Lys Ser Lys Ile

115                    120                    125

Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys Lys Met Lys His

130                    135                    140

Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser Lys Val

145                    150                    155                    160

Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile Ala Ser Pro Lys Lys Lys

165                    170                    175

Arg Lys Val Glu Ala Ser Gly Ser Gly Arg Ala Asp Ala Leu Asp Asp

180                    185                    190

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu

195                    200                    205

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu  
 210 215 220  
 Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Ile Asn Ser  
 225 230 235 240  
 Arg Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val  
 245 250 255  
 Glu Glu Asn Pro Gly Pro Val Ser Lys Gly Glu Glu Leu Phe Thr Gly  
 260 265 270  
 Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys  
 275 280 285  
 Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu  
 290 295 300  
 Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro  
 305 310 315 320  
 Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr  
 325 330 335  
 Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu  
 340 345 350  
 Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr  
 355 360 365  
 Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg  
 370 375 380  
 Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly  
 385 390 395 400  
 His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala  
 405 410 415  
 Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn  
 420 425 430  
 Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr  
 435 440 445  
 Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser

450                      455                      460  
 Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met  
 465                      470                      475                      480  
 Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp  
                             485                      490                      495  
 Glu Leu Tyr Lys  
                             500  
 <210> 136  
 <211> 693  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic

polypeptide"  
 <400> 136  
 Met Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn Phe Pro Asp Met  
 1                      5                      10                      15  
 Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr Leu Asn Pro Thr  
                             20                      25                      30  
 Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser Ser Met Ile Thr  
                             35                      40                      45  
 Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly Leu Asn Leu Pro  
                             50                      55                      60  
  
 Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser Arg Leu Ser Ile  
 65                      70                      75                      80  
 Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys Lys Arg Lys Phe  
                             85                      90                      95  
 Asp Thr Glu Thr Lys Asp Cys Asn Glu Lys Lys Lys Lys Met Thr Met  
                             100                      105                      110  
 Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu Lys Ser Lys Ile  
                             115                      120                      125  
  
 Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys Lys Met Lys His

130                      135                      140  
 Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser Lys Val  
 145                      150                      155                      160  
 Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His Val Arg Ala Arg Arg  
                          165                      170                      175  
 Gly Gln Ala Thr Asp Ser His Ser Ile Ala Glu Arg Val Arg Arg Glu  
                          180                      185                      190  
  
 Lys Ile Ser Glu Arg Met Lys Phe Leu Gln Asp Leu Val Pro Gly Cys  
                          195                      200                      205  
 Asp Lys Ile Thr Gly Lys Ala Gly Met Leu Asp Glu Ile Ile Asn Tyr  
                          210                      215                      220  
 Val Gln Ser Leu Gln Arg Gln Ile Glu Phe Leu Ser Met Lys Leu Ala  
 225                      230                      235                      240  
 Ile Val Asn Pro Arg Pro Asp Phe Asp Met Asp Asp Ile Phe Ala Lys  
                          245                      250                      255  
  
 Glu Val Ala Ser Thr Pro Met Thr Val Val Pro Ser Pro Glu Met Val  
                          260                      265                      270  
 Leu Ser Gly Tyr Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu  
                          275                      280                      285  
 Met Val Asn Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn  
                          290                      295                      300  
 Thr Ser Ser Asp Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser  
 305                      310                      315                      320  
  
 Met Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val Ala  
                          325                      330                      335  
 Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Ala Pro Pro Thr Asp  
                          340                      345                      350  
 Val Ser Leu Gly Asp Glu Leu His Leu Asp Gly Glu Asp Val Ala Met  
                          355                      360                      365  
 Ala His Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp  
                          370                      375                      380

Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr  
 385                      390                      395                      400

Gly Ala Leu Asp Met Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp  
                                  405                      410                      415

Ala Leu Gly Ile Asp Glu Tyr Gly Gly Glu Phe Pro Gly Ile Arg Arg  
                                  420                      425                      430

Ser Arg Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp  
                                  435                      440                      445

Val Glu Glu Asn Pro Gly Pro Val Ser Lys Gly Glu Glu Leu Phe Thr  
                                  450                      455                      460

Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His  
 465                      470                      475                      480

Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys  
                                  485                      490                      495

Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp  
                                  500                      505                      510

Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg  
                                  515                      520                      525

Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro  
                                  530                      535                      540

Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn  
 545                      550                      555                      560

Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn  
                                  565                      570                      575

Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu  
                                  580                      585                      590

Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met  
                                  595                      600                      605

Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His  
                                  610                      615                      620

Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn

625                            630                            635                            640  
  
 Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu  
  
                                 645                            650                            655  
 Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His  
  
                                 660                            665                            670  
 Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met  
  
                                 675                            680                            685  
  
 Asp Glu Leu Tyr Lys  
  
 690  
 <210> 137  
 <211> 792  
 <212> PRT  
 <213> Artificial Sequence  
  
 <220>  
 <221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
                                 polypeptide"  
 <400> 137  
  
 Met Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn Phe Pro Asp Met  
 1                            5                            10                            15  
 Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr Leu Asn Pro Thr  
  
                                 20                            25                            30  
 Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser Ser Met Ile Thr  
  
                                 35                            40                            45  
 Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly Leu Asn Leu Pro  
  
  
  
                                 50                            55                            60  
 Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser Arg Leu Ser Ile  
 65                            70                            75                            80  
 Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys Lys Arg Lys Phe  
  
                                 85                            90                            95  
 Asp Thr Glu Thr Lys Asp Cys Asn Glu Lys Lys Lys Lys Met Thr Met  
  
                                 100                            105                            110

Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu Lys Ser Lys Ile  
 115 120 125  
 Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys Lys Met Lys His  
 130 135 140  
 Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser Lys Val  
 145 150 155 160  
 Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His Val Arg Ala Arg Arg  
 165 170 175  
 Gly Gln Ala Thr Asp Ser His Ser Ile Ala Glu Arg Val Arg Arg Glu  
 180 185 190  
 Lys Ile Ser Glu Arg Met Lys Phe Leu Gln Asp Leu Val Pro Gly Cys  
 195 200 205  
 Asp Lys Ile Thr Gly Lys Ala Gly Met Leu Asp Glu Ile Ile Asn Tyr  
 210 215 220  
 Val Gln Ser Leu Gln Arg Gln Ile Glu Phe Leu Ser Met Lys Leu Ala  
 225 230 235 240  
 Ile Val Asn Pro Arg Pro Asp Phe Asp Met Asp Asp Ile Phe Ala Lys  
 245 250 255  
 Glu Val Ala Ser Thr Pro Met Thr Val Val Pro Ser Pro Glu Met Val  
 260 265 270  
 Leu Ser Gly Tyr Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu  
 275 280 285  
 Met Val Asn Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn  
 290 295 300  
 Thr Ser Ser Asp Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser  
 305 310 315 320  
 Met Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val Ala  
 325 330 335  
 Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Pro Ser Gly Gln Ile  
 340 345 350  
 Ser Asn Gln Ala Leu Ala Leu Ala Pro Ser Ser Ala Pro Val Leu Ala

355                            360                            365  
 Gln Thr Met Val Pro Ser Ser Ala Met Val Pro Leu Ala Gln Pro Pro

370                            375                            380  
 Ala Pro Ala Pro Val Leu Thr Pro Gly Pro Pro Gln Ser Leu Ser Ala  
 385                            390                            395                            400

Pro Val Pro Lys Ser Thr Gln Ala Gly Glu Gly Thr Leu Ser Glu Ala  
 405                            410                            415

Leu Leu His Leu Gln Phe Asp Ala Asp Glu Asp Leu Gly Ala Leu Leu  
 420                            425                            430

Gly Asn Ser Thr Asp Pro Gly Val Phe Thr Asp Leu Ala Ser Val Asp

435                            440                            445  
 Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln Gly Val Ser Met Ser His  
 450                            455                            460

Ser Thr Ala Glu Pro Met Leu Met Glu Tyr Pro Glu Ala Ile Thr Arg  
 465                            470                            475                            480

Leu Val Thr Gly Ser Gln Arg Pro Pro Asp Pro Ala Pro Thr Pro Leu  
 485                            490                            495

Gly Thr Ser Gly Leu Pro Asn Gly Leu Ser Gly Asp Glu Asp Phe Ser

500                            505                            510  
 Ser Ile Ala Asp Met Asp Phe Ser Ala Leu Leu Ser Gln Ile Ser Ser  
 515                            520                            525

Ser Gly Gln Ser Arg Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr  
 530                            535                            540

Cys Gly Asp Val Glu Glu Asn Pro Gly Pro Val Ser Lys Gly Glu Glu  
 545                            550                            555                            560

Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val

565                            570                            575  
 Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr  
 580                            585                            590

Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro  
 595                            600                            605

Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys  
 610                      615                      620  
 Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser  
  
 625                      630                      635                      640  
 Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp  
                                  645                      650                      655  
 Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr  
                                  660                      665                      670  
 Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly  
                                  675                      680                      685  
 Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val  
  
                                  690                      695                      700  
 Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys  
 705                      710                      715                      720  
 Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr  
                                  725                      730                      735  
 Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn  
                                  740                      745                      750  
 His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys  
  
                                  755                      760                      765  
 Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr  
                                  770                      775                      780  
 Leu Gly Met Asp Glu Leu Tyr Lys  
 785                      790  
 <210> 138  
 <211> 967  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
                                  polypeptide"  
 <220><221> MOD\_RES

<222> (158)..(159)  
<223> Any amino acid  
<220><221> MOD\_RES

<222> (192)..(193)  
<223> Any amino acid  
<220><221> MOD\_RES

<222> (226)..(227)  
<223> Any amino acid  
<220><221> MOD\_RES

<222> (260)..(261)  
<223> Any amino acid  
<220><221> MOD\_RES

<222> (294)..(295)  
<223> Any amino acid  
<220><221> MOD\_RES

<222> (328)..(329)  
<223> Any amino acid  
<220><221> MOD\_RES

<222> (362)..(363)  
<223> Any amino acid  
<220><221> MOD\_RES

<222> (396)..(397)  
<223> Any amino acid  
<220><221> MOD\_RES

<222> (430)..(431)  
<223> Any amino acid  
<220><221> MOD\_RES

<222> (464)..(465)  
<223> Any amino acid  
<220><221> MOD\_RES

<222> (498)..(499)  
<223> Any amino acid  
<220><221> MOD\_RES

<222> (532)..(533)

<223> Any amino acid

<220><221> MOD\_RES

<222> (566)..(567)

<223> Any amino acid

<400> 138

Met Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Val Asp Leu Arg Thr Leu

1                    5                    10                    15

Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser

                  20                    25                    30

Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His

                  35                    40                    45

Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val

                  50                    55                    60

Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His

65                    70                    75                    80

Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu

                  85                    90                    95

Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln

                  100                    105                    110

Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr

                  115                    120                    125

Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro

                  130                    135                    140

Leu Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly

145                    150                    155                    160

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys

                  165                    170                    175

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa

                  180                    185                    190

Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val

                  195                    200                    205

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 210 215 220

Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 225 230 235 240

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 245 250 255

Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 260 265 270

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 275 280 285

Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val  
 290 295 300

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305 310 315 320

Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu  
 325 330 335

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 340 345 350

Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala  
 355 360 365

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 370 375 380

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys  
 385 390 395 400

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 405 410 415

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly  
 420 425 430

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 435 440 445

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa

450                      455                      460  
 Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465                      470                      475                      480  
  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                                  485                      490                      495  
 Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
                                  500                      505                      510  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
                                  515                      520                      525  
 Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
                                  530                      535                      540  
  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545                      550                      555                      560  
 Val Ala Ile Ala Ser Xaa Xaa Gly Gly Arg Pro Ala Leu Glu Ser Ile  
                                  565                      570                      575  
 Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn  
                                  580                      585                      590  
 Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp  
                                  595                      600                      605  
  
 Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr  
                                  610                      615                      620  
 Asn Arg Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Pro  
 625                      630                      635                      640  
 Lys Lys Lys Arg Lys Val Glu Ala Ser Gly Ser Gly Arg Ala Asp Ala  
                                  645                      650                      655  
 Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp  
                                  660                      665                      670  
  
 Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu  
                                  675                      680                      685  
 Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu  
                                  690                      695                      700

Ile Asn Ser Arg Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys  
 705                      710                      715                      720  
 Gly Asp Val Glu Glu Asn Pro Gly Pro Val Ser Lys Gly Glu Glu Leu  
                                  725                      730                      735

Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn  
                                  740                      745                      750  
 Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr  
                                  755                      760                      765

Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val  
                                  770                      775                      780  
 Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe  
 785                      790                      795                      800

Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala  
                                  805                      810                      815  
 Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp  
                                  820                      825                      830  
 Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu  
                                  835                      840                      845  
 Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn  
                                  850                      855                      860

Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr  
 865                      870                      875                      880  
 Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile  
                                  885                      890                      895  
 Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln  
                                  900                      905                      910  
 Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His  
                                  915                      920                      925

Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg  
                                  930                      935                      940  
 Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu

945                    950                    955                    960

Gly Met Asp Glu Leu Tyr Lys

                         965

<210> 139

<211> 922

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
                         polypeptide"

<220><221> MOD\_RES

<222> (158)..(159)

<223> Any amino acid

<220><221> MOD\_RES

<222> (192)..(193)

<223> Any amino acid

<220><221> MOD\_RES

<222> (226)..(227)

<223> Any amino acid

<220><221> MOD\_RES

<222> (260)..(261)

<223> Any amino acid

<220><221> MOD\_RES

<222> (294)..(295)

<223> Any amino acid

<220><221> MOD\_RES

<222> (328)..(329)

<223> Any amino acid

<220><221> MOD\_RES

<222> (362)..(363)

<223> Any amino acid

<220><221> MOD\_RES

<222> (396)..(397)

<223> Any amino acid

<220><221> MOD\_RES

<222> (430)..(431)

<223> Any amino acid

<220><221> MOD\_RES

<222> (464)..(465)

<223> Any amino acid

<220><221> MOD\_RES

<222> (498)..(499)

<223> Any amino acid

<220><221> MOD\_RES

<222> (532)..(533)

<223> Any amino acid

<220><221> MOD\_RES

<222> (566)..(567)

<223> Any amino acid

<400> 139

Met Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Val Asp Leu Arg Thr Leu

1                    5                    10                    15

Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser

20                    25                    30

Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His

35                    40                    45

Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val

50                    55                    60

Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His

65                    70                    75                    80

Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu

85                    90                    95

Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln

100                    105                    110

Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr

115                    120                    125

Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro  
 130 135 140  
 Leu Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly  
 145 150 155 160  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 165 170 175  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa  
 180 185 190  
 Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 195 200 205  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 210 215 220  
 Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 225 230 235 240  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 245 250 255  
 Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 260 265 270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 275 280 285  
 Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val  
 290 295 300  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305 310 315 320  
 Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu  
 325 330 335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 340 345 350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala  
 355 360 365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly



Asn Arg Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Pro  
 625                      630                      635                      640  
 Lys Lys Lys Arg Lys Val Glu Ala Ser Pro Lys Lys Lys Arg Lys Val  
                                  645                      650                      655  
 Glu Ala Ser Gly Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala  
                                  660                      665                      670  
 Asp Tyr Leu Glu Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser  
                                  675                      680                      685  
 Met Leu Pro Gly Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala  
                                  690                      695                      700  
 Asp Tyr Leu Glu Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser  
 705                      710                      715                      720  
 Met Leu Pro Gly Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala  
                                  725                      730                      735  
 Asp Tyr Leu Glu Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser  
                                  740                      745                      750  
 Met Leu Pro Gly Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala  
                                  755                      760                      765  
 Asp Tyr Leu Glu Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser  
                                  770                      775                      780  
 Met Leu Pro Ser Arg Ser Arg Gly Ser Gly Glu Gly Arg Gly Ser Leu  
 785                      790                      795                      800  
 Leu Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro Ile Glu Lys Ser  
                                  805                      810                      815  
 Phe Val Ile Thr Asp Pro Arg Leu Pro Asp Tyr Pro Ile Ile Phe Ala  
                                  820                      825                      830  
 Ser Asp Gly Phe Leu Glu Leu Thr Glu Tyr Ser Arg Glu Glu Ile Met  
                                  835                      840                      845  
 Gly Arg Asn Ala Arg Phe Leu Gln Gly Pro Glu Thr Asp Gln Ala Thr  
                                  850                      855                      860  
 Val Gln Lys Ile Arg Asp Ala Ile Arg Asp Gln Arg Glu Thr Thr Val

865                    870                    875                    880  
 Gln Leu Ile Asn Tyr Thr Lys Ser Gly Lys Lys Phe Trp Asn Leu Leu  
                          885                    890                    895  
 His Leu Gln Pro Val Arg Asp Arg Lys Gly Gly Leu Gln Tyr Phe Ile  
                          900                    905                    910  
 Gly Val Gln Leu Val Gly Ser Asp His Val  
                          915                    920

<210> 140

<211> 983

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<220><221> MOD\_RES

<222> (158)..(159)

<223> Any amino acid

<220><221> MOD\_RES

<222> (192)..(193)

<223> Any amino acid

<220><221> MOD\_RES

<222> (226)..(227)

<223> Any amino acid

<220><221> MOD\_RES

<222> (260)..(261)

<223> Any amino acid

<220><221> MOD\_RES

<222> (294)..(295)

<223> Any amino acid

<220><221> MOD\_RES

<

<222> (328)..(329)

<223> Any amino acid

<220><221> MOD\_RES

<222> (362)..(363)  
 <223> Any amino acid  
 <220><221> MOD\_RES  
 <222> (396)..(397)  
 <223> Any amino acid  
 <220><221> MOD\_RES  
 <222> (430)..(431)  
 <223> Any amino acid  
 <220><221> MOD\_RES  
 <222> (464)..(465)  
 <223> Any amino acid  
 <220><221> MOD\_RES  
 <222> (498)..(499)  
 <223> Any amino acid  
 <220><221> MOD\_RES  
 <222> (532)..(533)  
 <223> Any amino acid  
 <220><221> MOD\_RES  
 <222> (566)..(567)  
 <223> Any amino acid  
 <400> 140

Met Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Val Asp Leu Arg Thr Leu

1	5	10	15
Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser			
	20	25	30
Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His			
	35	40	45
Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val			
	50	55	60
Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His			
65	70	75	80
Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu			
	85	90	95

Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln  
 100 105 110  
 Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr  
 115 120 125  
 Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro  
 130 135 140  
 Leu Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly  
 145 150 155 160  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 165 170 175  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa  
 180 185 190  
 Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 195 200 205  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 210 215 220  
 Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 225 230 235 240  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 245 250 255  
 Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 260 265 270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 275 280 285  
 Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val  
 290 295 300  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305 310 315 320  
 Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu  
 325 330 335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr

340 345 350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala  
 355 360 365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 370 375 380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys  
 385 390 395 400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 405 410 415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly  
 420 425 430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 435 440 445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa  
 450 455 460  
 Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465 470 475 480  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495  
 Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525  
 Ile Ala Ser Xaa Xaa Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560  
 Val Ala Ile Ala Ser Xaa Xaa Gly Gly Arg Pro Ala Leu Glu Ser Ile  
 565 570 575  
 Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn  
 580 585 590

Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp  
 595 600 605  
 Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr  
 610 615 620  
 Asn Arg Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Pro  
 625 630 635 640  
 Lys Lys Lys Arg Lys Val Glu Ala Ser Asn Gly Ala Ile Gly Gly Asp  
 645 650 655  
 Leu Leu Leu Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala  
 660 665 670  
 His Leu Lys Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe  
 675 680 685  
 Phe Ala Asp Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu  
 690 695 700  
 Ser Thr Ala Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val  
 705 710 715 720  
 Glu Gly Asp Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr  
 725 730 735  
 Gly Asn Phe Lys Lys Arg Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn  
 740 745 750  
 Glu Lys Lys Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu  
 755 760 765  
 Gly Glu Glu Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr  
 770 775 780  
 Lys Ser Ile Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn  
 785 790 795 800  
 Phe Ser Asn Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp  
 805 810 815  
 Tyr Ile His Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser  
 820 825 830  
 Ile Ala Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe

835                                    840                                    845  
 Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly  
 850                                    855                                    860  
 Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln Ile  
 865                                    870                                    875                                    880  
 Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro Asp Phe  
 885                                    890                                    895  
 Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr Pro Met Thr

900                                    905                                    910  
 Val Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr Ser His Glu Met  
 915                                    920                                    925  
 Val His Ser Gly Tyr Ser Ser Glu Met Val Asn Ser Gly Tyr Leu His  
 930                                    935                                    940  
 Val Asn Pro Met Gln Gln Val Asn Thr Ser Ser Asp Pro Leu Ser Cys  
 945                                    950                                    955                                    960  
 Phe Asn Asn Gly Glu Ala Pro Ser Met Trp Asp Ser His Val Gln Asn

965                                    970                                    975  
 Leu Tyr Gly Asn Leu Gly Val

980  
 <210> 141  
 <211> 830  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
                     polypeptide"

<400> 141  
 Met Lys Met Asp Lys Lys Thr Ile Val Trp Phe Arg Arg Asp Leu Arg  
 1                                    5                                    10                                    15  
 Ile Glu Asp Asn Pro Ala Leu Ala Ala Ala Ala His Glu Gly Ser Val  
 20                                    25                                    30

Phe Pro Val Phe Ile Trp Cys Pro Glu Glu Glu Gly Gln Phe Tyr Pro

	35	40	45	
Gly Arg Ala Ser Arg Trp Trp Met Lys Gln Ser Leu Ala His Leu Ser				
	50	55	60	
Gln Ser Leu Lys Ala Leu Gly Ser Asp Leu Thr Leu Ile Lys Thr His				
65	70	75	80	
Asn Thr Ile Ser Ala Ile Leu Asp Cys Ile Arg Val Thr Gly Ala Thr				
	85	90	95	
Lys Val Val Phe Asn His Leu Tyr Asp Pro Val Ser Leu Val Arg Asp				
	100	105	110	
His Thr Val Lys Glu Lys Leu Val Glu Arg Gly Ile Ser Val Gln Ser				
	115	120	125	
Tyr Asn Gly Asp Leu Leu Tyr Glu Pro Trp Glu Ile Tyr Cys Glu Lys				
	130	135	140	
Gly Lys Pro Phe Thr Ser Phe Asn Ser Tyr Trp Lys Lys Cys Leu Asp				
145	150	155	160	
Met Ser Ile Glu Ser Val Met Leu Pro Pro Pro Trp Arg Leu Met Pro				
	165	170	175	
Ile Thr Ala Ala Ala Glu Ala Ile Trp Ala Cys Ser Ile Glu Glu Leu				
	180	185	190	
Gly Leu Glu Asn Glu Ala Glu Lys Pro Ser Asn Ala Leu Leu Thr Arg				
	195	200	205	
Ala Trp Ser Pro Gly Trp Ser Asn Ala Asp Lys Leu Leu Asn Glu Phe				
	210	215	220	
Ile Glu Lys Gln Leu Ile Asp Tyr Ala Lys Asn Ser Lys Lys Val Val				
225	230	235	240	
Gly Asn Ser Thr Ser Leu Leu Ser Pro Tyr Leu His Phe Gly Glu Ile				
	245	250	255	
Ser Val Arg His Val Phe Gln Cys Ala Arg Met Lys Gln Ile Ile Trp				
	260	265	270	
Ala Arg Asp Lys Asn Ser Glu Gly Glu Glu Ser Ala Asp Leu Phe Leu				
	275	280	285	

Arg Gly Ile Gly Leu Arg Glu Tyr Ser Arg Tyr Ile Cys Phe Asn Phe  
 290 295 300

Pro Phe Thr His Glu Gln Ser Leu Leu Ser His Leu Arg Phe Phe Pro  
 305 310 315 320

Trp Asp Ala Asp Val Asp Lys Phe Lys Ala Trp Arg Gln Gly Arg Thr  
 325 330 335

Gly Tyr Pro Leu Val Asp Ala Gly Met Arg Glu Leu Trp Ala Thr Gly  
 340 345 350

Trp Met His Asn Arg Ile Arg Val Ile Val Ser Ser Phe Ala Val Lys  
 355 360 365

Phe Leu Leu Leu Pro Trp Lys Trp Gly Met Lys Tyr Phe Trp Asp Thr  
 370 375 380

Leu Leu Asp Ala Asp Leu Glu Cys Asp Ile Leu Gly Trp Gln Tyr Ile  
 385 390 395 400

Ser Gly Ser Ile Pro Asp Gly His Glu Leu Asp Arg Leu Asp Asn Pro  
 405 410 415

Ala Leu Gln Gly Ala Lys Tyr Asp Pro Glu Gly Glu Tyr Ile Arg Gln  
 420 425 430

Trp Leu Pro Glu Leu Ala Arg Leu Pro Thr Glu Trp Ile His His Pro  
 435 440 445

Trp Asp Ala Pro Leu Thr Val Leu Lys Ala Ser Gly Val Glu Leu Gly  
 450 455 460

Thr Asn Tyr Ala Lys Pro Ile Val Asp Ile Asp Thr Ala Arg Glu Leu  
 465 470 475 480

Leu Ala Lys Ala Ile Ser Arg Thr Arg Glu Ala Gln Ile Met Ile Gly  
 485 490 495

Ala Ala Pro Ala Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Gly  
 500 505 510

Ser Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly  
 515 520 525

Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala

530	535	540	
Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp			
545	550	555	560
Phe Asp Leu Asp Met Leu Ile Asn Ser Arg Gly Ser Gly Glu Gly Arg			
	565	570	575
Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro Val			
	580	585	590
Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu			
595	600	605	
Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly			
610	615	620	
Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr			
625	630	635	640
Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr			
	645	650	655
Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His			
660	665	670	
Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr			
675	680	685	
Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys			
690	695	700	
Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp			
705	710	715	720
Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr			
	725	730	735
Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile			
740	745	750	
Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln			
755	760	765	
Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val			
770	775	780	

Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys  
 785 790 795 800

Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr  
 805 810 815

Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Val  
 820 825 830

<210> 142

<211> 774

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 142

Met Lys Met Asp Lys Lys Thr Ile Val Trp Phe Arg Arg Asp Leu Arg  
 1 5 10 15

Ile Glu Asp Asn Pro Ala Leu Ala Ala Ala Ala His Glu Gly Ser Val  
 20 25 30

Phe Pro Val Phe Ile Trp Cys Pro Glu Glu Glu Gly Gln Phe Tyr Pro  
 35 40 45

Gly Arg Ala Ser Arg Trp Trp Met Lys Gln Ser Leu Ala His Leu Ser  
 50 55 60

Gln Ser Leu Lys Ala Leu Gly Ser Asp Leu Thr Leu Ile Lys Thr His  
 65 70 75 80

Asn Thr Ile Ser Ala Ile Leu Asp Cys Ile Arg Val Thr Gly Ala Thr  
 85 90 95

Lys Val Val Phe Asn His Leu Tyr Asp Pro Val Ser Leu Val Arg Asp  
 100 105 110

His Thr Val Lys Glu Lys Leu Val Glu Arg Gly Ile Ser Val Gln Ser  
 115 120 125

Tyr Asn Gly Asp Leu Leu Tyr Glu Pro Trp Glu Ile Tyr Cys Glu Lys

130	135	140	
Gly Lys Pro Phe Thr Ser Phe Asn Ser Tyr Trp Lys Lys Cys Leu Asp			
145	150	155	160
Met Ser Ile Glu Ser Val Met Leu Pro Pro Pro Trp Arg Leu Met Pro			
	165	170	175
Ile Thr Ala Ala Ala Glu Ala Ile Trp Ala Cys Ser Ile Glu Glu Leu			
	180	185	190
Gly Leu Glu Asn Glu Ala Glu Lys Pro Ser Asn Ala Leu Leu Thr Arg			
195	200	205	
Ala Trp Ser Pro Gly Trp Ser Asn Ala Asp Lys Leu Leu Asn Glu Phe			
210	215	220	
Ile Glu Lys Gln Leu Ile Asp Tyr Ala Lys Asn Ser Lys Lys Val Val			
225	230	235	240
Gly Asn Ser Thr Ser Leu Leu Ser Pro Tyr Leu His Phe Gly Glu Ile			
	245	250	255
Ser Val Arg His Val Phe Gln Cys Ala Arg Met Lys Gln Ile Ile Trp			
	260	265	270
Ala Arg Asp Lys Asn Ser Glu Gly Glu Glu Ser Ala Asp Leu Phe Leu			
275	280	285	
Arg Gly Ile Gly Leu Arg Glu Tyr Ser Arg Tyr Ile Cys Phe Asn Phe			
290	295	300	
Pro Phe Thr His Glu Gln Ser Leu Leu Ser His Leu Arg Phe Phe Pro			
305	310	315	320
Trp Asp Ala Asp Val Asp Lys Phe Lys Ala Trp Arg Gln Gly Arg Thr			
	325	330	335
Gly Tyr Pro Leu Val Asp Ala Gly Met Arg Glu Leu Trp Ala Thr Gly			
340	345	350	
Trp Met His Asn Arg Ile Arg Val Ile Val Ser Ser Phe Ala Val Lys			
355	360	365	
Phe Leu Leu Leu Pro Trp Lys Trp Gly Met Lys Tyr Phe Trp Asp Thr			
370	375	380	

Leu Leu Asp Ala Asp Leu Glu Cys Asp Ile Leu Gly Trp Gln Tyr Ile  
 385 390 395 400

Ser Gly Ser Ile Pro Asp Gly His Glu Leu Asp Arg Leu Asp Asn Pro  
 405 410 415

Ala Leu Gln Gly Ala Lys Tyr Asp Pro Glu Gly Glu Tyr Ile Arg Gln  
 420 425 430

Trp Leu Pro Glu Leu Ala Arg Leu Pro Thr Glu Trp Ile His His Pro  
 435 440 445

Trp Asp Ala Pro Leu Thr Val Leu Lys Ala Ser Gly Val Glu Leu Gly  
 450 455 460

Thr Asn Tyr Ala Lys Pro Ile Val Asp Ile Asp Thr Ala Arg Glu Leu  
 465 470 475 480

Leu Ala Lys Ala Ile Ser Arg Thr Arg Glu Ala Gln Ile Met Ile Gly  
 485 490 495

Ala Ala Pro Ala Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Gly  
 500 505 510

Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala Asp Tyr Leu Glu  
 515 520 525

Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser Met Leu Pro Gly  
 530 535 540

Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala Asp Tyr Leu Glu  
 545 550 555 560

Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser Met Leu Pro Gly  
 565 570 575

Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala Asp Tyr Leu Glu  
 580 585 590

Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser Met Leu Pro Gly  
 595 600 605

Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala Asp Tyr Leu Glu  
 610 615 620

Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser Met Leu Pro Ser

625                    630                    635                    640  
 Arg Ser Arg Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly  
                          645                    650                    655  
  
 Asp Val Glu Glu Asn Pro Gly Pro Ile Glu Lys Ser Phe Val Ile Thr  
                          660                    665                    670  
 Asp Pro Arg Leu Pro Asp Tyr Pro Ile Ile Phe Ala Ser Asp Gly Phe  
                          675                    680                    685  
 Leu Glu Leu Thr Glu Tyr Ser Arg Glu Glu Ile Met Gly Arg Asn Ala  
                          690                    695                    700  
 Arg Phe Leu Gln Gly Pro Glu Thr Asp Gln Ala Thr Val Gln Lys Ile  
 705                    710                    715                    720  
  
 Arg Asp Ala Ile Arg Asp Gln Arg Glu Thr Thr Val Gln Leu Ile Asn  
                          725                    730                    735  
 Tyr Thr Lys Ser Gly Lys Lys Phe Trp Asn Leu Leu His Leu Gln Pro  
                          740                    745                    750  
 Val Arg Asp Arg Lys Gly Gly Leu Gln Tyr Phe Ile Gly Val Gln Leu  
                          755                    760                    765  
 Val Gly Ser Asp His Val  
                          770  
 <210> 143  
 <211> 1356  
 <212> PRT  
 <213> Artificial Sequence  
  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
                          polypeptide"  
 <400> 143  
 Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala Phe  
 1                    5                    10                    15  
 Ser Ala Asp Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser Leu  
                          20                    25                    30  
 Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His His

Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu Arg  
 50 55 60  
 Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala Ala  
 65 70 75 80  
 Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln Pro  
 85 90 95  
 Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr  
 100 105 110  
  
 Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val  
 115 120 125  
 Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His  
 130 135 140  
 Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val  
 145 150 155 160  
 Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala  
 165 170 175  
  
 Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala  
 180 185 190  
 Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp  
 195 200 205  
 Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val  
 210 215 220  
 Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn  
 225 230 235 240  
  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 245 250 255  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 260 265 270  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly  
 275 280 285

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 290 295 300

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 305 310 315 320

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 325 330 335

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 340 345 350

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 355 360 365

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 370 375 380

Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 385 390 395 400

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 405 410 415

Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val  
 420 425 430

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 435 440 445

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 450 455 460

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 465 470 475 480

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala  
 485 490 495

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 500 505 510

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 515 520 525

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala

530                      535                      540  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 545                      550                      555                      560  
  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                             565                      570                      575  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
                             580                      585                      590  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
                             595                      600                      605  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                             610                      615                      620  
  
 Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 625                      630                      635                      640  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
                             645                      650                      655  
 Ile Ala Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala  
                             660                      665                      670  
 Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His  
                             675                      680                      685  
  
 Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val  
                             690                      695                      700  
 Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg  
 705                      710                      715                      720  
 Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala Asp His Ala Gln Val  
                             725                      730                      735  
 Val Arg Val Leu Gly Phe Phe Gln Cys His Ser His Pro Ala Gln Ala  
                             740                      745                      750  
  
 Phe Asp Asp Ala Met Thr Gln Phe Gly Met Ser Arg His Gly Leu Leu  
                             755                      760                      765  
 Gln Leu Phe Arg Arg Val Gly Val Thr Glu Leu Glu Ala Arg Ser Gly  
                             770                      775                      780

Thr Leu Pro Pro Ala Ser Gln Arg Trp Asp Arg Ile Leu Gln Ala Ser  
785                      790                      795                      800

Gly Met Lys Arg Ala Lys Pro Ser Pro Thr Ser Thr Gln Thr Pro Asp  
                                 805                      810                      815

Gln Ala Ser Leu His Ala Phe Ala Asp Ser Leu Glu Arg Asp Leu Asp  
                                 820                      825                      830

Ala Pro Ser Pro Met His Glu Gly Asp Gln Thr Arg Ala Ser Ala Ser  
                                 835                      840                      845

Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Lys Met Asp Lys Lys Thr  
                                 850                      855                      860

Ile Val Trp Phe Arg Arg Asp Leu Arg Ile Glu Asp Asn Pro Ala Leu  
865                      870                      875                      880

Ala Ala Ala Ala His Glu Gly Ser Val Phe Pro Val Phe Ile Trp Cys  
                                 885                      890                      895

Pro Glu Glu Glu Gly Gln Phe Tyr Pro Gly Arg Ala Ser Arg Trp Trp  
                                 900                      905                      910

Met Lys Gln Ser Leu Ala His Leu Ser Gln Ser Leu Lys Ala Leu Gly  
                                 915                      920                      925

Ser Asp Leu Thr Leu Ile Lys Thr His Asn Thr Ile Ser Ala Ile Leu  
                                 930                      935                      940

Asp Cys Ile Arg Val Thr Gly Ala Thr Lys Val Val Phe Asn His Leu  
945                      950                      955                      960

Tyr Asp Pro Val Ser Leu Val Arg Asp His Thr Val Lys Glu Lys Leu  
                                 965                      970                      975

Val Glu Arg Gly Ile Ser Val Gln Ser Tyr Asn Gly Asp Leu Leu Tyr  
                                 980                      985                      990

