



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2015년12월07일
(11) 등록번호 10-1574952
(24) 등록일자 2015년12월01일

(51) 국제특허분류(Int. Cl.)
C12Q 1/48 (2006.01) C12N 15/52 (2006.01)
C12P 19/04 (2006.01)
(21) 출원번호 10-2013-0039938
(22) 출원일자 2013년04월11일
심사청구일자 2013년04월11일
(65) 공개번호 10-2014-0122902
(43) 공개일자 2014년10월21일
(56) 선행기술조사문헌
KR1020080013909 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
서울대학교 산학협력단
서울특별시 관악구 관악로 1 (신림동)
(주)진켄
대전광역시 유성구 가정북로 26-69 (장동)
(72) 발명자
김병기
서울특별시 서초구 서초중앙로 188 아크로비스타
C동 1202호
최윤희
서울특별시 양천구 오목로 30 세진아파트 401호
(뒷면에 계속)
(74) 대리인
특허법인코리아나

전체 청구항 수 : 총 9 항

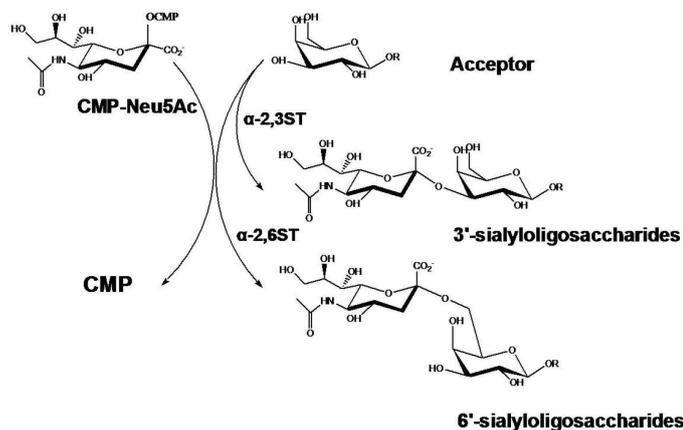
심사관 : 이현석

(54) 발명의 명칭 α -2,3 및 α -2,6 시알산 전이효소의 변이체 탐색 방법 및 시알릴올리고당 생산에의 응용

(57) 요약

본 발명은 시알산 전이효소 변이체의 탐색 방법과 변이체의 특성, 상기 시알산 전이효소 변이체 효소를 암호화하는 유전자 및 이들을 이용한 시알릴올리고당의 합성에 관한 것이다. 본 발명은 단백질 공학적 변이를 통해 개량된 파스테우렐라 유래의 α -2,3 시알산 전이효소와 포도박테리움 유래의 α -2,6 시알산 전이효소를 제공한다. 또한 본 발명의 시알산 전이효소 변이체는 3' - 및 6' -시알릴올리고당 생산 반응에서 부반응 감소, 생산성 증대 및 이로 인한 생산 비용 절감을 제공한다.

대표도 - 도1



(72) 발명자

김종훈

경기도 성남시 분당구 분당로201번길 17 효자촌현
대아파트 106동 1003호

김대회

대전광역시 서구 대덕대로 150 경성큰마을아파트
106동 907호

우진석

대전광역시 유성구 엑스포로 448 엑스포아파트 50
9동 1602호

명세서

청구범위

청구항 1

삭제

청구항 2

삭제

청구항 3

하기의 (a) 내지 (g) 에서 선택되는 어느 하나의 서열로 표시되는 α -2,3 시알산 전이효소 변이체 :

(a) 서열번호 1 의 아미노산 서열;

(b) 서열번호 2 의 아미노산 서열;

(c) 서열번호 3 의 아미노산 서열;

(d) 서열번호 4 의 아미노산 서열;

(e) 서열번호 5 의 아미노산 서열;

(f) 서열번호 1 내지 5 중 어느 하나의 아미노산 서열에서 313 번째의 아미노산이 세린, 아스파라긴, 히스티딘, 트레오닌, 티로신 또는 아스파르트산에서 선택되는 아미노산으로 치환되거나, 265 번째의 아미노산이 아스파라긴, 세린 또는 글리신에서 선택되는 아미노산으로 치환된 서열.

청구항 4

제 3 항에 따른 α -2,3 시알산 전이효소 변이체를 암호화 하는 DNA.

청구항 5

제 4 항에 있어서,

서열번호 6 내지 10 의 서열 중 어느 하나로 구성된 DNA.

청구항 6

제 4 항 또는 제 5 항에 따른 DNA 를 포함하는 재조합 DNA 벡터.

청구항 7

제 6 항에 따른 재조합 DNA 벡터로 형질전환된 숙주세포.

청구항 8

제 6 항에 따른 재조합 DNA 벡터로 형질전환된 숙주세포의 추출물.

청구항 9

제 6 항에 따른 재조합 DNA 벡터로 형질전환된 숙주세포 및 상기 숙주세포의 추출물로 이루어진 균에서 선택되는 어느 하나를 생축매로 사용한 3' -시알릴 올리고당의 제조 방법.

청구항 10

하기의 단계를 포함하는, 제 3 항에 따른 α -2,3 시알산 전이효소 변이체의 탐색 방법:

시알산 전이효소의 결정 구조 또는 모델 구조의 두 개의 기질 결합 부위로부터 5 내지 20 Å 이내에서 선택된

잔기 중에서, 다수 서열 정렬과 알라닌 스캐닝을 통해 포화변이를 수행할 기능적 잔기를 분석하는 단계.

청구항 11

제 10 항에 있어서, 20 내지 70 °C 의 온도 및 pH 6 내지 10 에서 pH 변화에 의한 지시약을 이용하여 알라닌 스캐닝 및 포화 변이체 탐색을 수행하는 것을 특징으로 하는 시알산 전이효소의 변이체 탐색 방법.

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

발명의 설명

기술 분야

[0001] 본 발명은 단백질 공학적 효소 변이를 통한 시알산 전이효소의 변이체 탐색 방법 및 제조된 변이체의 특성 분석과, 이들 시알산 전이효소 변이체를 이용하여 시알릴올리고당을 합성하는 것에 관한 발명이다.

배경 기술

[0002] 모유는 유아에게 필수적인 영양소를 제공하는 것뿐 아니라, 단순한 영양소의 개념을 넘어 다양한 건강상의 이익을 제공한다. 모유 올리고당은 기능성 성분으로 구성되어 있고, 우유보다 100-200 배 많은 1 L 당 5-10 g 의 올리고당을 포함하며, 지금까지 130 종류 이상의 모유 올리고당이 확인되었다. 올리고당의 함유량과 구조적 다양성은 우유와 달리 모유에 있어서 매우 특이적이다. 모유 올리고당 중, 시알릴올리고당의 3' -및 6' -시알릴락토오스는 우유보다 5-10 배 가량 많이 포함되어 있으며, 구체적으로, 0.1-0.3 g/L 의 3' -시알릴락토오스 및 0.3-0.5 g/L 의 6' -시알릴락토오스를 포함한다.

[0003] 시알릴올리고당의 인체 내 기능을 살펴보면, 첫 번째로 갱글리오사이드 (Ganglioside) 인 GM1 (Gal-GalNAc-3' -sialyllactose-ceramide), GM2 (GalNAc-3' -sialyllactose-ceramide) 및 GM3 (3' -sialyllactose-ceramide) 의 형태로 신경조직에서 생체 내 신호전달의 수용체로서 작용하여, 뇌의 발달과 기억력 형성에 관여한다. 두 번째로는, 백혈구 또는 골수구 등에서 시알릴 루이스 X 및/또는 시알릴 루이스 A (Sialyl-LewisX, sialyl-LewisA) 의 형태로 염증 부위로 백혈구를 유도하여 치료를 촉진하는 역할을 한다. 세 번째로는, 박테리아나 바이러스의 초기 감염 과정에서의 숙주세포의 세포표면에 부착을 막는 저해제로서 역할을 하는 연구가 진행되어왔다. 시알릴올리고당 중, 시알산이 α-2,3 와 α-2,6 로 연결된 시알릴락토오스는 각각 조류와 인간 인플루엔자 바이러스의 헤마글루티닌 (hemagglutinin) 에 대한 인식 물질이 될 수 있어, 바이러스가 숙주세포 표면의 당과 결합하여 침입하는 것을 경쟁적으로 저해할 수 있다. 시알릴락토오스의 다양한 기능에 관한 연구가 활발한 현재, 모유 내 존재하는 시알릴올리고당의 합성은 식품, 의약 또는 필터 등의 다양한 산업적 응용이 가능하며, 이들의 효소를 이용한 생물전환 공정이 활발히 연구되고 있다.

[0004] 다만, 시알릴락토오스를 포함한 시알릴올리고당을 생산할 수 있는 시알산 전이효소의 탐색 및 효소의 반응 특성에 관한 연구는 진행되어 왔으나, 발굴된 효소의 개량에 관한 연구는 미미한 수준이다.

[0005] 를루아르 (Leloir) 당전이효소 (glycosyltransferase) 는 뉴클레오티드 (nucleotide) 형태의 당을 공여체로 사용한다. 보통 이러한 뉴클레오티드 형태의 당은 매우 고가이기 때문에, 이 물질을 생산하기 위한 생물 전환

공정에 대한 연구가 많이 이루어져 왔다. 본 발명의 시알산 당전이 효소는 시티딘 일인산-N-아세틸뉴라민산을 공여체 기질로 사용하며, 이 물질을 합성하기 위한 생물전환 공정이 기존에 연구되었다.

[0006] 시티딘 일인산-N-아세틸뉴라민산의 합성은 N-아세틸글루코사민 (N-acetyl-D-glucosamine), 피루빈산염 (pyruvate) 및 시티딘 일인산 (CMP)의 기질로부터 N-아세틸글루코사민-2-에피머라아제, 아세테이트 키나아제, 시티딘 일인산 키나아제, N-아세틸뉴라민산 알돌라아제 및 시티딘 일인산 N-아세틸뉴라민산 합성 효소에 해당하는 다섯 가지 효소를 이용하여 합성하는 공정이 구축되었다. 최종적으로, 3' -, 6' -시알릴올리고당을 생산하기 위해, 시티딘 일인산-N-아세틸뉴라민산의 생산과 함께 α -2,3, α -2,6 시알산 전이효소를 이용하여 목적 물질을 생산하는 시스템이 구축되었다.

[0007] 이와 관련하여, (주)진캠에서는 시알릴올리고당 중, 시알릴락토오스를 대량생산하기 위해 목적 물질의 분리 및 정제 공정을 구축하였으며, 특허 3' -시알릴락토오스의 대량 생산을 시도하였다. 그러나 3' - 및 6' -시알릴락토오스를 경제적으로 대량생산하기 위해서는 목적 물질의 생산성 (productivity) 을 높이기 위한 α -2,3 및 α -2,6 전이효소의 기능 향상을 목표로 하는 효소 변이가 필요하다.

[0008] 시알산 전이효소는 시티딘 일인산-N-아세틸뉴라민산 (Cytidine monophosphate N-acetylneuraminic acid, CMP-Neu5Ac) 로부터 N-아세틸뉴라민산을 당 수용체인 갈락토스 부분에 α -2,3 또는 α -2,6 의 결합으로 전달하여 올리고당을 생성한다. 이 중, 구조로 보았을 때 GT-B 접힘에 해당하는 시알산 전이효소는 시알산을 전달하는 공여체 기질인 CMP-Neu5Ac 와 시알산 수용체인 락토오스 (Gal β 1,4Glc) 가 결합할 수 있는 두 개의 도메인으로 이루어진 기질 결합 포켓을 가지고 있으며, 시알산 전이효소의 개량 시에 포켓에 위치하는 아미노산 잔기들에 대해 단백질 공학적인 변이를 수행할 수 있다.

[0009] 본 발명에 이용된 α -2,3 시알산 전이효소는 Xi chen 그룹에 의해 처음으로 탐색되었고, 파스테우렐라 멀토시다 (*Pasteurella multocida*) 균주에서 유래하며, pH 에 따라 네 가지 효소 활성을 갖는 다기능 효소이다. 이 효소는 넓은 pH 범위 (pH 6.0-10.0) 에서 주요 반응인 갈락토스로의 α -2,3 시알산 전이 기능을 가지며, pH 4.5-7.0 사이에서는 α -2,3 전이보다 낮은 수준으로 α -2,6 시알산 전이의 활성을 갖는다. 또한 pH 5.0-5.5 에서 α -2,3 결합의 시알산을 분해하는 시알리다아제의 활성을 갖고, pH 5.5-6.5 에서 α -2,3 결합을 가지는 시알릴 갈락토사이드로부터 또 다른 갈락토사이드로의 트랜스 시알리다아제의 활성을 가진다. 이 시알산 전이효소는 대장균 내에서 가용성 단백질로의 발현이 용이하며, α -2,3 시알산 전이의 활성이 높고, 기질특이성이 넓어, 다양한 시알릴올리고당 생산에 용이한 장점이 있다.

[0010] 반면, 본 발명에 이용된 α -2,6 시알산 전이효소는 포토박테리움 담셀라 (*Photobacterium damsela*) 균주에서 유래하며, 당전이 효소의 구조 접힘과 서열로 볼 때, 파스테우렐라에서 유래한 α -2,3 시알산 전이효소와 같은 GT family 80 에 속한다. 포토박테리움 유래의 α -2,6 시알산 전이효소의 경우에도 최근 α -2,6 시알리다아제와 트랜스시알리다아제의 활성이 밝혀졌으나 이러한 부반응의 활성이 α -2,6시알산 전이의 활성보다 매우 미미하기 때문에 (150 배 이상), 시알산 전이의 활성이 대부분이라 볼 수 있다. α -2,6 시알산 전이효소의 경우 부반응이 거의 없고, α -2,6 시알산 전이의 활성이 대부분이며, 기질 특이성이 다양하다는 장점을 가진 반면, 효소 활성의 차이가 파스테우렐라 유래의 α -2,3 시알산 전이효소보다 5-6 배 이상 낮은 단점을 가지고 있다.

[0011] 따라서 α -2,3 및 α -2,6 결합을 가지는 다양한 시알릴올리고당의 생산 효율을 증가시키기 위해서는, 위에서 명시한 기질특이성이 다양한 α -2,3 및 α -2,6 시알산 전이효소의 기능이 향상된 변이체를 제조하고 이들을 시알릴올리고당 생산에 응용할 필요가 있다.

선행기술문헌

특허문헌

(특허문헌 0001) 대한민국 공개특허공보 제10-2008-0093087호(2008.10.20)

(특허문헌 0002) 대한민국 등록특허공보 제10-0886650호(2009.02.25).

발명의 내용

해결하려는 과제

- [0012] 본 발명은 α -2,3 및 α -2,6 시알산 전이효소의 단백질 공학적 변이를 통한 변이체를 제조하고, 변이체를 이용하여 시알릴올리고당을 효과적으로 생산하고자 한다.
- [0013] 또한, 본 발명의 α -2,3 시알산 전이효소의 변이체는 부반응의 생성을 억제함으로써, 회분식 일체형 반응을 가능케 하고, 촉매 반응을 빠르게 하여, 3' -시알릴락토오스를 포함한 다양한 3' -시알릴올리고당의 생산성 및 효율을 향상시키는 데에 목적이 있다.
- [0014] 또한, 본 발명의 α -2,6 시알산 전이효소의 변이체는 촉매 활성이 향상되어, 6' -시알릴락토오스를 포함한 다양한 6' -시알릴올리고당의 생산성을 향상시키는 데에 목적이 있다.

과제의 해결 수단

- [0015] 본 발명은 α -2,3 및 α -2,6 시알산 전이효소의 단백질 공학적 변이를 통한 변이체 탐색 방법과 변이체를 암호화하는 유전자 서열 정보를 제공한다.
- [0016] 또한, 이들 변이체의 특성을 분석하고 이들을 이용하여 시알릴올리고당을 생성하는 것을 제공한다. 즉, 본 발명은 단백질 공학적 변이를 통해 개량된 파스테우렐라 유래의 α -2,3 시알산 전이효소와 포토박테리움 유래의 α -2,6 시알산 전이효소를 이용하여 3' - 및 6' -시알릴올리고당 생산 반응에 적용하는 것이다.
- [0017] 또한, 본 발명은 시알산 전이효소의 변이체를 탐색하기 위한 하이브리드 (hybrid) 즉, 세미-래셔널 (semi-rational) 방법을 제공한다. 이 방법은 방향성 진화 (directed evolution) 와 논리적 예측 (rational design) 방법을 결합한 것으로서, 양질의 적은 수의 변이체 라이브러리만을 확보하고자 하기 위함이다. 이 방법은 단백질의 목적이 되는 부분을 선택하고 단백질의 서열, 구조와 기능 및 컴퓨터 프로그램을 이용하여 특정 아미노산의 잔기를 선택하여 변이를 수행하는 것을 말한다.
- [0018] 구체적으로, 본 발명은 하기를 제공한다.
- [0019] (1) 하기의 단계를 포함하는 시알산 전이효소의 변이체의 탐색 방법:
- [0020] 시알산 전이효소의 결정 구조 또는 모델 구조의 두 개의 기질 결합 부위로부터 5 내지 20 Å 이내에서 선택된 잔기 중에서, 다수 서열 정렬과 알라닌 스캐닝을 통해 포화변이를 수행할 기능적 잔기를 분석하는 단계.
- [0021] (2) (1) 에 있어서, 20 내지 70 °C 의 온도 및 pH 6 내지 10 에서 pH 변화에 의한 지시약을 이용하여 알라닌 스캐닝 및 포화 변이체 탐색을 수행하는 것을 특징으로 하는 시알산 전이효소의 변이체 탐색 방법.
- [0022] (3) 하기의 (a) 내지 (g) 에서 선택되는 어느 하나의 서열로 표시되는 α -2,3 시알산 전이효소 변이체 :
- [0023] (a) 서열번호 1 의 아미노산 서열;
- [0024] (b) 서열번호 2 의 아미노산 서열;
- [0025] (c) 서열번호 3 의 아미노산 서열;
- [0026] (d) 서열번호 4 의 아미노산 서열;
- [0027] (e) 서열번호 5 의 아미노산 서열;
- [0028] (f) 서열번호 1 내지 5 중 어느 하나의 아미노산 서열에서 313 번째 또는 265 번째의 아미노산이 친수성 아미노산으로 치환된 서열;
- [0029] (g) 서열번호 1 내지 5 중 어느 하나의 아미노산 서열에서 313 번째 또는 265 번째의 아미노산이 친수성 아미노산으로 치환된 서열로서, 서열번호 1 내지 5 와 97% 이상의 상동성을 가지는 아미노산 서열.
- [0030] (4) (3) 에 따른 α -2,3 시알산 전이효소 변이체를 암호화 하는 DNA.
- [0031] (5) (4) 에 있어서, 서열번호 6 내지 10 의 서열 중 어느 하나로 구성된 DNA.
- [0032] (6) (4) 또는 (5) 에 따른 DNA 를 포함하는 재조합 DNA 벡터.
- [0033] (7) (6) 에 따른 재조합 DNA 벡터로 형질전환된 숙주세포.
- [0034] (8) (6) 에 따른 재조합 DNA 벡터로 형질전환된 숙주세포의 추출물.
- [0035] (9) (6) 에 따른 재조합 DNA 벡터로 형질전환된 숙주세포 및 상기 숙주세포의 추출물로 이루어진 균에서 선택되

는 어느 하나를 생축매로 사용한 3' -시알릴 올리고당의 제조 방법.

- [0036] (10) 하기의 (a) 내지 (g) 에서 선택되는 어느 하나의 서열로 표시되는 α -2,6 시알산 전이효소 변이체 :
- [0037] (a) 서열번호 11 의 아미노산 서열;
- [0038] (b) 서열번호 12 의 아미노산 서열;
- [0039] (c) 서열번호 13 의 아미노산 서열;
- [0040] (d) 서열번호 14 의 아미노산 서열;
- [0041] (e) 서열번호 15 의 아미노산 서열;
- [0042] (f) 서열번호 11 내지 15 중 어느 하나의 아미노산 서열에서 411 번째의 아미노산이 작은 크기 또는 친수성 아미노산으로 치환되거나, 또는 433 번째의 아미노산이 친수성 아미노산으로 치환된 서열;
- [0043] (g) 서열번호 11 내지 15 중 어느 하나의 아미노산 서열에서 411 번째의 아미노산이 작은 크기 또는 친수성 아미노산으로 치환되거나, 또는 433 번째의 아미노산이 친수성 아미노산으로 치환된 서열로서, 서열번호 11 내지 15 와 55% 이상의 상동성을 가지는 아미노산 서열.
- [0044] (11) (10) 에 따른 α -2,6 시알산 전이효소 변이체를 암호화 하는 DNA.
- [0045] (12) (11) 에 있어서, 서열번호 16 내지 20 의 서열 중 어느 하나로 구성된 DNA.
- [0046] (13) (11) 또는 (12) 에 따른 DNA 를 포함하는 재조합 DNA 벡터.
- [0047] (14) (13) 에 따른 재조합 DNA 벡터로 형질전환된 숙주세포.
- [0048] (15) (13) 에 따른 재조합 DNA 벡터로 형질전환된 숙주세포의 추출물.
- [0049] (16) (13) 에 따른 재조합 DNA 벡터로 형질전환된 숙주세포 및 상기 숙주세포의 추출물로 이루어진 균에서 선택되는 어느 하나를 생축매로 사용한 6' -시알릴 올리고당의 제조 방법.

