



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 198 56 463 B4 2006.02.02**

(12)

Patentschrift

(21) Aktenzeichen: **198 56 463.5**
 (22) Anmeldetag: **26.11.1998**
 (43) Offenlegungstag: **08.06.2000**
 (45) Veröffentlichungstag
 der Patenterteilung: **02.02.2006**

(51) Int Cl.⁸: **C12N 15/86 (2006.01)**
C12N 7/00 (2006.01)
C12N 5/10 (2006.01)
A61K 48/00 (2006.01)

Innerhalb von drei Monaten nach Veröffentlichung der Patenterteilung kann nach § 59 Patentgesetz gegen das Patent Einspruch erhoben werden. Der Einspruch ist schriftlich zu erklären und zu begründen. Innerhalb der Einspruchsfrist ist eine Einspruchsgebühr in Höhe von 200 Euro zu entrichten (§ 6 Patentkostengesetz in Verbindung mit der Anlage zu § 2 Abs. 2 Patentkostengesetz).

(73) Patentinhaber:
Heinrich-Pette-Institut, 20251 Hamburg, DE

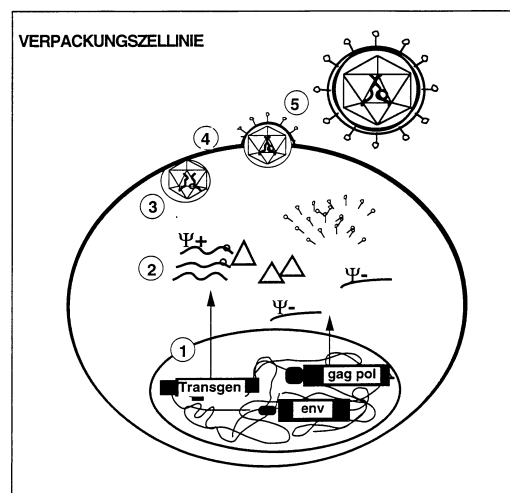
(74) Vertreter:
Uexküll & Stolberg, 22607 Hamburg

(72) Erfinder:
Laer, Meike-Dorothee von, 22529 Hamburg, DE;
Beyer, Winfried, 21035 Hamburg, DE

(56) Für die Beurteilung der Patentfähigkeit in Betracht
 gezogene Druckschriften:
Ally, B.A. et al., J. Immunol. 155(11), 1995, 5404-8;
Cosset, F.-L. et al., J. Virol. 69(12), 1995, 7430-6;

(54) Bezeichnung: **Retrovirale, mit LCMV pseudotypisierte Hybrid-Vektoren**

(57) Hauptanspruch: Retrovirale Verpackungszelle, dadurch gekennzeichnet, daß sie pseudotypisierte Virionen exprimiert, die LCMV-Glykoprotein in ihrer Hülle eingebaut enthalten.



Beschreibung

[0001] Die vorliegende Erfindung betrifft allgemein die Pseudotypisierung von Retroviren mit lymphozytärem Choriomeningitisvirus. Insbesondere betrifft die Erfindung die Pseudotypisierung in MLV-Verpackungszellen, die gegebenenfalls env-deletiert sind, oder in Verpackungszellen, die von Lentiviren abgeleitet sind. Vorzugsweise erfolgt die Pseudotypisierung durch Infektion mit LCMV oder einer vorzugsweise env-deletierten Mutante, oder durch Transfektion mit einem das gp-Gen des LCMV oder einen Teil desselben und gegebenenfalls zusätzlich das np-, das I- und/oder das z-Gen des LCMV enthaltenden Expressionsplasmid. Die Erfindung betrifft ferner die Verwendung derartiger Pseudotypen zur Infektion von Zellen, insbesondere die Verwendung in der Gentherapie.

Stand der Technik

[0002] Retrovirale Vektoren finden im Stand der Technik zunehmend Verwendung, beispielsweise für den Gentransfer in der gentechnischen bzw. medizinischen Forschung oder bei gentherapeutischen Ansätzen (vgl. z.B. C. Baum et al. in *Seminars in Oncology: Gene Therapy of Cancer: Translational approaches from preclinical studies to clinical implementations.*, eds. Gerson, & Lattime, Academic Press, 1998). Die retroviralen Vektoren sind meist von murinen Leukämieviren (MLV) abgeleitet und enthalten alle für die Integration notwendigen Sequenzen der LTR-Regionen und das für die Verpackung verantwortliche ψ -Element. Die für die Virusproteine codierenden Bereiche sind durch Fremdgene und deren Kontrollsequenzen ersetzt, die man in menschliche Zellen einbringen möchte. Die Vektoren werden in sogenannten Helferzelllinien (Verpackungszelllinien) exprimiert, die eine Kopie eines kompletten Retrovirusgenoms enthalten. Es synthetisiert alle für die Replikation und Infektion notwendigen Proteine, kann jedoch seine genomische Virus-RNA nicht in Partikel verpacken, da es einen Defekt in den ψ -Sequenzen aufweist. Werden die retroviralen Vektoren in diese Helferzellen eingebracht und transkribiert, kann die gebildete transgene mRNA durch die ihr eigene ψ -Region mit den Strukturproteinen des Helfervirus interagieren und zu Partikeln verpackt werden. Die rekombinanten Virionen, die keinerlei Erbinformation für Viruskomponenten besitzen, adsorbieren über ihre Oberflächenproteine an Zellen, die Capside werden in das Cytoplasma aufgenommen, und die transgene RNA wird in doppelsträngige DNA überschrieben und in das Wirtszellgenom integriert. Der Vorteil dieses Systems ist die stabile Integration der Fremdgene, die bei Teilungen auf die Tochterzellen weitergegeben werden. Nachteilig ist die Retroviren eigene unspezifische Integration an willkürlichen Stellen des Zellgenoms.

[0003] Retrovirale Vektoren vermitteln eine stabile, kolineare Integration (d.h. ohne Rekombinationen und Rearrangierung der kodierenden Sequenzen im Vektorgenom) und dadurch eine langfristige Expression des Transgens. Z. B. ist so möglich, mit Hilfe von retroviralen Vektoren Knochenmarkszellen zu transzifizieren, so dass sie LCMV-Glykoprotein auf ihrer Oberfläche exprimieren und dadurch T-Zell-Toleranz gegen dieses Protein induzieren (Ally et al., 1995, *J. Immunol* 155, 5404-8). Eine langfristige Genexpression ist sonst bislang nur noch durch die episomalen Herpesvirusvektoren oder die Adeno-associated virus-Vektoren (AAV-Vektoren) möglich.

[0004] Für die letztgenannten Vektorsysteme sind die Verpackungssysteme (Verpackungszelllinien) jedoch noch nicht optimiert. AAV-Vektoren weisen ferner eine geringe Verpackungskapazität (ca. 5 kb für AAV gegenüber ca. 10-12 kb für retrovirale Vektoren) auf.

[0005] Verpackungslinien exprimieren neben dem zu transferierenden Gen, dem Transgen, auch das Vektor-Genom, das retrovirale cis-Elemente enthält. Das genomische Vektortranskript kodiert also nicht für retrovirale Proteine, wird aber in den Verpackungslinien mit Hilfe der gag-, pol- und env-Genprodukte in ein infektiöses, aber nicht replikationsfähiges Virion eingebaut. Dieses Virion kann dann als retroviraler Vektor zum Transfer des in das Vektorgenom integrierten Transgens in die gewünschten Zielzellen eingesetzt werden, ohne daß es dort zur weiteren Vermehrung des Vektors kommt. Mit anderen Worten, der virale Vektor kann die Zielzellen nur infizieren, sich dort aber nicht weiter vermehren.

[0006] Die Entwicklung retroviraler Verpackungssysteme ist bereits weit fortgeschritten, und Vektorüberstände, die frei von replikationskompetenten Viren sind, können in großen Mengen unter GMP-Bedingungen (Good manufacturing practice; Richtlinie der Kommission zur Festlegung der Grundsätze und Leitlinien der guten Herstellungspraxis (GMP) für bestimmte Arzneimittel zur Anwendung beim Menschen (91/356/EWG) vom 13.6.91) hergestellt werden. Vektoren auf Basis des murinen Leukämievirus (MLV-Vektoren) wurden bereits mehrfach in klinischen Studien eingesetzt (P. Chu et al., *J. Mol. Med.* 76 (1998) 184-192).

[0007] Im Stand der Technik sind zwei grundsätzliche Typen retroviraler Verpackungssysteme bekannt (J.M.

Wilson, Clin. Exp. Immunol 107 Suppl. 1 (1997) 31-32; C. Baum et al. (1998), loc. cit.).

[0008] MLV-Verpackungszelllinien enthalten die retroviralen Gene gag, pol und env (**Abb. 1**), und die für die Verpackung der retroviralen RNA erforderliche Sequenzen sind deletiert (C. Baum et al. (1998), loc. cit.).

[0009] Der zweite Typ der bekannten Verpackungssysteme leitet sich von den Lentiviren ab (R. Carroll et al., J. Virol. 68 (1994) 6047-6051; P. Corbeau et al., Proc. Natl. Acad. Sci. USA 93 (1996) 14070-14075; L. Naldini et al., Science 272 (1996) 263-267; C. Parolin et al., J. Virol. 68 (1994) 3888-3895; J. Reiser et al., Proc. Natl. Acad. Sci. USA 93 (1996) 15266-15271; J.H. Richardson et al., J. Gen. Virol. 76 (1995) 691-696; T. Shimada et al., J. Clin. Invest. 88 (1991) 1043-1047). Lentiviren sind komplexe Retroviren, die zusätzlich zu den gag-, pol- und env-Genprodukten noch eine Reihe regulatorischer Gene exprimieren. Beispiele für Lentiviren, aus denen Verpackungssysteme abgeleitet wurden, sind das humane Immundefizienzvirus (HIV), das "Simian Immundefizienzvirus" (SIV) und das "Feline Immundefizienzvirus" (FIV). Der Aufbau der lentiviralen Verpackungssysteme ähnelt prinzipiell demjenigen der MLV-Vektoren.

[0010] Vorteil der lentiviralen Vektoren ist, daß sie auch ruhende Zellen infizieren können. Bei MLV-Vektoren hingegen kann das Vektorgenom nur während der Zellteilung in den Zellkern transportiert werden, d.h. wenn die Kernmembran aufgelöst ist. Allerdings weisen von Lentiviren abgeleitete Verpackungssysteme aufgrund des komplexen Aufbaus des lentiviralen Genoms Nachteile auf, die sich in einem vergleichsweise geringen Titer und einer geringeren Sicherheit äußern. Durch den komplexen Genomaufbau lassen sich cis- und trans-Elemente im Genom nicht klar voneinander trennen. In den Verpackungskonstrukten, die lentivirale gag-, pol- und env-Gene exprimieren, befinden sich daher auch wichtige cis-regulatorische Sequenzen (z.B. Teile des Verpackungs-Signals), die auch im Vektor-Genom enthalten sein müssen. Durch diese Homologien kann es zu Rekombinationen zwischen Vektorgenom und den Verpackungskonstrukten und damit zur Freisetzung replikationskompetenter Retroviren (z.B. einem HIV-Wildvirus, was hochgradig unerwünscht wäre) kommen, so daß diese Systeme nicht mit MLV-Verpackungslinien vergleichbar sind.

[0011] Alle bislang im Stand der Technik bekannten Vektorsysteme weisen ferner einige entscheidende Mängel auf, die einen erfolgreichen Einsatz in der Gentherapie verhindern: 1. Retrovirale Vektoren werden meist nur in unzureichenden Titern produziert und können durch die Instabilität ihrer Hüllproteine nicht weiter aufkonzentriert werden. 2. Vektorpartikel können aufgrund der Instabilität ihrer Hüllproteine nicht ohne Verlust der Infektiosität aufgereinigt werden. Eine solche Aufreinigung ist aber essentiell, da die Zellkulturüberstände, aus denen Vektoren geerntet werden, durch zelluläre Bestandteile verunreinigt sind. 3. Durch ihre Hüllproteine werden retrovirale Vektoren von humanem Serum-Komplement inaktiviert. 4. Der Rezeptor für das Hüllprotein der klassischen, amphotropen Vektoren wird auf fast allen in Betracht gezogenen Zelllinien exprimiert. Allerdings weisen viele primäre humane Zellen, wie Hepatozyten und hämatopoietische Stammzellen, die attraktive Ziele der Gentherapie sind, einen Mangel an funktionellen amphotropen Rezeptoren auf, wodurch eine Transduktion erschwert oder verhindert wird.

Aufgabenstellung

[0012] Aufgabe der vorliegenden Erfindung ist es daher, retrovirale Verpackungssysteme zur Verfügung zu stellen, die die Nachteile der im Stand der Technik bekannten Verpackungszelllinien nicht aufweisen.

[0013] Insbesondere ist es Aufgabe der vorliegenden Erfindung, Verpackungssysteme zur Verfügung zu stellen, die einen stabilen, retroviralen Transfer von Transgenen in die Zielzellen ermöglichen, d.h., zu einer stabilen Integration des Transgens in das Genom der Ziel- oder Wirtszellen und einer anschließenden stabilen Expression dieses Gens führen.

[0014] Die Aufgabe wird erfindungsgemäß dadurch gelöst, daß man Retroviren mit Lymphozytärem Choriomeningitisvirus (LCMV) pseudotypisiert.

[0015] Die vorliegende Erfindung betrifft daher ein rekombinantes Virion, das vorzugsweise mit einem oder mehreren Fremdgenen transfiziert ist, welches durch Pseudotypisierung des Viruspartikels mit Lymphozytärem Choriomeningitisvirus (LCMV) erhältlich ist.

[0016] Der Tropismus und auch die Stabilität eines Virus wird in erster Linie vom Hüllprotein bestimmt. Murine Retroviren können, neben den MLV-env kodierten Glykoproteinen, auch Hüllproteine anderer Virusarten in ihre Virushülle einbauen. Dadurch entstehen sogenannte Pseudotypen. Retrovirale Pseudotypvektoren entstehen durch Expression fremder viraler Hüllproteine in MLV-Verpackungslinien. Herkömmliche MLV-Verpackungszel-

linien enthalten die retroviralen Gene gag, pol und env. Sequenzen, die für die Verpackung retroviraler genomischer RNA notwendig sind, wurden deletiert. In solche Verpackungslinien wird ein Vektor eingebracht, der neben dem Gen, das transferiert werden soll, auch die retrovirale Verpackungssequenz und weitere retrovirale cis Elemente enthält (LTR, leader). Das retrovirale RNA-Genom wird mit Hilfe der Gag-, Pol- und Env-Genprodukte in ein infektiöses, aber nicht replikationsfähiges Virion eingebaut. Dieses Virion kann dann als retroviraler Vektor zur Transduktion von Zellen eingesetzt werden. Pseudotyp-Verpackungslinien enthalten zusätzlich das Hüllproteingen eines fremden Virus. So sind z. B. im Stand der Technik Verpackungszellen zur Pseudotypisierung retroviraler Vektoren (Moloney MLV) durch Glykoproteine des Amphitropen MLV oder „Cat endogenous virus“ RD 114 bekannt. Die erfindungsgemäßen Pseudotyp-Verpackungslinien enthalten das Hüllproteingen des LCMV, und es erfolgt eine Expression der LCMV-Glykoproteine.

[0017] Mit der vorliegenden Erfindung werden erstmals Vektorsysteme zur Verfügung gestellt, die in hohen Titern produziert werden und auf konzentriert werden können. Die erfindungsgemäßen Vektorpartikel lassen sich ferner ohne oder ohne wesentlichen Verlust der Infektiosität auf reinigen. Es hat sich im Rahmen der vorliegenden Erfindung überraschenderweise gezeigt, daß die erfindungsgemäße Pseudotypisierung für die Verpackungszellen nicht zelltoxisch ist. Es werden somit erstmals stabile Verpackungszelllinien (Verpackungssysteme) zur Verfügung gestellt, die einen stabilen, retroviralen Transfer von Transgenen in die Zielzellen ermöglichen, d.h., zu einer stabilen Integration des Transgens in das Genom der Ziel- oder Wirtszellen und einer anschließenden stabilen Expression dieses Gens führen.

[0018] Die erfindungsgemäßen Zelllinien zeichnen sich ferner durch ein breites, spezieübergreifendes Wirtszellspektrum (Zelltropismus) aus. Ein entscheidender Vorteil der vorliegenden Erfindung ist die Tatsache, daß einzelne Mutationen im Hüllprotein des LCMV zu einer Änderung des Tropismus von LCMV führen können. D.h., durch einzelne Punktmutationen in gp werden Viren, die eher Nervenzellen infizieren, zu Viren, die eher Lymphozyten infizieren oder zu solchen, die eher Monozyten infizieren.

[0019] Im Rahmen der vorliegenden Erfindung wird zur Pseudotypisierung LCMV eingesetzt. Dabei ist es möglich oder kann sogar bevorzugt sein, statt des LCMV-Wildtyps andere LCMV-Stämme einzusetzen. So können leichte Variationen in der gp-Nukleinsäuresequenz bzw. in der Aminosäure-Sequenz des exprimierten Hüllproteins in verschiedenen Stämmen von LCMV den Zelltropismus (Wirtszellspektrum) von LCMV erheblich verändern (M. Matloubian et al., J. Virol. 67 (1993) 7340-7349; M.N. Teng, J. Virol. 70 (1996) 8438-8443; King et al., J. Virol 64; 1990, 5611-5616). Solche Tropismusvarianten im Glykoprotein findet man für keinen der bisher bekannten retroviralen Vektorsysteme, und es wird erfindungsgemäß erstmals eine gezieltere Transduktion des gewünschten Zelltyps ermöglicht. Gemäß einer bevorzugten Ausführungsform der Erfindung kann es daher von Vorteil sein, Verpackungssysteme mit verschiedenen Glykoprotein-Varianten (GP-Varianten) für unterschiedliche Anwendungen anzulegen.

[0020] Im Rahmen der vorliegenden Erfindung wird vorzugsweise von den gp-Genen des eher neurotrophen LCMV-Stammes Armstrong, L(ARM) (L. Villarete et al., J. Virol. 68 (1994) 7490-7496) (für SEQ ID NO: 4 kodierender Bereich; vgl. Anlage zum Sequenzprotokoll, ZU SEQ ID NO: 3), und des eher hämatotropen Stammes WE (V. Romanowski et al., Virus Res. 3, (1985) 101-114) (SEQ ID NO: 1) ausgegangen. Erfindungsgemäß eingeschlossen sind auch Varianten (Tropismusvarianten) dieser beiden Stämme, bei denen im gp-Genprodukt einzelne Aminosäuren ausgetauscht sind, da dadurch der Tropismus des Virus verändert werden kann.

[0021] Da LCMV ein RNA-Virus ist, das im Vermehrungszyklus keine Kernphase hat, erscheint es sehr wahrscheinlich, daß sich "kryptische Spleiß-Regionen" in der RNA Sequenz befinden. Die Entfernung (Bereinigung) solcher Regionen kann ausgenutzt werden, um eine verbesserte Expression zu erzielen. Solche "spleißbereinigte" Varianten sind daher erfindungsgemäß ebenfalls eingeschlossen. Durch eine solche Optimierung der GP-Expression kann die Pseudotypisierung verbessert werden, wobei auf eine zusätzliche Unterstützung durch mindestens ein weiteres LCMV-Protein verzichtet werden kann.

[0022] Zur Expression von LCMV eignen sich allgemein Expressionsvektoren, die eine hohe stabile Genexpression in eukaryontischen Zellen ermöglichen. Die Wahl des Expressionsvektors ist für die Verpackung der retroviralen LCMV-Pseudotypen jedoch nur insoweit entscheidend, als er ein hohes und stabiles Expressionsniveau gewährleisten muß, d.h. ein Expressionsniveau, das hoch genug ist, um die Bildung von Pseudotypen zu ermöglichen, und das dauerhaft (stabil) ist, ohne daß es zum Abschalten des Promoters kommt.

[0023] Erfindungsgemäß besonders bevorzugt sind die folgenden beiden Expressionskassetten: (CMV-Promoter)--(β-Globin-Intron-2)--(gp)--(SV40 poly A-Signal) und

(EF-lalpha-Promoter)--(gp)--(poly-A-Signal des G-CSF-Genes)

(S. Mizushima, Nucleic Acids Res. 18 (1990) 5322, T. Uetsuki, J. Biol. Chem. 264 (1989) 5791-5798).

[0024] Die Sequenzen für die Bestandteile der Expressionskassetten sind im Sequenzprotokoll dargestellt oder allgemein bekannt:

Cytomegalovirus-Promoter (CMV-Promoter):

(M. Boshart et al., Cell 41 (1958) 521-530; F. Langle-Rouault et al., Virol. 72(7) 6181-5 (1998)) betaglobin-Intron-2:

(Jeffreys, A.J. et al., Cell 12 (1977) 1097-1108) SV40 poly A-Signal:

(M. Boshart et al., Cell 41 (1958) 521-530; F. Langle-Rouault et al., Virol. 72(7) 6181-5 (1998)) EF-lalpha-Promoter: SEQ ID NO: 9

(S. Mizushima, Nucleic Acids Res. 18 (1990) 5322, T. Uetsuki, J. Biol. Chem. 264 (1989) 5791-5798).

G-CSF poly A-Signal:

(S. Mizushima, Nucleic Acids Res. 18 (1990) 5322, T. Uetsuki, J. Biol. Chem. 264 (1989) 5791-5798).

gp (LCMV): vgl. SEQ ID NO: 1, 3, für SEQ ID NO: 4 kodierender Bereich (siehe auch Anlage zum Sequenzprotokoll).

[0025] Im Rahmen der vorliegenden Erfindung sind somit auch die o. g. Expressionskassetten eingeschlossen, wobei Änderungen in den jeweiligen Nukleinsäuresequenzen möglich sind, solange die Funktionalität der Expressionskassetten erhalten bleibt, d.h. deren erfindungsgemäße Verwendung die Pseudotypisierung der Verpackungszellen ermöglicht und auch die Transfektion der Zielzellen und die stabile Integration der Transgene in das Wirtsgenom nicht behindert.

[0026] Ferner zeigt auch ein episomaler EBV-Expressionsvektor (Epstein-Barr-Virus; vgl. F. Langle-Rouault et al., Virol. 72(7) 6181-5 (1998)) (pCep4) der Firma Invitrogen hohe Expression und kommt daher im Rahmen der vorliegenden Erfindung bevorzugt in Betracht.

[0027] Gegenstand der vorliegenden Erfindung ist somit ferner eine Verpackungszelle, die die retroviralen Gene gag (für SEQ ID NO: 12 kodierender Bereich; vgl. Anlage zum Sequenzprotokoll, ZU SEQ ID NO: 11), pol (für SEQ ID NO: 13 kodierender Bereich; vgl. Anlage zum Sequenzprotokoll, ZU SEQ ID NO: 11) und gegebenenfalls das retrovirale Gen env (für SEQ ID NO: 14 kodierender Bereich; vgl. Anlage zum Sequenzprotokoll, ZU SEQ ID NO: 11) und/oder regulatorische retrovirale Gene (im Falle lentiviraler Verpackungssysteme, s.u., z.B. das für das lentivirale Rev-Protein kodierende Gen, das ein Spleißen der retroviralen genomischen RNA verhindert) umfaßt und ferner das für die Glykoproteine GP-1 und GP-2 des LCMV kodierende Gen gp (für SEQ ID NO: 4 kodierender Bereich; vgl. Anlage zum Sequenzprotokoll, ZU SEQ ID NO: 3) oder einen Teil desselben umfaßt. Ferner eingeschlossen sind Nukleinsäuresequenzen, die Änderungen bzw. Abweichungen (Mutationen, Deletionen etc.) in den Sequenzen aufweisen (Derivate), solange bei der erfindungsgemäßen Verwendung die Pseudotypisierung der Verpackungszellen gewährleistet bleibt und auch die Transfektion der Zielzellen sowie die stabile Integration der Transgene in das Wirtsgenom nicht behindert wird. Dies schließt Fragmente der genannten Sequenzen ein. Im folgenden sollen diese Derivate stets mitumfaßt sein, wenn ein beliebiges Gen als solches erwähnt wird. Die erfindungsgemäße Verpackungszelle exprimiert pseudotypisierte Virionen, die LCMV-Glykoprotein in ihrer Hülle eingebaut enthalten.

[0028] Im Rahmen der vorliegenden Erfindung bezeichnet "GP" oder "GP-Protein" das GP-C-Vorläuferprotein, aus dem durch proteolytische Spaltung dann die GP-1 und GP-2 entstehen, die nachfolgend auch einfach als "LCMV-Glykoprotein" bezeichnet werden.

[0029] Erfindungsgemäß werden ferner Pseudotyp-Verpackungssysteme bereitgestellt, in denen neben dem gp-Genprodukt (SEQ ID NO: 4) ein oder mehrere weitere Gene von LCMV exprimiert werden, wie zum Beispiel das für das Nukleoprotein kodierende Gen np (für SEQ ID NO: 5 kodierender Bereich; vgl. Anlage zum Sequenzprotokoll, ZU SEQ ID NO: 3), das für ein Protein mit unbekannter Funktion kodierendes Gen z (für SEQ ID NO: 8 kodierender Bereich; vgl. Anlage zum Sequenzprotokoll, ZU SEQ ID NO: 6) und das für die RNA-Polymerase kodierende Gen 1 (für SEQ ID NO: 7 kodierender Bereich; vgl. Anlage zum Sequenzprotokoll, ZU SEQ ID NO: 6). Diese Gene können gemäß einer besonderen Ausführungsform der Erfindung entweder vom WE- oder Armstrong-Stamm des LCMV stammen.

[0030] In diesem Zusammenhang können entweder die vollständigen Sequenzen der Gene np, z und/oder 1 eingesetzt werden (SEQ ID NOs: s.o.) oder Teile derselben verwendet werden. Erfindungsgemäß eingeschlossen sind Nukleinsäuresequenzen, die Änderungen bzw. Abweichungen (Mutationen, Deletionen etc.) in den Sequenzen aufweisen (Derivate), solange die Pseudotypisierung der Verpackungszellen gewährleistet ist und

auch die Transfektion der Zielzellen und die stabile Integration der Transgene in das Wirtsgenom nicht behindert wird. Dies schließt Fragmente der genannten Sequenzen ein. Im folgenden sollen diese Derivate stets mitumfaßt sein, wenn ein beliebiges Gen als solches erwähnt wird.

[0031] Gegenstand der vorliegenden Erfindung ist daher ferner eine Verpackungszelle, die zusätzlich zu dem gp-Gen des LCMV mindestens ein Gen für das Nukleoprotein kodierendes Gen np, für die RNA-Polymerase kodierendes Gen 1 und für ein Protein mit unbekannter Funktion kodierendes Gen z des LCMV enthält.

[0032] Von den erfindungsgemäßen Verpackungszellen werden die MLV/LCMV-Pseudotypen produziert, d.h. rekombinante retrovirale Virionen, die in ihrer Hülle das LCMV-Glykoprotein eingebaut enthalten.

[0033] Für die Herstellung der rekombinanten Virionen geht man im Rahmen der vorliegenden Erfindung für die viralen Verpackungszelllinien vorzugsweise von allen Zelllinien aus, die hohe Titer retroviraler Vektoren produzieren. Bevorzugt eingesetzte Zelllinien sind NIH3T3, Te671, 293T, HT1080 (F.L. Cosset et al., J. Virol. 69 (1995) 7430-7436; D. Markowitz et al., Virology 167 (1988) 400-406; W.S.Pear et al., PNRS 90 (1993) 8392-8396). Die Wahl der Zelllinie ist für die spezifischen Vorteile der Erfindung jedoch unerheblich, da es sich gezeigt hat, daß GP in keiner bislang untersuchten Zelllinie toxisch wirkt. Sollten daher zukünftig Linien gefunden werden, die eine effizientere Vektorproduktion erlauben (d.h. stabiler Titer, der mindestens so hoch ist wie in den zuvor genannten Linien, > 10⁶/ml), können auch diese eingesetzt werden.

[0034] In den zur erfindungsgemäßen Pseudotypisierung vorzugsweise verwendeten Verpackungssystemen werden die gag- und pol-Gene des Moloney Stammes muriner Leukämieviren (MoMLV) exprimiert (gag: für SEQ ID NO: 12 kodierender Bereich; pol: SEQ ID NO: 11, Nukleotide 1970-5573; vgl. Anlage zum Sequenzprotokoll, ZU SEQ ID NO: 11). Erfindungsgemäß sind aber auch andere gag- und pol-Varianten von MLV eingeschlossen, solange sie die o.g. Vorteile für die Vektorproduktion aufweisen. Insbesondere sind erfindungsgemäß die o.g. Gen-Derivate eingeschlossen.

[0035] Im Rahmen der vorliegenden Erfindung wurde festgestellt, daß LCMV-GP auch lentivirale Nucleocapside pseudotypisiert (s. Beispiele). Gemäß einer besonderen Ausführungsform können die Verpackungssysteme daher auch die gag- und pol-Genprodukte von Lentiviren enthalten, d.h., es können lentivirale Verpackungssysteme (Verpackungszelllinien) eingesetzt werden. Dabei ist es unerheblich, von welchem Lentivirus sich das Verpackungssystem ableitet. Als lentivirale Verpackungszellen kommen im Rahmen der vorliegenden Erfindung beispielsweise von Humanem Immundefizienzvirus (HIV), Affen-Immundefizienzvirus (SIV) oder Feline Immundefizienzvirus (FIV) abgeleitete Zelllinien in Betracht. In diesem Zusammenhang könnte es für eine effiziente Produktion infektiöser Lentivirusvektoren erforderlich sein, zusätzlich akzessorische lentivirale Gene wie rev (für SEQ ID NO: 21 kodierender Bereich; vgl. Anlage zum Sequenzprotokoll, ZU SEQ ID NO: 15) oder tat (für SEQ ID NO: 20 kodierender Bereich; vgl. Anlage zum Sequenzprotokoll, ZU SEQ ID NO: 15) bei HIV-Vektoren zu exprimieren. Im Rahmen der vorliegenden Erfindung können LCMV-Proteine in allen lentiviralen Verpackungssystemen zur Pseudotypisierung eingesetzt werden.

[0036] Gegenstand der vorliegenden Erfindung ist somit ferner eine (virale) Verpackungszelle bzw. ein pseudotypisiertes Virion, bei dem das Virus aus der Familie der Retroviridae ist, insbesondere ein MLV-related virus oder Lentivirus ist. Gemäß einer besonderen Ausführungsform der Erfindung ist die retrovirale Verpackungszelle eine Verpackungszelle, die sich von MLV, HIV, SIV, und FIV ableitet.

[0037] Erfindungsgemäß wird ferner ein pseudotypisiertes Virion zur Verfügung gestellt, das durch Pseudotypisierung einer retroviralen Zelle von MLV, die kein ENV-Protein exprimiert, erhältlich ist, wobei zur Pseudotypisierung die defekte Mutante L(ARM) von LCMV verwendet wird. Alternativ können auch andere Varianten von LCMV eingesetzt werden.

[0038] Gegenstand der vorliegenden Erfindung ist ferner ein Verfahren zur Herstellung der erfindungsgemäßen Verpackungszellen/Verpackungszelllinien, bei dem man eine mit einem oder mehreren Fremdgenen transfizierte retrovirale Verpackungszelllinie nach bekannten Methoden (vgl. z.B. Maniatis, Sambrook, Fritsch: Molecular cloning, a laboratory manual; Cold Spring Harbor Laboratory Press, 1982) mit LCMV infiziert. Die Virionen, die LCMV-Glykoprotein in ihrer Hülle eingebaut enthalten, werden von den Verpackungslinien exprimiert. Gemäß einer besonderen Ausführungsform der Erfindung können die Pseudotyp-produzierenden Verpackungszellen hergestellt werden, indem man eine mit einem oder mehreren Fremdgenen transfizierte retrovirale Verpackungszelllinie mit einem Expressionsplasmid transfiziert, das das gp-Gen des LCMV oder einen Teil desselben und gegebenenfalls zusätzlich ein oder mehrere Gene aus der Gruppe bestehend aus np, 1 und z des LCMV enthält.

[0039] Die vorliegende Erfindung betrifft somit ferner ein Verfahren zur Herstellung retroviraler Pseudotyp-Vektoren, bei dem man zunächst retrovirale Verpackungszellen herstellt, die mindestens das gp-Gen des LCMV oder einen Teil desselben enthalten (s.o.), und diese anschließend unter Bedingungen kultiviert, die zur Produktion von Virionen geeignet sind.

[0040] Im Rahmen der vorliegenden Erfindung verwendet man als retrovirale Verpackungszelllinie vorzugsweise eine MLV-Verpackungszelllinie, die kein funktionelles ENV-Protein exprimiert, und als LCMV zur Pseudotypisierung die deletierte Mutante L(ARM). Als ENV-negative Verpackungszelllinie kann auch die von von Laer et al. (J. Virol. 72 (1998) 1424-1430) beschriebene Zelllinie verwendet werden.

[0041] Die erfindungsgemäßen Virionen bzw. Verpackungszelllinien lassen sich in vorteilhafter Weise zur Erzeugung viraler Pseudotypvektoren verwenden, die man in vorteilhafter Weise zur Transduktion von Zelllinien aber auch von primären Eukaryonten-Zellen zu Forschungszwecken (in vitro) oder im Rahmen der Gentherapie einsetzen kann.

[0042] Im Rahmen der vorliegenden Erfindung kommt jedoch auch die Verwendung der Virionen/Verpackungszellen zur Gentherapie in Betracht. Die Gentherapie kann in diesem Zusammenhang die Therapie von von infektiösen Erkrankungen wie AIDS und Neoplasien wie dem Mamma-Karzinom oder dem Melanom sowie anderen durch Gentherapie zugänglichen Erkrankungen umfassen. Im Rahmen der vorliegenden Erfindung kommt die Transduktion hämatopoietischer Stammzellen in Betracht.