Glu Pro Trp Glu Ile Tyr Cys Glu Lys Gly Lys Pro Phe Thr Ser Phe  
                                 995                      1000                      1005

Asn Ser Tyr Trp Lys Lys Cys Leu Asp Met Ser Ile Glu Ser Val  
                                 1010                      1015                      1020

Met Leu Pro Pro Pro Trp Arg Leu Met Pro Ile Thr Ala Ala Ala

1025	1030	1035
Glu Ala Ile Trp Ala Cys Ser	Ile Glu Glu Leu Gly	Leu Glu Asn
1040	1045	1050
Glu Ala Glu Lys Pro Ser Asn	Ala Leu Leu Thr Arg	Ala Trp Ser
1055	1060	1065
Pro Gly Trp Ser Asn Ala Asp	Lys Leu Leu Asn Glu	Phe Ile Glu
1070	1075	1080
Lys Gln Leu Ile Asp Tyr Ala	Lys Asn Ser Lys Lys	Val Val Gly
1085	1090	1095
Asn Ser Thr Ser Leu Leu Ser	Pro Tyr Leu His Phe	Gly Glu Ile
1100	1105	1110
Ser Val Arg His Val Phe Gln	Cys Ala Arg Met Lys	Gln Ile Ile
1115	1120	1125
Trp Ala Arg Asp Lys Asn Ser	Glu Gly Glu Glu Ser	Ala Asp Leu
1130	1135	1140
Phe Leu Arg Gly Ile Gly Leu	Arg Glu Tyr Ser Arg	Tyr Ile Cys
1145	1150	1155
Phe Asn Phe Pro Phe Thr His	Glu Gln Ser Leu Leu	Ser His Leu
1160	1165	1170
Arg Phe Phe Pro Trp Asp Ala	Asp Val Asp Lys Phe	Lys Ala Trp
1175	1180	1185
Arg Gln Gly Arg Thr Gly Tyr	Pro Leu Val Asp Ala	Gly Met Arg
1190	1195	1200
Glu Leu Trp Ala Thr Gly Trp	Met His Asn Arg Ile	Arg Val Ile
1205	1210	1215
Val Ser Ser Phe Ala Val Lys	Phe Leu Leu Leu Pro	Trp Lys Trp
1220	1225	1230
Gly Met Lys Tyr Phe Trp Asp	Thr Leu Leu Asp Ala	Asp Leu Glu
1235	1240	1245
Cys Asp Ile Leu Gly Trp Gln	Tyr Ile Ser Gly Ser	Ile Pro Asp
1250	1255	1260

Gly His Glu Leu Asp Arg Leu Asp Asn Pro Ala Leu Gln Gly Ala  
 1265 1270 1275  
 Lys Tyr Asp Pro Glu Gly Glu Tyr Ile Arg Gln Trp Leu Pro Glu  
 1280 1285 1290  
 Leu Ala Arg Leu Pro Thr Glu Trp Ile His His Pro Trp Asp Ala  
 1295 1300 1305

Pro Leu Thr Val Leu Lys Ala Ser Gly Val Glu Leu Gly Thr Asn  
 1310 1315 1320  
 Tyr Ala Lys Pro Ile Val Asp Ile Asp Thr Ala Arg Glu Leu Leu  
 1325 1330 1335  
 Ala Lys Ala Ile Ser Arg Thr Arg Glu Ala Gln Ile Met Ile Gly  
 1340 1345 1350

Ala Ala Pro  
 1355

<210> 144

<211> 1001

<212> PRT

<213> Artificial Sequence

<220><221> source

<223

> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 144

Met Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Pro Lys Lys Lys Arg  
 1 5 10 15  
 Lys Val Glu Ala Ser Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln  
 20 25 30  
 Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His  
 35 40 45  
 His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile Val Ala  
 50 55 60  
 Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln  
 65 70 75 80

Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly  
                   85                  90                  95  
 Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr  
                   100                  105                  110  
 Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln  
                   115                  120                  125  
 Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val  
                   130                  135                  140  
 His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro  
 145                  150                  155                  160  
 Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu  
                   165                  170                  175  
 Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu  
                   180                  185                  190  
 Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln  
                   195                  200                  205  
 Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
                   210                  215                  220  
 Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly  
 225                  230                  235                  240  
 Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln  
                   245                  250                  255  
 Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly  
                   260                  265                  270  
 Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu  
                   275                  280                  285  
 Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser  
                   290                  295                  300  
 Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro  
                   305                  310                  315                  320  
 Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile

325                      330                      335  
 Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu  
                          340                      345                      350  
 Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val  
                          355                      360                      365  
 Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln  
  
                          370                      375                      380  
 Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln  
 385                      390                      395                      400  
 Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr  
                          405                      410                      415  
 Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro  
                          420                      425                      430  
 Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu  
  
                          435                      440                      445  
 Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu  
                          450                      455                      460  
 Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln  
 465                      470                      475                      480  
 Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
                          485                      490                      495  
 Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly  
  
                          500                      505                      510  
 Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln  
                          515                      520                      525  
 Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile  
                          530                      535                      540  
 Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu  
 545                      550                      555                      560  
 Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser  
  
                          565                      570                      575

Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro  
                   580                                  585                                  590  
 Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile  
                   595                                  600                                  605  
 Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu  
                   610                                  615                                  620  
 Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val  
  
 625                                  630                                  635                                  640  
 Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln  
                   645                                  650                                  655  
 Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln  
                   660                                  665                                  670  
 Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr  
                   675                                  680                                  685  
 Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro  
  
 690                                  695                                  700  
 Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu  
 705                                  710                                  715                                  720  
 Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu  
                   725                                  730                                  735  
 Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln  
                   740                                  745                                  750  
 Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
  
 755                                  760                                  765  
 Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly  
                   770                                  775                                  780  
 Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro  
 785                                  790                                  795                                  800  
 Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu  
                   805                                  810                                  815  
 Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala

820 825 830  
 Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser  
 835 840 845  
 His Arg Val Ala Asp His Ala Gln Val Val Arg Val Leu Gly Phe Phe  
 850 855 860  
 Gln Cys His Ser His Pro Ala Gln Ala Phe Asp Asp Ala Met Thr Gln  
 865 870 875 880  
 Phe Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe Arg Arg Val Gly

885 890 895  
 Val Thr Glu Leu Glu Ala Arg Ser Gly Thr Leu Pro Pro Ala Ser Gln  
 900 905 910  
 Arg Trp Asp Arg Ile Leu Gln Ala Ser Gly Met Lys Arg Ala Lys Pro  
 915 920 925  
 Ser Pro Thr Ser Thr Gln Thr Pro Asp Gln Ala Ser Leu His Ala Phe  
 930 935 940  
 Ala Asp Ser Leu Glu Arg Asp Leu Asp Ala Pro Ser Pro Met His Glu

945 950 955 960  
 Gly Asp Gln Thr Arg Ala Ser Ala Ser Gly Ser Gly Met Asn Ile Gln  
 965 970 975  
 Met Leu Leu Glu Ala Ala Asp Tyr Leu Glu Arg Arg Glu Arg Glu Ala  
 980 985 990  
 Glu His Gly Tyr Ala Ser Met Leu Pro  
 995 1000

<210> 145

<211> 1099

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 145

Met Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Pro Lys Lys Lys Arg

1                    5                    10                    15  
 Lys Val Glu Ala Ser Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln  
                          20                    25                    30  
 Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His  
                          35                    40                    45  
 His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile Val Ala  
                          50                    55                    60  
  
 Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln  
 65                    70                    75                    80  
 Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly  
                          85                    90                    95  
 Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr  
                          100                    105                    110  
 Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln  
                          115                    120                    125  
  
 Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val  
                          130                    135                    140  
 His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro  
 145                    150                    155                    160  
 Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu  
                          165                    170                    175  
 Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu  
                          180                    185                    190  
  
 Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln  
                          195                    200                    205  
 Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
                          210                    215                    220  
 Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly  
 225                    230                    235                    240  
 Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln  
                          245                    250                    255

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly  
 260 265 270

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu  
 275 280 285

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser  
 290 295 300

Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro  
 305 310 315 320

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile  
 325 330 335

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu  
 340 345 350

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val  
 355 360 365

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln  
 370 375 380

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln  
 385 390 395 400

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr  
 405 410 415

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro  
 420 425 430

Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu  
 435 440 445

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu  
 450 455 460

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln  
 465 470 475 480

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
 485 490 495

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly

	500		505		510										
Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln
	515		520		525										
Ala	His	Gly	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile
	530		535		540										
Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu
545			550		555										560
Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser
	565		570		575										
Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro
	580		585		590										
Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Glu	Gln	Val	Val	Ala	Ile
	595		600		605										
Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu
	610		615		620										
Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Glu	Gln	Val	Val
625			630		635										640
Ala	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln
	645		650		655										
Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Glu	Gln
	660		665		670										
Val	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr
	675		680		685										
Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro
	690		695		700										
Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu
705			710		715										720
Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu
	725		730		735										
Thr	Pro	Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln
	740		745		750										

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
 755 760 765

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly  
 770 775 780

Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro  
 785 790 795 800

Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu  
 805 810 815

Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala  
 820 825 830

Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser  
 835 840 845

His Arg Val Ala Asp His Ala Gln Val Val Arg Val Leu Gly Phe Phe  
 850 855 860

Gln Cys His Ser His Pro Ala Gln Ala Phe Asp Asp Ala Met Thr Gln  
 865 870 875 880

Phe Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe Arg Arg Val Gly  
 885 890 895

Val Thr Glu Leu Glu Ala Arg Ser Gly Thr Leu Pro Pro Ala Ser Gln  
 900 905 910

Arg Trp Asp Arg Ile Leu Gln Ala Ser Gly Met Lys Arg Ala Lys Pro  
 915 920 925

Ser Pro Thr Ser Thr Gln Thr Pro Asp Gln Ala Ser Leu His Ala Phe  
 930 935 940

Ala Asp Ser Leu Glu Arg Asp Leu Asp Ala Pro Ser Pro Met His Glu  
 945 950 955 960

Gly Asp Gln Thr Arg Ala Ser Ala Ser Gly Ser Gly Met Asn Ile Gln  
 965 970 975

Met Leu Leu Glu Ala Ala Asp Tyr Leu Glu Arg Arg Glu Arg Glu Ala  
 980 985 990

Glu His Gly Tyr Ala Ser Met Leu Pro Gly Ser Gly Met Asn Ile Gln

995 1000 1005  
 Met Leu Leu Glu Ala Ala Asp Tyr Leu Glu Arg Arg Glu Arg Glu  
 1010 1015 1020

Ala Glu His Gly Tyr Ala Ser Met Leu Pro Gly Ser Gly Met Asn  
 1025 1030 1035

Ile Gln Met Leu Leu Glu Ala Ala Asp Tyr Leu Glu Arg Arg Glu  
 1040 1045 1050

Arg Glu Ala Glu His Gly Tyr Ala Ser Met Leu Pro Gly Ser Gly  
 1055 1060 1065

Met Asn Ile Gln Met Leu Leu Glu Ala Ala Asp Tyr Leu Glu Arg  
 1070 1075 1080

Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser Met Leu Pro Ser  
 1085 1090 1095

Arg

<210> 146

<211> 1187

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 146

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1 5 10 15  
 Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu

20 25 30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
 35 40 45

Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
 50 55 60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr

65                               70                               75                               80  
Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala

                                     85                               90                               95  
Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln

                                     100                               105                               110  
Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile

                                     115                               120                               125  
Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val

                                     130                               135                               140  
Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro

145                               150                               155                               160  
Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala

                                     165                               170                               175  
Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp

                                     180                               185                               190  
Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu

                                     195                               200                               205  
Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala

210                               215                               220  
Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn

225                               230                               235                               240  
Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala

                                     245                               250                               255  
Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg

                                     260                               265                               270  
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val

275                               280                               285  
Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val

290                               295                               300  
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu

305                               310                               315                               320

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
 325 330 335

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 340 345 350

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 355 360 365

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 370 375 380

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385 390 395 400

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 405 410 415

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
 420 425 430

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 435 440 445

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 450 455 460

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465 470 475 480

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val

565	570	575	
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu			
580	585	590	
Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu			
595	600	605	
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr			
610	615	620	
Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala			
625	630	635	640
Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly			
645	650	655	
Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys			
660	665	670	
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala			
675	680	685	
His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly			
690	695	700	
Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys			
705	710	715	720
Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn			
725	730	735	
Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val			
740	745	750	
Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala			
755	760	765	
Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu			
770	775	780	
Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val			
785	790	795	800
Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys			
805	810	815	

Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
 820 825 830

Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Gly Gly Gly Gly Ser  
 835 840 845

Gly Gly Gly Gly Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn  
 850 855 860

Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr  
 865 870 875 880

Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser  
 885 890 895

Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly  
 900 905 910

Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser  
 915 920 925

Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys  
 930 935 940

Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys  
 945 950 955 960

Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu  
 965 970 975

Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys  
 980 985 990

Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp  
 995 1000 1005

Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His  
 1010 1015 1020

Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala  
 1025 1030 1035

Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu  
 1040 1045 1050

Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly

1055                      1060                      1065  
 Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln

1070                      1075                      1080  
 Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro

1085                      1090                      1095  
 Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr

1100                      1105                      1110  
 Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr

1115                      1120                      1125  
 Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val Asn

1130                      1135                      1140  
 Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr Ser

1145                      1150                      1155  
 Ser Asp Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser Met

1160                      1165                      1170  
 Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val

1175                      1180                      1185

<210> 147

<211> 1188

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 147

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser

1                      5                      10                      15

Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu

20                      25                      30

Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser

35                      40                      45

Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp

50                        55                        60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65                                70                                75                                80  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
  
                                   85                                    90                                    95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
                                   100                                    105                                    110  
 Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
                                   115                                    120                                    125  
 Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
 130                                    135                                    140  
 Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
  
 145                                    150                                    155                                    160  
 Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
                                   165                                    170                                    175  
 Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
                                   180                                    185                                    190  
 Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu  
                                   195                                    200                                    205  
 Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
  
                                   210                                    215                                    220  
 Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
 225                                    230                                    235                                    240  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
                                   245                                    250                                    255  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
                                   260                                    265                                    270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
  
                                   275                                    280                                    285  
 Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
                                   290                                    295                                    300

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305                      310                      315                      320  
 Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
                                  325                      330                      335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
                                  340                      345                      350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
                                  355                      360                      365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                                  370                      375                      380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385                      390                      395                      400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                                  405                      410                      415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
                                  420                      425                      430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                                  435                      440                      445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
                                  450                      455                      460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
                                  465                      470                      475                      480  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                                  485                      490                      495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
                                  500                      505                      510  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
                                  515                      520                      525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
                                  530                      535                      540  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val

545                    550                    555                    560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
                          565                    570                    575  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
                          580                    585                    590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
  
                          595                    600                    605  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
                          610                    615                    620  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625                    630                    635                    640  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                          645                    650                    655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
  
                          660                    665                    670  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                          675                    680                    685  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
                          690                    695                    700  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705                    710                    715                    720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
  
                          725                    730                    735  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
                          740                    745                    750  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                          755                    760                    765  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
                          770                    775                    780  
 Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
  
 785                    790                    795                    800

Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805 810 815

Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
 820 825 830

Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Leu Tyr Pro Glu Arg  
 835 840 845

Leu Arg Arg Ile Leu Thr Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu  
 850 855 860

Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys  
 865 870 875 880

Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp  
 885 890 895

Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala  
 900 905 910

Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp  
 915 920 925

Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe  
 930 935 940

Lys Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala  
 945 950 955 960

Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu  
 965 970 975

Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile  
 980 985 990

Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn  
 995 1000 1005

Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile  
 1010 1015 1020

His Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile  
 1025 1030 1035

Ala Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe

1040                      1045                      1050  
 Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala  
 1055                      1060                      1065  
 Gly Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg  
 1070                      1075                      1080  
 Gln Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg  
 1085                      1090                      1095  
 Pro Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser

1100                      1105                      1110  
 Thr Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly  
 1115                      1120                      1125  
 Tyr Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val  
 1130                      1135                      1140  
 Asn Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr  
 1145                      1150                      1155  
 Ser Ser Asp Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser

1160                      1165                      1170  
 Met Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
 1175                      1180                      1185

<210> 148

<211> 1188

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 148

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1                      5                      10                      15  
 Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu

20                      25                      30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser

35                                  40                                  45  
 Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
 50                                  55                                  60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65                                  70                                  75                                  80  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
 85                                  90                                  95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
 100                                  105                                  110  
 Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
 115                                  120                                  125  
 Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
 130                                  135                                  140  
 Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
 145                                  150                                  155                                  160  
 Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
 165                                  170                                  175  
 Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
 180                                  185                                  190  
 Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu  
 195                                  200                                  205  
 Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
 210                                  215                                  220  
 Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
 225                                  230                                  235                                  240  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
 245                                  250                                  255  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 260                                  265                                  270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 275                                  280                                  285

Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
 290 295 300

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305 310 315 320

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
 325 330 335

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 340 345 350

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 355 360 365

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 370 375 380

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385 390 395 400

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 405 410 415

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
 420 425 430

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 435 440 445

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 450 455 460

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465 470 475 480

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg

530                      535                      540  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545                      550                      555                      560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
                          565                      570                      575  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
                          580                      585                      590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
  
                          595                      600                      605  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
                          610                      615                      620  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625                      630                      635                      640  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                          645                      650                      655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
  
                          660                      665                      670  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                          675                      680                      685  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
                          690                      695                      700  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705                      710                      715                      720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
  
                          725                      730                      735  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
                          740                      745                      750  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                          755                      760                      765  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
                          770                      775                      780

Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val

785                      790                      795                      800

Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys

                            805                      810                      815

Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile

                            820                      825                      830

Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Leu Gln Leu Pro Pro

                            835                      840                      845

Leu Glu Arg Leu Thr Leu Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu

850                      855                      860

Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys

865                      870                      875                      880

Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp

                            885                      890                      895

Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala

                            900                      905                      910

Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp

915                      920                      925

Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe

930                      935                      940

Lys Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala

945                      950                      955                      960

Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu

                            965                      970                      975

Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile

980                      985                      990

Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn

995                      1000                      1005

Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile

1010                      1015                      1020

His Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile

1025                      1030                      1035  
 Ala Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe

1040                      1045                      1050  
 Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala

1055                      1060                      1065  
 Gly Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg

1070                      1075                      1080  
 Gln Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg

1085                      1090                      1095  
 Pro Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser

1100                      1105                      1110  
 Thr Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly

1115                      1120                      1125  
 Tyr Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val

1130                      1135                      1140  
 Asn Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr

1145                      1150                      1155  
 Ser Ser Asp Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser

1160                      1165                      1170  
 Met Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val

1175                      1180                      1185

<210> 149

<211> 1185

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 149

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser

1                      5                      10                      15

Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
 20 25 30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
 35 40 45  
 Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
 50 55 60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65 70 75 80  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
 85 90 95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
 100 105 110  
 Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
 115 120 125  
 Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
 130 135 140  
 Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
 145 150 155 160  
 Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
 165 170 175  
 Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
 180 185 190  
 Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu  
 195 200 205  
 Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
 210 215 220  
 Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
 225 230 235 240  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
 245 250 255  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg

260 265 270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val

275 280 285  
 Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val

290 295 300  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu

305 310 315 320  
 Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu

325 330 335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr

340 345 350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala

355 360 365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

370 375 380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys

385 390 395 400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala

405 410 415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly

420 425 430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys

435 440 445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn

450 455 460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val

465 470 475 480  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala

485 490 495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu

500 505 510

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 725 730 735

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 740 745 750

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala

755                      760                      765  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
 770                      775                      780  
 Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
  
 785                      790                      795                      800  
 Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805                      810                      815  
 Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
 820                      825                      830  
 Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Leu Asp Leu Ala Ser  
 835                      840                      845  
 Leu Ile Leu Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn Phe Pro  
  
 850                      855                      860  
 Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr Leu Asn  
 865                      870                      875                      880  
 Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser Ser Met  
 885                      890                      895  
 Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly Leu Asn  
 900                      905                      910  
 Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser Arg Leu  
  
 915                      920                      925  
 Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys Ala Ala  
 930                      935                      940  
 Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys Lys Met  
 945                      950                      955                      960  
 Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu Lys Ser  
 965                      970                      975  
 Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys Lys Met  
  
 980                      985                      990  
 Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser  
 995                      1000                      1005

Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His Val Arg  
 1010 1015 1020  
 Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala Glu Arg  
 1025 1030 1035  
 Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu Gln Asp  
 1040 1045 1050  
 Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly Met Leu  
 1055 1060 1065  
 Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln Ile Glu  
 1070 1075 1080  
 Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro Asp Phe  
 1085 1090 1095  
 Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr Pro Met  
 1100 1105 1110  
 Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr Ser His  
 1115 1120 1125  
 Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val Asn Ser Gly  
 1130 1135 1140  
 Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr Ser Ser Asp  
 1145 1150 1155  
 Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser Met Trp Asp  
 1160 1165 1170  
 Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
 1175 1180 1185  
 <210> 150  
 <211> 1187  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 150

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1                    5                    10                    15  
 Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
                          20                    25                    30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
                          35                    40                    45  
 Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
                          50                    55                    60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65                    70                    75                    80  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
                          85                    90                    95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
                          100                    105                    110  
 Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
                          115                    120                    125  
 Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
                          130                    135                    140  
 Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
                          145                    150                    155                    160  
 Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
                          165                    170                    175  
 Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
                          180                    185                    190  
 Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu  
                          195                    200                    205  
 Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
                          210                    215                    220  
 Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
 225                    230                    235                    240  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala

245 250 255  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 260 265 270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
  
 275 280 285  
 Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
 290 295 300  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305 310 315 320  
 Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
 325 330 335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
  
 340 345 350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 355 360 365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 370 375 380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385 390 395 400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
  
 405 410 415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
 420 425 430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 435 440 445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 450 455 460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
  
 465 470 475 480  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 725 730 735

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val

740                                 745                                 750  
Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
755                                 760                                 765  
Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
770                                 775                                 780  
Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
785                                 790                                 795                                 800  
Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
805                                 810                                 815  
Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
820                                 825                                 830  
Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Leu Gln Lys Lys Leu  
835                                 840                                 845  
Glu Glu Leu Glu Leu Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn  
850                                 855                                 860  
Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr  
865                                 870                                 875                                 880  
Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser  
885                                 890                                 895  
Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly  
900                                 905                                 910  
Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser  
915                                 920                                 925  
Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys  
930                                 935                                 940  
Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys  
945                                 950                                 955                                 960  
Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu  
965                                 970                                 975  
Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys  
980                                 985                                 990

Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp  
           995                          1000                          1005  
 Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His  
           1010                          1015                          1020  
 Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala  
           1025                          1030                          1035  
 Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu  
  
           1040                          1045                          1050  
 Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly  
           1055                          1060                          1065  
 Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln  
           1070                          1075                          1080  
 Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro  
           1085                          1090                          1095  
 Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr  
  
           1100                          1105                          1110  
 Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr  
           1115                          1120                          1125  
 Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val Asn  
           1130                          1135                          1140  
 Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr Ser  
           1145                          1150                          1155  
 Ser Asp Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser Met  
  
           1160                          1165                          1170  
 Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
           1175                          1180                          1185  
 <210> 151  
 <211> 1159  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 151

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser

1                    5                    10                    15

Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu

20                    25                    30

Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser

35                    40                    45

Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp

50                    55                    60

Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr

65                    70                    75                    80

Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala

85                    90                    95

Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln

100                    105                    110

Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile

115                    120                    125

Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val

130                    135                    140

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro

145                    150                    155                    160

Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala

165                    170                    175

Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp

180                    185                    190

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu

195                    200                    205

Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala

210                    215                    220

Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn

225                      230                      235                      240  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
                               245                      250                      255  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
                               260                      265                      270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
  
                               275                      280                      285  
 Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
                               290                      295                      300  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305                      310                      315                      320  
 Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
                               325                      330                      335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
  
                               340                      345                      350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
                               355                      360                      365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                               370                      375                      380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385                      390                      395                      400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
  
                               405                      410                      415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
                               420                      425                      430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                               435                      440                      445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
                               450                      455                      460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
  
 465                      470                      475                      480

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn

725	730	735	
Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val			
740	745	750	
Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala			
755	760	765	
Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu			
770	775	780	
Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val			
785	790	795	800
Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys			
805	810	815	
Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile			
820	825	830	
Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Gly Gly Gly Gly Ser			
835	840	845	
Gly Gly Gly Gly Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn			
850	855	860	
Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr			
865	870	875	880
Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser			
885	890	895	
Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly			
900	905	910	
Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser			
915	920	925	
Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys			
930	935	940	
Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys			
945	950	955	960
Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu			
965	970	975	

Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys

980 985 990

Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp

995 1000 1005

Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His

1010 1015 1020

Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala

1025 1030 1035

Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu

1040 1045 1050

Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly

1055 1060 1065

Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln

1070 1075 1080

Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro

1085 1090 1095

Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr

1100 1105 1110

Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr

1115 1120 1125

Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val Asn

1130 1135 1140

Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr Ser

1145 1150 1155

Ser

<210> 152

<211> 1160

<212> PRT

<213

> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 152

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
1                   5                   10                   15

Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
                  20                   25                   30

Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
                  35                   40                   45

Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
50                   55                   60

Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
65                   70                   75                   80

Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
                  85                   90                   95

Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
                  100                   105                   110

Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
115                   120                   125

Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
130                   135                   140

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
145                   150                   155                   160

Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
                  165                   170                   175

Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
180                   185                   190

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu  
195                   200                   205

Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
210                   215                   220

Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn

225                      230                      235                      240  
  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
                                          245                                              250                                              255  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
                                          260                                              265                                              270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
                                          275                                              280                                              285  
 Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
                                          290                                              295                                              300  
  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305                                              310                                              315                                              320  
 Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
                                          325                                              330                                              335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
                                          340                                              345                                              350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
                                          355                                              360                                              365  
  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                                          370                                              375                                              380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385                                              390                                              395                                              400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                                          405                                              410                                              415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
                                          420                                              425                                              430  
  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                                          435                                              440                                              445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
                                          450                                              455                                              460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465                                              470                                              475                                              480

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn

725 730 735  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 740 745 750

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 755 760 765

Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
 770 775 780

Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
 785 790 795 800

Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805 810 815

Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
 820 825 830

Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Leu Tyr Pro Glu Arg  
 835 840 845

Leu Arg Arg Ile Leu Thr Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu  
 850 855 860

Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys  
 865 870 875 880

Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp  
 885 890 895

Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala  
 900 905 910

Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp  
 915 920 925

Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe  
 930 935 940

Lys Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala  
 945 950 955 960

Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu  
 965 970 975

Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile  
                   980                                  985                                  990  
 Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn  
                   995                                  1000                                  1005  
  
 Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile  
                   1010                                  1015                                  1020  
 His Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile  
                   1025                                  1030                                  1035  
 Ala Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe  
                   1040                                  1045                                  1050  
 Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala  
                   1055                                  1060                                  1065  
  
 Gly Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg  
                   1070                                  1075                                  1080  
 Gln Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg  
                   1085                                  1090                                  1095  
 Pro Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser  
                   1100                                  1105                                  1110  
 Thr Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly  
                   1115                                  1120                                  1125  
  
 Tyr Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val  
                   1130                                  1135                                  1140  
 Asn Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr  
                   1145                                  1150                                  1155  
 Ser Ser  
                   1160  
 <210> 153  
 <211> 1160  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 153

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser

1                    5                    10                    15  
 Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu

                  20                    25                    30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser

                  35                    40                    45  
 Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp

                  50                    55                    60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr

65                    70                    75                    80  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala

                  85                    90                    95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln

                  100                    105                    110  
 Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile

                  115                    120                    125  
 Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val

                  130                    135                    140  
 Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro

                  145                    150                    155                    160  
 Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala

                  165                    170                    175  
 Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp

                  180                    185                    190  
 Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu

                  195                    200                    205  
 Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala

                  210                    215                    220  
 Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn

225                    230                    235                    240  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
                                  245                    250                    255  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
  
                                  260                    265                    270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
                                  275                    280                    285  
 Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
                                  290                    295                    300  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305                    310                    315                    320  
 Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
  
                                  325                    330                    335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
                                  340                    345                    350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
                                  355                    360                    365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                                  370                    375                    380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
  
 385                    390                    395                    400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                                  405                    410                    415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
                                  420                    425                    430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                                  435                    440                    445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
  
                                  450                    455                    460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465                    470                    475                    480

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn

725                                  730                                  735

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val

740                                  745                                  750

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala

755                                  760                                  765

Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu

770                                  775                                  780

Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val

785                                  790                                  795                                  800

Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys

805                                  810                                  815

Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile

820                                  825                                  830

Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Ser Pro Lys Lys Lys

835                                  840                                  845

Arg Lys Val Glu Ala Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu

850                                  855                                  860

Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys

865                                  870                                  875                                  880

Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp

885                                  890                                  895

Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala

900                                  905                                  910

Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp

915                                  920                                  925

Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe

930                                  935                                  940

Lys Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala

945                                  950                                  955                                  960

Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu

965                                  970                                  975

Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile  
                   980                                  985                                  990  
 Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn  
                   995                                  1000                                  1005  
 Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile  
                   1010                                  1015                                  1020  
 His Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile  
                   1025                                  1030                                  1035  
 Ala Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe  
                   1040                                  1045                                  1050  
 Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala  
                   1055                                  1060                                  1065  
 Gly Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg  
                   1070                                  1075                                  1080  
 Gln Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg  
                   1085                                  1090                                  1095  
 Pro Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser  
                   1100                                  1105                                  1110  
 Thr Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly  
                   1115                                  1120                                  1125  
 Tyr Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val  
                   1130                                  1135                                  1140  
 Asn Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr  
                   1145                                  1150                                  1155  
 Ser Ser  
                   1160  
 <210> 154  
 <211> 1187  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 154

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser

1                    5                    10                    15

Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu

                  20                    25                    30

Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser

                  35                    40                    45

Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp

                  50                    55                    60

Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr

65                    70                    75                    80

Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala

                  85                    90                    95

Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln

                  100                    105                    110

Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile

                  115                    120                    125

Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val

                  130                    135                    140

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro

145                    150                    155                    160

Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala

                  165                    170                    175

Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp

                  180                    185                    190

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu

                  195                    200                    205

Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala

                  210                    215                    220

Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn

225                      230                      235                      240  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
                                          245                      250                      255  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
                                          260                      265                      270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
                                          275                      280                      285  
  
 Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
                                          290                      295                      300  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305                      310                      315                      320  
 Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
                                          325                      330                      335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
                                          340                      345                      350  
  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
                                          355                      360                      365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                                          370                      375                      380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385                      390                      395                      400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                                          405                      410                      415  
  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
                                          420                      425                      430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                                          435                      440                      445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
                                          450                      455                      460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465                      470                      475                      480

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540  
  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605  
  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670  
  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn

	725	730	735
Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val			
	740	745	750
Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala			
	755	760	765
Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu			
	770	775	780
Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val			
785	790	795	800
Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys			
	805	810	815
Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile			
	820	825	830
Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Gly Gly Gly Gly Ser			
	835	840	845
Gly Gly Gly Gly Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn			
	850	855	860
Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr			
865	870	875	880
Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser			
	885	890	895
Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly			
	900	905	910
Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser			
	915	920	925
Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys			
	930	935	940
Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys			
945	950	955	960
Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu			
	965	970	975

Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys  
 980 985 990

Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp  
 995 1000 1005

Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His  
 1010 1015 1020

Val Arg Ala Arg Ala Ala Gln Ala Thr Asp Ser His Ser Ile Ala  
 1025 1030 1035

Glu Ala Val Ala Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu  
 1040 1045 1050

Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly  
 1055 1060 1065

Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln  
 1070 1075 1080

Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro  
 1085 1090 1095

Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr  
 1100 1105 1110

Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr  
 1115 1120 1125

Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val Asn  
 1130 1135 1140

Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr Ser  
 1145 1150 1155

Ser Asp Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser Met  
 1160 1165 1170

Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
 1175 1180 1185

<210> 155

<211> 1179

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 155

Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala Phe Ser  
1                    5                    10                    15

Ala Asp Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser Leu Phe  
                  20                    25                    30

Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His His Thr  
                  35                    40                    45

Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu Arg Ala  
50                    55                    60

Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala Ala Arg  
65                    70                    75                    80

Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln Pro Ser  
                  85                    90                    95

Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser  
                  100                    105                    110

Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala  
115                    120                    125

Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile  
130                    135                    140

Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys  
145                    150                    155                    160

Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile  
                  165                    170                    175

Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu  
                  180                    185                    190

Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr  
                  195                    200                    205

Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val Glu

210                                  215                                  220  
 Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu  
 225                                  230                                  235                                  240  
 Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln  
                                          245                                  250                                  255  
 Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
                                          260                                  265                                  270  
 Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly  
                                          275                                  280                                  285  
  
 Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln  
                                          290                                  295                                  300  
 Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile  
 305                                  310                                  315                                  320  
 Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu  
                                          325                                  330                                  335  
 Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser  
                                          340                                  345                                  350  
  
 Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro  
                                          355                                  360                                  365  
 Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile  
                                          370                                  375                                  380  
 Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu  
 385                                  390                                  395                                  400  
 Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val  
                                          405                                  410                                  415  
  
 Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln  
                                          420                                  425                                  430  
 Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln  
                                          435                                  440                                  445  
 Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr  
                                          450                                  455                                  460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro  
 465 470 475 480

Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu  
 485 490 495

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu  
 500 505 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln  
 515 520 525

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
 530 535 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly  
 545 550 555 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln  
 565 570 575

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly  
 580 585 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu  
 595 600 605

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser  
 610 615 620

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro  
 625 630 635 640

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile  
 645 650 655

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu  
 660 665 670

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val  
 675 680 685

Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln  
 690 695 700

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln

705                    710                    715                    720  
 Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr  
                           725                    730                    735  
  
 Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro  
                           740                    745                    750  
 Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Arg Pro Ala Leu  
                           755                    760                    765  
 Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala  
                           770                    775                    780  
 Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro  
 785                    790                    795                    800  
  
 Ala Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile  
                           805                    810                    815  
 Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala  
                           820                    825                    830  
 Ala Ser Leu Tyr Pro Glu Arg Leu Arg Arg Ile Leu Thr Asn Gly Ala  
                           835                    840                    845  
 Ile Gly Gly Asp Leu Leu Leu Asn Phe Pro Asp Met Ser Val Leu Glu  
                           850                    855                    860  
  
 Arg Gln Arg Ala His Leu Lys Tyr Leu Asn Pro Thr Phe Asp Ser Pro  
 865                    870                    875                    880  
 Leu Ala Gly Phe Phe Ala Asp Ser Ser Met Ile Thr Gly Gly Glu Met  
                           885                    890                    895  
 Asp Ser Tyr Leu Ser Thr Ala Gly Leu Asn Leu Pro Met Met Tyr Gly  
                           900                    905                    910  
 Glu Thr Thr Val Glu Gly Asp Ser Arg Leu Ser Ile Ser Pro Glu Thr  
                           915                    920                    925  
  
 Thr Leu Gly Thr Gly Asn Phe Lys Ala Ala Lys Phe Asp Thr Glu Thr  
                           930                    935                    940  
 Lys Asp Cys Asn Glu Ala Ala Lys Lys Met Thr Met Asn Arg Asp Asp  
 945                    950                    955                    960