발명의 효과

- [0050] 본 발명의 변이체 제조 방법을 다른 당전이효소, 특히 GT-B 접힘 구조를 가지는 당전이효소에 적용할 수 있다. 또한, 3' -시알릴올리고당의 생산성 증대 및 생산 효율 증대 및 6' -시알릴올리고당의 생산성 증대의 효과를 가진다.
- [0051] 구체적으로, 촉매 기능이 향상된 α -2,3 시알산 전이효소의 변이체 사용을 통해 3' -시알릴락토오스를 포함한 3' -시알릴올리고당의 생산성 증대 및 이로 인한 생산 비용 절감이 가능하다. 또한 pH 4.5 내지 7.0 에서 가지는 α -2,6 시알산 전이의 부반응 감소로 인한 시티딘 일인산-N-아세틸뉴라민산 생산과의 회분식 일체형 반응을 가능하게 한다.
- [0052] 또한, 촉매 기능이 향상된 α -2,6 시알산 전이효소의 변이체 사용을 통해 일인산-N-아세틸뉴라민산과의 회분식 일체형 반응에서 속도 결정 단계가 되는 α -2,6 시알산 전이의 속도를 향상시킴으로써, 6' -시알릴올리고당의 생산성 증대 및 이로 인한 생산 비용 절감이 가능하다.
- [0053] 이에 따라, 상기 3' -, 6' -시알릴올리고당의 대량생산을 통해 식품, 의약품 또는 화장품 등의 다양한 활용이 가능하다.

도면의 간단한 설명

- [0054] 도 1 은 본 발명에서 시티딘 일인산-N-아세틸뉴라민산 (CMP-Neu5Ac) 과 갈락토스를 포함하는 수용체를 기질로 하여 α -2,3 및 α -2,6 시알산 전이효소를 이용한 3' -, 6' -시알릴올리고당의 합성 도식도를 나타낸다.
- 도 2 는 본 발명에서 탐색한 α -2,3 및 α -2,6 시알산 전이효소 각각의 단일 아미노산이 변화한 변이체의 야생형 대비 상대적인 고유 활성도 (specific activity) 를 나타낸다.
- 도 3 은 본 발명에서 탐색한 고유 활성도가 높은 단일 아미노산 치환 변이체 및 이들의 조합적 변이체의 동역학 변수 (kinetic parameter) 를 야생형의 값과 함께 나타낸 표이다. 도 3a 는 α -2,3 시알산 전이효소의 단일 아미노산 치환 변이체와 이들의 조합적 변이체의 동역학 변수를 나타내며, 도 3b 는 α -2,6 시알산 전이효소의

단일 아미노산 치환 변이체와 이들의 조합적 변이체의 동역학 변수를 나타낸다.

도 4 는 정제된 α -2,3 시알산 전이효소를 이용한 반응에서의 pH 4.5 내지 7.0 에서 3 '-시알릴락토오스의 생성 수율이 포화일 때의 6' -시알릴락토오스의 생성량을 나타낸 그래프이며, 야생형과 313 번째의 아르기닌이 각각 아스파라긴, 아스파르트산, 티로신, 트레오닌, 히스티딘으로 변환된 변이체와 조합 변이체인 R313N/T265S 와 R313H/T265S 를 이용한 6 '-시알릴락토오스의 정량 결과이다. 도 4a 는 pH 4.5 내지 pH 6.0 에서의 생성된 6' -시알릴락토오스의 양을, 도 4b 는 pH 6.5 내지 pH 7.0 에서의 생성된 6 '-시알릴락토오스의 양을 나타내며 단위는 켈 이다.

도 5 는 본 발명에서 탐색한 파스테우렐라 멀토시다 (*Pasteurella multocida*) 유래의 α -2,3 시알산 전이효소 변이체의 아미노산 서열을 나타낸 것으로, 도 5a 는 α -2,3 시알산 전이효소의 313 번째의 아르기닌이 아스파라긴으로 치환된 R313N 의 아미노산 서열 (서열번호 1), 도 5b 는 α -2,3 시알산 전이효소의 313 번째의 아르기닌 이 히스티딘으로 치환된 R313H 아미노산 서열 (서열번호 2), 도 5c 는 α -2,3 시알산 전이효소의 265 번째의 트레오닌이 세린으로 치환된 T265S 의 아미노산 서열 (서열번호 3), 도 5d 는 R313N 과 T265S 의 조합적 아미노산 서열 (서열번호 4), 도 5e 는 R313H 와 T265S 의 조합적 아미노산 서열 (서열번호 5) 이다. 여기서, 본 발명의 α -2,3 시알산 전이효소는 N 말단에 24 개의 아미노산이 제거된 형태이기 때문에 첫 메티오닌 서열부터 헤아릴 경우 메티오닌이 25 번이 된다.

도 6 은 본 발명에서 탐색한 α -2,3 시알산 전이효소 변이체의 DNA 서열을 나타낸 것으로, 도 6a 는 R313N 의 DNA 서열 (서열번호 6), 도 6b 는 R313H 의 DNA 서열 (서열번호 7), 도 6c 는 T265S 의 DNA 서열 (서열번호 8), 도 6d 는 R313N 과 T265S 의 조합적 DNA 서열 (서열번호 9), 도 6e 는 R313H 와 T265S 의 조합적 DNA 서열 (서열번호 10) 이다.

도 7 은 본 발명에서 탐색한 포토박테리움 담셀라 (*Photobacterium damsela*) 유래의 α -2,6 시알산 전이효소 변이체의 아미노산 서열을 나타낸 것으로, 도 7a 는 α -2,6 시알산 전이효소의 411 번째 이소류신이 트레오닌으로 치환된 I411T 의 아미노산 서열 (서열번호 11), 도 7b 는 α -2,6 시알산 전이효소의 433 번째의 류신이 세린으로 치환된 L433S 의 아미노산 서열 (서열번호 12), 도 7c 는 α -2,6 시알산 전이효소의 433 번째의 류신이 트레오닌으로 치환된 L433T 의 아미노산 서열 (서열번호 13), 도 7d 는 I411T 와 L433S 의 조합적 아미노산 서열 (서열번호 14), 도 7e 는 I411T 와 L433T 의 조합적 아미노산 서열 (서열번호 15) 이다.

도 8 은 본 발명에서 탐색한 α -2,6 시알산 전이효소 변이체의 DNA 서열을 나타낸 것으로, 도 8a 는 I411T 의 DNA 서열 (서열번호 16), 도 8b 는 L433S 의 DNA 서열 (서열번호 17), 도 8c 는 L433T 의 DNA 서열 (서열번호 18), 도 8d 는 I411T 와 L433S 의 조합적 DNA 서열 (서열번호 19), 도 8e 는 I411T 와 L433T 의 조합적 DNA 서열 (서열번호 20) 이다.

발명을 실시하기 위한 구체적인 내용

[0055] 본 발명에서 사용되는 용어는 당업계에서 통상적으로 사용되는 것으로 당업자라면 그 의미를 누구나 이해할 수 있을 것이나, 본 명세서에서 간략히 설명하면 다음과 같다:

[0056] (1) 시알산 전이효소는 를루아르 당전이효소로서, 시티딘 일인산-N-아세틸뉴라민산으로부터 N-아세틸뉴라민산을 수용체 당물질에 전이하는 효소를 의미한다. 또한 본 발명을 통해 생산한 α -2,3 및 α -2,6 시알산 전이효소의 변이체는 시티딘 일인산-N-아세틸뉴라민산 공여체 기질뿐 만 아니라 시티딘 일인산 디아미노뉴라민산 (CMP-KDN, CMP-deaminoneuraminic acid), 시티딘 일인산-N-글리콜릴뉴라민산 (CMP-Neu5Gc, CMP-N-glycolylneuraminic acid) 을 포함한 다양한 유도체 기질에 적용할 수 있다.

[0057] (2) α -2,3 시알산 전이효소는 Xi chen 그룹에 의해 처음으로 탐색되었고, 파스테우렐라 멀토시다 (*Pasteurella multocida*) 균주에서 유래하며, pH 에 따라 네 가지 효소 활성을 갖는 다기능 효소이다. 이 효소는 넓은 pH 범위 (pH 6.0-10.0) 에서 주요 반응인 갈락토스로의 α -2,3 시알산 전이 기능을 가지며, pH 4.5-7.0 사이에서는 α -2,3 전이보다 낮은 수준으로 α -2,6 시알산 전이의 활성을 갖는다. 또한 pH 5.0-5.5 에서 α -2,3 결합의 시알산을 분해하는 시알리다아제의 활성을 갖고, pH 5.5-6.5 에서는 α -2,3 결합을 가지는 시알릴 갈락토사이드로부터 또 다른 갈락토사이드로의 트랜스 시알리다아제의 활성을 가진다. 본 발명에서 사용한 α -2,3 시알산 전이효소는 N-말단에 24개의 아미노산이 제거된 형태로 첫 번째 아미노산의 번호가 25번이다. 이는 서열 번호 21 의 아미노산 서열 및 서열번호 22 의 DNA 서열을 나타낸다.

[0058] (3) α -2,6 시알산 전이효소는 포토박테리움 담셀라 (*Photobacterium damsela*) 균주에서 유래하며, 당전이 효

소의 구조 접힘과 서열로 볼 때, 파스테우렐라에서 유래한 α -2,3 시알산 전이효소와 같은 GT family 80 에 속한다. 본 발명에서 사용한 α -2,6 시알산 전이효소는 첫 메티오닌 서열부터 헤아릴 경우 메티오닌이 1 번이 된다. 이는 서열번호 23 의 아미노산 서열 및 서열번호 24 의 DNA 서열을 나타낸다.

- [0059] (4) 수용체 기질 중 하나인, 락토오스는 Gal β 1,4Glc (갈락토스와 글루코스가 β 1,4 결합으로 연결됨) 로 구성된 올리고당이다. 또한, 본 발명을 통해 생산한 α -2,3 및 α -2,6 시알산 전이효소의 변이체는 락토오스 수용체 기질뿐 만 아니라 다양한 락토오스 유도체 기질인, *N*-아세틸락토사민 (LacNAc), 아자이도 β -D-갈락토피라노실-(1-4)- β -D-글루코피라노사이드 (Lac β N3), 3-아자이도프로필 β -D-갈락토피라노실-(1-4)- β -D-글루코피라노사이드 (Lac β ProN3), 메틸 β -D-갈락토피라노실-(1-4)- β -D-글루코피라노사이드 (Lac β OMe) 등과 루이스X (LewisX, Gal β 1,4(Fuc α 1,3)GlcNAc), 루이스A (LewisA, Gal β 1,3(Fuc α 1,4)GlcNAc) 및 Lacto-*N*-tetraose (LNT, Gal β 1,3GlcNAc β 1,3Gal β 1,4Glc)를 포함하여 갈락토스 부분을 포함하는 다양한 올리고당 기질에 적용할 수 있다.
- [0060] (5) 3' - 및 6' -시알릴올리고당은 도 1 과 같이 *N*-아세틸뉴라민산 (시알산) 이 갈락토스 부분에 α -2,3 또는 α -2,6 결합으로 연결된 올리고당을 의미하며, 갈락토스 또는 글루코스에 다른 당이 더 결합할 수 있다.
- [0061] 또한, 본 발명을 통해 생산한 α -2,3 및 α -2,6 시알산 전이효소의 변이체는 3' -시알릴락토오스, 3' -시알릴락토사민, 시알릴 루이스X (Sialyl-LewisX), 시알릴 루이스A (Sialyl-LewisA), 6' -시알릴락토오스, 6' -시알릴락토사민 및 시알릴 LNT a, b, c (Sialyl-LNT, Neu5Ac α 2,3Gal β 1,3 GlcNAc β 1,3Gal β 1,4Glc/ Gal β 1,3(Neu5Ac α 2,6)GlcNAc β 1,3Gal β 1,4Glc/ Neu5Ac α 2,6Gal β 1,3GlcNAc β 1,3Gal β 1,4Glc) 를 포함하여 다양한 시알릴올리고당의 생산에 적용할 수 있다.
- [0062] (6) 3' - 및 6' -시알릴락토오스는 시알산이 락토오스의 갈락토스에 α -2,3 또는 α -2,6 결합으로 연결된 Neu5Ac α 2,3/2,6Gal β 1,4Glc 로 구성된 삼탄당 (triose) 물질을 의미한다.
- [0063] (7) 세포추출물은 시알산 전이효소가 발현된 본 발명의 미생물 추출물을 의미한다.
- [0064] (8) 전세포 반응은 특정 효소를 포함하는 세포를 파쇄하여 세포 내용물을 이용하거나 또는 효소를 분리 정제하지 않고 온전한 세포 전체를 이용한 반응을 의미한다.
- [0065] (9) 친수성 (hydrophilic) 아미노산은 물과 수소 결합을 이룰 수 있는 전기 음성도가 큰 원자 (산소, 질소, 황) 를 기능기 (functional group) 에 포함하고 있는 아미노산을 말하며, 세린 (serine), 트레오닌 (threonine), 시스테인 (cystein), 티로신 (tyrosine), 아스파르트산 (aspartic acid), 글루탐산 (glutamic acid), 아스파라긴 (asparagine), 글루타민 (glutamine), 히스티딘 (histidine), 라이신 (lysine) 또는 아르기닌 (arginine) 을 포함한다.
- [0066] (10) 작은 크기의 아미노산은 기능기가 작은 아미노산을 말하며, 글리신 (glycine), 알라닌 (alanine), 세린 (serine), 트레오닌 (threonine) 또는 시스테인 (cystein) 을 포함한다.
- [0067] (11) PCR 은 중합 효소 연쇄반응 (Polymerase Chain Reaction) 으로서, DNA 의 어떤 영역을 특이적으로 증폭시키는 방법을 의미 한다.
- [0068] (12) 위치 지정 돌연변이 (site directed mutagenesis) 는 유전자의 지정된 위치에 지정된 염기 배열의 변화를 도입하는 것을 말한다.
- [0069] (13) 포화 변이 (saturation mutagenesis) 는 유전자의 지정된 위치에 다양한 염기 배열의 변화를 도입하는 것을 말한다. 포화 변이는 주형 가닥에 결합하는 상보적인 서열의 프라이머 (primer) 상에 변이시키고자 하는 서열 대신 NNK 코돈 (codon) 을 삽입하여 PCR 을 통해 변이를 삽입시키는 것을 말한다. 이 때, NNK 코돈에서 N 은 뉴클레오티드의 A, T, G, C 를 의미하며, K 는 T, G 를 의미한다.
- [0070] (14) 벡터는 단일 가닥, 이중 가닥, 원형 또는 초나선 DNA 또는 RNA 로 이루어진 폴리뉴클레오티드를 의미하며, 제조할 단백질을 생산할 수 있도록 적절한 거리에 작동적으로 연결되어 있는 구성 요소들을 포함할 수 있다.
- [0071] 이러한 구성 요소에는 복제 오리진, 프로모터, 인핸서, 5' mRNA 리더 서열, 리보솜 결합 부위, 핵산 카세트, 종결 및 폴리아데닐화 부위, 또는 선별 가능한 표지 서식 등이 포함될 수 있으며, 상기 구성 요소들은 특이적인 용도에 따라 하나 또는 그 이상이 빠질 수도 있다. 핵산 카세트는 발현할 제조할 단백질의 삽입을 위한 제한효소 부위를 포함할 수 있다. 기능적 벡터에 있어서, 핵산 카세트는 번역 개시 및 종결 부위를 포함하는

발현될 핵산 서열을 함유하고, 필요에 따라, 벡터에 내에 두 종류의 카세트를 삽입할 수 있는 벡터를 사용하기도 하며, 상기 언급한 기능들이 부가적으로 서열화 될 수 있다.

- [0072] 재조합 벡터에 삽입된 유전자는 발현용 대장균 균주 BW25113 (DE3), 또는 BL21 (DE3) 등을 사용할 수 있으나, 삽입된 벡터의 종류에 따라 달라질 수 있다. 이러한 벡터 및 발현 균주는 당업자라면 용이하게 선택할 수 있다.
- [0073] (15) pH 지시약 (pH indicator) 은 적정을 하면서 중화점을 알기 위해, 혹은 수소 이온의 농도를 알기 위해서 주로 사용되는 것을 말한다. 지시약은 수소 이온 지수에 따라 산형 및 염기형으로 되어 색조가 다르며 이 영역을 변색 영역이라 부른다. 분광 광도법에 의해 흡광도에 따른 수소 이온의 농도를 측정할 수 있다. pH 지시약의 종류로서는 본 발명에서 사용한 크레졸 레드 (cresol red), 페놀 레드 (phenol red) 를 포함하여 pH 범위에 따라 티몰 블루 (thymol blue), 브로모페놀 블루 (bromophenol blue), 메틸 레드 (methyl red), 브로모티몰 블루 (bromothymol blue) 등이 있다.
- [0074] (16) 고유 활성도 (specific activity) 는 효소정제를 통해 불순물 및 다른 단백질을 제거한 순수한 단백질의 단위량당 활성을 나타내는 것으로 보통 1 분간에 1 μmol 의 기질 변화를 촉매하는 효소의 양을 1 단위로 하여 1mg 당 단위 수로 표시한다.
- [0075] (17) 효소반응 속도론 (enzyme kinetics)은 효소가 단위시간당 기질과 반응하여 전환시키는 속도를 다루는 것으로서 $v = V_{\text{max}}[S]/(K_m + [S])$ 의 속도식이 성립하며 반응속도 (v)의 기질농도 ([S]) 의존성을 조사하는 것으로 동역학 변수 (kinetic parameter)인 V_{max} 는 최대 반응 속도를 나타내고 V_{max} 를 효소의 몰농도 $[E]_0$ 로 제한한 것을 k_{cat} (turnover number: 효소 한 분자가 주어진 시간 안에 전환시키는 기질의 양)이라 한다. K_m 은 미카엘리스 (Michaelis) 상수로서 효소와 기질에 대한 겉보기 해리상수 (apparent dissociation constant)를 나타낸다. k_{cat}/K_m 은 k_{cat} 과 K_m 을 함께 나타낸 것을 말하며, 값이 높을수록 기질로부터 생성물로의 전환속도가 빠름을 의미한다.
- [0076] 한편, 본 발명에서 α -2,3 시알산 전이효소의 R313 의 단일 아미노산 치환 변이체 R313N 은 서열번호 1 의 아미노산 서열을 나타내며, 313 번째의 아미노산 위치에 친수성 아미노산 서열을 갖는 단백질 및 이 변이체 서열이 포함된 97% 이상의 상동성을 가지고, 시알산 전이의 활성을 가지는 효소도 모두 가능하다. 서열번호 1 의 단백질을 암호화 하는 DNA 는 서열번호 6 이며 상기의 아미노산을 암호화하는 모든 DNA 서열 또한 포함할 수 있다. 여기서, 본 발명의 α -2,3 시알산 전이효소는 N 말단에 24 개의 아미노산이 제거된 형태이기 때문에 첫 메티오닌 서열부터 헤아릴 경우 메티오닌이 25 번이 된다.
- [0077] 또한, α -2,3 시알산 전이효소의 R313 의 단일 아미노산 치환 변이체 R313H 은 서열번호 2 의 아미노산 서열을 나타내며, 313 번째의 아미노산 위치에 친수성 아미노산 서열을 갖는 단백질 및 이 변이체 서열이 포함된 97% 이상의 상동성을 가지고, 시알산 전이의 활성을 가지는 효소도 모두 가능하다. 서열번호 2 의 단백질을 암호화 하는 DNA 는 서열번호 7 이며 상기의 아미노산을 암호화하는 모든 DNA 서열 또한 포함될 수 있다.
- [0078] 또한, α -2,3 시알산 전이효소의 T265 의 단일 아미노산 치환 변이체 T265S 는 서열번호 3 의 아미노산 서열을 나타내며, 265 번째의 아미노산 위치에 친수성 아미노산 서열을 갖는 단백질 및 이 변이체 서열이 포함된 97% 이상의 상동성을 가지고, 시알산 전이의 활성을 가지는 효소도 모두 가능하다. 서열번호 3 의 단백질을 암호화 하는 DNA 는 서열번호 8 이며 상기의 아미노산을 암호화하는 모든 DNA 서열 또한 포함될 수 있다.
- [0079] 또한, α -2,3 시알산 전이효소의 R313N 과 T265S 의 조합적 변이체는 서열번호 4 의 아미노산 서열을 나타내며, R313H 와 T265S 의 조합적 변이체는 서열번호 5 의 아미노산 서열을 나타낸다. 또한 313 번째의 아미노산 위치 또는 265 번째의 아미노산 위치에 친수성 아미노산 서열을 갖는 단백질 및 이 변이체 서열이 포함된 97% 이상의 상동성을 가지고, 시알산 전이의 활성을 가지는 효소도 모두 가능하다. 서열번호 4 의 단백질을 암호화하는 DNA는 서열번호 9 이고, 서열번호 5 의 단백질을 암호화하는 DNA는 서열 10 이며, 상기의 아미노산을 암호화하는 모든 DNA 서열 또한 포함될 수 있다.
- [0080] 상기 명시한 α -2,3 시알산 전이효소의 변이체와 97% 이상의 상동성을 가지는 예로서는, 상기 변이체의 변이된 서열을 포함하고 있으며, α -2,3 시알산 전이효소의 활성을 갖는다고 명명, 또는 예측되는 서열로서, 파스테우렐라 (*Pasteurella*) 속 (genus), 특히 그 중 멀토시다 (*multocida*) 종 (species) 유래의 서열이 포함될 수 있다.
- [0081] 본 발명에서 α -2,6 시알산 전이효소의 I411 의 단일 아미노산 치환 변이체 I411T 는 서열번호 11 의 아미노산

서열을 나타내며, 411 번째의 아미노산 위치에 작은 크기 또는 친수성 아미노산 서열을 갖는 단백질 및 이 변이체 서열이 포함된 55% 이상의 상동성을 가지고, 시알산 전이의 활성을 가지는 효소도 모두 가능하다. 서열번호 11의 단백질을 암호화 하는 DNA는 서열번호 16이며 상기의 아미노산을 암호화하는 모든 DNA 서열 또한 포함할 수 있다.