[0043] Das rekombinante Virion bzw. die rekombinante Verpackungszelle (-zelllinie) der vorliegenden Erfindung umfaßt ein oder mehrere Transgene. Vorzugsweise sind diese Transgene Markergene, wie z.B. neo, lacZ oder enhanced green fluorescent protein (EGFP) und/oder therapeutisch einsetzbare Gene, wie z.B. die Suidigene Herpes Simplex Virus Thymidin Kinase (HSV-tk), Cytosin Deaminase (CD), und antiviral wirksame Sequenzen wie Ribozyme, antisense Sequenzen und transdominant-negativ wirkende Gene und in der Tumorthherapie einsetzbare Gene, wie das mdr-1 für den Schutz der hämatopoetischen Zellen bei Chemotherapie und Zytokin-Gene. Darüber hinaus können jedoch alle Transgene eingesetzt werden, die im Rahmen eines gezielten Gentransfers und der Expression des Transgens (der Transgene) in Zellen in vitro oder im Rahmen der Gentherapie von Interesse sein könnten.

[0044] Erfindungsgemäß eingeschlossen ist ferner ein Verfahren zur Herstellung eines pharmazeutischen Präparats zur Gentherapie, bei dem man erfindungsgemäße virale Pseudotypvektoren bzw. retrovirale Verpackungszellen gegebenenfalls mit pharmazeutisch verträglichen Hilfs- und/oder Trägerstoffen formuliert. Gegenstand der vorliegenden Erfindung ist ferner ein pharmazeutisches Präparat zur Gentherapie, das erfindungsgemäße retrovirale Verpackungszellen und gegebenenfalls pharmazeutisch verträgliche Hilfs- und/oder Trägerstoffe umfaßt.

[0045] Mit der vorliegenden Erfindung werden erstmals Vektorsysteme zur Verfügung gestellt, die in hohen Titern produziert werden und auf konzentriert werden können. Die erfindungsgemäßen Vektorpartikel lassen sich ferner ohne oder ohne wesentlichen Verlust der Infektiosität auf reinigen. Die erfindungsgemäßen Zelllinien zeichnen sich ferner durch ein breites, speziesübergreifendes Wirtszellspektrum (Zelltropismus) aus. Es hat sich im Rahmen der vorliegenden Erfindung überraschenderweise gezeigt, daß die erfindungsgemäße Pseudotypisierung für die Verpackungszellen nicht zelltoxisch ist. Es werden somit erstmals stabile Verpackungszelllinien (Verpackungssysteme) zur Verfügung gestellt, die einen stabilen, retroviralen Transfer von Transgenen in die Zielzellen ermöglichen, d.h., zu einer stabilen Integration des Transgens in das Genom der Ziel- oder Wirtszellen und einer anschließenden stabilen Expression dieses Gens führen.

[0046] Die vorliegende Erfindung wird nachfolgend anhand von Beispielen, Figuren und einem Sequenzprotokoll erläutert.

Ausführungsbeispiel

Materialien und Methoden

Zellen und Viren

[0047] Die env-negative Verpackungszelllinie TELCeB wurde von F.-L. Cosset (F.L. Cosset et al. J Virol. 69 (1995) 7430-7436) zur Verfügung gestellt. Die env-negative Zelllinie 293gp2 wurde bereits früher beschrieben (D. von Laer et al., J. Virol. 72 (1997) 1424-1430). Die Mausfibroblastenzelllinie Sc-1 wurde in Minimal Essential

Medium (Sigma, Deisenhofen, Deutschland), das mit 10% fetalem Kälberserum (FCS, PAN Systems, Aidenbach, Deutschland) angereichert worden war, kultiviert. Die humane Nierenzelllinie 293, die humane Hepatoma Linie HUH-7, die humane Fibroblastenzelllinie Te671 und TELCeB und die Mausfibroblastenzelllinie L-929 wurden in Dulbecco's Minimal Essential Medium (DMEM, Gibco, Paisley, Great Britain) kultiviert, das mit 10%-igem FCS angereichert worden war. Die humane hematopoietische Vorläuferzelllinie TF-1 wurde in Iscove's Modified Dulbecco's Medium (Gibco, Paisley, Großbritannien) gehalten, das mit 10%-igem FCS und IL-3 angereichert worden war. Konditioniertes Medium von NIH3T3-Zellen, die mit einem BPV-Vektor transfektiert worden waren, der das IL-3 Gen trägt, wurden als Quelle für IL-3 bei Konzentrationen verwendet, die für das maximale Wachstum von TF-1 notwendig sind (H. Karasuyama et al., Eur. J. Immunol. 18 (1988) 97-104). Die humane Vorläuferzelllinie K562 wurde in RPMI (Gibco) gehalten, das mit 10%-igem FCS angereichert worden war.

[0048] LCMV wurde am 10.11.1998 bei der European Collection of Cell Cultures (ECACC), Salisbury, Wiltshire SP4 OJG, Großbritannien, unter der Zugriffsnummer V98111005 nach dem Budapester Vertrag hinterlegt.

[0049] Der MESV-artige retrovirale Vektor, der das Neomycin phosphotransferase Gen (MP1N) trägt, wurde früher beschrieben (H.-G. Eckert, Blood 88 (1996) 3407-3415). Der amphotrophe Helfer war ein rekombinantes replikationskompetentes Moloney-MLV, bei dem Teile des pol- und der größte Teil des env-Genes mit dem von dem MLV Stamm 4070A [Mo-Ampho-MP, R320 (C. Münk, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 5837-5842)] als ein Sall- bis ClaI-Fragment ersetzt wurden. Der Virus wurde in Sc-1 propagiert. Der Plaquegereinigte WE Stamm des LCM Virus wurde in L-929 Zellen propagiert (T.M. Rivers, Virology 26 (1965) 270-282).

Durchflußzytometrische Analyse der LCMV-GP-Expression

[0050] Zur Analyse der Expression des LCMV Glykoproteins wurden 3×10^5 bis 10^6 Zellen geerntet, pelletiert und in 50µl einer 1:40 Verdünnung von Maus-Ascites resuspendiert, die einen murinen monoklonalen Antikörper gegen LCMV GP-1 enthielt (M. Bruns et al., Virology 130 (1983) 247-251). Nach 20-minütiger Inkubation auf Eis wurden die Zellen dreimal mit phosphatgepufferter Kochsalzlösung (PBS) gewaschen und dann für weitere 20 Minuten in einer 1:80 Verdünnung eines FITC-markierten Ziege Anti-Maus Antikörpers inkubiert (Dako, Glostrup, Dänemark). Nach drei abschließenden Waschschritten in PBS wurden die Zellen mittels eines FACScalibur Geräts analysiert (Becton Dickinson, Heidelberg).

Titration der Viren

[0051] Zur Bestimmung des Vektortiters wurden 5×10^4 Sc-1 Zellen mit einer fünffachen Verdünnung der Überstände in 24-Vertiefungen Gewebekulturplatten inokuliert. Für den retroviralen Neovektor wurde die Selektion nach 24 Stunden mit 400µg G418 pro ml (Trockengewicht GIBCO) initiiert. Das Medium wurde alle vier Tage ersetzt. Die Kolonien wurden nach zehn Tagen ausgewertet. Der Titer wurde ausgedrückt als G418 Resistenztransfereinheiten pro ml (GTU/ml). Für den retroviralen MFGnlsLacZ Vektor wurde x-Gal (5-Bromo-4-chloro-3-indolyl-β-D-galaktopyranosid) Färbung zwei Tage nach der Inokulation wie früher beschrieben (G.R. McGregor, Methods Mol. Biol. 7 (1989) 1-19) durchgeführt. Der Titer wurde ausgedrückt in LacZ Transfereinheiten (LTU) pro ml. Plaque-bildende Einheiten von LCMV wurden an L-929 Zellen wie früher beschrieben getestet (F. Lehmann-Grube, J. Gen. Virol. 37 (1977) 85-92).

[0052] MLV (LCMV) Pseudotypen wurden durch Preinkubation eines gleichen Volumens eines Virus-positiven Überstandes mit einem Anti-LCMV gp44 neutralisierendem monoklonalen Antikörper neutralisiert, der 1:100 verdünnt wurde (M. Bruns et al., Virology 130 (1983) 247-251). Die Titer wurden dann in LTU pro ml wie beschrieben bestimmt.

DNA Analyse

[0053] Herstellung von DNA und Southern Analyse wurden durchgeführt wie früher beschrieben (C. Stocking et al., Cell 53 (1988) 869-879). Die genomische DNA wurde mit HindIII verdaut, das einmal an dem 3' Ende des neo-Gens in dem MP1N-Vektor schneidet. Ein Fragment, das das vollständige neo-Gen enthielt, wurde als Sonde verwendet.

Herstellung und Reinigung des Virus

[0054] Die Virionen wurden durch Gradienten Ultrazentrifugation wie im Detail früher beschrieben gereinigt

(L. Martinez-Peralta et al., J. Gen. Virol. 55 (1981) 475-479). Kurz gesagt wurden infektiöse Zellkulturüberstände durch Zentrifugation bei niedrigen und hohen Geschwindigkeiten gereinigt. Das Virus wurde durch Ultrazentrifugation pelletiert und dann in einem 0-40 %-igen Urografin Gradienten (Schering AG, Berlin, Deutschland) gereinigt.

Beispiel 1

Infektion von TeLCeB mit LCMV-LacZ-Gentransfer auf Zielzellen mit Neutralisation des Vektors durch anti-GP mAb und Konzentrierung des Vektors im Gradienten

[0055] Wiederherstellung (rescue) eines Hüllprotein-negativen Murine Leukemia Virus Vector mit Lymphocytic Choriomeningitis Virus Um zu testen, ob LCMV einen Hüllprotein-negativen retroviralen Vektor wiederherstellen (mobilisieren, komplettieren) kann, wurde die env-negative Verpackungszelllinie TELCeB mit dem LCMV WE Stamm bei einer m.o.i. von 0.01 infiziert (m.o.i.: Infektionsmultiplizität; bezeichnet die Anzahl von Viren-Partikeln, mit denen eine Zelle infiziert wird). TELCeB stammen von der Humanfibroblastenzelllinie Te671 ab und enthalten gag- und pol-Gene sowie den retroviralen Vektor MFGn5LacZ (G.M. Crooks et al., Blood 82 (1993) 3290-3297). Nach Infektion mit LCMV wurde der Titer von LacZ transferring units (LTU) und von LCMV Wildtypvirus durch X-Gal-Färben der Mausfibroblasten-Zielzellen (Sc-1) und durch einen Plaqueversuch (in plaquebildenden Einheiten, PFU) gemessen. Zusätzlich wurde die Expression von LCMV-Glykoproteinen in den infizierten TELCeB durch durchflußzytometrische Analyse gemessen. Die Resultate sind in [Fig. 6](#) gezeigt. LTU wurden sechs Tage lang hergestellt, mit einem Maximum von 5×10^4 LTU pro ml an Tag 3. Der höchste Titer für LCMV Wildtypvirus betrug 3×10^8 an Tag 2. Die Produktion von PFU hatte bereits an Tag 3 abgenommen, als die maximale Produktion von LTU zusammen mit der Höchstexpression von LCMV Glykoprotein zu sehen war. Diese Diskrepanz könnte in der Herstellung von defekten interferierenden LCMV Partikeln begründet sein, die die Replikation von LCMV Wildtyp inhibieren, mit größter Wahrscheinlichkeit aber nicht die Ausschüttung von infektiösen retroviralen Vektorpartikeln. Während der Replikation von LCMV in den Verpackungszelllinien war kein offensichtlicher zytopathischer Effekt zu beobachten, obwohl hohe Spiegel an LCMV Glykoproteinen expremiert wurden (Daten nicht gezeigt).

[0056] Es wurde dann getestet, ob das infektiöse Virus, das durch die LCMV-infizierten TELCeB hergestellt wurde, den LacZ-Gentransfer spezifisch durch LCMV Glykoproteine vermittelte. Die Überstände wurden mit einem neutralisierenden Anti-LCMV gp44 monoklonalem Antikörper eine Stunde lang inkubiert. Dies führte zu einer mehr als dreifachen Log Reduzierung des LTU-Titers. Der amphotrophe Pseudotyp desselben retroviralen Vektors wurde nicht durch den anti-LCMV Antikörper neutralisiert (Tab. 1). Diese Daten zeigen, daß die MLV/LCMV chimären Virionen tatsächlich LCMV Glykoproteine auf ihrer Oberfläche trugen, die den Gentransfer in Abwesenheit von retroviralen Hüllproteinen durch den LCMV Rezeptor vermitteln können.

Tab. 1: Infektiöse retrovirale Vektor-Partikel, die von LCMV-infizierten Verpackungszelllinien produziert werden, lassen sich durch anti-LCMV-Glykoprotein monoklonale Antikörper (mAb) neutralisieren.

Virus	anti-LCMV mAb	Titer (LTU/ml)
Amphotroper Helfer	nein	$2 \cdot 10^5$
Amphotroper Helfer	ja	$2 \cdot 10^5$
LCMV	nein	$7 \cdot 10^4$
LCMV	ja	$3 \cdot 10^1$

[0057] MLV (LCMV) Pseudotypen behalten Infektiosität bei Konzentrierung mittels Ultrazentrifugation durch einen Gradienten: Amphotrophe Retroviren verlieren Infektivität nach Ultrazentrifugation, höchstwahrscheinlich aufgrund der Labilität der retroviralen Hüllglykoproteine (V. Moening et al., Virology 61 (1974) 100-111). Es wurde untersucht, ob die MLV(LCMV)-Pseudotypen stabiler sind. TELCeB wurden mit LCMV oder mit amphotropem Helfervirus infiziert. Titer des viralen Vektors im direkten Überstand betragen 5×10^4 LTU pro ml für beide Pseudotypen. Viren wurden aus 60 ml Überstand pelletiert und durch Ultrazentrifugation über einen 0-40%-igen Urografin-Gradienten gereinigt. Die Pseudotypentiter (in LTU pro ml) sind in [Fig. 7](#) gezeigt. Der gesamte erwartete Ertrag von 3×10^6 LTU für den LCMV Pseudotyp wurde vollständig erhalten, im Gegensatz

zu 1×10^3 LTU für den amphotropen Virus. Die reverse Transkriptase-Aktivität in den Banden zeigte, daß die Menge an Viruspartikeln, die aus dem Gradienten gewonnen wurde, für beide Pseudotypen ähnlich war (Daten nicht gezeigt). Verglichen mit den amphotropen Virionen war die Infektivität von MLV(LCMV)-Pseudotypen bei Ultrazentrifugation allerdings mindestens um einen Faktor 1000 stabiler.

[0058] LCMV Pseudotypen waren auch bei Lagerung bei 4°C stabil. Innerhalb des Beobachtungszeitraumes von drei Tagen war der Titerverlust zweifach niedriger als (im Vergleich zum Ausgangstiter des MLV(LCMV)). Ein Tiefrierzyklus (-80°C) und Auftauen führte zu einem Verlust an Pseudotypentiter, der zweifach niedriger lag.

Beispiel 2

Gag- und pol-Genprodukte sind erforderlich für die Verpackung von retroviralen RNA in die LCMV Glykoprotein Pseudotypen

[0059] Es wurde untersucht, ob die retrovirale RNA alleine in den LCMV verpackt werden könnte oder ob gag und pol Genprodukte erforderlich waren. 293-Zellen und 293gp2-Zellen, die letzteren enthielten gag und pol von MLV, wurden mit einem auf MLV-basierenden retroviralen Vektor transfiziert, der das neo-Gen (MP1N) enthielt, und Zelllinien, die den stabil integrierten Vektor enthielten, wurden durch G418-Selektion hergestellt (293MP1N und 293gp2MP1N; ein als SF23 bezeichneter Klon der Zelllinie 293gp2MP1N wurde am 05.11.1998 bei der Deutschen Sammlung von Mikroorganismen und Zellkulturen (DSMZ), Mascheroder Weg 1b, 38124 Braunschweig, Deutschland, unter der Zugriffsnummer DSM ACC2374 nach dem Budapester Vertrag hinterlegt). Diese Zellen (Massenkulturen) wurden dann entweder mit einem replikationskompetenten amphotropen Helfer oder mit dem LCMV-Wildtypvirus infiziert. Die Resultate sind in Tab. 2 gezeigt. Infektioser Vektor, der Neomycin-Resistenz transferierte, wurde von beiden Zelllinien nach Infektion mit dem amphotropen Helfer gewonnen. Nach Infektion mit LCMV allerdings produzierte nur 293gpMP1N infektiöse retrovirale Partikel, 293MP1N, die kein retrovirales Gag oder Pol exprimierten, hingegen nicht. Retrovirale genomische RNA wurde daher in Abwesenheit von gag- und pol-Genprodukten durch LCMV nicht in infektiöse Virionen verpackt.

Tab. 2: Gag- und/oder pol-Genprodukte sind für die Wiederherstellung (rescue) eines retroviralen Vektors durch LCMV essentiell

Zelllinie	freigesetzter Vektortiter* nach Infektion mit		
	LCMV	amphotrop. Helfer	(Kontrolle)
293MP1N	0	$1 \cdot 10^4$	0
293gp2MP1N	$2 \cdot 10^3$	$6 \cdot 10^4$	0

*Vektortiter werden in G418-resistenten Zellkolonien exprimiert, die man nach Inkulation von Sc-1 mit den viralen Überständen erhielt (G418 Transfereinheiten/ml)

Beispiel 3

Infektion von 293gpMP1N mit LCMV und stabiler Gentransfer auf L929 – Nachweis durch Southern-Blot

[0060] MLV (LCMV) Pseudotypen vermitteln Transfer und stabile Integration des retroviralen Vektor Genoms: Der Transfer von G418-Resistenz durch den retroviralen LCMV-Pseudotyp zeigte, daß das Markergen in das Wirtsgenom stabil integriert worden war. Um zu verifizieren, daß MLV(LCMV)-Pseudotypen stabile Transduktion mit Integration des Transgens in das Zielzellengenom vermitteln können, wurde ein retroviraler Vektor, der das Neomycin-Resistenzgen (neo) enthielt, durch LCMV-Infektion der env-negativen Verpackungszelllinie 293gp2MP1N wiederhergestellt. Die Titer wurden durch Transfer von G418-Resistenz auf Sc-1 Zellen gemessen und lagen zwischen 1×10^3 und 1×10^4 G418-Transfereinheiten (GTU) pro ml. Resistente Zellklone tauchten nach acht Tagen Selektion auf und wurden für weitere drei Wochen kultiviert. Die DNA von 12 G418-resis-

tenten Klonen wurde einer Southern Blot Analyse nach Restriktion mit HindIII unterworfen, einem Enzym, daß nur einmal schneidet, unter Verwendung einer Neo-Sonde. In 10 Klonen wurde eine Kopie des integrierten retroviralen Vektor-Genoms pro Zelle detektiert, und zwei Kopien in den übrigen zwei Klonen (Daten nicht gezeigt). Transduktion mit dem MLV(LCMV) Pseudotyp führte daher zu stabiler Integration des Transgens.

Beispiel 4

Expression von LCMV-Glykoprotein (LCMV-GP) in TeLCeB-L(Arm)

Material und Methoden

[0061] Die Anlage der env-negativen Verpackungslinie TeLCeb, die gag und pol von MLV sowie ein retrovirales Vektorgenom mit LacZ als Transgen enthält, wurde bereits ausführlich beschrieben (F.L).

[0062] Cosset et al., J. Virol. 69 (1995) 7430-7436). Die Titration der Vektorüberstände wurde wie bereits beschrieben auf 293-Zellen durch X-Gal Färbung vorgenommen (G.R. McGregor, Methods Mol. Biol. 7 (1989) 1-19). Die Zellen wurden in DMEM mit 10 % FCS kultiviert. Der L(Arm)-Stamm von LCMV entsteht nach mehreren Passagen von LCMV in L929-Zellen (M. Bruns et al., Virology 177 (1990) 615-624). LCMV-Nukleoprotein (LCMV-NP) von L(Arm) wurde durch Immunfluoreszenzfärbung der Zellen auf Objektträgern mit einem polyklonalen anti-LCMV Kaninchenserum nachgewiesen. Diese Standardmethode wurde bereits zuvor detailliert beschrieben (M. Bruns et al., Virology 177 (1990) 615-624). Zur Expression von LCMV-GP, wurde das gp-Gen in den episomalen EBV-Vektor pCep4 (Invitrogen), der ein Hygromycin-Resistenzgen trägt, kloniert.

Ergebnisse

[0063] Bei den Experimenten zur Pseudotypisierung durch die alleinige Expression von LCMV-GP in env-negativen retroviralen Verpackungslinien wurde mit den Expressionsplasmiden eine höhere GP-mRNA-Expression erzielt als bei LCM-Wildvirus-Infektion ([Fig. 4](#)). Dieses Ergebnis zeigt, daß das gleichzeitige Vorliegen noch mindestens eines weiteren LCMV-Genprodukts neben dem LCMV-Glykoprotein eine Steigerung der Glykoprotein-Produktion bewirkt bzw. die Bildung von Pseudotypen fördert. Um diese Schlußfolgerung direkt zu untermauern, wurde das ektop (von einem Plasmid) exprimierte LCMV-GP mit den LCMV Proteinen des L (ARM) Stammes von LCMV komplementiert. Diesem defekten Stamm fehlt das funktionelle Glykoprotein, und er bildet daher keine Plaques, ist nicht pathogen für Mäuse und breitet sich nur über mehrere Wochen innerhalb einer Zellkultur aus (LCM-Wildvirus dagegen innerhalb von 24 Stunden). Alle weiteren Genprodukte von L(ARM) (NP, L und Z) weisen keine nachweisbaren Defekte auf.

[0064] TeLCeb wurden mit L(Arm)-haltigem Zellkulturüberstand infiziert und anschließend 5 Wochen passagiert. Dieses ist erfahrungsgemäß der Zeitraum, den das defekte Virus benötigt, um alle Zellen einer Kultur zu infizieren. Die vollständige Infektion aller Zellen wurde durch Immunfluoreszenz-Färbung mit einem anti-LCMV Serum nachgewiesen. TeLCeb-L(Arm) wurden mit pCep-GP durch Elektroporation (Elektroporator der Firma Dr. Fischer, Heidelberg) transfiziert und 2 Wochen mit Hygromycin selektioniert. Als Kontrolle wurden Zellen mit pCep4 (ohne GP-Gen) transfiziert. In TeLCeb-L(Arm), die mit pCep-GP transfiziert waren, kam es nach der Selektion zur Produktion von Pseudotypen, die lacZ auf 293-Zellen übertrugen. Der Titer lag zwischen 10^2 und 10^3 /ml. Dieses Ergebnis zeigt eindeutig, daß das LCMV-GP im beschriebenen Expressionsplasmid funktionell war und retrovirale Vektoren pseudotypisieren kann.

Beispiel 5

Pseudotypisierung eines HIV-Vektors, der das green fluorescent protein (GFP) exprimiert

Material und Methoden

[0065] Der Lentivirale Vektor HIV-GFP leitet sich von dem infektiösen DNA-Klon pNL4-3 von HIV ab und ist bereits detailliert beschrieben ([Abb. 5](#)) (R.I. Connor et al., Virology 206 (1995) 935-944). Am Anfang von env wurde die NdeI-Schnittstelle aufgefüllt und religiert, wodurch sich das Leseraster verschiebt und kein funktionelles Hüllprotein synthetisiert wird. An Stelle von nev wurde außerdem das Gen für das green fluorescent protein (GFP) kloniert. Der Titer des verwendeten LCMV-WE-Stammes wurde durch einen Plaque Assay auf L929 bestimmt, der bereits detailliert beschrieben ist (F. Lehmann-Grube et al., J. Gen. Virol. 37 (1977) 85-92). Die Calciumphosphat-Transfektionen wurden auf 293 mit einem Standard-Protokoll durchgeführt (Maniatis, et al.: Molecular cloning, a laboratory manual; Cold Spring Harbor Laboratory Press, 1982).

Ergebnisse

[0066] 293 Zellen wurden mit einer m.o.i. von 0.1 des LCMV-WE-Stammes infiziert. Nach einer Stunde erfolgte die Transfektion mit HIV-GFP und nach zwei Tagen wurden die Überstände geerntet und auf 293 Zellen übertragen. Der Titer wurde durch die GFP-Expression in den 293-Zielzellen mittels Immunfluoreszenz ermittelt und lag zwischen 10^2 und 10^3 pro ml. Nach Transfektion mit HIV-GFP allein (ohne vorherige Infektion mit LCMV) kam es wie erwartet nicht zur Produktion von infektiösen Vektorpartikeln.

Beispiel 6

Untersuchung des Zelltropismus

[0067] MLV(LCMV) Pseudotypen infizieren verschiedene humane Zelllinien: Der Tropismus von MLV(LCMV) Pseudotypen wurde analysiert. Verschiedene humane Zelllinien, die von Zellen abstammten, die attraktive Ziele für Gentherapie sind, wie hämatopoetische Vorläuferzellen und Hepatozyten, wurden analysiert. Die Transfereffizienz bezogen auf Mausfibroblasten ist in Tab. 3 gezeigt. Alle analysierten Zelllinien waren für MLV(LCMV)-Pseudotypen empfänglich. Auch Hamsterzellen, die normalerweise resistent gegenüber einer Transduktion mit von MLV abstammenden Vektoren sind, konnten mit den MLV(LCMV)-Pseudotypen effizient transduziert werden.

Tab. 3: Wirtsspektrum von MLV(LCMV)-Pseudotypen

Zelllinie	Herkunft	Transduktions-effizienz
293	Epithel, Mensch	+ + +
K-562	myeloide Progenitorzellen, Mensch	+ + +
TF-1	myeloide Progenitorzellen, Mensch	+
HUH-1	Hepatom, Mensch	+ +
Jurkat	Lymphozyt, Mensch	+ +*
Sc-1	Fibroblast, Maus	+ + +
CHO	Epithel, Hamster	+ + +
Cf2Th	Thymus stroma, Hund	+ + +

*Die Pseudotypen wurden mit auf Lymphozyten passagiertem (adaptiertem) LCMV hergestellt. Pseudotypen aus Fibroblasten-passagiertem LCMV transduzieren Lymphozyten nicht.

FIGUREN:

[0068] **Fig. 1:** Retrovirale Verpackungslinien Die retroviralen Gene gag, pol und env sind stabil in das Genom der retroviralen Verpackungslinien integriert. Env wird in der Regel getrennt von gag und pol exprimiert. Außerdem wird das Vektorgenom mit dem Transgen und regulatorischen cis-Elementen exprimiert (1). Die genomische RNA des Vektor enthält ein Verpackungssignal, das auf den gag, pol und env Genen deletiert wurde. Im Zytoplasma verpacken die retroviralen Proteine daher selektiv das Vektorgenom (2). Ein komplettes retrovira-

les Nukleokapsid bildet sich (3), und wird durch Sprossung an der Zellmembran freigesetzt (4). Hierbei wird retrovirales Hüllprotein, bzw. bei den Pseudotypen ein fremdes virales Hüllprotein, mitgenommen. Außerhalb der Zelle reift das Virus durch proteolytische Spaltung der retroviralen Vorläuferproteine im Nukleokapsid.

[0069] **Fig. 2:** Die genomische RNA des LCM-Virus Das Genom von LCMV besteht aus 2 ambisense RNA Molekülen mit je zwei offenen Leserahmen. Die 4 Gene und ihre Orientierung sind angegeben.

[0070] **Fig. 3:** Die Herstellung von MLV(LCMV) Pseudotypen. Verpackungslinien, die keine viralen Hüllproteine produzieren, setzen auch keine infektiösen retroviralen Vektoren frei. Wird eine solche env-negative Verpackungslinie mit LCM-Virus infiziert, so können die Retroviren das LCMV-Glykoprotein in ihre Hülle einbauen. Es entstehen sogenannte Pseudotypen. Außerdem wird während der LCMV-Replikation auch LCM-Wildvirus freigesetzt.

[0071] **Fig. 4:** Diskrepanz von hoher mRNA und niedriger Proteinkonzentration bei ektopter GP-C-Expression. 293-Zellen wurden mit unterschiedlichen Expressionskonstrukten für LCMV Glykoprotein transfiziert (Ef-1 alpha Promoter, Alphavirus-Vektor von Invitrogen, CMV-Promoter+Beta-Globin-Intron, s.o.: bevorzugte Expressionskassetten). Nach 2 Tagen wurden parallel die gp-mRNA Konzentration im Northern Blot und die GP-1/GP-2 Protein-Mengen auf den Zellen mittels Durchflußzytometrie gemessen. Als Kontrolle dienten 293-Zellen, die mit LCMV Wild-Typ infiziert waren. Bei der Wildvirusinfektion lag die GP-1/GP-2 Protein-Expression 2-3 Logstufen über der Negativkontrolle, bei einer relativ schwachen mRNA-Bande im Northern Blot (obere Bande LCMV-genomische RNA, untere Bande gp-mRNA). Bei ektopter Expression lag trotz höherer mRNA-Konzentrationen als bei der Wildvirus-Infektion die Proteinkonzentration für GP-1/GP-2 nur maximal eine Log-Sufe über der Negativkontrolle.

[0072] **Fig. 5:** HIV-GFP

[0073] Der infektiöse DNA-Klon von HIV pNL4-3 wurde am Anfang des env Genes mit NdeI geschnitten, die Schnittstelle aufgefüllt und religiert (s). Hierdurch verschiebt sich der Leserahmen und es entsteht kein funktionelles env Genprodukt. Zwischen XhoI und BamHI wurde anstelle von nef das Gen für das "green fluorescent protein" kloniert (GFP).

[0074] **Fig. 6:** Wiederherstellung (rescue) des retroviralen Vektors MFGInsLacZ durch LCMV.

[0075] Die retrovirale env-negative Verpackungszelllinie TELCeB wurde durch LCMV bei einer m.o.i. von 0.01 infiziert. Zwischen Tag 1 und 7 nach Infektion wurden die Überstände täglich ersetzt, und es wurden die Titer des LCMV-Wildtyps und des LacZ-Vektors durch einen Plaque-Assay auf L-929 bzw. durch LacZ-Gentransfer auf Sc-1 ermittelt. Zusätzlich wurde ein Teil der LCMV-infizierten TELCeB-Zellen täglich mit einem gegen das LCMV-Glykoprotein GP-1 gerichteten monoklonalen Antikörper gefärbt und mittels Durchflußzytometrie analysiert. Es ist die mittlere Fluoreszenz gezeigt. O-O LCMV-Wildtyp-Titer; □-□ LacZ-Transfereinheiten; Δ-Δ mittlere Fluoreszenz der LCMV Glykoprotein GP-Expression.

[0076] **Fig. 7:** MLV(LCMV)-Pseudotypen bewahren ihre Infektivität nach Ultrazentrifugation.

[0077] TELCeB wurden mit LCMV oder amphotropem Helfervirus infiziert. Die Überstände wurden geerntet und eingefroren. Es wurden die MLV(LCMV)- und amphotropen Pseudotyp-Titer bestimmt. Gleiche Mengen des infektiösen Virus wurde mittels Ultrazentrifugation pelletiert und dann einer Reinigung an einem 0 % – 40 % Urografin-Gradienten unterworfen. Vektor-Titer und -dichten wurden in jeder Fraktion bestimmt. O-O amphotroper Pseudotyp; □-□ MLV(LCMV)-Pseudotyp.