Leu Val Glu Glu Gly Glu Glu Glu Lys Ser Lys Ile Thr Glu Gln Asn

965 970 975

Asn Gly Ser Thr Lys Ser Ile Lys Lys Met Lys His Lys Ala Lys Lys

980 985 990

Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser Lys Val Thr Lys Glu Leu

995 1000 1005

Glu Lys Thr Asp Tyr Ile His Val Arg Ala Arg Ala Ala Gln Ala

1010 1015 1020

Thr Asp Ser His Ser Ile Ala Glu Ala Val Ala Arg Glu Lys Ile

1025 1030 1035

Ser Glu Arg Met Lys Phe Leu Gln Asp Leu Val Pro Gly Cys Asp

1040 1045 1050

Lys Ile Thr Gly Lys Ala Gly Met Leu Asp Glu Ile Ile Asn Tyr

1055 1060 1065

Val Gln Ser Leu Gln Arg Gln Ile Glu Phe Leu Ser Met Lys Leu

1070 1075 1080

Ala Ile Val Asn Pro Arg Pro Asp Phe Asp Met Asp Asp Ile Phe

1085 1090 1095

Ala Lys Glu Val Ala Ser Thr Pro Met Thr Val Val Pro Ser Pro

1100 1105 1110

Glu Met Val Leu Ser Gly Tyr Ser His Glu Met Val His Ser Gly

1115 1120 1125

Tyr Ser Ser Glu Met Val Asn Ser Gly Tyr Leu His Val Asn Pro

1130 1135 1140

Met Gln Gln Val Asn Thr Ser Ser Asp Pro Leu Ser Cys Phe Asn

1145 1150 1155

Asn Gly Glu Ala Pro Ser Met Trp Asp Ser His Val Gln Asn Leu

1160 1165 1170

Tyr Gly Asn Leu Gly Val

1175

<210> 156

<211> 1163

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 156

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1                    5                    10                    15  
 Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
                   20                    25                    30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
                   35                    40                    45  
 Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
                   50                    55                    60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65                    70                    75                    80  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
                   85                    90                    95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
                   100                    105                    110  
 Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
                   115                    120                    125  
 Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
                   130                    135                    140  
 Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
 145                    150                    155                    160  
 Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
                   165                    170                    175  
 Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
                   180                    185                    190  
 Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu

195                                  200                                  205  
 Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
 210                                  215                                  220  
 Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
  
 225                                  230                                  235                                  240  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
 245                                  250                                  255  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 260                                  265                                  270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 275                                  280                                  285  
 Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
  
 290                                  295                                  300  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305                                  310                                  315                                  320  
 Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
 325                                  330                                  335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 340                                  345                                  350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
  
 355                                  360                                  365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 370                                  375                                  380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385                                  390                                  395                                  400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 405                                  410                                  415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
  
 420                                  425                                  430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 435                                  440                                  445

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 450 455 460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465 470 475 480  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly

690 695 700  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 725 730 735  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 740 745 750  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 755 760 765  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
 770 775 780  
 Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
 785 790 795 800  
 Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805 810 815  
 Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
 820 825 830  
 Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Gly Gly Gly Gly Ser  
 835 840 845  
 Gly Gly Gly Gly Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn  
 850 855 860  
 Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr  
 865 870 875 880  
 Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser  
 885 890 895  
 Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly  
 900 905 910  
 Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser  
 915 920 925  
 Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys  
 930 935 940

Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys  
 945                      950                      955                      960

Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu  
                                  965                      970                      975

Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys  
                                  980                      985                      990

Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp  
                                  995                      1000                      1005

Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His  
                                  1010                      1015                      1020

Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala  
                                  1025                      1030                      1035

Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu  
                                  1040                      1045                      1050

Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly  
                                  1055                      1060                      1065

Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gly  
                                  1070                      1075                      1080

Gly Ser Val Ala Ser Thr Pro Met Thr Val Val Pro Ser Pro Glu  
                                  1085                      1090                      1095

Met Val Leu Ser Gly Tyr Ser His Glu Met Val His Ser Gly Tyr  
                                  1100                      1105                      1110

Ser Ser Glu Met Val Asn Ser Gly Tyr Leu His Val Asn Pro Met  
                                  1115                      1120                      1125

Gln Gln Val Asn Thr Ser Ser Asp Pro Leu Ser Cys Phe Asn Asn  
                                  1130                      1135                      1140

Gly Glu Ala Pro Ser Met Trp Asp Ser His Val Gln Asn Leu Tyr  
                                  1145                      1150                      1155

Gly Asn Leu Gly Val  
                                  1160

<210> 157  
 <211> 1277

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 157

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser

1                    5                    10                    15

Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu

20                    25                    30

Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser

35                    40                    45

Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp

50                    55                    60

Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr

65                    70                    75                    80

Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala

85                    90                    95

Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln

100                    105                    110

Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile

115                    120                    125

Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val

130                    135                    140

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro

145                    150                    155                    160

Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala

165                    170                    175

Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp

180                    185                    190

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu

195                                  200                                  205  
 Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
 210                                  215                                  220  
 Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
 225                                  230                                  235                                  240  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
                                         245                                  250                                  255  
  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
                                         260                                  265                                  270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
                                         275                                  280                                  285  
 Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
                                         290                                  295                                  300  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305                                  310                                  315                                  320  
  
 Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
                                         325                                  330                                  335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
                                         340                                  345                                  350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
                                         355                                  360                                  365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                                         370                                  375                                  380  
  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385                                  390                                  395                                  400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                                         405                                  410                                  415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
                                         420                                  425                                  430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                                         435                                  440                                  445

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 450 455 460

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465 470 475 480

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly

690                                  695                                  700  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705                                  710                                  715                                  720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
                                          725                                  730                                  735  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
                                          740                                  745                                  750  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                                          755                                  760                                  765  
  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
                                          770                                  775                                  780  
 Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
 785                                  790                                  795                                  800  
 Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
                                          805                                  810                                  815  
 Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
                                          820                                  825                                  830  
  
 Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Leu Tyr Pro Glu Arg  
                                          835                                  840                                  845  
 Leu Arg Arg Ile Leu Thr Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu  
                                          850                                  855                                  860  
 Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys  
 865                                  870                                  875                                  880  
 Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp  
                                          885                                  890                                  895  
  
 Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala  
                                          900                                  905                                  910  
 Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp  
                                          915                                  920                                  925  
 Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe  
                                          930                                  935                                  940

Lys Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala  
 945                      950                      955                      960

Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu  
                                  965                      970                      975

Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile  
                                  980                      985                      990

Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn  
                                  995                      1000                      1005

Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile  
                                  1010                      1015                      1020

His Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile  
                                  1025                      1030                      1035

Ala Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe  
                                  1040                      1045                      1050

Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala  
                                  1055                      1060                      1065

Gly Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg  
                                  1070                      1075                      1080

Gly Gly Ser Gly Glu Glu Glu Lys Ser Lys Ile Thr Glu Gln Asn  
                                  1085                      1090                      1095

Asn Gly Ser Thr Lys Ser Ile Lys Lys Met Lys His Lys Ala Lys  
                                  1100                      1105                      1110

Lys Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser Lys Val Thr Lys  
                                  1115                      1120                      1125

Glu Leu Glu Lys Thr Asp Tyr Ile His Val Arg Ala Arg Arg Gly  
                                  1130                      1135                      1140

Gln Ala Thr Asp Ser His Ser Ile Ala Glu Arg Val Arg Arg Glu  
                                  1145                      1150                      1155

Lys Ile Ser Glu Arg Met Lys Phe Leu Gln Asp Leu Val Pro Gly  
                                  1160                      1165                      1170

Cys Asp Lys Ile Thr Gly Lys Ala Gly Met Leu Asp Glu Ile Ile

1175                      1180                      1185  
 Asn Tyr Val Gln Ser Leu Gln Arg Gly Gly Ser Val Ala Ser Thr  
 1190                      1195                      1200

Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr  
 1205                      1210                      1215

Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val Asn  
 1220                      1225                      1230

Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr Ser  
 1235                      1240                      1245

Ser Asp Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser Met  
 1250                      1255                      1260

Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
 1265                      1270                      1275

<210> 158

<211> 1158

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 158

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1                      5                      10                      15  
 Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
                     20                      25                      30

Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
                     35                      40                      45

Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
                     50                      55                      60

Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65                      70                      75                      80

Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala

	85	90	95
Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln			
	100	105	110
Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile			
	115	120	125
Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val			
	130	135	140
Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro			
145	150	155	160
Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala			
	165	170	175
Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp			
	180	185	190
Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu			
	195	200	205
Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala			
	210	215	220
Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn			
225	230	235	240
Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala			
	245	250	255
Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg			
	260	265	270
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val			
	275	280	285
Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val			
	290	295	300
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu			
305	310	315	320
Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu			
	325	330	335

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 340 345 350  
  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 355 360 365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 370 375 380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385 390 395 400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 405 410 415  
  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
 420 425 430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 435 440 445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 450 455 460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465 470 475 480  
  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540  
  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu

580 585 590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 725 730 735

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 740 745 750

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 755 760 765

Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
 770 775 780

Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
 785 790 795 800

Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805 810 815

Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
 820 825 830

Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Gly Gly Gly Gly Ser  
 835 840 845

Gly Gly Gly Gly Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn  
 850 855 860

Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr  
 865 870 875 880

Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser  
 885 890 895

Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly  
 900 905 910

Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser  
 915 920 925

Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys  
 930 935 940

Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys  
 945 950 955 960

Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu  
 965 970 975

Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys  
 980 985 990

Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp  
 995 1000 1005

Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His  
 1010 1015 1020

Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala  
 1025 1030 1035

Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu  
 1040 1045 1050

Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly  
 1055 1060 1065

Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln

1070                      1075                      1080  
 Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro

1085                      1090                      1095  
 Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr  
 1100                      1105                      1110

Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr  
 1115                      1120                      1125

Gly Gly Ser Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser  
 1130                      1135                      1140

Met Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
 1145                      1150                      1155

<210> 159

<211> 1159

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 159

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser

1                      5                      10                      15

Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu

20                      25                      30

Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser

35                      40                      45

Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp

50                      55                      60

Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr

65                      70                      75                      80

Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala

85                      90                      95

Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln

	100		105		110
Val Asp Leu Arg Thr	Leu Gly Tyr Ser	Gln Gln Gln	Glu Lys Ile		
	115		120		125
Lys Pro Lys Val Arg Ser Thr	Val Ala Gln His His	Glu Ala Leu Val			
	130		135		140
Gly His Gly Phe Thr His Ala His Ile	Val Ala Leu Ser	Gln His Pro			
145		150		155	160
Ala Ala Leu Gly Thr Val Ala Val Lys Tyr	Gln Asp Met	Ile Ala Ala			
	165		170		175
Leu Pro Glu Ala Thr His Glu Ala Ile	Val Gly Val Gly	Lys Gln Trp			
	180		185		190
Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr	Val Ala Gly Glu Leu				
	195		200		205
Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala					
	210		215		220
Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn					
225		230		235	240
Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala					
	245		250		255
Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg					
	260		265		270
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val					
	275		280		285
Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val					
	290		295		300
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu					
305		310		315	320
Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu					
	325		330		335
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr					
	340		345		350

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 355 360 365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 370 375 380  
  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385 390 395 400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 405 410 415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
 420 425 430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 435 440 445  
  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 450 455 460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465 470 475 480  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510  
  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575  
  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu

595                                  600                                  605  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610                                  615                                  620  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625                                  630                                  635                                  640  
  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645                                  650                                  655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660                                  665                                  670  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675                                  680                                  685  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690                                  695                                  700  
  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705                                  710                                  715                                  720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 725                                  730                                  735  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 740                                  745                                  750  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 755                                  760                                  765  
  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
 770                                  775                                  780  
 Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
 785                                  790                                  795                                  800  
 Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805                                  810                                  815  
 Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
 820                                  825                                  830  
  
 Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Leu Tyr Pro Glu Arg  
 835                                  840                                  845

Leu Arg Arg Ile Leu Thr Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu  
 850 855 860  
 Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys  
 865 870 875 880  
 Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp  
 885 890 895  
  
 Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala  
 900 905 910  
 Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp  
 915 920 925  
 Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe  
 930 935 940  
 Lys Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala  
 945 950 955 960  
  
 Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu  
 965 970 975  
 Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile  
 980 985 990  
 Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn  
 995 1000 1005  
 Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile  
 1010 1015 1020  
  
 His Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile  
 1025 1030 1035  
 Ala Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe  
 1040 1045 1050  
 Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala  
 1055 1060 1065  
 Gly Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg  
 1070 1075 1080  
  
 Gln Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg

1085                      1090                      1095  
 Pro Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser  
 1100                      1105                      1110  
 Thr Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly  
 1115                      1120                      1125  
 Tyr Gly Gly Ser Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro  
 1130                      1135                      1140

Ser Met Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly  
 1145                      1150                      1155  
 Val

<210> 160

<211> 1188

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 160

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1                      5                      10                      15  
 Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
                     20                      25                      30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
                     35                      40                      45  
 Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
                     50                      55                      60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65                      70                      75                      80  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
                     85                      90                      95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln

100 105 110  
Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile  
115 120 125  
Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
130 135 140  
Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
145 150 155 160  
Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
165 170 175  
Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
180 185 190  
Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu  
195 200 205  
Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
210 215 220  
Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
225 230 235 240  
Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
245 250 255  
Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
260 265 270  
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
275 280 285  
Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
290 295 300  
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
305 310 315 320  
Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
325 330 335  
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
340 345 350

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 355 360 365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 370 375 380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385 390 395 400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 405 410 415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
 420 425 430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 435 440 445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 450 455 460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465 470 475 480  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu

595                                      600                                      605  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610                                      615                                      620  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625                                      630                                      635                                      640  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645                                      650                                      655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
  
 660                                      665                                      670  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675                                      680                                      685  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690                                      695                                      700  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705                                      710                                      715                                      720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
  
 725                                      730                                      735  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 740                                      745                                      750  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 755                                      760                                      765  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
 770                                      775                                      780  
 Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
  
 785                                      790                                      795                                      800  
 Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805                                      810                                      815  
 Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
 820                                      825                                      830  
 Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Ser Pro Lys Lys Lys  
 835                                      840                                      845

Arg Lys Val Glu Ala Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu

850 855 860

Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys

865 870 875 880

Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp

885 890 895

Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala

900 905 910

Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp

915 920 925

Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe

930 935 940

Lys Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala

945 950 955 960

Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu

965 970 975

Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile

980 985 990

Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn

995 1000 1005

Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile

1010 1015 1020

His Val Arg Ala Arg Ala Ala Gln Ala Thr Asp Ser His Ser Ile

1025 1030 1035

Ala Glu Ala Val Ala Arg Glu Lys Ile Ser Glu Arg Met Lys Phe

1040 1045 1050

Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala

1055 1060 1065

Gly Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg

1070 1075 1080

Gln Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg

1085                      1090                      1095  
 Pro Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser

1100                      1105                      1110  
 Thr Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly

1115                      1120                      1125  
 Tyr Ser His Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val

1130                      1135                      1140  
 Asn Ser Gly Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr

1145                      1150                      1155  
 Ser Ser Asp Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser

1160                      1165                      1170  
 Met Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val

1175                      1180                      1185

<210> 161

<211> 1164

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 161

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1                      5                      10                      15  
 Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu

20                      25                      30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser

35                      40                      45  
 Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp

50                      55                      60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr

65                      70                      75                      80

Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
85 90 95  
Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
100 105 110  
Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
115 120 125  
Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
130 135 140  
Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
145 150 155 160  
Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
165 170 175  
Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
180 185 190  
Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu  
195 200 205  
Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
210 215 220  
Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
225 230 235 240  
Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
245 250 255  
Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
260 265 270  
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
275 280 285  
Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
290 295 300  
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
305 310 315 320  
Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu

325 330 335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr

340 345 350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala

355 360 365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

370 375 380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys

385 390 395 400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala

405 410 415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly

420 425 430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys

435 440 445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn

450 455 460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val

465 470 475 480  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala

485 490 495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu

500 505 510  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala

515 520 525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg

530 535 540  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val

545 550 555 560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val

565 570 575

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 725 730 735

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 740 745 750

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 755 760 765

Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
 770 775 780

Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
 785 790 795 800

Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805 810 815

Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile

820                      825                      830  
 Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Ser Pro Lys Lys Lys  
 835                      840                      845  
 Arg Lys Val Glu Ala Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu  
  
 850                      855                      860  
 Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys  
 865                      870                      875                      880  
 Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp  
 885                      890                      895  
 Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala  
 900                      905                      910  
 Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp  
  
 915                      920                      925  
 Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe  
 930                      935                      940  
 Lys Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala  
 945                      950                      955                      960  
 Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu  
 965                      970                      975  
 Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile  
  
 980                      985                      990  
 Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn  
 995                      1000                      1005  
 Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile  
 1010                      1015                      1020  
 His Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile  
 1025                      1030                      1035  
 Ala Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe  
  
 1040                      1045                      1050  
 Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala  
 1055                      1060                      1065

Gly Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg  
 1070 1075 1080

Gly Gly Ser Val Ala Ser Thr Pro Met Thr Val Val Pro Ser Pro  
 1085 1090 1095

Glu Met Val Leu Ser Gly Tyr Ser His Glu Met Val His Ser Gly  
 1100 1105 1110

Tyr Ser Ser Glu Met Val Asn Ser Gly Tyr Leu His Val Asn Pro  
 1115 1120 1125

Met Gln Gln Val Asn Thr Ser Ser Asp Pro Leu Ser Cys Phe Asn  
 1130 1135 1140

Asn Gly Glu Ala Pro Ser Met Trp Asp Ser His Val Gln Asn Leu  
 1145 1150 1155

Tyr Gly Asn Leu Gly Val  
 1160

<210> 162

<211> 1159

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 162

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1 5 10 15

Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
 20 25 30

Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
 35 40 45

Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
 50 55 60

Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65 70 75 80

Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
85 90 95

Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
100 105 110

Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
115 120 125

Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
130 135 140

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
145 150 155 160

Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
165 170 175

Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
180 185 190

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu  
195 200 205

Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
210 215 220

Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
225 230 235 240

Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
245 250 255

Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
260 265 270

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
275 280 285

Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
290 295 300

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
305 310 315 320

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu

	325		330		335
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr					
	340		345		350
Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala					
	355		360		365
Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly					
	370		375		380
Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys					
385		390		395	400
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala					
	405		410		415
His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly					
	420		425		430
Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys					
	435		440		445
Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn					
	450		455		460
Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val					
465		470		475	480
Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala					
	485		490		495
Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu					
	500		505		510
Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala					
	515		520		525
Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg					
	530		535		540
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val					
545		550		555	560
Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val					
	565		570		575

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 725 730 735

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 740 745 750

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 755 760 765

Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
 770 775 780

Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
 785 790 795 800

Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805 810 815

Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile

820                                825                                830  
 Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Leu Tyr Pro Glu Arg  
                                   835                                840                                845  
 Leu Arg Arg Ile Leu Thr Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu  
                                   850                                855                                860  
 Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys  
 865                                870                                875                                880  
  
 Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp  
                                   885                                890                                895  
 Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala  
                                   900                                905                                910  
 Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp  
                                   915                                920                                925  
 Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe  
                                   930                                935                                940  
  
 Lys Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala  
 945                                950                                955                                960  
 Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu  
                                   965                                970                                975  
 Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile  
                                   980                                985                                990  
 Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn  
                                   995                                1000                                1005  
  
 Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile  
                                   1010                                1015                                1020  
 His Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile  
                                   1025                                1030                                1035  
 Ala Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe  
                                   1040                                1045                                1050  
 Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala  
                                   1055                                1060                                1065

Gly Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg  
 1070 1075 1080  
 Gln Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg  
 1085 1090 1095  
 Pro Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser  
 1100 1105 1110  
 Thr Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly  
 1115 1120 1125

Tyr Gly Gly Ser Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro  
 1130 1135 1140  
 Ser Met Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly  
 1145 1150 1155

Val

<210> 163

<211> 1198

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 163

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser

1	5	10	15
Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu			
	20	25	30
Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser			
	35	40	45
Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp			
	50	55	60
Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr			
65	70	75	80

Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
85 90 95

Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
100 105 110

Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
115 120 125

Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
130 135 140

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
145 150 155 160

Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
165 170 175

Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
180 185 190

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu  
195 200 205

Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
210 215 220

Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
225 230 235 240

Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
245 250 255

Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
260 265 270

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
275 280 285

Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
290 295 300

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
305 310 315 320

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu

	325	330	335
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr			
	340	345	350
Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala			
	355	360	365
Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly			
	370	375	380
Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys			
385	390	395	400
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala			
	405	410	415
His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly			
	420	425	430
Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys			
	435	440	445
Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn			
450	455	460	
Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val			
465	470	475	480
Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala			
	485	490	495
Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu			
	500	505	510
Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala			
515	520	525	
Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg			
530	535	540	
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val			
545	550	555	560
Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val			
	565	570	575

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 725 730 735  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 740 745 750  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 755 760 765  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
 770 775 780  
 Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
 785 790 795 800  
 Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805 810 815  
 Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile

820                      825                      830  
 Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Gly Gly Gly Gly Ser  
  
 835                      840                      845  
 Gly Gly Gly Gly Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn  
 850                      855                      860  
 Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr  
 865                      870                      875                      880  
 Leu Leu Tyr Pro Glu Arg Leu Arg Arg Ile Leu Thr Asn Pro Thr Phe  
 885                      890                      895  
 Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser Ser Met Ile Thr Gly  
  
 900                      905                      910  
 Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly Leu Asn Leu Pro Met  
 915                      920                      925  
 Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser Arg Leu Ser Ile Ser  
 930                      935                      940  
 Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys Ala Ala Lys Phe Asp  
 945                      950                      955                      960  
 Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys Lys Met Thr Met Asn  
  
 965                      970                      975  
 Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu Lys Ser Lys Ile Thr  
 980                      985                      990  
 Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys Lys Met Lys His Lys  
 995                      1000                      1005  
 Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser Lys Val  
 1010                      1015                      1020  
 Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His Val Arg Ala Arg  
  
 1025                      1030                      1035  
 Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala Glu Arg Val Arg  
 1040                      1045                      1050  
 Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu Gln Asp Leu Val  
 1055                      1060                      1065

Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly Met Leu Asp Glu  
 1070 1075 1080

Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln Ile Glu Phe Leu  
 1085 1090 1095

Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro Asp Phe Asp Met  
 1100 1105 1110

Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr Pro Met Thr Val  
 1115 1120 1125

Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr Ser His Glu Met  
 1130 1135 1140

Val His Ser Gly Tyr Ser Ser Glu Met Val Asn Ser Gly Tyr Leu  
 1145 1150 1155

His Val Asn Pro Met Gln Gln Val Asn Thr Ser Ser Asp Pro Leu  
 1160 1165 1170

Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser Met Trp Asp Ser His  
 1175 1180 1185

Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
 1190 1195

<210> 164

<211> 1198

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 164

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1 5 10 15

Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
 20 25 30

Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
 35 40 45

Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
 50 55 60  
  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65 70 75 80  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
 85 90 95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
 100 105 110  
 Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
 115 120 125  
  
 Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
 130 135 140  
 Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
 145 150 155 160  
 Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
 165 170 175  
 Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
 180 185 190  
  
 Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu  
 195 200 205  
 Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
 210 215 220  
 Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
 225 230 235 240  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
 245 250 255  
  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 260 265 270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 275 280 285  
 Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val

290                      295                      300  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305                      310                      315                      320  
  
 Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
                                  325                      330                      335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
                                  340                      345                      350  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
                                  355                      360                      365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                                  370                      375                      380  
  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385                      390                      395                      400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                                  405                      410                      415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
                                  420                      425                      430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                                  435                      440                      445  
  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
                                  450                      455                      460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465                      470                      475                      480  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                                  485                      490                      495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
                                  500                      505                      510  
  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
                                  515                      520                      525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
                                  530                      535                      540

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545                      550                      555                      560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
                                  565                      570                      575  
  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
                                  580                      585                      590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
                                  595                      600                      605  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
                                  610                      615                      620  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625                      630                      635                      640  
  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                                  645                      650                      655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
                                  660                      665                      670  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                                  675                      680                      685  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
                                  690                      695                      700  
  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705                      710                      715                      720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
                                  725                      730                      735  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
                                  740                      745                      750  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                                  755                      760                      765  
  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
                                  770                      775                      780  
 Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val

785                      790                      795                      800  
Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
                                         805                      810                      815  
Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
                                         820                      825                      830

Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Gly Gly Gly Gly Ser  
                                         835                      840                      845  
Gly Gly Gly Gly Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn  
                                         850                      855                      860  
Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr  
865                      870                      875                      880  
Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser  
                                         885                      890                      895

Ser Met Ile Thr Gly Gly Glu Met Asp Leu Tyr Pro Glu Arg Leu Arg  
                                         900                      905                      910  
Arg Ile Leu Thr Ser Tyr Leu Ser Thr Ala Gly Leu Asn Leu Pro Met  
                                         915                      920                      925  
Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser Arg Leu Ser Ile Ser  
                                         930                      935                      940  
Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys Ala Ala Lys Phe Asp  
945                      950                      955                      960

Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys Lys Met Thr Met Asn  
                                         965                      970                      975  
Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu Lys Ser Lys Ile Thr  
                                         980                      985                      990  
Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys Lys Met Lys His Lys  
                                         995                      1000                      1005  
Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser Lys Val  
                                         1010                      1015                      1020

Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His Val Arg Ala Arg  
                                         1025                      1030                      1035

Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala Glu Arg Val Arg  
 1040 1045 1050  
 Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu Gln Asp Leu Val  
 1055 1060 1065  
 Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly Met Leu Asp Glu  
 1070 1075 1080

Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln Ile Glu Phe Leu  
 1085 1090 1095  
 Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro Asp Phe Asp Met  
 1100 1105 1110  
 Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr Pro Met Thr Val  
 1115 1120 1125  
 Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr Ser His Glu Met  
 1130 1135 1140

Val His Ser Gly Tyr Ser Ser Glu Met Val Asn Ser Gly Tyr Leu  
 1145 1150 1155  
 His Val Asn Pro Met Gln Gln Val Asn Thr Ser Ser Asp Pro Leu  
 1160 1165 1170  
 Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser Met Trp Asp Ser His  
 1175 1180 1185  
 Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
 1190 1195

<210> 165

<211> 1198

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 165

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1 5 10 15

Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
 20 25 30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
 35 40 45  
  
 Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
 50 55 60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65 70 75 80  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
 85 90 95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
 100 105 110  
  
 Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
 115 120 125  
 Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
 130 135 140  
 Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
 145 150 155 160  
 Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
 165 170 175  
  
 Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
 180 185 190  
 Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu  
 195 200 205  
 Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala  
 210 215 220  
 Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn  
 225 230 235 240  
  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
 245 250 255  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg

260	265	270
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val		
275	280	285
Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val		
290	295	300
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu		
305	310	315
Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu		
325	330	335
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr		
340	345	350
Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala		
355	360	365
Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly		
370	375	380
Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys		
385	390	395
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala		
405	410	415
His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly		
420	425	430
Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys		
435	440	445
Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn		
450	455	460
Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val		
465	470	475
Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala		
485	490	495
Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu		
500	505	510

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 725 730 735

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 740 745 750

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala

755 760 765  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
 770 775 780  
 Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
 785 790 795 800  
 Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805 810 815

Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
 820 825 830  
 Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Gly Gly Gly Gly Ser  
 835 840 845  
 Gly Gly Gly Gly Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn  
 850 855 860  
 Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr  
 865 870 875 880

Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser  
 885 890 895  
 Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly  
 900 905 910  
 Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Leu Tyr  
 915 920 925  
 Pro Glu Arg Leu Arg Arg Ile Leu Thr Asp Ser Arg Leu Ser Ile Ser  
 930 935 940

Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys Ala Ala Lys Phe Asp  
 945 950 955 960  
 Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys Lys Met Thr Met Asn  
 965 970 975  
 Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu Lys Ser Lys Ile Thr  
 980 985 990  
 Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys Lys Met Lys His Lys  
 995 1000 1005

Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser Lys Val  
 1010 1015 1020

Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His Val Arg Ala Arg  
 1025 1030 1035

Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala Glu Arg Val Arg  
 1040 1045 1050

Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu Gln Asp Leu Val  
 1055 1060 1065

Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly Met Leu Asp Glu  
 1070 1075 1080

Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln Ile Glu Phe Leu  
 1085 1090 1095

Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro Asp Phe Asp Met  
 1100 1105 1110

Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr Pro Met Thr Val  
 1115 1120 1125

Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr Ser His Glu Met  
 1130 1135 1140

Val His Ser Gly Tyr Ser Ser Glu Met Val Asn Ser Gly Tyr Leu  
 1145 1150 1155

His Val Asn Pro Met Gln Gln Val Asn Thr Ser Ser Asp Pro Leu  
 1160 1165 1170

Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser Met Trp Asp Ser His  
 1175 1180 1185

Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
 1190 1195

<210> 166

<211> 1200

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 166

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser

1                    5                    10                    15

Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu

                  20                    25                    30

Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser

                  35                    40                    45

Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp

                  50                    55                    60

Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr

65                    70                    75                    80

Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala

                  85                    90                    95

Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln

                  100                    105                    110

Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile

                  115                    120                    125

Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val

                  130                    135                    140

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro

145                    150                    155                    160

Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala

                  165                    170                    175

Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp

                  180                    185                    190

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu

                  195                    200                    205

Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala

                  210                    215                    220

Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn

225                      230                      235                      240  
 Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala  
                                  245                      250                      255  
 Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
                                  260                      265                      270  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
                                  275                      280                      285  
  
 Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val  
                                  290                      295                      300  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 305                      310                      315                      320  
 Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu  
                                  325                      330                      335  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
                                  340                      345                      350  
  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
                                  355                      360                      365  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                                  370                      375                      380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys  
 385                      390                      395                      400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                                  405                      410                      415  
  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
                                  420                      425                      430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                                  435                      440                      445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
                                  450                      455                      460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465                      470                      475                      480

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540  
  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605  
  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670  
  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 690 695 700  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn

	725	730	735
Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val			
	740	745	750
Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala			
	755	760	765
Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu			
	770	775	780
Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val			
785	790	795	800
Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys			
	805	810	815
Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile			
	820	825	830
Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Gly Gly Gly Gly Ser			
	835	840	845
Gly Gly Gly Gly Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn			
	850	855	860
Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr			
865	870	875	880
Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser			
	885	890	895
Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly			
	900	905	910
Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser			
	915	920	925
Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys			
	930	935	940
Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Ala			
945	950	955	960
Ala Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Leu			
	965	970	975

Tyr Pro Glu Arg Leu Arg Arg Ile Leu Thr Glu Glu Glu Lys Ser Lys  
 980 985 990

Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys Lys Met Lys  
 995 1000 1005

His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser  
 1010 1015 1020

Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His Val Arg  
 1025 1030 1035

Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala Glu Arg  
 1040 1045 1050

Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu Gln Asp  
 1055 1060 1065

Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly Met Leu  
 1070 1075 1080

Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln Ile Glu  
 1085 1090 1095

Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro Asp Phe  
 1100 1105 1110

Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr Pro Met  
 1115 1120 1125

Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr Ser His  
 1130 1135 1140

Glu Met Val His Ser Gly Tyr Ser Ser Glu Met Val Asn Ser Gly  
 1145 1150 1155

Tyr Leu His Val Asn Pro Met Gln Gln Val Asn Thr Ser Ser Asp  
 1160 1165 1170

Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser Met Trp Asp  
 1175 1180 1185

Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
 1190 1195 1200

<210> 167

<211> 1198  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 167  
 Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1                    5                    10                    15  
  
 Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
                   20                    25                    30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
                   35                    40                    45  
 Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
                   50                    55                    60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65                    70                    75                    80  
  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
                   85                    90                    95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
                   100                    105                    110  
 Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
                   115                    120                    125  
 Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
                   130                    135                    140  
  
 Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
 145                    150                    155                    160  
 Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala  
                   165                    170                    175  
 Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp  
                   180                    185                    190  
 Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu

195	200	205	
Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala			
210	215	220	
Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn			
225	230	235	240
Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala			
	245	250	255
Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg			
	260	265	270
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val			
	275	280	285
Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val			
	290	295	300
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu			
305	310	315	320
Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu			
	325	330	335
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr			
	340	345	350
Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala			
	355	360	365
Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly			
	370	375	380
Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys			
385	390	395	400
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala			
	405	410	415
His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly			
	420	425	430
Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys			
	435	440	445

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 450 455 460

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465 470 475 480

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
 660 665 670

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
 675 680 685

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly

690 695 700  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 705 710 715 720  
  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 725 730 735  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 740 745 750  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 755 760 765  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
 770 775 780  
  
 Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
 785 790 795 800  
 Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
 805 810 815  
 Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
 820 825 830  
 Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Gly Gly Gly Gly Ser  
 835 840 845  
  
 Gly Gly Gly Gly Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn  
 850 855 860  
 Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr  
 865 870 875 880  
 Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser  
 885 890 895  
 Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly  
 900 905 910  
  
 Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser  
 915 920 925  
 Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys  
 930 935 940

Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys  
 945                      950                      955                      960

Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu  
                                  965                      970                      975

Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys  
                                  980                      985                      990

Lys Leu Tyr Pro Glu Arg Leu Arg Arg Ile Leu Thr Met Lys His Lys  
                                  995                      1000                      1005

Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp Ser Ser Lys Val  
                                  1010                      1015                      1020

Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His Val Arg Ala Arg  
                                  1025                      1030                      1035

Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala Glu Arg Val Arg  
                                  1040                      1045                      1050

Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu Gln Asp Leu Val  
                                  1055                      1060                      1065

Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly Met Leu Asp Glu  
                                  1070                      1075                      1080

Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln Ile Glu Phe Leu  
                                  1085                      1090                      1095

Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro Asp Phe Asp Met  
                                  1100                      1105                      1110

Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr Pro Met Thr Val  
                                  1115                      1120                      1125

Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr Ser His Glu Met  
                                  1130                      1135                      1140

Val His Ser Gly Tyr Ser Ser Glu Met Val Asn Ser Gly Tyr Leu  
                                  1145                      1150                      1155

His Val Asn Pro Met Gln Gln Val Asn Thr Ser Ser Asp Pro Leu  
                                  1160                      1165                      1170

Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser Met Trp Asp Ser His

1175                      1180                      1185  
 Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
 1190                      1195  
 <210> 168  
 <211> 1198  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"  
  
 <400> 168  
 Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
 1                      5                      10                      15  
 Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
                     20                      25                      30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
                     35                      40                      45  
 Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
                     50                      55                      60  
  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
 65                      70                      75                      80  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
                     85                      90                      95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
                     100                      105                      110  
 Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
                     115                      120                      125  
  
 Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val  
                     130                      135                      140  
 Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro  
 145                      150                      155                      160  
 Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala

	165	170	175
Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp			
	180	185	190
Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu			
	195	200	205
Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala			
	210	215	220
Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn			
225	230	235	240
Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala			
	245	250	255
Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg			
	260	265	270
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val			
	275	280	285
Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val			
	290	295	300
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu			
305	310	315	320
Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu			
	325	330	335
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr			
	340	345	350
Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala			
	355	360	365
Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly			
	370	375	380
Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys			
385	390	395	400
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala			
	405	410	415

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
 420 425 430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
 435 440 445  
  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 450 455 460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465 470 475 480  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
 485 490 495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
 500 505 510  
  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
 515 520 525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530 535 540  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545 550 555 560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
 565 570 575  
  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
 580 585 590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
 595 600 605  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610 615 620  
 Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625 630 635 640  
  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 645 650 655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys

660	665	670
Gln Ala Leu Glu Thr Val	Gln Arg Leu Leu Pro Val	Leu Cys Gln Ala
675	680	685
His Gly Leu Thr Pro Glu	Gln Val Val Ala Ile Ala Ser	Asn Gly Gly
690	695	700
Gly Lys Gln Ala Leu Glu Thr Val	Gln Arg Leu Leu Pro Val	Leu Cys
705	710	715
720	725	730
Gln Ala His Gly Leu Thr Pro Glu	Gln Val Val Ala Ile Ala Ser	Asn
735	740	745
Ile Gly Gly Lys Gln Ala Leu Glu Thr Val	Gln Arg Leu Leu Pro Val	
750	755	760
Leu Cys Gln Ala His Gly Leu Thr Pro Glu	Gln Val Val Ala Ile Ala	
765	770	775
Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser	Ile Val Ala Gln Leu	
780	785	790
Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr	Asn Asp His Leu Val	
795	800	805
Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp	Ala Val Lys Lys	
810	815	820
Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr	Asn Arg Arg Ile	
825	830	835
Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Gly Gly Gly Gly Ser		
840	845	850
Gly Gly Gly Gly Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu Asn		
855	860	865
Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys Tyr		
870	875	880
Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp Ser		
885	890	895
Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala Gly		
900	905	910

Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp Ser  
 915 920 925  
 Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe Lys  
 930 935 940  
 Ala Ala Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Ala Ala Lys  
 945 950 955 960  
  
 Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu Glu  
 965 970 975  
 Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile Lys  
 980 985 990  
 Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn Asp  
 995 1000 1005  
 Ser Ser Lys Val Thr Leu Tyr Pro Glu Arg Leu Arg Arg Ile Leu  
 1010 1015 1020  
  
 Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile His Val Arg Ala Arg  
 1025 1030 1035  
 Arg Gly Gln Ala Thr Asp Ser His Ser Ile Ala Glu Arg Val Arg  
 1040 1045 1050  
 Arg Glu Lys Ile Ser Glu Arg Met Lys Phe Leu Gln Asp Leu Val  
 1055 1060 1065  
 Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala Gly Met Leu Asp Glu  
 1070 1075 1080  
  
 Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg Gln Ile Glu Phe Leu  
 1085 1090 1095  
 Ser Met Lys Leu Ala Ile Val Asn Pro Arg Pro Asp Phe Asp Met  
 1100 1105 1110  
 Asp Asp Ile Phe Ala Lys Glu Val Ala Ser Thr Pro Met Thr Val  
 1115 1120 1125  
 Val Pro Ser Pro Glu Met Val Leu Ser Gly Tyr Ser His Glu Met  
 1130 1135 1140  
  
 Val His Ser Gly Tyr Ser Ser Glu Met Val Asn Ser Gly Tyr Leu

1145                      1150                      1155  
 His Val Asn Pro Met Gln Gln Val Asn Thr Ser Ser Asp Pro Leu  
 1160                      1165                      1170  
 Ser Cys Phe Asn Asn Gly Glu Ala Pro Ser Met Trp Asp Ser His  
 1175                      1180                      1185  
 Val Gln Asn Leu Tyr Gly Asn Leu Gly Val  
 1190                      1195  
 <210> 169  
 <211> 1165  
 <212> PRT  
  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"  
 <400> 169  
 Met Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro  
 1                      5                      10                      15  
 Ser Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp  
                     20                      25                      30  
 Leu Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp  
                     35                      40                      45  
  
 Ser Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu  
                     50                      55                      60  
 Trp Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro  
 65                      70                      75                      80  
 Thr Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro  
                     85                      90                      95  
 Ala Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala  
                     100                      105                      110  
  
 Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys  
                     115                      120                      125  
 Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu

130                      135                      140  
 Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His  
 145                      150                      155                      160  
 Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala  
                          165                      170                      175  
  
 Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln  
                          180                      185                      190  
 Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu  
                          195                      200                      205  
 Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile  
                          210                      215                      220  
 Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg  
 225                      230                      235                      240  
  
 Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val  
                          245                      250                      255  
 Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln  
                          260                      265                      270  
 Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln  
                          275                      280                      285  
 Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr  
                          290                      295                      300  
  
 Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro  
 305                      310                      315                      320  
 Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu  
                          325                      330                      335  
 Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu  
                          340                      345                      350  
 Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln  
                          355                      360                      365  
  
 Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
                          370                      375                      380

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly  
 385                    390                    395                    400  
 Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln  
                          405                    410                    415  
 Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn  
                          420                    425                    430  
  
 Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu  
                          435                    440                    445  
 Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser  
                          450                    455                    460  
 Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro  
 465                    470                    475                    480  
 Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile  
                          485                    490                    495  
  
 Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu  
                          500                    505                    510  
 Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val  
                          515                    520                    525  
 Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln  
                          530                    535                    540  
 Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln  
 545                    550                    555                    560  
  
 Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr  
                          565                    570                    575  
 Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro  
                          580                    585                    590  
 Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu  
                          595                    600                    605  
 Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu  
                          610                    615                    620  
  
 Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln

625                                  630                                  635                                  640  
 Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His  
                                                         645                                  650                                  655  
 Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly  
                                 660                                  665                                  670  
 Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln  
                                 675                                  680                                  685  
  
 Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly  
                                 690                                  695                                  700  
 Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu  
 705                                  710                                  715                                  720  
 Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser  
                                 725                                  730                                  735  
 Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro  
                                 740                                  745                                  750  
  
 Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile  
                                 755                                  760                                  765  
 Ala Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln  
                                 770                                  775                                  780  
 Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu  
 785                                  790                                  795                                  800  
 Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys  
                                 805                                  810                                  815  
  
 Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg  
                                 820                                  825                                  830  
 Ile Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Ser Pro Lys Lys  
                                 835                                  840                                  845  
 Lys Arg Lys Val Glu Ala Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu  
                                 850                                  855                                  860  
 Leu Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu  
 865                                  870                                  875                                  880

Lys Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala  
 885 890 895

Asp Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr  
 900 905 910

Ala Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly  
 915 920 925

Asp Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn  
 930 935 940

Phe Lys Lys Arg Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Lys  
 945 950 955 960

Lys Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu  
 965 970 975

Glu Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser  
 980 985 990

Ile Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser  
 995 1000 1005

Asn Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr  
 1010 1015 1020

Ile His Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser  
 1025 1030 1035

Ile Ala Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys  
 1040 1045 1050

Phe Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys  
 1055 1060 1065

Ala Gly Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln  
 1070 1075 1080

Arg Gly Gly Ser Val Ala Ser Thr Pro Met Thr Val Val Pro Ser  
 1085 1090 1095

Pro Glu Met Val Leu Ser Gly Tyr Ser His Glu Met Val His Ser  
 1100 1105 1110

Gly Tyr Ser Ser Glu Met Val Asn Ser Gly Tyr Leu His Val Asn

1115                      1120                      1125  
 Pro Met Gln Gln Val Asn Thr Ser Ser Asp Pro Leu Ser Cys Phe  
 1130                      1135                      1140  
 Asn Asn Gly Glu Ala Pro Ser Met Trp Asp Ser His Val Gln Asn  
 1145                      1150                      1155  
 Leu Tyr Gly Asn Leu Gly Val  
 1160                      1165  
 <210> 170  
 <211> 1159  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"  
 <400> 170  
 Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Arg Thr Arg Leu Pro Ser  
  
 1                      5                      10                      15  
 Pro Pro Ala Pro Ser Pro Ala Phe Ser Ala Asp Ser Phe Ser Asp Leu  
                     20                      25                      30  
 Leu Arg Gln Phe Asp Pro Ser Leu Phe Asn Thr Ser Leu Phe Asp Ser  
                     35                      40                      45  
 Leu Pro Pro Phe Gly Ala His His Thr Glu Ala Ala Thr Gly Glu Trp  
                     50                      55                      60  
 Asp Glu Val Gln Ser Gly Leu Arg Ala Ala Asp Ala Pro Pro Pro Thr  
  
 65                      70                      75                      80  
 Met Arg Val Ala Val Thr Ala Ala Arg Pro Pro Arg Ala Lys Pro Ala  
                     85                      90                      95  
 Pro Arg Arg Arg Ala Ala Gln Pro Ser Asp Ala Ser Pro Ala Ala Gln  
                     100                      105                      110  
 Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile  
                     115                      120                      125

Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val

130 135 140

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro

145 150 155 160

Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala

165 170 175

Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp

180 185 190

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu

195 200 205

Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala

210 215 220

Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn

225 230 235 240

Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala

245 250 255

Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg

260 265 270

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val

275 280 285

Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val

290 295 300

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu

305 310 315 320

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu

325 330 335

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr

340 345 350

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala

355 360 365

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

370                    375                    380  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys

385                    390                    395                    400  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                          405                    410                    415  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Asn Gly  
                          420                    425                    430  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                          435                    440                    445  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn

450                    455                    460  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
 465                    470                    475                    480  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                          485                    490                    495  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
                          500                    505                    510  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala

515                    520                    525  
 Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 530                    535                    540  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
 545                    550                    555                    560  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
                          565                    570                    575  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu

580                    585                    590  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
                          595                    600                    605  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 610                    615                    620

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala  
 625                      630                      635                      640  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
  
                                 645                      650                      655  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
                                 660                      665                      670  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                                 675                      680                      685  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
                                 690                      695                      700  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
  
 705                      710                      715                      720  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
                                 725                      730                      735  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
                                 740                      745                      750  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                                 755                      760                      765  
 Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu  
  
                                 770                      775                      780  
 Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val  
 785                      790                      795                      800  
 Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys  
                                 805                      810                      815  
 Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile  
                                 820                      825                      830  
 Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Ser Pro Lys Lys Lys  
  
                                 835                      840                      845  
 Arg Lys Val Glu Ala Ser Asn Gly Ala Ile Gly Gly Asp Leu Leu Leu  
                                 850                      855                      860  
 Asn Phe Pro Asp Met Ser Val Leu Glu Arg Gln Arg Ala His Leu Lys

865                      870                      875                      880  
 Tyr Leu Asn Pro Thr Phe Asp Ser Pro Leu Ala Gly Phe Phe Ala Asp  
                                  885                      890                      895  
 Ser Ser Met Ile Thr Gly Gly Glu Met Asp Ser Tyr Leu Ser Thr Ala  
  
                                  900                      905                      910  
 Gly Leu Asn Leu Pro Met Met Tyr Gly Glu Thr Thr Val Glu Gly Asp  
                                  915                      920                      925  
 Ser Arg Leu Ser Ile Ser Pro Glu Thr Thr Leu Gly Thr Gly Asn Phe  
                                  930                      935                      940  
 Lys Lys Arg Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Lys Lys  
 945                      950                      955                      960  
 Lys Lys Met Thr Met Asn Arg Asp Asp Leu Val Glu Glu Gly Glu Glu  
  
                                  965                      970                      975  
 Glu Lys Ser Lys Ile Thr Glu Gln Asn Asn Gly Ser Thr Lys Ser Ile  
                                  980                      985                      990  
 Lys Lys Met Lys His Lys Ala Lys Lys Glu Glu Asn Asn Phe Ser Asn  
                                  995                      1000                      1005  
 Asp Ser Ser Lys Val Thr Lys Glu Leu Glu Lys Thr Asp Tyr Ile  
                                  1010                      1015                      1020  
 His Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Ser His Ser Ile  
  
                                  1025                      1030                      1035  
 Ala Glu Arg Val Arg Arg Glu Lys Ile Ser Glu Arg Met Lys Phe  
                                  1040                      1045                      1050  
 Leu Gln Asp Leu Val Pro Gly Cys Asp Lys Ile Thr Gly Lys Ala  
                                  1055                      1060                      1065  
 Gly Met Leu Asp Glu Ile Ile Asn Tyr Val Gln Ser Leu Gln Arg  
                                  1070                      1075                      1080  
 Gln Ile Glu Phe Leu Ser Met Lys Leu Ala Ile Val Asn Pro Arg  
  
                                  1085                      1090                      1095  
 Pro Asp Phe Asp Met Asp Asp Ile Phe Ala Lys Glu Val Ala Ser  
                                  1100                      1105                      1110

Thr Pro Met Thr Val Val Pro Ser Pro Glu Met Val Leu Ser Gly  
 1115 1120 1125

Tyr Gly Gly Ser Pro Leu Ser Cys Phe Asn Asn Gly Glu Ala Pro  
 1130 1135 1140

Ser Met Trp Asp Ser His Val Gln Asn Leu Tyr Gly Asn Leu Gly  
 1145 1150 1155

Val

<210> 171

<211> 846

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 171

Met Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys  
 1 5 10 15  
 Lys Lys Met Asp Lys Lys Thr Ile Val Trp Phe Arg Arg Asp Leu Arg  
 20 25 30

Ile Glu Asp Asn Pro Ala Leu Ala Ala Ala Ala His Glu Gly Ser Val  
 35 40 45

Phe Pro Val Phe Ile Trp Cys Pro Glu Glu Glu Gly Gln Phe Tyr Pro  
 50 55 60

Gly Arg Ala Ser Arg Trp Trp Met Lys Gln Ser Leu Ala His Leu Ser  
 65 70 75 80

Gln Ser Leu Lys Ala Leu Gly Ser Asp Leu Thr Leu Ile Lys Thr His  
 85 90 95

Asn Thr Ile Ser Ala Ile Leu Asp Cys Ile Arg Val Thr Gly Ala Thr  
 100 105 110

Lys Val Val Phe Asn His Leu Tyr Asp Pro Val Ser Leu Val Arg Asp  
 115 120 125

His Thr Val Lys Glu Lys Leu Val Glu Arg Gly Ile Ser Val Gln Ser  
 130 135 140  
 Tyr Asn Gly Asp Leu Leu Tyr Glu Pro Trp Glu Ile Tyr Cys Glu Lys  
 145 150 155 160  
  
 Gly Lys Pro Phe Thr Ser Phe Asn Ser Tyr Trp Lys Lys Cys Leu Asp  
 165 170 175  
 Met Ser Ile Glu Ser Val Met Leu Pro Pro Pro Trp Arg Leu Met Pro  
 180 185 190  
 Ile Thr Ala Ala Ala Glu Ala Ile Trp Ala Cys Ser Ile Glu Glu Leu  
 195 200 205  
 Gly Leu Glu Asn Glu Ala Glu Lys Pro Ser Asn Ala Leu Leu Thr Arg  
 210 215 220  
  
 Ala Trp Ser Pro Gly Trp Ser Asn Ala Asp Lys Leu Leu Asn Glu Phe  
 225 230 235 240  
 Ile Glu Lys Gln Leu Ile Asp Tyr Ala Lys Asn Ser Lys Lys Val Val  
 245 250 255  
 Gly Asn Ser Thr Ser Leu Leu Ser Pro Tyr Leu His Phe Gly Glu Ile  
 260 265 270  
 Ser Val Arg His Val Phe Gln Cys Ala Arg Met Lys Gln Ile Ile Trp  
 275 280 285  
  
 Ala Arg Asp Lys Asn Ser Glu Gly Glu Glu Ser Ala Asp Leu Phe Leu  
 290 295 300  
 Arg Gly Ile Gly Leu Arg Glu Tyr Ser Arg Tyr Ile Cys Phe Asn Phe  
 305 310 315 320  
 Pro Phe Thr His Glu Gln Ser Leu Leu Ser His Leu Arg Phe Phe Pro  
 325 330 335  
 Trp Asp Ala Asp Val Asp Lys Phe Lys Ala Trp Arg Gln Gly Arg Thr  
 340 345 350  
  
 Gly Tyr Pro Leu Val Asp Ala Gly Met Arg Glu Leu Trp Ala Thr Gly  
 355 360 365  
 Trp Met His Asn Arg Ile Arg Val Ile Val Ser Ser Phe Ala Val Lys

370                                  375                                  380  
 Phe Leu Leu Leu Pro Trp Lys Trp Gly Met Lys Tyr Phe Trp Asp Thr  
 385                                  390                                  395                                  400  
 Leu Leu Asp Ala Asp Leu Glu Cys Asp Ile Leu Gly Trp Gln Tyr Ile  
                                          405                                  410                                  415  
  
 Ser Gly Ser Ile Pro Asp Gly His Glu Leu Asp Arg Leu Asp Asn Pro  
                                          420                                  425                                  430  
 Ala Leu Gln Gly Ala Lys Tyr Asp Pro Glu Gly Glu Tyr Ile Arg Gln  
                                          435                                  440                                  445  
 Trp Leu Pro Glu Leu Ala Arg Leu Pro Thr Glu Trp Ile His His Pro  
                                          450                                  455                                  460  
 Trp Asp Ala Pro Leu Thr Val Leu Lys Ala Ser Gly Val Glu Leu Gly  
 465                                  470                                  475                                  480  
  
 Thr Asn Tyr Ala Lys Pro Ile Val Asp Ile Asp Thr Ala Arg Glu Leu  
                                          485                                  490                                  495  
 Leu Ala Lys Ala Ile Ser Arg Thr Arg Glu Ala Gln Ile Met Ile Gly  
                                          500                                  505                                  510  
 Ala Ala Pro Ala Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Gly  
                                          515                                  520                                  525  
 Ser Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly  
                                          530                                  535                                  540  
  
 Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala  
 545                                  550                                  555                                  560  
 Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp  
                                          565                                  570                                  575  
 Phe Asp Leu Asp Met Leu Ile Asn Ser Arg Gly Ser Gly Glu Gly Arg  
                                          580                                  585                                  590  
 Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro Val  
                                          595                                  600                                  605  
  
 Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu  
                                          610                                  615                                  620

Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly  
 625                      630                      635                      640  
 Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr  
                                  645                      650                      655  
 Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr  
                                  660                      665                      670  
  
 Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His  
                                  675                      680                      685  
 Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr  
                                  690                      695                      700  
 Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys  
 705                      710                      715                      720  
 Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp  
                                  725                      730                      735  
  
 Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr  
                                  740                      745                      750  
 Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile  
                                  755                      760                      765  
 Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln  
                                  770                      775                      780  
 Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val  
 785                      790                      795                      800  
  
 Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys  
                                  805                      810                      815  
 Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr  
                                  820                      825                      830  
 Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Val  
                                  835                      840                      845  
 <210> 172  
 <211> 840  
 <212> PRT  
 <213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 172

Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Lys Met Asp Lys Lys

1                    5                    10                    15

Thr Ile Val Trp Phe Arg Arg Asp Leu Arg Ile Glu Asp Asn Pro Ala

                  20                    25                    30

Leu Ala Ala Ala Ala His Glu Gly Ser Val Phe Pro Val Phe Ile Trp

                  35                    40                    45

Cys Pro Glu Glu Glu Gly Gln Phe Tyr Pro Gly Arg Ala Ser Arg Trp

                  50                    55                    60

Trp Met Lys Gln Ser Leu Ala His Leu Ser Gln Ser Leu Lys Ala Ala

65                    70                    75                    80

Gly Ser Asp Ala Thr Leu Ile Lys Thr His Asn Thr Ile Ser Ala Ile

                  85                    90                    95

Leu Asp Cys Ile Arg Val Thr Gly Ala Thr Lys Val Val Phe Asn His

                  100                    105                    110

Leu Tyr Asp Pro Val Ser Leu Val Arg Asp His Thr Val Lys Glu Lys

                  115                    120                    125

Leu Val Glu Arg Gly Ile Ser Val Gln Ser Tyr Asn Gly Asp Leu Leu

                  130                    135                    140

Tyr Glu Pro Trp Glu Ile Tyr Cys Glu Lys Gly Lys Pro Phe Thr Ser

145                    150                    155                    160

Phe Asn Ser Tyr Trp Lys Lys Cys Leu Asp Met Ser Ile Glu Ser Val

                  165                    170                    175

Met Leu Pro Pro Pro Trp Arg Leu Met Pro Ile Thr Ala Ala Ala Glu

                  180                    185                    190

Ala Ile Trp Ala Cys Ser Ile Glu Glu Leu Gly Leu Glu Asn Glu Ala

                  195                    200                    205

Glu Lys Pro Ser Asn Ala Leu Leu Thr Arg Ala Trp Ser Pro Gly Trp

210	215	220	
Ser Asn Ala Asp Lys Leu	Leu Asn Glu Phe Ile	Glu Lys Gln Leu Ile	
225	230	235	240
Asp Tyr Ala Lys Asn Ser	Lys Lys Val Val	Gly Asn Ser Thr	Ser Leu
	245	250	255
Leu Ser Pro Tyr Leu His	Phe Gly Glu Ile	Ser Val Arg His	Val Phe
	260	265	270
Gln Cys Ala Arg Met Lys	Gln Ile Ile Trp Ala	Arg Asp Lys Asn Ser	
	275	280	285
Glu Gly Glu Glu Ser Ala	Asp Leu Phe Leu Arg	Gly Ile Gly Leu Arg	
	290	295	300
Glu Tyr Ser Arg Tyr Ile	Cys Phe Asn Phe Pro	Phe Thr His Glu Gln	
305	310	315	320
Ser Leu Leu Ser His Leu	Arg Phe Phe Pro Trp	Asp Ala Asp Val Asp	
	325	330	335
Lys Phe Lys Ala Trp Arg	Gln Gly Arg Thr Gly	Tyr Pro Leu Val Asp	
	340	345	350
Ala Gly Met Arg Glu Leu	Trp Ala Thr Gly Trp	Met His Asn Arg Ile	
	355	360	365
Arg Val Ile Val Ser Ser	Phe Ala Val Lys Phe	Leu Leu Leu Pro Trp	
	370	375	380
Lys Trp Gly Met Lys Tyr	Phe Trp Asp Thr Leu	Leu Asp Ala Asp Leu	
385	390	395	400
Glu Cys Asp Ile Leu Gly	Trp Gln Tyr Ile Ser	Gly Ser Ile Pro Asp	
	405	410	415
Gly His Glu Leu Asp Arg	Leu Asp Asn Pro Ala	Leu Gln Gly Ala Lys	
	420	425	430
Tyr Asp Pro Glu Gly Glu	Tyr Ile Arg Gln Trp	Leu Pro Glu Leu Ala	
	435	440	445
Arg Leu Pro Thr Glu Trp	Ile His His Pro Trp	Asp Ala Pro Leu Thr	
	450	455	460

Val Leu Lys Ala Ser Gly Val Glu Leu Gly Thr Asn Tyr Ala Lys Pro  
 465                      470                      475                      480

Ile Val Asp Ile Asp Thr Ala Arg Glu Leu Leu Ala Lys Ala Ile Ser  
                                  485                      490                      495

Arg Thr Arg Glu Ala Gln Ile Met Ile Gly Ala Ala Pro Ala Ser Pro  
                                  500                      505                      510

Lys Lys Lys Arg Lys Val Glu Ala Ser Gly Ser Gly Arg Ala Asp Ala  
                                  515                      520                      525

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp  
                                  530                      535                      540

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu  
 545                      550                      555                      560

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu  
                                  565                      570                      575

Ile Asn Ser Arg Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys  
                                  580                      585                      590

Gly Asp Val Glu Glu Asn Pro Gly Pro Val Ser Lys Gly Glu Glu Leu  
                                  595                      600                      605

Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn  
                                  610                      615                      620

Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr  
 625                      630                      635                      640

Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val  
                                  645                      650                      655

Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe  
                                  660                      665                      670

Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala  
                                  675                      680                      685

Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp  
                                  690                      695                      700

Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu

705                    710                    715                    720  
 Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn  
                          725                    730                    735  
 Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr  
                          740                    745                    750  
 Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile  
                          755                    760                    765

Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln  
                          770                    775                    780  
 Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His  
 785                    790                    795                    800  
 Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg  
                          805                    810                    815  
 Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu  
                          820                    825                    830

Gly Met Asp Glu Leu Tyr Lys Val  
                          835                    840

<210> 173

<211> 841

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 173

Met Lys Met Asp Lys Lys Thr Ile Val Trp Phe Arg Arg Asp Leu Arg  
 1                    5                    10                    15  
 Ile Glu Asp Asn Pro Ala Leu Ala Ala Ala Ala His Glu Gly Ser Val  
                          20                    25                    30  
 Phe Pro Val Phe Ile Trp Cys Pro Glu Glu Glu Gly Gln Phe Tyr Pro  
                          35                    40                    45  
 Gly Arg Ala Ser Arg Trp Trp Met Lys Gln Ser Leu Ala His Leu Ser

50                      55                      60  
 Gln Ser Leu Lys Ala Leu Gly Ser Asp Leu Thr Leu Ile Lys Thr His  
 65                      70                      75                      80  
 Asn Thr Ile Ser Ala Ile Leu Asp Cys Ile Arg Val Thr Gly Ala Thr  
                                  85                      90                      95  
 Lys Val Val Phe Asn His Leu Tyr Asp Pro Val Ser Leu Val Arg Asp  
  
                                  100                      105                      110  
 His Thr Val Lys Glu Lys Leu Val Glu Arg Gly Ile Ser Val Gln Ser  
                                  115                      120                      125  
 Tyr Asn Gly Asp Leu Leu Tyr Glu Pro Trp Glu Ile Tyr Cys Glu Lys  
                                  130                      135                      140  
 Gly Lys Pro Phe Thr Ser Phe Asn Ser Tyr Trp Lys Lys Cys Leu Asp  
 145                      150                      155                      160  
 Met Ser Ile Glu Ser Val Met Leu Pro Pro Pro Trp Arg Leu Met Pro  
  
                                  165                      170                      175  
 Ile Thr Ala Ala Ala Glu Ala Ile Trp Ala Cys Ser Ile Glu Glu Leu  
                                  180                      185                      190  
 Gly Leu Glu Asn Glu Ala Glu Lys Pro Ser Asn Ala Leu Leu Thr Arg  
                                  195                      200                      205  
 Ala Trp Ser Pro Gly Trp Ser Asn Ala Asp Lys Leu Leu Asn Glu Phe  
                                  210                      215                      220  
 Ile Glu Lys Gln Leu Ile Asp Tyr Ala Lys Asn Ser Lys Lys Val Val  
  
 225                      230                      235                      240  
 Gly Asn Ser Thr Ser Leu Leu Ser Pro Tyr Leu His Phe Gly Glu Ile  
                                  245                      250                      255  
 Ser Val Arg His Val Phe Gln Cys Ala Arg Met Lys Gln Ile Ile Trp  
                                  260                      265                      270  
 Ala Arg Asp Lys Asn Ser Glu Gly Glu Glu Ser Ala Asp Leu Phe Leu  
                                  275                      280                      285  
 Arg Gly Ile Gly Leu Arg Glu Tyr Ser Arg Tyr Ile Cys Phe Asn Phe  
  
 290                      295                      300

Pro Phe Thr His Glu Gln Ser Leu Leu Ser His Leu Arg Phe Phe Pro  
 305                      310                      315                      320  
 Trp Asp Ala Asp Val Asp Lys Phe Lys Ala Trp Arg Gln Gly Arg Thr  
                                  325                      330                      335  
 Gly Tyr Pro Leu Val Asp Ala Gly Met Arg Glu Leu Trp Ala Thr Gly  
                                  340                      345                      350  
 Trp Met His Asn Arg Ile Arg Val Ile Val Ser Ser Phe Ala Val Lys  
                                  355                      360                      365  
 Phe Leu Leu Leu Pro Trp Lys Trp Gly Met Lys Tyr Phe Trp Asp Thr  
                                  370                      375                      380  
 Leu Leu Asp Ala Asp Leu Glu Cys Asp Ile Leu Gly Trp Gln Tyr Ile  
 385                      390                      395                      400  
 Ser Gly Ser Ile Pro Asp Gly His Glu Leu Asp Arg Leu Asp Asn Pro  
                                  405                      410                      415  
 Ala Leu Gln Gly Ala Lys Tyr Asp Pro Glu Gly Glu Tyr Ile Arg Gln  
                                  420                      425                      430  
 Trp Leu Pro Glu Leu Ala Arg Leu Pro Thr Glu Trp Ile His His Pro  
                                  435                      440                      445  
 Trp Asp Ala Pro Leu Thr Val Leu Lys Ala Ser Gly Val Glu Leu Gly  
                                  450                      455                      460  
 Thr Asn Tyr Ala Lys Pro Ile Val Asp Ile Asp Thr Ala Arg Glu Leu  
 465                      470                      475                      480  
 Leu Ala Lys Ala Ile Ser Arg Thr Arg Glu Ala Gln Ile Met Ile Gly  
                                  485                      490                      495  
 Ala Ala Pro Ala Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Gly  
                                  500                      505                      510  
 Ser Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly  
                                  515                      520                      525  
 Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala  
                                  530                      535                      540  
 Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp

545                    550                    555                    560  
 Phe Asp Leu Asp Met Leu Ile Asn Ser Pro Lys Lys Lys Arg Lys Val  
                          565                    570                    575  
 Glu Ala Ser Ser Arg Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr  
                          580                    585                    590  
 Cys Gly Asp Val Glu Glu Asn Pro Gly Pro Val Ser Lys Gly Glu Glu  
                          595                    600                    605  
 Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val  
  
                          610                    615                    620  
 Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr  
 625                    630                    635                    640  
 Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro  
                          645                    650                    655  
 Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys  
                          660                    665                    670  
 Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser  
  
                          675                    680                    685  
 Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp  
                          690                    695                    700  
 Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr  
 705                    710                    715                    720  
 Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly  
                          725                    730                    735  
 Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val  
  
                          740                    745                    750  
 Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys  
                          755                    760                    765  
 Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr  
                          770                    775                    780  
 Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn  
 785                    790                    795                    800

His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys

805 810 815

Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr

820 825 830

Leu Gly Met Asp Glu Leu Tyr Lys Val

835 840

<210> 174

<211> 954

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 174

Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala Phe

1 5 10 15

Ser Ala Asp Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser Leu

20 25 30

Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His His

35 40 45

Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu Arg

50 55 60

Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala Ala

65 70 75 80

Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln Pro

85 90 95

Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr

100 105 110

Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val

115 120 125

Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His

130                      135                      140  
 Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val  
 145                      150                      155                      160  
 Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala  
                                  165                      170                      175  
 Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala  
                                  180                      185                      190  
 Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp  
  
                                  195                      200                      205  
 Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val  
                                  210                      215                      220  
 Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn  
 225                      230                      235                      240  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asn Gly Gly Lys  
                                  245                      250                      255  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
  
                                  260                      265                      270  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly  
                                  275                      280                      285  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                                  290                      295                      300  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
 305                      310                      315                      320  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
  
                                  325                      330                      335  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                                  340                      345                      350  
 Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu  
                                  355                      360                      365  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
                                  370                      375                      380

Ile Ala Ser His Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg  
 385                      390                      395                      400  
 Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val  
                                  405                      410                      415  
 Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val  
                                  420                      425                      430  
 Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu  
                                  435                      440                      445  
 Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu  
  
                                  450                      455                      460  
 Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr  
 465                      470                      475                      480  
 Pro Glu Gln Val Val Ala Ile Ala Ser His Asn Gly Gly Lys Gln Ala  
                                  485                      490                      495  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                                  500                      505                      510  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys  
  
                                  515                      520                      525  
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala  
                                  530                      535                      540  
 His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly  
 545                      550                      555                      560  
 Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys  
                                  565                      570                      575  
 Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn  
  
                                  580                      585                      590  
 Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val  
                                  595                      600                      605  
 Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala  
                                  610                      615                      620  
 Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu

625                    630                    635                    640  
 Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala  
  
                          645                    650                    655  
 Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala  
                          660                    665                    670  
 Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His  
                          675                    680                    685  
 Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val  
                          690                    695                    700  
 Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg  
  
 705                    710                    715                    720  
 Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala Ala Ser Met Ala Asn  
                          725                    730                    735  
 Ser Glu Ser Ser Ser Ser Pro Val Asn Glu Glu Glu Asn Ser Gln Arg  
                          740                    745                    750  
 Ile Ser Thr Leu His His Gln Thr Met Pro Ser Asp Leu Thr Gln Asp  
                          755                    760                    765  
 Glu Phe Thr Gln Leu Ser Gln Ser Ile Ala Glu Phe His Thr Tyr Gln  
  
                          770                    775                    780  
 Leu Gly Asn Gly Arg Cys Ser Ser Leu Leu Ala Gln Arg Ile His Ala  
 785                    790                    795                    800  
 Pro Pro Glu Thr Val Trp Ser Val Val Arg Arg Phe Asp Arg Pro Gln  
                          805                    810                    815  
 Ile Tyr Lys His Phe Ile Lys Ser Cys Asn Val Ser Glu Asp Phe Glu  
                          820                    825                    830  
 Met Arg Val Gly Cys Thr Arg Asp Val Asn Val Ile Ser Gly Leu Pro  
  
                          835                    840                    845  
 Ala Asn Thr Ser Arg Glu Arg Leu Asp Leu Leu Asp Asp Asp Arg Arg  
                          850                    855                    860  
 Val Thr Gly Phe Ser Ile Thr Gly Gly Glu His Arg Leu Arg Asn Tyr  
 865                    870                    875                    880

Lys Ser Val Thr Thr Val His Arg Phe Glu Lys Glu Glu Glu Glu Glu  
 885 890 895

Arg Ile Trp Thr Val Val Leu Glu Ser Tyr Val Val Asp Val Pro Glu  
 900 905 910

Gly Asn Ser Glu Glu Asp Thr Arg Leu Phe Ala Asp Thr Val Ile Arg  
 915 920 925

Leu Asn Leu Gln Lys Leu Ala Ser Ile Thr Glu Ala Met Asn Arg Asn  
 930 935 940

Asn Asn Asn Asn Asn Ser Ser Gln Val Arg  
 945 950

<210> 175  
 <211> 379  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 175

Met Val Pro Leu Tyr Gly Phe Thr Ser Ile Cys Gly Arg Arg Pro Glu  
 1 5 10 15

Met Glu Ala Ala Val Ser Thr Ile Pro Arg Phe Leu Gln Ser Ser Ser  
 20 25 30

Gly Ser Met Leu Asp Gly Arg Phe Asp Pro Gln Ser Ala Ala His Phe  
 35 40 45

Phe Gly Val Tyr Asp Gly His Gly Gly Ser Gln Val Ala Asn Tyr Cys  
 50 55 60

Arg Glu Arg Met His Leu Ala Leu Ala Glu Glu Ile Ala Lys Glu Lys  
 65 70 75 80

Pro Met Leu Cys Asp Gly Asp Thr Trp Leu Glu Lys Trp Lys Lys Ala  
 85 90 95

Leu Phe Asn Ser Phe Leu Arg Val Asp Ser Glu Ile Glu Ser Val Ala  
 100 105 110

Pro Glu Thr Val Gly Ser Thr Ser Val Val Ala Val Val Phe Pro Ser  
 115 120 125

His Ile Phe Val Ala Asn Cys Gly Asp Ser Arg Ala Val Leu Cys Arg  
 130 135 140

Gly Lys Thr Ala Leu Pro Leu Ser Val Asp His Lys Pro Asp Arg Glu  
 145 150 155 160

Asp Glu Ala Ala Arg Ile Glu Ala Ala Gly Gly Lys Val Ile Gln Trp  
 165 170 175

Asn Gly Ala Arg Val Phe Gly Val Leu Ala Met Ser Arg Ser Ile Gly  
 180 185 190

Asp Arg Tyr Leu Lys Pro Ser Ile Ile Pro Asp Pro Glu Val Thr Ala  
 195 200 205

Val Lys Arg Val Lys Glu Asp Asp Cys Leu Ile Leu Ala Ser Asp Gly  
 210 215 220

Val Trp Asp Val Met Thr Asp Glu Glu Ala Cys Glu Met Ala Arg Lys  
 225 230 235 240

Arg Ile Leu Leu Trp His Lys Lys Asn Ala Val Ala Gly Asp Ala Ser  
 245 250 255

Leu Leu Ala Asp Glu Arg Arg Lys Glu Gly Lys Asp Pro Ala Ala Met  
 260 265 270

Ser Ala Ala Glu Tyr Leu Ser Lys Leu Ala Ile Gln Arg Gly Ser Lys  
 275 280 285

Asp Asn Ile Ser Val Val Val Val Asp Leu Lys Pro Arg Arg Lys Leu  
 290 295 300

Lys Ser Lys Pro Leu Asn Ala Ser Pro Lys Lys Lys Arg Lys Val Glu  
 305 310 315 320

Ala Ser Gly Ser Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp  
 325 330 335

Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly  
 340 345 350

Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala

355                          360                          365  
 Leu Asp Asp Phe Asp Leu Asp Met Leu Ile Asn  
 370                          375  
 <210> 176  
 <211> 1453  
  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
                           polypeptide"  
 <400> 176  
 Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val  
 1                          5                          10                          15  
 Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe  
                           20                          25                          30  
 Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile  
                           35                          40                          45  
  
 Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu  
                           50                          55                          60  
 Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys  
 65                          70                          75                          80  
 Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser  
                           85                          90                          95  
 Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys  
                           100                          105                          110  
  