[0082] 또한, α -2,6 시알산 전이효소의 L433의 단일 아미노산 치환 변이체 L433S은 서열번호 12의 아미노산 서열을 나타내고, L433T는 서열번호 13의 아미노산 서열을 나타내며, 433 번째의 아미노산 위치에 친수성 아미노산 서열을 갖는 단백질 및 이 변이체 서열이 포함된 55% 이상의 상동성을 가지고 시알산 전이의 활성을 가지는 효소도 모두 가능하다. 서열번호 12 및 13의 단백질을 암호화 하는 DNA는 서열번호 17 및 18이며 상기의 아미노산을 암호화하는 모든 DNA 서열 또한 포함될 수 있다.

[0083] 또한, α -2,6 시알산 전이효소의 I411T과 L433S의 조합적 변이체는 서열번호 14의 아미노산 서열을 나타내며, I411T와 L433T의 조합적 변이체는 서열번호 15의 아미노산 서열을 나타낸다. 또한 411 번째의 아미노산 위치 또는 433 번째의 아미노산 위치에 작은 크기 또는 친수성 아미노산 서열을 갖는 단백질 및 이 변이체 서열이 포함된 55% 이상의 상동성을 가지고, 시알산 전이의 활성을 가지는 효소도 모두 가능하다. 서열번호 14의 단백질을 암호화 하는 DNA는 서열번호 19이고, 서열번호 15의 단백질을 암호화 하는 DNA는 서열번호 20이며, 상기의 아미노산을 암호화하는 모든 DNA 서열 또한 포함될 수 있다.

[0084] 상기 명시한 α -2,6 시알산 전이효소의 변이체와 55% 이상의 상동성을 가지는 예로서는, 상기 변이체의 변이된 서열을 포함하고 있으며 α -2,6 시알산 전이효소의 활성을 갖는다고 명명, 또는 예측되는 서열로서, 포토박테리움 (*Photobacterium*) 속, 특히 그 중 담셀라 (*damseale*), 레이오그나티 (*leiognathi*) 중 유래의 서열이 포함될 수 있다. 또한 본 발명의 α -2,6 시알산 전이효소와 55%의 상동성을 가짐으로 인해 본 발명의 단백질 모델 구조를 형성하는 주형 단백질이 되는 포토박테리움 Jt-Ish-224 α -2,6 시알산 전이효소의 서열이 포함될 수 있다.

[0085] 시알산 전이효소의 변이 목적 부분의 선택 및 변이 수행

[0086] 본 발명의 α -2,3 시알산 전이효소는 결정 구조로부터 기질 결합 포켓부분을 확인하였으며, α -2,6 시알산 전이효소는 다른 포토박테리움 유래의 α -2,6 시알산 전이효소의 결정 구조를 주형으로 한 모델 구조로부터 기질 결합 부분을 확인하였다. 시티딘 일인산-N-아세틸뉴라민산과 수용체 기질 결합 부위로부터 5 내지 20 Å 이내에 위치하는 잔기들을 각각의 α -2,3 시알산 전이효소와 α -2,6 시알산 전이효소로부터 선택하였다. 본 발명에서는 생물 정보학의 서열 정보를 이용한 다수 서열 정렬 (multiple sequence alignment) 을 수행하였고 진화론적인 압력에도 불구하고 단백질 구조 내에 특정 위치의 아미노산 잔기를 보존하고 있는 것은 그 단백질에 있어서 구조와 기능상 매우 중요한 역할을 하고 특히 촉매 과정에 있어서 직접적인 역할을 할 가능성이 높기 때문에 이들은 변이 잔기로서 배제하였다.

[0087] 기질 결합 잔기들 중에서 위에서 선택된 잔기들 중, 포화 변이를 수행할 기능적 잔기 (functional residues) 를 선택하기 위해서 알라닌으로의 위치 지정 돌연변이를 수행하였다. 알라닌으로의 치환은 잔기가 제거된 것과 같아, 특정 잔기가 기질과의 작용으로 인해 중요한 촉매 활성에 기여를 하는지의 여부를 해석할 수 있다. 알라닌으로 치환 이후에 효소 활성을 비색법으로 측정하여 야생주와의 활성의 차이를 비교하였다. 시알산 전이의 활성을 잃은 치환 잔기들은 기질에 대한 촉매 활성을 중요하게 조절하는 잔기임을 확인하였다. 본 발명에서는 이와 같이 시알산 전이의 활성을 잃은 알라닌 치환 변이체들은 이들의 본래 잔기가 촉매 활성에 중요한 역할을 하는 강한 가능성으로 인해 ‘바꿀 수 없는’ 잔기로 인식하였다.

[0088] 또한 본 발명에서는 야생주와 비교하여 발현된 단백질의 접힘 (folding) 정도를 유지하는 알라닌 치환 변이체를 선택하였다. 결론적으로 각 α -2,3 및 α -2,6 시알산 전이효소의 알라닌 치환 변이체 중에서 야생주에 비해 30% 이상, 바람직하게는 50% 이상, 더욱 바람직하게는 60% 이상의 활성을 나타내고 단백질의 폴딩을 유지하는 잔기를 다음 단계인 포화 변이를 수행할 기능적 잔기로 선택하였다. 촉매 반응에 있어서 기질과의 상호작용에 필수적으로 기여하는, 즉 시알산 전이의 주된 활성에 기여하는 잔기들은 그대로 두고, 효소의 폴딩 및 본래 활성을 유지하는 알라닌 치환 변이체의 잔기에 포화 변이를 수행함으로써 효소의 ‘중립적 움직임’ (neutral drift) 을 유도할 수 있다. 이는 포화 변이를 통해서 야생주보다 기질에 더욱 적절하게 맞는 (fitting) 활성형의 효소-기질 복합체를 생성하는 것을 의미한다.

[0089] 본 발명에서는 상기한 시알산 전이효소의 변이체 탐색 방법을 통하여, 서열번호 1 내지 5의 아미노산 서열, 서열번호 1 내지 5 중 어느 하나의 아미노산 서열에서 313 번째 또는 265 번째의 아미노산이 친수성 아미노산으로

치환된 서열 및 서열번호 1 내지 5 중 어느 하나의 아미노산 서열에서 313 번째 또는 265 번째의 아미노산이 친수성 아미노산으로 치환된 서열로서, 서열번호 1 내지 5 와 97% 이상의 상동성을 가지는 아미노산 서열에서 선택되는 어느 하나의 서열로 표시되는 α -2,3 시알산 전이효소 변이체를 선택하였고, 서열번호 11 내지 15 의 아미노산 서열, 서열번호 11 내지 15 중 어느 하나의 아미노산 서열에서 411 번째의 아미노산이 작은 크기 또는 친수성 아미노산으로 치환되거나, 또는 433 번째의 아미노산이 친수성 아미노산으로 치환된 서열 및 서열번호 11 내지 15 중 어느 하나의 아미노산 서열에서 411 번째의 아미노산이 작은 크기 또는 친수성 아미노산으로 치환되거나, 또는 433 번째의 아미노산이 친수성 아미노산으로 치환된 서열로서, 서열번호 11 내지 15 와 55% 이상의 상동성을 가지는 아미노산 서열에서 선택되는 어느 하나의 서열로 표시되는 α -2,6 시알산 전이효소 변이체를 선택하였다.

[0090]

시알산 전이효소 변이체의 고유 활성도

[0091]

α -2,3 및 α -2,6 시알산 전이효소의 단일 아미노산 치환 변이체의 고유 활성도는 각각 동일한 양의 단백질을 사용하여 상기 pH 지시약을 이용한 효소 활성 분석 방법을 통해 분석할 수 있다.

[0092]

α -2,3 시알산 전이효소 변이체의 경우, R313 의 아르기닌은 락토오스의 글루코스 근처의 루프 (loop) 상에 위치하고 있다. 상기 R313 의 아르기닌이 알라닌과 글리신과 같은 작은 크기 아미노산으로 치환된 변이체는 야생주에 비해 중립적인 활성을 나타내고 세린, 트레오닌, 티로신, 아스파르트산, 아스파라긴 또는 히스티딘과 같은 친수성 아미노산으로 치환된 변이체들은 야생주에 비해 약 1.5 배 이상의 활성을 나타낸다. 또한, 시티딘 일인산-N-아세틸뉴라민산의 20 Å 이내에 위치하는 T265 의 트레오닌이 글리신과 같은 작은 크기 아미노산으로 치환된 변이체는 야생주에 비해 중립적인 활성을 나타내고 세린 또는 아스파라긴과 같은 친수성 아미노산으로 치환된 변이체들은 야생주에 비해 높은 활성을 나타낸다. 또한, 상기 R313 의 치환 및 T265 의 치환이 조합된 변이체는 단일 변이체보다 더욱 높은 활성을 나타낸다.

[0093]

한편, α -2,6 시알산 전이효소의 경우, I411 과 L433 은 시티딘 일인산-N-아세틸뉴라민산으로부터 5 내지 20 Å 거리 내에 위치하고 있다. 상기 L433 의 류신이 작은 크기 아미노산 또는 친수성 아미노산으로 치환된 변이체들은 야생주에 비해 증가된 활성을 나타내며, 특히 세린 또는 트레오닌으로 치환된 변이체는 야생주에 비해 약 3 배의 증가된 활성을 나타낸다. 또한, I411 의 이소류신이 작은 크기 아미노산 또는 친수성 아미노산으로 치환된 변이체들은 야생주에 비해 증가된 활성을 나타내며, 특히 트레오닌으로 치환된 변이체는 야생주에 비해 약 2 배의 증가된 활성을 나타낸다. 또한, 상기 I411 의 치환 및 L433 의 치환이 조합된 변이체는 단일 변이체보다 더욱 높은 활성을 나타내며, I411T/L433T 조합 변이체의 경우 야생주에 비해 약 5 배의 증가된 활성을 나타낸다.

[0094]

시알산 전이효소 변이체의 동역학 변수

[0095]

기질 공여체와 수용체에 대해 각각의 변이가 미치는 영향을 알아보기 위해, 단일 아미노산 치환 변이체와 조합적 변이체에 대해 동역학 변수를 측정한다. 동역학변수의 측정은 상기 비색법을 이용한 변이체 탐색 방법을 통해 분석할 수 있다.

[0096]

α -2,3 시알산 전이효소의 R313 및/또는 T265 이 친수성 아미노산으로 치환된 변이체들은 시티딘 일인산-N-아세틸뉴라민산과 락토오스에 대해서 k_{cat} 값이 증가하고, 두 기질에 대해 k_{cat}/K_m 이 증가한다. 특히, α -2,3 시알산 전이효소의 단일 아미노산 치환 변이체인 R313N 과 R313H 는 시티딘 일인산-N-아세틸뉴라민산과 락토오스에 대해서 k_{cat} 값이 증가하고, 두 기질에 대해 k_{cat}/K_m 은, R313N이 야생주 대비 약 1.4 배, R313H이 약 1.2 배 증가한다. 조합 변이체인 R313N/T265S 와 R313H/T265S 는 두 기질에 대해 k_{cat} 값이 증가하고, 두 조합 변이체는 시티딘 일인산-N-아세틸뉴라민산에 대해 k_{cat}/K_m 이 약 1.6 배 증가한다. 또한 R313N/T265S 와 R313H/T265S 는 락토오스에 대해서 k_{cat}/K_m 이 야생주 대비 각각 약 1.5배, 약 1.8배 증가한다.

[0097]

α -2,6 시알산 전이효소의 I411 및/또는 L433 이 작은 크기 아미노산 또는 친수성 아미노산으로 치환된 변이체들은 시티딘 일인산-N-아세틸뉴라민산과 락토오스에 대해서 k_{cat} 값이 증가한다. 특히, 수용체 기질인 락토오스에 대해서, 모든 변이체가 기질과의 결합력은 감소하는데 반해 k_{cat} 값이 증가하며, 야생주에 비해 k_{cat} 값이 6 배에서 최대 27 배까지 증가한다. 이에 대해, 단일 아미노산 치환 변이체인 I411T는 시티딘 일인산-N-아세틸뉴라민산과 락토오스에 대해서 k_{cat}/K_m 이 각각 야생주 대비 약 2.4 배, 약 1.8 배 증가한다. 또한, L433S 와 L433T 의 k_{cat}/K_m 은 시티딘 일인산-N-아세틸뉴라민산에 대해서는 각각 약 2.6 배, 약 6.7 배

증가하며, 락토오스에 대해서는 약 3 배, 약 2.6 배씩 증가한다. 조합 변이체 중, I411T/L433T 는 시티딘 일인산-N-아세틸뉴라민산과 락토오스에 대해 k_{cat}/K_m 이 야생주 대비 각각 약 8 배, 약 3.9 배 증가한다.

[0098] α -2,3 시알산 전이효소 변이주의 α -2,6 시알산 전이 부반응 활성

[0099] 본 발명에서는 또한 α -2,3 시알산 전이효소의 313 번째 아미노산 위치에 아르기닌을 대체하여, 다른 친수성 아미노산 (N, D, Y, T, H) 으로 전환됐을 때, 효소의 고유 활성도가 증가하는 것을 확인함과 함께, 이들 변이체에 대해서 α -2,6 시알산 전이 부반응을 확인할 수 있다.

[0100] α -2,3 시알산 전이효소 변이체의 α -2,6 시알산 전이의 부반응성은 pH 4.5 내지 6.0 에서 6 '-시알릴락토오스의 생성량이 야생주에 비해 4 내지 30 배 감소하며, pH 6.5 내지 pH 7.0 에서는 6' -시알릴락토오스의 생성량이 없거나 현저히 감소한다.

[0101] [실시예]

[0102] 본 발명의 구체적인 방법을 실시예로서 상세히 설명하나, 본 발명의 기술적 범위가 이들 실시예에 한정되는 것은 아니다.

[0103] pH 지시약을 이용한 효소 활성 분석 방법

[0104] 당전이 효소의 비색법 (colorimetric method) 을 통한 활성 측정은 당 공여체와 수용체간의 글리코시드 결합 (glycosidic bond) 이 형성될 때 발생하는 수소 이온 (proton)에 의한 pH의 변화를 측정하는 방법으로서 시알릴 올리고당이 생산성과 비례한다.

[0105] 본 발명에서는 α -2,3 시알산 전이효소의 최적 활성 pH 에 의해 크레졸 레드 (cresol red) 와 α -2,6 시알산 전이효소의 최적 활성 pH 에 의한 페놀 레드 (phenol red) 를 사용하였다. 효소 활성 분석은 1 내지 5 mM 시티딘 일인산-N-아세틸뉴라민산과 락토오스, 0.1 내지 1 mM pH 지시약을 1 내지 10 mM 트리스 (Tris-HCl) 완충 용액에 섞어 총 100 μ L 부피로 96-정 (well) 플레이트에서 진행하였다. 반응은 5 내지 10 분간 진행하였으며, 크레졸 레드의 경우 노란색으로 변하는 405 nm 에서의 흡광도의 증가를, 페놀 레드의 경우 붉은 색이 감소하는 560 nm 에서의 흡광도의 감소를 30초의 간격을 두고 스펙트럼 기계를 이용하여 분석하였다. 발생하는 수소이온의 농도는 염산 (HCl) 의 농도에 따른 검정 곡선에 의해 계산되었다. 효소 반응은 전체 반응 부피의 10% 의 효소를 이용하였으며, 효소에 기질 혼합물을 넣음과 함께 시작되었다. 효소 활성 (Unit) 은 상온에서 분당 글리코시드 결합이 형성되는 동안 발생하는 1 μ mol e 의 수소 이온을 생성하는데 필요한 효소의 양으로 정의하였다.

[0106] 포화 변이 수행 및 비색법을 이용한 변이체 탐색 방법

[0107] α -2,3 시알산 전이효소의 아미노산 위치 313 및 265 에 해당하는 AGA 및 ACC 서열을 무작위 하게 치환한 NNK 서열 (N 은 A, C, G 또는 T 이고, K 는 G 또는 T 인 서열) 이 도입된 프라이머를 이용하여 벡터 전체를 PCR하여 라이브러리를 구축하였다. 본 발명의 α -2,3 시알산 전이효소는 N 말단에 24 개의 아미노산이 제거된 형태이기 때문에 첫 메티오닌 서열부터 헤아릴 경우 메티오닌이 25 번이 된다.

[0108] α -2,6 시알산 전이효소의 아미노산 위치 411 및 433 에 해당하는 ATT 및 CTG 서열을 무작위 하게 치환한 NNK 서열 (N 은 A, C, G 또는 T 이고, K 는 G 또는 T 인 서열) 이 도입된 프라이머를 이용하여 벡터 전체를 PCR 하여 라이브러리를 구축하였다. 본 발명의 α -2,6 시알산 전이효소는 첫 메티오닌 서열부터 헤아릴 경우 메티오닌이 1 번이 된다.

[0109] 벡터 서열을 포함한 시알산 전이효소의 증폭된 유전자는 본래의 플라스미드 (plasmid) 를 제거하기 위해 Dpn I 효소 처리 후, 대장균 DH5 α 에 형질전환시켰다. 발생된 모든 콜로니 (colony) 로부터 변이 유전자를 추출하여, 이를 대장균 BW25113 (DE3)에 형질전환 하였다. 형질전환된 각각의 콜로니를 96-정 상에서 암피실린이 포함된 LB 배지 500 μ L 에 접종하여 30 내지 37 $^{\circ}$ C 에서 18 내지 24 시간 동안 진탕 배양 후, 배양액 일부를 100 μ g mL⁻¹ 암피실린과 IPTG (Isopropyl β -D-1-thiogalactopyranoside) 가 포함된 새로운 LB 배지 500 μ L 에 접종하여 18 내지 30 $^{\circ}$ C 에서 18 내지 40 시간을 배양하였다. 배양된 세포들은 원심분리 후, 세포를 1 내지 10 mM 트리스 완충용액 100 μ L 에 재부유 시킨 다음 이 중 10 μ L 의 전세포를 변이체 탐색 반응에 사용하거나 또는 세포를 50 μ L 의 버그 부스터 (BugBuster) 단백질 추출 시약으로 재부유 시켜 원심분리 후에 세포 추출물을 얻어 이 중 10 μ L 를 변이체 탐색 반응에 사용하였다. 90 μ L 의 반응 용액은 1 내지 10 mM 트리스 완충용액, 1 내지 5 mM 시티딘 일인산-N-아세틸뉴라민산과 락토오스, 0.1 내지 1 mM pH 지시약을 포함하며 시알산

전이효소의 전세포 또는 세포 추출액 10 μ L 를 첨가함과 동시에 반응을 진행하여 1 분의 시간 간격으로 10-30 분간의 반응 속도를 야생주와 비교하여 관찰하였다.

[0110] 시알산 전이 효소의 정제

[0111] 대장균 BW25113 (DE3) 에 형질전환된 시알산 전이효소의 야생주 및 변이체는 50 mL 의 배양 부피로 인두서인 IPTG 를 이용하여 발현 한 이후 Ni-NTA 컬럼을 이용하여 순수 단백질을 정제하였다. 먼저 단백질 발현 이후, 음과 파쇄기로 세포를 파쇄한 후, 원심분리 후 세포 추출액을 얻었다. 5 mM 이미다졸 (imidazole) 과 300 mM 염화나트륨이 첨가된 50 mM 트리스 완충용액 (pH 8.0) 으로 평형화 시킨 컬럼에 세포 추출액을 넣어 0°C 에서 1 시간 동안 니켈 수지 (resin) 와 결합을 시켰다. 이후 수지에 결합하지 못한 단백질을 흘려버리고 50 mM 이미다졸이 포함된 트리스 완충용액으로 비특이적으로 결합된 다른 단백질들을 제거하였다. 마지막으로 250 mM 이미다졸이 포함된 트리스 완충용액으로 원하는 단백질을 용출하였다. 용출된 단백질은 이미다졸을 제거하기 위해, 여과 컬럼을 이용한 탈염과정을 수행하여 최종적으로 활성 있는 단백질을 얻었으며 브래드포드 (Bradford) 단백질 정량 키트를 사용하여 단백질량을 측정하여 동일한 양의 단백질을 사용하여 반응한 후, 고유 활성도, 동역학 변수 및 α -2,3 시알산 전이효소 변이주의 α -2,6 시알산 전이 부반응 활성을 측정하였다.

실시예 1

[0112] 시알산 전이효소 변이체의 고유 활성도 측정

[0113] α -2,3 및 α -2,6 시알산 전이효소의 단일 아미노산 치환 변이체의 고유 활성도는 각각 동일한 양의 단백질을 사용하여 상기 pH 지시약을 이용한 효소 활성 분석 방법을 통해 분석하였으며, 반응은 5 내지 10 분간 진행하여 초기 수용체 기질 농도 대비 10 내지 25% 의 전환 수율을 나타냈을 때의 효소 mg 당 활성 (unit) 으로 계산하였으며, 이를 도 2 에 나타내었다.

[0114] α -2,3 시알산 전이효소 변이체의 경우, R313 의 아르기닌은 락토오스의 글루코스 근처의 루프 (loop) 상에 위치하고 있다. R313 의 변이체들에 대해, 알라닌과 글리신과 같은 작은 크기 아미노산으로의 치환된 변이체는 야생주에 비해 중립적인 활성을 나타냈고 세린, 트레오닌, 티로신, 아스파르트산, 아스파라긴 또는 히스티딘과 같은 친수성 아미노산으로 치환된 변이체들은 야생주에 비해 1.5 배 이상의 활성을 나타냈다.

[0115] 시티딘 일인산-N-아세틸뉴라민산의 20 Å 이내에 위치하는 T265 변이체에 대해서는 글리신, 세린 또는 아스파라긴으로 치환된 변이체가 야생주와 비슷하거나 높은 활성을 가진 것으로 탐색되었다.

[0116] R313 변이체에 대해 야생형 대비 상대적인 고유 활성도를 비교한 결과, 도 2 와 같으며, 결과적으로 R313 은 상대적으로 다양한 변이체들을 받아들일 수 있었으며, 그 중 R313N 의 고유 활성도가 야생주 대비 231% 로서 단일 변이체 중 가장 높았다. 또한 T265 변이체에 대한 야생형 대비 상대적인 고유 활성도를 도 2 에 나타냈다.