SEQUENZPROTOKOLL

(1) ALLGEMEINE ANGABEN:

(i) ANMELDER:

- (A) NAME: Heinrich-Pette-Institut
- (B) STRASSE: Martinistrasse 52
- (C) ORT: Hamburg
- (D) BUNDESLAND: Hamburg
- (E) LAND: Deutschland
- (F) POSTLEITZAHL: 20251

(ii) BEZEICHNUNG DER ERFINDUNG: Retrovirale, mit LCMV pseudotypisierte Hybrid-Vektoren

(iii) ANZAHL DER SEQUENZEN: 24

(iv) COMPUTER-LESBARE FASSUNG:

- (A) DATENTRÄGER: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

(2) ANGABEN ZU SEQ ID NO: 1:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 3375 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

CGCACCGGG	ATCCTAGGCT	TTTTGGATTG	CGCTTTCCTT	TAGGACAAC	GGGTGCTGGA	60
TTCTATCCAG	TAAAAGGATG	GGTCAGATTG	TGACAATGTT	TGAGGCTTTG	CCTCACATCA	120
TTGATGAGGT	CATCAACATT	GTCATTATTG	TGCTCATTAT	AATCACGAGC	ATCAAAGCTG	180
TGTACAATTT	CGCCACCTGT	GGGATATTAG	CACTGGTCAG	CTTCCTTTTT	CTGGCTGGTA	240
GGTCCTGTGG	CATGTACGGC	CTTAATGGTC	CCGATATCTA	TAAAGGGGTT	TACCAGTTCA	300
AATCAGTGGA	GTTTGATATG	TCTCACTTAA	ATCTGACGAT	GCCCAATGCG	TGCTCAGTCA	360
ACAACTCTCA	TCACTACATC	AGTATGGGAA	GCTCTGGACT	GGAGCCAAC	TTCACCAACG	420
ACTCCATCCT	TAATCACAAC	TTCTGCAACT	TAACCTCCGC	TCTCAACAAA	AAGTCTTTTG	480
ACCATACACT	CATGAGTATA	GTCTCGAGTC	TACACCTCAG	TATCAGAGGG	AATTCCAAC	540
ACAAAGCAGT	GTCTTGTGAT	TTTAACAATG	GCATCACCAT	TCAATACAAC	TTGTCATCTT	600
CGGACCCACA	GAGCGCCATG	AGCCAGTGTA	GGACTTTCAG	AGGTAGAGTC	TTGGACATGT	660
TTAGAAGTGC	CTTTGGAGGA	AAGTACATGA	GAAGTGGCTG	GGGCTGGACA	GGTTCAGATG	720
GCAAGACCAC	TTGGTGCAGC	CAAACAAGCT	ATCAGTACCT	AATCATACAA	AACAGGACTT	780
GGGAAAACCA	CTGTAGATAT	GCAGGCCCTT	TTGGGATGTC	TAGAATCCTC	TTTGCTCAGG	840
AAAAGACAAA	GTTTCTCACT	AGGAGACTTT	CAGGCACATT	CACCTGGACC	CTGTCAGACT	900
CCTCAGGAGT	AGAAAATCCA	GGTGGTTATT	GCCTGACCAA	ATGGATGATC	CTTGCTGCAG	960
AGCTCAAATG	TTTTGGGAAT	ACAGCTGTTG	CAAAATGTAA	TGTCAATCAT	GATGAAGAGT	1020
TCTGTGACAT	GCTACGACTA	ATTGATTACA	ACAAGGCTGC	CCTGAGTAAG	TTCAAGCAAG	1080
ATGTAGAGTC	TGCCTTGCAT	GTATTCAAAA	CAACATTA	TTCTCTGATT	TCCGATCAGC	1140
TGTTGATGAG	GAATCATCTA	AGAGATCTAA	TGGGGGTACC	ATACTGTAAT	TACTCAAAGT	1200
TCTGGTATCT	GGAACATGCT	AAGACTGGTG	AGACTAGTGT	ACCCAAGTGT	TGGCTTGTC	1260
CTAATGGCTC	CTACTTGAAT	GAGACCCATT	TTAGTGATCA	AATCGAACAA	GAAGCAGATA	1320
ACATGATCAC	AGAGATGTTG	AGGAAGGACT	ACATAAAAAG	ACAAGGGAGT	ACTCCTTTAG	1380
CCTTAATGGA	TCTTTTGATG	TTTTCAACAT	CAGCATACTT	GATCAGCATC	TTTCTGCATT	1440
TTGTGAGGAT	ACCAACACAT	AGACACATAA	AGGGCGGTTT	ATGTCCAAAG	CCACATCGCT	1500
TGACCAACAA	GGGGATCTGT	AGTTGTGGTG	CATTCAAGGT	GCCTGGTGTA	AAAACATCT	1560
GGAAAAGACG	CTGATCAGCA	GCGCCTCCCT	GACTCTCCAC	CTCGAAAGAG	GTGGAGAGTC	1620

```

AGGGAGGCC AGCGGGTCTT AGAGTGTAC AACATTGGGT CCTCTGAAGA TCAAATCATG 1680
TGGCAGGATG TTGTGAACGG TCTTTAGATC AGGGAGTCTT GCCTTGAAG CACTCTCAAA 1740
GATGATGCAG TCCATGAGTG CACAGTGTGG GGTGATTTCT TTCTTCTTTT TGTCTCTCAC 1800
TACCCAGTG TGCATTTTGC ATAGCCAGCC ATATTTGTCC CACACTTTAT CTTCATATTC 1860
TCTTGAGGCC TCCTTAGTCA TCTCAACATC AATGAGTTTT ATGTCCCTTC TATTCTGTGA 1920
GTCCAGAAGC TTTCTGATGT CATCAGAACC TTGACAGCTC AAGACCATCC CTTGTGGGAG 1980
AGCACCTATA ACTGATGAGG TCAGCCCAGC CTGTGCATTG AAGAGGTCAG CAAGATCCAT 2040
GCCGTGTGAA TACTTGGAGT CCTGTCTGAA TTGCTTCTGG TCCGTAGGTT CTCTGTAAAA 2100
ATGTATGAAT TGCCCATTTT GTGGTTGAAA TATTGCTATC TCCACTGGAT CATTGAACCT 2160
GCCTTCAATG TCAATCCATG TGGGAGCATT GGGATCAATC CCTCCCATCA AGTCTTTCAA 2220
CAGCATTGTT TGACTGTAAC TCAAGCCCAC CTGAGGTGGG CCTGCTGCTC CAGGCACTGG 2280
CCTAGATGAG TTGGCCACAA GTTTTTTCATT TGTGAGATCA ATTGTCGTGT TCTCCCATGC 2340
TCTCCCCACA ACTGACGTTT TACAGGCTAT GTATGGCCAT CCTTCACCTG AAAGACAGAC 2400
TTTATAAAG ATGTTTTTCAT AAGGATTTCT ATCCCCAACT TGATCTGAGA CAAACATGTT 2460
GAGTTTCTTC TTGGCCCCAA GGACTGCTTT TAGGAGATCC TCACTATTGC TTGGTTTGAT 2520
CAAAATAGAT TCCAGCATGT TCCCTCCATG TAGCAGAGCT GCCCCCGCTT TCACAGCCGC 2580
ACCAAGACTG AAATTATAAC CAGAGATATT GATACTAGAT TGCTGTTTCA TAATGACCCC 2640
CAGAACTGGG TGTTTATCTT TTAGCCTTTC TAGGTCACTG AGATTGGGGT ATTTGACTGT 2700
GTAAAGTAAG CCAAGGTCTG TGAGTGCCTG CACAACATCA TTGAGTGGGG TCTGTGACTG 2760
TTTTGCCATG CAAGCCATTG TCAGGCTTGG CATTGTGCCG AACTGATTGT TCAGAAGTGA 2820
TGAGTCCTTC ACATCCCAA CCCTTACTAC ACCACTTGCA CCCTGCTGAG GTCTTCTCAT 2880
CCCAACCATT TGCAGTATTT GGGATCTCTG ATCAAGTTGT TGTGCTGTCA AATTTCCCAT 2940
GTAGACTCCA GAAGCTTGAG GCCTCTCAGT TCTCATAATT TTGGCCTTCA GCTTCTCAAG 3000
ATCAGCTGCA AGGGTCATCA ATTCCTCTGC ACTAAGTCTT CCCACTTTCA GAACATTTTT 3060
CTTTGATGTA GACTTCGGAT CAACAAGAGA ATGCACAGTC TGGTTAAGAC TCCTGAGTCT 3120
CTGCAAGTCT TTATCGTCCC TCCTTTCCTT TCTCATGATC CTCTGAACGT TGCTGACTTC 3180
AGAAAAGTCC AACCCATTTA GAAGACTGGT TCGTCTTTG ATGACGGCAG CCTTTACATC 3240
TGATGTAAAA CCCTGCAACT CCCTCCTCAA CGCCTGTGTC CACTGAAAGC TTTTGACTTC 3300
TTTGGACAAA GACATTTTGT CACACAATGA ATTTCCAAAT AAAAGCGCAA TCAAATGCCT 3360
AGGATCCACT GTGCG 3375

```

(2) ANGABEN ZU SEQ ID NO: 2:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 498 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "S protein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

```

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

```

Thr	Asn	Asp	Ser	Ile	Leu	Asn	His	Asn	Phe	Cys	Asn	Leu	Thr	Ser	Ala
		115					120					125			
Leu	Asn	Lys	Lys	Ser	Phe	Asp	His	Thr	Leu	Met	Ser	Ile	Val	Ser	Ser
	130					135					140				
Leu	His	Leu	Ser	Ile	Arg	Gly	Asn	Ser	Asn	Tyr	Lys	Ala	Val	Ser	Cys
145					150					155					160
Asp	Phe	Asn	Asn	Gly	Ile	Thr	Ile	Gln	Tyr	Asn	Leu	Ser	Ser	Ser	Asp
				165					170						175
Pro	Gln	Ser	Ala	Met	Ser	Gln	Cys	Arg	Thr	Phe	Arg	Gly	Arg	Val	Leu
			180					185					190		
Asp	Met	Phe	Arg	Thr	Ala	Phe	Gly	Gly	Lys	Tyr	Met	Arg	Ser	Gly	Trp
		195					200					205			
Gly	Trp	Thr	Gly	Ser	Asp	Gly	Lys	Thr	Thr	Trp	Cys	Ser	Gln	Thr	Ser
	210					215					220				
Tyr	Gln	Tyr	Leu	Ile	Ile	Gln	Asn	Arg	Thr	Trp	Glu	Asn	His	Cys	Arg
225				230						235					240
Tyr	Ala	Gly	Pro	Phe	Gly	Met	Ser	Arg	Ile	Leu	Phe	Ala	Gln	Glu	Lys
				245					250					255	
Thr	Lys	Phe	Leu	Thr	Arg	Arg	Leu	Ser	Gly	Thr	Phe	Thr	Trp	Thr	Leu
			260					265					270		
Ser	Asp	Ser	Ser	Gly	Val	Glu	Asn	Pro	Gly	Gly	Tyr	Cys	Leu	Thr	Lys
		275					280					285			
Trp	Met	Ile	Leu	Ala	Ala	Glu	Leu	Lys	Cys	Phe	Gly	Asn	Thr	Ala	Val
	290					295					300				
Ala	Lys	Cys	Asn	Val	Asn	His	Asp	Glu	Glu	Phe	Cys	Asp	Met	Leu	Arg
305					310					315					320
Leu	Ile	Asp	Tyr	Asn	Lys	Ala	Ala	Leu	Ser	Lys	Phe	Lys	Gln	Asp	Val
				325					330					335	
Glu	Ser	Ala	Leu	His	Val	Phe	Lys	Thr	Thr	Leu	Asn	Ser	Leu	Ile	Ser
			340					345					350		
Asp	Gln	Leu	Leu	Met	Arg	Asn	His	Leu	Arg	Asp	Leu	Met	Gly	Val	Pro
		355					360					365			
Tyr	Cys	Asn	Tyr	Ser	Lys	Phe	Trp	Tyr	Leu	Glu	His	Ala	Lys	Thr	Gly
	370					375					380				
Glu	Thr	Ser	Val	Pro	Lys	Cys	Trp	Leu	Val	Thr	Asn	Gly	Ser	Tyr	Leu
385					390					395					400
Asn	Glu	Thr	His	Phe	Ser	Asp	Gln	Ile	Glu	Gln	Glu	Ala	Asp	Asn	Met
				405					410					415	
Ile	Thr	Glu	Met	Leu	Arg	Lys	Asp	Tyr	Ile	Lys	Arg	Gln	Gly	Ser	Thr
			420					425					430		
Pro	Leu	Ala	Leu	Met	Asp	Leu	Leu	Met	Phe	Ser	Thr	Ser	Ala	Tyr	Leu
		435					440					445			
Ile	Ser	Ile	Phe	Leu	His	Phe	Val	Arg	Ile	Pro	Thr	His	Arg	His	Ile
	450					455					460				
Lys	Gly	Gly	Ser	Cys	Pro	Lys	Pro	His	Arg	Leu	Thr	Asn	Lys	Gly	Ile
465					470					475					480
Cys	Ser	Cys	Gly	Ala	Phe	Lys	Val	Pro	Gly	Val	Lys	Thr	Ile	Trp	Lys
				485					490					495	
Arg	Arg														

(2) ANGABEN ZU SEQ ID NO: 3:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 3376 Basenpaare
 - (B) ART: Nucleotid
 - (C) STRANGFORM: Doppelstrang
 - (D) TOPOLOGIE: linear

- (ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

CGCACCGGGG	ATCCTAGGCT	TTTTGGATTG	CGCTTTCCTC	TAGATCAACT	GGGTGTCAGG	60
CCCTATCCTA	CAGAAGGATG	GGTCAGATTG	TGACAATGTT	TGAGGCTCTG	CCTCACATCA	120
TCGATGAGGT	GATCAACATT	GTCATTATTG	TGCTTATCGT	GATCACGGGT	ATCAAGGCTG	180
TCTACAATTT	TGCCACCTGT	GGGATATTCG	CATTGATCAG	TTTCCTACTT	CTGGCTGGCA	240
GGTCCTGTGG	CATGTACGGT	CTTAAGGGAC	CCGACATTTA	CAAAGGAGTT	TACCAATTTA	300
AGTCAGTGGA	GTTTGATATG	TCACATCTGA	ACCTGACCAT	GCCCAACGCA	TGTTACAGCCA	360
ACAAGACCCA	CCATTACATC	AGTATGGGGA	CTTCTGGACT	AGAATTGACC	TTCACCAATG	420
ATTCCATCAT	CAGTCACAAC	TTTTGCAATC	TGACCTCTGC	CTTCAACAAA	AAGACCTTTG	480
ACCACACACT	CATGAGTATA	GTTTCGAGCC	TACACCTCAG	TATCAGAGGG	AACTCCAAC	540
ATAAGGCAGT	ATCCTGCGAC	TTCAACAATG	GCATAACCAT	CCAATACAAC	TTGACATTCT	600
CAGATCGACA	AAGTGCTCAG	AGCCAGTGTA	GAACCTTCAG	AGGTAGAGTC	CTAGATATGT	660
TTAGAACTGC	CTTCGGGGGG	AAATACATGA	GGAGTGGCTG	GGGCTGGACA	GGCTCAGATG	720
GCAAGACCAC	CTGGTGTAGC	CAGACGAGTT	ACCAATACCT	GATTATACAA	AATAGAACCT	780
GGGAAAACCA	CTGCACATAT	GCAGGTCCCT	TTGGGATGTC	CAGGATTCTC	CTTTCCCAAG	840
AGAAGACTAA	GTTCTTCACT	AGGAGACTAG	CGGGCACATT	CACCTGGACT	TTGTCAGACT	900
CTTCAGGGGT	GGAGAATCCA	GGTGGTTATT	GCCTGACCAA	ATGGATGATT	CTTGCTGCAG	960
AGCTTAAGTG	TTTCGGGAAC	ACAGCAGTTG	CGAAATGCAA	TGTAAATCAT	GATGCCGAAT	1020
TCTGTGACAT	CTGCGGACTA	ATTGACTACA	ACAAGGCTGC	TTTGAGTAAG	TTCAAAGAGG	1080
ACGTAGAATC	TGCCTTGCAC	TTATTCAAAA	CAACAGTGAA	TTCTTTGATT	TCAGATCAAC	1140
TACTGATGAG	GAACCACTTG	AGAGATCTGA	TGGGGGTGCC	ATATTGCAAT	TACTCAAAGT	1200
TTTGGTACCT	AGAACATGCA	AAGACCGGCG	AACTAGTGT	CCCCAAGTGC	TGGCTTGTC	1260
CCAATGGTTC	TTACTTAAAT	GAGACCCACT	TCAGTGATCA	AATCGAACAG	GAAGCCGATA	1320
ACATGATTAC	AGAGATGTTG	AGGAAGGATT	ACATAAAGAG	GCAGGGGAGT	ACCCCCCTAG	1380
CATTGATGGA	CCTTCTGATG	TTTTCCACAT	CTGCATATCT	AGTCAGCATC	TTCTCTCACC	1440
TTGTCAAAAT	ACCAACACAC	AGGCACATAA	AAGGTGGCTC	ATGTCCAAAAG	CCACACCGAT	1500
TAACCAACAA	AGGAATTTGT	AGTTGTGGTG	CATTTAAGGT	GCCTGGTGTA	AAAACCGTCT	1560
GGAAAAGACG	CTGAAGAACA	GCGCCTCCCT	GACTCTCCAC	CTCGAAAGAG	GTGGAGAGTC	1620
AGGGAGGCC	AGAGGGTCTT	AGAGTGTAC	AACATTTGGG	CCTCTAAAAA	TTAGGTTCATG	1680
TGGCAGAATG	TTGTGAACAG	TTTTTCAGATC	TGGGAGCCTT	GCTTTGGAGG	CGCTTTCAA	1740
AATGATGCAG	TCCATGACAG	CACAGTGCAG	GGTGATCTCT	TTCTTCTTTT	TGTCCTTAC	1800
TATTCCAGTA	TGCATCTTAC	ACAACCAGCC	ATATTTGTCC	CACACTTTGT	CTTCATACTC	1860
CCTCGAAGCT	TCCCTGGTCA	TTTCAACATC	GATAAGCTTA	ATGTCCTTCC	TATTCTGTGA	1920
GTCCAGAAGC	TTTCTGATGT	CATCGGAGCC	TTGACAGCTT	AGAACCATCC	CCTGCGGAAG	1980
AGCACCTATA	ACTGACGAGG	TCAACCCGGG	TTGCGCATTG	AAGAGGTCGG	CAAGATCCAT	2040
GCCGTGTGAG	TACTTGGAA	CTTGCTTGAA	TTGTTTTTGA	TCAACGGGTT	CCCTGTAAAA	2100
GTGTATGAAC	TGCCCGTTCT	GTGGTTGGAA	AATTGCTATT	TCCACTGGAT	CATTAATCT	2160
ACCCTCAATG	TCAATCCATG	TAGGAGCGTT	GGGGTCAATT	CCTCCCATGA	GGTCTTTTAA	2220
AAGCATTGTC	TGGCTGTAGC	TTAAGCCCAC	CTGAGGTGGA	CCTGCTGCTC	CAGGCGCTGG	2280
CCTGGGTGAA	TTGACTGCAG	GTTTCTCGCT	TGTGAGATCA	ATTGTTGTGT	TTTCCCATGC	2340
TCTCCCCACA	ATCGATGTTT	TACAAGCTAT	GTATGGCCAT	CCTTCACCTG	AAAGGCAAAC	2400
TTTATAGAGG	ATGTTTTTCAT	AAGGGTTCC	GTCCCCAAT	TGGTCTGAAA	CAAACATGTT	2460
GAGTTTTCTC	TTGGCCCCGA	GAAGTGCCTT	CAAGAGGTC	TCGCTGTTGC	TTGGCTTGAT	2520
CAAAATTGAC	TCTAACATGT	TACCCCCATC	CAACAGGGCT	GCCCCCTGCCT	TCACGGCAGC	2580
ACCAAGACTA	AAGTTATAGC	CAGAAATGTT	GATGCTGGAC	TGCTGTTTCC	TGATGACCCC	2640
CAGAAGTGGG	TGCTTGTCTT	TCAGCCTTTC	AAGATCATT	AGATTTGGAT	ACTTGACTGT	2700
GTAAGCAAG	CCAAGGCTG	TGAGCGCTTG	TACAACGTCA	TTGAGCGGAG	TCTGTGACTG	2760
TTTGCCATC	CAAGCCATAG	TTAGACTTGG	CATTGTGCCA	AATTGATTGT	TCAAAAGTGA	2820
TGAGTCTTTC	ACATCCCAA	CTCTTACCAC	ACCACTTGCA	CCCTGCTGAG	GCTTTCTCAT	2880
CCCAACTATC	TGTAGGATCT	GAGATCTTTG	GTCTAGTTGC	TGTGTTGTTA	AGTTCCCCAT	2940
ATATACCCCT	GAAGCCTGGG	GCCTTTCAGA	CCTCATGATC	TTGGCCTTCA	GCTTCTCAAG	3000
GTCAGCCGCA	AGAGACATCA	GTTCTTCTGC	ACTGAGCCTC	CCCACTTTC	AAACATTCTT	3060
CTTTGATGTT	GACTTTAAAT	CCACAAGAGA	ATGTACAGTC	TGGTTGAGAC	TTCTGAGTCT	3120
CTGTAGTCT	TTGTCATCTC	TCTTTTCTT	CCTCATGATC	CTCTGAACAT	TGCTGACCTC	3180
AGAGAAGTCC	AACCCATTCA	GAAGGTTGGT	TGCATCCTTA	ATGACAGCAG	CCTTCACATC	3240
TGATGTGAAG	CTCTGCAATT	CTCTTCTCAA	TGCTTGCCTC	CATTGGAAGC	TCTTAACTTC	3300
CTTAGACAAG	GACATCTTGT	TGCTCAATGG	TTTCTCAAGA	CAAATGCGCA	ATCAAATGCC	3360
TAGGATCCAC	TGTGCG					3376

(2) ANGABEN ZU SEQ ID NO: 4:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 498 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "envelope glycoprotein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

```

Met Gly Gln Ile Val Thr Met Phe Glu Ala Leu Pro His Ile Ile Asp
1          5          10          15
Glu Val Ile Asn Ile Val Ile Ile Val Leu Ile Val Ile Thr Gly Ile
20          25          30
Lys Ala Val Tyr Asn Phe Ala Thr Cys Gly Ile Phe Ala Leu Ile Ser
35          40          45
Phe Leu Leu Leu Ala Gly Arg Ser Cys Gly Met Tyr Gly Leu Lys Gly
50          55          60
Pro Asp Ile Tyr Lys Gly Val Tyr Gln Phe Lys Ser Val Glu Phe Asp
65          70          75          80
Met Ser His Leu Asn Leu Thr Met Pro Asn Ala Cys Ser Ala Asn Asn
85          90          95
Ser His His Tyr Ile Ser Met Gly Thr Ser Gly Leu Glu Leu Thr Phe
100         105         110
Thr Asn Asp Ser Ile Ile Ser His Asn Phe Cys Asn Leu Thr Ser Ala
115         120         125
Phe Asn Lys Lys Thr Phe Asp His Thr Leu Met Ser Ile Val Ser Ser
130         135         140
Leu His Leu Ser Ile Arg Gly Asn Ser Asn Tyr Lys Ala Val Ser Cys
145         150         155         160
Asp Phe Asn Asn Gly Ile Thr Ile Gln Tyr Asn Leu Thr Phe Ser Asp
165         170         175
Arg Gln Ser Ala Gln Ser Gln Cys Arg Thr Phe Arg Gly Arg Val Leu
180         185         190
Asp Met Phe Arg Thr Ala Phe Gly Gly Lys Tyr Met Arg Ser Gly Trp
195         200         205
Gly Trp Thr Gly Ser Asp Gly Lys Thr Thr Trp Cys Ser Gln Thr Ser
210         215         220
Tyr Gln Tyr Leu Ile Ile Gln Asn Arg Thr Trp Glu Asn His Cys Thr
225         230         235         240
Tyr Ala Gly Pro Phe Gly Met Ser Arg Ile Leu Leu Ser Gln Glu Lys
245         250         255
Thr Lys Phe Phe Thr Arg Arg Leu Ala Gly Thr Phe Thr Trp Thr Leu
260         265         270
Ser Asp Ser Ser Gly Val Glu Asn Pro Gly Gly Tyr Cys Leu Thr Lys
275         280         285
Trp Met Ile Leu Ala Ala Glu Leu Lys Cys Phe Gly Asn Thr Ala Val
290         295         300
Ala Lys Cys Asn Val Asn His Asp Ala Glu Phe Cys Asp Met Leu Arg
305         310         315         320
Leu Ile Asp Tyr Asn Lys Ala Ala Leu Ser Lys Phe Lys Glu Asp Val
325         330         335
Glu Ser Ala Leu His Leu Phe Lys Thr Thr Val Asn Ser Leu Ile Ser
340         345         350
Asp Gln Leu Leu Met Arg Asn His Leu Arg Asp Leu Met Gly Val Pro
355         360         365
Tyr Cys Asn Tyr Ser Lys Phe Trp Tyr Leu Glu His Ala Lys Thr Gly
370         375         380

```

Glu Thr Ser Val Pro Lys Cys Trp Leu Val Thr Asn Gly Ser Tyr Leu
 385 390 395 400
 Asn Glu Thr His Phe Ser Asp Gln Ile Glu Gln Glu Ala Asp Asn Met
 405 410 415
 Ile Thr Glu Met Leu Arg Lys Asp Tyr Ile Lys Arg Gln Gly Ser Thr
 420 425 430
 Pro Leu Ala Leu Met Asp Leu Leu Met Phe Ser Thr Ser Ala Tyr Leu
 435 440 445
 Val Ser Ile Phe Leu His Leu Val Lys Ile Pro Thr His Arg His Ile
 450 455 460
 Lys Gly Gly Ser Cys Pro Lys Pro His Arg Leu Thr Asn Lys Gly Ile
 465 470 475 480
 Cys Ser Cys Gly Ala Phe Lys Val Pro Gly Val Lys Thr Val Trp Lys
 485 490 495
 Arg Arg

(2) ANGABEN ZU SEQ ID NO: 5:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 558 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "nucleoprotein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

Met Ser Leu Ser Lys Glu Val Lys Ser Phe Gln Trp Thr Gln Ala Leu
 1 5 10 15
 Arg Arg Glu Leu Gln Ser Phe Thr Ser Asp Val Lys Ala Ala Val Ile
 20 25 30
 Lys Asp Ala Thr Asn Leu Leu Asn Gly Leu Asp Phe Ser Glu Val Ser
 35 40 45
 Asn Val Gln Arg Ile Met Arg Lys Glu Lys Arg Asp Asp Lys Asp Leu
 50 55 60
 Gln Arg Leu Arg Ser Leu Asn Gln Thr Val His Ser Leu Val Asp Leu
 65 70 75 80
 Lys Ser Thr Ser Lys Lys Asn Val Leu Lys Val Gly Arg Leu Ser Ala
 85 90 95
 Glu Glu Leu Met Ser Leu Ala Ala Asp Leu Glu Lys Leu Lys Ala Lys
 100 105 110
 Ile Met Arg Ser Glu Arg Pro Gln Ala Ser Gly Val Tyr Met Gly Asn
 115 120 125
 Leu Thr Thr Gln Gln Leu Asp Gln Arg Ser Gln Ile Leu Gln Ile Val
 130 135 140
 Gly Met Arg Lys Pro Gln Gln Gly Ala Ser Gly Val Val Arg Val Trp
 145 150 155 160
 Asp Val Lys Asp Ser Ser Leu Leu Asn Asn Gln Phe Gly Thr Met Pro
 165 170 175
 Ser Leu Thr Met Ala Cys Met Ala Lys Gln Ser Gln Thr Pro Leu Asn
 180 185 190
 Asp Val Val Gln Ala Leu Thr Asp Leu Gly Leu Leu Tyr Thr Val Lys
 195 200 205
 Tyr Pro Asn Leu Asn Asp Leu Glu Arg Leu Lys Asp Lys His Pro Val
 210 215 220
 Leu Gly Val Ile Thr Glu Gln Gln Ser Ser Ile Asn Ile Ser Gly Tyr
 225 230 235 240

Asn Phe Ser Leu Gly Ala Ala Val Lys Ala Gly Ala Ala Leu Leu Asp
 245 250 255
 Gly Gly Asn Met Leu Glu Ser Ile Leu Ile Lys Pro Ser Asn Ser Glu
 260 265 270
 Asp Leu Leu Lys Ala Val Leu Gly Ala Lys Arg Lys Leu Asn Met Phe
 275 280 285
 Val Ser Asp Gln Val Gly Asp Arg Asn Pro Tyr Glu Asn Ile Leu Tyr
 290 295 300
 Lys Val Cys Leu Ser Gly Glu Gly Trp Pro Tyr Ile Ala Cys Arg Thr
 305 310 315 320
 Ser Ile Val Gly Arg Ala Trp Glu Asn Thr Thr Ile Asp Leu Thr Ser
 325 330 335
 Glu Lys Pro Ala Val Asn Ser Pro Arg Pro Ala Pro Gly Ala Ala Gly
 340 345 350
 Pro Pro Gln Val Gly Leu Ser Tyr Ser Gln Thr Met Leu Lys Asp
 355 360 365
 Leu Met Gly Gly Ile Asp Pro Asn Ala Pro Thr Trp Ile Asp Ile Glu
 370 375 380
 Gly Arg Phe Asn Asp Pro Val Glu Ile Ala Ile Phe Gln Pro Gln Asn
 385 390 395 400
 Gly Gln Phe Ile His Phe Tyr Arg Glu Pro Val Asp Gln Lys Gln Phe
 405 410 415
 Lys Gln Asp Ser Lys Tyr Ser His Gly Met Asp Leu Ala Asp Leu Phe
 420 425 430
 Asn Ala Gln Pro Gly Leu Thr Ser Ser Val Ile Gly Ala Leu Pro Gln
 435 440 445
 Gly Met Val Leu Ser Cys Gln Gly Ser Asp Asp Ile Arg Lys Leu Leu
 450 455 460
 Asp Ser Gln Asn Arg Lys Asp Ile Lys Leu Ile Asp Val Glu Met Thr
 465 470 475 480
 Arg Glu Ala Ser Arg Glu Tyr Glu Asp Lys Val Trp Asp Lys Tyr Gly
 485 490 495
 Trp Leu Cys Lys Met His Thr Gly Ile Val Arg Asp Lys Lys Lys
 500 505 510
 Glu Ile Thr Pro His Cys Ala Leu Met Asp Cys Ile Ile Phe Glu Ser
 515 520 525
 Ala Ser Lys Ala Arg Leu Pro Asp Leu Lys Thr Val His Asn Ile Leu
 530 535 540
 Pro His Asp Leu Ile Phe Arg Gly Pro Asn Val Val Thr Leu
 545 550 555

(2) ANGABEN ZU SEQ ID NO: 6:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 6680 Basenpaare
 - (B) ART: Nucleotid
 - (C) STRANGFORM: Doppelstrang
 - (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

CGCACCGAGG	ATCCTAGGCT	TTTTGATGCG	CAATGGATGA	AATCATCTCA	GAATTGAGAG	60
AGTTATGTTT	AAACTATATA	GAACAGGATG	AGAGGTTGTC	AAGGCAGAAA	CTCAACTTTC	120
TGGGACAAAG	GGAACCCAGA	ATGGTTCTGA	TTGAGGGACT	CAAGTTGCTG	TCACGCTGCA	180
TTGAAATAGA	CAGTGCAGAC	AAGAGTGGCT	GCACACACAA	CCACGACGAT	AAGTCTGTGG	240
AAACAATTTT	GGTGGAGTCT	GGAATTGTAT	GCCCAGGACT	ACCACTTATC	ATTCCTGATG	300
GTTACAAGCT	GATAGACAAT	TCTCTCATTC	TTCTTGAGTG	TTTTGTTAGG	AGCTCACCAG	360
CCAGTTTTGA	GAAGAAATTT	ATAGAGGACA	CTAACAAATT	GGCATGCATC	AGGGAAGACC	420
TTGCTGTTGC	GGGTGTCACA	TTAGTTCCAA	TAGTAGATGG	TCGTTGTGAT	TATGATAATA	480