 His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr  
                           115                          120                          125  
 His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp  
                           130                          135                          140  
 Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His  
 145                          150                          155                          160  
 Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro

165                                  170                                  175  
 Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr  
                                           180                                  185                                  190  
 Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala  
                                           195                                  200                                  205  
 Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn  
                                           210                                  215                                  220  
 Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn  
 225                                  230                                  235                                  240  
  
 Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe  
                                           245                                  250                                  255  
 Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp  
                                           260                                  265                                  270  
 Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp  
                                           275                                  280                                  285  
 Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp  
                                           290                                  295                                  300  
  
 Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser  
 305                                  310                                  315                                  320  
 Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys  
                                           325                                  330                                  335  
 Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe  
                                           340                                  345                                  350  
 Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser  
                                           355                                  360                                  365  
  
 Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp  
                                           370                                  375                                  380  
 Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg  
 385                                  390                                  395                                  400  
 Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu  
                                           405                                  410                                  415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe  
 420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile  
 435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp  
 450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu  
 465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr  
 485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser  
 500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys  
 515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln  
 530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr  
 545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp  
 565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly  
 580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp  
 595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr  
 610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala  
 625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr  
 645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp

660                          665                          670  
 Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe  
 675                          680                          685  
  
 Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe  
 690                          695                          700  
 Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu  
 705                          710                          715                          720  
 His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly  
 725                          730                          735  
 Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly  
 740                          745                          750  
  
 Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln  
 755                          760                          765  
 Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile  
 770                          775                          780  
 Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro  
 785                          790                          795                          800  
 Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu  
 805                          810                          815  
  
 Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg  
 820                          825                          830  
 Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys  
 835                          840                          845  
 Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg  
 850                          855                          860  
 Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys  
 865                          870                          875                          880  
  
 Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys  
 885                          890                          895  
 Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp  
 900                          905                          910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr  
 915 920 925  
 Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp  
 930 935 940  
  
 Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser  
 945 950 955 960  
 Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg  
 965 970 975  
 Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val  
 980 985 990  
 Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe  
 995 1000 1005  
  
 Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala  
 1010 1015 1020  
 Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe  
 1025 1030 1035  
 Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala  
 1040 1045 1050  
 Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu  
 1055 1060 1065  
  
 Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val  
 1070 1075 1080  
 Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr  
 1085 1090 1095  
 Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys  
 1100 1105 1110  
 Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro  
 1115 1120 1125  
  
 Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val  
 1130 1135 1140  
 Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys

1145	1150	1155
Ser Val Lys Glu Leu Leu Gly	Ile Thr Ile Met Glu	Arg Ser Ser
1160	1165	1170
Phe Glu Lys Asn Pro Ile Asp	Phe Leu Glu Ala Lys	Gly Tyr Lys
1175	1180	1185
Glu Val Lys Lys Asp Leu Ile	Ile Lys Leu Pro Lys	Tyr Ser Leu
1190	1195	1200
Phe Glu Leu Glu Asn Gly Arg	Lys Arg Met Leu Ala	Ser Ala Gly
1205	1210	1215
Glu Leu Gln Lys Gly Asn Glu	Leu Ala Leu Pro Ser	Lys Tyr Val
1220	1225	1230
Asn Phe Leu Tyr Leu Ala Ser	His Tyr Glu Lys Leu	Lys Gly Ser
1235	1240	1245
Pro Glu Asp Asn Glu Gln Lys	Gln Leu Phe Val Glu	Gln His Lys
1250	1255	1260
His Tyr Leu Asp Glu Ile Ile	Glu Gln Ile Ser Glu	Phe Ser Lys
1265	1270	1275
Arg Val Ile Leu Ala Asp Ala	Asn Leu Asp Lys Val	Leu Ser Ala
1280	1285	1290
Tyr Asn Lys His Arg Asp Lys	Pro Ile Arg Glu Gln	Ala Glu Asn
1295	1300	1305
Ile Ile His Leu Phe Thr Leu	Thr Asn Leu Gly Ala	Pro Ala Ala
1310	1315	1320
Phe Lys Tyr Phe Asp Thr Thr	Ile Asp Arg Lys Arg	Tyr Thr Ser
1325	1330	1335
Thr Lys Glu Val Leu Asp Ala	Thr Leu Ile His Gln	Ser Ile Thr
1340	1345	1350
Gly Leu Tyr Glu Thr Arg Ile	Asp Leu Ser Gln Leu	Gly Gly Asp
1355	1360	1365
Ser Ala Gly Gly Gly Gly Ser	Gly Gly Gly Gly Ser	Gly Gly Gly
1370	1375	1380

Gly Ser Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Gly Ser  
 1385 1390 1395  
 Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly  
 1400 1405 1410  
 Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp  
 1415 1420 1425

Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu  
 1430 1435 1440  
 Asp Asp Phe Asp Leu Asp Met Leu Ile Asn  
 1445 1450

<210> 177

<211> 1569

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 177

Met Gly Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala Asp Tyr  
 1 5 10 15

Leu Glu Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser Met Leu  
 20 25 30

Pro Gly Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala Asp Tyr  
 35 40 45

Leu Glu Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser Met Leu  
 50 55 60

Pro Gly Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala Asp Tyr  
 65 70 75 80

Leu Glu Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser Met Leu  
 85 90 95

Pro Gly Ser Gly Met Asn Ile Gln Met Leu Leu Glu Ala Ala Asp Tyr  
 100 105 110

Leu Glu Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser Met Leu  
 115 120 125  
 Pro Ser Arg Ser Ala Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala  
 130 135 140  
  
 Asp Tyr Lys Asp Asp Asp Asp Lys Ser Arg Leu Glu Pro Gly Glu Lys  
 145 150 155 160  
 Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Ala  
 165 170 175  
 Leu Thr Arg His Gln Arg Thr His Thr Arg Asp Lys Lys Tyr Ser Ile  
 180 185 190  
 Gly Leu Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp  
 195 200 205  
  
 Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp  
 210 215 220  
 Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser  
 225 230 235 240  
 Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg  
 245 250 255  
 Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser  
 260 265 270  
  
 Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu  
 275 280 285  
 Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe  
 290 295 300  
 Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile  
 305 310 315 320  
 Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu  
 325 330 335  
  
 Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His  
 340 345 350  
 Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys

355                                  360                                  365  
 Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn  
 370                                  375                                  380  
 Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg  
 385                                  390                                  395                                  400  
  
 Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly  
 405                                  410                                  415  
 Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly  
 420                                  425                                  430  
 Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys  
 435                                  440                                  445  
 Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu  
 450                                  455                                  460  
  
 Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn  
 465                                  470                                  475                                  480  
 Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu  
 485                                  490                                  495  
 Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu  
 500                                  505                                  510  
 His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu  
 515                                  520                                  525  
  
 Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr  
 530                                  535                                  540  
 Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe  
 545                                  550                                  555                                  560  
 Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val  
 565                                  570                                  575  
 Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn  
 580                                  585                                  590  
  
 Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu  
 595                                  600                                  605

Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys  
 610 615 620  
 Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu  
 625 630 635 640  
 Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu  
 645 650 655  
  
 Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser  
 660 665 670  
 Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro  
 675 680 685  
 Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr  
 690 695 700  
 Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg  
 705 710 715 720  
  
 Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu  
 725 730 735  
 Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp  
 740 745 750  
 Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val  
 755 760 765  
 Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys  
 770 775 780  
  
 Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile  
 785 790 795 800  
 Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met  
 805 810 815  
 Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val  
 820 825 830  
 Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser  
 835 840 845  
  
 Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile

850                      855                      860  
 Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln  
 865                      870                      875                      880  
 Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala  
                                  885                      890                      895  
 Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu  
                                  900                      905                      910  
  
 Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val  
                                  915                      920                      925  
 Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile  
                                  930                      935                      940  
 Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys  
 945                      950                      955                      960  
 Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu  
                                  965                      970                      975  
  
 Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln  
                                  980                      985                      990  
 Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr  
                                  995                      1000                      1005  
 Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val  
                                  1010                      1015                      1020  
 Asp Ala Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp  
                                  1025                      1030                      1035  
  
 Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp  
                                  1040                      1045                      1050  
 Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp  
                                  1055                      1060                      1065  
 Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp  
                                  1070                      1075                      1080  
 Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys  
                                  1085                      1090                      1095

Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr  
 1100 1105 1110

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr  
 1115 1120 1125

Asp Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu  
 1130 1135 1140

Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr  
 1145 1150 1155

Lys Val Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr  
 1160 1165 1170

Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys  
 1175 1180 1185

Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val  
 1190 1195 1200

Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr  
 1205 1210 1215

Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr  
 1220 1225 1230

Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile  
 1235 1240 1245

Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg  
 1250 1255 1260

Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln Val Asn  
 1265 1270 1275

Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe Ser Lys Glu  
 1280 1285 1290

Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys  
 1295 1300 1305

Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr  
 1310 1315 1320

Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys

1325	1330	1335
Ser Lys Lys Leu Lys Ser Val	Lys Glu Leu Leu Gly	Ile Thr Ile
1340	1345	1350
Met Glu Arg Ser Ser Phe Glu	Lys Asn Pro Ile Asp	Phe Leu Glu
1355	1360	1365
Ala Lys Gly Tyr Lys Glu Val	Lys Lys Asp Leu Ile	Ile Lys Leu
1370	1375	1380
Pro Lys Tyr Ser Leu Phe Glu	Leu Glu Asn Gly Arg	Lys Arg Met
1385	1390	1395
Leu Ala Ser Ala Gly Glu Leu	Gln Lys Gly Asn Glu	Leu Ala Leu
1400	1405	1410
Pro Ser Lys Tyr Val Asn Phe	Leu Tyr Leu Ala Ser	His Tyr Glu
1415	1420	1425
Lys Leu Lys Gly Ser Pro Glu	Asp Asn Glu Gln Lys	Gln Leu Phe
1430	1435	1440
Val Glu Gln His Lys His Tyr	Leu Asp Glu Ile Ile	Glu Gln Ile
1445	1450	1455
Ser Glu Phe Ser Lys Arg Val	Ile Leu Ala Asp Ala	Asn Leu Asp
1460	1465	1470
Lys Val Leu Ser Ala Tyr Asn	Lys His Arg Asp Lys	Pro Ile Arg
1475	1480	1485
Glu Gln Ala Glu Asn Ile Ile	His Leu Phe Thr Leu	Thr Asn Leu
1490	1495	1500
Gly Ala Pro Ala Ala Phe Lys	Tyr Phe Asp Thr Thr	Ile Asp Arg
1505	1510	1515
Lys Arg Tyr Thr Ser Thr Lys	Glu Val Leu Asp Ala	Thr Leu Ile
1520	1525	1530
His Gln Ser Ile Thr Gly Leu	Tyr Glu Thr Arg Ile	Asp Leu Ser
1535	1540	1545
Gln Leu Gly Gly Asp Lys Arg	Pro Ala Ala Thr Lys	Lys Ala Gly
1550	1555	1560

Gln Ala Lys Lys Lys Lys

1565

<210> 178

<211> 84

<212> PRT

<213> Artificial Sequence

<220

><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 178

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Asn Gly

1                    5                    10                    15

Leu Met Glu Asp Pro Met Lys Val Tyr Lys Asp Arg Gln Phe Met Asn

                  20                    25                    30

Val Trp Thr Asp His Glu Lys Glu Ile Phe Lys Asp Lys Phe Ile Gln

                  35                    40                    45

His Pro Lys Asn Phe Gly Leu Ile Ala Ser Tyr Leu Glu Arg Lys Ser

                  50                    55                    60

Val Pro Asp Cys Val Leu Tyr Tyr Tyr Leu Thr Lys Lys Asn Glu Asn

65                    70                    75                    80

Tyr Lys Glu Phe

<210> 179

<211> 288

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 179

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Gly Asn

1                    5                    10                    15

Leu Met Ile Ser Phe Leu Lys Lys Asp Thr Gln Ser Ile Thr Leu Glu  
 20 25 30  
 Glu Leu Ala Lys Ile Ile Lys Lys Cys Lys His Val Val Ala Leu Thr  
 35 40 45  
 Gly Ser Gly Thr Ser Ala Glu Ser Asn Ile Pro Ser Phe Arg Gly Ser  
 50 55 60  
 Ser Asn Ser Ile Trp Ser Lys Tyr Asp Pro Arg Ile Tyr Gly Thr Ile  
 65 70 75 80  
  
 Trp Gly Phe Trp Lys Tyr Pro Glu Lys Ile Trp Glu Val Ile Arg Asp  
 85 90 95  
 Ile Ser Ser Asp Tyr Glu Ile Glu Ile Asn Asn Gly His Val Ala Leu  
 100 105 110  
 Ser Thr Leu Glu Ser Leu Gly Tyr Leu Lys Ser Val Val Thr Gln Asn  
 115 120 125  
 Val Asp Gly Leu His Glu Ala Ser Gly Asn Thr Lys Val Ile Ser Leu  
 130 135 140  
  
 His Gly Asn Val Phe Glu Ala Val Cys Cys Thr Cys Asn Lys Ile Val  
 145 150 155 160  
 Lys Leu Asn Lys Ile Met Leu Gln Lys Thr Ser His Phe Met His Gln  
 165 170 175  
 Leu Pro Pro Glu Cys Pro Cys Gly Gly Ile Phe Lys Pro Asn Ile Ile  
 180 185 190  
 Leu Phe Gly Glu Val Val Ser Ser Asp Leu Leu Lys Glu Ala Glu Glu  
 195 200 205  
  
 Glu Ile Ala Lys Cys Asp Leu Leu Leu Val Ile Gly Thr Ser Ser Thr  
 210 215 220  
 Val Ser Thr Ala Thr Asn Leu Cys His Phe Ala Cys Lys Lys Lys Lys  
 225 230 235 240  
 Lys Ile Val Glu Ile Asn Ile Ser Lys Thr Tyr Ile Thr Asn Lys Met  
 245 250 255  
 Ser Asp Tyr His Val Cys Ala Lys Phe Ser Glu Leu Thr Lys Val Ala

260 265 270

Asn Ile Leu Lys Gly Ser Ser Glu Lys Asn Lys Lys Ile Met Glu Phe

275 280 285

<210> 180

<211> 346

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 180

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Glu Lys

1 5 10 15

Ala Phe Arg Pro His Phe Phe Asn His Gly Lys Pro Asp Ala Asn Pro

20 25 30

Lys Glu Lys Lys Asn Cys His Trp Cys Gln Ile Arg Ser Phe Ala Thr

35 40 45

His Ala Gln Leu Pro Ile Ser Ile Val Asn Arg Glu Asp Asp Ala Phe

50 55 60

Leu Asn Pro Asn Phe Arg Phe Ile Asp His Ser Ile Ile Gly Lys Asn

65 70 75 80

Val Pro Val Ala Asp Gln Ser Phe Arg Val Gly Cys Ser Cys Ala Ser

85 90 95

Asp Glu Glu Cys Met Tyr Ser Thr Cys Gln Cys Leu Asp Glu Met Ala

100 105 110

Pro Asp Ser Asp Glu Glu Ala Asp Pro Tyr Thr Arg Lys Lys Arg Phe

115 120 125

Ala Tyr Tyr Ser Gln Gly Ala Lys Lys Gly Leu Leu Arg Asp Arg Val

130 135 140

Leu Gln Ser Gln Glu Pro Ile Tyr Glu Cys His Gln Gly Cys Ala Cys

145 150 155 160

Ser Lys Asp Cys Pro Asn Arg Val Val Glu Arg Gly Arg Thr Val Pro  
 165 170 175  
 Leu Gln Ile Phe Arg Thr Lys Asp Arg Gly Trp Gly Val Lys Cys Pro  
 180 185 190  
 Val Asn Ile Lys Arg Gly Gln Phe Val Asp Arg Tyr Leu Gly Glu Ile  
 195 200 205  
 Ile Thr Ser Glu Glu Ala Asp Arg Arg Arg Ala Glu Ser Thr Ile Ala  
 210 215 220

Arg Arg Lys Asp Val Tyr Leu Phe Ala Leu Asp Lys Phe Ser Asp Pro  
 225 230 235 240  
 Asp Ser Leu Asp Pro Leu Leu Ala Gly Gln Pro Leu Glu Val Asp Gly  
 245 250 255  
 Glu Tyr Met Ser Gly Pro Thr Arg Phe Ile Asn His Ser Cys Asp Pro  
 260 265 270  
 Asn Met Ala Ile Phe Ala Arg Val Gly Asp His Ala Asp Lys His Ile  
 275 280 285

His Asp Leu Ala Leu Phe Ala Ile Lys Asp Ile Pro Lys Gly Thr Glu  
 290 295 300  
 Leu Thr Phe Asp Tyr Val Asn Gly Leu Thr Gly Leu Glu Ser Asp Ala  
 305 310 315 320  
 His Asp Pro Ser Lys Ile Ser Glu Met Thr Lys Cys Leu Cys Gly Thr  
 325 330 335  
 Ala Lys Cys Arg Gly Tyr Leu Trp Glu Phe  
 340 345

<210> 181  
 <211> 306  
 <212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 181

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Thr Glu Met  
 1                    5                    10                    15  
 Ser Val Arg Lys Ile Ala Ala His Met Lys Ser Asn Pro Asn Ala Lys  
                   20                    25                    30  
 Val Ile Phe Met Val Gly Ala Gly Ile Ser Thr Ser Cys Gly Ile Pro  
                   35                    40                    45

Asp Phe Arg Ser Pro Gly Thr Gly Leu Tyr His Asn Leu Ala Arg Leu  
                   50                    55                    60  
 Lys Leu Pro Tyr Pro Glu Ala Val Phe Asp Val Asp Phe Phe Gln Ser  
 65                    70                    75                    80  
 Asp Pro Leu Pro Phe Tyr Thr Leu Ala Lys Glu Leu Tyr Pro Gly Asn  
                   85                    90                    95  
 Phe Arg Pro Ser Lys Phe His Tyr Leu Leu Lys Leu Phe Gln Asp Lys  
                   100                    105                    110

Asp Val Leu Lys Arg Val Tyr Thr Gln Asn Ile Asp Thr Leu Glu Arg  
                   115                    120                    125  
 Gln Ala Gly Val Lys Asp Asp Leu Ile Ile Glu Ala His Gly Ser Phe  
                   130                    135                    140  
 Ala His Cys His Cys Ile Gly Cys Gly Lys Val Tyr Pro Pro Gln Val  
 145                    150                    155                    160  
 Phe Lys Ser Lys Leu Ala Glu His Pro Ile Lys Asp Phe Val Lys Cys  
                   165                    170                    175

Asp Val Cys Gly Glu Leu Val Lys Pro Ala Ile Val Phe Phe Gly Glu  
                   180                    185                    190  
 Asp Leu Pro Asp Ser Phe Ser Glu Thr Trp Leu Asn Asp Ser Glu Trp  
                   195                    200                    205  
 Leu Arg Glu Lys Ile Thr Thr Ser Gly Lys His Pro Gln Gln Pro Leu  
                   210                    215                    220  
 Val Ile Val Val Gly Thr Ser Leu Ala Val Tyr Pro Phe Ala Ser Leu  
 225                    230                    235                    240

Pro Glu Glu Ile Pro Arg Lys Val Lys Arg Val Leu Cys Asn Leu Glu

245 250 255  
 Thr Val Gly Asp Phe Lys Ala Asn Lys Arg Pro Thr Asp Leu Ile Val  
 260 265 270  
 His Gln Tyr Ser Asp Glu Phe Ala Glu Gln Leu Val Glu Glu Leu Gly  
 275 280 285  
 Trp Gln Glu Asp Phe Glu Lys Ile Leu Thr Ala Gln Gly Gly Met Gly  
 290 295 300

Glu Phe

305

<210> 182

<211> 272

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 182

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Val Gly  
 1 5 10 15  
 Ala Gly Ile Ser Thr Pro Ser Gly Ile Pro Asp Phe Arg Ser Pro Gly  
 20 25 30  
 Ser Gly Leu Tyr Ser Asn Leu Gln Gln Tyr Asp Leu Pro Tyr Pro Glu

35 40 45  
 Ala Ile Phe Glu Leu Pro Phe Phe Phe His Asn Pro Lys Pro Phe Phe  
 50 55 60

Thr Leu Ala Lys Glu Leu Tyr Pro Gly Asn Tyr Lys Pro Asn Val Thr  
 65 70 75 80

His Tyr Phe Leu Arg Leu Leu His Asp Lys Gly Leu Leu Leu Arg Leu  
 85 90 95

Tyr Thr Gln Asn Ile Asp Gly Leu Glu Arg Val Ser Gly Ile Pro Ala

100 105 110  
 Ser Lys Leu Val Glu Ala His Gly Thr Phe Ala Ser Ala Thr Cys Thr

115 120 125  
 Val Cys Gln Arg Pro Phe Pro Gly Glu Asp Ile Arg Ala Asp Val Met  
 130 135 140  
 Ala Asp Arg Val Pro Arg Cys Pro Val Cys Thr Gly Val Val Lys Pro  
 145 150 155 160  
 Asp Ile Val Phe Phe Gly Glu Pro Leu Pro Gln Arg Phe Leu Leu His

165 170 175  
 Val Val Asp Phe Pro Met Ala Asp Leu Leu Leu Ile Leu Gly Thr Ser  
 180 185 190  
 Leu Glu Val Glu Pro Phe Ala Ser Leu Thr Glu Ala Val Arg Ser Ser  
 195 200 205  
 Val Pro Arg Leu Leu Ile Asn Arg Asp Leu Val Gly Pro Leu Ala Trp  
 210 215 220  
 His Pro Arg Ser Arg Asp Val Ala Gln Leu Gly Asp Val Val His Gly

225 230 235 240  
 Val Glu Ser Leu Val Glu Leu Leu Gly Trp Thr Glu Glu Met Arg Asp  
 245 250 255  
 Leu Val Gln Arg Glu Thr Gly Lys Leu Asp Gly Pro Asp Lys Glu Phe  
 260 265 270

<210> 183

<211> 344

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 183

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Ala Ala

1 5 10 15  
 Ala Ala Asn Ser Gly Ser Ser Leu Pro Leu Phe Asp Cys Pro Thr Trp  
 20 25 30  
 Ala Gly Lys Pro Pro Pro Gly Leu His Leu Asp Val Val Lys Gly Asp



Ser Gln Pro His Gly Ile His Gly Thr Ala Leu Ile Gly Gly Leu Pro  
 290 295 300

Met Pro Tyr Pro Asn Leu Ala Pro Asp Val Asp Leu Thr Pro Val Val  
 305 310 315 320

Pro Ser Ala Val Asn Met Asn Pro Ala Pro Asn Pro Ala Val Tyr Asn  
 325 330 335

Pro Glu Ala Val Asn Glu Glu Phe  
 340

<210> 184

<211> 234

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 184

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Thr Thr  
 1 5 10 15

Asn Ser Thr Gln Asp Thr Leu Tyr Leu Ser Leu His Gly Gly Ile Asp  
 20 25 30

Ser Ala Ile Pro Tyr Pro Val Arg Arg Val Glu Gln Leu Leu Gln Phe  
 35 40 45

Ser Phe Leu Pro Glu Leu Gln Phe Gln Asn Ala Ala Val Lys Gln Arg  
 50 55 60

Ile Gln Arg Leu Cys Tyr Arg Glu Glu Lys Arg Leu Ala Val Ser Ser  
 65 70 75 80

Leu Ala Lys Trp Leu Gly Gln Leu His Lys Gln Arg Leu Arg Ala Pro  
 85 90 95

Lys Asn Pro Pro Val Ala Ile Cys Trp Ile Asn Ser Tyr Val Gly Tyr  
 100 105 110

Gly Val Phe Ala Arg Glu Ser Ile Pro Ala Trp Ser Tyr Ile Gly Glu  
 115 120 125

Tyr Thr Gly Ile Leu Arg Arg Arg Gln Ala Leu Trp Leu Asp Glu Asn  
 130 135 140  
 Asp Tyr Cys Phe Arg Tyr Pro Val Pro Arg Tyr Ser Phe Arg Tyr Phe  
 145 150 155 160  
 Thr Ile Asp Ser Gly Met Gln Gly Asn Val Thr Arg Phe Ile Asn His  
 165 170 175  
 Ser Asp Asn Pro Asn Leu Glu Ala Ile Gly Ala Phe Glu Asn Gly Ile  
 180 185 190  
 Phe His Ile Ile Ile Arg Ala Ile Lys Asp Ile Leu Pro Gly Glu Glu  
 195 200 205  
 Leu Cys Tyr His Tyr Gly Pro Leu Tyr Trp Lys His Arg Lys Lys Arg  
 210 215 220  
 Glu Glu Phe Val Pro Gln Glu Glu Glu Phe  
 225 230  
 <210> 185  
 <211> 233  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 185  
 Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Ala Arg Tyr  
 1 5 10 15  
 Leu Gly Asn Thr Val Asp Leu Ser Ser Phe Asp Phe Arg Thr Gly Lys  
 20 25 30  
 Met Met Pro Ser Lys Leu Gln Lys Asn Lys Gln Arg Leu Arg Asn Asp  
 35 40 45  
 Pro Leu Asn Gln Asn Lys Gly Lys Pro Asp Leu Asn Thr Thr Leu Pro  
 50 55 60  
 Ile Arg Gln Thr Ala Ser Ile Phe Lys Gln Pro Val Thr Lys Val Thr  
 65 70 75 80

Asn His Pro Ser Asn Lys Val Lys Ser Asp Pro Gln Arg Met Asn Glu  
 85 90 95  
 Gln Pro Arg Gln Leu Phe Trp Glu Lys Arg Leu Gln Gly Leu Ser Ala  
 100 105 110  
 Ser Asp Val Thr Glu Gln Ile Ile Lys Thr Met Glu Leu Pro Lys Gly  
 115 120 125  
 Leu Gln Gly Val Gly Pro Gly Ser Asn Asp Glu Thr Leu Leu Ser Ala  
 130 135 140  
 Val Ala Ser Ala Leu His Thr Ser Ser Ala Pro Ile Thr Gly Gln Val  
 145 150 155 160  
 Ser Ala Ala Val Glu Lys Asn Pro Ala Val Trp Leu Asn Thr Ser Gln  
 165 170 175  
 Pro Leu Cys Lys Ala Phe Ile Val Thr Asp Glu Asp Ile Arg Lys Gln  
 180 185 190  
 Glu Glu Arg Val Gln Gln Val Arg Lys Ile Leu Glu Asp Ala Leu Met  
 195 200 205  
 Ala Asp Ile Leu Ser Arg Ala Ala Asp Thr Glu Glu Met Asp Ile Glu  
 210 215 220  
 Met Asp Ser Gly Asp Glu Ala Glu Phe  
 225 230  
 <210> 186  
 <211> 346  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 186  
 Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Pro Ser  
 1 5 10 15  
 Leu Asp Asp Ile Leu Lys Pro Val Ala Glu Ala Val Lys Asn Gly Lys

20 25 30  
 Lys Val Thr Phe Phe Asn Gly Ala Gly Ile Ser Thr Gly Ala Gly Ile  
 35 40 45  
 Pro Asp Phe Arg Ser Pro Asp Thr Gly Leu Tyr Ala Asn Leu Ala Lys  
 50 55 60  
 Leu Asn Leu Pro Phe Ala Glu Ala Val Phe Asp Ile Asp Phe Phe Lys  
 65 70 75 80  
 Glu Asp Pro Lys Pro Phe Tyr Thr Leu Ala Glu Glu Leu Tyr Pro Gly  
  
 85 90 95  
 Asn Phe Ala Pro Thr Lys Phe His His Phe Ile Lys Leu Leu Gln Asp  
 100 105 110  
 Gln Gly Ser Leu Lys Arg Val Tyr Thr Gln Asn Ile Asp Thr Leu Glu  
 115 120 125  
 Arg Leu Ala Gly Val Glu Asp Lys Tyr Ile Val Glu Ala His Gly Ser  
 130 135 140  
 Phe Ala Ser Asn His Cys Val Asp Cys His Lys Glu Met Thr Thr Glu  
  
 145 150 155 160  
 Thr Leu Lys Thr Tyr Met Lys Asp Lys Lys Ile Pro Ser Cys Gln His  
 165 170 175  
 Cys Glu Gly Tyr Val Lys Pro Asp Ile Val Phe Phe Gly Glu Gly Leu  
 180 185 190  
 Pro Val Lys Phe Phe Asp Leu Trp Glu Asp Asp Cys Glu Asp Val Glu  
 195 200 205  
 Val Ala Ile Val Ala Gly Thr Ser Leu Thr Val Phe Pro Phe Ala Ser  
  
 210 215 220  
 Leu Pro Gly Glu Val Asn Lys Lys Cys Leu Arg Val Leu Val Asn Lys  
 225 230 235 240  
 Glu Lys Val Gly Thr Phe Lys His Glu Pro Arg Lys Ser Asp Ile Ile  
 245 250 255  
 Ala Leu His Asp Cys Asp Ile Val Ala Glu Arg Leu Cys Thr Leu Leu  
 260 265 270

Gly Leu Asp Asp Lys Leu Asn Glu Val Tyr Glu Lys Glu Lys Ile Lys

275 280 285

Tyr Ser Lys Ala Glu Thr Lys Glu Ile Lys Met His Glu Ile Glu Asp

290 295 300

Lys Leu Lys Glu Glu Ala His Leu Lys Glu Asp Lys His Thr Thr Lys

305 310 315 320

Val Asp Lys Lys Glu Lys Gln Asn Asp Ala Asn Asp Lys Glu Leu Glu

325 330 335

Gln Leu Ile Asp Lys Ala Lys Ala Glu Phe

340 345

<210> 187

<211> 350

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 187

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Glu Asn

1 5 10 15

Arg Ala Leu Asp Pro Gly Thr Arg Asp Ser Tyr Gly Ala Thr Ser His

20 25 30

Leu Pro Asn Lys Gly Ala Leu Ala Lys Val Lys Asn Asn Phe Lys Asp

35 40 45

Leu Met Ser Lys Leu Thr Glu Gly Gln Tyr Val Leu Cys Arg Trp Thr

50 55 60

Asp Gly Leu Tyr Tyr Leu Gly Lys Ile Lys Arg Val Ser Ser Ser Lys

65 70 75 80

Gln Ser Cys Leu Val Thr Phe Glu Asp Asn Ser Lys Tyr Trp Val Leu

85 90 95

Trp Lys Asp Ile Gln His Ala Gly Val Pro Gly Glu Glu Pro Lys Cys

100	105	110	
Asn Ile Cys Leu Gly Lys Thr Ser Gly Pro Leu Asn Glu Ile Leu Ile			
115	120	125	
Cys Gly Lys Cys Gly Leu Gly Tyr His Gln Gln Cys His Ile Pro Ile			
130	135	140	
Ala Gly Ser Ala Asp Gln Pro Leu Leu Thr Pro Trp Phe Cys Arg Arg			
145	150	155	160
Cys Ile Phe Ala Leu Ala Val Arg Lys Gly Gly Ala Leu Lys Lys Gly			
165	170	175	
Ala Ile Ala Arg Thr Leu Gln Ala Val Lys Met Val Leu Ser Tyr Gln			
180	185	190	
Pro Glu Glu Leu Glu Trp Asp Ser Pro His Arg Thr Asn Gln Gln Gln			
195	200	205	
Cys Tyr Cys Tyr Cys Gly Gly Pro Gly Glu Trp Tyr Leu Arg Met Leu			
210	215	220	
Gln Cys Tyr Arg Cys Arg Gln Trp Phe His Glu Ala Cys Thr Gln Cys			
225	230	235	240
Leu Asn Glu Pro Met Met Phe Gly Asp Arg Phe Tyr Leu Phe Phe Cys			
245	250	255	
Ser Val Cys Asn Gln Gly Pro Gly Gly Ser Gly Ser Asp Ser Ser Ala			
260	265	270	
Glu Gly Ala Ser Val Pro Glu Arg Pro Asp Glu Gly Ile Asp Ser His			
275	280	285	
Thr Phe Glu Ser Ile Ser Glu Asp Asp Ser Ser Leu Ser His Leu Lys			
290	295	300	
Ser Ser Ile Thr Asn Tyr Phe Gly Ala Ala Gly Arg Leu Ala Cys Gly			
305	310	315	320
Glu Lys Tyr Gln Val Leu Ala Arg Arg Val Thr Pro Glu Gly Lys Val			
325	330	335	
Gln Tyr Leu Val Glu Trp Glu Gly Thr Thr Pro Tyr Glu Phe			
340	345	350	

<210> 188  
 <211> 362  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 188

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Leu His  
 1                    5                    10                    15  
 Thr Thr Gln Leu Tyr Gln His Val Pro Glu Thr Arg Trp Pro Ile Val  
                   20                    25                    30  
 Tyr Ser Pro Arg Tyr Asn Ile Thr Phe Met Gly Leu Glu Lys Leu His  
                   35                    40                    45  
 Pro Phe Asp Ala Gly Lys Trp Gly Lys Val Ile Asn Phe Leu Lys Glu  
                   50                    55                    60  
  
 Glu Lys Leu Leu Ser Asp Ser Met Leu Val Glu Ala Arg Glu Ala Ser  
 65                    70                    75                    80  
 Glu Glu Asp Leu Leu Val Val His Thr Arg Arg Tyr Leu Asn Glu Leu  
                   85                    90                    95  
 Lys Trp Ser Phe Ala Val Ala Thr Ile Thr Glu Ile Pro Pro Val Ile  
                   100                    105                    110  
 Phe Leu Pro Asn Phe Leu Val Gln Arg Lys Val Leu Arg Pro Leu Arg  
                   115                    120                    125  
  
 Thr Gln Thr Gly Gly Thr Ile Met Ala Gly Lys Leu Ala Val Glu Arg  
                   130                    135                    140  
 Gly Trp Ala Ile Asn Val Gly Gly Gly Phe His His Cys Ser Ser Asp  
 145                    150                    155                    160  
 Arg Gly Gly Gly Phe Cys Ala Tyr Ala Asp Ile Thr Leu Ala Ile Lys  
                   165                    170                    175  
 Phe Leu Phe Glu Arg Val Glu Gly Ile Ser Arg Ala Thr Ile Ile Asp

180 185 190

Leu Asp Ala His Gln Gly Asn Gly His Glu Arg Asp Phe Met Asp Asp

195 200 205

Lys Arg Val Tyr Ile Met Asp Val Tyr Asn Arg His Ile Tyr Pro Gly

210 215 220

Asp Arg Phe Ala Lys Gln Ala Ile Arg Arg Lys Val Glu Leu Glu Trp

225 230 235 240

Gly Thr Glu Asp Asp Glu Tyr Leu Asp Lys Val Glu Arg Asn Ile Lys

245 250 255

Lys Ser Leu Gln Glu His Leu Pro Asp Val Val Val Tyr Asn Ala Gly

260 265 270

Thr Asp Ile Leu Glu Gly Asp Arg Leu Gly Gly Leu Ser Ile Ser Pro

275 280 285

Ala Gly Ile Val Lys Arg Asp Glu Leu Val Phe Arg Met Val Arg Gly

290 295 300

Arg Arg Val Pro Ile Leu Met Val Thr Ser Gly Gly Tyr Gln Lys Arg

305 310 315 320

Thr Ala Arg Ile Ile Ala Asp Ser Ile Leu Asn Leu Phe Gly Leu Gly

325 330 335

Leu Ile Gly Pro Glu Ser Pro Ser Val Ser Ala Gln Asn Ser Asp Thr

340 345 350

Pro Leu Leu Pro Pro Ala Val Pro Glu Phe

355 360

<210> 189

<211> 315

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 189

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Pro Leu  
 1                    5                    10                    15  
 Gln Ile Val His His Pro Asp Tyr Asp Ala Gly Phe Ala Thr Asn His  
                   20                    25                    30  
 Arg Phe Pro Met Ser Lys Tyr Pro Leu Leu Met Glu Ala Leu Arg Ala  
                   35                    40                    45  
 Arg Gly Leu Ala Ser Pro Asp Ala Leu Asn Thr Thr Glu Pro Ala Pro  
                   50                    55                    60