[0117] 반면, α -2,6 시알산 전이효소의 경우, I411 과 L433 은 시티딘 일인산-N-아세틸뉴라민산으로부터 5 내지 20 Å 거리 내에 위치하고 있다. L433 의 변이체들 중에서 L433S 와 L433T가 야생형 대비 3 배의 증가된 활성을 나타냈다. I411 의 변이체들 중에서는 I411T 가 야생형 대비 2 배의 활성을 나타내는 것으로 탐색되었다. 야생형 α -2,6 시알산 전이효소의 경우 pET15b 벡터보다 pET28a 벡터에서의 발현이 증가되는 바, 탐색된 변이체들을 pET28a 벡터에 클로닝하여 이들의 고유 활성도를 확인하였다. 변이체의 고유 활성도를 도 2 (b) 에 나타냈다.

[0118] 본 발명에서는 또한 높은 고유 활성도를 가지는 α -2,3 시알산 전이효소의 R313 의 단일 아미노산 치환 변이체와 T265 의 단일 아미노산 치환 변이체에 대해서 각각의 조합적인 변이체를 만들었으며, 조합적 변이체 중 R313H/T265S 와 R313N/T265S 가 야생주 대비 237%, 216% 의 가장 높은 고유 활성도를 나타냈다.

[0119] 본 발명에서는 높은 고유 활성도를 가지는 α -2,6 시알산 전이효소의 I411T 의 단일 아미노산 치환 변이체와 L433S 및 L433T 의 단일 아미노산 치환 변이체에 대해서 각각의 조합적인 변이체를 만들었으며, 조합적 변이체 중 I411T/L433S 와 I411T/L433T 가 야생주 대비 194%, 510% 의 높은 고유 활성도를 나타냈다.

실시예 2

[0120] 시알산 전이효소 변이체의 동역학 변수 측정

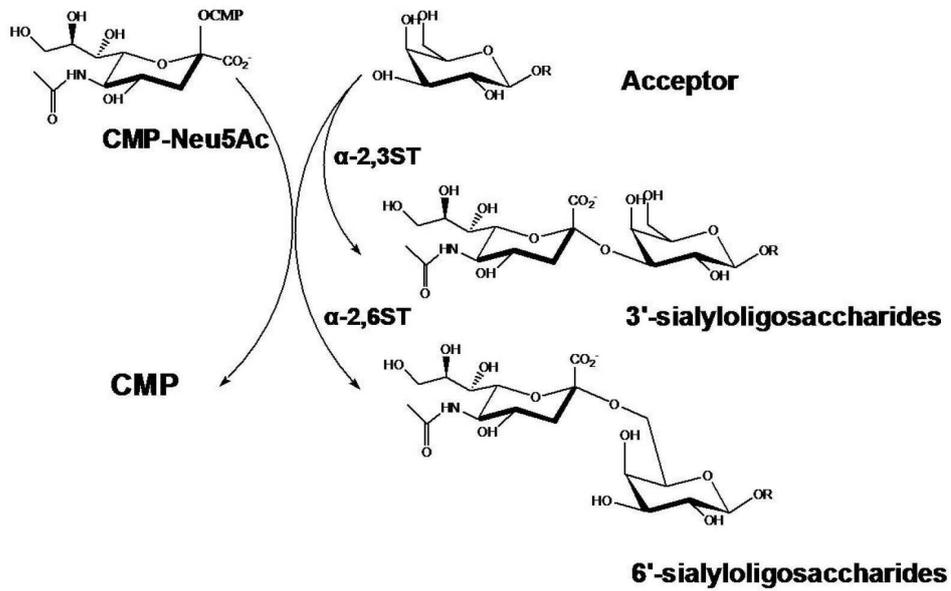
- [0121] 동역학변수의 측정은 상기 비색법을 이용한 변이체 탐색 방법을 통해 분석하였고, 반응은 실온에서 5 내지 10 분간 수행하여 30 초마다의 간격을 두고 초기 수용체 기질 농도 대비 10 내지 25%의 전환 수율을 나타냈을 때의 초기 반응속도를 측정하였다. 동역학 변수의 측정은 공여체 기질인 시티딘 일인산-*N*-아세틸뉴라민산과 수용체 기질인 락토오스 두 가지에 대해 모두 측정하였으며 기질 농도는 0.1에서 30 mM까지의 범위를 사용하였다. k_{cat} 과 K_m 의 동역학 변수는 시그마 플롯 (SigmaPlot) 프로그램을 사용하여 미카엘리스-멘텐 (Michaelis-Menten) 방정식의 비선형 회귀분석을 통해 얻었다. α -2,3 및 α -2,6 시알산 전이효소의 야생주와 변이체에 대한 동역학변수는 도 3에 나타내었다.
- [0122] α -2,3 시알산 전이효소의 단일 아미노산 치환 변이체인 R313N과 R313H는 시티딘 일인산-*N*-아세틸뉴라민산과 락토오스에 대해서 k_{cat} 값을 증가시켰고, 두 기질에 대해 k_{cat}/K_m 은 R313N은 야생주 대비 1.4배, R313H는 1.2배 증가하였다. 조합 변이체인 R313N/T265S와 R313H/T265S는 두 기질에 대해 k_{cat} 값을 증가시켰고 두 조합 변이체는 시티딘 일인산-*N*-아세틸뉴라민산에 대해 k_{cat}/K_m 을 1.6배 증가시켰다. 또한 락토오스에 대해서는 R313N/T265S와 R313H/T265S가 k_{cat}/K_m 을 야생주 대비 각각 1.5배, 1.8배 증가시켰다.
- [0123] α -2,6 시알산 전이효소의 변이체의 경우, 두 기질에 대해서 모두 k_{cat} 값을 증가시켰다. 특히, 수용체 기질인 락토오스에 대해서, 모든 변이체가 기질과의 결합력은 감소하는데 반해 k_{cat} 값을 증가시키며, 야생주에 비해 k_{cat} 값이 6배에서 최대 27배까지 증가하였다. 단일 아미노산 치환 변이체인 I411T는 시티딘 일인산-*N*-아세틸뉴라민산과 락토오스에 대해서 k_{cat}/K_m 이 각각 야생주 대비 2.4배, 1.8배 증가하였다. 또한, L433S와 L433T의 k_{cat}/K_m 은 시티딘 일인산-*N*-아세틸뉴라민산에 대해서는 각각 2.6배, 6.7배 증가하였으며, 락토오스에 대해서는 3배, 2.6배씩 증가하였다. 조합 변이체 중, I411T/L433T는 I411T/L433S보다 활성 증가의 효과가 컸으며 시티딘 일인산-*N*-아세틸뉴라민산과 락토오스에 대해 각각 k_{cat}/K_m 이 야생주 대비 8배, 3.9배 증가하였다.

실시예 3

- [0124] α -2,3 시알산 전이효소 변이주의 α -2,6 시알산 전이 부반응 활성
- [0125] 단백질 정제를 마친 α -2,3 시알산 전이효소의 R313 변이체인 R313N, R313D, R313Y, R313T, R313H, R313N/T265S, R313H/T265S에 대해 α -2,6 시알산 전이 부반응 활성을 측정하였다. 효소 반응은 10 mM 시티딘 일인산-*N*-아세틸뉴라민산과 5 mM 락토오스가 포함된 50 mM MES의 pH 4.5 내지 pH 7의 완충 용액상에서 동일한 양의 정제된 효소를 넣고 상온에서 30 분간 반응하였다. 반응 후, 1 분간 열을 가하여 단백질을 변성시키고 원심 분리 후 제거, 상등액 상에 생성된 3'-시알릴락토오스 및 6'-시알릴락토오스는 Bio-LC를 통해 정량분석하였다. α -2,3 시알산 전이효소 변이체의 α -2,6 시알산 전이의 부반응성을 측정한 결과, pH 4.5 내지 6.0에서는 6'-시알릴락토오스의 생성량이 야생주에 비해 4-30배 감소하였으며, pH 6.5 내지 pH 7.0에서는 6'-시알릴락토오스의 생성량이 R313Y (15배 감소)를 제외하고, 모두 사라진 것을 확인할 수 있었으며, 이는 도 4에 나타냈다.

도면

도면1



도면2

효소 (α -2,3PST)	상대적인 고유 활성도 (%)	효소 (α -2,6PdST)	상대적인 고유 활성도 (%)
WT	100	WT	100
R313N	231	I411T	198
R313H	146	L433S	289
R313T	129	L433T	296
R313Y	125	I411T/L433S	194
R313D	108	I411T/L433T	510
T265N	126		
T265S	116		
T265G	94		
R313N/T265S	216		
R313H/T265S	237		

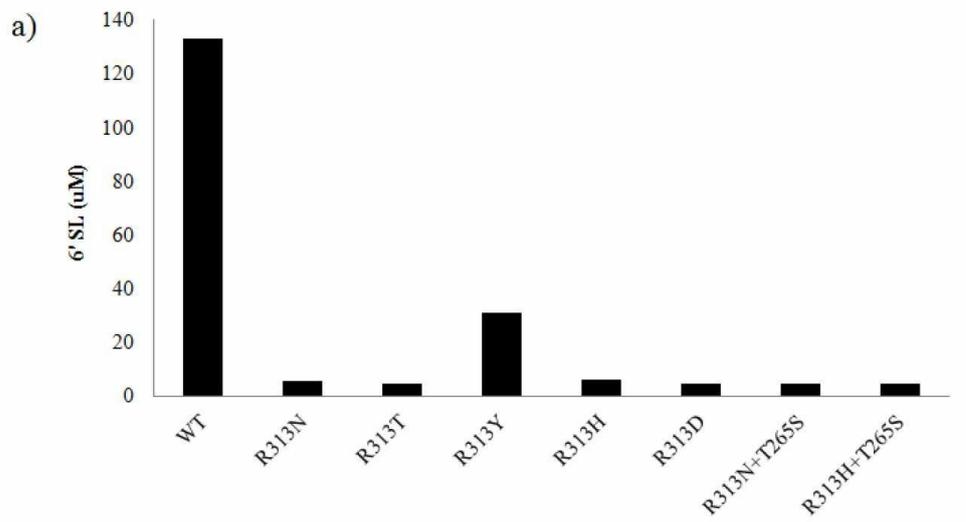
도면3a

효소 (α -2,3 PST)	CMP-Neu5Ac ^a			Lac ^b		
	k_{cat} (s ⁻¹)	K_m (mM)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)	k_{cat} (s ⁻¹)	K_m (mM)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)
WT	12.9	1.83	7.01	57.4	2.31	24.8
R313N	23.9	2.51	9.51	89.9	2.55	35.2
R313H	26.8	3.24	8.26	80.5	2.68	30.1
T265S	22.1	2.23	9.91	63.2	2.43	26.1
R313N/T265S	36.9	3.33	11.1	82.4	2.21	37.3
R313H/T265S	43.4	3.94	11.0	73.9	1.65	44.8

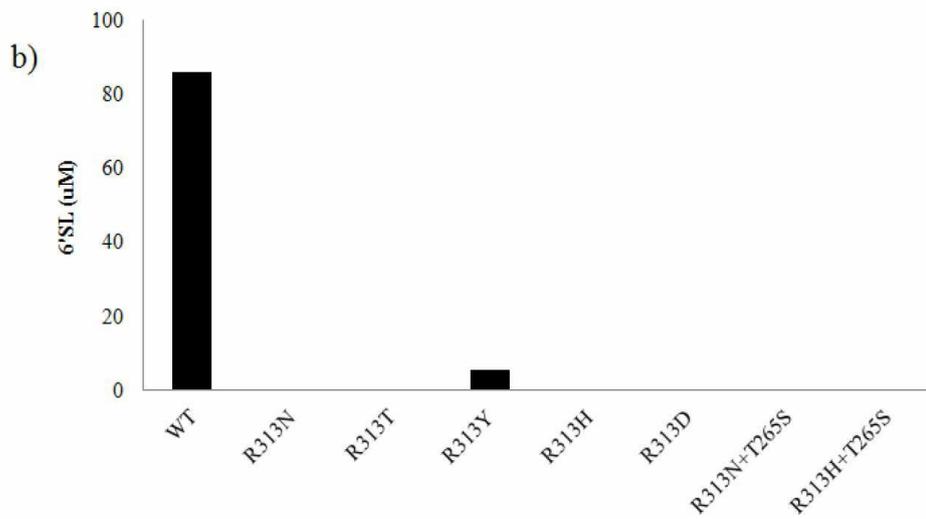
도면3b

효소 (α -2,6 PdST)	CMP-Neu5Ac ^a			Lac ^b		
	k_{cat} (s ⁻¹)	K_m (mM)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)	k_{cat} (s ⁻¹)	K_m (mM)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)
WT	3.99	5.47	0.73	4.45	9.03	0.49
I411T	11.1	6.36	1.75	25.7	28.4	0.90
L433S	18.3	9.54	1.92	85.4	59.1	1.45
L433T	10.4	2.11	4.90	73.2	56.7	1.29
I411T/L433S	17.4	11.9	1.47	113	78.5	1.31
I411T/L433T	18.1	3.09	5.86	119	57.0	1.90

도면4a



도면4b



도면5a

(a)
 MKTITLYLDPASLPALNQLMDFQNNEDKTHPRIFGLSRFKIPDNII TQYQNIHFVELKDNR PTEALFTILDQY PGNIEL
 NIHLNIAHSVQLIRPILAYRFKHLDRVSIQQNLNLYDDGSM EYVDLEKEENKDISAEIKQAEKQLSHYLLTGKIKFDNPTI
 ARYVWQSAPV KYHFLS TDYFEKAEFLQPLKEYLAENYQKMDWTAYQQLTPEQQAFYLT LVGFNDEVKQSLEVQQAKFIF
 TGTTTWEQNTDVREYYAQQLNLLNHFQAEGDLF IG DHYKIYFKGHPNGGEINDYI LNNAKNITNIPANISFEVLMMTG
 LLPDKVGGVASSLYFSLPKEKISHIIFTSNKQVKS KEDALNPNYVKVMRRLGII DESQVIFWDSLKQLGGG

도면5b

(b)
MKTTITLYLDPASLPA LNQLMDF TQNNEDKTHPRIFGLSRFKIPDNII TQYQNIHFVELKDNRPTEALFTILDQYPGNIEL
NIHLNIAHSVQLIRPILAYRFKHLDRVSIQQLNLYDDGSM EYVDLEKEENKDISAEIKQAEKQLSHYLLTGKIKFDNPTI
ARYVWQSAFPVKYHFLSTDYFEKAEFLQPLKEYLAENYQKMDWTAYQQLTPEQQAFYLLTVGFNDEVKQSLVQQA KFI F
TGTTTWE GNTDVREYYAQQQLNLLNHFTQAE GDLF IG DHYKIYFKGHPHGGEINDYI LNNAKNITNIPANISFEVLM MTG
LLPDKVGGVASSLYFSLPKEKISHIIFT SNKQVKS KEDALNNPYVKVMRRLGIIDESQVIFWDSLKQLGGG

도면5c

(c)
MKTTITLYLDPASLPA LNQLMDF TQNNEDKTHPRIFGLSRFKIPDNII TQYQNIHFVELKDNRPTEALFTILDQYPGNIEL
NIHLNIAHSVQLIRPILAYRFKHLDRVSIQQLNLYDDGSM EYVDLEKEENKDISAEIKQAEKQLSHYLLTGKIKFDNPTI
ARYVWQSAFPVKYHFLSTDYFEKAEFLQPLKEYLAENYQKMDWTAYQQLTPEQQAFYLLTVGFNDEVKQSLVQQA KFI F
SGTTTWE GNTDVREYYAQQQLNLLNHFTQAE GDLF IG DHYKIYFKGHPRGGEINDYI LNNAKNITNIPANISFEVLM MTG
LLPDKVGGVASSLYFSLPKEKISHIIFT SNKQVKS KEDALNNPYVKVMRRLGIIDESQVIFWDSLKQLGGG

도면5d

(d)
MKTTITLYLDPASLPA LNQLMDF TQNNEDKTHPRIFGLSRFKIPDNII TQYQNIHFVELKDNRPTEALFTILDQYPGNIEL
NIHLNIAHSVQLIRPILAYRFKHLDRVSIQQLNLYDDGSM EYVDLEKEENKDISAEIKQAEKQLSHYLLTGKIKFDNPTI
ARYVWQSAFPVKYHFLSTDYFEKAEFLQPLKEYLAENYQKMDWTAYQQLTPEQQAFYLLTVGFNDEVKQSLVQQA KFI F
SGTTTWE GNTDVREYYAQQQLNLLNHFTQAE GDLF IG DHYKIYFKGHPNGGEINDYI LNNAKNITNIPANISFEVLM MTG
LLPDKVGGVASSLYFSLPKEKISHIIFT SNKQVKS KEDALNNPYVKVMRRLGIIDESQVIFWDSLKQLGGG

도면5e

(e)
MKTTITLYLDPASLPA LNQLMDF TQNNEDKTHPRIFGLSRFKIPDNII TQYQNIHFVELKDNRPTEALFTILDQYPGNIEL
NIHLNIAHSVQLIRPILAYRFKHLDRVSIQQLNLYDDGSM EYVDLEKEENKDISAEIKQAEKQLSHYLLTGKIKFDNPTI
ARYVWQSAFPVKYHFLSTDYFEKAEFLQPLKEYLAENYQKMDWTAYQQLTPEQQAFYLLTVGFNDEVKQSLVQQA KFI F
SGTTTWE GNTDVREYYAQQQLNLLNHFTQAE GDLF IG DHYKIYFKGHPHGGEINDYI LNNAKNITNIPANISFEVLM MTG
LLPDKVGGVASSLYFSLPKEKISHIIFT SNKQVKS KEDALNNPYVKVMRRLGIIDESQVIFWDSLKQLGGG

도면6a

(a)
ATGAAAACAATCACGCTGTATTAGATCCTGCCTCCTTACCGGCATTAAATCAGCTGATGGACTTTACGCAAAATAATGA
AGATAAAACACATCCACGATTTTGGTCTTTCTCGCTTTAAAATCCCTGACAACATTTATACACAGTATCAAAATATCC
ATTTTCGTCGAACCTCAAAGATAATCGTCCCCTGAAGCACTTTTACGATTTTAGATCAATACCCCTGGTAACATGAGTTA
AATATACACTTAAATATGCTCATTCGGTCAATTAATTCGTCGATTTGGCATATCGTTTAAACATTTAGATCGTGT
ATCAATCAGCAGTTAAATCTTATGACGATGGCTCAATGGAATATGTTGATTTAGAAAAAGAGAAAATAAGATATTT
CCGCAGAAATAAGCAAGCAGAAAAACAACCTTCTCACTATTTGCTTACTGGCAAAATAAAATTTGATAACCCAACTATT
GCTCGTTATGCTGGCAATCCGCGTCCCAGTAAAATATCATTTTAAAGTACAGACTATTTGAAAAAGCCGAATTTT
ACAACCACTAAAAGAAATTTAGCAGAAAAATCAAAAAATGGACTGGACTGCTTACCAACAGCTGACTCCAGAACAGC
AAGCATTCTACTTAAACATTTGGTGGCTTCAATGACGAAGTCAAGCAGTCGCTAGAAGTGCAACAAGCTAAATTTATCTTT
ACCGGCACGACAACCTGGGAAGGAAATACCGATGTCGAGAATACTACGCACAGCAACAACCTAATTTACTTTAATCACTT
TACCAAGCTGAGGGCGATTTATTTATGGTGATCATTAATAAATCTACTTTAAAGGGCATCCTAATGGTGGTGAAATTA
ATGACTACATTTCTGAACAATGCTAAAAATATCACC AATATCCCTGCCAATATTTCTTTGAAGTATTGATGATGACAGGC
TTATTAACCTGATAAAGTGGGTGGTGTGCAAGTTCCTGATTTCTCCTTACCAAAAAGAAAAATAGCCATATATTTT
CACATCGAATAACAAGTGAAAAGCAAAGAAGATGCGCTAAATAATCCGATGTTAAGGTATGCGCTCGTTTAGGTATAA
TTGACGAATCACAAGTCATCTTTGGACAGTTTAAAACAGTTGGGTGGAGGT

도면6b

(b)

ATGAAAACAATCACGCTGATTTAGATCCTGCCTCCTTACCGGCATTAAATCAGCTGATGGACTTTACGCAAAAT AATGA
 AGATAAAACACATCCACGTAATTTTGGTCTTTCTCGCTTTAAAATCCCTGACAACATTAATACACAGTATCAAAATATCC
 ATTTTCGTCGAACCTCAAAGATAATCGTCCCCTGAAGCACTTTTACGATTTTAGATCAATACCCCTGGTAACATTGAGTTA
 AATATACACTTAAATATGCTCATTCGGTCAATTAATTCGTCCGATTTTGGCATATCGTTTAAACATTTAGATCGTGT
 ATCAATTCAGCAGTTAAATCTTATGACGATGGCTCAATGGAATATGTTGATTTAGAAAAAGAAGAAAAATAAAGATATTT
 CCGCAGAAATTAAGCAAGCAGAAAAACAACCTTCTCACTATTTGCTTACTGGCAAAATAAAATTTGATAACCCAACATAT
 GCTCGTTATGCTGGCAATCCGCTTCCCAGTAAAATATCATTTTTAAGTACAGACTATTTGAAAAAGCCGAATTTTT
 ACAACCACTAAAAGAATATTTAGCAGAAAAATATCAAAAAATGGACTGGACTGCTTACCAACAGCTGACTCCAGAACAGC
 AAGCATTCTACTTAACATTTGGTAGGCTTCAATGACGAAGTCAAGCAGTCGCTAGAAGTGCAACAAGCTAAATTTATCTTT
 AGTGGCAGCACAACCTGGGAAGGAAATACCGATGTGCGAGAATACTACGCACAGCAACAACCTAAATTTACTTAATCACTT
 TACCCAAGCTGAGGGCGATTTATTTATGGTGATCATTAATAAATCTACTTTAAAGGGCATCCTCATGGTGGTGAATTA
 ATGACTACATTTCTGAACAATGCTAAAAATATCACCAATATCCCTGCCAATATTTCTTTGAAGTATTGATGATGACAGGC
 TTATTACCTGATAAAGTGGTGGTGTGCAAGTTCACTGTATTTCTCCTTACCAAAAGAAAAAATAGCCATATTTATTT
 CACATCGAATAACAAGTGAAGCAAGCAAGAAAGATGCGCTAAATAATCCGATGTTAAGGTATGCTCGTTAGGTATAA
 TTGACGAATCACAAAGTCATCTTTGGGACAGTTTAAAACAGTTGGGTGGAGGT