GTTTTATGCC	AGAGTGGGCA	AACTTCAAAT	TTAGAGACCT	TTTATTCAAA	CTTTTGGAGT	540
ATTCTAACCA	AAATGAGAAA	GTCTTTGAAG	AGTCTGAATA	TTTTAGACTC	TGTGAGTCCC	600
TGAAGACTAC	TATCGACAAG	CGTCCGGTA	TGGACTCTAT	GAAAATTCTG	AAAGATGCGA	660
GGTCAACTCA	CAATGATGAA	ATTATGAGGA	TGTGCCACGA	AGGCATCAAC	CCCAACATGA	720
GCTGTGATGA	TGTGGTTTTT	GGAATAAACT	CTCTTTTCAG	CAGGTTTAGA	AGAGATTTAG	780
AAAGTGGGAA	ATTAAGAGAG	AACTTTTCTG	AAGTAAACCC	TGAAGGCTTG	ATCAAGGAAT	840
TCTCTGAGCT	CTATGAAAA	CTTGCTGATA	GTGATGATAT	CTTAACATTA	AGCAGGGAGG	900
CAGTCGAATC	CTGTCCTTTG	ATGAGATTCA	TAACCTGAGA	GACCCATGGG	CACGAAAGGG	960
GAAGTGAGAC	TAGCACTGAA	TATGAGAGGC	TCCTCTCTAT	GTTAAACAAA	GTCAAGAGTT	1020
TGAAACTGTT	GAATACTAGA	AGGAGACAGT	TGTTAAATCT	GGATGTTTTG	TGTCTTTCCT	1080
CATTGATAAA	ACAGTCGAAA	TTCAAAGGGT	TAAAAAATGA	TAAACACTGG	GTGGGTTGTT	1140
GCTATAGTAG	TGTGAATGAT	AGGCTGGTAA	GCTTTCACAG	CACTAAAGAG	GAGTTCATTA	1200
GACTTTTGAG	GAATAGAAAA	AAGTCAAAGG	TGTTTAGAAA	GGTGTCTTTT	GAGGAATTGT	1260
TTAGGGCGTC	TATTAGTGAG	TTCATTGCAA	AAATTCAAAA	ATGCCTGTTA	GTGGTGGGAC	1320
TGAGTTTCGA	GCATTACGGA	CTGTCTGAAC	ACCTTGAGCA	AGAATGCCAC	ATACCATTCA	1380
CTGAATTTGA	GAACCTTATG	AAAATTGGAG	CTCACCCGAT	AATGTATTAT	ACGAAGTTTG	1440
AAGATTACAA	TTTCCAACCC	AGCACAGAGC	AGCTGAAGAA	CATACAGAGC	CTGAGAAGAT	1500
TATCATCTGT	TTGTCTGGCC	TAAACAAACA	GTATGAAAAC	TAGCTCAGTT	GCTAGACTAA	1560
GGCAAAATCA	AATAGGGTCT	GTGAGATATC	AAGTGGTAGA	ATGCAAAGAA	GTGTTTTGCC	1620
AAGTAATAAA	ACTGGACTCT	GAAGAATACC	ACCTATTATA	CCAGAAGACT	GGAGAATCTT	1680
CAAGGTGCTA	CTCCATACAA	GGCCCGGATG	GTCATTTAAT	TTCTTCTTAT	GCAGATCTTA	1740
AAAGGTTCTT	TTTACCAATT	TTTTCAGATG	AGGTCTTATA	CAATATGATA	GACATCATGA	1800
TTTCATGGAT	TAGATCATGT	CCTGATTTGA	AAGACTGTCT	CACCGACATT	GAGGTTGCAC	1860
TGAGGACCCCT	ATTGTTGCTA	ATGCTCACCA	ACCCAACAAA	GAGAAATCAA	AAGCAGGTAC	1920
AGAGTGTGAG	ATATTTGGTG	ATGGCAATAG	TGTCAGATTT	TTCATCTACA	TCATTAATGG	1980
ATAAGTTTAC	GGAGGATCTG	ATCACACCTG	CTGAGAAGTT	GGTGTATAAG	CTGCTTAGAT	2040
TCCTAATAAAA	AACTATTTTT	GGTACTGGTG	AGAAGGTGTT	GTTGAGTGCA	AAATTTAAAT	2100
TTATGTTGAA	TGTGTCATAC	CTGTGTCATT	TGATCACAAA	GGAGACCCCT	GACAGGCTAA	2160
CAGATCAGAT	AAAATGTTTT	GAAAAGTTCT	TTGAGCCCAA	AAGTCAATTT	GGTTTTTTTTG	2220
TCAACCCCAA	GGAAAGCAATC	ACTCCTGAGG	AAGAATGTGT	GTTCTATGAG	CAAATGAAGA	2280
GATTCACTAG	TAAAGAAATF	GACTGTCAGC	ATACAACCTC	AGGTGTTAAT	CTGGAAGCCT	2340
TTAGCCATAA	GGTGCTTTCA	TTTAAACAAC	GCACTTTAAT	TTTCAAAGGA	GAGAAGAAGC	2400
TAAACAGCCT	AGATCCCATG	ACTAACTCTG	GATGTCCGAC	AGCATTAGAT	CTTGCTAGTA	2460
ACAAAAGTGT	GGTGGTTAAT	AAGCATCTAA	ATGGAGAACG	ACTTCTGGAA	TATGACTTTA	2520
ACAAATTGCT	TGTTAGTGCT	GTGAGTCAAA	TTACGGAGAG	TTTCGTAAGA	AAACAAAAGT	2580
ATAAGTTGAG	CCACTCAGAC	TATGAATATA	AAGTTCCCAA	GTTAGTCTCT	AGATTGGTCA	2640
TCGGTTCCAA	GGGAGAAGAG	ACAGGGAGAT	CGGAAGACAA	CCTGGCAGAA	ATATGTTTTG	2700
ATGGAGAAGA	AGAGACAAGC	TTCTTCAAAA	GTCTCGAAGA	AAAGGTCAAC	ACCACAATAG	2760
CACGGTACAG	AAGAGGTAGG	AGGGCCAATG	ACAAAGGAGA	TGGAGAAAAA	CTTACAAAAT	2820
CAAAAGGACT	ACATCATTTA	CAGCTTATTC	TAACAGGGAA	GATGGCTCAC	TTAAGAAAAG	2880
TTATCTTGTC	AGAAATATCT	TTCCATTTAG	TAGAAGACTT	TGACCCATCA	TGTCTAACCA	2940
ATGATGACAT	GAAATTTATC	TGTGAGGCTG	TTGAGGGTTC	CACAGAGCTG	TCACCTTTGT	3000
ATTTACCTC	AGTCATTAAT	GATCAGTGTG	GCCTCGATGA	GATGGCAAAA	AACCTTTGTA	3060
GAAAGTTCTT	TTCTGAGAAT	GATTGGTTTT	CTTGCATGAA	GATGATTCTG	TTGCAATGA	3120
ATGCAAATGC	GTACTCAGGG	AAATACAGGC	ATATGCAAAG	GCAAGGCTTG	AATTTCAAAT	3180
TTGACTGGGA	CAAACCTGGAA	GAAGACGTGA	GAATCAGTGA	GAGGGAAAAGT	AATTCAGTGT	3240
CCCTTAGTAA	AGCTCTGTCG	TTGACAAAAAT	GTATGAGTGC	TGCTTTGAAA	AATCTGTGCT	3300
TCTACTCAGA	AGAATCACCA	ACATCATACA	CCTCAGTAGG	TCCTGACTCT	GGAAGGCTGA	3360
AATTTGCACT	ATCTTATAAA	GAGCAGGTTG	GGGGAAAATG	AGAACTCTAT	ATTGGAGATT	3420
TGAGGACAAA	AATGTTTACA	AGGTTAATAG	AAGATTATTT	TGAGTCTTTT	TCAAGTTTCT	3480
TTTCAGGCTC	CTGTTTAAAC	AATGATAAAG	AATTTGAAAA	TGCAATCTTG	TCAATGACTA	3540
TCAATGTGCG	GGAAAGGTTT	TTAAACTATA	GTATGGATCA	CAGCAAATGG	GGACCAATGA	3600
TGTGCCCAAT	TTTGTCTTTA	ATGTTTCTAC	AAAATCTCAA	ACTAGGTGAT	GACCAGTATG	3660
TGCGTTCCGG	GAAAGATCAT	GTTAGCACTT	TGTTAACTTG	GCACATGCAT	AAGCTTGTGC	3720
AGGTCCCCTT	TCCTGTTGTG	AATGCAATGA	TGAAATCATA	TGTCAAGTCG	AAGCTAAAAAC	3780
TTCTCAGGGG	TTCAGAAAACA	ACTGTTACTG	AGAGAATTTT	CAGACAATAT	TTTGAATGG	3840
GGATAGTGCC	ATCCCATATA	TCCAGCCTTA	TTGATATGGG	GCAGGGAATC	TTGCATAATG	3900
CTTCTGACTT	CTATGGTTTT	CTTAGCGAGA	GGTTCATCAA	CTACTGCATT	GGTGTATCT	3960
TTGGCGAAAG	ACCAGAGGCT	TACACATCAA	GTGATGATCA	GATCACTTTA	TTTGATAGGA	4020
GGCTGAGTGA	CCTGGTTGTA	AGTGATCCGG	AGGAAGTCTT	TGTCCTGTTG	GAATTCCAAT	4080
CTCATCTGAG	CGGCTTGTTA	AACAAATTTA	TCAGCCCAA	AAGTGTGGCT	GGGAGGTTCC	4140
CTGCAGAATT	TAAATCTAGA	TTCTATGTAT	GGGGGGAGGA	AGTCCCTCTT	CTCACAAAAGT	4200
TTGTATCTGC	AGCGCTACAC	AATGTCAAGT	GTAAGAGGCC	ACATCAACTT	TGTGAAAACAA	4260
TAGATACAAAT	TGCAGATCAA	GCCATCGCAA	ATGGCGTCCC	AGTCTCCCTA	GTTAATAGTA	4320

TCCAAAGGAG	AACACTGGAC	CTCCTAAAGT	ATGCCAATTT	CCCTTTGGAT	CCATTTCTAC	4380
TGAATACCAA	CACTGATGTG	AAAGATTGGC	TGGATGGTTC	TAGAGGTTAC	AGAATACAAA	4440
GACTCATTGA	GGAAGTGTGT	CCTAATGAAA	CAAAGGTTGT	AAGAAAGCTT	GTAAGGAAAC	4500
TGCATCATAA	GCTCAAAAAT	GGTGAATTTA	ATGAAGAATT	TTTCTTAGAC	CTATTTAACA	4560
GAGATAAAAC	GGAGGCCATT	CTTCAATTGG	GAGACCTCCT	CGGTCTTGAA	GAAGATCTGA	4620
ATCAGTTAGC	AGATGTTAAC	TGGTTGAATT	TGAATGAAAT	GTTCCCATTA	AGGATGGTTT	4680
TAAGACAAAA	GGTGGTTTAT	CCATCAGTGA	TGACTTTCCA	AGAGGAAAGA	ATCCCATCAT	4740
TGATCAAGAC	ACTCCAGAAC	AAACTTTGTA	GTAATTTTCC	AAGGGGTGCA	CAGAAGCTGC	4800
TGTCAGAAGC	AATCAACAAG	TCAGCTTTCC	AGAGTTGTAT	CTCATCTGGC	TTTATAGGCC	4860
TTTGCAAAAAC	TCTAGGAAGC	AGGTGTGTGA	AAAACAAAAA	TAGGGAAAAT	CTGTATATCA	4920
AAAAGCTGCT	TGAGGATCTA	ACCACAGATG	ATCATGTGAC	AAGAGTTTGC	AATCGGGATG	4980
GTATAACGCT	GTACATTTGT	GACAAACAGT	CTCATCCAGA	AGCCCACCGT	GATCATATAT	5040
GCCTTTTAAG	GCCTCTTCTT	TGGGACTACA	TTTGTATTTT	ATTGAGCAAC	TCTTTTGAGT	5100
TGGGTGTTTG	GGTCCTAGCA	GAACCGACCA	AAGGGAAGAA	TAACAGTGAG	AACCTAACTC	5160
TTAAGCACTT	AAACCCATGT	GATTATGTAG	CAAGAAAGCC	TGAGAGCTCA	AGGCTACTGG	5220
AGGACAAAAGT	GAATTTGAAC	CAAGTGATTC	AATCTGTGAG	GCGGCTATAT	CCCAAGACTT	5280
TTGAGGATCA	GCTTCTTCCA	TTTATGTCTG	ACATGAGCTC	AAAAAACATG	AGGTGGAGTC	5340
CCAGAATTAA	ATTCCTTGAC	CTCTGTGTTT	TAATTGATAT	TAACTCAGAA	TCCTTGTCAC	5400
TCATTTCTCA	TGTTGTTAAG	TGGAAAAGGG	ATGAACATTA	CACTGTTCTG	TTTTCTGACC	5460
TTGCCAATTC	TCATCAGCGA	TCTGACTCCA	GTCTGGTTGA	TGAATTTGTT	GTTAGCACGA	5520
GGGATGTCTG	CAAGAACTTC	TTAAAACAGG	TGTATTTTGA	ATCATTGTTT	CGAGAATTTG	5580
TTGCAACAAC	CAGGACATTA	GGCAATTTTT	CATGTTCCC	TCATAAAGAA	ATGATGCCAT	5640
CTGAAGATGG	TGCTGAGGCA	CTGGGCCCTT	TTCAATCATT	TGCTCAAAG	GTGGTGAACA	5700
AAAATGTGGA	GAGGCCTATG	TTTAGGAATG	ATTTGCAGTT	TGGTTTTGGG	TGGTCTCTT	5760
ACCGAATGGG	AGATGTTGTG	TGTAATGCTG	CCATGTTGAT	TAGGCAGGGC	CTGACAAACC	5820
CAAAGGCATT	TAAATCCTTA	AAGGATCTGT	GGGACTACAT	GCTCAACTAC	ACAAAAGGGG	5880
TATTGGAGTT	TTCAATTTCA	GTGGACTTTA	CGCACAATCA	GAATAATACT	GAATGTTTAA	5940
GGAATTTTTC	ATTGATATTC	TTGGTTAGGT	GCCAATTACA	GAATCCAGGT	GTGGCTGAAC	6000
TTTTATCATG	CTCTCACCTC	TTTAAGGGTG	AGATAGATAG	AAGAATGTTG	GATGAATGCC	6060
TCCACTTACT	GAGGACAGAC	TCTGTCTTCA	AGGTGAACGA	TGGTGTCTTT	GATATCAGAT	6120
CTGAAGAGTT	TGAGGATTAC	ATGGAAGATC	CCTTGATACT	TGGTGATTCT	CTTGAGCTTG	6180
AGTTGTTGGG	CTCCAAAAGA	ATACTGGATG	GGATTAGATC	TATTGACTTT	GAGAGAGTTG	6240
GACCTGAGTG	GGAGCCTGTG	CCACTGACTG	TAAAGATGGG	TGCCCTTTT	GAAGGAAGAA	6300
ACCTTGTTCCA	AAATATCATT	GTGAAGCTGG	AGACCAAGGA	CATGAAAGTC	TTTCTAGCAG	6360
GACTTGAGGG	CTATGAAAAG	ATTAGTGATG	TCCTTGGGAA	CCTCTTCTCG	CATCGATTCA	6420
GAAGTGGTGA	ACATTTGTTG	GGTTCAGAGA	TAAGTGTAAT	CCTCCAGGAA	CTATGTATAG	6480
ACAGATCTAT	TCTGCTGATT	CCACTGTGCG	TTTTGCCAGA	CTGGTTCGCC	TTTAAGGATT	6540
GCAGACTTTG	TTTTAGCAAA	TCTAGGAGCA	CTTTGATGTA	TGAAATAGTG	GGGGGCAGGT	6600
TTAGACTCAA	GGGGAGGTCC	TGCGACGATT	GGCTAGGCGG	GTCGGTGGCC	GAGGACATCG	6660
ACTGATGGGC	ATCTCCTGGG					6680

(2) ANGABEN ZU SEQ ID NO: 7:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 2210 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "L protein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

Met	Asp	Glu	Ile	Ile	Ser	Glu	Leu	Arg	Glu	Leu	Cys	Leu	Asn	Tyr	Ile
1				5					10					15	
Glu	Gln	Asp	Glu	Arg	Leu	Ser	Arg	Gln	Lys	Leu	Asn	Phe	Leu	Gly	Gln
			20					25					30		

Arg Glu Pro Arg Met Val Leu Ile Glu Gly Leu Lys Leu Leu Ser Arg
 35 40 45
 Cys Ile Glu Ile Asp Ser Ala Asp Lys Ser Gly Cys Thr His Asn His
 50 55 60
 Asp Asp Lys Ser Val Glu Thr Ile Leu Val Glu Ser Gly Ile Val Cys
 65 70 75 80
 Pro Gly Leu Pro Leu Ile Ile Pro Asp Gly Tyr Lys Leu Ile Asp Asn
 85 90 95
 Ser Leu Ile Leu Leu Glu Cys Phe Val Arg Ser Ser Pro Ala Ser Phe
 100 105 110
 Glu Lys Lys Phe Ile Glu Asp Thr Asn Lys Leu Ala Cys Ile Arg Glu
 115 120 125
 Asp Leu Ala Val Ala Gly Val Thr Leu Val Pro Ile Val Asp Gly Arg
 130 135 140
 Cys Asp Tyr Asp Asn Ser Phe Met Pro Glu Trp Ala Asn Phe Lys Phe
 145 150 155 160
 Arg Asp Leu Leu Phe Lys Leu Leu Glu Tyr Ser Asn Gln Asn Glu Lys
 165 170 175
 Val Phe Glu Glu Ser Glu Tyr Phe Arg Leu Cys Glu Ser Leu Lys Thr
 180 185 190
 Thr Ile Asp Lys Arg Ser Gly Met Asp Ser Met Lys Ile Leu Lys Asp
 195 200 205
 Ala Arg Ser Thr His Asn Asp Glu Ile Met Arg Met Cys His Glu Gly
 210 215 220
 Ile Asn Pro Asn Met Ser Cys Asp Asp Val Val Phe Gly Ile Asn Ser
 225 230 235 240
 Leu Phe Ser Arg Phe Arg Arg Asp Leu Glu Ser Gly Lys Leu Lys Arg
 245 250 255
 Asn Phe Gln Lys Val Asn Pro Glu Gly Leu Ile Lys Glu Phe Ser Glu
 260 265 270
 Leu Tyr Glu Asn Leu Ala Asp Ser Asp Asp Ile Leu Thr Leu Ser Arg
 275 280 285
 Glu Ala Val Glu Ser Cys Pro Leu Met Arg Phe Ile Thr Ala Glu Thr
 290 295 300
 His Gly His Glu Arg Gly Ser Glu Thr Ser Thr Glu Tyr Glu Arg Leu
 305 310 315 320
 Leu Ser Met Leu Asn Lys Val Lys Ser Leu Lys Leu Leu Asn Thr Arg
 325 330 335
 Arg Arg Gln Leu Leu Asn Leu Asp Val Leu Cys Leu Ser Ser Leu Ile
 340 345 350
 Lys Gln Ser Lys Phe Lys Gly Leu Lys Asn Asp Lys His Trp Val Gly
 355 360 365
 Cys Cys Tyr Ser Ser Val Asn Asp Arg Leu Val Ser Phe His Ser Thr
 370 375 380
 Lys Glu Glu Phe Ile Arg Leu Leu Arg Asn Arg Lys Lys Ser Lys Val
 385 390 395 400
 Phe Arg Lys Val Ser Phe Glu Glu Leu Phe Arg Ala Ser Ile Ser Glu
 405 410 415
 Phe Ile Ala Lys Ile Gln Lys Cys Leu Leu Val Val Gly Leu Ser Phe
 420 425 430
 Glu His Tyr Gly Leu Ser Glu His Leu Glu Gln Glu Cys His Ile Pro
 435 440 445
 Phe Thr Glu Phe Glu Asn Phe Met Lys Ile Gly Ala His Pro Ile Met
 450 455 460
 Tyr Tyr Thr Lys Phe Glu Asp Tyr Asn Phe Gln Pro Ser Thr Glu Gln
 465 470 475 480
 Leu Lys Asn Ile Gln Ser Leu Arg Arg Leu Ser Ser Val Cys Leu Ala
 485 490 495
 Leu Thr Asn Ser Met Lys Thr Ser Ser Val Ala Arg Leu Arg Gln Asn
 500 505 510
 Gln Ile Gly Ser Val Arg Tyr Gln Val Val Glu Cys Lys Glu Val Phe
 515 520 525
 Cys Gln Val Ile Lys Leu Asp Ser Glu Glu Tyr His Leu Leu Tyr Gln
 530 535 540

Lys Thr Gly Glu Ser Ser Arg Cys Tyr Ser Ile Gln Gly Pro Asp Gly
 545 550 555
 His Leu Ile Ser Phe Tyr Ala Asp Pro Lys Arg Phe Phe Leu Pro Ile
 565 570 575
 Phe Ser Asp Glu Val Leu Tyr Asn Met Ile Asp Ile Met Ile Ser Trp
 580 585 590
 Ile Arg Ser Cys Pro Asp Leu Lys Asp Cys Leu Thr Asp Ile Glu Val
 595 600 605
 Ala Leu Arg Thr Leu Leu Leu Leu Met Leu Thr Asn Pro Thr Lys Arg
 610 615 620
 Asn Gln Lys Gln Val Gln Ser Val Arg Tyr Leu Val Met Ala Ile Val
 625 630 635 640
 Ser Asp Phe Ser Ser Thr Ser Leu Met Asp Lys Leu Arg Glu Asp Leu
 645 650 655
 Ile Thr Pro Ala Glu Lys Val Val Tyr Lys Leu Leu Arg Phe Leu Ile
 660 665 670
 Lys Thr Ile Phe Gly Thr Gly Glu Lys Val Leu Leu Ser Ala Lys Phe
 675 680 685
 Lys Phe Met Leu Asn Val Ser Tyr Leu Cys His Leu Ile Thr Lys Glu
 690 695 700
 Thr Pro Asp Arg Leu Thr Asp Gln Ile Lys Cys Phe Glu Lys Phe Phe
 705 710 715 720
 Glu Pro Lys Ser Gln Phe Gly Phe Phe Val Asn Pro Lys Glu Ala Ile
 725 730 735
 Thr Pro Glu Glu Glu Cys Val Phe Tyr Glu Gln Met Lys Arg Phe Thr
 740 745 750
 Ser Lys Glu Ile Asp Cys Gln His Thr Thr Pro Gly Val Asn Leu Glu
 755 760 765
 Ala Phe Ser Leu Met Val Ser Ser Phe Asn Asn Gly Thr Leu Ile Phe
 770 775 780
 Lys Gly Glu Lys Lys Leu Asn Ser Leu Asp Pro Met Thr Asn Ser Gly
 785 790 795 800
 Cys Ala Thr Ala Leu Asp Leu Ala Ser Asn Lys Ser Val Val Val Asn
 805 810 815
 Lys His Leu Asn Gly Glu Arg Leu Leu Glu Tyr Asp Phe Asn Lys Leu
 820 825 830
 Leu Val Ser Ala Val Ser Gln Ile Thr Glu Ser Phe Val Arg Lys Gln
 835 840 845
 Lys Tyr Lys Leu Ser His Ser Asp Tyr Glu Tyr Lys Val Ser Lys Leu
 850 855 860
 Val Ser Arg Leu Val Ile Gly Ser Lys Gly Glu Glu Thr Gly Arg Ser
 865 870 875 880
 Glu Asp Asn Leu Ala Glu Ile Cys Phe Asp Gly Glu Glu Glu Thr Ser
 885 890 895
 Phe Phe Lys Ser Leu Glu Glu Lys Val Asn Thr Thr Ile Ala Arg Tyr
 900 905 910
 Arg Arg Gly Arg Arg Ala Asn Asp Lys Gly Asp Gly Glu Lys Leu Thr
 915 920 925
 Asn Thr Lys Gly Leu His His Leu Gln Leu Ile Leu Thr Gly Lys Met
 930 935 940
 Ala His Leu Arg Lys Val Ile Leu Ser Glu Ile Ser Phe His Leu Val
 945 950 955 960
 Glu Asp Phe Asp Pro Ser Cys Leu Thr Asn Asp Asp Met Lys Phe Ile
 965 970 975
 Cys Glu Ala Val Glu Gly Ser Thr Glu Leu Ser Pro Leu Tyr Phe Thr
 980 985 990
 Ser Val Ile Lys Asp Gln Cys Gly Leu Asp Glu Met Ala Lys Asn Leu
 995 1000 1005
 Cys Arg Lys Phe Phe Ser Glu Asn Asp Trp Phe Ser Cys Met Lys Met
 1010 1015 1020
 Ile Leu Leu Gln Met Asn Ala Asn Ala Tyr Ser Gly Lys Tyr Arg His
 1025 1030 1035 1040
 Met Gln Arg Gln Gly Leu Asn Phe Lys Phe Asp Trp Asp Lys Leu Glu
 1045 1050 1055

Glu Asp Val Arg Ile Ser Glu Arg Glu Ser Asn Ser Glu Ser Leu Ser
 1060 1065 1070
 Lys Ala Leu Ser Leu Thr Lys Cys Met Ser Ala Ala Leu Lys Asn Leu
 1075 1080 1085
 Cys Phe Tyr Ser Glu Glu Ser Pro Thr Ser Tyr Thr Ser Val Gly Pro
 1090 1095 1100
 Asp Ser Gly Arg Leu Lys Phe Ala Leu Ser Tyr Lys Glu Gln Val Gly
 1105 1110 1115 1120
 Gly Asn Arg Glu Leu Tyr Ile Gly Asp Leu Arg Thr Lys Met Phe Thr
 1125 1130 1135
 Arg Leu Ile Glu Asp Tyr Phe Glu Ser Phe Ser Ser Phe Phe Ser Gly
 1140 1145 1150
 Ser Cys Leu Asn Asn Asp Lys Glu Phe Glu Asn Ala Ile Leu Ser Met
 1155 1160 1165
 Thr Ile Asn Val Arg Glu Gly Phe Leu Asn Tyr Ser Met Asp His Ser
 1170 1175 1180
 Lys Trp Gly Pro Met Met Cys Pro Phe Leu Phe Leu Met Phe Leu Gln
 1185 1190 1195 1200
 Asn Leu Lys Leu Gly Asp Asp Gln Tyr Val Arg Ser Gly Lys Asp His
 1205 1210 1215
 Val Ser Thr Leu Leu Thr Trp His Met His Lys Leu Val Glu Val Pro
 1220 1225 1230
 Phe Pro Val Val Asn Ala Met Met Lys Ser Tyr Val Lys Ser Lys Leu
 1235 1240 1245
 Lys Leu Leu Arg Gly Ser Glu Thr Thr Val Thr Glu Arg Ile Phe Arg
 1250 1255 1260
 Gln Tyr Phe Glu Met Gly Ile Val Pro Ser His Ile Ser Ser Leu Ile
 1265 1270 1275 1280
 Asp Met Gly Gln Gly Ile Leu His Asn Ala Ser Asp Phe Tyr Gly Leu
 1285 1290 1295
 Leu Ser Glu Arg Phe Ile Asn Tyr Cys Ile Gly Val Ile Phe Gly Glu
 1300 1305 1310
 Arg Pro Glu Ala Tyr Thr Ser Ser Asp Asp Gln Ile Thr Leu Phe Asp
 1315 1320 1325
 Arg Arg Leu Ser Asp Leu Val Val Ser Asp Pro Glu Glu Val Leu Val
 1330 1335 1340
 Leu Leu Glu Phe Gln Ser His Leu Ser Gly Leu Asn Lys Phe Ile
 1345 1350 1355 1360
 Ser Pro Lys Ser Val Ala Gly Arg Phe Ala Ala Glu Phe Lys Ser Arg
 1365 1370 1375
 Phe Tyr Val Trp Gly Glu Glu Val Pro Leu Leu Thr Lys Phe Val Ser
 1380 1385 1390
 Ala Ala Leu His Asn Val Lys Cys Lys Glu Pro His Gln Leu Cys Glu
 1395 1400 1405
 Thr Ile Asp Thr Ile Ala Asp Gln Ala Ile Ala Asn Gly Val Pro Val
 1410 1415 1420
 Ser Leu Val Asn Ser Ile Gln Arg Arg Thr Leu Asp Leu Leu Lys Tyr
 1425 1430 1435 1440
 Ala Asn Phe Pro Leu Asp Pro Phe Leu Leu Asn Thr Asn Thr Asp Val
 1445 1450 1455
 Lys Asp Trp Leu Asp Gly Ser Arg Gly Tyr Arg Ile Gln Arg Leu Ile
 1460 1465 1470
 Glu Glu Leu Cys Pro Asn Glu Thr Lys Val Val Arg Lys Leu Val Arg
 1475 1480 1485
 Lys Leu His His Lys Leu Lys Asn Gly Glu Phe Asn Glu Glu Phe Phe
 1490 1495 1500
 Leu Asp Leu Phe Asn Arg Asp Lys Thr Glu Ala Ile Leu Gln Leu Gly
 1505 1510 1515 1520
 Asp Leu Leu Gly Leu Glu Glu Asp Leu Asn Gln Leu Ala Asp Val Asn
 1525 1530 1535
 Trp Leu Asn Leu Asn Glu Met Phe Pro Leu Arg Met Val Leu Arg Gln
 1540 1545 1550
 Lys Val Val Tyr Pro Ser Val Met Thr Phe Gln Glu Glu Arg Ile Pro
 1555 1560 1565

Ser Leu Ile Lys Thr Leu Gln Asn Lys Leu Cys Ser Lys Phe Thr Arg
 1570 1575 1580
 Gly Ala Gln Lys Leu Leu Ser Glu Ala Ile Asn Lys Ser Ala Phe Gln
 1585 1590 1595 1600
 Ser Cys Ile Ser Ser Gly Phe Ile Gly Leu Cys Lys Thr Leu Gly Ser
 1605 1610 1615
 Arg Cys Val Arg Asn Lys Asn Arg Glu Asn Leu Tyr Ile Lys Lys Leu
 1620 1625 1630
 Leu Glu Asp Leu Thr Thr Asp Asp His Val Thr Arg Val Cys Asn Arg
 1635 1640 1645
 Asp Gly Ile Thr Leu Tyr Ile Cys Asp Lys Gln Ser His Pro Glu Ala
 1650 1655 1660
 His Arg Asp His Ile Cys Leu Leu Arg Pro Leu Leu Trp Asp Tyr Ile
 1665 1670 1675 1680
 Cys Ile Ser Leu Ser Asn Ser Phe Glu Leu Gly Val Trp Val Leu Ala
 1685 1690 1695
 Glu Pro Thr Lys Gly Lys Asn Asn Ser Glu Asn Leu Thr Leu Lys His
 1700 1705 1710
 Leu Asn Pro Cys Asp Tyr Val Ala Arg Lys Pro Glu Ser Ser Arg Leu
 1715 1720 1725
 Leu Glu Asp Lys Val Asn Leu Asn Gln Val Ile Gln Ser Val Arg Arg
 1730 1735 1740
 Leu Tyr Pro Lys Ile Phe Glu Asp Gln Leu Leu Pro Phe Met Ser Asp
 1745 1750 1755 1760
 Met Ser Ser Lys Asn Met Arg Trp Ser Pro Arg Ile Lys Phe Leu Asp
 1765 1770 1775
 Leu Cys Val Leu Ile Asp Ile Asn Ser Glu Ser Leu Ser Leu Ile Ser
 1780 1785 1790
 His Val Val Lys Trp Lys Arg Asp Glu His Tyr Thr Val Leu Phe Ser
 1795 1800 1805
 Asp Leu Ala Asn Ser His Gln Arg Ser Asp Ser Ser Leu Val Asp Glu
 1810 1815 1820
 Phe Val Val Ser Thr Arg Asp Val Cys Lys Asn Phe Leu Lys Gln Val
 1825 1830 1835 1840
 Tyr Phe Glu Ser Phe Val Arg Glu Phe Val Ala Thr Thr Arg Thr Leu
 1845 1850 1855
 Gly Asn Phe Ser Trp Phe Pro His Lys Glu Met Met Pro Ser Glu Asp
 1860 1865 1870
 Gly Ala Glu Ala Leu Gly Pro Phe Gln Ser Phe Val Ser Lys Val Val
 1875 1880 1885
 Asn Lys Asn Val Glu Arg Pro Met Phe Arg Asn Asp Leu Gln Phe Gly
 1890 1895 1900
 Phe Gly Trp Phe Ser Tyr Arg Met Gly Asp Val Val Cys Asn Ala Ala
 1905 1910 1915 1920
 Met Leu Ile Arg Gln Gly Leu Thr Asn Pro Lys Ala Phe Lys Ser Leu
 1925 1930 1935
 Lys Asp Leu Trp Asp Tyr Met Leu Asn Tyr Thr Lys Gly Val Leu Glu
 1940 1945 1950
 Phe Ser Ile Ser Val Asp Phe Thr His Asn Gln Asn Asn Thr Asp Cys
 1955 1960 1965
 Leu Arg Lys Phe Ser Leu Ile Phe Leu Val Arg Cys Gln Leu Gln Asn
 1970 1975 1980
 Pro Gly Val Ala Glu Leu Leu Ser Cys Ser His Leu Phe Lys Gly Glu
 1985 1990 1995 2000
 Ile Asp Arg Arg Met Leu Asp Glu Cys Leu His Leu Leu Arg Thr Asp
 2005 2010 2015
 Ser Val Phe Lys Val Asn Asp Gly Val Phe Asp Ile Arg Ser Glu Glu
 2020 2025 2030
 Phe Glu Asp Tyr Met Glu Asp Pro Leu Ile Leu Gly Asp Ser Leu Glu
 2035 2040 2045
 Leu Glu Leu Leu Gly Ser Lys Arg Ile Leu Asp Gly Ile Arg Ser Ile
 2050 2055 2060
 Asp Phe Glu Arg Val Gly Pro Glu Trp Glu Pro Val Pro Leu Thr Val
 2065 2070 2075 2080

Lys Met Gly Ala Leu Phe Glu Gly Arg Asn Leu Val Gln Asn Ile Ile
 2085 2090 2095
 Val Lys Leu Glu Thr Lys Asp Met Lys Val Phe Leu Ala Gly Leu Glu
 2100 2105 2110
 Gly Tyr Glu Lys Ile Ser Asp Val Leu Gly Asn Leu Phe Leu His Arg
 2115 2120 2125
 Phe Arg Thr Gly Glu His Leu Leu Gly Ser Glu Ile Ser Val Ile Leu
 2130 2135 2140
 Gln Glu Leu Cys Ile Asp Arg Ser Ile Leu Leu Ile Pro Leu Ser Leu
 2145 2150 2155 2160
 Leu Pro Asp Trp Phe Ala Phe Lys Asp Cys Arg Leu Cys Phe Ser Lys
 2165 2170 2175
 Ser Arg Ser Thr Leu Met Tyr Glu Ile Val Gly Gly Arg Phe Arg Leu
 2180 2185 2190
 Lys Gly Arg Ser Cys Asp Asp Trp Leu Gly Gly Ser Val Ala Glu Asp
 2195 2200 2205
 Ile Asp
 2210

(2) ANGABEN ZU SEQ ID NO: 8:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 95 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "unknown protein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8:

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

(2) ANGABEN ZU SEQ ID NO: 9:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 4695 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 9:

CCCCGGCTGG	GCTGAGACCC	GCAGAGGAAG	ACGCTCTAGG	GATTTGTCCC	GGACTAGCGA	60
GATGGCAAGG	CTGAGGACGG	GAGGCTGATT	GAGAGGCGAA	GGTACACCCT	AACTCAATA	120
CAACCTTTGG	AGCTAAGCCA	GCAATGGTAG	AGGGAAGATT	CTGCACGTCC	CTTCCAGGCG	180
GCCTCCCCGT	CACCACCCCC	CCCAACCCGC	CCCACCAGGA	GCTGAGAGTA	ATTCATACAA	240
AAGGACTCGC	CCCTGCCTTG	GGGAATCCCA	GGGACCGTCG	TTAAACTCCC	ACTAACGTAG	300
AACCCAGAGA	TCGCTGCGTT	CCCGCCCCCT	CACCCGCCCG	CTCTCGTCAT	CACTGAGGTG	360
GAGAAGAGCA	TGCGTGAGGC	TCCGGTGCCC	GTCAGTGGGC	AGAGCGCACA	TCGCCACAG	420
TCCCCGAGAA	GTTGGGGGGA	GGGTTCGGCA	ATTGAACCGG	TGCCTAGAGA	AGGTGGCGCG	480
GGGTAAACTG	GGAAAGTGAT	GTCGTGTA	GGCTCCGCTT	TTTTCCCGAG	GGTGGGGGAG	540
AACCGTATAT	AAGTGCAGTA	GTCGCCGTGA	ACGTTCTTTT	TCGCAACGGG	TTTGCCGCCA	600
GAACACAGGT	AAGTGCCTTG	TGTGTTTCCC	GCGGGCCTGG	CCTCTTTACG	GGTTATGGCC	660
CTTCGGTGCC	TTGAATTACT	TCCACGCCCC	TGGCTGCAGT	ACGTGATTCT	TGATCCCAG	720
CTTCGGGTTG	GAAGTGGGTG	GGAGAGTTCC	AGGCCTTCGG	CCTAAGGAGC	CCCTTCGCTT	780
CGTGCTTGAG	TTGAGGCCCTG	GCCTGGGCGC	TGGGGCCGCC	GCGTGCGAAT	CTGGTGGCAC	840
CTTCGGCGCT	GTCTCGCTGC	TTTCGATAAG	TCTCTAGCCA	TTTAAAATTT	TTGATGACCT	900
GCTGCGACGC	TTTTTTTCTG	GCAAGATAGT	CTTGTAATG	CGGGCCAAGA	TCTGCACACT	960
GGTATTTCCG	TTTTTGGGGC	CGCGGGCGGC	GACGGGGCCC	GTGCGTCCCA	GCGCACATGT	1020
TCGGCGAGGC	GGGGCCTGCG	AGCGCGGCCA	CCGAGAATCG	GACGGGGGTA	GTCTCAAGCT	1080
GGCCGGCCTG	CTCTGGTGCC	TGGCCTCGCG	CCGCCGTGTA	TCGCCCGCC	CTGGGCGGCA	1140
AGGCTGGCCC	GGTCGGCACC	AGTTGCGTGA	CGGAAAGAT	GGCCGCTTCC	CGGCCCTGCT	1200
GCAGGGAGCT	CAAAATGGAG	GACGCGGCGC	TCGGGAGAGC	GGGCGGGTGA	GTCACCCACA	1260
CAAAGGAAAA	GGGCCTTTCC	GTCTCAGCC	GTGCTTCAT	GTGACTCCAC	GGAGTACCGG	1320
GCGCCGTCCA	GGCACCTCGA	TTAGTTCTCG	AGCTTTTGA	GTACGTCGTC	TTTAGGTTGG	1380
GGGGAGGGGT	TTTATGCGAT	GGAGTTTCCC	CACACTGAGT	GGGTGGAGAC	TGAAGTTAGG	1440
CCAGCTTGCC	ACTTGATGTA	ATTCTCCTTG	GAATTTGCC	TTTTGAGTT	TGGATCTTGG	1500
TTCATTCTCA	AGCCTCAGAC	AGTGGTTCAA	AGTTTTTTTC	TTCCATTTC	GGTGTCTGTA	1560
AACTACCCC	TAAAAGCCAA	AATGGGAAAG	GAAAAGACTC	ATATCAACAT	TGTCGTCAAT	1620
GGACACGTAG	ATTCCGGCAA	GTCCACCACT	ACTGGCCATC	TGATCTATAA	ATGCGGTGGC	1680
ATCGACAAAA	GAACCACTGA	AAAATTTGAG	AAGGAGGCTG	CTGAGGTATG	TTAATACCA	1740
GAAAGGAAA	GATCAACTAA	AATGAGTTTT	ACCAGCAGAA	TCATTAGGTG	ATTTCCCCAG	1800
AACTAGTGAG	TGGTTTAGAT	CTGAATGCTA	ATAGTTAAGA	CCTTACTTAT	GAAATAATTT	1860
TGCTTTTGGT	GACTTCTGTA	ATCGTATTGC	TAGTGAGTAG	ATTTGGATGT	TAATAGTTAA	1920
GATCCTACTT	ATAAAAGTTT	GATTTTTGGT	TGCTTCTGTA	ACCCAAAGTG	ACCAAAATCA	1980
CTTTGGACTT	GGAGTTGTAA	AGTGGAAACT	GCCAATTAAG	GGCTGGGGAC	AAGGAAATFG	2040
AAGCTGGAGT	TTGTGTTTTA	GTAACCAAGT	AACGACTCTT	AATCCTTACA	GATGGGAAAG	2100
GGCTCCTTCA	AGTATGCCTG	GGTCTTGGAT	AACTGAAAG	CTGAGCGTGA	ACGTGGTATC	2160
ACCATTGATA	TCTCCTTG TG	GAAATTTGAG	ACCAGCAAGT	ACTATGTGAC	TATCATTTGAT	2220
GCCCCAGGAC	ACAGAGACTT	TATCAAAAAC	ATGATTACAG	GGACATCTCA	GGTTGGGATT	2280
AATAATTTCTA	GGTTTCTTTA	TCCCAAAGG	CTTGCTTTGT	ACACTGGTTT	TGTCATTTGG	2340
AGAGTTGACA	GGGATATGTC	TTTGCTTTCT	TTAAAGGCTG	ACTGTGCTGT	CCTGATTGTT	2400
GCTGCTGGTG	TTGGTGAATT	TGAAGCTGGT	ATCTCCAAGA	ATGGGCAGAC	CCGAGAGCAT	2460
GCCCTTTCTG	CTTACACACT	GGGTGTGAAA	CAACTAATTG	TCGGTGTTAA	CAAAATGGAT	2520
TCCACTGAGC	CACCTACAG	CCAGAAGAGA	TATGAGAAA	TTGTTAAGGA	AGTCAGCACT	2580
TACATTAAGA	AAATTTGGCTA	CAACCCCGAC	ACAGTAGCAT	TTGTGCCAAT	TTCTGGTTGG	2640
AATGGTGACA	ACATGCTGGA	GCCAAGTGCT	AACGTAAGTG	GCTTTCAAGA	CCATTGTTAA	2700
AAAGCTCTGG	GAATGGCGAT	TTCATGCTTA	CACAAATTGG	CATGCTTG TG	TTTCAGATGC	2760
CTTGGTTCAA	GGGATGGAAA	GTCACCCGTA	AGGATGGCAA	TGCCAGTGGA	ACCACGCTGC	2820
TTGAGGCTCT	GGACTGCATC	CTACCACCAA	CTCGTCCAAC	TGACAAGCCC	TTGCGCCTGC	2880
CTCTCCAGGA	TGTCTACAAA	ATTGGTGTA	AGTTGGCTGT	AAACAAAGTT	GAATTTGAGT	2940
TGATAGAGTA	CTGTCTGCCT	TCATAGGTAT	TTAGTATGCT	GTAATATTTT	TTAGGTATTG	3000
GTAAGTTTCC	TGTTGGCCGA	GTGGAGACTG	GTGTTCTCAA	ACCCGGTATG	GTGGTACCTT	3060
TTGCTCCAGT	CAACGTTACA	ACGGAAGTAA	AATCTGTGCA	AATGCACCAT	GAAGCTTTGA	3120
GTGAAGCTCT	TCCTGGGGAC	AATGTGGGCT	TCAATGTCAA	GAATGTGTCT	GTC AAGGATG	3180
TTCGTCTGTTG	CAACGTTGCT	GGTGACAGCA	AAAATGACCC	ACCAATGGAA	GCAGCTGGCT	3240
TCACTGCTCA	GGTAACAATT	TAAAGTAACA	TAACTTATT	GCAGAGGCTA	AAGTCATTTG	3300
AGACTTTGGA	TTTGCCTGTA	ATGCAAATCT	TTTTTCCAAG	GTGATTATCC	TGAACCATCC	3360
AGGCCAAATA	AGCGCCGGCT	ATGCCCCCTGT	ATTGGATTGC	CACACGGCTC	ACATTGCATG	3420
CAAGTTTGCT	GAGCTGAAGG	AAAAGATTGA	TCGCCGTTCT	GGTAAAAAGC	TGGAAGATGG	3480
CCCTAAATTC	TTGAAGTCTG	GTGATGCTGC	CATTGTTGAT	ATGGTTCCTG	GCAAGCCCAT	3540
GTGTGTTGAG	AGCTTCTCAG	ACTATCCACC	TTTGGGTAAG	GATGACTACT	TAAATGTA	3600
AAAGTTGTGT	TAAAGATGAA	AAATACAAC	GAACAGTACT	TTGGGTAATA	ATTAACTTT	3660
TTTTAATAG	GTCGCTTTGC	TGTTCTGAT	ATGAGACAGA	CAGTTGCGGT	GGGTGTCATC	3720

AAAGCAGTGG	ACAAGAAGGC	TGCTGGAGCT	GGCAAGGTCA	CCAAGTCTGC	CCAGAAAGCT	3780
CAGAAGGCTA	AATGAATATT	ATCCCTAATA	CCTGCCACCC	CACTCTTAAT	CAGTGGTGGA	3840
AGAACGGTCT	CAGAACTGTT	TGTTTCAATT	GGCCATTTAA	GTTTAGTAGT	AAAAGACTGG	3900
TTAATGATAA	CAATGCATCG	TAAAACCTTC	AGAAGGAAAG	GAGAATGTTT	TGTGGACCAC	3960
TTTGGTTTTT	TTTTTTGCGT	GTGGCAGTTT	TAAGTTATTA	GTTTTTAAAA	TCAGTACTTT	4020
TTAATGGAAA	CAACTTGACC	AAAAATTTGT	CACAGAATTT	TGAGACCCAT	TAAAAAAGTT	4080
AAATGAGAAA	CCTGTGTGTT	CCTTTGGTCA	ACACCGAGAC	ATTTAGGTGA	AAGACATCTA	4140
ATTCTGGTTT	TACGAATCTG	GAAACTTCTT	GAAAATGTAA	TTCTTGAGTT	AACACTTCTG	4200
GGTGGAGAAT	AGGGTTGTTT	TCCCCCACA	TAATTGGAAG	GGGAAGGAAT	ATCATTTAAA	4260
GCTATGGGAG	GGTTTCTTTG	ATTACAACAC	TGGAGAGAAA	TGCAGCATGT	TGCTGATTGC	4320
CTGTCACTAA	AACAGGCCAA	AAACTGAGTC	CCTGGGTTGC	ATAGAAAGCT	TCATGTTGCT	4380
AAACCAATGT	TAAGTGAATC	TTTGAAACA	AAATGTTTCC	AAATTACTGG	GATGTGCATG	4440
TTGAAACGTG	GGTAAAAATG	ACTGGGCAGT	GAAAGTTGAC	TATTTGCCAT	GACATAAGAA	4500
ATAAGTGTAG	TGGCTAGTGT	ACACCCTATG	AGTGGAAGGG	TCCATTTTGA	AGTCAGTGGA	4560
GTAAGCTTTA	TGCCATTTTG	ATGGTTTCAC	AAGTTCTATT	GAGTGCTATT	CAGAATAGGA	4620
ACAAGGTTCT	AATAGAAAAA	GATGGCAATT	TGAAGTAGCT	ATAAAATTAG	ACTAATTACA	4680
TTGCTTTTCT	CCGAC					4695

(2) ANGABEN ZU SEQ ID NO: 10:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 462 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "elongation factor EF-1-alpha"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:

Met	Gly	Lys	Glu	Lys	Thr	His	Ile	Asn	Ile	Val	Val	Ile	Gly	His	Val
1				5					10					15	
Asp	Ser	Gly	Lys	Ser	Thr	Thr	Thr	Gly	His	Leu	Ile	Tyr	Lys	Cys	Gly
			20					25					30		
Gly	Ile	Asp	Lys	Arg	Thr	Ile	Glu	Lys	Phe	Glu	Lys	Glu	Ala	Ala	Glu
		35					40					45			
Met	Gly	Lys	Gly	Ser	Phe	Lys	Tyr	Ala	Trp	Val	Leu	Asp	Lys	Leu	Lys
	50					55					60				
Ala	Glu	Arg	Glu	Arg	Gly	Ile	Thr	Ile	Asp	Ile	Ser	Leu	Trp	Lys	Phe
65					70					75					80
Glu	Thr	Ser	Lys	Tyr	Tyr	Val	Thr	Ile	Ile	Asp	Ala	Pro	Gly	His	Arg
				85					90					95	
Asp	Phe	Ile	Lys	Asn	Met	Ile	Thr	Gly	Thr	Ser	Gln	Ala	Asp	Cys	Ala
			100					105					110		
Val	Leu	Ile	Val	Ala	Ala	Gly	Val	Gly	Glu	Phe	Glu	Ala	Gly	Ile	Ser
			115					120					125		
Lys	Asn	Gly	Gln	Thr	Arg	Glu	His	Ala	Leu	Leu	Ala	Tyr	Thr	Leu	Gly
	130					135						140			
Val	Lys	Gln	Leu	Ile	Val	Gly	Val	Asn	Lys	Met	Asp	Ser	Thr	Glu	Pro
145					150					155					160
Pro	Tyr	Ser	Gln	Lys	Arg	Tyr	Glu	Glu	Ile	Val	Lys	Glu	Val	Ser	Thr
				165					170					175	
Tyr	Ile	Lys	Lys	Ile	Gly	Tyr	Asn	Pro	Asp	Thr	Val	Ala	Phe	Val	Pro
			180					185					190		
Ile	Ser	Gly	Trp	Asn	Gly	Asp	Asn	Met	Leu	Glu	Pro	Ser	Ala	Asn	Met
		195					200						205		

Pro Trp Phe Lys Gly Trp Lys Val Thr Arg Lys Asp Gly Asn Ala Ser
 210 215 220
 Gly Thr Thr Leu Leu Glu Ala Leu Asp Cys Ile Leu Pro Pro Thr Arg
 225 230 235
 Pro Thr Asp Lys Pro Leu Arg Leu Pro Leu Gln Asp Val Tyr Lys Ile
 245 250 255
 Gly Gly Ile Gly Thr Val Pro Val Gly Arg Val Glu Thr Gly Val Leu
 260 265 270
 Lys Pro Gly Met Val Val Thr Phe Ala Pro Val Asn Val Thr Thr Glu
 275 280 285
 Val Lys Ser Val Glu Met His His Glu Ala Leu Ser Glu Ala Leu Pro
 290 295 300
 Gly Asp Asn Val Gly Phe Asn Val Lys Asn Val Ser Val Lys Asp Val
 305 310 315
 Arg Arg Gly Asn Val Ala Gly Asp Ser Lys Asn Asp Pro Pro Met Glu
 325 330 335
 Ala Ala Gly Phe Thr Ala Gln Val Ile Ile Leu Asn His Pro Gly Gln
 340 345 350
 Ile Ser Ala Gly Tyr Ala Pro Val Leu Asp Cys His Thr Ala His Ile
 355 360 365
 Ala Cys Lys Phe Ala Glu Leu Lys Glu Lys Ile Asp Arg Arg Ser Gly
 370 375 380
 Lys Lys Leu Glu Asp Gly Pro Lys Phe Leu Lys Ser Gly Asp Ala Ala
 385 390 395 400
 Ile Val Asp Met Val Pro Gly Lys Pro Met Cys Val Glu Ser Phe Ser
 405 410 415
 Asp Tyr Pro Pro Leu Gly Arg Phe Ala Val Arg Asp Met Arg Gln Thr
 420 425 430
 Val Ala Val Gly Val Ile Lys Ala Val Asp Lys Lys Ala Ala Gly Ala
 435 440 445
 Gly Lys Val Thr Lys Ser Ala Gln Lys Ala Gln Lys Ala Lys
 450 455 460

(2) ANGABEN ZU SEQ ID NO: 11:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 8332 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:

TATGCGCCTG	CGTCGGTACT	AGTTAGCTAA	CTAGCTCTGT	ATCTGGCGGA	CCCGTGGTGG	60
AACTGACGAG	TTCGGAACAC	CCGGCCGCAA	CCCTGGGAGA	CGTCCCAGGG	ACTTCGGGGG	120
CCGTTTTTGT	GGCCCCGACCT	GAGTCCAAAA	ATCCCGATCG	TTTTGGACTC	TTTGGTGCAC	180
CCCCCTTAGA	GGAGGGATAT	GTGGTTCTGG	TAGGAGACGA	GAACCTAAAA	CAGTTCCCGC	240
CTCCGTCTGA	ATTTTTGCTT	TCGGTTTGGG	ACCGAAGCCG	CGCCGCGCGT	CTTGTCTGCT	300
GCAGCATCGT	TCTGTGTTGT	CTCTGTCTGA	CTGTGTTTCT	GTATTTGTCT	GAGAATATGG	360
GCCAGACTGT	TACCACTCCC	TTAAGTTTGA	CCTTAGGTCA	CTGGAAAGAT	GTCGAGCGGA	420
TCGCTCACAA	CCAGTCGGTA	GATGTCAAGA	AGAGACGTTG	GGTTACCTTC	TGCTCTGCAG	480
AATGGCCAAC	CTTTAACGTC	GGATGGCCGC	GAGACGGCAC	CTTTAACCGA	GACCTCATCA	540
CCCAGGTTAA	GATCAAGGTC	TTTTCACCTG	GCCCGCATGG	ACACCCAGAC	CAGGTCCCCT	600
ACATCGTGAC	CTGGGAAGCC	TTGGCTTTTG	ACCCCCCTCC	CTGGGTCAAG	CCCTTTGTAC	660
ACCCTAAGCC	TCCGCCTCCT	CTTCCTCCAT	CCGCCCCGTC	TCTCCCCCTT	GAACCTCCTC	720
GTTGACCCC	GCCTCGATCC	TCCCTTTATC	CAGCCCTCAC	TCCTTCTCTA	GGCGCCAAAC	780
CTAAACCTCA	AGTTCTTTCT	GACAGTGGGG	GGCCGCTCAT	CGACCTACTT	ACAGAAGACC	840
CCCCGCCTTA	TAGGGACCCA	AGACCACCCC	CTTCCGACAG	GGACGGAAAT	GGTGGAGAAG	900

CGACCCCTGC	GGGAGAGGCA	CCGGACCCCT	CCCCAATGGC	ATCTCGCCTA	CGTGGGAGAC	960
GGGAGCCCC	TGTGGCCGAC	TCCACTACCT	CGCAGGCATT	CCCCCTCCGC	GCAGGAGGAA	1020
ACGGACAGCT	TCAATACTGG	CCGTTCTCCT	CTTCTGACCT	TTACAACCTGG	AAAAATAATA	1080
ACCCCTCTTT	TTCTGAAGAT	CCAGGTA AAC	TGACAGCTCT	GATCGAGTCT	GTTCTCATCA	1140
CCCATCAGCC	CACCTGGGAC	GACTGTGAGC	AGCTGTTGGG	GACTCTGCTG	ACCGGAGAAG	1200
AAAAACAACG	GGTGCTCTTA	GAGGCTAGAA	AGGCGGTGCG	GGGCGATGAT	GGGCGCCCCA	1260
CTCAACTGCC	CAATGAAGTC	GATGCCGCTT	TTCCCCTCGA	GCGCCCAGAC	TGGGATTACA	1320
CCACCCAGGC	AGGTAGGAAC	CACCTAGTCC	ACTATCGCCA	GTTGCTCCTA	GCGGGTCTCC	1380
AAAACGCGGG	CAGAAGCCCC	ACCAATTTGG	CCAAGGTAAA	AGGAATAACA	CAAGGGCCCA	1440
ATGAGTCTCC	CTCGGCCTTC	CTAGAGAGAC	TTAAGGAAGC	CTATCGCAGG	TACTACTCCTT	1500
ATGACCCTGA	GGACCCAGGG	CAAGAAACTA	ATGTGTCTAT	GTCTTTTCATT	TGGCAGTCTG	1560
CCCCAGACAT	TGGGAGAAAG	TTAGAGAGGT	TAGAAGATTT	AAAAACAAG	ACGCTTGGAG	1620
ATTTGGTTAG	AGAGGCAGAA	AAGATCTTTA	ATAAACGAGA	AACCCCGGAA	GAAAGAGAGG	1680
AACGTATCAG	GAGAGAAACA	GAGGAAAAAG	AAGAACGCCC	TAGGACAGAG	GATGAGCAGA	1740
AAGAGAAAGA	AAGAGATCGT	AGGAGACATA	GAGAGATGAG	CAAGCTATTG	GCCACTGTCC	1800
TTAGTGGACA	GAACAGGAT	AGACAGGGAG	GAGAACGAAG	GAGGTCCCAA	CTCGATCGCG	1860
ACCAAGTGTG	CTACTGCAAA	AAAAAGGGGC	ACTGGGCTAA	AGATTGTCCC	AAGAAACCAC	1920
GAGGACCTCG	GGGACCAAGA	CCCCAGACCT	CCCTCCTGAC	CCTAGATGAC	TAGGGAGGTC	1980
AGGGTCAGGA	GGCCCCCCTT	GAACCCAGGA	TAACCCCTCAA	AGTCGGGGGG	CAACCCGTCA	2040
CCTTCCTGGT	AGATACTGGG	GCCCAACACT	CCGTGCTGAC	CCAAAATCCT	GGACCCCTAA	2100
GTGATAAGTC	TGCCTGGGTC	CAAGGGGCTA	CTGGAGGAAA	GCGGTATCGC	TGGACCACGG	2160
ATCGCAAAGT	GCATCTAGCT	ACCGGTAAGG	TCACCCACTC	TTTCCTCCAT	GTACCAGACT	2220
GTCCCTATCC	TCTGTTAGGA	AGAGATTTGC	TGACTAAACT	AAAAGCCCAA	ATCCACTTTG	2280
AGGGATCAGG	AGCTCAGGTT	ATGGGACCAA	TGGGGCAGCC	CCTGCAAGTG	TTGACCCTAA	2340
ATATAGAAGA	TGAGCATCGG	CTACATGAGA	CCTCAAAAAGA	GCCAGATGTT	TCTCTAGGGT	2400
CCACATGGCT	GTCTGATTTT	CCTCAGGCCT	GGGCGGAAAC	CGGGGGCATG	GGACTGGCAG	2460
TTCGCCAACG	TCCTCTGATC	ATACCTCTGA	AAGCAACCTC	TACCCCGTGT	TCCATAAAAC	2520
AATACCCCAT	GTCACAAGAA	GCCAGACTGG	GGATCAAGCC	CCACATACAG	AGACTGTTGG	2580
ACCAGGGAAT	ACTGGTACCC	TGCCAGTCCC	CCTGGAACAC	GCCCCTGCTA	CCCGTTAAGA	2640
AACCAGGGAC	TAATGATTAT	AGGCCTGTCC	AGGATCTGAG	AGAAGTCAAC	AAGCGGGTGG	2700
AAGACATCCA	CCCCACCGTG	CCCAACCCTT	ACAACCTCTT	GAGCGGGCTC	CCACCGTCCC	2760
ACCAGTGGTA	CACTGTGCTT	GATTTAAAGG	ATGCCTTTTT	CTGCCTGAGA	CTCCACCCCA	2820
CCAGTCAGCC	TCTCTTCGCT	TTTGAGTGGA	GAGATCCAGA	GATGGGAATC	TCAGGACAAT	2880
TGACCTGGAC	CAGACTCCCA	CAGGGTTTCA	AAAACAGTCC	CACCCTGTTT	GATGAGGCAC	2940
TGCACAGAGA	CCTAGCAGAC	TTCCGGATCC	AGCACCCAGA	CTTGATCCTG	CTACAGTACG	3000
TGGATGACTT	ACTGCTGGCC	GCCACTTCTG	AGCTAGACTG	CCAACAAGGT	ACTCGGGCCC	3060
TGTTACAAAC	CCTAGGGAAC	CTCGGGTATC	GGGCCTCGGC	CAAGAAAAGCC	CAAATTTGCC	3120
AGAAACAGGT	CAAGTATCTG	GGGTATCTTC	TAAAAGAGGG	TCAGAGATGG	CTGACTGAGG	3180
CCAGAAAAGA	GACTGTGATG	GGGCAGCCTA	CTCCGAAGAC	CCCTCGACAA	CTAAGGGAGT	3240
TCCTAGGGAC	GGCAGGCTTC	TGTCGCCTCT	GGATCCCTGG	GTTTGCAGAA	ATGGCAGCCC	3300
CCTTGTAACC	TCTCACCAAA	ACGGGGACTC	TGTTTAAATTG	GGGCCAGAC	CAACAAAAGG	3360
CCTATCAAGA	AATCAAGCAA	GCTCTTCTAA	CTGCCCCAGC	CCTGGGGTTG	CCAGATTTGA	3420
CTAAGCCCTT	TGAACTCTTT	GTCGACGAGA	AGCAGGGCTA	CGCCAAAGGT	GTCCCTAACG	3480
AAAAACTGGG	ACCTTGGCGT	CGGCCGGTGG	CCTACCTGTC	CAAAAAGCTA	GACCCAGTAG	3540
CAGCTGGGTG	GCCCCCTTGC	CTACGGATGG	TAGCAGCCAT	TGCCGTACTG	ACAAAAGGATG	3600
CAGGCAAGCT	AACCATGGGA	CAGCCACTAG	TCATTCTGGC	CCCCCATGCA	GTAGAGGCAC	3660
TAGTCAAACA	ACCCCCCGAC	CGCTGGCTTT	CCAACGCCCG	GATGACTCAC	TATCAGGCCT	3720
TGCTTTTGG	CACGGACCGG	GTCCAGTTCC	GACCGGTGGT	AGCCCTGAAC	CCGGCTACGC	3780
TGCTCCCACT	GCCTGAGGAA	GGGCTGCAAC	ACAACCTGCCT	TGATATCCTG	GCCGAAGCCC	3840
ACGGAACCCG	ACCCGACCTA	ACGGACCAGC	CGCTCCGAGA	CGCCGACCAC	ACCTGGTACA	3900
CGGATGGAAG	CAGTCTCTTA	CAAGAGGGAC	AGCGTAAGGC	GGGAGCTGCG	GTGACCACCG	3960
AGACCGAGGT	AATCTGGGCT	AAAGCCCTGC	CAGCCGGGAC	ATCCGCTCAG	CGGGCTGAAC	4020
TGATAGCACT	CACCCAGGCC	CTAAAGATGG	CAGAAGGTAA	GAAGCTAAAT	GTTTATACTG	4080
ATAGCCGTTA	TGCTTTTGCT	ACTGCCATA	TCCATGGAGA	AATATACAGA	AGGCGTGGGT	4140
TGCTCACATC	AGAAGGCAAA	GAGATCAAAA	ATAAAGACGA	GATCTTGGCC	CTACTAAAAG	4200
CCCTCTTTCT	GCCCCAAAAGA	CTTAGCATAA	TCCATTGTCC	AGGACATCAA	AAGGGACACA	4260
GCGCCGAGGC	TAGAGGCAAC	CGGATGGCTG	ACCAAGCGGC	CCGAAAAGGCA	GCCATCACAG	4320
AGACTCCAGA	CACCTCTACC	CTCCTCATAG	AAAATTCATC	ACCCTACACC	TCAGAACATT	4380
TTCATTACAC	AGTGA CTGAT	ATAAAGGACC	TAACCAAGTT	GGGGGCCATT	TATGATAAAA	4440
CAAAGAAGTA	TTGGGTCTAC	CAAGGAAAAC	CTGTGATGCC	TGACCAGTTT	ACTTTTGAAT	4500
TATTAGACTT	TCTTCATCAG	CTGACTCACC	TCAGTTCTC	AAAAATGAAG	GCTCTCCTAG	4560
AGAGAAGCCA	CAGTCCCTAC	TACATGCTGA	ACCGGATCG	AACACTCAA	AATATCACTG	4620
AGACCTGCAA	AGCTTGTCGA	CAAGTCAACG	CCAGCAAGTC	TGCCGTAAAA	CAGGGA ACTA	4680
GGGTCCGCGG	GCATCGGCCC	GGCACTCATT	GGGAGATCGA	TTTCACCGAG	ATAAAGCCCC	4740

GATTGTATGG	CTATAAATAT	CTTCTAGTTT	TTATAGATAC	CTTTTCTGGC	TGGATAGAAG	4800
CCTTCCCAAC	CAAGAAAGAA	ACCGCCAAGG	TCGTAACCAA	GAAGCTACTA	GAGGAGATCT	4860
TCCCCAGGTT	CGGCATGCCT	CAGGTATTGG	GAACTGACAA	TGGGCCTGCC	TTCGTCTCCA	4920
AGGTGAGTCA	GACAGTGGCC	GATCTGTTGG	GGATTGATTG	GAAATTACAT	TGTGCATACA	4980
GACCCCAAAG	CTCAGGCCAG	GTAGAAAGAA	TGAATAGAAC	CATCAAGGAG	ACTTTAACTA	5040
AATTAACGCT	TGCAACTGGC	TCTAGAGACT	GGGTGCTCCT	ACTCCCCTTA	GCCCTGTACC	5100
GAGCCCAGAA	CACGCCGGGC	CCCCATGGCC	TCACCCCAT	TGAGATCTTA	TATGGGGCAC	5160
CCCCGCCCTC	TGTAAACTTC	CCTGACCCTG	ACATGACAAG	AGTTACTAAC	AGCCCCCTCT	5220
TCCAAGCTCA	CTTACAGGCT	CTCTACTTAG	TCCAGCACGA	AGTCTGGAGA	CCTCTGGCGG	5280
CAGCCTACCA	AGAACAACCTG	GACCGACCGG	TGGTACCTCA	CCCTTACCGA	GTCGGCGACA	5340
CAGTGTGGGT	CCGCCGACAC	CAGACTAAGA	ACCTAGAACC	TCGCTGGAAA	GGACCTTACA	5400
CAGTCTTGCT	GACCACCCCC	ACCGCCCTCA	AAGTAGACGG	CATCGCAGCT	TGGATACACG	5460
CCGCCACGCT	GAAGGCTGCC	GACCCCGGGG	GTGGACCATC	CTCTAGACTG	ACATGGCGCG	5520
TTCAACGCTC	TCAAAAACCC	TTAAAAATAA	GGTTAACCCG	CGAGGCCCCC	TAATCCCCTT	5580
AATTCTTCTG	ATGCTCAGAG	GGTTCAGTAC	TGCTTCGCCC	GGCTCCAGTC	CTCATCAAGT	5640
CTATAATATC	ACCTGGGAGG	TAACCAATGG	AGATCGGGAG	ACGGTATGGG	CAACTTCTGG	5700
CAACCACCCT	CTGTGGACCT	GGTGGCCTGA	CCTTACCCCA	GATTTATGTA	TGTTAGCCCA	5760
CCATGGACCA	TCTTATTGGG	GGCTAGAATA	TCAATCCCCT	TTTTCTTCTC	CCCCGGGGCC	5820
CCCTTGTTGC	TCAGGGGGCA	GCAGCCCAGG	CTGTTCCAGA	GACTGCGAAG	AACCTTTAAC	5880
CTCCCTCAAC	CCTCGGTGCA	ACACTGCCTG	GAACAGACTC	AAGCTAGACC	AGACAACCTA	5940
TAAATCAAAT	GAGGGATTTT	ATGTTTGCCC	CGGGCCCCAC	CGCCCCGAG	AATCCAAGTC	6000
ATGTGGGGGT	CCAGACTCCT	TCTACTGTGC	CTATTGGGGC	TGTGAGACAA	CCGGTAGAGC	6060
TTACTGGAAG	CCCTCCTCAT	CATGGGATTT	CATCACAGTA	AACAACAATC	TCACCTCTGA	6120
CCAGGCTGTC	CAGGTATGCA	AAGATAATAA	GTGGTGCAAC	CCCTTAGTTA	TTCGGTTTAC	6180
AGACGCCGGG	AGACGGGTTA	CTTCCTGGAC	CACAGGACAT	TACTGGGGCT	TACGTTTGTA	6240
TGTCTCCGGA	CAAGATCCAG	GGCTTACATT	TGGGATCCGA	CTCAGATACC	AAAATCTAGG	6300
ACCCCGCGTC	CCAATAGGGC	CAAACCCCGT	TCTGGCAGAC	CAACAGCCAC	TCTCCAAGCC	6360
CAAACCTGTT	AAGTCGCCTT	CAGTCACCAA	ACCACCCAGT	GGGACTCCTC	TCTCCCCTAC	6420
CCAACCTCCA	CCGGCGGGAA	CGGAAAATAG	GCTGCTAAAC	TTAGTAGACG	GAGCCTACCA	6480
AGCCCTCAAC	CTCACCAGTC	CTGACAAAAC	CCAAGAGTGC	TGGTTGTGTC	TAGTAGCGGG	6540
ACCCCCCTAC	TACGAAGGGG	TTGCCGTCTC	GGGTACCTAC	TCCAACCATA	CCTCTGCTCC	6600
AGCCCACTGC	TCCGTGGCCT	CCCAACACAA	GTTGACCCTG	TCCGAAGTGA	CCGGACAGGG	6660
ACTCTGCATA	GGAGCAGTTT	CCAAAACACA	TCAGGCCCTA	TGTAATACCA	CCCAGACAAG	6720
CAGTCGAGGG	TCCTATTATC	TAGTTGCCCC	TACAGGTACC	ATGTGGGCTT	GTAGTACCGG	6780
GCTTACTCCA	TGCATCTCCA	CCACCATACT	GAACCTTACC	ACTGATTATT	GTGTTCTTGT	6840
CGAACTCTGG	CCAAGAGTCA	CCTATCATTC	CCCCAGCTAT	GTTTACGGCC	TGTTTGAGAG	6900
ATCCAACCGA	CACAAAAGAG	AACCGGTGTC	GTTAACCCCTG	GCCCTATTAT	TGGGTGGACT	6960
AACCATGGGG	GGAATTGCCG	CTGGAATAGG	AACAGGGACT	ACTGCTCTAA	TGGCCACTCA	7020
GCAATTCCAG	CAGTCCAAG	CCGCAGTACA	GGATGATCTC	AGGGAGGTTG	AAAAATCAAT	7080
CTCTAACCTA	GAAAAGTCTC	TCACTTCCCT	GTCTGAAGTT	GTCCTACAGA	ATCGAAGGGG	7140
CCTAGACTTG	TTATTTCTAA	AAGAAGGAGG	GCTGTGTGCT	GCTCTAAAAG	AAGAATGTTG	7200
CTTCTATGCG	GACCACACAG	GACTAGTGAG	AGACAGCATG	GCCAAATTGA	GAGAGAGGCT	7260
TAATCAGAGA	CAGAAACTGT	TTGAGTCAAC	TCAAGGATGG	TTTGAGGGAC	TGTTTAAACAG	7320
ATCCCTTGG	TTTACCATTG	TGATATCTAC	CATTATGGGA	CCCCTCATTG	TACTCCTAAT	7380
GATTTTGCTC	TTCGGACCCT	GCATTCTTAA	TCGATTAGTC	CAATTTGTTA	AAGACAGGAT	7440
ATCAGTGGTC	CAGGCTCTAG	TTTTGACTCA	ACAATATCAC	CAGCTGAAGC	CTATAGAGTA	7500
CGAGCCATAG	ATAAAATAAA	AGATTTTATT	TAGTCTCCAG	AAAAAGGGGG	GAATGAAAGA	7560
CCCCACCTGT	AGGTTTGGCA	AGCTAGCTTA	AGTAACGCCA	TTTTGCAAGG	CATGGAAAAA	7620
TACATAACTG	AGAATAGAGA	AGTTCAGATC	AAGGTCAGGA	ACAGATGGAA	CAGCTGAATA	7680
TGGGCCAAAC	AGGATATCTG	TGGTAAGCAG	TTCTTGCCCC	GGCTCAGGGC	CAAGAACAGA	7740
TGGAACAGCT	GAATATGGGC	CAAACAGGAT	ATCTGTGGTA	AGCAGTTCCCT	GCCCCGGCTC	7800
AGGGCCAAGA	ACAGATGGTC	CCCAGATGCG	GTCCAGCCCT	CAGCAGTTTC	TAGAGAACCA	7860
TCAGATGTTT	CCAGGGTGCC	CCAAGGACCT	GAAATGACCC	TGTGCCTTAT	TTGAACTAAC	7920
CAATCAGTTC	GCTTCTCGCT	TCTGTTCCGG	CGCTTCTGCT	CCCCGAGCTC	AATAAAAGAG	7980
CCCACAACCC	CTCACTCGGG	GCGCCAGTCC	TCCGATTGAC	TGAGTCGCCC	GGGTACCCGT	8040
GTATCCAATA	AACCCCTCTG	CAGTTGCAGC	GCCAGTCCCT	CGATTGACTG	AGTCGCCCGG	8100
GTACCCGTGT	ATCCAATAAA	CCCTCTTGCA	GTTGCATCCG	ACTTGTGGTC	TCGCTGTTCC	8160
TTGGGAGGGT	CTCCTCTGAG	TGATTGACTA	CCCGTCAGCG	GGGTCTTTTC	ATTTGGGGGC	8220
TCGTCCGGGA	TCGGGAGACC	CCTGCCCAGG	GACCACCGAC	CCACCACCGG	GAGGTAAGCT	8280
GGCCAGCAAC	TTATCTGTGT	CTGTCCGATT	GTCTAGTGTC	TATGACTGAT	TT	8332