Ala Ser Trp Leu Lys Leu Ala His Ala Ala Asp Tyr Val Asp Gln Val  
 65                    70                    75                    80  
 Ile Ser Cys Ser Val Pro Glu Lys Ile Glu Arg Glu Ile Gly Phe Pro  
                   85                    90                    95  
 Val Gly Pro Arg Val Ser Leu Arg Ala Gln Leu Ala Thr Gly Gly Thr  
                   100                    105                    110  
 Ile Leu Ala Ala Arg Leu Ala Leu Arg His Gly Ile Ala Cys Asn Thr  
                   115                    120                    125

Ala Gly Gly Ser His His Ala Arg Arg Ala Gln Gly Ala Gly Phe Cys  
                   130                    135                    140  
 Thr Phe Asn Asp Val Ala Val Ala Ser Leu Val Leu Leu Asp Glu Gly  
 145                    150                    155                    160  
 Ala Ala Gln Asn Ile Leu Val Val Asp Leu Asp Val His Gln Gly Asp  
                   165                    170                    175  
 Gly Thr Ala Asp Ile Leu Ser Asp Glu Pro Gly Val Phe Thr Phe Ser  
                   180                    185                    190

Met His Gly Glu Arg Asn Tyr Pro Val Arg Lys Ile Ala Ser Asp Leu  
                   195                    200                    205  
 Asp Ile Ala Leu Pro Asp Gly Thr Gly Asp Ala Ala Tyr Leu Arg Arg  
                   210                    215                    220  
 Leu Ala Thr Ile Leu Pro Glu Leu Ser Ala Arg Ala Arg Trp Asp Ile  
 225                    230                    235                    240  
 Val Phe Tyr Asn Ala Gly Val Asp Val His Ala Glu Asp Arg Leu Gly

245 250 255

Arg Leu Ala Leu Ser Asn Gly Gly Leu Arg Ala Arg Asp Glu Met Val

260 265 270

Ile Gly His Phe Arg Ala Leu Gly Ile Pro Val Cys Gly Val Ile Gly

275 280 285

Gly Gly Tyr Ser Thr Asp Val Pro Ala Leu Ala Ser Arg His Ala Ile

290 295 300

Leu Phe Glu Val Ala Ser Thr Tyr Ala Glu Phe

305 310 315

<210> 190

<211> 134

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 190

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Phe Asn

1 5 10 15

Asp Arg Val Ile Val Lys Lys Ser Pro Leu Gly Gly Tyr Gly Val Phe

20 25 30

Ala Arg Lys Ser Phe Glu Lys Gly Glu Leu Val Glu Glu Cys Leu Cys

35 40 45

Ile Val Arg His Asn Asp Asp Trp Gly Thr Ala Leu Glu Asp Tyr Leu

50 55 60

Phe Ser Arg Lys Asn Met Ser Ala Met Ala Leu Gly Phe Gly Ala Ile

65 70 75 80

Phe Asn His Ser Lys Asp Pro Asn Ala Arg His Glu Leu Thr Ala Gly

85 90 95

Leu Lys Arg Met Arg Ile Phe Thr Ile Lys Pro Ile Ala Ile Gly Glu

100 105 110

Glu Ile Thr Ile Ser Tyr Gly Asp Asp Tyr Trp Leu Ser Arg Pro Arg  
                   115                                  120                                  125  
 Leu Thr Gln Asn Glu Phe  
                   130  
 <210> 191  
 <211> 282  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
                   polypeptide"  
 <400> 191  
 Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Asp Ile Ser  
 1                  5                                  10                                  15  
 Gly Gly Leu Glu Phe Lys Gly Ile Pro Ala Thr Asn Arg Val Asp Asp  
                                   20                                  25                                  30  
 Ser Pro Val Ser Pro Thr Ser Gly Phe Thr Tyr Ile Lys Ser Leu Ile  
                   35                                  40                                  45  
 Ile Glu Pro Asn Val Ile Ile Pro Lys Ser Ser Thr Gly Cys Asn Cys  
                   50                                  55                                  60  
 Arg Gly Ser Cys Thr Asp Ser Lys Lys Cys Ala Cys Ala Lys Leu Asn  
 65                                  70                                  75                                  80  
 Gly Gly Asn Phe Pro Tyr Val Asp Leu Asn Asp Gly Arg Leu Ile Glu  
                                   85                                  90                                  95  
 Ser Arg Asp Val Val Phe Glu Cys Gly Pro His Cys Gly Cys Gly Pro  
                   100                                  105                                  110  
 Lys Cys Val Asn Arg Thr Ser Gln Lys Arg Leu Arg Phe Asn Leu Glu  
                   115                                  120                                  125  
 Val Phe Arg Ser Ala Lys Lys Gly Trp Ala Val Arg Ser Trp Glu Tyr  
                   130                                  135                                  140  
 Ile Pro Ala Gly Ser Pro Val Cys Glu Tyr Ile Gly Val Val Arg Arg  
                                   145                                  150                                  155                                  160

Thr Ala Asp Val Asp Thr Ile Ser Asp Asn Glu Tyr Ile Phe Glu Ile  
 165 170 175  
 Asp Cys Gln Gln Thr Met Gln Gly Leu Gly Gly Arg Gln Arg Arg Leu  
 180 185 190  
 Arg Asp Val Ala Val Pro Met Asn Asn Gly Val Ser Gln Ser Ser Glu  
 195 200 205  
 Asp Glu Asn Ala Pro Glu Phe Cys Ile Asp Ala Gly Ser Thr Gly Asn  
 210 215 220  
 Phe Ala Arg Phe Ile Asn His Ser Cys Glu Pro Asn Leu Phe Val Gln  
 225 230 235 240  
 Cys Val Leu Ser Ser His Gln Asp Ile Arg Leu Ala Arg Val Val Leu  
 245 250 255  
 Phe Ala Ala Asp Asn Ile Ser Pro Met Gln Glu Leu Thr Tyr Asp Tyr  
 260 265 270  
 Gly Tyr Ala Leu Asp Ser Val His Glu Phe  
 275 280

<210> 192

<211> 319

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 192

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Ala Ser Arg  
 1 5 10 15  
 Arg Thr Gly Glu Phe Leu Arg Asp Ala Gln Ala Pro Ser Arg Trp Leu  
 20 25 30  
 Lys Arg Ser Lys Thr Gly Gln Asp Asp Gly Ala Phe Cys Leu Glu Thr  
 35 40 45  
 Trp Leu Ala Gly Ala Gly Asp Asp Ala Ala Gly Gly Glu Arg Gly Arg  
 50 55 60

Asp Arg Glu Gly Ala Ala Asp Lys Ala Lys Gln Arg Glu Glu Arg Arg  
 65                   70                   75                   80  
 Gln Lys Glu Leu Glu Glu Arg Phe Glu Glu Met Lys Val Glu Phe Glu  
                   85                   90                   95  
 Glu Lys Ala Gln Arg Met Ile Ala Arg Arg Ala Ala Leu Thr Gly Glu  
                   100                   105                   110  
  
 Ile Tyr Ser Asp Gly Lys Gly Ser Lys Lys Pro Arg Val Pro Ser Leu  
                   115                   120                   125  
 Pro Glu Asn Asp Asp Asp Ala Leu Ile Glu Ile Ile Ile Asp Pro Glu  
                   130                   135                   140  
 Gln Gly Ile Leu Lys Trp Pro Leu Ser Val Met Ser Ile Arg Gln Arg  
 145                   150                   155                   160  
 Thr Val Ile Tyr Gln Glu Cys Leu Arg Arg Asp Leu Thr Ala Cys Ile  
                   165                   170                   175  
  
 His Leu Thr Lys Val Pro Gly Lys Gly Arg Ala Val Phe Ala Ala Asp  
                   180                   185                   190  
 Thr Ile Leu Lys Asp Asp Phe Val Val Glu Tyr Lys Gly Glu Leu Cys  
                   195                   200                   205  
 Ser Glu Arg Glu Ala Arg Glu Arg Glu Gln Arg Tyr Asn Arg Ser Lys  
                   210                   215                   220  
 Val Pro Met Gly Ser Phe Met Phe Tyr Phe Lys Asn Gly Ser Arg Met  
 225                   230                   235                   240  
  
 Met Ala Ile Asp Ala Thr Asp Glu Lys Gln Asp Phe Gly Pro Ala Arg  
                   245                   250                   255  
 Leu Ile Asn His Ser Arg Arg Asn Pro Asn Met Thr Pro Arg Ala Ile  
                   260                   265                   270  
 Thr Leu Gly Asp Phe Asn Ser Glu Pro Arg Leu Ile Phe Val Ala Arg  
                   275                   280                   285  
 Arg Asn Ile Glu Lys Gly Glu Glu Leu Leu Val Asp Tyr Gly Glu Arg  
                   290                   295                   300  
  
 Asp Pro Asp Val Ile Lys Glu His Pro Trp Leu Asn Ser Glu Phe

305                    310                    315

<210> 193

<211> 304

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 193

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Ser Val  
 1                    5                    10                    15

Asn Tyr Ala Ala Gly Leu Ser Pro Tyr Ala Asp Lys Gly Lys Cys Gly  
                   20                    25                    30

Leu Pro Glu Ile Phe Asp Pro Pro Glu Glu Leu Glu Arg Lys Val Trp  
                   35                    40                    45

Glu Leu Ala Arg Leu Val Trp Gln Ser Ser Ser Val Val Phe His Thr  
                   50                    55                    60

Gly Ala Gly Ile Ser Thr Ala Ser Gly Ile Pro Asp Phe Arg Gly Pro  
 65                    70                    75                    80

His Gly Val Trp Thr Met Glu Glu Arg Gly Leu Ala Pro Lys Phe Asp  
                   85                    90                    95

Thr Thr Phe Glu Ser Ala Arg Pro Thr Gln Thr His Met Ala Leu Val  
                   100                    105                    110

Gln Leu Glu Arg Val Gly Leu Leu Arg Phe Leu Val Ser Gln Asn Val  
                   115                    120                    125

Asp Gly Leu His Val Arg Ser Gly Phe Pro Arg Asp Lys Leu Ala Glu  
                   130                    135                    140

Leu His Gly Asn Met Phe Val Glu Glu Cys Ala Lys Cys Lys Thr Gln  
 145                    150                    155                    160

Tyr Val Arg Asp Thr Val Val Gly Thr Met Gly Leu Lys Ala Thr Gly  
                   165                    170                    175

Arg Leu Cys Thr Val Ala Lys Ala Arg Gly Leu Arg Ala Cys Arg Gly

180 185 190  
 Glu Leu Arg Asp Thr Ile Leu Asp Trp Glu Asp Ser Leu Pro Asp Arg

195 200 205  
 Asp Leu Ala Leu Ala Asp Glu Ala Ser Arg Asn Ala Asp Leu Ser Ile  
 210 215 220

Thr Leu Gly Thr Ser Leu Gln Ile Arg Pro Ser Gly Asn Leu Pro Leu  
 225 230 235 240

Ala Thr Lys Arg Arg Gly Gly Arg Leu Val Ile Val Asn Leu Gln Pro  
 245 250 255

Thr Lys His Asp Arg His Ala Asp Leu Arg Ile His Gly Tyr Val Asp  
 260 265 270

Glu Val Met Thr Arg Leu Met Lys His Leu Gly Leu Glu Ile Pro Ala  
 275 280 285

Trp Asp Gly Pro Arg Val Leu Glu Arg Ala Leu Pro Pro Leu Glu Phe  
 290 295 300

<210> 194

<211> 257

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 194

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Lys Val  
 1 5 10 15

Ala Ala Lys Lys Leu Ala Thr Ser Arg Met Arg Lys Asp Arg Ala Ala  
 20 25 30

Ala Ala Ser Pro Ser Ser Asp Ile Glu Asn Ser Glu Asn Pro Ser Ser  
 35 40 45

Leu Ala Ser His Ser Ser Ser Ser Gly Arg Met Thr Pro Ser Lys Asn  
 50 55 60

Thr Arg Ser Arg Lys Gly Val Ser Val Lys Asp Val Ser Asn His Lys

65                      70                      75                      80  
 Ile Thr Glu Phe Phe Gln Val Arg Arg Ser Asn Arg Lys Thr Ser Lys  
                                          85                      90                      95

Gln Ile Ser Asp Glu Ala Lys His Ala Leu Arg Asp Thr Val Leu Lys  
                                          100                      105                      110

Gly Thr Asn Glu Arg Leu Leu Glu Val Tyr Lys Asp Val Val Lys Gly  
                                          115                      120                      125

Arg Gly Ile Arg Thr Lys Val Asn Phe Glu Lys Gly Asp Phe Val Val  
                                          130                      135                      140

Glu Tyr Arg Gly Val Met Met Glu Tyr Ser Glu Ala Lys Val Ile Glu  
 145                      150                      155                      160

Glu Gln Tyr Ser Asn Asp Glu Glu Ile Gly Ser Tyr Met Tyr Phe Phe  
                                          165                      170                      175

Glu His Asn Asn Lys Lys Trp Cys Ile Asp Ala Thr Lys Glu Ser Pro  
                                          180                      185                      190

Trp Lys Gly Arg Leu Ile Asn His Ser Val Leu Arg Pro Asn Leu Lys  
                                          195                      200                      205

Thr Lys Val Val Glu Ile Asp Gly Ser His His Leu Ile Leu Val Ala  
                                          210                      215                      220

Arg Arg Gln Ile Ala Gln Gly Glu Glu Leu Leu Tyr Asp Tyr Gly Asp  
 225                      230                      235                      240

Arg Ser Ala Glu Thr Ile Ala Lys Asn Pro Trp Leu Val Asn Thr Glu  
                                          245                      250                      255

Phe

<210> 195

<211> 310

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 195

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Val Arg Thr

1                    5                    10                    15

Glu Lys Ile Ile Cys Arg Asp Val Ala Arg Gly Tyr Glu Asn Val Pro

20                    25                    30

Ile Pro Cys Val Asn Gly Val Asp Gly Glu Pro Cys Pro Glu Asp Tyr

35                    40                    45

Lys Tyr Ile Ser Glu Asn Cys Glu Thr Ser Thr Met Asn Ile Asp Arg

50                    55                    60

Asn Ile Thr His Leu Gln His Cys Thr Cys Val Asp Asp Cys Ser Ser

65                    70                    75                    80

Ser Asn Cys Leu Cys Gly Gln Leu Ser Ile Arg Cys Trp Tyr Asp Lys

85                    90                    95

Asp Gly Arg Leu Leu Gln Glu Phe Asn Lys Ile Glu Pro Pro Leu Ile

100                    105                    110

Phe Glu Cys Asn Gln Ala Cys Ser Cys Trp Arg Ser Cys Lys Asn Arg

115                    120                    125

Val Val Gln Ser Gly Ile Lys Val Arg Leu Gln Leu Tyr Arg Thr Ala

130                    135                    140

Lys Met Gly Trp Gly Val Arg Ala Leu Gln Thr Ile Pro Gln Gly Thr

145                    150                    155                    160

Phe Ile Cys Glu Tyr Val Gly Glu Leu Ile Ser Asp Ala Glu Ala Asp

165                    170                    175

Val Arg Glu Asp Asp Ser Tyr Leu Phe Asp Leu Asp Asn Lys Asp Gly

180                    185                    190

Glu Val Tyr Cys Ile Asp Ala Arg Tyr Tyr Gly Asn Ile Ser Arg Phe

195                    200                    205

Ile Asn His Leu Cys Asp Pro Asn Ile Ile Pro Val Arg Val Phe Met

210                    215                    220

Leu His Gln Asp Leu Arg Phe Pro Arg Ile Ala Phe Phe Ser Ser Arg

225                    230                    235                    240

Asp Ile Arg Thr Gly Glu Glu Leu Gly Phe Asp Tyr Gly Asp Arg Phe  
 245 250 255

Trp Asp Ile Lys Ser Lys Tyr Phe Thr Cys Gln Cys Gly Ser Glu Lys  
 260 265 270

Cys Lys His Ser Ala Glu Ala Ile Ala Leu Glu Gln Ser Arg Leu Ala  
 275 280 285

Arg Leu Asp Pro His Pro Glu Leu Leu Pro Asp Leu Ser Ser Leu Pro  
 290 295 300

Pro Ile Asn Thr Glu Phe  
 305 310

<210> 196  
 <211> 289  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 196

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Ser Ser Ser  
 1 5 10 15

Met Ala Asp Phe Arg Lys Phe Phe Ala Lys Ala Lys His Ile Val Ile  
 20 25 30

Ile Ser Gly Ala Gly Val Ser Ala Glu Ser Gly Val Pro Thr Phe Arg  
 35 40 45

Gly Ala Gly Gly Tyr Trp Arg Lys Trp Gln Ala Gln Asp Leu Ala Thr  
 50 55 60

Pro Leu Ala Phe Ala His Asn Pro Ser Arg Val Trp Glu Phe Tyr His  
 65 70 75 80

Tyr Arg Arg Glu Val Met Gly Ser Lys Glu Pro Asn Ala Gly His Arg  
 85 90 95

Ala Ile Ala Glu Cys Glu Thr Arg Leu Gly Lys Gln Gly Arg Arg Val  
 100 105 110

Val Val Ile Thr Gln Asn Ile Asp Glu Leu His Arg Lys Ala Gly Thr  
 115 120 125

Lys Asn Leu Leu Glu Ile His Gly Ser Leu Phe Lys Thr Arg Cys Thr  
 130 135 140

Ser Cys Gly Val Val Ala Glu Asn Tyr Lys Ser Pro Ile Cys Pro Ala  
 145 150 155 160

Leu Ser Gly Lys Gly Ala Pro Glu Pro Gly Thr Gln Asp Ala Ser Ile  
 165 170 175

Pro Val Glu Lys Leu Pro Arg Cys Glu Glu Ala Gly Cys Gly Gly Leu  
 180 185 190

Leu Arg Pro His Val Val Trp Phe Gly Glu Asn Leu Asp Pro Ala Ile  
 195 200 205

Leu Glu Glu Val Asp Arg Glu Leu Ala His Cys Asp Leu Cys Leu Val  
 210 215 220

Val Gly Thr Ser Ser Val Val Tyr Pro Ala Ala Met Phe Ala Pro Gln  
 225 230 235 240

Val Ala Ala Arg Gly Val Pro Val Ala Glu Phe Asn Thr Glu Thr Thr  
 245 250 255

Pro Ala Thr Asn Arg Phe Arg Phe His Phe Gln Gly Pro Cys Gly Thr  
 260 265 270

Thr Leu Pro Glu Ala Leu Ala Cys His Glu Asn Glu Thr Val Ser Glu  
 275 280 285

Phe

<210> 197

<211> 340

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 197

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Ser Arg

1                    5                    10                    15

Val Val Lys Pro Lys Val Ala Ser Met Glu Glu Met Ala Ala Phe His

                  20                    25                    30

Thr Asp Ala Tyr Leu Gln His Leu His Lys Val Ser Glu Glu Gly Asp

                  35                    40                    45

Asn Asp Asp Pro Glu Thr Leu Glu Tyr Gly Leu Gly Tyr Asp Cys Pro

                  50                    55                    60

Ile Thr Glu Gly Ile Tyr Asp Tyr Ala Ala Ala Val Gly Gly Ala Thr

65                    70                    75                    80

Leu Thr Ala Ala Glu Gln Leu Ile Glu Gly Lys Thr Arg Ile Ala Val

                  85                    90                    95

Asn Trp Pro Gly Gly Trp His His Ala Lys Lys Asp Glu Ala Ser Gly

                  100                    105                    110

Phe Cys Tyr Leu Asn Asp Ala Val Leu Gly Ile Leu Lys Leu Arg Glu

                  115                    120                    125

Lys Phe Asp Arg Val Leu Tyr Val Asp Met Asp Leu His His Gly Asp

130                    135                    140

Gly Val Glu Asp Ala Phe Ser Phe Thr Ser Lys Val Met Thr Val Ser

145                    150                    155                    160

Leu His Lys Phe Ser Pro Gly Phe Phe Pro Gly Thr Gly Asp Val Ser

                  165                    170                    175

Asp Ile Gly Leu Gly Lys Gly Arg Tyr Tyr Ser Ile Asn Val Pro Leu

                  180                    185                    190

Gln Asp Gly Ile Gln Asp Asp Lys Tyr Tyr Gln Ile Cys Glu Gly Val

195                    200                    205

Leu Lys Glu Val Phe Thr Thr Phe Asn Pro Glu Ala Val Val Leu Gln

210                    215                    220

Leu Gly Ala Asp Thr Ile Ala Gly Asp Pro Met Cys Ser Phe Asn Met

225                    230                    235                    240

Thr Pro Glu Gly Ile Gly Lys Cys Leu Lys Tyr Val Leu Gln Trp Gln

245 250 255  
 Leu Pro Thr Leu Ile Leu Gly Gly Gly Tyr His Leu Pro Asn Thr

260 265 270  
 Ala Arg Cys Trp Thr Tyr Leu Thr Ala Leu Ile Val Gly Arg Thr Leu

275 280 285  
 Ser Ser Glu Ile Pro Asp His Glu Phe Phe Thr Glu Tyr Gly Pro Asp

290 295 300  
 Tyr Val Leu Glu Ile Thr Pro Ser Cys Arg Pro Asp Arg Asn Asp Thr

305 310 315 320  
 Gln Lys Val Gln Glu Ile Leu Gln Ser Ile Lys Gly Asn Leu Lys Arg

325 330 335  
 Val Val Glu Phe

340

<210> 198

<211> 134

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 198

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Lys Glu  
 1 5 10 15

Gly Glu Asn Asn Lys Pro Arg Glu Lys Ser Glu Gly Asn Lys Arg Lys  
 20 25 30

Ser Ser Phe Ser Asn Ser Ala Asp Asp Ile Lys Ser Lys Lys Lys Arg  
 35 40 45

Glu Gln Ser Asn Asp Ile Ala Arg Gly Phe Glu Arg Gly Leu Glu Pro  
 50 55 60

Glu Lys Ile Ile Gly Ala Thr Asp Ser Cys Gly Asp Leu Met Phe Leu  
 65 70 75 80

Met Lys Trp Lys Asp Thr Asp Glu Ala Asp Leu Val Leu Ala Lys Glu

85 90 95

Ala Asn Val Lys Cys Pro Gln Ile Val Ile Ala Phe Tyr Glu Glu Arg

100 105 110

Leu Thr Trp His Ala Tyr Pro Glu Asp Ala Glu Asn Lys Glu Lys Glu

115 120 125

Ser Ala Lys Ser Glu Phe

130

<210> 199

<211> 226

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 199

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Glu Phe

1 5 10 15

Trp Gly Ile Glu Val Lys Ser Gly Lys Pro Val Thr Val Thr Pro Glu

20 25 30

Glu Gly Ile Leu Ile His Val Ser Gln Ala Ser Leu Gly Glu Cys Lys

35 40 45

Asn Lys Lys Gly Glu Phe Val Pro Leu His Val Lys Val Gly Asn Gln

50 55 60

Asn Leu Val Leu Gly Thr Leu Ser Thr Glu Asn Ile Pro Gln Leu Phe

65 70 75 80

Cys Asp Leu Val Phe Asp Lys Glu Phe Glu Leu Ser His Thr Trp Gly

85 90 95

Lys Gly Ser Val Tyr Phe Val Gly Tyr Lys Thr Pro Asn Ile Glu Pro

100 105 110

Gln Gly Tyr Ser Glu Glu Glu Glu Glu Glu Glu Val Pro Ala

115 120 125

Gly Asn Ala Ala Lys Ala Val Ala Lys Pro Lys Ala Lys Pro Ala Glu

130 135 140

Val Lys Pro Ala Val Asp Asp Glu Glu Asp Glu Ser Asp Ser Asp Gly

145 150 155 160

Met Asp Glu Asp Asp Ser Asp Gly Glu Asp Ser Glu Glu Glu Glu Pro

165 170 175

Thr Pro Lys Lys Pro Ala Ser Ser Lys Lys Arg Ala Asn Glu Thr Thr

180 185 190

Pro Lys Ala Pro Val Ser Ala Lys Lys Ala Lys Val Ala Val Thr Pro

195 200 205

Gln Lys Thr Asp Glu Lys Lys Lys Gly Gly Lys Ala Ala Asn Gln Ser

210 215 220

Glu Phe

225

<210> 200

<211> 108

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 200

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Ser Arg

1 5 10 15

Arg Lys Gln Ala Lys Pro Gln His Phe Gln Ser Asp Pro Glu Val Ala

20 25 30

Ser Leu Pro Arg Arg Asp Gly Asp Thr Glu Lys Gly Gln Pro Ser Arg

35 40 45

Pro Thr Lys Ser Lys Asp Ala His Val Cys Gly Arg Cys Cys Ala Glu

50 55 60

Phe Phe Glu Leu Ser Asp Leu Leu Leu His Lys Lys Ser Cys Thr Lys

65                                 70                                 75                                 80  
  
 Asn Gln Leu Val Leu Ile Val Asn Glu Ser Pro Ala Ser Pro Ala Lys  
  
                                   85                                 90                                 95  
 Thr Phe Pro Pro Gly Pro Ser Leu Asn Asp Glu Phe  
  
                                   100                                 105  
  
 <210> 201  
 <211> 224  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
                                   polypeptide"  
 <400> 201  
 Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Ser Cys Asp  
 1                                 5                                 10                                 15  
  
  
 Ser Thr Asn Ala Ala Ile Ala Lys Gln Ala Leu Lys Lys Pro Ile Lys  
  
                                   20                                 25                                 30  
 Gly Lys Gln Ala Pro Arg Lys Lys Ala Gln Gly Lys Thr Gln Gln Asn  
  
                                   35                                 40                                 45  
 Arg Lys Leu Thr Asp Phe Tyr Pro Val Arg Arg Ser Ser Arg Lys Ser  
  
                                   50                                 55                                 60  
 Lys Ala Glu Leu Gln Ser Glu Glu Arg Lys Arg Ile Asp Glu Leu Ile  
 65                                 70                                 75                                 80  
  
  
 Glu Ser Gly Lys Glu Glu Gly Met Lys Ile Asp Leu Ile Asp Gly Lys  
  
                                   85                                 90                                 95  
 Gly Arg Gly Val Ile Ala Thr Lys Gln Phe Ser Arg Gly Asp Phe Val  
  
                                   100                                 105                                 110  
 Val Glu Tyr His Gly Asp Leu Ile Glu Ile Thr Asp Ala Lys Lys Arg  
  
                                   115                                 120                                 125  
 Glu Ala Leu Tyr Ala Gln Asp Pro Ser Thr Gly Cys Tyr Met Tyr Tyr  
  
                                   130                                 135                                 140

Phe Gln Tyr Leu Ser Lys Thr Tyr Cys Val Asp Ala Thr Arg Glu Thr  
 145                      150                      155                      160  
 Asn Arg Leu Gly Arg Leu Ile Asn His Ser Lys Cys Gly Asn Cys Gln  
                                  165                      170                      175  
 Thr Lys Leu His Asp Ile Asp Gly Val Pro His Leu Ile Leu Ile Ala  
                                  180                      185                      190  
 Ser Arg Asp Ile Ala Ala Gly Glu Glu Leu Leu Tyr Asp Tyr Gly Asp  
                                  195                      200                      205

Arg Ser Lys Ala Ser Ile Glu Ala Phe Pro Trp Leu Lys His Glu Phe  
                                  210                      215                      220

<210> 202

<211> 337

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 202

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Arg Arg Val  
 1                      5                      10                      15  
 Ala Tyr Phe Tyr Asp Ala Asp Val Gly Asn Tyr Ala Tyr Gly Ala Gly  
                                  20                      25                      30

His Pro Met Lys Pro His Arg Ile Arg Met Ala His Ser Leu Ile Met  
                                  35                      40                      45

Asn Tyr Gly Leu Tyr Lys Lys Met Glu Ile Tyr Arg Ala Lys Pro Ala  
                                  50                      55                      60

Thr Lys Gln Glu Met Cys Gln Phe His Thr Asp Glu Tyr Ile Asp Phe  
 65                      70                      75                      80

Leu Ser Arg Val Thr Pro Asp Asn Leu Glu Met Phe Lys Arg Glu Ser  
                                  85                      90                      95

Val Lys Phe Asn Val Gly Asp Asp Cys Pro Val Phe Asp Gly Leu Tyr  
                                  100                      105                      110

Glu Tyr Cys Ser Ile Ser Gly Gly Gly Ser Met Glu Gly Ala Ala Arg  
 115 120 125  
 Leu Asn Arg Gly Lys Cys Asp Val Ala Val Asn Tyr Ala Gly Gly Leu  
 130 135 140  
 His His Ala Lys Lys Ser Glu Ala Ser Gly Phe Cys Tyr Leu Asn Asp  
 145 150 155 160  
  
 Ile Val Leu Gly Ile Ile Glu Leu Leu Arg Tyr His Pro Arg Val Leu  
 165 170 175  
 Tyr Ile Asp Ile Asp Val His His Gly Asp Gly Val Glu Glu Ala Phe  
 180 185 190  
 Tyr Thr Thr Asp Arg Val Met Thr Cys Ser Phe His Lys Tyr Gly Glu  
 195 200 205  
 Phe Phe Pro Gly Thr Gly Glu Leu Arg Asp Ile Gly Val Gly Ala Gly  
 210 215 220  
  
 Lys Asn Tyr Ala Val Asn Val Pro Leu Arg Asp Gly Ile Asp Asp Ala  
 225 230 235 240  
 Thr Tyr Arg Ser Val Phe Glu Pro Val Ile Lys Lys Ile Met Glu Trp  
 245 250 255  
 Tyr Gln Pro Ser Ala Val Val Leu Gln Cys Gly Gly Asp Ser Leu Ser  
 260 265 270  
 Gly Asp Arg Leu Gly Cys Phe Asn Leu Ser Met Glu Gly His Ala Asn  
 275 280 285  
  
 Cys Val Asn Tyr Val Lys Ser Phe Gly Ile Pro Met Met Val Val Gly  
 290 295 300  
 Gly Gly Gly Tyr Thr Met Arg Asn Val Ala Arg Thr Trp Cys Phe Glu  
 305 310 315 320  
 Thr Gly Leu Leu Asn Asn Val Val Leu Asp Lys Asp Leu Pro Tyr Glu  
 325 330 335  
 Phe

<210> 203

<211> 257

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 203

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Glu Lys

1                    5                    10                    15

Pro Arg Val Leu Val Leu Thr Gly Ala Gly Ile Ser Ala Glu Ser Gly

                  20                    25                    30

Ile Arg Thr Phe Arg Ala Ala Asp Gly Leu Trp Glu Glu His Arg Val

                  35                    40                    45

Glu Asp Val Ala Thr Pro Glu Gly Phe Asp Arg Asp Pro Glu Leu Val

50                    55                    60

Gln Ala Phe Tyr Asn Ala Arg Arg Arg Gln Leu Gln Gln Pro Glu Ile

65                    70                    75                    80

Gln Pro Asn Ala Ala His Leu Ala Leu Ala Lys Leu Gln Asp Ala Leu

                  85                    90                    95

Gly Asp Arg Phe Leu Leu Val Thr Gln Asn Ile Asp Asn Leu His Glu

                  100                    105                    110

Arg Ala Gly Asn Thr Asn Val Ile His Met His Gly Glu Leu Leu Lys

                  115                    120                    125

Val Arg Cys Ser Gln Ser Gly Gln Val Leu Asp Trp Thr Gly Asp Val

                  130                    135                    140

Thr Pro Glu Asp Lys Cys His Cys Cys Gln Phe Pro Ala Pro Leu Arg

145                    150                    155                    160

Pro His Val Val Trp Phe Gly Glu Met Pro Leu Gly Met Asp Glu Ile

                  165                    170                    175

Tyr Met Ala Leu Ser Met Ala Asp Ile Phe Ile Ala Ile Gly Thr Ser

                  180                    185                    190

Gly His Val Tyr Pro Ala Ala Gly Phe Val His Glu Ala Lys Leu His

195                      200                      205  
 Gly Ala His Thr Val Glu Leu Asn Leu Glu Pro Ser Gln Val Gly Asn  
 210                      215                      220  
 Glu Phe Ala Glu Lys Tyr Tyr Gly Pro Ala Ser Gln Val Val Pro Glu  
 225                      230                      235                      240  
 Phe Val Glu Lys Leu Leu Lys Gly Leu Lys Ala Gly Ser Ile Ala Glu  
 245                      250                      255

Phe

<210> 204

<211> 349

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 204

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Asn Leu Lys  
 1                      5                      10                      15  
 Cys Val Arg Ile Leu Lys Gln Phe His Lys Asp Leu Glu Arg Glu Leu  
 20                      25                      30  
 Leu Arg Arg His His Arg Ser Lys Thr Pro Arg His Leu Asp Pro Ser  
 35                      40                      45  
 Leu Ala Asn Tyr Leu Val Gln Lys Ala Lys Gln Arg Arg Ala Leu Arg  
 50                      55                      60  
 Arg Trp Glu Gln Glu Leu Asn Ala Lys Arg Ser His Leu Gly Arg Ile  
 65                      70                      75                      80  
 Thr Val Glu Asn Glu Val Asp Leu Asp Gly Pro Pro Arg Ala Phe Val  
 85                      90                      95  
 Tyr Ile Asn Glu Tyr Arg Val Gly Glu Gly Ile Thr Leu Asn Gln Val  
 100                      105                      110  
 Ala Val Gly Cys Glu Cys Gln Asp Cys Leu Trp Ala Pro Thr Gly Gly

115                            120                            125  
 Cys Cys Pro Gly Ala Ser Leu His Lys Phe Ala Tyr Asn Asp Gln Gly  
 130                            135                            140  
 Gln Val Arg Leu Arg Ala Gly Leu Pro Ile Tyr Glu Cys Asn Ser Arg  
 145                            150                            155                            160  
 Cys Arg Cys Gly Tyr Asp Cys Pro Asn Arg Val Val Gln Lys Gly Ile  
  
 165                            170                            175  
 Arg Tyr Asp Leu Cys Ile Phe Arg Thr Asp Asp Gly Arg Gly Trp Gly  
 180                            185                            190  
 Val Arg Thr Leu Glu Lys Ile Arg Lys Asn Ser Phe Val Met Glu Tyr  
 195                            200                            205  
 Val Gly Glu Ile Ile Thr Ser Glu Glu Ala Glu Arg Arg Gly Gln Ile  
 210                            215                            220  
 Tyr Asp Arg Gln Gly Ala Thr Tyr Leu Phe Asp Leu Asp Tyr Val Glu  
  
 225                            230                            235                            240  
 Asp Val Tyr Thr Val Asp Ala Ala Tyr Tyr Gly Asn Ile Ser His Phe  
 245                            250                            255  
 Val Asn His Ser Cys Asp Pro Asn Leu Gln Val Tyr Asn Val Phe Ile  
 260                            265                            270  
 Asp Asn Leu Asp Glu Arg Leu Pro Arg Ile Ala Phe Phe Ala Thr Arg  
 275                            280                            285  
 Thr Ile Arg Ala Gly Glu Glu Leu Thr Phe Asp Tyr Asn Met Gln Val  
  