도면6c

(c)

ATGAAAACAATCACGCTGATTTAGATCCTGCCTCCTTACCGGCATTAAATCAGCTGATGGACTTTACGCAAAAT AATGA
 AGATAAAACACATCCACGTAATTTTGGTCTTTCTCGCTTTAAAATCCCTGACAACATTAATACACAGTATCAAAATATCC
 ATTTTCGTCGAACCTCAAAGATAATCGTCCCCTGAAGCACTTTTACGATTTTAGATCAATACCCCTGGTAACATTGAGTTA
 AATATACACTTAAATATGCTCATTCGGTCAATTAATTCGTCCGATTTTGGCATATCGTTTAAACATTTAGATCGTGT
 ATCAATTCAGCAGTTAAATCTTATGACGATGGCTCAATGGAATATGTTGATTTAGAAAAAGAAGAAAAATAAAGATATTT
 CCGCAGAAATTAAGCAAGCAGAAAAACAACCTTCTCACTATTTGCTTACTGGCAAAATAAAATTTGATAACCCAACATAT
 GCTCGTTATGCTGGCAATCCGCTTCCCAGTAAAATATCATTTTTAAGTACAGACTATTTGAAAAAGCCGAATTTTT
 ACAACCACTAAAAGAATATTTAGCAGAAAAATATCAAAAAATGGACTGGACTGCTTACCAACAGCTGACTCCAGAACAGC
 AAGCATTCTACTTAACATTTGGTAGGCTTCAATGACGAAGTCAAGCAGTCGCTAGAAGTGCAACAAGCTAAATTTATCTTT
 AGTGGCAGCACAACCTGGGAAGGAAATACCGATGTGCGAGAATACTACGCACAGCAACAACCTAAATTTACTTAATCACTT
 TACCCAAGCTGAGGGCGATTTATTTATGGTGATCATTAATAAATCTACTTTAAAGGGCATCCTAGAGGTGGTGAATTA
 ATGACTACATTTCTGAACAATGCTAAAAATATCACCAATATCCCTGCCAATATTTCTTTGAAGTATTGATGATGACAGGC
 TTATTACCTGATAAAGTGGTGGTGTGCAAGTTCACTGTATTTCTCCTTACCAAAAGAAAAAATAGCCATATTTATTT
 CACATCGAATAACAAGTGAAGCAAGCAAGAAAGATGCGCTAAATAATCCGATGTTAAGGTATGCTCGTTAGGTATAA
 TTGACGAATCACAAAGTCATCTTTGGGACAGTTTAAAACAGTTGGGTGGAGGT

도면6d

(d)

ATGAAAACAATCACGCTGATTTAGATCCTGCCTCCTTACCGGCATTAAATCAGCTGATGGACTTTACGCAAAAT AATGA
 AGATAAAACACATCCACGTAATTTTGGTCTTTCTCGCTTTAAAATCCCTGACAACATTAATACACAGTATCAAAATATCC
 ATTTTCGTCGAACCTCAAAGATAATCGTCCCCTGAAGCACTTTTACGATTTTAGATCAATACCCCTGGTAACATTGAGTTA
 AATATACACTTAAATATGCTCATTCGGTCAATTAATTCGTCCGATTTTGGCATATCGTTTAAACATTTAGATCGTGT
 ATCAATTCAGCAGTTAAATCTTATGACGATGGCTCAATGGAATATGTTGATTTAGAAAAAGAAGAAAAATAAAGATATTT
 CCGCAGAAATTAAGCAAGCAGAAAAACAACCTTCTCACTATTTGCTTACTGGCAAAATAAAATTTGATAACCCAACATAT
 GCTCGTTATGCTGGCAATCCGCTTCCCAGTAAAATATCATTTTTAAGTACAGACTATTTGAAAAAGCCGAATTTTT
 ACAACCACTAAAAGAATATTTAGCAGAAAAATATCAAAAAATGGACTGGACTGCTTACCAACAGCTGACTCCAGAACAGC
 AAGCATTCTACTTAACATTTGGTAGGCTTCAATGACGAAGTCAAGCAGTCGCTAGAAGTGCAACAAGCTAAATTTATCTTT
 AGCGGCAGCACAACCTGGGAAGGAAATACCGATGTGCGAGAATACTACGCACAGCAACAACCTAAATTTACTTAATCACTT
 TACCCAAGCTGAGGGCGATTTATTTATGGTGATCATTAATAAATCTACTTTAAAGGGCATCCTAATGGTGGTGAATTA
 ATGACTACATTTCTGAACAATGCTAAAAATATCACCAATATCCCTGCCAATATTTCTTTGAAGTATTGATGATGACAGGC
 TTATTACCTGATAAAGTGGTGGTGTGCAAGTTCACTGTATTTCTCCTTACCAAAAGAAAAAATAGCCATATTTATTT
 CACATCGAATAACAAGTGAAGCAAGCAAGAAAGATGCGCTAAATAATCCGATGTTAAGGTATGCTCGTTAGGTATAA
 TTGACGAATCACAAAGTCATCTTTGGGACAGTTTAAAACAGTTGGGTGGAGGT

도면6e

(e)

ATGAAAACAATCACGCTGATTTAGATCCTGCCTCCTACCGCATTAAATCAGCTGATGGACTTTACGCAAAATAATGA
 AGATAAACACATCCACGATTTTTGGTCTTTCTCGCTTTAAATCCCTGACAACATATTTACACAGTATCAAAATATCC
 ATTTTCGTCGAACCTCAAAGATAATCGTCCCACTGAAGCACTTTTACGATTTTATAGATCAATACCCTGGTAACATTGAGTTA
 AATATACACTTAAATATGCTCATTCCGTTCAATTAATTCGTCGATTTTGGCATATCGTTTTAAACATTTAGATCGTGT
 ATCAATTCAGCAGTTAAATCTTTATGACGATGGCTCAATGGAATATGTTGATTTAGAAAAAGAAGAAAAAAGATATTT
 CCGCAGAAATTAAGCAAGCAGAAAAACAACCTTCTCACTATTGCTTACTGGCAAAATAAAATTTGATAACCCAATAT
 GCTCGTTATGCTGGCAATCCGCGTTCCCACTAAAATATCATTTTTAAGTACAGACTATTTGAAAAAGCCGAATTTTT
 ACAACCACTAAAAGAATATTTAGCAGAAAAATATCAAAAAATGGACTGGACTGCTTACCAACAGCTGACTCCAGAACAGC
 AAGCATTCTACTTAACATTTGGTAGGCTCAATGACGAAGTCAAGCAGTCGCTAGAAGTGCAACAAGCTAAATTTATCTTT
 AGCGGCACGACAACCTGGGAAGGAAATACCGATGTGCGAGAATACTACGCACAGCAACAACCTTAATTTACTTTAACTCTT
 TACCAAGCTGAGGGCGATTTATTTATGGTGATCATTATAAAATCTACTTTAAAGGGCATCCTCATGGTGGTGAATTA
 ATGACTACATCTGAACAATGCTAAAAATATCACCAATATCCCTGCCAATATTTCTTTGAAGTATTGATGATGACAGG
 TTATACCTGATAAAGTGGTGGTGTGCAAGTCTACTGATTTCTCCTTACCAAAAGAAAAATAGCCATATATTTTT
 CACATCGAATAACAAGTGAAGCAAGCAAGATGCGCTAAAATAATCCGATGTTAAGGTGATGCTCGTTTAGGTATAA
 TTGACGAATCACAAAGTCATCTTTGGGACAGTTAAAACAGTTGGGTGGAGGT

도면7a

(a)

MCNSDNTSLKETVSSNSADVVEETETYQLTPIDAPSSFLSHSWEQTCGTPILNESDKQAISDFVAPPELKQDEKYCFTFKG
 ITGDHRYITNTTTLVVAPTLEVYIDHASLPSLQQLIHIIQAKDEYPSNQRFSWKRVTVADNANKLNIHTYPLKGNNTS
 PEMVAAIDEYAQSKNRLNIEFYTNTAHVFNNLPPIIQPLYNNEKVKISHISLYDDGSSEYVSLYQWKDTPNKIETLEGEV
 SLLANYLAGTSPDAPKGMGNRYNWHKLYDTDYFPLREDYLDVEANLHDLRDYLGSSAKQMPWDEF AKLSDSQQTLFLDIV
 GFDKEQLQQQYSQSPLPNFIPTGTTWAGGETKEYYAQQQVNVINNAINETS PYYLGKDYDLFFKGPAGGVINDIILGS
 FDPMINIPAKTSEFVLMMTDMLPDTVAGIASSLYFTIPADKVNFI VFTSSDITDREBALKSPLVQVMLTLGIVKEKDVLF
 FWA

도면7b

(b)

MCNSDNTSLKETVSSNSADVVEETETYQLTPIDAPSSFLSHSWEQTCGTPILNESDKQAISDFVAPPELKQDEKYCFTFKG
 ITGDHRYITNTTTLVVAPTLEVYIDHASLPSLQQLIHIIQAKDEYPSNQRFSWKRVTVADNANKLNIHTYPLKGNNTS
 PEMVAAIDEYAQSKNRLNIEFYTNTAHVFNNLPPIIQPLYNNEKVKISHISLYDDGSSEYVSLYQWKDTPNKIETLEGEV
 SLLANYLAGTSPDAPKGMGNRYNWHKLYDTDYFPLREDYLDVEANLHDLRDYLGSSAKQMPWDEF AKLSDSQQTLFLDIV
 GFDKEQLQQQYSQSPLPNFIPTGTTWAGGETKEYYAQQQVNVINNAINETS PYYLGKDYDLFFKGPAGGVINDIILGS
 FDPMINIPAKISEFVLMMTDMLPDTVAGIASSYFTIPADKVNFI VFTSSDITDREBALKSPLVQVMLTLGIVKEKDVLF
 FWA

도면7c

(c)

MCNSDNTSLKETVSSNSADVVEETETYQLTPIDAPSSFLSHSWEQTCGTPILNESDKQAISDFVAPPELKQDEKYCFTFKG
 ITGDHRYITNTTTLVVAPTLEVYIDHASLPSLQQLIHIIQAKDEYPSNQRFSWKRVTVADNANKLNIHTYPLKGNNTS
 PEMVAAIDEYAQSKNRLNIEFYTNTAHVFNNLPPIIQPLYNNEKVKISHISLYDDGSSEYVSLYQWKDTPNKIETLEGEV
 SLLANYLAGTSPDAPKGMGNRYNWHKLYDTDYFPLREDYLDVEANLHDLRDYLGSSAKQMPWDEF AKLSDSQQTLFLDIV
 GFDKEQLQQQYSQSPLPNFIPTGTTWAGGETKEYYAQQQVNVINNAINETS PYYLGKDYDLFFKGPAGGVINDIILGS
 FDPMINIPAKISEFVLMMTDMLPDTVAGIASSYFTIPADKVNFI VFTSSDITDREBALKSPLVQVMLTLGIVKEKDVLF
 FWA

도면7d

(d)

MCNSDNTSLKETVSSNSADVVEETETYQLTPIDAPSSFLSHSWEQTCGTPILNESDKQAISDFVAPPELKQDEKYCFTFKG
 ITGDHRYITNTTTLVVAPTLEVYIDHASLPSLQQLIHIIQAKDEYPSNQRFSWKRVTVADNANKLNIHTYPLKGNNTS
 PEMVAAIDEYAQSKNRLNIEFYTNTAHVFNNLPPIIQPLYNNEKVKISHISLYDDGSSEYVSLYQWKDTPNKIETLEGEV
 SLLANYLAGTSPDAPKGMGNRYNWHKLYDTDYFPLREDYLDVEANLHDLRDYLGSSAKQMPWDEF AKLSDSQQTLFLDIV
 GFDKEQLQQQYSQSPLPNFIPTGTTWAGGETKEYYAQQQVNVINNAINETS PYYLGKDYDLFFKGPAGGVINDIILGS
 FDPMINIPAKTSEFVLMMTDMLPDTVAGIASSYFTIPADKVNFI VFTSSDITDREBALKSPLVQVMLTLGIVKEKDVLF
 FWA

도면7e

(e)
 MCNSDNT SLKETVSSNSADVVEETETYQLT PIDAPS SFLSHSWBQTCGTPILNESDKQAI S PDFVAPELKQDEKYCFTFKG
 ITGDHRYITNTT LTVVAPT LEVYI DHASLPSLQQLIHII QAKDEY PSNQRVFSWKRVTV DADNANKLNIHTY PLKGNNTS
 PEMVAAI DEYAQSKNRLNI EBYTNT AHVFNPLPIIQ PLYNNEKVKI SHI SLYDDGS SEYVSLYQWKDT PNKIETLEGEV
 SLLANYLAGTSPDA PKGMGNRYNWHKLYD TDYF LRE DYLDVEANLHDLRDY LGS SAKQMEWDEF AKLSDSQQLFLFDIV
 GF DKEQLQQYSQS PLPNFI FTGTTWAGGETKEYYAQQQVNVINNAINETS PYYLKDYDLFF KGH PGGVINDII LGS
 F PDMINI PAKTSEVFLMMDMLPDTVAGIASSTYFTI PADKVN I VFTSSDTITDREALKS PLVQVMLT LGIVKEKEDVL
 FWA

도면8a

(a)
 ATGTGTAATAGTGACAATACCAGCTTGAAAGAAACGGTAAGCTCTAATTCTGCAGATGTAGTAGAAACGAGAACTTACCA
 ACTGACACCGGATGATGCTCCTAGCTCTTTTTTATCTCATCTCTGGGAGCAAAACATGTGGCACACCTATCTTGAATGAAA
 GTGACAAGCAAGOGATATCTTTTGATTTTGTGCTCCAGAGTTAAAGCAAGATGAAAAGTATGTGTTACTTTTTAAAGGT
 ATTACAGGCGATCATAGGTATATCACAAATACAACATTAAGTGTGTGTCACCTACGCTAGAAGTTACATCGATCATGC
 ATCCTTACCATCGCTACAGCAGCTTATCCACATTATTCAAGCAAAAAGATGAATACCCAAGTAATCAACGTTTTGTCTCTT
 GGAAGCGTGTAACTGTTGATGCTGATAATGCCAATAAGTTAAACATTCATCTTATCCATTAAGGCAATAATACCTCA
 CCAGAAATGGTGGCAGCGATTGATGAGTATGCTCAGAGCAAAAATCGATTGAATATAGAGTTCTATACAAATACAGCTCA
 TGTTTTAATAATTTACCACCTATTATCAACCTTTATATAAACAAGAGAAGGTGAAAATTTCTCATATTAGTTTGTATG
 ATGATGGTTCTTCTGAATATGTAAGTTTATATCAATGGAAGATACACCAAAATAAGATAGAAACATTAGAAGGTGAAGTA
 TCGCTCTTGCTAATTATTTAGCAGGAACATCTCCGGATGCAACAAAAGGAATGGGAAATCGTTATAACTGGCATAAAT
 ATATGACACTGATTATTACTTTTTGCGCGAAGATTACCTTGACGTTGAAGCAAACTACATGATTACGTTGATTATTTAG
 GCTCTTCCGCAAAGCAATGCCATGGGATGAATTTGCTAAATATCTGATTCTCAGCAAACTATTTTTAGATATTGTG
 GGTTTTGATAAAGAGCAATTGCAACAACAATATTCACAATCCCCACTACCAAATTTATTTTTACCGGCACAACAACCTTG
 GGCTGGGGGGAAACGAAAGAGTATATGCTCAGCAACAAGTAAATGTGATTAAATGCGATCAATGAAACTAGCCCTT
 ATTTATTTAGGTAAGACTACGATCTATTTTCAAGGGGCATCCTGCTGGTGGCGTTATTAAAGCATCATCTTGGGAGC
 TTCCTGATATGATCAATATCCAGCCAAGACTTCTTATTGAGGTTCTTGTGATGATGACGGATATGTTGCCCTGATACAGTAGC
 TGGTATTGCGAGCTCTCTGACTTCACAATTCCTGCCGATAAAGTAAATTTTTATTGATTTACTTCACTGACACTATTA
 CTGATCGTGAAGAGGCTCTTAAATCACCATTAGTACAAGTGATGCTAACGTTGGGTATTGTTAAAGAAAAGATGTTCTG
 TTCTGGGCT

도면8b

(b)
 ATGTGTAATAGTGACAATACCAGCTTGAAAGAAACGGTAAGCTCTAATTCTGCAGATGTAGTAGAAACGAGAACTTACCA
 ACTGACACCGGATGATGCTCCTAGCTCTTTTTTATCTCATCTCTGGGAGCAAAACATGTGGCACACCTATCTTGAATGAAA
 GTGACAAGCAAGOGATATCTTTTGATTTTGTGCTCCAGAGTTAAAGCAAGATGAAAAGTATGTGTTACTTTTTAAAGGT
 ATTACAGGCGATCATAGGTATATCACAAATACAACATTAAGTGTGTGTCACCTACGCTAGAAGTTACATCGATCATGC
 ATCCTTACCATCGCTACAGCAGCTTATCCACATTATTCAAGCAAAAAGATGAATACCCAAGTAATCAACGTTTTGTCTCTT
 GGAAGCGTGTAACTGTTGATGCTGATAATGCCAATAAGTTAAACATTCATCTTATCCATTAAGGCAATAATACCTCA
 CCAGAAATGGTGGCAGCGATTGATGAGTATGCTCAGAGCAAAAATCGATTGAATATAGAGTTCTATACAAATACAGCTCA
 TGTTTTAATAATTTACCACCTATTATCAACCTTTATATAAACAAGAGAAGGTGAAAATTTCTCATATTAGTTTGTATG
 ATGATGGTTCTTCTGAATATGTAAGTTTATATCAATGGAAGATACACCAAAATAAGATAGAAACATTAGAAGGTGAAGTA
 TCGCTCTTGCTAATTATTTAGCAGGAACATCTCCGGATGCAACAAAAGGAATGGGAAATCGTTATAACTGGCATAAAT
 ATATGACACTGATTATTACTTTTTGCGCGAAGATTACCTTGACGTTGAAGCAAACTACATGATTACGTTGATTATTTAG
 GCTCTTCCGCAAAGCAATGCCATGGGATGAATTTGCTAAATATCTGATTCTCAGCAAACTATTTTTAGATATTGTG
 GGTTTTGATAAAGAGCAATTGCAACAACAATATTCACAATCCCCACTACCAAATTTATTTTTACCGGCACAACAACCTTG
 GGCTGGGGGGAAACGAAAGAGTATATGCTCAGCAACAAGTAAATGTGATTAAATGCGATCAATGAAACTAGCCCTT
 ATTTATTTAGGTAAGACTACGATCTATTTTCAAGGGGCATCCTGCTGGTGGCGTTATTAAAGCATCATCTTGGGAGC
 TTCCTGATATGATCAATATCCAGCCAAGACTTCTTATTGAGGTTCTTGTGATGATGACGGATATGTTGCCCTGATACAGTAGC
 TGGTATTGCGAGCTCTGTTACTTCACAATTCCTGCCGATAAAGTAAATTTTTATTGATTTACTTCACTGACACTATTA
 CTGATCGTGAAGAGGCTCTTAAATCACCATTAGTACAAGTGATGCTAACGTTGGGTATTGTTAAAGAAAAGATGTTCTG
 TTCTGGGCT

도면8c

(c)

ATGTGTAATAGTGACAATACCAGCTTGAAAGAAACGGTAAGCTCTAATTCTGCAGATGTAGTAGAAACAGAACTTACCA
 ACTGACACCGATTGATGCTCCTAGCTCTTTTATCTCATCTCTGGGAGCAAACATGTGGCACACCTATCTTGAATGAAA
 GTGACAAGCAAGCGATATCTTTTGATTTTGTGCTCCAGAGTTAAAGCAAGATGAAAAGTATGTGTTTACTTTTAAAGGT
 ATTACAGGCGATCATAGGTATATCACAATACAACATTAACCTGTTGTGTCACCTACGCTAGAAGTTACATCGATCATGC
 ATCCTTACCATCGCTACAGCAGCTTATCCACATTATTCAAGCAAAGATGAATACCCAAGTAATCAACGTTTTGTCTCTT
 GGAAGCGTGTAACTGTTGATGCTGATAATGCCAATAAGTTAAACATTCACTTATCCATTAAGGCAATAATACCTCA
 CCAGAAATGGTGGCAGCGATTGATGAGTATGCTCAGAGCAAAAATCGATTGAATATAGAGTTCTATACAAATACAGCTCA
 TGTTTTTAATAATTTACCACCTATTTCAACCTTTATATAATAACGAGAAGGTGAAAATTTCTCATATAGTTTGTATG
 ATGATGGTCTTCTGAATATGTAAGTTTATATCAATGGAAGATACACCAAATAAGATAGAAACATTAGAAGGTGAAGTA
 TCGTCTCTGCTAATTAATTTAGCAGGAACATCTCCGGATGCACCAAAGGAATGGGAAATCGTTATAACTGGCATAAAT
 ATATGACACTGATTATTTACTTTTTGCGCGAAGATTACCTTGACGTTGAAGCAAACCTACATGATTTACGTTGATTATTTAG
 GCTCTCCGCAAAGCAAATGCCATGGGATGAATTTGCTAAATATCTGATTTCTCAGCAAACACTATTTTAGATATTGTG
 GGTTTTGATAAAGAGCAATGCAACAACAATATTCACAATCCCCACTACCAAACCTTATTTTACCGGCACAACAACCTG
 GGCTGGGGGGAAACGAAAGAGTATATGCTCAGCAACAAGTAAATGTGATTAAATGCGATCAATGAACTAGCCCTT
 ATTTATTTAGGTAAGACTACGATCTATTTTCAAGGGGCATCTGCTGGTGGCGTTATTACGACATCATCTTGGAAAGC
 TTCCCTGATATGATCAATATCCAGCCAAGACTTCTTGGAGGCTTGATGATGACGGATATGTTGCCTGATACAGTAGC
 TGGTATTGCGAGCTCTAGCTACTTCACAATTCCTGCCGATAAAGTTAATTTTATGATTTACTTCACTGACACTATTA
 CTGATCGTGAAGAGGCTCTTAAATCACCATTAGTACAAGTGATGCTAACGTTGGGTATTGTTAAAGAAAAGATGTTCTG
 TTCTGGGCT