(2) ANGABEN ZU SEQ ID NO: 12:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 538 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "gag protein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:

```

Met Gly Gln Thr Val Thr Thr Pro Leu Ser Leu Thr Leu Gly His Trp
 1      5      10      15
Lys Asp Val Glu Arg Ile Ala His Asn Gln Ser Val Asp Val Lys Lys
 20      25      30
Arg Arg Trp Val Thr Phe Cys Ser Ala Glu Trp Pro Thr Phe Asn Val
 35      40      45
Gly Trp Pro Arg Asp Gly Thr Phe Asn Arg Asp Leu Ile Thr Gln Val
 50      55      60
Lys Ile Lys Val Phe Ser Pro Gly Pro His Gly His Pro Asp Gln Val
 65      70      75
Pro Tyr Ile Val Thr Trp Glu Ala Leu Ala Phe Asp Pro Pro Pro Trp
 85      90      95
Val Lys Pro Phe Val His Pro Lys Pro Pro Pro Pro Leu Pro Pro Ser
100     105     110
Ala Pro Ser Leu Pro Leu Glu Pro Pro Arg Ser Thr Pro Pro Arg Ser
115
Ser Leu Tyr Pro Ala Leu Thr Pro Ser Leu Gly Ala Lys Pro Lys Pro
130     135     140
Gln Val Leu Ser Asp Ser Gly Gly Pro Leu Ile Asp Leu Leu Thr Glu
145     150     155     160
Asp Pro Pro Pro Tyr Arg Asp Pro Arg Pro Pro Pro Ser Asp Arg Asp
165     170     175
Gly Asn Gly Gly Glu Ala Thr Pro Ala Gly Glu Ala Pro Asp Pro Ser
180     185     190
Pro Met Ala Ser Arg Leu Arg Gly Arg Arg Glu Pro Pro Val Ala Asp
195     200     205
Ser Thr Thr Ser Gln Ala Phe Pro Leu Arg Ala Gly Gly Asn Gly Gln
210     215     220
Leu Gln Tyr Trp Pro Phe Ser Ser Ser Asp Leu Tyr Asn Trp Lys Asn
225     230     235     240
Asn Asn Pro Ser Phe Ser Glu Asp Pro Gly Lys Leu Thr Ala Leu Ile
245     250     255
Glu Ser Val Leu Ile Thr His Gln Pro Thr Trp Asp Asp Cys Gln Gln
260     265     270
Leu Leu Gly Thr Leu Leu Thr Gly Glu Glu Lys Gln Arg Val Leu Leu
275     280     285
Glu Ala Arg Lys Ala Val Arg Gly Asp Asp Gly Arg Pro Thr Gln Leu
290     295     300
Pro Asn Glu Val Asp Ala Ala Phe Pro Leu Glu Arg Pro Asp Trp Asp
305     310     315     320
Tyr Thr Thr Gln Ala Gly Arg Asn His Leu Val His Tyr Arg Gln Leu
325     330     335
Leu Leu Ala Gly Leu Gln Asn Ala Gly Arg Ser Pro Thr Asn Leu Ala
340     345     350
Lys Val Lys Gly Ile Thr Gln Gly Pro Asn Glu Ser Pro Ser Ala Phe
355     360     365
Leu Glu Arg Leu Lys Glu Ala Tyr Arg Arg Tyr Thr Pro Tyr Asp Pro
370     375     380

```

Glu Asp Pro Gly Gln Glu Thr Asn Val Ser Met Ser Phe Ile Trp Gln
 385 390
 Ser Ala Pro Asp Ile Gly Arg Lys Leu Glu Arg Leu Glu Asp Leu Lys
 405 410 415
 Asn Lys Thr Leu Gly Asp Leu Val Arg Glu Ala Glu Lys Ile Phe Asn
 420 425 430
 Lys Arg Glu Thr Pro Glu Glu Arg Glu Glu Arg Ile Arg Arg Glu Thr
 435 440 445
 Glu Glu Lys Glu Glu Arg Arg Thr Glu Asp Glu Gln Lys Glu Lys
 450 455 460
 Glu Arg Asp Arg Arg Arg His Arg Glu Met Ser Lys Leu Leu Ala Thr
 465 470 475 480
 Val Val Ser Gly Gln Lys Gln Asp Arg Gln Gly Gly Glu Arg Arg Arg
 485 490 495
 Ser Gln Leu Asp Arg Asp Gln Cys Ala Tyr Cys Lys Glu Lys Gly His
 500 505 510
 Trp Ala Lys Asp Cys Pro Lys Lys Pro Arg Gly Pro Arg Gly Pro Arg
 515 520 525
 Pro Gln Thr Ser Leu Leu Thr Leu Asp Asp
 530 535

(2) ANGABEN ZU SEQ ID NO: 13:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1737 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "gag-pol protein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:

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

Pro Met Ala Ser Arg Leu Arg Gly Arg Arg Glu Pro Pro Val Ala Asp
 195 200 205
 Ser Thr Thr Ser Gln Ala Phe Pro Leu Arg Ala Gly Gly Asn Gly Gln
 210 215 220
 Leu Gln Tyr Trp Pro Phe Ser Ser Ser Asp Leu Tyr Asn Trp Lys Asn
 225 230 235 240
 Asn Asn Pro Ser Phe Ser Glu Asp Pro Gly Lys Leu Thr Ala Leu Ile
 245 250 255
 Glu Ser Val Leu Ile Thr His Gln Pro Thr Trp Asp Asp Cys Gln Gln
 260 265 270
 Leu Leu Gly Thr Leu Leu Thr Gly Glu Glu Lys Gln Arg Val Leu Leu
 275 280 285
 Glu Ala Arg Lys Ala Val Arg Gly Asp Asp Gly Arg Pro Thr Gln Leu
 290 295 300
 Pro Asn Glu Val Asp Ala Ala Phe Pro Leu Glu Arg Pro Asp Trp Asp
 305 310 315 320
 Tyr Thr Thr Gln Ala Gly Arg Asn His Leu Val His Tyr Arg Gln Leu
 325 330 335
 Leu Leu Ala Gly Leu Gln Asn Ala Gly Arg Ser Pro Thr Asn Leu Ala
 340 345 350
 Lys Val Lys Gly Ile Thr Gln Gly Pro Asn Glu Ser Pro Ser Ala Phe
 355 360 365
 Leu Glu Arg Leu Lys Glu Ala Tyr Arg Arg Tyr Thr Pro Tyr Asp Pro
 370 375 380
 Glu Asp Pro Gly Gln Glu Thr Asn Val Ser Met Ser Phe Ile Trp Gln
 385 390 395 400
 Ser Ala Pro Asp Ile Gly Arg Lys Leu Glu Arg Leu Glu Asp Leu Lys
 405 410 415
 Asn Lys Thr Leu Gly Asp Leu Val Arg Glu Ala Glu Lys Ile Phe Asn
 420 425 430
 Lys Arg Glu Thr Pro Glu Glu Arg Glu Glu Arg Ile Arg Arg Glu Thr
 435 440 445
 Glu Glu Lys Glu Glu Arg Arg Thr Glu Asp Glu Gln Lys Glu Lys
 450 455 460
 Glu Arg Asp Arg Arg Arg His Arg Glu Met Ser Lys Leu Leu Ala Thr
 465 470 475 480
 Val Val Ser Gly Gln Lys Gln Asp Arg Gln Gly Gly Glu Arg Arg Arg
 485 490 495
 Ser Gln Leu Asp Arg Asp Gln Cys Ala Tyr Cys Lys Glu Lys Gly His
 500 505 510
 Trp Ala Lys Asp Cys Pro Lys Lys Pro Arg Gly Pro Arg Gly Pro Arg
 515 520 525
 Pro Gln Thr Ser Leu Leu Thr Leu Asp Asp Gly Gly Gln Gly Gln Glu
 530 535 540
 Pro Pro Pro Glu Pro Arg Ile Thr Leu Lys Val Gly Gly Gln Pro Val
 545 550 555 560
 Thr Phe Leu Val Asp Thr Gly Ala Gln His Ser Val Leu Thr Gln Asn
 565 570 575
 Pro Gly Pro Leu Ser Asp Lys Ser Ala Trp Val Gln Gly Ala Thr Gly
 580 585 590
 Gly Lys Arg Tyr Arg Trp Thr Thr Asp Arg Lys Val His Leu Ala Thr
 595 600 605
 Gly Lys Val Thr His Ser Phe Leu His Val Pro Asp Cys Pro Tyr Pro
 610 615 620
 Leu Leu Gly Arg Asp Leu Leu Thr Lys Leu Lys Ala Gln Ile His Phe
 625 630 635 640
 Glu Gly Ser Gly Ala Gln Val Met Gly Pro Met Gly Gln Pro Leu Gln
 645 650 655
 Val Leu Thr Leu Asn Ile Glu Asp Glu His Arg Leu His Glu Thr Ser
 660 665 670
 Lys Glu Pro Asp Val Ser Leu Gly Ser Thr Trp Leu Ser Asp Phe Pro
 675 680 685
 Gln Ala Trp Ala Glu Thr Gly Gly Met Gly Leu Ala Val Arg Gln Ala
 690 695 700

Pro Leu Ile Ile Pro Leu Lys Ala Thr Ser Thr Pro Val Ser Ile Lys
 705 710 715 720
 Gln Tyr Pro Met Ser Gln Glu Ala Arg Leu Gly Ile Lys Pro His Ile
 725 730 735
 Gln Arg Leu Leu Asp Gln Gly Ile Leu Val Pro Cys Gln Ser Pro Trp
 740 745 750
 Asn Thr Pro Leu Leu Pro Val Lys Lys Pro Gly Thr Asn Asp Tyr Arg
 755 760 765
 Pro Val Gln Asp Leu Arg Glu Val Asn Lys Arg Val Glu Asp Ile His
 770 775 780
 Pro Thr Val Pro Asn Pro Tyr Asn Leu Leu Ser Gly Leu Pro Pro Ser
 785 790 795 800
 His Gln Trp Tyr Thr Val Leu Asp Leu Lys Asp Ala Phe Phe Cys Leu
 805 810 815
 Arg Leu His Pro Thr Ser Gln Pro Leu Phe Ala Phe Glu Trp Arg Asp
 820 825 830
 Pro Glu Met Gly Ile Ser Gly Gln Leu Thr Trp Thr Arg Leu Pro Gln
 835 840 845
 Gly Phe Lys Asn Ser Pro Thr Leu Phe Asp Glu Ala Leu His Arg Asp
 850 855 860
 Leu Ala Asp Phe Arg Ile Gln His Pro Asp Leu Ile Leu Leu Gln Tyr
 865 870 875 880
 Val Asp Asp Leu Leu Leu Ala Ala Thr Ser Glu Leu Asp Cys Gln Gln
 885 890 895
 Gly Thr Arg Ala Leu Leu Gln Thr Leu Gly Asn Leu Gly Tyr Arg Ala
 900 905 910
 Ser Ala Lys Lys Ala Gln Ile Cys Gln Lys Gln Val Lys Tyr Leu Gly
 915 920 925
 Tyr Leu Leu Lys Glu Gly Gln Arg Trp Leu Thr Glu Ala Arg Lys Glu
 930 935 940
 Thr Val Met Gly Gln Pro Thr Pro Lys Thr Pro Arg Gln Leu Arg Glu
 945 950 955 960
 Phe Leu Gly Thr Ala Gly Phe Cys Arg Leu Trp Ile Pro Gly Phe Ala
 965 970 975
 Glu Met Ala Ala Pro Leu Tyr Pro Leu Thr Lys Thr Gly Thr Leu Phe
 980 985 990
 Asn Trp Gly Pro Asp Gln Gln Lys Ala Tyr Gln Glu Ile Lys Gln Ala
 995 1000 1005
 Leu Leu Thr Ala Pro Ala Leu Gly Leu Pro Asp Leu Thr Lys Pro Phe
 1010 1015 1020
 Glu Leu Phe Val Asp Glu Lys Gln Gly Tyr Ala Lys Gly Val Leu Thr
 1025 1030 1035 1040
 Gln Lys Leu Gly Pro Trp Arg Arg Pro Val Ala Tyr Leu Ser Lys Lys
 1045 1050 1055
 Leu Asp Pro Val Ala Ala Gly Trp Pro Pro Cys Leu Arg Met Val Ala
 1060 1065 1070
 Ala Ile Ala Val Leu Thr Lys Asp Ala Gly Lys Leu Thr Met Gly Gln
 1075 1080 1085
 Pro Leu Val Ile Leu Ala Pro His Ala Val Glu Ala Leu Val Lys Gln
 1090 1095 1100
 Pro Pro Asp Arg Trp Leu Ser Asn Ala Arg Met Thr His Tyr Gln Ala
 1105 1110 1115 1120
 Leu Leu Leu Asp Thr Asp Arg Val Gln Phe Gly Pro Val Val Ala Leu
 1125 1130 1135
 Asn Pro Ala Thr Leu Leu Pro Leu Pro Glu Glu Gly Leu Gln His Asn
 1140 1145 1150
 Cys Leu Asp Ile Leu Ala Glu Ala His Gly Thr Arg Pro Asp Leu Thr
 1155 1160 1165
 Asp Gln Pro Leu Pro Asp Ala Asp His Thr Trp Tyr Thr Asp Gly Ser
 1170 1175 1180
 Ser Leu Leu Gln Glu Gly Gln Arg Lys Ala Gly Ala Ala Val Thr Thr
 1185 1190 1195 1200
 Glu Thr Glu Val Ile Trp Ala Lys Ala Leu Pro Ala Gly Thr Ser Ala
 1205 1210 1215

Gln Arg Ala Glu Leu Ile Ala Leu Thr Gln Ala Leu Lys Met Ala Glu
 1220 1225 1230
 Gly Lys Lys Leu Asn Val Tyr Thr Asp Ser Arg Tyr Ala Phe Ala Thr
 1235 1240 1245
 Ala His Ile His Gly Glu Ile Tyr Arg Arg Arg Gly Leu Leu Thr Ser
 1250 1255 1260
 Glu Gly Lys Glu Ile Lys Asn Lys Asp Glu Ile Leu Ala Leu Leu Lys
 1265 1270 1275 1280
 Ala Leu Phe Leu Pro Lys Arg Leu Ser Ile Ile His Cys Pro Gly His
 1285 1290 1295
 Gln Lys Gly His Ser Ala Glu Ala Arg Gly Asn Arg Met Ala Asp Gln
 1300 1305 1310
 Ala Ala Arg Lys Ala Ala Ile Thr Glu Thr Pro Asp Thr Ser Thr Leu
 1315 1320 1325
 Leu Ile Glu Asn Ser Ser Pro Tyr Thr Ser Glu His Phe His Tyr Thr
 1330 1335 1340
 Val Thr Asp Ile Lys Asp Leu Thr Lys Leu Gly Ala Ile Tyr Asp Lys
 1345 1350 1355 1360
 Thr Lys Lys Tyr Trp Val Tyr Gln Gly Lys Pro Val Met Pro Asp Gln
 1365 1370 1375
 Phe Thr Phe Glu Leu Leu Asp Phe Leu His Gln Leu Thr His Leu Ser
 1380 1385 1390
 Phe Ser Lys Met Lys Ala Leu Leu Glu Arg Ser His Ser Pro Tyr Tyr
 1395 1400 1405
 Met Leu Asn Arg Asp Arg Thr Leu Lys Asn Ile Thr Glu Thr Cys Lys
 1410 1415 1420
 Ala Cys Ala Gln Val Asn Ala Ser Lys Ser Ala Val Lys Gln Gly Thr
 1425 1430 1435 1440
 Arg Val Arg Gly His Arg Pro Gly Thr His Trp Glu Ile Asp Phe Thr
 1445 1450 1455
 Glu Ile Lys Pro Gly Leu Tyr Gly Tyr Lys Tyr Leu Leu Val Phe Ile
 1460 1465 1470
 Asp Thr Phe Ser Gly Trp Ile Glu Ala Phe Pro Thr Lys Lys Glu Thr
 1475 1480 1485
 Ala Lys Val Val Thr Lys Lys Leu Leu Glu Glu Ile Phe Pro Arg Phe
 1490 1495 1500
 Gly Met Pro Gln Val Leu Gly Thr Asp Asn Gly Pro Ala Phe Val Ser
 1505 1510 1515 1520
 Lys Val Ser Gln Thr Val Ala Asp Leu Leu Gly Ile Asp Trp Lys Leu
 1525 1530 1535
 His Cys Ala Tyr Arg Pro Gln Ser Ser Gly Gln Val Glu Arg Met Asn
 1540 1545 1550
 Arg Thr Ile Lys Glu Thr Leu Thr Lys Leu Thr Leu Ala Thr Gly Ser
 1555 1560 1565
 Arg Asp Trp Val Leu Leu Leu Pro Leu Ala Leu Tyr Arg Ala Arg Asn
 1570 1575 1580
 Thr Pro Gly Pro His Gly Leu Thr Pro Tyr Glu Ile Leu Tyr Gly Ala
 1585 1590 1595 1600
 Pro Pro Pro Leu Val Asn Phe Pro Asp Pro Asp Met Thr Arg Val Thr
 1605 1610 1615
 Asn Ser Pro Ser Leu Gln Ala His Leu Gln Ala Leu Tyr Leu Val Gln
 1620 1625 1630
 His Glu Val Trp Arg Pro Leu Ala Ala Ala Tyr Gln Glu Gln Leu Asp
 1635 1640 1645
 Arg Pro Val Val Pro His Pro Tyr Arg Val Gly Asp Thr Val Trp Val
 1650 1655 1660
 Arg Arg His Gln Thr Lys Asn Leu Glu Pro Arg Trp Lys Gly Pro Tyr
 1665 1670 1675 1680
 Thr Val Leu Leu Thr Thr Pro Thr Ala Leu Lys Val Asp Gly Ile Ala
 1685 1690 1695
 Ala Trp Ile His Ala Ala His Val Lys Ala Ala Asp Pro Gly Gly Gly
 1700 1705 1710
 Pro Ser Ser Arg Leu Thr Trp Arg Val Gln Arg Ser Gln Asn Pro Leu
 1715 1720 1725

Lys Ile Arg Leu Thr Arg Glu Ala Pro
1730 1735

(2) ANGABEN ZU SEQ ID NO: 14:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 665 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "env protein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 14:

Met	Ala	Arg	Ser	Thr	Leu	Ser	Lys	Pro	Leu	Lys	Asn	Lys	Val	Asn	Pro
1				5					10					15	
Arg	Gly	Pro	Leu	Ile	Pro	Leu	Ile	Leu	Leu	Met	Leu	Arg	Gly	Val	Ser
			20					25					30		
Thr	Ala	Ser	Pro	Gly	Ser	Ser	Pro	His	Gln	Val	Tyr	Asn	Ile	Thr	Trp
		35					40					45			
Glu	Val	Thr	Asn	Gly	Asp	Arg	Glu	Thr	Val	Trp	Ala	Thr	Ser	Gly	Asn
	50					55					60				
His	Pro	Leu	Trp	Thr	Trp	Trp	Pro	Asp	Leu	Thr	Pro	Asp	Leu	Cys	Met
65				70						75					80
Leu	Ala	His	His	Gly	Pro	Ser	Tyr	Trp	Gly	Leu	Glu	Tyr	Gln	Ser	Pro
				85					90					95	
Phe	Ser	Ser	Pro	Pro	Gly	Pro	Pro	Cys	Cys	Ser	Gly	Gly	Ser	Ser	Pro
			100					105					110		
Gly	Cys	Ser	Arg	Asp	Cys	Glu	Glu	Pro	Leu	Thr	Ser	Leu	Thr	Pro	Arg
		115				120						125			
Cys	Asn	Thr	Ala	Trp	Asn	Arg	Leu	Lys	Leu	Asp	Gln	Thr	Thr	His	Lys
	130					135					140				
Ser	Asn	Glu	Gly	Phe	Tyr	Val	Cys	Pro	Gly	Pro	His	Arg	Pro	Arg	Glu
145					150					155					160
Ser	Lys	Ser	Cys	Gly	Gly	Pro	Asp	Ser	Phe	Tyr	Cys	Ala	Tyr	Trp	Gly
				165					170					175	
Cys	Glu	Thr	Thr	Gly	Arg	Ala	Tyr	Trp	Lys	Pro	Ser	Ser	Ser	Trp	Asp
			180					185					190		
Phe	Ile	Thr	Val	Asn	Asn	Asn	Leu	Thr	Ser	Asp	Gln	Ala	Val	Gln	Val
			195				200					205			
Cys	Lys	Asp	Asn	Lys	Trp	Cys	Asn	Pro	Leu	Val	Ile	Arg	Phe	Thr	Asp
	210					215					220				
Ala	Gly	Arg	Arg	Val	Thr	Ser	Trp	Thr	Thr	Gly	His	Tyr	Trp	Gly	Leu
225					230					235					240
Arg	Leu	Tyr	Val	Ser	Gly	Gln	Asp	Pro	Gly	Leu	Thr	Phe	Gly	Ile	Arg
				245					250					255	
Leu	Arg	Tyr	Gln	Asn	Leu	Gly	Pro	Arg	Val	Pro	Ile	Gly	Pro	Asn	Pro
			260					265					270		
Val	Leu	Ala	Asp	Gln	Gln	Pro	Leu	Ser	Lys	Pro	Lys	Pro	Val	Lys	Ser
		275					280					285			
Pro	Ser	Val	Thr	Lys	Pro	Pro	Ser	Gly	Thr	Pro	Leu	Ser	Pro	Thr	Gln
	290					295					300				
Leu	Pro	Pro	Ala	Gly	Thr	Glu	Asn	Arg	Leu	Leu	Asn	Leu	Val	Asp	Gly
305					310					315					320
Ala	Tyr	Gln	Ala	Leu	Asn	Leu	Thr	Ser	Pro	Asp	Lys	Thr	Gln	Glu	Cys
				325					330					335	

Trp Leu Cys Leu Val Ala Gly Pro Pro Tyr Tyr Glu Gly Val Ala Val
 340 345 350
 Leu Gly Thr Tyr Ser Asn His Thr Ser Ala Pro Ala Asn Cys Ser Val
 355 360 365
 Ala Ser Gln His Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu
 370 375 380
 Cys Ile Gly Ala Val Pro Lys Thr His Gln Ala Leu Cys Asn Thr Thr
 385 390 395 400
 Gln Thr Ser Ser Arg Gly Ser Tyr Tyr Leu Val Ala Pro Thr Gly Thr
 405 410 415
 Met Trp Ala Cys Ser Thr Gly Leu Thr Pro Cys Ile Ser Thr Thr Ile
 420 425 430
 Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu Leu Trp Pro Arg
 435 440 445
 Val Thr Tyr His Ser Pro Ser Tyr Val Tyr Gly Leu Phe Glu Arg Ser
 450 455 460
 Asn Arg His Lys Arg Glu Pro Val Ser Leu Thr Leu Ala Leu Leu Leu
 465 470 475 480
 Gly Gly Leu Thr Met Gly Gly Ile Ala Ala Gly Ile Gly Thr Gly Thr
 485 490 495
 Thr Ala Leu Met Ala Thr Gln Gln Phe Gln Gln Leu Gln Ala Ala Val
 500 505 510
 Gln Asp Asp Leu Arg Glu Val Glu Lys Ser Ile Ser Asn Leu Glu Lys
 515 520 525
 Ser Leu Thr Ser Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu
 530 535 540
 Asp Leu Leu Phe Leu Lys Glu Gly Gly Leu Cys Ala Ala Leu Lys Glu
 545 550 555 560
 Glu Cys Cys Phe Tyr Ala Asp His Thr Gly Leu Val Arg Asp Ser Met
 565 570 575
 Ala Lys Leu Arg Glu Arg Leu Asn Gln Arg Gln Lys Leu Phe Glu Ser
 580 585 590
 Thr Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser Pro Trp Phe Thr
 595 600 605
 Thr Leu Ile Ser Thr Ile Met Gly Pro Leu Ile Val Leu Leu Met Ile
 610 615 620
 Leu Leu Phe Gly Pro Cys Ile Leu Asn Arg Leu Val Gln Phe Val Lys
 625 630 635 640
 Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr Gln Gln Tyr His
 645 650 655
 Gln Leu Lys Pro Ile Glu Tyr Glu Pro
 660 665

(2) ANGABEN ZU SEQ ID NO: 15:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 9709 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 15:

TGGAAGGGCT	AATTTGGTCC	CAAAAAGAC	AAGAGATCCT	TGATCTGTGG	ATCTACCACA	60
CACAAGGCTA	CTTCCCTGAT	TGGCAGAACT	ACACACCAGG	GCCAGGGATC	AGATATCCAC	120
TGACCTTTGG	ATGGTGCTTC	AAGTTAGTAC	CAGTTGAACC	AGAGCAAGTA	GAAGAGGCCA	180
AATAAGGAGA	GAAGAACAGC	TTGTTACACC	CTATGAGCCA	GCATGGGATG	GAGGACCCGG	240
AGGGAGAAGT	ATTAGTGTGG	AAGTTTGACA	GCCTCCTAGC	ATTTTCGTAC	ATGGCCCCGAG	300
AGCTGCATCC	GGAGTACTAC	AAAGACTGCT	GACATCGAGC	TTTCTACAAG	GGACTTTCCG	360

CTGGGACTT	TCCAGGGAGG	TGTGGCCTGG	GCGGGACTGG	GGAGTGGCGA	GCCCTCAGAT	420
GCTACATATA	AGCAGCTGCT	TTTTGCCTGT	ACTGGGTCTC	TCTGGTTAGA	CCAGATCTGA	480
GCCTGGGAGC	TCTCTGGCTA	ACTAGGGAAC	CCACTGCTTA	AGCCTCAATA	AAGCTTGCTT	540
TGAGTGCTCA	AAGTAGTGTG	TGCCCCGTCTG	TTGTGTGACT	CTGGTAACTA	GAGATCCCTC	600
AGACCCTTTT	AGTCAGTGTG	GAAAATCTCT	AGCAGTGGCG	CCCGAACAGG	GACTTGAAG	660
CGAAAGTAAA	GCCAGAGGAG	ATCTCTCGAC	GCAGGACTCG	GCTTGCTGAA	GCGCGCACGG	720
CAAGAGGCGA	GGGGCGGCGA	CTGGTGAGTA	CGCCAAAAAT	TTTACTAGC	GGAGGCTAGA	780
AGGAGAGAGA	TGGGTGCGAG	AGCGTCCGTA	TTAAGCGGGG	GAGAATTAGA	TAAATGGGAA	840
AAAATTCGGT	TAAGGCCAGG	GGGAAAGAAA	CAATATAAAC	TAAAAATAT	AGTATGGGCA	900
AGCAGGGAGC	TAGAACGATT	CGCAGTTAAT	CCTGGCCTTT	TAGAGACATC	AGAAGGCTGT	960
AGACAAATAC	TGGGACAGCT	ACAACCATCC	CCTCAGACAG	GATCAGAAGA	ACTTAGATCA	1020
TTATATAATA	CAATAGCAGT	CCTCTATTGT	GTGCATCAAA	GGATAGATGT	AAAAGACACC	1080
AAGGAAGCCT	TAGATAAGAT	AGAGGAAGAG	CAAAAACAAA	GTAAGAAAAA	GGCACAGCAA	1140
GCAGCAGCTG	ACACAGGAAA	CAACAGCCAG	GTCAGCCAAA	ATTACCCTAT	AGTGCAGAAC	1200
CTCCAGGGGC	AAATGGTACA	TCAGGCCATA	TCACCTAGAA	CTTTAAATGC	ATGGGTAAAA	1260
GTAGTAGAAG	AGAAGGCTTT	CAGCCCAGAA	GTAATACCCA	TGTTTTCAGC	ATTATCAGAA	1320
GGAGCCACCC	CACAAGATTT	AAATACCATG	CTAAACACAG	TGGGGGGACA	TCAAGCAGCC	1380
ATGCAAATGT	TAAAAGAGAC	CATCAATGAG	GAAGCTGCAG	AATGGGATAG	ATTGCATCCA	1440
GTGCATGCAG	GGCCTATTGC	ACCAGGCCAG	ATGAGAGAAC	CAAGGGGAAG	TGACATAGCA	1500
GGAACACTA	GTACCCTTCA	GGAAACAAATA	GGATGGATGA	CACATAATCC	ACCTATCCCA	1560
GTAGGAGAAA	TCTATAAAAAG	ATGGATAATC	CTGGGATTA	ATAAAATAGT	AAGAATGTAT	1620
AGCCCTACCA	GCATCTTGGA	CATAAGACAA	GGACCAAAGG	AACCCTTTAG	AGACTATGTA	1680
GACCGATTCT	ATAAACTCT	AAGAGCCGAG	CAAGCTTCAC	AAGAGGTAAA	AAATTGGATG	1740
ACAGAAACCT	TGTTGGTCCA	AAATGCGAAC	CCAGATTGTA	AGACTATTTT	AAAAGCATTG	1800
GGACCAGGAG	CGACACTAGA	AGAAATGATG	ACAGCATGTC	AGGGAGTGGG	GGGACCCGGC	1860
CATAAAGCAA	GAGTTTTGGC	TGAAGCAATG	AGCCAAGTAA	CAAATCCAGC	TACCATAATG	1920
ATACAGAAAG	GCAATTTTTAG	GAACCAAAGA	AAGACTGTTA	AGTGTTTCAA	TTGTGGCAAA	1980
GAAGGGCACA	TAGCCAAAAA	TTGCAGGGCC	CCTAGGAAAA	AGGGCTGTTG	GAAATGTGGA	2040
AAGGAAGGAC	ACCAAATGAA	AGATTGTACT	GAGAGACAGG	CTAATTTTTT	AGGGAAGATC	2100
TGGCCTTCCC	ACAAGGGAAG	GCCAGGGAAT	TTTCTTCAGA	GCAGACCAGA	GCCAACAGCC	2160
CCACCAGAAG	AGAGCTTCAG	GTTTGGGGAA	GAGACAACAA	CTCCCTCTCA	GAAGCAGGAG	2220
CCGATAGACA	AGGAACTGTA	TCCTTTAGCT	TCCCTCAGAT	CACTCTTTGG	CAGCGACCCC	2280
TCGTACAAAT	AAAGATAGGG	GGGCAATTA	AGGAAGCTCT	ATTAGATACA	GGAGCAGATG	2340
ATACAGTATT	AGAAGAAATG	AATTTGCCAG	GAAGATGGAA	ACCAAAAATG	ATAGGGGGAA	2400
TTGGAGGTTT	TATCAAAGTA	GGACAGTATG	ATCAGATACT	CATAGAAATC	TGCGGACATA	2460
AAGCTATAGG	TACAGTATTA	GTAGGACCTA	CACCTGTCAA	CATAATTGGA	AGAAATCTGT	2520
TGACTCAGAT	TGGCTGCACT	TTAAATTTTT	CCATTAGTCC	TATTGAGACT	GTACCAGTAA	2580
AATTAAGCC	AGGAATGGAT	GGCCCAAAAG	TTAAACAATG	GCCATTGACA	GAAGAAAAAA	2640
TAAAAGCATT	AGTAGAAATT	TGTACAGAAA	TGGAAAAGGA	AGGAAAAAAT	TCAAAAATTTG	2700
GGCCTGAAAA	TCCATACAAT	ACTCCAGTAT	TTGCCATAAA	GAAAAAAGAC	AGTACTAAAT	2760
GGAGAAAATT	AGTAGATTTT	AGAGAATTTA	ATAAGAGAAC	TCAAGATTTT	TGGGAAGTTC	2820
AATTAGGAAT	ACCACATCCT	GCAGGGTTAA	AACAGAAAAA	ATCAGTAACA	GTACTGGATG	2880
TGGGCGATGC	ATATTTTTTCA	GTTCCCTTAG	ATAAAGACTT	CAGGAAGTAT	ACTGCATTTA	2940
CCATACCTAG	TATAACAAT	GAGACACCAG	GGATTAGATA	TCAGTACAAT	GTGCTTCCAC	3000
AGGGATGGAA	AGGATCACCA	GCAATATTCC	AGTGTAGCAT	GACAAAAATC	TTAGAGCCTT	3060
TTAGAAAACA	AAATCCAGAC	ATAGTCATCT	ATCAATACAT	GGATGATTTG	TATGTAGGAT	3120
CTGACTTAGA	AATAGGGCAG	CATAGAACAA	AAATAGAGGA	ACTGAGACAA	CATCTGTTGA	3180
GGTGGGGATT	TACCACACCA	GACAAAAAAC	ATCAGAAAGA	ACCTCCATT	CTTTGGATGG	3240
GTTATGAACT	CCATCTTGAT	AAATGGACAG	TACAGCCTAT	AGTGCTGCCA	GAAAAGGACA	3300
GCTGGACTGT	CAATGCACATA	CAGAAATTAG	TGGGAAAAAT	GAATTGGGCA	AGTCAGATTT	3360
ATGCAGGGAT	TAAAGTAAGG	CAATTATGTA	AACCTTTAG	GGGAACCAAA	GCACTAACAG	3420
AAGTAGTACC	ACTAACAGAA	GAAGCAGAGC	TAGAACTGGC	AGAAAAACAGG	GAGATTCTAA	3480
AAGAACCGGT	ACATGGAGTG	TATTATGACC	CATCAAAAAG	CTTAATAGCA	GAATACAGA	3540
AGCAGGGGCA	AGGCCAATGG	ACATATCAAA	TTTATCAAGA	GCCATTTAAA	AATCTGAAAA	3600
CAGGAAAATA	TGCAAGAATG	AAGGGTGCCC	ACACTAATGA	TGTGAAACAA	TTAACAGAGG	3660
CAGTACAAAA	AATAGCCACA	GAAAGCATAG	TAATATGGGG	AAAGACTCCT	AAATTTAAAT	3720
TACCCATACA	AAAGGAAAAA	TGGGAAGCAT	GGTGGACAGA	GTATTGGCAA	GCCACCTGGA	3780
TTCTGATAGT	GGAGTTTGT	AATACCCCTC	CCTTAGTGAA	GTTATGGTAC	CAGTTAGAGA	3840
AAGAACCCAT	AATAGGAGCA	GAAACTTTCT	ATGTAGATGG	GGCAGCCAAT	AGGGAAACTA	3900
AATTAGGAAA	AGCAGGATAT	GTAACGACA	GAGGAAGACA	AAAAGTTGTC	CCCCTAACGG	3960
ACACAACAAA	TCAGAAGACT	GAGTTACAAG	CAATTCATCT	AGCTTTGCAG	GATTCGGGAT	4020
TAGAAGTAAA	CATAGTGACA	GACTCACAAT	ATGCATTGGG	AATCATTCAA	GCACAACCAG	4080
ATAAGAGTGA	ATCAGAGTTA	GTCAGTCAAA	TAATAGAGCA	GTTAATAAAA	AAGGAAAAAG	4140
TCTACCTGGC	ATGGGTACCA	GCACACAAA	GAATTGGAGG	AAATGAACAA	GTAGATGGGT	4200