 290                            295                            300  
 Asp Pro Val Asp Met Glu Ser Thr Arg Met Asp Ser Asn Phe Gly Leu  
 305                            310                            315                            320  
 Ala Gly Leu Pro Gly Ser Pro Lys Lys Arg Val Arg Ile Glu Cys Lys  
 325                            330                            335  
 Cys Gly Thr Glu Ser Cys Arg Lys Tyr Leu Phe Glu Phe  
 340                            345  
 <210> 205  
 <211> 235  
 <212> PRT

<213> Artificial Sequence

<220><221

> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 205

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Ser Asn Ser  
 1                    5                    10                    15  
 Trp Glu Glu Gly Ser Ser Gly Ser Ser Ser Asp Glu Glu His Gly Gly  
                   20                    25                    30  
 Gly Gly Met Arg Val Gly Pro Gln Tyr Gln Ala Val Val Pro Asp Phe  
                   35                    40                    45  
 Asp Pro Ala Lys Leu Ala Arg Arg Ser Gln Glu Arg Asp Asn Leu Gly  
  
                   50                    55                    60  
 Met Leu Val Trp Ser Pro Asn Gln Asn Leu Ser Glu Ala Lys Leu Asp  
 65                    70                    75                    80  
 Glu Tyr Ile Ala Ile Ala Lys Glu Lys His Gly Tyr Asn Met Glu Gln  
                   85                    90                    95  
 Ala Leu Gly Met Leu Phe Trp His Lys His Asn Ile Glu Lys Ser Leu  
                   100                    105                    110  
 Ala Asp Leu Pro Asn Phe Thr Pro Phe Pro Asp Glu Trp Thr Val Glu  
  
                   115                    120                    125  
 Asp Lys Val Leu Phe Glu Gln Ala Phe Ser Phe His Gly Lys Thr Phe  
                   130                    135                    140  
 His Arg Ile Gln Gln Met Leu Pro Asp Lys Ser Ile Ala Ser Leu Val  
 145                    150                    155                    160  
 Lys Phe Tyr Tyr Ser Trp Lys Lys Thr Arg Thr Lys Thr Ser Val Met  
                   165                    170                    175  
 Asp Arg His Ala Arg Lys Gln Lys Arg Glu Arg Glu Glu Ser Glu Asp  
  
                   180                    185                    190  
 Glu Leu Glu Glu Ala Asn Gly Asn Asn Pro Ile Asp Ile Glu Val Asp  
                   195                    200                    205

Gln Asn Lys Glu Ser Lys Lys Glu Val Pro Pro Thr Glu Thr Val Pro  
 210 215 220

Gln Val Lys Lys Glu Lys His Ser Thr Glu Phe  
 225 230 235

<210> 206  
 <211> 343  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223>  
 > /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 206

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Tyr Lys Glu  
 1 5 10 15  
 Ser Val His Leu Glu Thr Tyr Pro Lys Glu Arg Ala Thr Glu Gly Ile  
 20 25 30  
 Ala Met Glu Ile Asp Tyr Ala Ser Cys Lys Arg Leu Gly Ser Ser Tyr  
 35 40 45  
 Arg Ala Leu Pro Lys Ser Tyr Gln Gln Pro Lys Cys Thr Gly Arg Thr  
 50 55 60  
 Pro Leu Cys Lys Glu Val Leu Asn Asp Thr Trp Val Ser Phe Pro Ser  
 65 70 75 80  
 Trp Ser Glu Asp Ser Thr Phe Val Ser Ser Lys Lys Thr Gln Tyr Glu  
 85 90 95  
 Glu His Ile Tyr Arg Cys Glu Asp Glu Arg Phe Glu Leu Asp Val Val  
 100 105 110  
 Leu Glu Thr Asn Leu Ala Thr Ile Arg Val Leu Glu Ala Ile Gln Lys  
 115 120 125  
 Lys Leu Ser Arg Leu Ser Ala Glu Glu Gln Ala Lys Phe Arg Leu Asp  
 130 135 140  
 Asn Thr Leu Gly Gly Thr Ser Glu Val Ile His Arg Lys Ala Leu Gln  
 145 150 155 160

Arg Ile Tyr Ala Asp Lys Ala Ala Asp Ile Ile Asp Gly Leu Arg Lys  
 165 170 175  
 Asn Pro Ser Ile Ala Val Pro Ile Val Leu Lys Arg Leu Lys Met Lys  
 180 185 190  
 Glu Glu Glu Trp Arg Glu Ala Gln Arg Gly Phe Asn Lys Val Trp Arg  
 195 200 205  
 Glu Gln Asn Glu Lys Tyr Tyr Leu Lys Ser Leu Asp His Gln Gly Ile  
 210 215 220  
 Asn Phe Lys Gln Asn Asp Thr Lys Val Leu Arg Ser Lys Ser Leu Leu  
 225 230 235 240  
 Asn Glu Ile Glu Ser Ile Tyr Asp Glu Arg Gln Glu Gln Ala Thr Glu  
 245 250 255  
 Glu Asn Ala Gly Val Pro Val Gly Pro His Leu Ser Leu Ala Tyr Glu  
 260 265 270  
 Asp Lys Gln Ile Leu Glu Asp Ala Ala Ala Leu Ile Ile His His Val  
 275 280 285  
 Lys Arg Gln Thr Gly Ile Gln Lys Glu Asp Lys Tyr Lys Ile Lys Gln  
 290 295 300  
 Ile Met His His Phe Ile Pro Asp Leu Leu Phe Ala Gln Arg Gly Asp  
 305 310 315 320  
 Leu Ser Asp Val Glu Glu Glu Glu Glu Glu Met Asp Val Asp Glu  
 325 330 335  
 Ala Thr Gly Ala Val Glu Phe  
 340  
 <210> 207  
 <211> 328  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 207

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Gln Ser Ala  
 1 5 10 15

Tyr Leu His Val Ser Leu Ala Arg Ile Ser Asp Glu Asp Cys Cys Ala  
 20 25 30

Asn Cys Lys Gly Asn Cys Leu Ser Ala Asp Phe Pro Cys Thr Cys Ala  
 35 40 45

Arg Glu Thr Ser Gly Glu Tyr Ala Tyr Thr Lys Glu Gly Leu Leu Lys  
 50 55 60

Glu Lys Phe Leu Asp Thr Cys Leu Lys Met Lys Lys Glu Pro Asp Ser  
 65 70 75 80

Phe Pro Lys Val Tyr Cys Lys Asp Cys Pro Leu Glu Arg Asp His Asp  
 85 90 95

Lys Gly Thr Tyr Gly Lys Cys Asp Gly His Leu Ile Arg Lys Phe Ile  
 100 105 110

Lys Glu Cys Trp Arg Lys Cys Gly Cys Asp Met Gln Cys Gly Asn Arg  
 115 120 125

Val Val Gln Arg Gly Ile Arg Cys Gln Leu Gln Val Tyr Phe Thr Gln  
 130 135 140

Glu Gly Lys Gly Trp Gly Leu Arg Thr Leu Gln Asp Leu Pro Lys Gly  
 145 150 155 160

Thr Phe Ile Cys Glu Tyr Ile Gly Glu Ile Leu Thr Asn Thr Glu Leu  
 165 170 175

Tyr Asp Arg Asn Val Arg Ser Ser Ser Glu Arg His Thr Tyr Pro Val  
 180 185 190

Thr Leu Asp Ala Asp Trp Gly Ser Glu Lys Asp Leu Lys Asp Glu Glu  
 195 200 205

Ala Leu Cys Leu Asp Ala Thr Ile Cys Gly Asn Val Ala Arg Phe Ile  
 210 215 220

Asn His Arg Cys Glu Asp Ala Asn Met Ile Asp Ile Pro Ile Glu Ile  
 225 230 235 240

Glu Thr Pro Asp Arg His Tyr Tyr His Ile Ala Phe Phe Thr Leu Arg

245 250 255  
 Asp Val Lys Ala Met Asp Glu Leu Thr Trp Asp Tyr Met Ile Asp Phe  
 260 265 270

Asn Asp Lys Ser His Pro Val Lys Ala Phe Arg Cys Cys Cys Gly Ser  
 275 280 285

Glu Ser Cys Arg Asp Arg Lys Ile Lys Gly Ser Gln Gly Lys Ser Ile  
 290 295 300

Glu Arg Arg Lys Ile Val Ser Ala Lys Lys Gln Gln Gly Ser Lys Glu  
 305 310 315 320

Val Ser Lys Lys Arg Lys Glu Phe  
 325

<210> 208

<211> 301

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 208

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Val Gln Val  
 1 5 10 15

Lys Arg Val Leu Glu Lys Ser Pro Gly Lys Leu Leu Val Lys Met Pro  
 20 25 30

Phe Gln Ala Ser Pro Gly Gly Lys Gly Glu Gly Gly Gly Ala Thr Thr  
 35 40 45

Ser Ala Gln Val Met Val Ile Lys Arg Pro Gly Arg Lys Arg Lys Ala  
 50 55 60

Glu Ala Asp Pro Gln Ala Ile Pro Lys Lys Arg Gly Arg Lys Pro Gly  
 65 70 75 80

Ser Val Val Ala Ala Ala Ala Glu Ala Lys Lys Lys Ala Val Lys  
 85 90 95

Glu Ser Ser Ile Arg Ser Val Gln Glu Thr Val Leu Pro Ile Lys Lys



polypeptide"

<400> 209

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Ala Pro

1                    5                    10                    15

Lys Arg Gln Ser Ala Ile Leu Pro Gln Pro Lys Lys Pro Arg Pro Ala

                  20                    25                    30

Ala Ala Pro Lys Leu Glu Asp Lys Ser Ala Ser Pro Gly Leu Pro Lys

                  35                    40                    45

Gly Glu Lys Glu Gln Gln Glu Ala Ile Glu His Ile Asp Glu Val Gln

                  50                    55                    60

Asn Glu Ile Asp Arg Leu Asn Glu Gln Ala Ser Glu Glu Ile Leu Lys

65                    70                    75                    80

Val Glu Gln Lys Tyr Asn Lys Leu Arg Gln Pro Phe Phe Gln Lys Arg

                  85                    90                    95

Ser Glu Leu Ile Ala Lys Ile Pro Asn Phe Trp Val Thr Thr Phe Val

                  100                    105                    110

Asn His Pro Gln Val Ser Ala Leu Leu Gly Glu Glu Asp Glu Glu Ala

                  115                    120                    125

Leu His Tyr Leu Thr Arg Val Glu Val Thr Glu Phe Glu Asp Ile Lys

                  130                    135                    140

Ser Gly Tyr Arg Ile Asp Phe Tyr Phe Asp Glu Asn Pro Tyr Phe Glu

145                    150                    155                    160

Asn Lys Val Leu Ser Lys Glu Phe His Leu Asn Glu Ser Gly Asp Pro

                  165                    170                    175

Ser Ser Lys Ser Thr Glu Ile Lys Trp Lys Ser Gly Lys Asp Leu Thr

                  180                    185                    190

Lys Arg Ser Ser Gln Thr Gln Asn Lys Ala Ser Arg Lys Arg Gln His

                  195                    200                    205

Glu Glu Pro Glu Ser Phe Phe Thr Trp Phe Thr Asp His Ser Asp Ala

                  210                    215                    220

Gly Ala Asp Glu Leu Gly Glu Val Ile Lys Asp Asp Ile Trp Pro Asn

225                            230                            235                            240

Pro Leu Gln Tyr Tyr Leu Val Pro Asp Met Asp Asp Glu Glu Gly Glu

                                 245                            250                            255

Ala Glu Asp Asp Asp Asp Asp Asp Glu Glu Glu Gly Leu Glu Asp

                                 260                            265                            270

Ile Asp Glu Glu Gly Asp Glu Asp Glu Gly Glu Glu Asp Asp Asp Glu

                                 275                            280                            285

Asp Glu Gly Glu Glu Gly Glu Glu Asp Glu Gly Glu Asp Asp Glu Phe

                                 290                            295                            300

<210> 210

<211> 303

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 210

Ala Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Met Gln Leu

1                            5                            10                            15

His Glu Gln Ile Ala Asn Ile Ser Val Thr Phe Asn Asp Ile Pro Arg

                                 20                            25                            30

Ser Asp His Ser Met Thr Pro Thr Glu Leu Cys Tyr Phe Asp Asp Phe

                                 35                            40                            45

Ala Thr Thr Leu Val Val Asp Ser Val Leu Asn Phe Thr Thr His Lys

                                 50                            55                            60

Met Ser Lys Lys Arg Arg Tyr Leu Tyr Gln Asp Glu Tyr Arg Thr Ala

65                            70                            75                            80

Arg Thr Val Met Lys Thr Phe Arg Glu Gln Arg Asp Trp Thr Asn Ala

                                 85                            90                            95

Ile Tyr Gly Leu Leu Thr Leu Arg Ser Val Ser His Phe Leu Ser Lys

                                 100                            105                            110

Leu Pro Pro Asn Lys Leu Phe Glu Phe Arg Asp His Ile Val Arg Phe  
 115 120 125  
 Leu Asn Met Phe Ile Leu Asp Ser Gly Tyr Thr Ile Gln Glu Cys Lys  
 130 135 140  
 Arg Tyr Ser Gln Glu Gly His Gln Gly Ala Lys Leu Val Ser Thr Gly  
 145 150 155 160  
 Val Trp Ser Arg Gly Asp Lys Ile Glu Arg Leu Ser Gly Val Val Cys  
 165 170 175

Leu Leu Ser Ser Glu Asp Glu Asp Ser Ile Leu Ala Gln Glu Gly Ser  
 180 185 190  
 Asp Phe Ser Val Met Tyr Ser Thr Arg Lys Arg Cys Ser Thr Leu Trp  
 195 200 205  
 Leu Gly Pro Gly Ala Tyr Ile Asn His Asp Cys Arg Pro Thr Cys Glu  
 210 215 220  
 Phe Val Ser His Gly Ser Thr Ala His Ile Arg Val Leu Arg Asp Met  
 225 230 235 240

Val Pro Gly Asp Glu Ile Thr Cys Phe Tyr Gly Ser Glu Phe Phe Gly  
 245 250 255  
 Pro Asn Asn Ile Asp Cys Glu Cys Cys Thr Cys Glu Lys Asn Met Asn  
 260 265 270  
 Gly Ala Phe Ser Tyr Leu Arg Gly Asn Glu Asn Ala Glu Pro Ile Ile  
 275 280 285  
 Ser Glu Lys Lys Thr Lys Tyr Glu Leu Arg Ser Arg Ser Glu Phe  
 290 295 300

<210> 211  
 <211> 14164  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polynucleotide"  
 <400> 211

caactttgta tagaaaagtt ggctccgaat tcgcccttca ggtccgaggt tctagacgag 60  
 ttactccct atcagtata gagaacgat tcgagtttac tccctatcag tgatagagaa 120  
 cgatgtcga gtttactccc tatcagtgat agagaacgta tgtcgagttt actccctatc 180  
 agtgatagag aacgtatgtc gagtttatcc ctatcagtga tagagaacgt atgtcgagtt 240  
 tactccctat cagtgataga gaacgtatgt cgaggtaggc gtgtacggtg ggaggcctat 300  
  
 ataagcagag ctcgtttagt gaaccgtcag atcgcaaagg gcgaattcga cccaagtttg 360  
 tacagccacc atggactata aggaccacga cggagactac aaggatcatg atattgatta 420  
 caaagacgat gacgataaga tggcccaaaa gaagaagcgg aaggtcggta tccacggagt 480  
 cccagcagcc gacaagaagt acagcatcgg cctggacatc ggcaccaact ctgtgggctg 540  
 ggccgtgatc accgacgagt acaaggtgcc cagcaagaaa ttcaaggtgc tgggcaacac 600  
 cgaccggcac agcatcaaga agaacctgat cggagccctg ctgttcgaca gggcgaaac 660  
 agccgaggcc acccggctga agagaaccgc cagaagaaga tacaccagac ggaagaaccg 720  
  
 gatctgctat ctgcaagaga tcttcagcaa cgagatggcc aaggtggacg acagcttctt 780  
 ccacagactg gaagagtctt tcttggtgga agaggataag aagcacgagc ggcaccccat 840  
 cttcggcaac atcgtggacg aggtggccta ccacgagaag taccacacca tctaccacct 900  
 gagaaagaaa ctggtggaca gcaccgaaa ggccgacctg cggctgatct atctggcctt 960  
 ggcccacatg atcaagtcc ggggcccactt cctgatcgag ggcgacctga accccgaaa 1020  
 cagcgacgtg gacaagctgt tcatccagct ggtgcagacc tacaaccagc tgttcgagga 1080  
 aaaccccatc aacccagcg gcgtggacgc caaggccatc ctgtctgcca gactgagcaa 1140  
  
 gagcagacgg ctgaaaaatc tgatcgccca gctgcccggc gagaagaaga atggcctggt 1200  
 cggcaacctg attgccctga gcctgggcct gaccccaac ttcaagagca acttcgacct 1260  
 ggccgaggat gccaaactgc agctgagcaa ggacacctac gacgacgacc tggacaacct 1320  
 gctggcccag atcggcgacc agtaccccga cctgtttctg gccgccaaga acctgtccga 1380  
 cgccatcctg ctgagcgaca tcttgagagt gaacaccgag atcaccaagg cccccctgag 1440  
 cgctctatg atcaagagat acgacgagca ccaccaggac ctgacctgc tgaagctct 1500  
 cgtgcggcag cagctgcctg agaagtataa agagattttc ttcgaccaga gcaagaacgg 1560  
  
 ctacgccggc tacattgacg gcggagccag ccaggaagag ttctacaagt tcatcaagcc 1620  
 catcctgaa aagatggacg gcaccgagga actgctcgtg aagctgaaca gagaggacct 1680  
 gctgcggaag cagcggacct tcgacaacgg cagcatcccc caccagatcc acctgggaga 1740  
 gctgcacgcc attctgcggc ggcaggaaga tttttacca ttctgaagg acaaccggga 1800  
 aaagatcgag aagatcctga ccttcgcat ccctactac gtgggcctc tggccagggg 1860

aaacagcaga ttgcctgga tgaccagaaa gagcgaggaa accatcaccc cctggaactt 1920  
 cgaggaagtg gtggacaagg gcgcttccgc ccagagcttc atcgagcgga tgaccaactt 1980  
  
 cgataagaac ctgcccacg agaaggtgct gcccagcac agcctgctgt acgagtactt 2040  
 caccgtgtat aacgagctga ccaaagtga atacgtgacc gagggaatga gaaagcccgc 2100  
 cttcctgagc ggcgagcaga aaaaggccat cgtggacctg ctgttcaaga ccaaccggaa 2160  
 agtgaccgtg aagcagctga aagaggacta cttcaagaaa atcgagtgtc tcgactccgt 2220  
 ggaaatctcc ggcgtggaag atcggttcaa gcctcctg ggcacatacc acgatctgct 2280  
 gaaaattatc aaggacaagg acttctgga caatgaggaa aacgaggaca ttctggaaga 2340  
 tatcgtgctg accctgacac tgtttgagga cagagagatg atcgaggaac ggctgaaaac 2400  
  
 ctatgccac ctgttcgacg acaaagtgat gaagcagctg aagcggcgga gatacaccgg 2460  
 ctggggcagg ctgagccgga agctgatcaa cggcatccgg gacaagcagt ccggcaagac 2520  
 aatcctggat ttctgaagt cgcacggctt cgccaacaga aacttcatgc agctgatcca 2580  
 cgacgacagc ctgaccttta aagaggacat ccagaaagcc cagggtgtccg gccagggcga 2640  
 tagcctgcaac gagcacattg ccaatctggc cggcagcccc gccattaaga agggcatcct 2700  
 gcagacagtg aaggtgggtg acgagctcgt gaaagtgatg ggccggcaca agcccagaaa 2760  
 catcgtgatc gaaatggcca gagagaacca gaccaccag aaggacaga agaacagccg 2820  
  
 cgagagaatg aagcggatcg aagagggcat caaagagctg ggcagccaga tcctgaaaga 2880  
 acaccccgtg gaaaacaccc agctgcagaa cgagaagctg tactgtact acctgcagaa 2940  
 tgggcgggat atgtacgtgg accaggaact ggacatcaac cggetgtccg actacgatgt 3000  
 ggaccatata gtgcctcaga gctttctgaa ggacgactcc atcgacaaca aggtgtctgac 3060  
 cagaagcagc aagaaccggg gcaagagcga caactgtccc tccgaagagg tcgtgaagaa 3120  
 gatgaagaac tactggcggc agctgctgaa cgccaagctg attaccaga gaaagttcga 3180  
 caatctgacc aaggccgaga gagggcgcct gagcgaactg gataaggccg gcttcatcaa 3240  
  
 gagacagctg gtggaaaccc ggcagatcac aaagcacgtg gcacagatcc tggactcccg 3300  
 gatgaacact aagtacgacg agaatgacaa gctgatccgg gaagtgaaag tgatcacctt 3360  
 gaagtccaag ctggtgtccg atttccgaa ggatttccag ttttcaaag tgcgcgagat 3420  
 caacaactac caccacgcc acgacgccta cctgaacgcc gtcgtgggaa ccgacctgat 3480  
 caaaaagtac cctaaagctgg aaagcgagtt cgtgtacggc gactacaagg tgtacgacgt 3540  
 gcggaagatg atcgccaaga gcgagcagga aatcggaag gctaccgcca agtacttctt 3600

ctacagcaac atcatgaact ttttcaagac cgagattacc ctggccaacg gcgagatccg 3660

gaagcggcct ctgatcgaga caaacggcga aaccggggag atcgtgtggg ataagggccg 3720

ggattttgcc accgtgcgga aagtgctgag catgcccaca gtgaatatcg tgaanaagac 3780

cgaggtgcag acaggcggct tcagcaaaga gtctatcctg cccaagagga acagcgataa 3840

gctgatcgcc agaaagaagg actgggaccc taagaagtac ggcggcttcg acagccccac 3900

cgtggcctat tctgtgctgg tgggtggccaa agtggaaaag ggcaagtcca agaaactgaa 3960

gagtgtgaaa gagctgctgg ggatcacat catggaaaga agcagcttcg agaagaatcc 4020

catcgacttt ctggaagcca agggctacaa agaagtgaaa aaggacctga tcatcaagct 4080

gcctaagtac tcctgttcg agctggaaaa cgcccgaag agaatgctgg cctctgccgg 4140

cgaactgcag aagggaacg aactggccct gcctccaaa tatgtgaact tcctgtacct 4200

ggccagccac tatgagaagc tgaagggtc ccccaggat aatgagcaga aacagctgtt 4260

tgtggaacag cacaagcact acctggacga gatcatcgag cagatcagcg agttctcaa 4320

gagagtgatc ctggccgacg ctaatctgga caaagtctg tccgctaca acaagcaccg 4380

ggataagccc atcagagagc aggcagagaa tatcaccac ctgtttacc tgaccaatct 4440

gggagcccct gccgccttca agtactttga caccaccatc gaccggaaga ggtacaccag 4500

caccaaagag gtgctggacg ccacctgat ccaccagagc atcaccggcc tgtacgagac 4560

acggatcgac ctgtctcagc tgggaggcga caaaaggccg gcggccacga aaaaggccgg 4620

ccaggcaaaa aagaaaaagg aattctctag aggcagtgga gagggcagag gaagtctgct 4680

aacatcggtg gacgtcgagg agaatcctgg cccagtgagc aaggcgagg agctgttcac 4740

cgggtgggtg cccatcctgg tcgagctgga cggcgacgta aacggccaca agttcagcgt 4800

gtccgagcag ggcgagggcg atgccaccta cggcaagctg acctgaagt tcatctgac 4860

caccggcaag ctgccctgac cctggcccac cctcgtgacc acctgacct acggcgtgca 4920

gtgcttcagc cgctacccc accacatgaa gcagcacgac ttcttcaagt ccgcatgcc 4980

cgaaggctac gtccaggagc gcaccatctt ctcaaggac gacggcaact acaagaccg 5040

cgccgaggtg aagtctgagg gcgacaccct ggtgaaccgc atcgagctga agggcatcga 5100

cttcaaggag gacggcaaca tcctggggca caagctggag tacaactaca acagccacaa 5160

cgtctatata atggccgaca agcagaagaa cggcatcaag gtgaacttca agatccgcca 5220

caacatcgag gacggcagcg tgcagctcgc cgaccactac cagcagaaca ccccatcgg 5280

cgacggcccc gtgctgctgc ccgacaacca ctacctgagc acccagtcg cctgagcaa 5340

agacccaac gagaagcgcg atcacatggt cctgctggag ttcgtgaccg cgcgggat 5400  
 cactctcggc atggacgagc tgtacaagct cgaggggaagc ggagctacta acttcagcct 5460  
 gctgaagcag gctggcgacg tggaggagaa ccctggacct atgtctagc tggacaagag 5520  
 caaagtcata aacggagctc tggaaactact caatggtgtc ggtatcgaag gcctgacgac 5580  
 aaggaactc gctcaaaagc tgggagttga gcagcctacc ctgtactggc acgtgaagaa 5640  
 caagcgggcc ctgctcgatg ccctgccaat cgagatgctg gacaggcatc ataccactt 5700  
 ctgccccctg gaaggcgagt catggcaaga ctttctgagg aacaacgcca agtcataccg 5760  
  
 ctgtgtcttc ctctcacatc gcgacggggc taaagtgcac ctggcaccg gcccaacaga 5820  
 gaaacagtac gaaaccctgg aaaatcagct cgcgttctctg tgtcagcaag gcttctccct 5880  
 ggagaacgca ctgtacgctc tgtccgccgt gggccacttt acactgggct gcgtattgga 5940  
 ggaacaggag catcaagtag caaaagagga aagagagaca cctaccaccg attctatgcc 6000  
 cccacttctg agacaagcaa ttgagctggt cgaccggcag ggagccgaac ctgccttctt 6060  
 tttcgccctg gaactaatca tatgtggcct ggagaaacag ctaaagtgcg aaagcggcgg 6120  
 gccgaccgac gcccttgacg attttgactt agacatgctc ccagccgatg cccttgacga 6180  
  
 ttttgacctt gacatgctcc ccgggtaatg taaaaagtg tgaattccgg caattcgata 6240  
 tcaagcttat cgataatcaa cctctggatt acaaaatttg taaaagattg actggtatc 6300  
 ttaactatgt tgcctctttt acgctatgtg gatagctgc ttaaatgcct ttgtatcatg 6360  
 ctattgcttc ccgtatggct ttcattttct cctccttgta taaatcctgg ttgctgtctc 6420  
 tttatgagga gttgtggccc gttgtcaggc aacgtggcgt ggtgtgact gtgtttgctg 6480  
 acgcaacccc cactggttgg ggcattgcca ccacctgtca gctcctttcc gggactttcg 6540  
 ctttccccct ccctattgcc acggcggaac tcacgcccgc ctgccttgcc cgtgctgga 6600  
  
 caggggctcg gctgttgggc actgacaatt ccgtggtgtt gtcggggaaa tcacgtcct 6660  
 ttccttggct gctcgcctgt gttgccacct ggattctgcg cgggacgtcc tctgctacg 6720  
 tccttcggc cctcaatcca gcggaccttc ctccccggg cctgctgccg gctctcggc 6780  
 ctcttcggc tcttcgctt cgccctcaga cgagtcggat ctcctttgg gccgctccc 6840  
 cgcatcgata ccgtcgacct cgagacctag aaaaacatgg agcaatcaca agtagcaata 6900  
 cagcagctac caatgctgat tgtgctggc tagaagcaca agaggaggag gaggtgggtt 6960  
 ttccagtcac acctcaggta cctttaagac caatgactta caaggcagct gtagatctta 7020  
  
 gccacttttt aaaagaaaag gggggactgg aagggtaat tcactccca cgaagacaag 7080  
 atactcttga tctgtggatc taccacacac aaggctactt ccctgattgg cagaactaca 7140  
 caccagggcc agggatcaga tatccactga cctttggatg gtgctacaag ctagtaccag 7200

ttgagcaaga gaaggtagaa gaagccaatg aaggagagaa cacccgcttg ttacaccctg 7260  
 tgagcctgca tgggatggat gaccggaga gagaagtatt agagtggagg ttgacagcc 7320  
 gcctagcatt tcatcacatg gcccgagagc tgcacccgga ctgtactggg tctctctggt 7380  
 tagaccagat ctgagcctgg gagctctctg gctaactagg gaaccactg cttaaacctc 7440  
  
 aataaagctt gccttgagtg cttcaagtag tgtgtgccg tctgttgtg gactctggtg 7500  
 actagagate cctcagacc ttttagtcag tgtggaaaat ctctagcagg gcccgtttaa 7560  
 acccgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt ttgccctcc 7620  
 cccgtgcctt ccttgacctt ggaagtgcc actcccactg tcctttccta ataaaatgag 7680  
 gaaattgcat cgcattgtct gagtaggtgt cattctattc tggggggtgg ggtggggcag 7740  
 gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct 7800  
 atggcttctg aggcggaag aaccagctgg ggctctaggg ggtatccca cgcgcctgt 7860  
  
 agcggcgcac taagcgcggc ggggtgtgtg gttacgcga gcgtgaccg tacacttgcc 7920  
 agcgccttag cgcctctcc tttcgtttc ttccttctt tctctgccac gttcgcggc 7980  
 tttcccctg aagctctaaa tcgggggctc ctttagggg tccgatttag tgctttacgg 8040  
 cacctcgacc caaaaaact tgattagggt gatggttcac gtagtgggcc atcgccctga 8100  
 tagacggttt ttcgccctt gacgttgag tccacgttct ttaatagtg actctgttc 8160  
 caaactggaa caaacctca ccctatctcg gtctattctt ttgattata agggattttg 8220  
 ccgatttcgg cctattggtt aaaaaatgag ctgatttaac aaaaattta cgcgaattaa 8280  
  
 ttctgtggaa tgtgtgtcag ttaggggtg gaaagteccc aggetccca gcaggcagaa 8340  
 gtagtcaaag catgcatctc aattagtcag caaccagggt tggaaagtcc ccaggctccc 8400  
 cagcaggcag aagtatgcaa agcatgcatc tcaattagtc agcaaccata gtcccgcctc 8460  
 taactccgc catcccgcct ctaactccgc ccagttccgc ccattctcgc ccccatgget 8520  
 gactaatttt tttattttat gcagaggccg aggccctc tgcctctgag ctattccaga 8580  
 agtagtgagg aggccttttt ggaggcctag gcttttgcaa aaagctccc ggagcttgta 8640  
 tatccatttt cggatctgat cagcacgtgt tgacaattaa tcatcgcat agtatatcg 8700  
  
 catagtataa tacgacaagg tgaggaacta aaccatggcc aagttgacca gtgccgttcc 8760  
 ggtgctcacc gcgcgcgacg tcgccggagc ggtcgagttc tggaccgacc ggctcgggtt 8820  
 ctcccgggac ttcgtggagg acgacttcgc cgggtgtgtc cgggacgacg tgaccctgtt 8880  
 catcagcgcg gtccaggacc aggtgggtgcc ggacaacacc ctggcctggg tgtgggtgcg 8940  
 cggcctggac gagctgtacg ccgagtgttc ggaggtcgtg tccacgaact tccgggacgc 9000  
 ctccgggccc gccatgaccg agatcggcga gcagccgtgg gggcgggagt tcgccctgcg 9060

cgaccggcc ggcaactgcg tgcacttcgt ggccgaggag caggactgac acgtgctacg 9120

agatttcgat tccaccgccc ccttctatga aaggttgggc ttcggaatcg ttttccggga 9180

cgccggctgg atgatectcc agcgcgggga tctcatgctg gagttcttcg cccaccccaa 9240

cttgtttatt gcagcttata atggttacia ataaagcaat agcatcacia atttcacaaa 9300

taaagcattt ttttcactgc attctagtgt tggtttgc cc aaactcatca atgtatctta 9360

tcatgtctgt ataccgtcga cctctagctg gagcttggcg taatcatggt catagctgtt 9420

tcctgtgtga aattgttacc cgctcacaat tccacacaac atacgagccg gaagcataaa 9480

gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca ttaattgcgt tgcgctcact 9540

gcccgtttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc 9600

ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgtcgcg 9660

ctcggctggt cgctgcggc gagcggatc agctcactca aaggcggtaa tacggttatc 9720

cacagaatca ggggataacg caggaagaa catgtgagca aaaggccagc aaaaggccag 9780

gaaccgtaaa aaggccgctg tgcctggcgtt ttccatagg ctccgcccc ctgacgagca 9840

tcacaaaaat cgacgtcaa gtcagagggt gcgaaaccg acaggactat aaagatacca 9900

ggcgtttccc ctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg 9960

atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag 10020

gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccctg 10080

tcagcccgac cgctgcgctt tatccggtaa ctatcgtctt gactccaacc cggtaaagaca 10140

cgacttatcg ccaactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 10200

cggtgtctaca gacttcttga agtgggtggc taactacggc tacactagaa gaacagtatt 10260

tggatctgc gctctgtcga agccagttac cttcggaaaa agagtggta gctcttgatc 10320

cgcaaaaaa accaccgctg gtagecgggtg tttttttgtt tgcaagcagc agattacgcg 10380

cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg 10440

gaacgaaac tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta 10500

gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaaactg 10560

gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg 10620

ttcatccata gttgctgac tccccgctg gtatagataact acgatcggg agggcttacc 10680

atctggcccc agtgcctcaa tgataccgcg agaccacgc tcaccggctc cagatttacc 10740

agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc 10800

ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag 10860  
 tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat 10920  
 ggcttcattc agctccggtt cccaacgata aaggcgagtt acatgatccc ccatgtttgtg 10980  
 caaaaaagcg gttagctcct tcggctctcc gatcgtttgc agaagtaagt tggcccagat 11040  
 gttatcactc atggttatgg cagcactgca taattctctt actgcatgc catccgtaag 11100  
 atgcttttct gtgactggtg agtactcaac caagtattc tgagaatagt gtatgcggcg 11160  
 accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt 11220

aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct 11280  
 gttgagatcc agttcgatgt aaccactcgc tgcaccaac tgatcttcag catcttttac 11340  
 tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggat 11400  
 aaggcgacac cggaaatgtt gaatactcat actcttctt tttcaatatt attgaagcat 11460  
 ttatcagggt tattgtctca tgagcggata catattgaa tgtatttaga aaaataaaca 11520  
 aataggggtt ccgcgacat tccccgaaa agtgccacct gacgctgacg gatcgggaga 11580  
 tctcccgatc cctatggtg cactctcagt acaatctgct ctgatgccgc atagttaagc 11640

cagtatctgc tcctgcttg tgtgtggag gtcgctgagt agtgccgag caaaattaa 11700  
 gctacaacia ggcaaggctt gaccgacaat tgcatgaaga atctgcttag gtttaggcgt 11760  
 tttgcgctgc ttcgcatgt acgggccaga tatacgcgtt gacattgatt attgactagt 11820  
 tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga gttccgctt 11880  
 acataactta cggtaaatgg cccgctggc tgaccgcca acgacccccg cccattgacg 11940  
 tcaataatga cgtatgtcc catagtaacg ccaatagga ctttccattg acgtcaatgg 12000  
 gtggagtatt tacggtaaac tgcccactg gcagtacatc aagtgtatca tatgccaagt 12060

acgcccccta ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc ccagtacatg 12120  
 accttatggg actttctac ttggcagtac atctacgtat tagtcatcgc tattaccatg 12180  
 gtgatcggtt tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt 12240  
 ccaagtctcc acccattga cgtcaatggg agttttgtt ggcacaaaa tcaacgggac 12300  
 tttccaaaat gtcgtaacia ctccgcccc ttgacgcaa tgggcggtag gcgtgtacgg 12360  
 tgggaggtct atataagcag cgcgttttgc ctgtaactggg tctctctggt tagaccagat 12420  
 ctgagcctgg gagctctctg gctaactagg gaaccactg cttaaacctc aataaagctt 12480

gccttgagtg cttcaagtag tgtgtgcccg tctgttgtgt gactctgta actagagatc 12540  
 cctcagacce ttttagtcag tgtggaaaat ctctagcagt ggcgcccgaa cagggacttg 12600  
 aaagcгааг gгааaccaga ggagctctct cgacgcagga ctcggcttgc tgaagcgcgc 12660

acggcaagag gcgagggcg gcgactggtg agtacgcaa aaatTTTgac tagcggaggc 12720  
 tagaaggaga gagatgggtg cgagagcgtc agtattaagc gggggagaat tagatcgca 12780  
 tgggaaaaaa ttcggttaag gccaggggga aagaaaaaat ataaattaaa acatatagta 12840  
 tgggcaagca gggagctaga acgattcgca gTTaatcctg gcctgttaga aacatcagaa 12900  
  
 ggctgtagac aaatactggg acagctacaa ccatcccttc agacaggatc agaagaactt 12960  
 agatcattat ataatacagt agcaaccctc tattgtgtgc atcaaaggat agagataaaa 13020  
 gacaccaagg aagctttaga caagatagag gaagagcaaa acaaaagtaa gaccaccgca 13080  
 cagcaagcgg ccgctgatct tcagacctgg aggaggagat atgagggaca attggagaag 13140  
 tgaattatat aaatataaag tagtaaaaat tgaaccatta ggagtagcac ccaccaaggc 13200  
 aaagagaaga gtggtgcaga gaaaaaaag agcagtggga ataggagctt tgttccttgg 13260  
 gttcttggga gcagcaggaa gcactatggg cgcagcgtca atgacgctga cggtagcaggc 13320  
  
 cagacaatta ttgtctggtg tagtgcagca gcagaacaat ttgctgaggg ctattgaggc 13380  
 gcaacagcat ctgttgcaac tcacagtctg gggcatcaag cagctccagg caagaatcct 13440  
 ggctgtggaa agatacctaa aggatcaaca gctcctgggg atttggggtt gctctggaaa 13500  
 actcatttgc accactgctg tgccttggaa tgctagtgg agtaataaat ctctggaaca 13560  
 gatttggat cacacgacct ggatggagtg ggacagagaa attaacaatt acacaagctt 13620  
 aatacactcc ttaattgaag aatcgcaaaa ccagcaagaa aagaatgaac aagaattatt 13680  
 ggaattagat aaatgggcaa gtttgtggaa ttggttTaac ataacaatt ggctgtggta 13740  
  
 tataaaatta ttcataatga tagtaggagg cttggtaggt ttaagaatag tttttgctgt 13800  
 actttctata gtgaatagag ttaggcaggg atattcacca ttatcgtttc agaccacct 13860  
 cccaacccc aggggaccgc acaggcccga aggaatagaa gaagaaggtg gagagagaga 13920  
 cagagacaga tccattcgat tagtgaacgg atcggcactg cgtgcgcaa ttctgcagac 13980  
 aaatggcagt attcatccac aatTTTaaa gaaaagggg gattggggg tacagtgcag 14040  
 gggaaagaat agtagacata atagcaacag acatacaaac taaagaatta caaaaacaaa 14100  
 ttacaaaaat tcaaaatTTT cgggtttatt acagggacag cagagatcca gtttggTtaa 14160  
  
 ttaa 14164  
 <210> 212  
 <211> 14  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
oligonucleotide"