도면8d

(d)

ATGTGTAATAGTGACAATACCAGCTTGAAAGAAACGGTAAGCTCTAATTCTGCAGATGTAGTAGAAACAGAACTTACCA
 ACTGACACCGATTGATGCTCCTAGCTCTTTTATCTCATCTCTGGGAGCAAACATGTGGCACACCTATCTTGAATGAAA
 GTGACAAGCAAGCGATATCTTTTGATTTTGTGCTCCAGAGTTAAAGCAAGATGAAAAGTATGTGTTTACTTTTAAAGGT
 ATTACAGGCGATCATAGGTATATCACAATACAACATTAACCTGTTGTGTCACCTACGCTAGAAGTTACATCGATCATGC
 ATCCTTACCATCGCTACAGCAGCTTATCCACATTATTCAAGCAAAGATGAATACCCAAGTAATCAACGTTTTGTCTCTT
 GGAAGCGTGTAACTGTTGATGCTGATAATGCCAATAAGTTAAACATTCACTTATCCATTAAGGCAATAATACCTCA
 CCAGAAATGGTGGCAGCGATTGATGAGTATGCTCAGAGCAAAAATCGATTGAATATAGAGTTCTATACAAATACAGCTCA
 TGTTTTTAATAATTTACCACCTATTTCAACCTTTATATAATAACGAGAAGGTGAAAATTTCTCATATAGTTTGTATG
 ATGATGGTCTTCTGAATATGTAAGTTTATATCAATGGAAGATACACCAAATAAGATAGAAACATTAGAAGGTGAAGTA
 TCGTCTCTGCTAATTAATTTAGCAGGAACATCTCCGGATGCACCAAAGGAATGGGAAATCGTTATAACTGGCATAAAT
 ATATGACACTGATTATTTACTTTTTGCGCGAAGATTACCTTGACGTTGAAGCAAACCTACATGATTTACGTTGATTATTTAG
 GCTCTCCGCAAAGCAAATGCCATGGGATGAATTTGCTAAATATCTGATTTCTCAGCAAACACTATTTTAGATATTGTG
 GGTTTTGATAAAGAGCAATGCAACAACAATATTCACAATCCCCACTACCAAACCTTATTTTACCGGCACAACAACCTG
 GGCTGGGGGGAAACGAAAGAGTATATGCTCAGCAACAAGTAAATGTGATTAAATGCGATCAATGAACTAGCCCTT
 ATTTATTTAGGTAAGACTACGATCTATTTTCAAGGGGCATCTGCTGGTGGCGTTATTACGACATCATCTTGGAAAGC
 TTCCCTGATATGATCAATATCCAGCCAAGACTTCTTGGAGGCTTGATGATGACGGATATGTTGCCTGATACAGTAGC
 TGGTATTGCGAGCTCTAGCTACTTCACAATTCCTGCCGATAAAGTTAATTTTATGATTTACTTCACTGACACTATTA
 CTGATCGTGAAGAGGCTCTTAAATCACCATTAGTACAAGTGATGCTAACGTTGGGTATTGTTAAAGAAAAGATGTTCTG
 TTCTGGGCT

도면8e

(e)

ATGTGTAATAGTGACAATACCAGCTTGAAAGAAACGGTAAGCTCTAATTCTGCAGATGTAGTAGAAACAGAACTTACCA
 ACTGACACCGATTGATGCTCCTAGCTCTTTTATCTCATCTCTGGGAGCAAACATGTGGCACACCTATCTTGAATGAAA
 GTGACAAGCAAGCGATATCTTTTGATTTTGTGCTCCAGAGTTAAAGCAAGATGAAAAGTATGTGTTTACTTTTAAAGGT
 ATTACAGGCGATCATAGGTATATCACAATACAACATTAACCTGTTGTGTCACCTACGCTAGAAGTTACATCGATCATGC
 ATCCTTACCATCGCTACAGCAGCTTATCCACATTATTCAAGCAAAGATGAATACCCAAGTAATCAACGTTTTGTCTCTT
 GGAAGCGTGTAACTGTTGATGCTGATAATGCCAATAAGTTAAACATTCACTTATCCATTAAGGCAATAATACCTCA
 CCAGAAATGGTGGCAGCGATTGATGAGTATGCTCAGAGCAAAAATCGATTGAATATAGAGTTCTATACAAATACAGCTCA
 TGTTTTTAATAATTTACCACCTATTTCAACCTTTATATAATAACGAGAAGGTGAAAATTTCTCATATAGTTTGTATG
 ATGATGGTCTTCTGAATATGTAAGTTTATATCAATGGAAGATACACCAAATAAGATAGAAACATTAGAAGGTGAAGTA
 TCGTCTCTGCTAATTAATTTAGCAGGAACATCTCCGGATGCACCAAAGGAATGGGAAATCGTTATAACTGGCATAAAT
 ATATGACACTGATTATTTACTTTTTGCGCGAAGATTACCTTGACGTTGAAGCAAACCTACATGATTTACGTTGATTATTTAG
 GCTCTCCGCAAAGCAAATGCCATGGGATGAATTTGCTAAATATCTGATTTCTCAGCAAACACTATTTTAGATATTGTG
 GGTTTTGATAAAGAGCAATGCAACAACAATATTCACAATCCCCACTACCAAACCTTATTTTACCGGCACAACAACCTG
 GGCTGGGGGGAAACGAAAGAGTATATGCTCAGCAACAAGTAAATGTGATTAAATGCGATCAATGAACTAGCCCTT
 ATTTATTTAGGTAAGACTACGATCTATTTTCAAGGGGCATCTGCTGGTGGCGTTATTACGACATCATCTTGGAAAGC
 TTCCCTGATATGATCAATATCCAGCCAAGACTTCTTGGAGGCTTGATGATGACGGATATGTTGCCTGATACAGTAGC
 TGGTATTGCGAGCTCTAGCTACTTCACAATTCCTGCCGATAAAGTTAATTTTATGATTTACTTCACTGACACTATTA
 CTGATCGTGAAGAGGCTCTTAAATCACCATTAGTACAAGTGATGCTAACGTTGGGTATTGTTAAAGAAAAGATGTTCTG
 TTCTGGGCT

서열목록

- <110> Seoul National University R&DB Foundation
GeneChem Inc.
- <120> METHOD FOR SCREENING α -2,3 AND α -2,6 SIALYLTRANSFERASE VARIANTS
AND THEIR APPLICATION FOR SYNTHESIS OF SIALYLOLIGOSACCHARIDES
- <130> Y13KP-021
- <160> 24
- <170> KopatentIn 2.0
- <210> 1
- <211> 391
- <212> PRT
- <213> Artificial Sequence
- <220><223> Artificial
- <400> 1

Met Lys Thr Ile Thr Leu Tyr Leu Asp Pro Ala Ser Leu Pro Ala Leu
 1 5 10 15

Asn Gln Leu Met Asp Phe Thr Gln Asn Asn Glu Asp Lys Thr His Pro
 20 25 30

Arg Ile Phe Gly Leu Ser Arg Phe Lys Ile Pro Asp Asn Ile Ile Thr
 35 40 45

Gln Tyr Gln Asn Ile His Phe Val Glu Leu Lys Asp Asn Arg Pro Thr
 50 55 60

Glu Ala Leu Phe Thr Ile Leu Asp Gln Tyr Pro Gly Asn Ile Glu Leu
 65 70 75 80

Asn Ile His Leu Asn Ile Ala His Ser Val Gln Leu Ile Arg Pro Ile

 85 90 95

Leu Ala Tyr Arg Phe Lys His Leu Asp Arg Val Ser Ile Gln Gln Leu
 100 105 110

Asn Leu Tyr Asp Asp Gly Ser Met Glu Tyr Val Asp Leu Glu Lys Glu
 115 120 125

Glu Asn Lys Asp Ile Ser Ala Glu Ile Lys Gln Ala Glu Lys Gln Leu
 130 135 140

Ser His Tyr Leu Leu Thr Gly Lys Ile Lys Phe Asp Asn Pro Thr Ile

145 150 155 160
 Ala Arg Tyr Val Trp Gln Ser Ala Phe Pro Val Lys Tyr His Phe Leu
 165 170 175
 Ser Thr Asp Tyr Phe Glu Lys Ala Glu Phe Leu Gln Pro Leu Lys Glu
 180 185 190
 Tyr Leu Ala Glu Asn Tyr Gln Lys Met Asp Trp Thr Ala Tyr Gln Gln
 195 200 205
 Leu Thr Pro Glu Gln Gln Ala Phe Tyr Leu Thr Leu Val Gly Phe Asn
 210 215 220
 Asp Glu Val Lys Gln Ser Leu Glu Val Gln Gln Ala Lys Phe Ile Phe

 225 230 235 240
 Thr Gly Thr Thr Thr Trp Glu Gly Asn Thr Asp Val Arg Glu Tyr Tyr
 245 250 255
 Ala Gln Gln Gln Leu Asn Leu Leu Asn His Phe Thr Gln Ala Glu Gly
 260 265 270
 Asp Leu Phe Ile Gly Asp His Tyr Lys Ile Tyr Phe Lys Gly His Pro
 275 280 285
 Asn Gly Gly Glu Ile Asn Asp Tyr Ile Leu Asn Asn Ala Lys Asn Ile
 290 295 300

 Thr Asn Ile Pro Ala Asn Ile Ser Phe Glu Val Leu Met Met Thr Gly
 305 310 315 320
 Leu Leu Pro Asp Lys Val Gly Gly Val Ala Ser Ser Leu Tyr Phe Ser
 325 330 335
 Leu Pro Lys Glu Lys Ile Ser His Ile Ile Phe Thr Ser Asn Lys Gln
 340 345 350
 Val Lys Ser Lys Glu Asp Ala Leu Asn Asn Pro Tyr Val Lys Val Met
 355 360 365
 Arg Arg Leu Gly Ile Ile Asp Glu Ser Gln Val Ile Phe Trp Asp Ser

 370 375 380
 Leu Lys Gln Leu Gly Gly Gly
 385 390

<210> 2
 <211> 391
 <212> PRT
 <213> Artificial Sequence
 <220><223> Artificial
 <400> 2
 Met Lys Thr Ile Thr Leu Tyr Leu Asp Pro Ala Ser Leu Pro Ala Leu
 1 5 10 15
 Asn Gln Leu Met Asp Phe Thr Gln Asn Asn Glu Asp Lys Thr His Pro
 20 25 30
 Arg Ile Phe Gly Leu Ser Arg Phe Lys Ile Pro Asp Asn Ile Ile Thr
 35 40 45
 Gln Tyr Gln Asn Ile His Phe Val Glu Leu Lys Asp Asn Arg Pro Thr
 50 55 60
 Glu Ala Leu Phe Thr Ile Leu Asp Gln Tyr Pro Gly Asn Ile Glu Leu
 65 70 75 80
 Asn Ile His Leu Asn Ile Ala His Ser Val Gln Leu Ile Arg Pro Ile
 85 90 95
 Leu Ala Tyr Arg Phe Lys His Leu Asp Arg Val Ser Ile Gln Gln Leu
 100 105 110

 Asn Leu Tyr Asp Asp Gly Ser Met Glu Tyr Val Asp Leu Glu Lys Glu
 115 120 125
 Glu Asn Lys Asp Ile Ser Ala Glu Ile Lys Gln Ala Glu Lys Gln Leu
 130 135 140
 Ser His Tyr Leu Leu Thr Gly Lys Ile Lys Phe Asp Asn Pro Thr Ile
 145 150 155 160
 Ala Arg Tyr Val Trp Gln Ser Ala Phe Pro Val Lys Tyr His Phe Leu
 165 170 175
 Ser Thr Asp Tyr Phe Glu Lys Ala Glu Phe Leu Gln Pro Leu Lys Glu
 180 185 190
 Tyr Leu Ala Glu Asn Tyr Gln Lys Met Asp Trp Thr Ala Tyr Gln Gln
 195 200 205

Leu Thr Pro Glu Gln Gln Ala Phe Tyr Leu Thr Leu Val Gly Phe Asn
 210 215 220

Asp Glu Val Lys Gln Ser Leu Glu Val Gln Gln Ala Lys Phe Ile Phe
 225 230 235 240

Thr Gly Thr Thr Thr Trp Glu Gly Asn Thr Asp Val Arg Glu Tyr Tyr
 245 250 255

Ala Gln Gln Gln Leu Asn Leu Leu Asn His Phe Thr Gln Ala Glu Gly
 260 265 270

Asp Leu Phe Ile Gly Asp His Tyr Lys Ile Tyr Phe Lys Gly His Pro
 275 280 285

His Gly Gly Glu Ile Asn Asp Tyr Ile Leu Asn Asn Ala Lys Asn Ile
 290 295 300

Thr Asn Ile Pro Ala Asn Ile Ser Phe Glu Val Leu Met Met Thr Gly
 305 310 315 320

Leu Leu Pro Asp Lys Val Gly Gly Val Ala Ser Ser Leu Tyr Phe Ser

325 330 335

Leu Pro Lys Glu Lys Ile Ser His Ile Ile Phe Thr Ser Asn Lys Gln
 340 345 350

Val Lys Ser Lys Glu Asp Ala Leu Asn Asn Pro Tyr Val Lys Val Met
 355 360 365

Arg Arg Leu Gly Ile Ile Asp Glu Ser Gln Val Ile Phe Trp Asp Ser
 370 375 380

Leu Lys Gln Leu Gly Gly Gly
 385 390

<210> 3

<211> 391

<212>

> PRT

<213> Artificial Sequence

<220><223> Artificial

<400> 3

Met Lys Thr Ile Thr Leu Tyr Leu Asp Pro Ala Ser Leu Pro Ala Leu
 1 5 10 15

Asn Gln Leu Met Asp Phe Thr Gln Asn Asn Glu Asp Lys Thr His Pro
 20 25 30
 Arg Ile Phe Gly Leu Ser Arg Phe Lys Ile Pro Asp Asn Ile Ile Thr
 35 40 45
 Gln Tyr Gln Asn Ile His Phe Val Glu Leu Lys Asp Asn Arg Pro Thr
 50 55 60
 Glu Ala Leu Phe Thr Ile Leu Asp Gln Tyr Pro Gly Asn Ile Glu Leu
 65 70 75 80
 Asn Ile His Leu Asn Ile Ala His Ser Val Gln Leu Ile Arg Pro Ile
 85 90 95
 Leu Ala Tyr Arg Phe Lys His Leu Asp Arg Val Ser Ile Gln Gln Leu
 100 105 110
 Asn Leu Tyr Asp Asp Gly Ser Met Glu Tyr Val Asp Leu Glu Lys Glu
 115 120 125

 Glu Asn Lys Asp Ile Ser Ala Glu Ile Lys Gln Ala Glu Lys Gln Leu
 130 135 140
 Ser His Tyr Leu Leu Thr Gly Lys Ile Lys Phe Asp Asn Pro Thr Ile
 145 150 155 160
 Ala Arg Tyr Val Trp Gln Ser Ala Phe Pro Val Lys Tyr His Phe Leu
 165 170 175
 Ser Thr Asp Tyr Phe Glu Lys Ala Glu Phe Leu Gln Pro Leu Lys Glu
 180 185 190
 Tyr Leu Ala Glu Asn Tyr Gln Lys Met Asp Trp Thr Ala Tyr Gln Gln
 195 200 205
 Leu Thr Pro Glu Gln Gln Ala Phe Tyr Leu Thr Leu Val Gly Phe Asn
 210 215 220
 Asp Glu Val Lys Gln Ser Leu Glu Val Gln Gln Ala Lys Phe Ile Phe
 225 230 235 240
 Ser Gly Thr Thr Thr Trp Glu Gly Asn Thr Asp Val Arg Glu Tyr Tyr
 245 250 255
 Ala Gln Gln Gln Leu Asn Leu Leu Asn His Phe Thr Gln Ala Glu Gly

Leu Leu Pro Asp Lys Val Gly Gly Val Ala Ser Ser Leu Tyr Phe Ser
 325 330 335

Leu Pro Lys Glu Lys Ile Ser His Ile Ile Phe Thr Ser Asn Lys Gln
 340 345 350

Val Lys Ser Lys Glu Asp Ala Leu Asn Asn Pro Tyr Val Lys Val Met
 355 360 365

Arg Arg Leu Gly Ile Ile Asp Glu Ser Gln Val Ile Phe Trp Asp Ser
 370 375 380

Leu Lys Gln Leu Gly Gly Gly
 385 390

<210> 5

<211> 391

<212> PRT

<213> Artificial Sequence

<220><223> Artificial

<400> 5

Met Lys Thr Ile Thr Leu Tyr Leu Asp Pro Ala Ser Leu Pro Ala Leu
 1 5 10 15

Asn Gln Leu Met Asp Phe Thr Gln Asn Asn Glu Asp Lys Thr His Pro
 20 25 30

Arg Ile Phe Gly Leu Ser Arg Phe Lys Ile Pro Asp Asn Ile Ile Thr
 35 40 45

Gln Tyr Gln Asn Ile His Phe Val Glu Leu Lys Asp Asn Arg Pro Thr
 50 55 60

Glu Ala Leu Phe Thr Ile Leu Asp Gln Tyr Pro Gly Asn Ile Glu Leu
 65 70 75 80

Asn Ile His Leu Asn Ile Ala His Ser Val Gln Leu Ile Arg Pro Ile
 85 90 95

Leu Ala Tyr Arg Phe Lys His Leu Asp Arg Val Ser Ile Gln Gln Leu
 100 105 110

Asn Leu Tyr Asp Asp Gly Ser Met Glu Tyr Val Asp Leu Glu Lys Glu
 115 120 125

Glu Asn Lys Asp Ile Ser Ala Glu Ile Lys Gln Ala Glu Lys Gln Leu
 130 135 140
 Ser His Tyr Leu Leu Thr Gly Lys Ile Lys Phe Asp Asn Pro Thr Ile
 145 150 155 160
 Ala Arg Tyr Val Trp Gln Ser Ala Phe Pro Val Lys Tyr His Phe Leu
 165 170 175

 Ser Thr Asp Tyr Phe Glu Lys Ala Glu Phe Leu Gln Pro Leu Lys Glu
 180 185 190
 Tyr Leu Ala Glu Asn Tyr Gln Lys Met Asp Trp Thr Ala Tyr Gln Gln
 195 200 205
 Leu Thr Pro Glu Gln Gln Ala Phe Tyr Leu Thr Leu Val Gly Phe Asn
 210 215 220
 Asp Glu Val Lys Gln Ser Leu Glu Val Gln Gln Ala Lys Phe Ile Phe
 225 230 235 240
 Ser Gly Thr Thr Thr Trp Glu Gly Asn Thr Asp Val Arg Glu Tyr Tyr

 245 250 255
 Ala Gln Gln Gln Leu Asn Leu Leu Asn His Phe Thr Gln Ala Glu Gly
 260 265 270
 Asp Leu Phe Ile Gly Asp His Tyr Lys Ile Tyr Phe Lys Gly His Pro
 275 280 285
 His Gly Gly Glu Ile Asn Asp Tyr Ile Leu Asn Asn Ala Lys Asn Ile
 290 295 300
 Thr Asn Ile Pro Ala Asn Ile Ser Phe Glu Val Leu Met Met Thr Gly
 305 310 315 320

 Leu Leu Pro Asp Lys Val Gly Gly Val Ala Ser Ser Leu Tyr Phe Ser
 325 330 335
 Leu Pro Lys Glu Lys Ile Ser His Ile Ile Phe Thr Ser Asn Lys Gln
 340 345 350
 Val Lys Ser Lys Glu Asp Ala Leu Asn Asn Pro Tyr Val Lys Val Met
 355 360 365
 Arg Arg Leu Gly Ile Ile Asp Glu Ser Gln Val Ile Phe Trp Asp Ser
 370 375 380