TGGTCAGTGC	TGGAATCAGG	AAAGTACTAT	TTTTAGATGG	AATAGATAAG	GCCCAAGAAG	4260
AACATGAGAA	ATATCACAGT	AATTGGAGAG	CAATGGCTAG	TGATTTTAAAC	CTACCACCTG	4320
TAGTAGCAAA	AGAAATAGTA	GCCAGCTGTG	ATAAATGTCA	GCTAAAAGGG	GAAGCCATGC	4380
ATGGACAAGT	AGACTGTAGC	CCAGGAATAT	GGCAGCTAGA	TTGTACACAT	TTAGAAGGAA	4440
AAGTTATCTT	GGTAGCAGTT	CATGTAGCCA	TGTGGATATAT	AGAAGCAGAA	GTAATTCCAG	4500
CAGAGACAGG	GCAAGAAACA	GCATACTTCC	TCTTAAAATT	AGCAGGAAGA	TGGCCAGTAA	4560
AAACAGTACA	TACAGACAAT	GGCAGCAATT	TCACCAGTAC	TACAGTTAAG	GCCGCCTGTT	4620
GGTGGGCGGG	GATCAAGCAG	GAATTTGGCA	TTCCCTACAA	TCCCCAAAGT	CAAGGAGTAA	4680
TAGAATCTAT	GAATAAAGAA	TTAAAGAAAA	TTATAGGACA	GGTAAGAGAT	CAGGCTGAAC	4740
ATCTTAAGAC	AGCAGTACAA	ATGGCAGTAT	TCATCCACAA	TTTTAAAAGA	AAAGGGGGGA	4800
TTGGGGGGTA	CAGTGCAGGG	AAAAGAATAG	TAGACATAAT	AGCAACAGAC	ATACAAACTA	4860
AAGAATTTACA	AAAACAAATT	ACAAAAATTC	AAAAATTTTCG	GGTTTATTAC	AGGGACAGCA	4920
GAGATCCAGT	TTGGAAAAGGA	CCAGCAAAGC	TCCTCTGGAA	AGGTGAAGGG	GCAGTAGTAA	4980
TACAAGATAA	TAGTGACATA	AAAGTAGTGC	CAAGAAGAAA	AGCAAAGATC	ATCAGGGATT	5040
ATGGAAAACA	GATGGCAGGT	GATGATTGTG	TGGCAAGTAG	ACAGGATGAG	GATTAACACA	5100
TGGAAAAGAT	TAGTAAAAACA	CCATATGTAT	ATTCAAGGA	AAGCTAAGGA	CTGGTTTTAT	5160
AGACATCACT	ATGAAAGTAC	TAATCCAAAA	ATAAGTTTCAG	AAGTACACAT	CCCCTAGGG	5220
GATGCTAAAT	TAGTAAATAC	AACATATTGG	GGTCTGCATA	CAGGAGAAAAG	AGACTGGCAT	5280
TTGGGTCCAGG	GAGTCTCCAT	AGAATGGAGG	AAAAAGAGAT	ATAGCACACA	AGTAGACCCT	5340
GACCTAGCAG	ACCAACTAAT	TCATCTGCAC	TATTTTGATT	GTTTTTCAGA	ATCTGCTATA	5400
AGAAATACCA	TATTAGGACG	TATAGTTAGT	CCTAGGTGTG	AATATCAAGC	AGGACATAAC	5460
AAGGTAGGAT	CTCTACAGTA	CTTGGCACTA	GCAGCATTAA	TAAAACCCAAA	ACAGATAAAG	5520
CCACCTTTGC	CTAGTGTTAG	GAAACTGACA	GAGGACAGAT	GGAAACAAGCC	CCAGAAGACC	5580
AAGGGCCACA	GAGGGAGCCA	TACAAATGAAT	GGACACTAGA	GCTTTTAGAG	GAACTTAAGA	5640
GTGAAGCTGT	TAGACATTTT	CCTAGGATAT	GGCTCCATAA	CTTAGGACAA	CATATCTATG	5700
AAACTTACGG	GGTACTTGG	GCAGGAGTGG	AAGCCATAAT	AAGAATTCTG	CAACAACCTG	5760
TGTTTATCCA	TTTCAGAATT	GGGTGTCGAC	ATAGCAGAAT	AGGCGTTACT	CGACAGAGGA	5820
GAGCAAGAAA	TGGAGCCAGT	AGATCCTAGA	CTAGAGCCCT	GGAAGCATCC	AGGAAGTCAG	5880
CCTAAAACCTG	CTTGTAACCA	TTGCTATTGT	AAAAAGTGTT	GCTTTTCATTG	CCAAGTTTGT	5940
TTCATGACAA	AAGCCTTAGG	CATCTCCTAT	GGCAGGAAGA	AGCGGAGACA	GCGACGAAGA	6000
GCTCATCAGA	ACAGTCAGAC	TCATCAAGCT	TCTCTATCAA	AGCAGTAAGT	AGTACATGTA	6060
ATGCAACCTA	TAATAGTAGC	AATAGTAGCA	TTAGTAGTAG	CAATAATAAT	AGCAATAGTT	6120
GTGTGGTCCA	TAGTAATCAT	AGAATATAGG	AAAATATTTAA	GACAAAGAAA	AATAGACAGG	6180
TTAATTGATA	GACTAATAGA	AAGAGCAGAA	GACAGTGGCA	ATGAGAGTGA	AGGAGAAGTA	6240
TCAGCACTTG	TGGAGATGGG	GGTGGAAATG	GGGCACCATG	CTCCTTGGGA	TATTGATGAT	6300
CTGTAGTCTG	ACAGAAAAAT	TGTGGGTCAC	AGTCTATTAT	GGGGTACCTG	TGTGGAAGGA	6360
AGCAACCACC	ACTCTATTTT	GTGCATCAGA	TGCTAAAGCA	TATGATACAG	AGGTACATAA	6420
TGTTTGGGCC	ACACATGCCT	GTGTACCCAC	AGACCCCAAC	CCACAAGAAG	TAGTATTGGT	6480
AAATGTGACA	GAAAATTTTA	ACATGTGGAA	AAATGACATG	GTAGAACAGA	TGCATGAGGA	6540
TATAATCAGT	TTATGGGATC	AAAGCCTAAA	GCCATGTGTA	AAATTAACCC	CACTCTGTGT	6600
TAGTTTAAAG	TGCACTGATT	TGAAGAATGA	TACTAATACC	AATAGTAGTA	GCGGGAGAAT	6660
GATAATGGAG	AAAGGAGAGA	TAAAAAACTG	CTCTTTCAAT	ATCAGCACAA	GCATAAGAGA	6720
TAAGGTGCAG	AAAGAATATG	CATTCTTTTA	TAAACTTGAT	ATAGTACCAA	TAGATAATAC	6780
CAGCTATAGG	TTGATAAGTT	GTAACACCTC	AGTCATTACA	CAGGCCTGTC	CAAAGGTATC	6840
CTTTGAGCCA	ATTCCCATAC	ATTATTGTGC	CCCGGCTGGT	TTTGCGATTG	TAAAATGTAA	6900
TAATAAGACG	TTCAATGGAA	CAGGACCATG	TACAAATGTC	AGCACAGTAC	AATGTACACA	6960
TGGAATCAGG	CCAGTAGTAT	CAACTCAACT	GCTGTAAAT	GGCAGTCTAG	CAGAAGAAGA	7020
TGTAGTAATT	AGATCTGCCA	ATTTACAGAG	CAATGCTAAA	ACCATAATAG	TACAGCTGAA	7080
CACATCTGTA	GAAATTAATT	GTACAAGACC	CAACAACAAT	ACAAGAAAAA	GTATCCGTAT	7140
CCAGAGGGGA	CCAGGGAGAG	CATTTGTTAC	AATAGGAAAA	ATAGGAAATA	TGAGACAAGC	7200
ACATTGTAAC	ATTAGTAGAG	CAAAAATGGAA	TGCCACTTTA	AAACAGATAG	CTAGCAAATT	7260
AAGAGAACAA	TTTGAAAATA	ATAAAAACAAT	AATCTTTAAG	CAATCCTCAG	GAGGGGACCC	7320
AGAAATTTGTA	AGCCACAGTT	TTAATTGTGG	AGGGGAATTT	TTCTACTGTA	ATTCAACACA	7380
ACTGTTTAAAT	AGTACTTGGT	TTAATAGTAC	TTGGAGTACT	GAAGGTCAA	ATAACTCTGA	7440
AGGAAGTGAC	ACAATCACAC	TCCCATGCAG	AATAAAACAA	TTTATAAACA	TGTGGCAGGA	7500
AGTAGGAAAA	GCAATGTATG	CCCTCCCAT	CAGTGGACAA	ATTAGATGTT	CATCAAATAT	7560
TACTGGGCTG	CTATTAACAA	GAGATGGTGG	TAATAACAAC	AATGGGTCCG	AGATCTTCAG	7620
ACCTGGAGGA	GGCGATATGA	GGGACAATTG	GAGAAGTGAA	TTATATAAAT	ATAAAGTAGT	7680
AAAAATTGAA	CCATTAGGAG	TAGCACCCAC	CAAGGCAAAG	AGAAGAGTGG	TGCAGAGAGA	7740
AAAAAGAGCA	TGGGAATAG	GAGCTTTGTT	CCTTGGGTTT	TTGGGAGCAG	CAGGAAGCAC	7800
TATGGGCTGC	ACGTCAATGA	CGCTGACGGT	ACAGGCCAGA	CAATTATTGT	CTGATATAGT	7860
GCAGCAGCAG	AACAATTTGC	TGAGGGCTAT	TGAGGCGCAA	CAGCATCTGT	TGCAACTCAC	7920
AGTCTGGGGC	ATCAAACAGC	TCCAGGCAAG	AATCCTGGCT	GTGGAAAAGAT	ACCTAAAGGA	7980
TCAACAGCTC	CTGGGGATTT	GGGGTTGCTC	TGAAAAACTC	ATTTGCACCA	CTGCTGTGCC	8040

TTGGAATGCT	AGTTGGAGTA	ATAAATCTCT	GGAACAGATT	TGGAATAACA	TGACCTGGAT	8100
GGAGTGGGAC	AGAGAAATTA	ACAATTACAC	AAGCTTAATA	CACTCCTTAA	TTGAAGAATC	8160
GCAAAACCAG	CAAGAAAAGA	ATGAACAAGA	ATTATTGGAA	TTAGATAAAT	GGGCAAGTTT	8220
GTGGAATTGG	TTTAACATAA	CAAATTGGCT	GTGGTATATA	AAATTATTCA	TAATGATAGT	8280
AGGAGGCTTG	GTAGGTTTAA	GAATAGTTTT	TGCTGTACTT	TCTATAGTGA	ATAGAGTTAG	8340
GCAGGGATAT	TCACCATTAT	CGTTTCAGAC	CCACCTCCCA	ATCCCGAGGG	GACCCGACAG	8400
GCCCCAAGGA	ATAGAAGAAG	AAGGTGGAGA	GAGAGACAGA	GACAGATCCA	TTCGATTAGT	8460
GAACGGATCC	TTAGCACTTA	TCTGGGACGA	TCTGCGGAGC	CTGTGCCTCT	TCAGCTACCA	8520
CCGCTTGAGA	GACTTACTCT	TGATTGTAAC	GAGGATTGTG	GAACTTCTGG	GACGCAGGGG	8580
GTGGGAAGCC	CTCAAATATT	GGTGGAAATCT	CCTACAGTAT	TGGAGTCAGG	AACTAAAGAA	8640
TAGTGCTGTT	AACTTGCTCA	ATGCCACAGC	CATAGCAGTA	GCTGAGGGGA	CAGATAGGGT	8700
TATAGAAGTA	TTACAAGCAG	CTTATAGAGC	TATTGCCCAC	ATACCTAGAA	GAATAAGACA	8760
GGGCTTGGA	AGGATTTTGC	TATAAGATGG	GTGGCAAGTG	GTCAAAAAGT	AGTGTGATTG	8820
GATGGCCTGC	TGTAAGGGAA	AGAATGAGAC	GAGCTGAGCC	AGCAGCAGAT	GGGGTGGGAG	8880
CAGTATCTCG	AGACCTAGAA	AAACATGGAG	CAATCACAAG	TAGCAATACA	GCAGCTAACA	8940
ATGCTGCTTG	TGCCTGGCTA	GAAGCACAAG	AGGAGGAAGA	GGTGGGTTTT	CCAGTCACAC	9000
CTCAGGTACC	TTTAAGACCA	ATGACTTACA	AGGCAGCTGT	AGATCTTAGC	CACTTTTTTAA	9060
AAGAAAAGGG	GGGACTGGAA	GGGCTAATTC	ACTCCCAAAG	AAGACAAGAT	ATCCTTGATC	9120
TGTGGATCTA	CCACACACAA	GGCTACTTCC	CTGATTGGCA	GAACCTACACA	CCAGGGCCAG	9180
GGGTCAGATA	TCCACTGACC	TTTGGATGGT	GCTACAAGCT	AGTACCAGTT	GAGCCAGATA	9240
AGGTAGAAGA	GGCCAATAAA	GGAGAGAACA	CCAGCTTGTT	ACACCCTGTG	AGCCTGCATG	9300
GAATGGATGA	CCCTGAGAGA	GAAGTGTTAG	AGTGGAGGTT	TGACAGCCGC	CTAGCATTTT	9360
ATCACGTGGC	CCGAGAGCTG	CATCCGGAGT	ACTTCAAGAA	CTGCTGACAT	CGAGCTTGCT	9420
ACAAGGGACT	TTCCGCTGGG	GACTTTCCAG	GGAGGCGTGG	CCTGGGCGGG	ACTGGGGAGT	9480
GGCGAGCCCT	CAGATGCTGC	ATATAAGCAG	CTGCTTTTTG	CCTGTACTGG	GTCTCTCTGG	9540
TTAGACCAGA	TCTGAGCCTG	GGAGCTCTCT	GGCTAACTAG	GGAACCCACT	GCTTAAGCCT	9600
CAATAAAGCT	TGCCTTGAGT	GCTTCAAGTA	GTGTGTGCC	GTCTGTGTG	TGACTCTGGT	9660
AACTAGAGAT	CCCTCAGACC	CTTTTAGTCA	GTGTGGAAAA	TCTCTAGCA		9709

(2) ANGABEN ZU SEQ ID NO: 16:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 500 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "gag polyprotein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 16:

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

```

Ser  Gln  Asn  Tyr  Pro  Ile  Val  Gln  Asn  Leu  Gln  Gly  Gln  Met  Val  His
   130                135                140
Gln  Ala  Ile  Ser  Pro  Arg  Thr  Leu  Asn  Ala  Trp  Val  Lys  Val  Val  Glu
145                150                155
Glu  Lys  Ala  Phe  Ser  Pro  Glu  Val  Ile  Pro  Met  Phe  Ser  Ala  Leu  Ser
   165                170                175
Glu  Gly  Ala  Thr  Pro  Gln  Asp  Leu  Asn  Thr  Met  Leu  Asn  Thr  Val  Gly
   180                185                190
Gly  His  Gln  Ala  Ala  Met  Gln  Met  Leu  Lys  Glu  Thr  Ile  Asn  Glu  Glu
   195                200                205
Ala  Ala  Glu  Trp  Asp  Arg  Leu  His  Pro  Val  His  Ala  Gly  Pro  Ile  Ala
   210                215                220
Pro  Gly  Gln  Met  Arg  Glu  Pro  Arg  Gly  Ser  Asp  Ile  Ala  Gly  Thr  Thr
225                230                235
Ser  Thr  Leu  Gln  Glu  Gln  Ile  Gly  Trp  Met  Thr  His  Asn  Pro  Pro  Ile
   245                250                255
Pro  Val  Gly  Glu  Ile  Tyr  Lys  Arg  Trp  Ile  Ile  Leu  Gly  Leu  Asn  Lys
   260                265                270
Ile  Val  Arg  Met  Tyr  Ser  Pro  Thr  Ser  Ile  Leu  Asp  Ile  Arg  Gln  Gly
   275                280                285
Pro  Lys  Glu  Pro  Phe  Arg  Asp  Tyr  Val  Asp  Arg  Phe  Tyr  Lys  Thr  Leu
   290                295                300
Arg  Ala  Glu  Gln  Ala  Ser  Gln  Glu  Val  Lys  Asn  Trp  Met  Thr  Glu  Thr
305                310                315
Leu  Leu  Val  Gln  Asn  Ala  Asn  Pro  Asp  Cys  Lys  Thr  Ile  Leu  Lys  Ala
   325                330                335
Leu  Gly  Pro  Gly  Ala  Thr  Leu  Glu  Glu  Met  Met  Thr  Ala  Cys  Gln  Gly
   340                345                350
Val  Gly  Gly  Pro  Gly  His  Lys  Ala  Arg  Val  Leu  Ala  Glu  Ala  Met  Ser
   355                360                365
Gln  Val  Thr  Asn  Pro  Ala  Thr  Ile  Met  Ile  Gln  Lys  Gly  Asn  Phe  Arg
   370                375                380
Asn  Gln  Arg  Lys  Thr  Val  Lys  Cys  Phe  Asn  Cys  Gly  Lys  Glu  Gly  His
385                390                395
Ile  Ala  Lys  Asn  Cys  Arg  Ala  Pro  Arg  Lys  Lys  Gly  Cys  Trp  Lys  Cys
   405                410                415
Gly  Lys  Glu  Gly  His  Gln  Met  Lys  Asp  Cys  Thr  Glu  Arg  Gln  Ala  Asn
   420                425                430
Phe  Leu  Gly  Lys  Ile  Trp  Pro  Ser  His  Lys  Gly  Arg  Pro  Gly  Asn  Phe
   435                440                445
Leu  Gln  Ser  Arg  Pro  Glu  Pro  Thr  Ala  Pro  Pro  Glu  Glu  Ser  Phe  Arg
   450                455                460
Phe  Gly  Glu  Glu  Thr  Thr  Thr  Pro  Ser  Gln  Lys  Gln  Glu  Pro  Ile  Asp
465                470                475
Lys  Glu  Leu  Tyr  Pro  Leu  Ala  Ser  Leu  Arg  Ser  Leu  Phe  Gly  Ser  Asp
   485                490                495
Pro  Ser  Ser  Gln
   500

```

(2) ANGABEN ZU SEQ ID NO: 17:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1003 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "pol polyprotein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 17:

Phe Phe Arg Glu Asp Leu Ala Phe Pro Gln Gly Lys Ala Arg Glu Phe
1 5 10 15
Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Gln
20 25 30
Val Trp Gly Arg Asp Asn Asn Ser Leu Ser Glu Ala Gly Ala Asp Arg
35 40 45
Gln Gly Thr Val Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg
50 55 60
Pro Leu Val Thr Ile Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu
65 70 75 80
Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Glu Met Asn Leu Pro Gly
85 90 95
Arg Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val
100 105 110
Gly Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly His Lys Ala Ile
115 120 125
Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn
130 135 140
Leu Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile
145 150 155 160
Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
165 170 175
Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile
180 185 190
Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
195 200 205
Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
210 215 220
Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
225 230 235 240
Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
245 250 255
Gln Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
260 265 270
Val Pro Leu Asp Lys Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
275 280 285
Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
290 295 300
Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr
305 310 315 320
Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr
325 330 335
Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
340 345 350
His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly
355 360 365
Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Phe Leu Trp
370 375 380
Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val
385 390 395 400
Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val
405 410 415
Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys Val Arg
420 425 430
Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Val
435 440 445
Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
450 455 460
Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
465 470 475 480
Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile
485 490 495

Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met
 500 505 510
 Lys Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln
 515 520 525
 Lys Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 530 535 540
 Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Ala Trp Trp Thr Glu Tyr
 545 550 555 560
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 565 570 575
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Ile Gly Ala
 580 585 590
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 595 600 605
 Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Pro Leu
 610 615 620
 Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala
 625 630 635 640
 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 645 650 655
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu
 660 665 670
 Val Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu
 675 680 685
 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 690 695 700
 Gly Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile
 705 710 715 720
 Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala
 725 730 735
 Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val
 740 745 750
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln
 755 760 765
 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu
 770 775 780
 Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
 785 790 795 800
 Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu
 805 810 815
 Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Val His Thr Asp Asn
 820 825 830
 Gly Ser Asn Phe Thr Ser Thr Thr Val Lys Ala Ala Cys Trp Trp Ala
 835 840 845
 Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
 850 855 860
 Val Ile Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
 865 870 875 880
 Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
 885 890 895
 Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
 900 905 910
 Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu
 915 920 925
 Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
 930 935 940
 Ser Arg Asp Pro Val Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
 945 950 955 960
 Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
 965 970 975
 Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
 980 985 990
 Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp
 995 1000

(2) ANGABEN ZU SEQ ID NO: 18:

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 192 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: linear
- (ii) ART DES MOLEKÜLS: Protein
- (ix) MERKMALE:
 (D) SONSTIGE ANGABEN: /product= "vif protein"
- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 18:

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

(2) ANGABEN ZU SEQ ID NO: 19:

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 96 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: linear
- (ii) ART DES MOLEKÜLS: Protein
- (ix) MERKMALE:
 (D) SONSTIGE ANGABEN: /product= "vpr protein"
- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 19:

Met	Glu	Gln	Ala	Pro	Glu	Asp	Gln	Gly	Pro	Gln	Arg	Glu	Pro	Tyr	Asn
1				5					10					15	

```

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

```

(2) ANGABEN ZU SEQ ID NO: 20:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 86 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "tat protein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 20:

```

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

```

(2) ANGABEN ZU SEQ ID NO: 21:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 116 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "rev protein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 21:

```

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

```

```

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

```

(2) ANGABEN ZU SEQ ID NO: 22:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 81 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "vpu protein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 22:

```

Met Gln Pro Ile Ile Val Ala Ile Val Ala Leu Val Val Ala Ile Ile
1                5                10                15
Ile Ala Ile Val Val Trp Ser Ile Val Ile Ile Glu Tyr Arg Lys Ile
                20                25                30
Leu Arg Gln Arg Lys Ile Asp Arg Leu Ile Asp Arg Leu Ile Glu Arg
                35                40                45
Ala Glu Asp Ser Gly Asn Glu Ser Glu Gly Glu Val Ser Ala Leu Val
50                55                60
Glu Met Gly Val Glu Met Gly His His Ala Pro Trp Asp Ile Asp Asp
65                70                75                80
Leu

```

(2) ANGABEN ZU SEQ ID NO: 23:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 854 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "envelope polyprotein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 23:

Met	Arg	Val	Lys	Glu	Lys	Tyr	Gln	His	Leu	Trp	Arg	Trp	Gly	Trp	Lys
1				5					10					15	
Trp	Gly	Thr	Met	Leu	Leu	Gly	Ile	Leu	Met	Ile	Cys	Ser	Ala	Thr	Glu
			20					25					30		
Lys	Leu	Trp	Val	Thr	Val	Tyr	Tyr	Gly	Val	Pro	Val	Trp	Lys	Glu	Ala
		35				40						45			
Thr	Thr	Thr	Leu	Phe	Cys	Ala	Ser	Asp	Ala	Lys	Ala	Tyr	Asp	Thr	Glu
	50				55						60				
Val	His	Asn	Val	Trp	Ala	Thr	His	Ala	Cys	Val	Pro	Thr	Asp	Pro	Asn
65					70					75				80	
Pro	Gln	Glu	Val	Val	Leu	Val	Asn	Val	Thr	Glu	Asn	Phe	Asn	Met	Trp
				85					90					95	
Lys	Asn	Asp	Met	Val	Glu	Gln	Met	His	Glu	Asp	Ile	Ile	Ser	Leu	Trp
			100					105					110		
Asp	Gln	Ser	Leu	Lys	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	Val	Ser
		115					120					125			
Leu	Lys	Cys	Thr	Asp	Leu	Lys	Asn	Asp	Thr	Asn	Thr	Asn	Ser	Ser	Ser
	130					135					140				
Gly	Arg	Met	Ile	Met	Glu	Lys	Gly	Glu	Ile	Lys	Asn	Cys	Ser	Phe	Asn
145					150					155				160	
Ile	Ser	Thr	Ser	Ile	Arg	Asp	Lys	Val	Gln	Lys	Glu	Tyr	Ala	Phe	Phe
				165					170					175	
Tyr	Lys	Leu	Asp	Ile	Val	Pro	Ile	Asp	Asn	Thr	Ser	Tyr	Arg	Leu	Ile
			180					185					190		
Ser	Cys	Asn	Thr	Ser	Val	Ile	Thr	Gln	Ala	Cys	Pro	Lys	Val	Ser	Phe
		195					200					205			
Glu	Pro	Ile	Pro	Ile	His	Tyr	Cys	Ala	Pro	Ala	Gly	Phe	Ala	Ile	Leu
	210				215						220				
Lys	Cys	Asn	Asn	Lys	Thr	Phe	Asn	Gly	Thr	Gly	Pro	Cys	Thr	Asn	Val
225					230					235					240
Ser	Thr	Val	Gln	Cys	Thr	His	Gly	Ile	Arg	Pro	Val	Val	Ser	Thr	Gln
				245					250					255	
Leu	Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu	Glu	Asp	Val	Val	Ile	Arg	Ser
			260					265					270		
Ala	Asn	Phe	Thr	Asp	Asn	Ala	Lys	Thr	Ile	Ile	Val	Gln	Leu	Asn	Thr
		275					280					285			
Ser	Val	Glu	Ile	Asn	Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser
	290					295					300				
Ile	Arg	Ile	Gln	Arg	Gly	Pro	Gly	Arg	Ala	Phe	Val	Thr	Ile	Gly	Lys
305					310					315					320
Ile	Gly	Asn	Met	Arg	Gln	Ala	His	Cys	Asn	Ile	Ser	Arg	Ala	Lys	Trp
				325					330					335	
Asn	Ala	Thr	Leu	Lys	Gln	Ile	Ala	Ser	Lys	Leu	Arg	Glu	Gln	Phe	Gly
			340					345					350		
Asn	Asn	Lys	Thr	Ile	Ile	Phe	Lys	Gln	Ser	Ser	Gly	Gly	Asp	Pro	Glu
		355					360					365			
Ile	Val	Thr	His	Ser	Phe	Asn	Cys	Gly	Gly	Glu	Phe	Phe	Tyr	Cys	Asn
	370					375					380				
Ser	Thr	Gln	Leu	Phe	Asn	Ser	Thr	Trp	Phe	Asn	Ser	Thr	Trp	Ser	Thr
385					390					395				400	
Glu	Gly	Ser	Asn	Asn	Thr	Glu	Gly	Ser	Asp	Thr	Ile	Thr	Leu	Pro	Cys
				405					410					415	
Arg	Ile	Lys	Gln	Phe	Ile	Asn	Met	Trp	Gln	Glu	Val	Gly	Lys	Ala	Met
			420					425					430		
Tyr	Ala	Pro	Pro	Ile	Ser	Gly	Gln	Ile	Arg	Cys	Ser	Ser	Asn	Ile	Thr
	435					440						445			
Gly	Leu	Leu	Leu	Thr	Arg	Asp	Gly	Gly	Asn	Asn	Asn	Asn	Gly	Ser	Glu
	450					455						460			
Ile	Phe	Arg	Pro	Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	Arg	Ser	Glu	
465					470				475					480	
Leu	Tyr	Lys	Tyr	Lys	Val	Val	Lys	Ile	Glu	Pro	Leu	Gly	Val	Ala	Pro
				485					490					495	

Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala Val Gly
 500 505 510
 Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met
 515 520 525
 Gly Cys Thr Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser
 530 535 540
 Asp Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln
 545 550 555 560
 Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala
 565 570 575
 Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly
 580 585 590
 Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp
 595 600 605
 Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp Asn Asn Met
 610 615 620
 Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile
 625 630 635 640
 His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln
 645 650 655
 Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asn
 660 665 670
 Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu Phe Ile Met Ile Val Gly
 675 680 685
 Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile Val Asn
 690 695 700
 Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His Leu Pro
 705 710 715 720
 Ile Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu Gly Gly
 725 730 735
 Glu Arg Asp Arg Asp Arg Ser Ile Arg Leu Val Asn Gly Ser Leu Ala
 740 745 750
 Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg
 755 760 765
 Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu Leu Leu Gly
 770 775 780
 Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu Gln Tyr
 785 790 795 800
 Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Asn Leu Leu Asn Ala Thr
 805 810 815
 Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val Leu Gln
 820 825 830
 Ala Ala Tyr Arg Ala Ile Arg His Ile Pro Arg Arg Ile Arg Gln Gly
 835 840 845
 Leu Glu Arg Ile Leu Leu
 850

(2) ANGABEN ZU SEQ ID NO: 24:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 206 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (D) SONSTIGE ANGABEN: /product= "nef protein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 24:

Met	Gly	Gly	Lys	Trp	Ser	Lys	Ser	Ser	Val	Ile	Gly	Trp	Pro	Ala	Val
1				5					10					15	
Arg	Glu	Arg	Met	Arg	Arg	Ala	Glu	Pro	Ala	Ala	Asp	Gly	Val	Gly	Ala
			20					25					30		
Val	Ser	Arg	Asp	Leu	Glu	Lys	His	Gly	Ala	Ile	Thr	Ser	Ser	Asn	Thr
		35					40					45			
Ala	Ala	Asn	Asn	Ala	Ala	Cys	Ala	Trp	Leu	Glu	Ala	Gln	Glu	Glu	Glu
	50					55					60				
Glu	Val	Gly	Phe	Pro	Val	Thr	Pro	Gln	Val	Pro	Leu	Arg	Pro	Met	Thr
65					70					75					80
Tyr	Lys	Ala	Ala	Val	Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly
				85					90					95	
Leu	Glu	Gly	Leu	Ile	His	Ser	Gln	Arg	Arg	Gln	Asp	Ile	Leu	Asp	Leu
			100					105					110		
Trp	Ile	Tyr	His	Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr
		115					120					125			
Pro	Gly	Pro	Gly	Val	Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Tyr	Lys
	130					135					140				
Leu	Val	Pro	Val	Glu	Pro	Asp	Lys	Val	Glu	Glu	Ala	Asn	Lys	Gly	Glu
145					150					155					160
Asn	Thr	Ser	Leu	Leu	His	Pro	Val	Ser	Leu	His	Gly	Met	Asp	Asp	Pro
				165					170					175	
Glu	Arg	Glu	Val	Leu	Glu	Trp	Arg	Phe	Asp	Ser	Arg	Leu	Ala	Phe	His
			180					185					190		
His	Val	Ala	Arg	Glu	Leu	His	Pro	Glu	Tyr	Phe	Lys	Asn	Cys		
		195					200						205		