<400> 212

tttattccct gacc

14

<210> 213

<211> 34

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 213

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys

1                    5                    10                    15

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala

                  20                    25                    30

His Gly

<210> 214

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 214

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 215

<211> 32

<212> PRT

<213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 215  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 216  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 216  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 217  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221>  
 > source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 217  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 218  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 218  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 219  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 219  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Glu Gln His Gly  
                   20                    25                    30

<210> 220  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 220  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 221  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
                   polypeptide"  
 <400> 221

Leu Thr Leu Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210  
 > 222  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
                   polypeptide"  
 <400> 222

Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 223  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 223

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210> 224

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 224

Leu Thr Pro Asn Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 225

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 225

Leu Thr Leu Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210> 226

<211> 32

<212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 226  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                          20                    25                    30

<210> 227  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 227  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                          20                    25                    30

<210>  
 > 228  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 228  
 Leu Thr Leu Asp Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

20 25 30

<210> 229

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 229

Leu Thr Gln Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

20 25 30

<210> 230

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 230

Leu Ser Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

20 25 30

<210> 231

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 231

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Leu Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 232

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
                  polypeptide"

<400> 232

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Leu Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 233

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
                  polypeptide"

<400> 233

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Arg Gln Ala His Gly

                  20                    25                    30

<210

> 234

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 234

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Asn Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 235

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 235

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Ala Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 236

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 236

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 237

<211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 237  
 Leu Thr Leu Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 238  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 238  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Cys Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 239  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 239  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Gln Leu Leu Pro Val Leu Cys Glu Gln His Gly  
                   20                  25                  30

<210

> 240

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 240

Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Gly Gly Arg Pro Ala  
   1          5          10          15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                  25                  30

<210> 241

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 241

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala  
   1          5          10          15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                  25                  30

<210> 242

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 242

Leu Thr Pro Asn Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210> 243

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 243

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Gly Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 244

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 244

Leu Thr Leu Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 245

<211> 32

<212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 245  
 Leu Thr Pro Ala Gln Ala Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210>  
 > 246  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 246  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Asn Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 247  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 247  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Leu Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

20 25 30

<210> 248

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 248

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Thr Leu Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

20 25 30

<210> 249

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 249

Leu Thr Pro Asp Gln Val Val Thr Ile Ala Ser Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

20 25 30

<210> 250

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 250

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Arg Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 251

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 251

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Asn Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210

> 252

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 252

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Thr His Gly  
                   20                    25                    30

<210> 253

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 253

Leu Pro Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 254

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 254

Leu Thr Ser Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 255

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 255

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Glu Gln His Gly  
                   20                    25                    30

<210> 256  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 256  
 Leu Ile Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30  
 <210> 257  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 257  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Met Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210>  
 > 258  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 258  
 Leu Thr Arg Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                          20                    25                    30

<210> 259

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 259

Leu Thr Pro Asp Gln Val Val Ala Thr Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                          20                    25                    30

<210> 260

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 260

Leu Ile Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                          20                    25                    30

<210> 261

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 261

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asn His Gly

                  20                    25                    30

<210> 262

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 262

Leu Thr Leu Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Lys Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 263

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 263

Leu Thr Pro Asp Gln Leu Val Ala Ile Ala Asn Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210

> 264

<211> 32

<212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 264  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Gly His Gly  
                   20                    25                    30

<210> 265  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 265  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Glu His Gly  
                   20                    25                    30

<210> 266  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
  
 <400> 266  
 Leu Thr Leu Asp Lys Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

20 25 30

<210> 267

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 267

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

20 25 30

<210> 268

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 268

Leu Thr Pro Asp Lys Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

20 25 30

<210> 269

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 269

Leu Thr Gln Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Tyr Gln Asp His Gly  
                   20                    25                    30

<210

> 270

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 270

Leu Thr Pro Ala Gln Val Val Ala Ile Val Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 271

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 271

Leu Thr Pro Asp Lys Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 272

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 272

Leu Thr Gln Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                   5                   10                   15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                   25                   30

<210> 273

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 273

Leu Thr Pro Asp Gln Val Met Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                   5                   10                   15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                   25                   30

<210> 274

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 274

Leu Thr Thr Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                   5                   10                   15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                   25                   30

<210> 275

<211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 275  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala  
 1            5            10            15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
              20            25            30

<210  
 > 276  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 276  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1            5            10            15  
 Leu Glu Thr Val Gln Arg Leu Leu Leu Val Leu Cys Gln Ala His Gly  
              20            25            30

<210> 277  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 277  
 Leu Thr Gln Glu Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1            5            10            15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                  25                  30

<210> 278

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 278

Leu Thr Pro Asp Gln Val Val Thr Ile Ala Asn Gly Gly Lys Gln Ala  
   1                  5                  10                  15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                  25                  30

<210> 279

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 279

Leu Ser Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

  1                  5                  10                  15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys His Asp His Gly  
                   20                  25                  30

<210> 280

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 280

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Met Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210> 281

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
                  polypeptide"

<400> 281

Leu Ile Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210

> 282

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
                  polypeptide"

<400> 282

Leu Thr Pro Val Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 283

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 283

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Lys Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 284

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 284

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Met Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 285

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 285

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Phe Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 286  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 286  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Gln Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30  
 <210> 287  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 287  
 Leu Thr Pro Ala Gln Val Val Ala Leu Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30  
 <210>  
 > 288  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 288  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Arg Pro Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Glu Gln His Gly  
                          20                    25                    30

<210> 289

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 289

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Ala Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                          20                    25                    30

<210> 290

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 290

Leu Thr Gln Val Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                          20                    25                    30

<210> 291

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 291

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Arg Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 292

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 292

Leu Pro Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 293

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 293

Leu Thr Leu Asp Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210

> 294

<211> 32

<212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 294  
 Leu Ser Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Leu Gln Arg Leu Leu Pro Val Leu Cys Gln Thr His Ala  
                   20                    25                    30

<210> 295  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 295  
 Leu Asn Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 296  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
  
 <400> 296  
 Leu Thr Pro Asp Gln Val Met Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

20 25 30

<210> 297

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 297

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Arg Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

20 25 30

<210> 298

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 298

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Thr

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

20 25 30

<210> 299

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 299

Leu Thr Pro Asp Gln Val Met Thr Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210

> 300

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 300

Leu Thr Pro Ala Gln Val Val Thr Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 301

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 301

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Arg Ala His Gly  
                   20                    25                    30

<210> 302

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 302

Leu Ser Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 303

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 303

Leu Thr Pro Asp Gln Val Val Gly Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 304

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 304

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala Asn Gly  
                   20                    25                    30

<210> 305

<211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 305  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Thr His Gly  
                   20                    25                    30

<210  
 > 306  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 306  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Met Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 307  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 307  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Met Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
 20 25 30

<210> 308

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 308

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1 5 10 15  
 Leu Ala Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
 20 25 30

<210> 309

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 309

Leu Thr Pro Asp Gln Val Val Thr Ile Ala Ser Gly Gly Lys Gln Ala  
 1 5 10 15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 20 25 30

<210> 310

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 310

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Thr Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 311

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
                  polypeptide"

<400> 311

Met Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210

> 312

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
                  polypeptide"

<400> 312

Leu Ala Pro Asp Gln Val Val Ala Val Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 313

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 313

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Lys Thr Val Gln Gln Leu Leu Pro Val Leu Cys Glu Gln His Gly  
                   20                    25                    30

<210> 314

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 314

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Arg Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 315

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 315

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Gln Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 316  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 316  
 Leu Thr Pro Asp Gln Val Leu Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Leu Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 317  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 317  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Arg Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210>  
 > 318  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 318  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Met Gln Arg Leu Leu Pro Val Leu Cys Arg Ala His Gly  
                          20                    25                    30

<210> 319

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 319

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Met Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                          20                    25                    30

<210> 320

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 320

Leu Thr Thr Asp Gln Val Val Thr Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                          20                    25                    30

<210> 321

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 321

Leu Thr Pro Thr Gln Val Met Ala Ile Ala Asn Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 322

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 322

Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210> 323

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 323

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Met Leu Cys Gln Asp His Gly

                  20                    25                    30

<210

> 324

<211> 32

<212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 324  
 Leu Thr Ser Ala Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 325  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 325  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Gln Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 326  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
  
 <400> 326  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Ala Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

20 25 30

<210> 327

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 327

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Met Leu Cys Gln Ala His Gly

20 25 30

<210> 328

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 328

Leu Thr Leu Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala Arg Gly

20 25 30

<210> 329

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 329

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Leu Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210

> 330

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 330

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asn His Gly  
                   20                    25                    30

<210> 331

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 331

Leu Thr Pro Asp Gln Val Val Thr Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Met Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 332

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 332

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Arg Val Gln Arg Leu Leu Pro Val Leu Cys Glu Gln His Gly  
                   20                    25                    30

<210> 333

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 333

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Cys Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Arg Gln Ala His Gly  
                   20                    25                    30

<210> 334

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 334

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Arg Asp His Gly  
                   20                    25                    30

<210> 335

<211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 335  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Met Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210  
 > 336  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 336  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Cys Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Arg His Ala His Gly  
                   20                    25                    30

<210> 337  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 337  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln His His Gly  
                   20                  25                  30

<210> 338

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 338

Leu Ile Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
   1                  5                  10                  15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln His His Gly  
                   20                  25                  30

<210> 339

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 339

Leu Thr Arg Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

  1                  5                  10                  15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Glu Gln His Gly  
                   20                  25                  30

<210> 340

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 340

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala

1                    5                    10                    15

Val Gly Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210> 341

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 341

Leu Thr Leu Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Glu Gln His Gly

                  20                    25                    30

<210

> 342

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 342

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Met Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 343

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 343

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Met Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 344

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 344

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Lys Gln His Gly  
                   20                    25                    30

<210> 345

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 345

Leu Thr Leu Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Thr His Gly  
                   20                    25                    30

<210> 346  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 346  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Ala Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                          20                    25                    30

<210> 347  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 347  
 Leu Thr Pro Ala Gln Val Val Thr Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Glu Gln His Gly  
                          20                    25                    30

<210>  
 > 348  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 348  
 Leu Thr Pro Ala Gln Val Met Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                           20                    25                    30

<210> 349

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 349

Leu Thr Arg Glu Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Arg Gln Ala His Gly  
                           20                    25                    30

<210> 350

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 350

Leu Thr Leu Ala Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                           20                    25                    30

<210> 351

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 351

Leu Thr Leu Glu Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210> 352

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 352

Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Glu Gln His Gly

                  20                    25                    30

<210> 353

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 353

Leu Ser Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210

> 354

<211> 32

<212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 354  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln His His Gly  
                   20                    25                    30

<210> 355  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 355  
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Arg Gln Ala His Gly  
                   20                    25                    30

<210> 356  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 356  
 Leu Ser Gln Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

20 25 30

<210> 357

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 357

Leu Pro Pro Glu Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

20 25 30

<210> 358

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 358

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Ala Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

20 25 30

<210> 359

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 359

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Glu His Gly  
                   20                    25                    30

<210

> 360

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 360

Leu Thr Leu Asp Gln Val Ala Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 361

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 361

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Val Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 362

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 362

Leu Ile Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 363

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 363

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Arg Gln Ala His Gly  
                   20                    25                    30

<210> 364

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 364

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Thr His Gly  
                   20                    25                    30

<210> 365

<211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 365  
 Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210  
 > 366  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 366  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Val Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 367  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 367  
 Leu Ser Pro Asp Gln Val Val Thr Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Leu Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
 20 25 30

<210> 368

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 368

Leu Thr Pro Val Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
 20 25 30

<210> 369

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 369

Leu Thr Leu Asp Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Thr His Gly  
 20 25 30

<210> 370

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 370

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Cys Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Arg Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210> 371

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
                  polypeptide"

<400> 371

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Phe Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210

> 372

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
                  polypeptide"

<400> 372

Leu Pro Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210> 373

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 373

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Phe Gln Glu His Gly  
                   20                    25                    30

<210> 374

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 374

Leu Thr Pro Ala Lys Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 375

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 375

Leu Thr Pro Val Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Ala Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 376  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 376  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Gly Leu Cys Gln Asp His Gly

20                    25                    30

<210> 377  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 377  
 Leu Thr Leu Ala Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

20                    25                    30

<210>  
 > 378  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 378  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Thr Val Leu Cys Gln Asp His Gly  
                          20                    25                    30

<210> 379

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 379

Leu Pro Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                          20                    25                    30

<210> 380

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 380

Leu Thr Pro Ala Gln Ala Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                          20                    25                    30

<210> 381

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 381

Leu Thr Pro Ala Gln Val Val Ala Ile Val Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Thr His Gly

                  20                    25                    30

<210> 382

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 382

Leu Thr Pro Asp Gln Val Val Ala Val Ala Gly Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 383

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 383

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Gly Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210

> 384

<211> 32

<212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 384  
 Leu Pro Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Glu Ala His Gly  
                   20                    25                    30

<210> 385  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 385  
 Leu Thr Thr Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 386  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
  
 <400> 386  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Val Pro Val Leu Cys Gln Asp His Gly

20 25 30

<210> 387

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 387

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Thr His Ala

20 25 30

<210> 388

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 388

Leu Thr Leu Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Thr His Gly

20 25 30

<210> 389

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 389

Leu Thr Pro Asn Gln Leu Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210

> 390

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 390

Leu Ser Pro Ala Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 391

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 391

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Val Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 392

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 392

Leu Thr Pro Asp Gln Val Met Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 393

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 393

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Gly Gly Arg Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 394

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 394

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Trp Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 395

<211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 395  
 Leu Thr Pro Asp Lys Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210  
 > 396  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 396  
 Leu Thr Pro Ala Gln Val Met Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 397  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 397  
 Leu Thr Gln Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala Asn Gly  
                   20                  25                  30

<210> 398

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 398

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Pro Ala  
   1                  5                  10                  15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Glu Gln His Gly  
                   20                  25                  30

<210> 399

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 399

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Ser Gly Lys Gln Ala

  1                  5                  10                  15

Leu Glu Thr Met Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                  25                  30

<210> 400

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 400

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Arg Gln Asp His Gly

                  20                    25                    30

<210> 401

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
                  polypeptide"

<400> 401

Leu Thr Pro Tyr Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210

> 402

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
                  polypeptide"

<400> 402

Leu Thr Pro Tyr Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210> 403

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 403

Leu Thr Leu Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Glu His Gly  
                   20                    25                    30

<210> 404

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 404

Leu Thr Leu Glu Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Leu Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 405

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 405

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Arg Arg Leu Leu Gln Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 406  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 406  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1            5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Arg Gln Asp His Gly

20                    25                    30

<210> 407  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 407  
 Leu Thr Pro Asp Gln Val Val Ser Ile Ala Asn Gly Gly Lys Gln Ala  
 1            5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

20                    25                    30

<210  
 > 408  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 408  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Thr His Gly  
                          20                    25                    30

<210> 409

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 409

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Lys Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                          20                    25                    30

<210> 410

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 410

Leu Thr Thr Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                          20                    25                    30

<210> 411

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 411

Leu Ile Pro Gln Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 412

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 412

Leu Thr Leu Thr Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 413

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 413

Leu Thr Pro Thr Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210

> 414

<211> 32

<212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 414  
 Leu Thr Pro Thr Gln Val Met Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 415  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 415  
 Leu Thr Pro Asp Gln Val Val Ala Val Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 416  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"  
 <400> 416  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

20 25 30

<210> 417  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 417  
 Leu Thr Pro Gly Gln Val Val Ala Ile Ala Ser Gly Gly Lys Arg Ala

1 5 10 15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
 20 25 30

<210> 418  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 418  
 Leu Thr Pro Asp Gln Val Val Val Ile Ala Ser Gly Gly Lys Gln Ala  
 1 5 10 15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

20 25 30

<210> 419  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 419

Leu Pro Pro Asp Gln Val Val Ala Ile Ala Ser Gly Ser Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210

> 420

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 420

Leu Thr Pro Asp Gln Val Val Thr Ile Ala Asn Gly Ser Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 421

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 421

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Gln Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 422

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 422

Leu Thr Pro Asp His Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 423

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 423

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Gln Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 424

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 424

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Arg Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Glu Gln His Gly  
                   20                    25                    30

<210> 425

<211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 425  
 Leu His Pro Gly Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210  
 > 426  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 426  
 Leu Thr Leu Asp Gln Val Val Ser Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 427  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 427  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Ala Leu Cys Gln Asp His Gly  
                   20                  25                  30

<210> 428

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 428

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Pro Ala  
   1                  5                  10                  15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Glu Gln His Gly  
                   20                  25                  30

<210> 429

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 429

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

  1                  5                  10                  15

Leu Lys Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                  25                  30

<210> 430

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
           polypeptide"

<400> 430

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Arg Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210> 431

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 431

Leu Asn Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

                  20                    25                    30

<210

> 432

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polypeptide"

<400> 432

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15

Leu Glu Thr Val Lys Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

                  20                    25                    30

<210> 433

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 433

Leu Thr Leu Asp Gln Val Val Ala Ile Ala Asn Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 434

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 434

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Arg Asp His Gly  
                   20                    25                    30

<210> 435

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 435

Leu Thr Pro Ala Gln Val Leu Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Thr Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 436  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 436  
 Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Met Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly

20                    25                    30

<210> 437  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 437  
 Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Gly Leu Cys Gln Ala His Gly

20                    25                    30

<210>  
 > 438  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 438  
 Leu Thr Arg Glu Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Arg Gln Ala His Gly  
                          20                    25                    30

<210> 439

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 439

Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Val His Gly  
                          20                    25                    30

<210> 440

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 440

Leu Thr Pro Asn Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Leu Val Leu Cys Gln Asp His Gly  
                          20                    25                    30

<210> 441

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 441

Leu Thr Pro Asp Gln Val Met Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly  
                   20                    25                    30

<210> 442

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 442

Leu Thr Arg Glu Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 443

<211> 33

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 443

Leu Ser Thr Ala Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala

1                    5                    10                    15  
 Leu Glu Gly Ile Gly Glu Gln Leu Leu Lys Leu Arg Thr Ala Pro Tyr  
                   20                    25                    30  
 Gly

<210> 444  
 <211> 33  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 444  
 Leu Ser Thr Ala Gln Val Val Ala Val Ala Ser Gly Gly Lys Pro Ala  
 1                    5                    10                    15  
 Leu Glu Ala Val Arg Ala Gln Leu Leu Ala Leu Arg Ala Ala Pro Tyr  
                   20                    25                    30  
 Gly

<210> 445  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"  
 <400> 445  
 Leu Thr Gln Val Gln Val Val Ala Ile Ala Ser Gly Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 446  
 <211> 32  
 <212> PRT  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 446

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Gly Lys Gln Ala  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 447

<211> 32

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 polypeptide"

<400> 447

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Gly Gly Lys Arg Ala  
  
 1                    5                    10                    15  
 Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly  
                   20                    25                    30

<210> 448

<211> 14

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"

<400> 448

tttattccct gaca

14

<210> 449

<211> 18

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"

<400> 449  
 tcggcccctg ccggccca 18  
 <210> 450  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <400> 450  
 tgctgcct ccaggctcct 20  
 <210> 451  
 <211> 40  
 <212> DNA  
 <213> Homo sapiens  
 <400>  
 > 451  
 aaacggaagg gcctgagtcc gagcagaaga agaagtttta 40  
 <210> 452  
 <211> 40  
 <212> DNA  
 <213> Homo sapiens  
 <400> 452  
 aggaggaagg gcctgagtcc gagcagaaga agaagggctc 40  
 <210> 453  
 <211> 26  
 <212> RNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <400> 453  
 gaguccgagc agaagaagaa guuuua 26  
 <210> 454  
 <211> 39

<212  
 > DNA  
 <213> Homo sapiens  
 <400> 454  
 aggaggaagg gcctgagtcc gagcagaaga gaagggctc 39  
 <210> 455  
 <211> 39  
 <212> DNA  
 <213> Homo sapiens  
 <400> 455  
 ggaggaaggg cctgagtccg agcagaagaa gaagggctc 39  
 <210> 456  
 <211> 38  
 <212> DNA  
 <213> Homo sapiens  
 <400> 456  
 ggaggaaggg cctgagtccg agcagaagag aagggctc 38  
 <210> 457  
 <211> 40  
 <212> DNA  
 <213> Homo sapiens  
 <400> 457  
 ggaggaaggg cctgagtccg agcagaagaa agaagggctc 40  
 <210> 458  
 <211> 37  
 <212> DNA  
 <213> Homo sapiens  
 <400> 458  
 ggaggaaggg cctgagtccg agcagaagga agggctc 37  
 <210> 459  
 <211> 36  
 <212> DNA  
 <213> Homo sapiens  
 <400> 459

ggaggaaggg cctgagtccg agcagaagaa gggctc 36

<210> 460

<211> 33

<212> DNA

<213> Homo sapiens

<400> 460

ggaggaaggg cctgagtccg agcagaaggg ctc 33

<210> 461

<211> 33

<212> DNA

<213> Homo sapiens

<400> 461

ggaggaaggg cctgagcccg agcagaaggg ctc 33

<210> 462

<211> 122

<212> DNA

<213> Homo sapiens

<400> 462

ggaggaaggg cctgagtccg agcagaagaa gaagggtcc catcacatca accggtggcg 60

cattgccacg aagcaggcca atggggagga catcgatgtc acctccaatg actagggtgg 120

gc 122

<210> 463

<211> 48

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
oligonucleotide"

<220><221> modified\_base

<222> (3)..(32)

<223> a, c, u, g, unknown or other

<400> 463

acnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnguuiuaga gcuaugcu 48

<210> 464  
 <211> 67  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <220><221> source  
 <223> /note="Description of Combined DNA/RNA Molecule: Synthetic  
 oligonucleotide"  
 <400> 464  
 agcauagcaa guaaaaauaa ggctaguccg uuaucaacu gaaaaagugg caccgagucg 60  
 gugcuuu 67

<210> 465  
 <211> 62  
 <212> RNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <220><221> modified\_base  
 <222> (1)..(20)  
 <223> a, c, u, g, unknown or other  
 <400> 465  
 nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaaagc aaguuuuuuu aaggcuaguc 60  
 cg 62

<210> 466  
 <211> 48  
 <212> DNA  
 <213> Homo sapiens  
 <400> 466  
 ctggaggagg aaggcctga gtccgagcag aagaagaagg gctcccat 48

<210> 467

<211> 30  
 <212> RNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <400> 467  
 gaguccgagc agaagaagaa guuuuagagc 30  
 <210> 468  
 <211> 20  
 <212> RNA  
 <213> Homo sapiens  
 <400> 468  
 gaguccgagc agaagaagau 20  
 <210> 469  
 <211> 20  
 <212> RNA  
 <213> Homo sapiens  
 <400> 469  
 gaguccgagc agaagaagua 20  
 <210> 470  
 <211> 20  
 <212> RNA  
 <213> Homo sapiens  
 <400> 470  
 gaguccgagc agaagaacaa 20  
 <210> 471  
 <211> 20  
 <212> RNA  
 <213> Homo sapiens  
 <400> 471  
 gaguccgagc agaaugaa 20  
 <210> 472  
 <211> 20

<212> RNA  
 <213> Homo sapiens  
 <400> 472  
 gaguccgagc agaaguagaa 20  
 <210> 473  
 <211> 20  
 <212> RNA  
 <213> Homo sapiens  
 <400> 473  
 gaguccgagc agaugaagaa 20  
  
 <210> 474  
 <211> 20  
 <212> RNA  
 <213> Homo sapiens  
 <400> 474  
 gaguccgagc acaagaagaa 20  
 <210> 475  
 <211> 20  
 <212> RNA  
 <213> Homo sapiens  
 <400> 475  
 gaguccgagg agaagaagaa 20  
 <210> 476  
 <211> 20  
 <212> RNA  
 <213> Homo sapiens  
 <400> 476  
 gaguccgugc agaagaagaa 20  
 <210> 477  
 <211> 20  
 <212> RNA  
 <213> Homo sapiens  
 <400> 477

gagucggagc agaagaagaa 20

<210> 478  
 <211> 20  
 <212> RNA  
 <213> Homo sapiens  
 <400> 478  
 gagaccgagc agaagaagaa 20  
 <210> 479  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <400> 479  
 aatgacaagc ttgctagcgg tggg 24  
 <210> 480  
 <211> 39  
 <212> DNA  
 <213> Homo sapiens  
 <400> 480  
 aaaacggaag ggcctgagtc cgagcagaag aagaagttt 39  
 <210> 481  
 <211> 39  
 <212> DNA  
 <213> Homo sapiens  
 <400> 481  
 aaacaggggc cgagattggg tgttcagggc agaggtttt 39  
 <210> 482  
 <211> 38  
 <212> DNA  
 <213> Homo sapiens  
 <400> 482

aaaacggaag ggctgagtc cgagcagaag aagaagtt 38

<210> 483

<211> 40

<212> DNA

<213> Homo sapiens

<400> 483

aacggaggga ggggcacaga tgagaaactc agggttttag 40

<210> 484

<211> 38

<212> DNA

<213> Homo sapiens

<400> 484

agcccttctt ctctgctcg gactcaggcc ctctctcc 38

<210> 485

<211> 40

<212> DNA

<213> Homo sapiens

<400> 485

cagggaggga ggggcacaga tgagaaactc aggaggcccc 40

<210> 486

<211> 80

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
oligonucleotide"

<400> 486

ggcaatgccc caccggttga tgtgatggga gcccttctag gaggccccca gacgagccac 60

tggggcctca aactcaggc 80

<210> 487

<211> 98

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
oligonucleotide"

<400> 487

ggacgaaaca ccggaacat tcaaacagc atagcaagt aaaataaggc tagtccgtta 60

tcaactgaa aaagtggcac cgagtcggtg cttttttt 98

<210> 488

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
polynucleotide"

<400> 488

ggacgaaaca ccggtagtat taagtattgt tttatggctg ataaatttct ttgaatttct 60

ccttgattat ttgtataaaa agttataaaa taatcttggt ggaaccattc aaaacagcat 120

agcaagttaa aataaggcta gtccgttata aacttgaaaa agtggcaccc agtcggtgct 180

tttttt 186

<210> 489

<211> 73

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
oligonucleotide"

<400> 489

tgaatggtcc caaacggaa gggcctgagt ccgagcagaa gaagaagttt tagagctatg 60

ctgttttgaa tgg 73

<210> 490

<211> 95

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic  
oligonucleotide"  
<400> 490  
gggttttaga gctatgctgt tttgaatggt cccaaaacgg gtcttcgaga agacgtttta 60  
gagctatgct gttttgaatg gtcccaaac ttttt 95  
<210> 491  
<211> 36  
<212> DNA  
<213> Artificial Sequence  
<220><221> source  
<223> /note="Description of Artificial Sequence: Synthetic  
oligonucleotide"  
<220><221>  
> modified\_base  
<222> (5)..(34)  
<223> a, c, t, g, unknown or other  
<400> 491  
aaacnnnnnn nnnnnnnnnn nnnnnnnnnn nnnngt 36  
<210> 492  
<211> 36  
<212> DNA  
<213> Artificial Sequence  
<220><221> source  
<223> /note="Description of Artificial Sequence: Synthetic  
oligonucleotide"  
<220><221> modified\_base  
<222> (7)..(36)  
<223> a, c, t, g, unknown or other  
<400> 492  
taaaacnnnn nnnnnnnnnn nnnnnnnnnn nnnnnn 36  
<210> 493  
<211> 84  
<212> DNA  
<213> Artificial Sequence

<220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <400> 493  
 giggaagga cgaacaccg ggtcttcgag aagacctgtt ttagagctag aatagcaag 60  
 ttaaaataag gctagtcctg tttt 84  
 <210> 494  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <220><221> modified\_base  
  
 <222> (6)..(24)  
 <223> a, c, t, g, unknown or other  
 <400> 494  
 caccgnnnnn nnnnnnnnnn nnnn 24  
 <210> 495  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <220><221> modified\_base  
 <222> (5)..(23)  
 <223> a, c, t, g, unknown or other  
 <400> 495  
 aaacnnnnnn nnnnnnnnnn nnnn 24  
 <210> 496  
 <211> 69  
 <212> DNA

<213> Homo sapiens  
 <400> 496  
 ctggctctcc acctctctgc cctgaacacc caatctcggc ccctctcgcc acctctctgc 60  
 atttctggt 69  
 <210> 497  
 <211> 138  
 <212> DNA  
 <213> Mus sp.  
 <400> 497  
 acccaagcac tgagtgccat tagctaaatg cataggttac cacccacagg tgccaggggc 60  
 ctttcccaaa gttcccagcc ctttctcaa ctttctctgg cccagaggct ttcccatgtg 120  
 tgtggtgga ccctttga 138  
 <210> 498  
 <211> 46  
 <212> RNA  
  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <220><221> modified\_base  
 <222> (1)..(19)  
 <223> a, c, u, g, unknown or other  
 <400> 498  
 nnnnnnnnnn nnnnnnnng uuauuguacu cucaagauuu auuuuu 46  
 <210> 499  
 <211> 91  
 <212> RNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <400> 499  
 guuacuuaaa ucuugcagaa gcuacaaaga uaaggcuuca ugccgaauc aacaccugu 60

cauuuuauagg caggguuuu ucguuuuuu a 91

<210> 500

<211> 70

<212> DNA

<213> Homo sapiens

<400> 500

ttttctagtg ctgagtttct gtgactcctc tacattctac ttctctgtgt ttctgtatac 60

tacctctcc 70

<210> 501

<211> 17

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 501

Lys Arg Lys Phe Asp Thr Glu Thr Lys Asp Cys Asn Glu Lys Lys Lys

1 5 10 15

Lys

<210> 502

<211> 10

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 502

Leu Lys Ala Leu Gly Ser Asp Leu Thr Leu

1 5 10

<210> 503

<211> 4

<212> PRT

<213> Artificial Sequence

<220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 503

Arg Arg Gly Gln

1

<210> 504

<211> 5

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 504

Arg Val Arg Arg Glu

1                      5

<210> 505

<211> 14

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 505

tctgcaagag tagg

14

<210> 506

<211> 14

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 506

tgcactccac cttg 14  
 <210> 507  
 <211> 14  
 <212> DNA  
 <213> Artificial Sequence  
 <220><221> source  
 <223> /note="Description of Artificial Sequence: Synthetic  
 oligonucleotide"  
 <400> 507  
 tcaagcagct tctc 14  
 <210> 508  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
 <400> 508  
 gcccccttc atgcaaaacc cgg 23  
 <210> 509  
 <211> 20  
 <212> DNA  
 <213> Homo sapiens  
 <400> 509  
 gcccccttc atgcaaaacc 20  
 <210> 510  
 <211> 276  
 <212> DNA  
 <213> Homo sapiens  
 <400> 510  
 cgcccatggc cgcgcgcgt ccacacaact caccggagtc cgcgccttgc gccgccgacc 60  
 agttcgcagc tccgcgccac ggcagccagt ctctgagtgg aagagagcgc agcccggcca 120  
 ccggacctac ttactcgct tgctgattgt ctatTTTTgc gttacaact tttctaagaa 180  
 cttttgtata caaaggaact ttttaaaaaa gacgcttcca agttatattt aatccaaaga 240  
 agaaggatct cggccaattt ggggttttgg gttttg 276  
 <210> 511

<211> 151  
 <212> DNA  
 <213> Homo sapiens  
 <400> 511  
 tccccgcgc ggccggcggc gcgggaggcc cgccccctt tcatgcaaaa cccggcagcg 60  
 aggctgggct cgagtggagg agccgccg cgctgattgg tcctagaaa cccatttatt 120  
 ccctgacagc ccccgtcaca tggatggtt g t 151  
 <210> 512  
 <211> 30  
 <212> DNA  
 <213> Homo sapiens  
 <400>  
 > 512  
 gaagccgccg ggcccgagc aaacttcggg 30  
 <210> 513  
 <211> 30  
 <212> DNA  
 <213> Homo sapiens  
 <400> 513  
 ccctacatgt tgtactacct ctgcctcgac 30