<213> Artificial Sequence

<220><223> Artificial

<400> 7

atgaaaacaa tcacgctgta tttagatcct gcctccttac cggcattaaa tcagctgatg 60
 gactttacgc aaaataatga agataaaaca catccacgta tttttggtct ttctcgcttt 120
 aaaatccctg acaacattat tacacagtat caaaatatcc atttcgtcga actcaaagat 180
 aatcgtccca ctgaagcact ttttacgatt ttagatcaat accctggtaa cattgagtta 240
 aatatacact taaatattgc tcattccgtt caattaattc gtccgatttt ggcatatcgt 300

tttaaacatt tagatcgtgt atcaattcag cagttaaadc tttatgacga tggetcaatg 360
 gaatatgttg atttagaaaa agaagaaaat aaagatattt ccgcagaaat taagcaagca 420
 gaaaaacaac tttctcacta tttgcttact ggcaaaataa aatttgataa cccaactatt 480
 gctcgttatg tctggcaatc cgcgttccca gtaaaatadc atttttaag tacagactat 540
 tttgaaaaag ccgaattttt acaaccacta aaagaatatt tagcagaaaa ttatcaaaaa 600
 atggactgga ctgcttacca acagctgact ccagaacagc aagcattcta cttaacattg 660
 gtaggcttca atgacgaagt caagcagtcg ctagaagtgc aacaagctaa atttatcttt 720

accggcacga caacttggga aggaaatacc gatgtgagag aatactacgc acagcaacaa 780
 cttaatftac ttaatcactt tacccaagct gagggcgatt tatttattgg tgatcattat 840
 aaaatctact ttaaaggcca tcctcatggt ggtgaaatta atgactacat tctgaacaat 900
 gctaaaaata tcaccaatat ccctgccaat atttcctttg aagtattgat gatgacaggc 960
 ttattacctg ataaagtggg tgggtgtgca agttcactgt atttctcctt accaaaagaa 1020
 aaaattagcc atattatttt cacatcgaat aaacaagtga aaagcaaaga agatgcgcta 1080
 aataatccgt atgtaagggt catgcgtcgt ttaggtataa ttgacgaatc acaagtcac 1140

ttttgggaca gtttaaaaca gttgggtgga ggt 1173

<210> 8

<211> 1173

<212> DNA

<213> Artificial Sequence

<220><223> Artificial

<400> 8

atgaaaacaa tcacgctgta tttagatcct gcctccttac cggcattaaa tcagctgatg 60
 gactttacgc aaaataatga agataaaaca catccacgta tttttggtct ttctcgcttt 120
 aaaatccctg acaacattat tacacagtat caaaatatcc atttcgtcga actcaaagat 180

aatcgtccca ctgaagcact ttttacgatt ttagatcaat accctggtaa cattgagtta 240
aatatacact taaatattgc tcattccggt caattaattc gtccgatttt ggcatatcgt 300

ttaaacatt tagatcgtgt atcaattcag cagttaaadc tttatgacga tggctcaatg 360
gaatatgttg atttagaaaa agaagaaaat aaagatattt ccgcagaaat taagcaagca 420
gaaaaacaac tttctcacta tttgcttact ggcaaaaataa aatttgataa cccaactatt 480
gctcgttatg tctggcaatc cgcgttccca gtaaaatadc attttttaag tacagactat 540
tttgaaaaag ccgaattttt acaaccacta aaagaatatt tagcagaaaa ttatcaaaaa 600
atggactgga ctgcttacca acagctgact ccagaacagc aagcattcta cttaacattg 660
gtaggcttca atgacgaagt caagcagtcg ctagaagtgc aacaagctaa atttatcttt 720

agtggcacga caacttggga aggaaatacc gatgtgcgag aatactacgc acagcaacaa 780
cttaatttac ttaatcactt tacccaagct gagggcgatt tatttattgg tgatcattat 840
aaaatctact ttaaagggca tcctagaggt ggtgaaatta atgactacat tctgaacaat 900
gctaaaaata tcaccaatat ccctgccaat atttcctttg aagtattgat gatgacaggc 960
ttattacctg ataaagtggg tgggtgttga agttcactgt atttctcctt accaaaagaa 1020
aaaattagcc atattatttt cacatcgaat aaacaagtga aaagcaaaga agatgcgcta 1080
aataatccgt atgtaaggt catgcgtcgt ttaggtataa ttgacgaatc acaagtcac 1140

ttttgggaca gtttaaaaca gttgggtgga ggt 1173

<210> 9
<211> 1173
<212> DNA
<213> Artificial Sequence
<220><223> Artificial
<400> 9

atgaaaacaa tcacgctgta ttagatcct gcctccttac cggcattaaa tcagctgatg 60
gactttacgc aaaataatga agataaaaaca catccacgta tttttggtct ttctcgttt 120
aaaatccctg acaacattat tacacagtat caaaatatcc atttcgtcga actcaaagat 180
aatcgtccca ctgaagcact ttttacgatt ttagatcaat accctggtaa cattgagtta 240
aatatacact taaatattgc tcattccggt caattaattc gtccgatttt ggcatatcgt 300

ttaaacatt tagatcgtgt atcaattcag cagttaaadc tttatgacga tggctcaatg 360
gaatatgttg atttagaaaa agaagaaaat aaagatattt ccgcagaaat taagcaagca 420
gaaaaacaac tttctcacta tttgcttact ggcaaaaataa aatttgataa cccaactatt 480

gctcgttatg tctggcaatc cgcgttccca gtaaaatatic attttttaag tacagactat 540
 tttgaaaaag ccgaattttt acaaccacta aaagaatatt tagcagaaaa ttatcaaaaa 600
 atggactgga ctgcttacca acagctgact ccagaacagc aagcattcta cttaacattg 660
 gtaggcttca atgacgaagt caagcagtcg ctagaagtgc aacaagctaa atttatcttt 720

agcggcacga caacttggga aggaaatacc gatgtgcgag aatactacgc acagcaacaa 780
 cttaatattac ttaatcactt tacccaagct gagggcgatt tatttattgg tgatcattat 840
 aaaatctact ttaaagggca tcctaattgg ggtgaaatta atgactacat tctgaacaat 900
 gctaaaaata tcaccaatat ccctgccaat atttcctttg aagtattgat gatgacaggc 960
 ttattacctg ataaagtggg tgggtttgca agttcactgt atttctcctt accaaaagaa 1020
 aaaattagcc atattatfff cacatcgaat aaacaagtga aaagcaaaga agatgcgcta 1080
 aataatccgt atgtaaggt catgcgctgt ttaggtataa ttgacgaatc acaagtcac 1140

ttttgggaca gtttaaaaca gttgggtgga ggt 1173

<210> 10
 <211> 1173
 <212> DNA
 <213> Artificial Sequence
 <220><223> Artificial
 <400> 10

atgaaaaaca tcacgtgta tttagatcct gcctccttac cggcattaaa tcagctgatg 60
 gactttacgc aaaataatga agataaaaca catccacgta tttttggtct ttctcgcttt 120
 aaaatccctg acaacattat tacacagtat caaaatatcc atttcgtcga actcaaagat 180
 aatcgtccca ctgaagcact tttfacgatt ttagatcaat accctggtaa cattgagtta 240
 aatatacact taaatattgc tcattccggt caattaattc gtccgatttt ggcatatcgt 300

tttaaacatt tagatcgtgt atcaattcag cagttaaatic tttatgacga tggctcaatg 360
 gaatatgttg atttagaaaa agaagaaaat aaagatattt ccgcagaaat taagcaagca 420
 gaaaaacaac tttctacta tttgcttact ggcaaaataa aatttgataa cccaactatt 480
 gctcgttatg tctggcaatc cgcgttccca gtaaaatatic attttttaag tacagactat 540
 tttgaaaaag ccgaattttt acaaccacta aaagaatatt tagcagaaaa ttatcaaaaa 600
 atggactgga ctgcttacca acagctgact ccagaacagc aagcattcta cttaacattg 660
 gtaggcttca atgacgaagt caagcagtcg ctagaagtgc aacaagctaa atttatcttt 720

agcggcacga caacttggga aggaaatacc gatgtgcgag aatactacgc acagcaacaa 780

cttaatttac ttaatcactt tacccaagct gagggcgatt tatttattgg tgatcattat 840
 aaaatctact ttaaaggga tcctcatggt ggtgaaatta atgactacat tctgaacaat 900
 gctaaaaata tcaccaatat ccctgccaat atttcctttg aagtattgat gatgacaggc 960
 ttattacctg ataaagtggg tgggtgtgca agttcactgt atttctcctt accaaaagaa 1020
 aaaattagcc atattatfff cacatcgaat aaacaagtga aaagcaaaga agatgcgcta 1080
 aataatccgt atgtaaggt catgcgtcgt ttaggtataa ttgacgaatc acaagtcac 1140

ttttgggaca gtttaaaaca gttgggtgga ggt 1173

<210> 11

<211> 483

<212> PRT

<213> Artificial Sequence

<220><223> Artificial

<400> 11

Met Cys Asn Ser Asp Asn Thr Ser Leu Lys Glu Thr Val Ser Ser Asn

1 5 10 15

Ser Ala Asp Val Val Glu Thr Glu Thr Tyr Gln Leu Thr Pro Ile Asp

20 25 30

Ala Pro Ser Ser Phe Leu Ser His Ser Trp Glu Gln Thr Cys Gly Thr

35 40 45

Pro Ile Leu Asn Glu Ser Asp Lys Gln Ala Ile Ser Phe Asp Phe Val

50 55 60

Ala Pro Glu Leu Lys Gln Asp Glu Lys Tyr Cys Phe Thr Phe Lys Gly

65 70 75 80

Ile Thr Gly Asp His Arg Tyr Ile Thr Asn Thr Thr Leu Thr Val Val

85 90 95

Ala Pro Thr Leu Glu Val Tyr Ile Asp His Ala Ser Leu Pro Ser Leu

100 105 110

Gln Gln Leu Ile His Ile Ile Gln Ala Lys Asp Glu Tyr Pro Ser Asn

115 120 125

Gln Arg Phe Val Ser Trp Lys Arg Val Thr Val Asp Ala Asp Asn Ala

130 135 140

Asn Lys Leu Asn Ile His Thr Tyr Pro Leu Lys Gly Asn Asn Thr Ser

145 150 155 160
 Pro Glu Met Val Ala Ala Ile Asp Glu Tyr Ala Gln Ser Lys Asn Arg
 165 170 175
 Leu Asn Ile Glu Phe Tyr Thr Asn Thr Ala His Val Phe Asn Asn Leu
 180 185 190
 Pro Pro Ile Ile Gln Pro Leu Tyr Asn Asn Glu Lys Val Lys Ile Ser
 195 200 205
 His Ile Ser Leu Tyr Asp Asp Gly Ser Ser Glu Tyr Val Ser Leu Tyr
 210 215 220
 Gln Trp Lys Asp Thr Pro Asn Lys Ile Glu Thr Leu Glu Gly Glu Val
 225 230 235 240
 Ser Leu Leu Ala Asn Tyr Leu Ala Gly Thr Ser Pro Asp Ala Pro Lys
 245 250 255

 Gly Met Gly Asn Arg Tyr Asn Trp His Lys Leu Tyr Asp Thr Asp Tyr
 260 265 270
 Tyr Phe Leu Arg Glu Asp Tyr Leu Asp Val Glu Ala Asn Leu His Asp
 275 280 285
 Leu Arg Asp Tyr Leu Gly Ser Ser Ala Lys Gln Met Pro Trp Asp Glu
 290 295 300
 Phe Ala Lys Leu Ser Asp Ser Gln Gln Thr Leu Phe Leu Asp Ile Val
 305 310 315 320
 Gly Phe Asp Lys Glu Gln Leu Gln Gln Gln Tyr Ser Gln Ser Pro Leu
 325 330 335
 Pro Asn Phe Ile Phe Thr Gly Thr Thr Thr Trp Ala Gly Gly Glu Thr
 340 345 350
 Lys Glu Tyr Tyr Ala Gln Gln Gln Val Asn Val Ile Asn Asn Ala Ile
 355 360 365
 Asn Glu Thr Ser Pro Tyr Tyr Leu Gly Lys Asp Tyr Asp Leu Phe Phe
 370 375 380
 Lys Gly His Pro Ala Gly Gly Val Ile Asn Asp Ile Ile Leu Gly Ser
 385 390 395 400

Phe Pro Asp Met Ile Asn Ile Pro Ala Lys Thr Ser Phe Glu Val Leu
 405 410 415

Met Met Thr Asp Met Leu Pro Asp Thr Val Ala Gly Ile Ala Ser Ser
 420 425 430

Leu Tyr Phe Thr Ile Pro Ala Asp Lys Val Asn Phe Ile Val Phe Thr
 435 440 445

Ser Ser Asp Thr Ile Thr Asp Arg Glu Glu Ala Leu Lys Ser Pro Leu
 450 455 460

Val Gln Val Met Leu Thr Leu Gly Ile Val Lys Glu Lys Asp Val Leu
 465 470 475 480

Phe Trp Ala

<210> 12

<211> 483

<212> PRT

<213> Artificial Sequence

<220><223> Artificial

<400> 12

Met Cys Asn Ser Asp Asn Thr Ser Leu Lys Glu Thr Val Ser Ser Asn
 1 5 10 15

Ser Ala Asp Val Val Glu Thr Glu Thr Tyr Gln Leu Thr Pro Ile Asp
 20 25 30

Ala Pro Ser Ser Phe Leu Ser His Ser Trp Glu Gln Thr Cys Gly Thr
 35 40 45

Pro Ile Leu Asn Glu Ser Asp Lys Gln Ala Ile Ser Phe Asp Phe Val
 50 55 60

Ala Pro Glu Leu Lys Gln Asp Glu Lys Tyr Cys Phe Thr Phe Lys Gly
 65 70 75 80

Ile Thr Gly Asp His Arg Tyr Ile Thr Asn Thr Thr Leu Thr Val Val
 85 90 95

Ala Pro Thr Leu Glu Val Tyr Ile Asp His Ala Ser Leu Pro Ser Leu
 100 105 110

Gln Gln Leu Ile His Ile Ile Gln Ala Lys Asp Glu Tyr Pro Ser Asn
 115 120 125

Gln Arg Phe Val Ser Trp Lys Arg Val Thr Val Asp Ala Asp Asn Ala
 130 135 140

Asn Lys Leu Asn Ile His Thr Tyr Pro Leu Lys Gly Asn Asn Thr Ser
 145 150 155 160

Pro Glu Met Val Ala Ala Ile Asp Glu Tyr Ala Gln Ser Lys Asn Arg
 165 170 175

Leu Asn Ile Glu Phe Tyr Thr Asn Thr Ala His Val Phe Asn Asn Leu
 180 185 190

Pro Pro Ile Ile Gln Pro Leu Tyr Asn Asn Glu Lys Val Lys Ile Ser
 195 200 205

His Ile Ser Leu Tyr Asp Asp Gly Ser Ser Glu Tyr Val Ser Leu Tyr
 210 215 220

Gln Trp Lys Asp Thr Pro Asn Lys Ile Glu Thr Leu Glu Gly Glu Val
 225 230 235 240

Ser Leu Leu Ala Asn Tyr Leu Ala Gly Thr Ser Pro Asp Ala Pro Lys
 245 250 255

Gly Met Gly Asn Arg Tyr Asn Trp His Lys Leu Tyr Asp Thr Asp Tyr
 260 265 270

Tyr Phe Leu Arg Glu Asp Tyr Leu Asp Val Glu Ala Asn Leu His Asp
 275 280 285

Leu Arg Asp Tyr Leu Gly Ser Ser Ala Lys Gln Met Pro Trp Asp Glu
 290 295 300

Phe Ala Lys Leu Ser Asp Ser Gln Gln Thr Leu Phe Leu Asp Ile Val
 305 310 315 320

Gly Phe Asp Lys Glu Gln Leu Gln Gln Gln Tyr Ser Gln Ser Pro Leu
 325 330 335

Pro Asn Phe Ile Phe Thr Gly Thr Thr Thr Trp Ala Gly Gly Glu Thr
 340 345 350

Lys Glu Tyr Tyr Ala Gln Gln Gln Val Asn Val Ile Asn Asn Ala Ile
 355 360 365

Asn Glu Thr Ser Pro Tyr Tyr Leu Gly Lys Asp Tyr Asp Leu Phe Phe
 370 375 380

Lys Gly His Pro Ala Gly Gly Val Ile Asn Asp Ile Ile Leu Gly Ser
 385 390 395 400

Phe Pro Asp Met Ile Asn Ile Pro Ala Lys Ile Ser Phe Glu Val Leu
 405 410 415

Met Met Thr Asp Met Leu Pro Asp Thr Val Ala Gly Ile Ala Ser Ser
 420 425 430

Ser Tyr Phe Thr Ile Pro Ala Asp Lys Val Asn Phe Ile Val Phe Thr
 435 440 445

Ser Ser Asp Thr Ile Thr Asp Arg Glu Glu Ala Leu Lys Ser Pro Leu
 450 455 460

Val Gln Val Met Leu Thr Leu Gly Ile Val Lys Glu Lys Asp Val Leu

465 470 475 480

Phe Trp Ala

<210> 13

<211> 483

<212> PRT

<213> Artificial Sequence

<220><223> Artificial

<400> 13

Met Cys Asn Ser Asp Asn Thr Ser Leu Lys Glu Thr Val Ser Ser Asn
 1 5 10 15

Ser Ala Asp Val Val Glu Thr Glu Thr Tyr Gln Leu Thr Pro Ile Asp
 20 25 30

Ala Pro Ser Ser Phe Leu Ser His Ser Trp Glu Gln Thr Cys Gly Thr

35 40 45

Pro Ile Leu Asn Glu Ser Asp Lys Gln Ala Ile Ser Phe Asp Phe Val
 50 55 60

Ala Pro Glu Leu Lys Gln Asp Glu Lys Tyr Cys Phe Thr Phe Lys Gly
 65 70 75 80

290 295 300
 Phe Ala Lys Leu Ser Asp Ser Gln Gln Thr Leu Phe Leu Asp Ile Val
 305 310 315 320
 Gly Phe Asp Lys Glu Gln Leu Gln Gln Gln Tyr Ser Gln Ser Pro Leu

 325 330 335
 Pro Asn Phe Ile Phe Thr Gly Thr Thr Thr Trp Ala Gly Gly Glu Thr
 340 345 350
 Lys Glu Tyr Tyr Ala Gln Gln Gln Val Asn Val Ile Asn Asn Ala Ile
 355 360 365
 Asn Glu Thr Ser Pro Tyr Tyr Leu Gly Lys Asp Tyr Asp Leu Phe Phe
 370 375 380
 Lys Gly His Pro Ala Gly Gly Val Ile Asn Asp Ile Ile Leu Gly Ser
 385 390 395 400

Phe Pro Asp Met Ile Asn Ile Pro Ala Lys Thr Ser Phe Glu Val Leu
 405 410 415
 Met Met Thr Asp Met Leu Pro Asp Thr Val Ala Gly Ile Ala Ser Ser
 420 425 430
 Ser Tyr Phe Thr Ile Pro Ala Asp Lys Val Asn Phe Ile Val Phe Thr
 435 440 445
 Ser Ser Asp Thr Ile Thr Asp Arg Glu Glu Ala Leu Lys Ser Pro Leu
 450 455 460
 Val Gln Val Met Leu Thr Leu Gly Ile Val Lys Glu Lys Asp Val Leu

465 470 475 480
 Phe Trp Ala

<210> 15

<211> 483

<212> PRT

<213> Artificial Sequence

<220><223> Artificial

<400> 15

Met Cys Asn Ser Asp Asn Thr Ser Leu Lys Glu Thr Val Ser Ser Asn

1 5 10 15
 Ser Ala Asp Val Val Glu Thr Glu Thr Tyr Gln Leu Thr Pro Ile Asp
 20 25 30
 Ala Pro Ser Ser Phe Leu Ser His Ser Trp Glu Gln Thr Cys Gly Thr
 35 40 45
 Pro Ile Leu Asn Glu Ser Asp Lys Gln Ala Ile Ser Phe Asp Phe Val
 50 55 60
 Ala Pro Glu Leu Lys Gln Asp Glu Lys Tyr Cys Phe Thr Phe Lys Gly
 65 70 75 80
 Ile Thr Gly Asp His Arg Tyr Ile Thr Asn Thr Thr Leu Thr Val Val
 85 90 95
 Ala Pro Thr Leu Glu Val Tyr Ile Asp His Ala Ser Leu Pro Ser Leu
 100 105 110

 Gln Gln Leu Ile His Ile Ile Gln Ala Lys Asp Glu Tyr Pro Ser Asn
 115 120 125
 Gln Arg Phe Val Ser Trp Lys Arg Val Thr Val Asp Ala Asp Asn Ala
 130 135 140
 Asn Lys Leu Asn Ile His Thr Tyr Pro Leu Lys Gly Asn Asn Thr Ser
 145 150 155 160
 Pro Glu Met Val Ala Ala Ile Asp Glu Tyr Ala Gln Ser Lys Asn Arg
 165 170 175
 Leu Asn Ile Glu Phe Tyr Thr Asn Thr Ala His Val Phe Asn Asn Leu

 180 185 190
 Pro Pro Ile Ile Gln Pro Leu Tyr Asn Asn Glu Lys Val Lys Ile Ser
 195 200 205
 His Ile Ser Leu Tyr Asp Asp Gly Ser Ser Glu Tyr Val Ser Leu Tyr
 210 215 220
 Gln Trp Lys Asp Thr Pro Asn Lys Ile Glu Thr Leu Glu Gly Glu Val
 225 230 235 240
 Ser Leu Leu Ala Asn Tyr Leu Ala Gly Thr Ser Pro Asp Ala Pro Lys
 245 250 255

Gly Met Gly Asn Arg Tyr Asn Trp His Lys Leu Tyr Asp Thr Asp Tyr
 260 265 270

Tyr Phe Leu Arg Glu Asp Tyr Leu Asp Val Glu Ala Asn Leu His Asp
 275 280 285

Leu Arg Asp Tyr Leu Gly Ser Ser Ala Lys Gln Met Pro Trp Asp Glu
 290 295 300

Phe Ala Lys Leu Ser Asp Ser Gln Gln Thr Leu Phe Leu Asp Ile Val
 305 310 315 320

Gly Phe Asp Lys Glu Gln Leu Gln Gln Gln Tyr Ser Gln Ser Pro Leu
 325 330 335

Pro Asn Phe Ile Phe Thr Gly Thr Thr Thr Trp Ala Gly Gly Glu Thr
 340 345 350

Lys Glu Tyr Tyr Ala Gln Gln Gln Val Asn Val Ile Asn Asn Ala Ile
 355 360 365

Asn Glu Thr Ser Pro Tyr Tyr Leu Gly Lys Asp Tyr Asp Leu Phe Phe
 370 375 380

Lys Gly His Pro Ala Gly Gly Val Ile Asn Asp Ile Ile Leu Gly Ser
 385 390 395 400

Phe Pro Asp Met Ile Asn Ile Pro Ala Lys Thr Ser Phe Glu Val Leu
 405 410 415

Met Met Thr Asp Met Leu Pro Asp Thr Val Ala Gly Ile Ala Ser Ser
 420 425 430

Thr Tyr Phe Thr Ile Pro Ala Asp Lys Val Asn Phe Ile Val Phe Thr
 435 440 445

Ser Ser Asp Thr Ile Thr Asp Arg Glu Glu Ala Leu Lys Ser Pro Leu
 450 455 460

Val Gln Val Met Leu Thr Leu Gly Ile Val Lys Glu Lys Asp Val Leu
 465 470 475 480

Phe Trp Ala

<210> 16

<211> 1449

<212> DNA

<213> Artificial Sequence

<220><223> Artificial

<400> 16

atgtgtaata gtgacaatac cagcttgaaa gaaacggtaa gctctaattc tgcagatgta 60
 gtagaacag aaacttacca actgacaccg attgatgctc ctgactcttt tttatctcat 120
 tcttgggagc aaacatgtgg cacacctatc ttgaatgaaa gtgacaagca agcgatatct 180
 tttgattttg ttgctccaga gttaaagcaa gatgaaaagt attgttttac ttttaaaggt 240

attacaggcg atcataggtat tatcacaaat acaacattaa ctgittgttg acctacgcta 300
 gaagtttaca tcatatgctc atccttacca tcgctacagc agcttatcca cattattcaa 360
 gcaaaagatg aatacccaag taatcaacgt tttgtctctt ggaagcgtgt aactgttgat 420
 gctgataatg ccaataagtt aaacattcat acttatccat taaaaggcaa taatacctca 480
 ccagaaatgg tggcagcgat tgatgagtat gctcagagca aaaatcgatt gaatatagag 540
 ttctatacaa atacagctca tgtttttaat aatttaccac ctattattca acctttatat 600
 aataacgaga aggtgaaaat ttctcatatt agtttgtatg atgatggttc ttctgaatat 660

gtaagtttat atcaatggaa agatacacca aataagatag aaacattaga aggtgaagta 720
 tcgcttcttg ctaattatct agcaggaaca tctccggatg caccaaaaagg aatgggaaat 780
 cgttataact ggcataaatt atatgacact gattattact ttttgcgca agattacctt 840
 gacgttgaag caaacctaca tgatttacgt gattatttag gctcttccgc aaagcaaatg 900
 ccatgggatg aatttgctaa attatctgat tctcagcaaa cactatcttt agatattgtg 960
 ggttttgata aagagcaatt gcaacaacaa tattcacaat cccactacc aaactttatt 1020
 tttaccggca caacaacttg ggctgggggg gaaacgaaag agtattatgc tcagcaacaa 1080

gtaaatgta ttaataatgc gatcaatgaa actagccctt attatttagg taaagactac 1140
 gatctatctt tcaaggggca tcctgctggt ggcgttatta acgacatcat tcttgggaagc 1200
 ttcctgata tgatcaatat tccagccaag acttcatttg aggtcttgat gatgacggat 1260
 atgttgccctg atacagtagc tggatttgcg agctctctgt acttcacaat tcctgccgat 1320
 aaagtttaatt ttattgtatt tacttcatct gacactatta ctgatcgtga agaggctctt 1380
 aatcaccat tagtacaagt gatgctaacg ttgggtattg ttaaagaaaa agatgttctg 1440
 ttctgggct 1449

<210> 17

<211> 1449

<212> DNA

<213> Artificial Sequence

<220><223> Artificial

<400> 17

atgtgtaata gtgacaatac cagcttgaaa gaaacggtaa gctctaattc tgcagatgta 60

gtagaaacag aaacttacca actgacaccg attgatgctc ctgactcttt tttatctcat 120

tcttgggagc aaacatgtgg cacacctatc ttgaatgaaa gtgacaagca agcgatatct 180

tttgattttg ttgctccaga gttaaagcaa gatgaaaagt attgttttac ttttaaaggt 240

attacaggcg atcataggta taccacaaat acaacattaa ctgttgttgc acctacgcta 300

gaagtttaca tcgatcatgc atccttacca tcgctacagc agcttatcca cattattcaa 360

gcaaaagatg aatacccaag taatcaacgt tttgtctctt ggaagcgtgt aactgttgat 420

gctgataatg ccaataagtt aaacattcat acttatccat taaaaggcaa taatacctca 480

ccagaaatgg tggcagcgat tgatgagtat gctcagagca aaaatcgatt gaatatagag 540

ttctatacaa atacagctca tgtttttaat aatttaccac ctattattca acctttatat 600

aataacgaga aggtgaaaat ttctcatatt agtttgtatg atgatggttc ttctgaatat 660

gtaagtttat atcaatggaa agatacacca aataagatag aaacattaga aggtgaagta 720

tcgcttcttg ctaattatct agcaggaaca tctccggatg caccaaaagg aatgggaaat 780

cgttataact ggcataaatt atatgacact gattattact ttttgcgcga agattacctt 840

gacgttgaag caaacctaca tgatttacgt gattatttag gctcttccgc aaagcaaatg 900

ccatgggatg aatttgctaa attatctgat tctcagcaaa cactatcttt agatattgtg 960

ggttttgata aagagcaatt gcaacaacaa tattcacaat cccactacc aaactttatt 1020

tttaccggca caacaacttg ggctgggggg gaaacgaaag agtattatgc tcagcaacaa 1080

gtaaatgtga ttaataatgc gatcaatgaa actagccctt attatttagg taaagactac 1140

gatctatctt tcaaggggca tcctgctggt ggcgttatta acgacatcat tcttgggaagc 1200

ttccctgata tgatcaatat tccagccaag atttcatttg aggtcttgat gatgacggat 1260

atgttgccctg atacagtagc tggatttgcg agctctagtt acttcacaat tcctgccgat 1320

aaagtttaatt ttattgtatt tacttcatct gacactatta ctgatcgtga agaggctctt 1380

aaatcaccat tagtacaagt gatgctaacg ttgggtattg ttaaagaaaa agatgttctg 1440

ttctgggct 1449

<210> 18

<211> 1449

<212> DNA

<213> Artificial Sequence

<220><223> Artificial

<400> 18

atgtgtaata gtgacaatac cagcttgaaa gaaacggtaa gctctaattc tgcagatgta 60

gtagaaacag aaacttacca actgacaccg attgatgctc ctgactcttt tttatctcat 120

tcttgggagc aaacatgtgg cacacctatc ttgaatgaaa gtgacaagca agcgatatct 180

tttgattttg ttgctccaga gttaaagcaa gatgaaaagt attgttttac ttttaaaggt 240

attacaggcg atcataggtat taccacaat acaacattaa ctgttgttgc acctacgcta 300

gaagtttaca tcatgatgc atccttacca tcgctacagc agcttatcca cattattcaa 360

gcaaaagatg aatacccaag taatcaacgt tttgtctctt ggaagcgtgt aactgttgat 420

gctgataatg ccaataagtt aaacattcat acttatccat taaaaggcaa taatacctca 480

ccagaaatgg tggcagcgat tgatgagtat gctcagagca aaaatcgatt gaatatagag 540

ttctatacaa atacagctca tgtttttaat aatttaccac ctattattca acctttatat 600

aataacgaga aggtgaaaat ttctcatatt agtttgtatg atgatggttc ttctgaatat 660

gtaagtttat atcaatggaa agatacacca aataagatag aacattaga aggtgaagta 720

tcgcttcttg ctaattattt agcaggaaca tctccggatg caccaaaagg aatgggaaat 780

cgttataact ggcataaatt atatgacact gattattact ttttgcgcga agattacctt 840

gacgttgaag caaacctaca tgatttacgt gattatttag gctcttccgc aaagcaaatg 900

ccatgggatg aatttgctaa attatctgat tctcagcaaa cactattttt agatattgtg 960

ggttttgata aagagcaatt gcaacaacaa tattcacaat cccactacc aaactttatt 1020

tttaccggca caacaacttg ggctgggggg gaaacgaaag agtattatgc tcagcaacaa 1080

gtaaatgtga ttaataatgc gatcaatgaa actagccctt attatttagg taaagactac 1140

gatctatfff tcaaggggca tctgctggtt ggcgttatta acgacatcat tcttgggaagc 1200

ttccctgata tgatcaatat tccagccaag atttcatttg aggtcttgat gatgacggat 1260

atgttgcctg atacagtagc tggatttgcg agctctacgt acttcacaat tcctgccgat 1320

aaagttaatt ttattgtatt tacttcatct gacactatta ctgatcgtga agaggctctt 1380

aaatcaccat tagtacaagt gatgctaacg ttgggtattg ttaaagaaaa agatgttctg 1440

ttctgggct 1449

<210> 19

<211> 1449

<212> DNA

<213> Artificial Sequence

<220><223> Artificial

<400> 19

atgtgtaata gtgacaatac cagcttgaaa gaaacggtaa gctctaattc tgcagatgta 60
 gtagaaacag aaacttacca actgacaccg attgatgctc ctagctcttt tttatctcat 120
 tcttgggagc aaacatgtgg cacacctatc ttgaatgaaa gtgacaagca agcgatatct 180

tttgattttg ttgctccaga gttaaagcaa gatgaaaagt attgttttac ttttaaaggt 240
 attacaggcg atcataghta taccacaat acaacattaa ctgttgttgc acctacgcta 300
 gaagtttaca tcatcatgc atccttacca tctctacagc agcttatcca cattattcaa 360
 gcaaaagatg aatacccaag taatcaacgt tttgtctctt ggaagcgtgt aactgttgat 420
 gctgataatg ccaataagtt aaacattcat acttatccat taaaaggcaa taatacctca 480
 ccagaaatgg tggcagcgat tgatgagtat gctcagagca aaaatcgatt gaatatagag 540
 ttctatacaa atacagctca tgtttttaat aatttaccac ctattattca acctttatat 600

aataacgaga aggtgaaaat ttctcatatt agtttgtatg atgatggttc ttctgaatat 660
 gtaagtttat atcaatggaa agatacacca aataagatag aacattaga aggtgaagta 720
 tctctctctg ctaattatct agcaggaaca tctccggatg caccaaaagg aatgggaaat 780
 cgttataact ggcataaatt atatgacact gattattact ttttgcgcga agattacctt 840
 gacgttgaag caaacctaca tgatttacgt gattatctag gctcttccgc aaagcaaatg 900
 ccatgggatg aatttgctaa attatctgat tctcagcaaa cactatcttt agatattgtg 960
 ggttttgata aagagcaatt gcaacaacaa tattcacaat cccactacc aaactttatt 1020

tttaccggca caacaacttg ggctgggggg gaaacgaaag agtattatgc tcagcaacaa 1080
 gtaaatgtga ttaataatgc gatcaatgaa actagccctt attatctagg taaagactac 1140
 gatctatctt tcaaggggca tctgctggtt ggcgttatta acgacatcat tcttgggaagc 1200
 ttcctgata tgatcaatat tccagccaag acttcatttg aggtcttgat gatgacggat 1260
 atgttgccctg atacagtagc tggatttgcg agctctagtt acttcacaat tctgcccgat 1320
 aaagttaatt ttattgtatt tacttcatct gacactatta ctgatcgtga agaggctctt 1380
 aaatcaccat tagtacaagt gatgctaacg ttgggtattg ttaaagaaaa agatgttctg 1440

ttctgggct 1449

<210> 20

<211> 1449

<212> DNA

<213> Artificial Sequence

<220><223> Artificial

<400> 20

atgtgtaata gtgacaatac cagcttgaaa gaaacggtaa gctctaattc tgcagatgta 60
 gtagaaacag aaacttacca actgacaccg attgatgctc ctgactcttt tttatctcat 120
 tcttgggagc aaacatgtgg cacacctatc ttgaatgaaa gtgacaagca agcgatatct 180
 tttgattttg ttgctccaga gttaaagcaa gatgaaaagt attgttttac ttttaaaggt 240
 attacaggcg atcataggta tatcacaaat acaacattaa ctgttgttgc acctacgcta 300

gaagtttaca tcgatcatgc atccttacca tcgctacagc agcttatcca cattattcaa 360
 gcaaaagatg aatacccaag taatcaacgt tttgtctctt ggaagcgtgt aactgttgat 420
 gctgataatg ccaataagtt aaacattcat acttatccat taaaaggcaa taatacctca 480
 ccagaaatgg tggcagcgat tgatgagtat gctcagagca aaaatcgatt gaatatagag 540
 ttctatacaa atacagctca tgtttttaat aatttaccac ctattattca acctttatat 600
 aataacgaga aggtgaaaat ttctcatatt agtttgtatg atgatggttc ttctgaatat 660
 gtaagtttat atcaatggaa agatacacca aataagatag aaacattaga aggtgaagta 720

tcgcttcttg ctaattattht agcaggaaca tctccggatg caccaaaagg aatgggaaat 780
 cgttataact ggcataaatt atatgacact gattattact ttttgcgcga agattacctt 840
 gacgttgaag caaacctaca tgatttacgt gattatthtag gctcttccgc aaagcaaatg 900
 ccatgggatg aatttgctaa attatctgat tctcagcaaa cactatthtt agatattgtg 960
 ggttttgata aagagcaatt gcaacaacaa tattcacaat cccactacc aaactttatt 1020
 tttaccggca caacaacttg ggctgggggg gaaacgaaag agtattatgc tcagcaacaa 1080
 gtaaatgtga ttaataatgc gatcaatgaa actagccctt attatthtagg taaagactac 1140

gatctattht tcaaggggca tcctgctggg ggcgttatta acgacatcat tcttgggaagc 1200
 ttcctgata tgatcaatat tccagccaag acttcatttg aggtcttgat gatgacggat 1260
 atgttgccctg atacagttagc tggatttgcg agctctacct acttcacaat tcctgccgat 1320
 aaagtttaatt ttattgtatt tacttcatct gacactatta ctgatcgtga agaggctctt 1380
 aaatcccat tagtacaagt gatgctaacg ttgggtattg ttaaagaaaa agatgttctg 1440
 ttctgggct 1449

<210> 21

<211> 391

<212> PRT

<213> Pasteurella multocida

<400> 21

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

1 5 10 15

Asn Gln Leu Met Asp Phe Thr Gln Asn Asn Glu Asp Lys Thr His Pro

20 25 30

Arg Ile Phe Gly Leu Ser Arg Phe Lys Ile Pro Asp Asn Ile Ile Thr

35 40 45

Gln Tyr Gln Asn Ile His Phe Val Glu Leu Lys Asp Asn Arg Pro Thr

50 55 60

Glu Ala Leu Phe Thr Ile Leu Asp Gln Tyr Pro Gly Asn Ile Glu Leu

65 70 75 80

Asn Ile His Leu Asn Ile Ala His Ser Val Gln Leu Ile Arg Pro Ile

85 90 95

Leu Ala Tyr Arg Phe Lys His Leu Asp Arg Val Ser Ile Gln Gln Leu

100 105 110

Asn Leu Tyr Asp Asp Gly Ser Met Glu Tyr Val Asp Leu Glu Lys Glu

115 120 125

Glu Asn Lys Asp Ile Ser Ala Glu Ile Lys Gln Ala Glu Lys Gln Leu

130 135 140

Ser His Tyr Leu Leu Thr Gly Lys Ile Lys Phe Asp Asn Pro Thr Ile

145 150 155 160

Ala Arg Tyr Val Trp Gln Ser Ala Phe Pro Val Lys Tyr His Phe Leu

165 170 175

Ser Thr Asp Tyr Phe Glu Lys Ala Glu Phe Leu Gln Pro Leu Lys Glu

180 185 190

Tyr Leu Ala Glu Asn Tyr Gln Lys Met Asp Trp Thr Ala Tyr Gln Gln

195 200 205

Leu Thr Pro Glu Gln Gln Ala Phe Tyr Leu Thr Leu Val Gly Phe Asn

210 215 220

Asp Glu Val Lys Gln Ser Leu Glu Val Gln Gln Ala Lys Phe Ile Phe

225 230 235 240

Thr Gly Thr Thr Thr Trp Glu Gly Asn Thr Asp Val Arg Glu Tyr Tyr
 245 250 255
 Ala Gln Gln Gln Leu Asn Leu Leu Asn His Phe Thr Gln Ala Glu Gly
 260 265 270
 Asp Leu Phe Ile Gly Asp His Tyr Lys Ile Tyr Phe Lys Gly His Pro
 275 280 285
 Arg Gly Gly Glu Ile Asn Asp Tyr Ile Leu Asn Asn Ala Lys Asn Ile
 290 295 300
 Thr Asn Ile Pro Ala Asn Ile Ser Phe Glu Val Leu Met Met Thr Gly
 305 310 315 320
 Leu Leu Pro Asp Lys Val Gly Gly Val Ala Ser Ser Leu Tyr Phe Ser
 325 330 335
 Leu Pro Lys Glu Lys Ile Ser His Ile Ile Phe Thr Ser Asn Lys Gln
 340 345 350

Val Lys Ser Lys Glu Asp Ala Leu Asn Asn Pro Tyr Val Lys Val Met
 355 360 365
 Arg Arg Leu Gly Ile Ile Asp Glu Ser Gln Val Ile Phe Trp Asp Ser
 370 375 380
 Leu Lys Gln Leu Gly Gly Gly
 385 390

<210> 22

<211> 1173

<212> DNA

<213> Pasteurella multocida

<400> 22

atgaaaacaa tcacgctgta tttagatcct gcctccttac cggcattaaa tcagctgatg 60
 gactttacgc aaaataatga agataaaaaca catccacgta tttttggctct ttctcgcttt 120

aaaatccctg acaacattat tacacagtat caaaatatcc atttcgtcga actcaaagat 180
 aatcgtccca ctgaagcact ttttacgatt ttagatcaat accctggtaa cattgagtta 240
 aatatacact taaatattgc tcattccgtt caattaattc gtcgatttt ggcatatcgt 300
 tttaaacatt tagatcgtgt atcaattcag cagttaaatc tttatgacga tggetcaatg 360
 gaatatgttg atttagaaaa agaagaaaat aaagatatit cgcagaaat taagcaagca 420

gaaaaacaac tttctcacta tttgcttact ggcaaaataa aatttgataa cccaactatt 480
gctcgttatg tctggcaatc cgcgttccca gtaaaatatac atttttaag tacagactat 540

tttgaaaaag ccgaattttt acaaccacta aaagaatatt tagcagaaaa ttatcaaaaa 600
atggactgga ctgcttacca acagctgact ccagaacagc aagcattcta cttaacattg 660
gtaggcttca atgacgaagt caagcagtcg ctagaagtgc aacaagctaa atttatcttt 720
accggcacga caacttggga aggaatacc gatgtgcgag aatactacgc acagcaacaa 780
cttaatttac ttaatcactt tacccaagct gagggcgatt tatttattgg tgatcattat 840
aaaatctact ttaaagggca tcctagaggt ggtgaaatta atgactacat tctgaacaat 900
gctaaaaata tcaccaatat ccctgccaat atttcctttg aagtattgat gatgacaggc 960

ttattacctg ataaagtggg tgggtttgca agttcactgt atttctcctt accaaaagaa 1020
aaaattagcc atattatfff cacatcgaat aaacaagtga aaagcaaaga agatgcgcta 1080
aataatccgt atgtaaggt catgcgtcgt ttaggtataa ttgacgaatc acaagtcac 1140
ttttgggaca gtttaaaaca gttgggtgga ggt 1173

<210> 23

<211> 483

<212> PRT

<213> Photobacterium damsela

<400> 23

Met Cys Asn Ser Asp Asn Thr Ser Leu Lys Glu Thr Val Ser Ser Asn

1 5 10 15

Ser Ala Asp Val Val Glu Thr Glu Thr Tyr Gln Leu Thr Pro Ile Asp

20 25 30

Ala Pro Ser Ser Phe Leu Ser His Ser Trp Glu Gln Thr Cys Gly Thr

35 40 45

Pro Ile Leu Asn Glu Ser Asp Lys Gln Ala Ile Ser Phe Asp Phe Val

50 55 60

Ala Pro Glu Leu Lys Gln Asp Glu Lys Tyr Cys Phe Thr Phe Lys Gly

65 70 75 80

Ile Thr Gly Asp His Arg Tyr Ile Thr Asn Thr Thr Leu Thr Val Val

85 90 95

Ala Pro Thr Leu Glu Val Tyr Ile Asp His Ala Ser Leu Pro Ser Leu

ttctatacaa atacagctca tgtttttaat aatttaccac ctattattca acctttatat	600
aataacgaga aggtgaaaat ttctcatatt agtttgtatg atgatggttc ttctgaatat	660
gtaagtttat atcaatggaa agatacacca aataagatag aaacattaga aggtgaagta	720
tcgcttcttg ctaattatit agcaggaaca tctccggatg caccaaaagg aatgggaaat	780
cgttataact ggcataaatt atatgacct gattattact ttttgcgcga agattacctt	840
gacgttgaag caaacctaca tgatttacgt gattatttag gctcttccgc aaagcaaatg	900
ccatgggatg aatttgctaa attatctgat tctcagcaaa cactatitit agatattgtg	960
ggttttgata aagagcaatt gcaacaacaa tattcacaat cccactacc aaactttatt	1020
tttaccggca caacaacttg ggctgggggg gaaacgaaag agtattatgc tcagcaacaa	1080
gtaaatgtga ttaataatgc gatcaatgaa actagccctt attatttagg taaagactac	1140
gatctatit tcaaggggca tcttctggt ggcgttatta acgacatcat tcttgaagc	1200
ttccctgata tgatcaatat tccagccaag atitcatttg aggtcttgat gatgacggat	1260
atgttgcctg atacagtagc tggatttgcg agctctctgt acttcacaat tcttccgat	1320
aaagttaatt ttattgtatt tacttcatct gacactatta ctgatcgtga agaggctctt	1380
aaatcaccat tagtacaagt gatgctaacg ttgggtattg ttaaagaaaa agatgttctg	1440
ttctgggct	1449