ANLAGE ZUM SEQUENZPROTOKOLL

ZU SEQ ID NO: 1

LOCUS LCVSRNA 3375 bp ss-RNA VRL 15-JUN-1989
 DEFINITION Lymphocytic choriomeningitis virus S RNA, complete cds.
 ACCESSION M22138
 NID g331379
 KEYWORDS S RNA; small RNA.
 SOURCE Lymphocytic choriomeningitis virus (strain WE), cDNA to viral RNA.
 ORGANISM Lymphocytic choriomeningitis virus
 Viruses; ssRNA negative-strand viruses; Arenaviridae; Arenavirus;
 1-LCMV-LASV complex.
 REFERENCE 1 (bases 1 to 3375)
 AUTHORS Romanowski,V., Matsuura,Y. and Bishop,D.H.L.
 TITLE Complete sequence of the S RNA of Lymphocytic choriomeningitis
 virus (WE strain) compared to that of Pichinde arenavirus
 JOURNAL Virus Res. 3, 101-114 (1985)
 MEDLINE 86046554
 FEATURES Location/Qualifiers
 source 1..3375
 /organism="Lymphocytic choriomeningitis virus"
 /db_xref="taxon:11623"
 CDS 78..1574
 /note="S protein"
 /codon_start=1
 /db_xref="PID:g331380"
 /translation=SEQ ID NO: 2
 BASE COUNT 881 a 786 c 725 g 983 t

ZU SEQ ID NO: 3

LOCUS LCVGPNP 3376 bp ss-RNA VRL 15-JUN-1989
 DEFINITION Lymphocytic choriomeningitis virus envelope glycoprotein (GP-C)
 and nucleoprotein (NP) genes, complete cds.
 ACCESSION M20869
 NID g331358
 KEYWORDS envelope protein; nucleoprotein.
 SOURCE Lymphocytic choriomeningitis virus (strain Armstrong 53b), cDNA
 to viral RNA.
 ORGANISM Lymphocytic choriomeningitis virus
 Viruses; ssRNA negative-strand viruses; Arenaviridae; Arenavirus;
 1-LCMV-LASV complex.
 REFERENCE 1 (bases 1 to 3376)
 AUTHORS Salvato,M., Shimomaye,E., Southern,P. and Oldstone,M.B.A.
 TITLE Virus-lymphocyte interactions: IV. Molecular characterization of
 LCMV Armstrong (CTL+) small genomic segment and that of its
 variant, clone 13 (CTL-)
 JOURNAL Virology 164, 517-522 (1988)
 MEDLINE 88219540
 FEATURES Location/Qualifiers
 source 1..3376
 /organism="Lymphocytic choriomeningitis virus"
 /db_xref="taxon:11623"
 CDS 78..1574
 /note="envelope glycoprotein"
 /codon_start=1
 /db_xref="PID:g331359"
 /translation=SEQ ID NO: 4
 variation 856
 /note="t in ARM 53b; c in ARM 53b Clone 13"

variation 1298
 /note="t in ARM 53b; c in ARM 53b Clone 13"
 CDS complement(1639..3315)
 /note="nucleoprotein"
 /codon_start=1
 /db_xref="PID:g331360"
 /translation=SEQ ID NO: 5
 BASE COUNT 868 a 809 c 748 g 951 t

ZU SEQ ID NO: 6

LOCUS LCVLPY 6680 bp ss-RNA VRL 17-MAY-1995
 DEFINITION Lymphocytic choriomeningitis virus putative RNA polymerase L gene, complete cds.
 ACCESSION J04331
 NID g331368
 KEYWORDS L protein; RNA polymerase.
 SOURCE Lymphocytic choriomeningitis virus (strain Armstrong 53b) RNA.
 ORGANISM Lymphocytic choriomeningitis virus
 Viruses; ssRNA negative-strand viruses; Arenaviridae; Arenavirus; 1-LCMV-LASV complex.
 REFERENCE 1 (bases 1 to 6680)
 AUTHORS Salvato,M., Shimomaye,E. and Oldstone,M.B.
 TITLE The primary structure of the lymphocytic choriomeningitis virus L gene encodes a putative RNA polymerase
 JOURNAL Virology 169 (2), 377-384 (1989)
 MEDLINE 89204909
 COMMENT Draft entry and computer-readable sequence of [1] kindly submitted by M.Salvato, 18-JAN-1989.
 FEATURES Location/Qualifiers
 source 1..6680
 /organism="Lymphocytic choriomeningitis virus"
 /strain="Armstrong 53b"
 /db_xref="taxon:11623"
 CDS 33..6665
 /codon_start=1
 /product="L protein"
 /db_xref="PID:g331369"
 /translation=SEQ ID NO: 7
 CDS complement(6371..6658)
 /note="ORF; putative"
 /codon_start=1
 /product="unknown protein"
 /db_xref="PID:g808709"
 /translation=SEQ ID NO: 8
 BASE COUNT 2065 a 1153 c 1543 g 1919 t

ZU SEQ ID NO: 9

LOCUS HUMEF1A 4695 bp DNA PRI 07-NOV-1994
 DEFINITION Human elongation factor EF-1-alpha gene, complete cds.
 ACCESSION J04617 J04616
 NID g181962
 KEYWORDS elongation factor.
 SOURCE Human placenta DNA, clone pEFG1, and fibroblast cell line GM 637, cDNA to mRNA, (library of H.Okayama), clone pAN7.
 ORGANISM Homo sapiens
 Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4695)
 AUTHORS Uetsuki,T., Naito,A., Nagata,S. and Kaziro,Y.
 TITLE Isolation and characterization of the human chromosomal gene for

polypeptide chain elongation factor-1 alpha
 JOURNAL J. Biol. Chem. 264 (10), 5791-5798 (1989)
 MEDLINE 89174636
 COMMENT Draft entry and computer-readable sequence for [1] kindly provided
 by S.Nagata, 20-JAN-1989.

FEATURES Location/Qualifiers

source 1..4695
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="Unassigned"

misc_binding 205..214
 /bound_moiety="Sp1"

misc_binding 320..328
 /bound_moiety="Sp1"

misc_binding 332..340
 /bound_moiety="Sp1"

TATA_signal 546..552

prim_transcript 576..4087
 /note="EF-1-alpha mRNA and introns"

intron 609..1551
 /note="EF-1-alpha intron A"

misc_binding 983..992
 /bound_moiety="Sp1"

misc_binding 1026..1034
 /bound_moiety="Sp1"

misc_binding 1122..1131
 /bound_moiety="Sp1"

misc_binding 1132..1141
 /bound_moiety="Sp1"

misc_binding 1240..1249
 /bound_moiety="Sp1"

misc_binding 1302..1308
 /bound_moiety="Ap1"

exon <1582..1725
 /gene="EEF1A"
 /note="elongation factor EF-1-alpha"
 /number=2

gene join(1582..1725,2092..2271,2377..2673,2757..2907,
 2995..3251,3341..3575,3671..3795)
 /gene="EEF1A"

CDS join(1582..1725,2092..2271,2377..2673,2757..2907,
 2995..3251,3341..3575,3671..3795)
 /gene="EEF1A"
 /note="elongation factor EF-1-alpha"
 /codon_start=1
 /db_xref="GDB:G00-118-791"
 /db_xref="PID:g181963"
 /translation=SEQ ID NO: 10

intron 1726..2091
 /note="EF-1-alpha intron B"

exon 2092..2271
 /gene="EEF1A"
 /number=3

intron 2272..2376
 /note="EF-1-alpha intron C"

exon 2377..2673
 /gene="EEF1A"
 /number=4

intron 2674..2756
 /note="EF-1-alpha intron D"

exon 2757..2907
 /gene="EEF1A"
 /number=5

intron 2908..2994

```

                /note="EF-1-alpha intron E"
exon            2995..3251
                /gene="EEF1A"
                /number=6
intron          3252..3340
                /note="EF-1-alpha intron F"
exon            3341..3575
                /gene="EEF1A"
                /number=7
intron          3576..3670
                /note="EF-1-alpha intron G"
exon            3671..>3795
                /gene="EEF1A"
                /note="elongation factor EF-1-alpha"
                /number=8
BASE COUNT     1200 a      989 c      1235 g      1271 t

```

ZU SEQ ID NO: 11

```

LOCUS           AF033811      8332 bp      RNA           VRL           05-FEB-1998
DEFINITION     Moloney murine leukemia virus, complete genome.
ACCESSION      AF033811
NID            g2801468
KEYWORDS       .
SOURCE         Moloney murine leukemia virus.
  ORGANISM     Moloney murine leukemia virus
               Viruses; Retroid viruses; Retroviridae; Mammalian type C
               retroviruses; 1-Mammalian type C virus group.
REFERENCE      1 (bases 1 to 8332)
AUTHORS        Petropoulos,C.J.
TITLE          Appendix 2: Retroviral taxonomy, protein structure, sequences, and
               genetic maps
JOURNAL        (in) Coffin,J.M. (Ed.);
               RETROVIRUSES: 757;
               Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York,
               NY, USA (1997)
REFERENCE      2 (bases 1 to 8332)
AUTHORS        Chappay,C.
TITLE          Direct Submission
JOURNAL        Submitted (12-NOV-1997) NIH, NLM, Rockville Pike, Bethesda, MD
               20894, USA
FEATURES       Location/Qualifiers
  source        1..8332
               /organism="Moloney murine leukemia virus"
               /db_xref="taxon:11801"
  mRNA         1..8332
               /gene="gag-pol"
  mRNA         join(1..205,5491..8332)
               /gene="env"
               /product="gPr80"
  5'UTR        69..145
  gene         357..1973
               /gene="gag"
  CDS          357..1973
               /gene="gag"
               /codon_start=1
               /product="Pr65"
               /db_xref="PID:g2801469"
               /translation=SEQ ID NO: 12
  CDS          join(357..1970,1974..5573)
               /gene="gag-pol"
               /codon_start=1

```

```

/product="Pr180"
/db_xref="PID:g2801471"
/translation=SEQ ID NO: 13
mat_peptide 360..749
/gene="gag"
mat_peptide /product="p15 MA"
750..1001
/gene="gag"
mat_peptide /product="pp12"
1002..1790
/gene="gag"
mat_peptide /product="p30 CA"
1791..1958
/gene="gag"
mat_peptide /product="p10 NC"
join(1959..1970,1974..2336)
/gene="gag"
gene /product="p14 PR"
1970..5573
/gene="pol"
mat_peptide 2337..4349
/gene="pol"
mat_peptide /product="p80 RT"
4350..5570
/gene="pol"
CDS /product="p46 IN"
5513..7510
/gene="env"
/codon_start=1
/product="gPr80"
/db_xref="PID:g2801470"
/translation=SEQ ID NO: 14
mat_peptide 5612..6919
/product="gp70 SU"
mat_peptide 6920..7507
/product="p15E"
mat_peptide 6920..7507
/product="p12E TM"
3'UTR 7818..8264
BASE COUNT 2143 a 2395 c 2025 g 1769 t

```

ZU SEQ ID NO: 15

```

LOCUS      HIVNL43          9709 bp ss-RNA          VRL          15-JUN-1989
DEFINITION Human immunodeficiency virus type 1, NY5/BRU (LAV-1) recombinant
clone pNL4-3.
ACCESSION M19921
NID       g328415
KEYWORDS
SOURCE    Human immunodeficiency virus type 1 (HIV-1), NY5/BRU (LAV-1)
recombinant clone pNL4-3.
ORGANISM  Human immunodeficiency virus type 1
Viruses; Retrovird viruses; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE 1 (bases 1 to 9709)
AUTHORS  Adachi,A., Gendelman,H.E., Koenig,S., Folks,T., Willey,R.,
Rabson,A. and Martin,M.A.
TITLE    Production of acquired immunodeficiency syndrome-associated
retrovirus in human and nonhuman cells transfected with an
infectious molecular clone
JOURNAL  J. Virol. 59, 284-291 (1986)
MEDLINE  86281827
REFERENCE 2 (bases 1 to 9709)
AUTHORS  Buckler,C.E., Buckler-White,A.J., Willey,R.L. and McCoy,J.

```


JOURNAL Unpublished (1988)
REFERENCE 3 (sites)
AUTHORS Dai,L.C., Littaua,R., Takahashi,K. and Ennis,F.A.
TITLE Mutation of human immunodeficiency virus type 1 at amino acid 585
on gp41 resultis in loss of killing by CD8+ A24-restricted
cytotoxic T lymphocytes
J. Virol. 66, 3151-3154 (1992)
MEDLINE 92219406
COMMENT [3] sites; revisions of [3].
Clean copy of sequence [3] kindly provided by Chuck Buckler,
NIAID, Bethesda, MD, 24-JUN-1988. The construction of pNL4-3 has
been described in [1]. pNL4-3 is a recombinant (infectious)
proviral clone that contains DNA from HIV isolates NY5 (5' half)
and BRU (3'half). The site of recombination is the EcoRI site at
positions 5743-5748.
The length and sequence of the vpr coding region corresponds to
that of the BRU, SC, SF2, MAL and ELI isolates. The vpr coding
region of these isolates is about 18 amino acid residues longer
than the vpr coding region of the IIIb isolates. In HIVNL43, this
shift is due to a single base deletion (with respect to the
IIIb's) at position 5770. The sequence at this position is
'atttc' in HIVNL43 and 'atattc' in HIVHXB2.
The original BRU clone, sequenced by Wain-Hobson, et al. (Cell 40,
9-17 (1985)), and the BRU portion of the pNL4-3 recombinant clone
are different clones from the same BRU isolate.
Two of the revisions reported in the FEATURES produced changes in
amino acid sequences. The revision at position 2421 changes one
amino acid residue from 'R' to 'G' in the pol coding region. The
revision at positions 8995-9000 changes three amino acid residues
from 'AHT' to 'VTP' in the nef coding region.

FEATURES Location/Qualifiers
source 1..9709
/organism="Human immunodeficiency virus type 1"
/db_xref="taxon:11676"
LTR 1..634
/note="5' LTR"
repeat_region 454..550
/note="R repeat 5' copy"
prim_transcript 455..9626
/note="tat, rev, nef subgenomic mRNA"
intron 744..5776
/note="tat, rev, nef mRNA intron 1"
CDS 790..2292
/note="gag polyprotein"
/codon_start=1
/db_xref="PID:g328418"
/translation=SEQ ID NO: 16
CDS 2085..5096
/partial
/note="pol polyprotein (NH2-terminus uncertain)"
/codon_start=1
/db_xref="PID:g328419"
/translation=SEQ ID NO: 17
CDS 5041..5619
/note="vif protein"
/codon_start=1
/db_xref="PID:g328420"
/translation=SEQ ID NO: 18
CDS 5559..5849
/note="vpr protein"
/codon_start=1
/db_xref="PID:g328421"
/translation=SEQ ID NO: 19
misc_feature 5743..5748

```

CDS      /note="EcoRI site of recombination"
         join(5830..6044,8369..8414)
         /note="tat protein"
         /codon_start=1
         /db_xref="PID:g328416"
         /translation=SEQ ID NO: 20
CDS      join(5969..6044,8369..8643)
         /note="rev protein"
         /codon_start=1
         /db_xref="PID:g328417"
         /translation=SEQ ID NO: 21
intron   6045..8368
         /note="tat cds intron 2"
intron   6045..8368
         /note="tat, rev, nef mRNA intron 2"
intron   6045..8368
         /note="rev cds intron 2"
CDS      6061..6306
         /note="vpu protein"
         /codon_start=1
         /db_xref="PID:g328422"
         /translation=SEQ ID NO: 22
CDS      6221..8785
         /note="envelope polyprotein"
         /codon_start=1
         /db_xref="PID:g328423"
         /translation=SEQ ID NO: 23
CDS      8787..9407
         /note="nef protein"
         /codon_start=1
         /db_xref="PID:g328424"
         /translation=SEQ ID NO: 24

```

Patentansprüche

1. Retrovirale Verpackungszelle, **dadurch gekennzeichnet**, daß sie pseudotypisierte Virionen exprimiert, die LCMV-Glykoprotein in ihrer Hülle eingebaut enthalten.
2. Retrovirale Verpackungszelle nach Anspruch 1, dadurch gekennzeichnet, daß sie die retroviralen Gene gag und pol und ferner das für die Glykoproteine GP-1 und GP-2 des LCMV kodierende Gen gp oder einen Teil desselben enthält, wobei der Teil des gp-Gens von LCMV die Pseudotypisierung der Verpackungszellen gewährleistet und die Transfektion der Zielzellen und die stabile Integration von Transgenen nicht behindert.
3. Retrovirale Verpackungszelle nach den Ansprüchen 1 und 2, dadurch gekennzeichnet, daß sie ferner das retrovirale Gen env und/oder regulatorische retrovirale Gene enthält.
4. Retrovirale Verpackungszelle nach den Ansprüchen 1 bis 3, dadurch gekennzeichnet, daß sie ferner mindestens ein für das Nukleoprotein kodierendes Gen np, für die RNA-Polymerase kodierendes Gen 1 und/oder für ein Protein mit unbekannter Funktion kodierendes Gen z des LCMV enthält.
5. Retrovirale Verpackungszelle nach Anspruch 4, dadurch gekennzeichnet, daß das Retrovirus MLV-related virus oder Lentivirus ist.
6. Retrovirale Verpackungszelle nach Anspruch 5, dadurch gekennzeichnet, daß das Retrovirus MLV, HIV, SIV oder FIV ist.
7. Retrovirale Verpackungszelle nach den Ansprüchen 1 bis 6, dadurch gekennzeichnet, daß sie ein oder mehrere Transgene enthält.
8. Retrovirale Verpackungszelle nach Anspruch 7, dadurch gekennzeichnet, daß das Transgen/die Transgene neo, lacZ, EGFP, Ribozyme, mdr-1, das Suizidgen Herpes Simplex Virus Thymidin Kinase (HSV-tk) und/oder das Suizidgen Cytosin-Deaminase (CD) ist/sind.
9. Retrovirale Verpackungszelle nach den Ansprüchen 7 und 8, dadurch gekennzeichnet, daß das Transgen/die Transgene Markergene, antiviral wirksame Sequenzen, antisense Sequenzen, transdominant-negativ wirkende Gene und/oder Zytokin-Gene ist/sind.

10. Retrovirale Verpackungszelle nach den Ansprüchen 1 bis 9, dadurch gekennzeichnet, daß das Retrovirus MLV ist, das kein Env-Protein exprimiert und LCMV die defekte Mutante L(ARM) ist.
11. Retrovirale Verpackungszelle nach den Ansprüchen 1 bis 9, dadurch gekennzeichnet, daß sie durch Transfektion einer Verpackungszelle, die die retroviralen Gene gag und pol enthält, mit einem Expressionsplasmid erhältlich ist, das das gp-Gen des LCMV oder einen Teil desselben enthält, wobei der Teil des gp-Gens von LCMV die Pseudotypisierung der Verpackungszellen gewährleistet und die Transfektion der Zielzellen und die stabile Integration von Transgenen nicht behindert.
12. Retrovirale Verpackungszelle nach Anspruch 11, dadurch gekennzeichnet, daß sie ferner das retrovirale Gen env, regulatorische retrovirale Gene und/oder das np-Gen, das I-Gen und/oder das z-Gen des LCMV oder Teile derselben enthält.
13. Pseudotypisiertes Virion, dadurch gekennzeichnet, daß es von einer Verpackungszelle nach den Ansprüchen 1 bis 6 exprimiert ist.
14. Retroviraler Pseudotyp-Vektor, erhältlich durch Kultivieren von Verpackungszellen nach den Ansprüchen 1 und 7 bis 12.
15. Verfahren zur Herstellung einer retroviralen Verpackungszelle nach den Ansprüchen 1 und 7 bis 12, dadurch gekennzeichnet, daß man eine mit einem oder mehreren Transgenen transfizierte retrovirale Verpackungszelle, die die retroviralen Gene gag und pol enthält, durch Infektion mit LCMV pseudotypisiert.
16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, daß die retrovirale Verpackungszelle ferner das retrovirale Gen env und/oder regulatorische retrovirale Gene enthält.
17. Verfahren nach den Ansprüchen 15 und 16, dadurch gekennzeichnet, daß man als LCMV die deletierte Mutante L(ARM) verwendet.
18. Verfahren zur Herstellung einer retroviralen Verpackungszelle nach den Ansprüchen 1 und 7 bis 12, dadurch gekennzeichnet, daß man eine mit einem oder mehreren Fremdgenen transfizierte retrovirale Verpackungszelle, die die retroviralen Gene gag und pol enthält, mit einem Expressionsplasmid transfiziert, das das gp-Gen des LCMV oder einen Teil desselben enthält, wobei der Teil des gp-Gens von LCMV die Pseudotypisierung der Verpackungszellen gewährleistet und die Transfektion der Zielzellen und die stabile Integration von Transgenen nicht behindert.
19. Verfahren nach Anspruch 18, dadurch gekennzeichnet, daß die retrovirale Verpackungszelle ferner das retrovirale Gen env, regulatorische retrovirale Gene und/oder ein oder mehrere Gene np, 1 und/oder z des LCMV enthält.
20. Verwendung einer retroviralen Verpackungszelle nach den Ansprüchen 1 und 7 bis 12 zur in vitro-Infektion von Zellen und zur Expression des Transgens in diesen Zellen.
21. Verwendung einer retroviralen Verpackungszelle nach den Ansprüchen 1 und 7 bis 12 zur Gentherapie.
22. Verwendung nach den Ansprüchen 20 und 21 zur Transfektion hämatopoietischer Stammzellen.
23. Verwendung nach Anspruch 21 zur Therapie von infektiösen Erkrankungen, insbesondere AIDS, und Neoplasien, insbesondere des Mamma-Karzinoms, oder des Melanoms sowie von anderen durch Gentherapie zugänglichen Erkrankungen.
24. Verwendung nach den Ansprüchen 20 bis 22, bei der man die Verpackungszelle mit pharmazeutisch verträglichen Hilfs- und/oder Trägerstoffen formuliert.
25. Pharmazeutisches Präparat, dadurch gekennzeichnet, daß es retrovirale Verpackungszellen nach einem der Ansprüche 1 und 7 bis 12 enthält.
26. Pharmazeutisches Präparat nach Anspruch 25, dadurch gekennzeichnet, daß es ferner pharmazeutisch verträgliche Hilfs- und/oder Trägerstoffen enthält.

Es folgen 7 Blatt Zeichnungen

Fig. 1

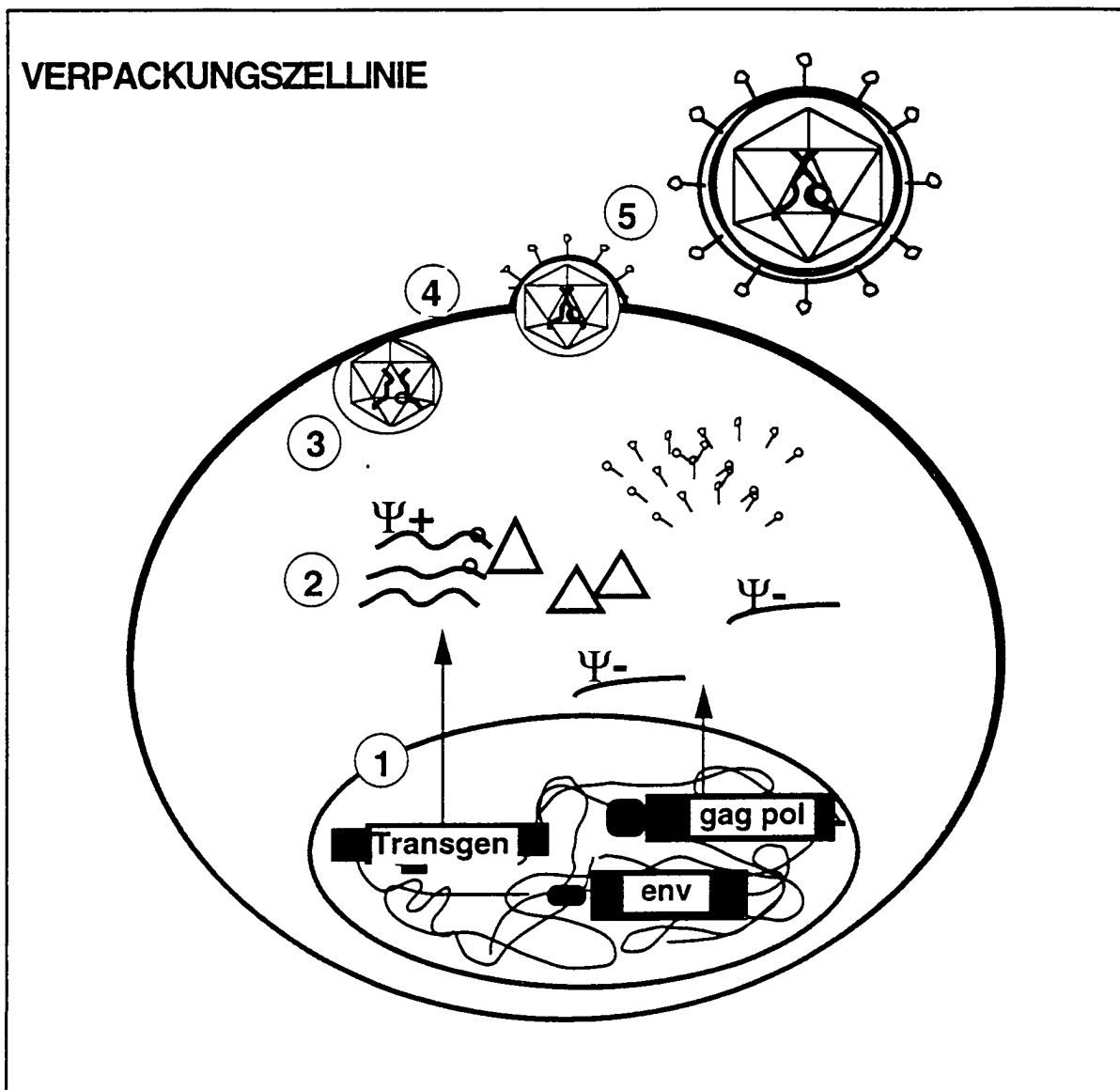
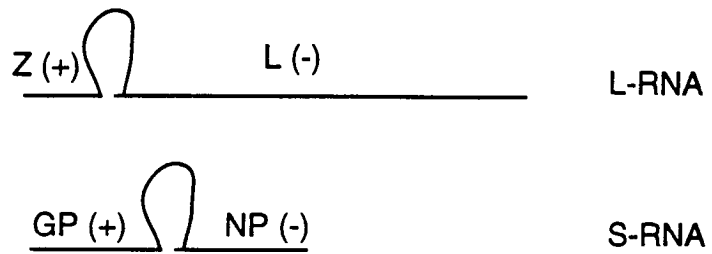
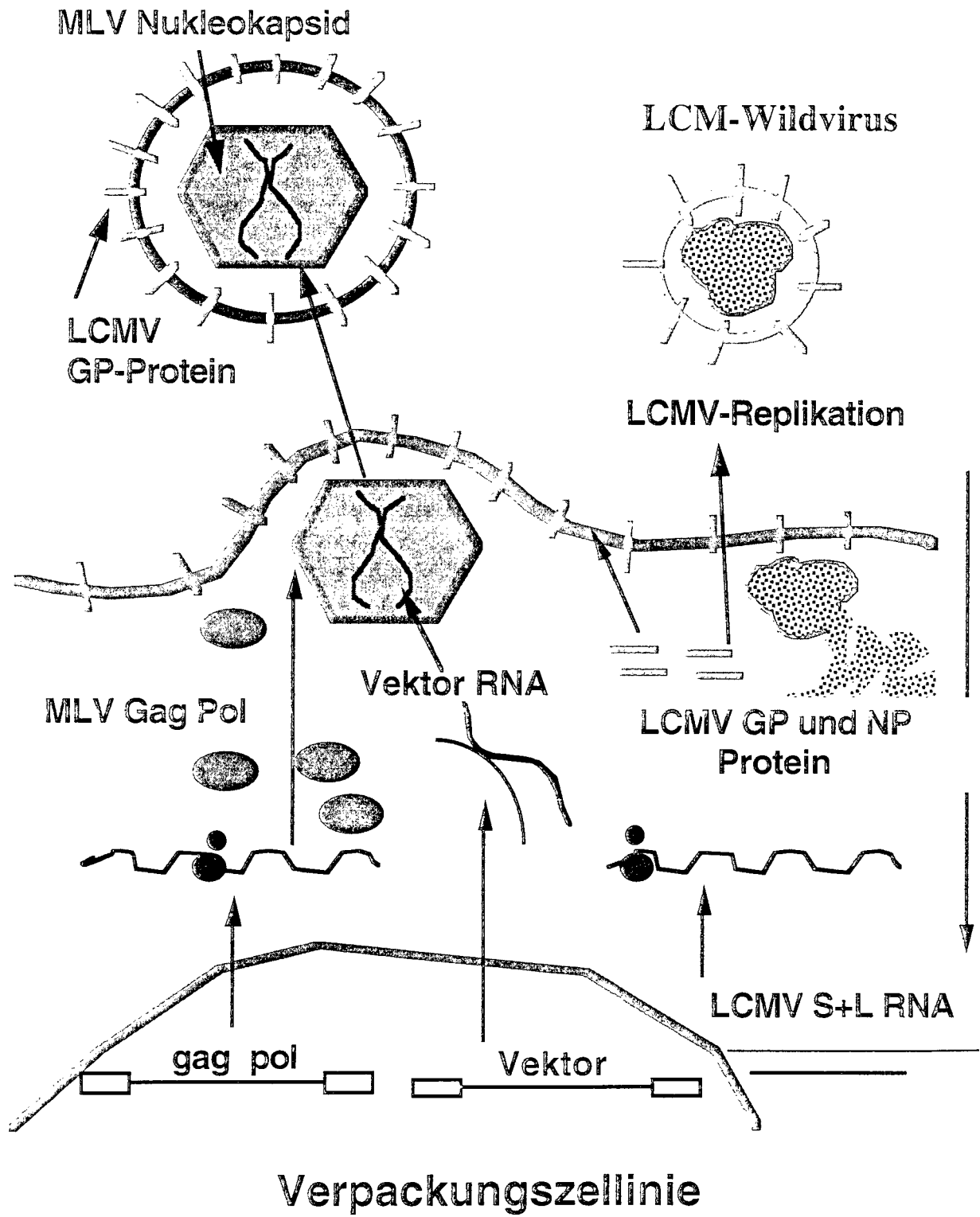


Fig. 2



MLV/LCMV-Pseudotyp

Fig. 3



Ektope Expression des LCMV-GP

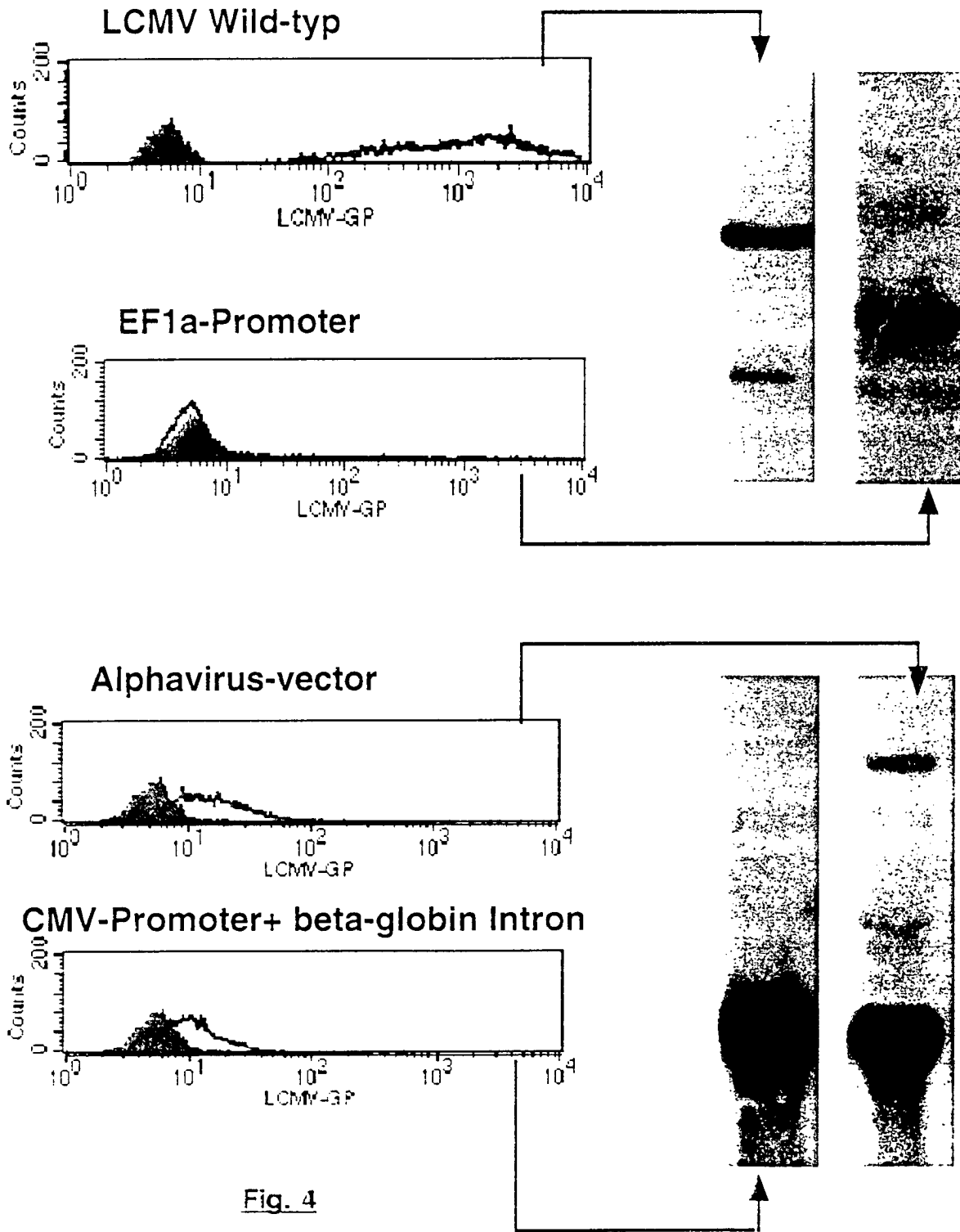
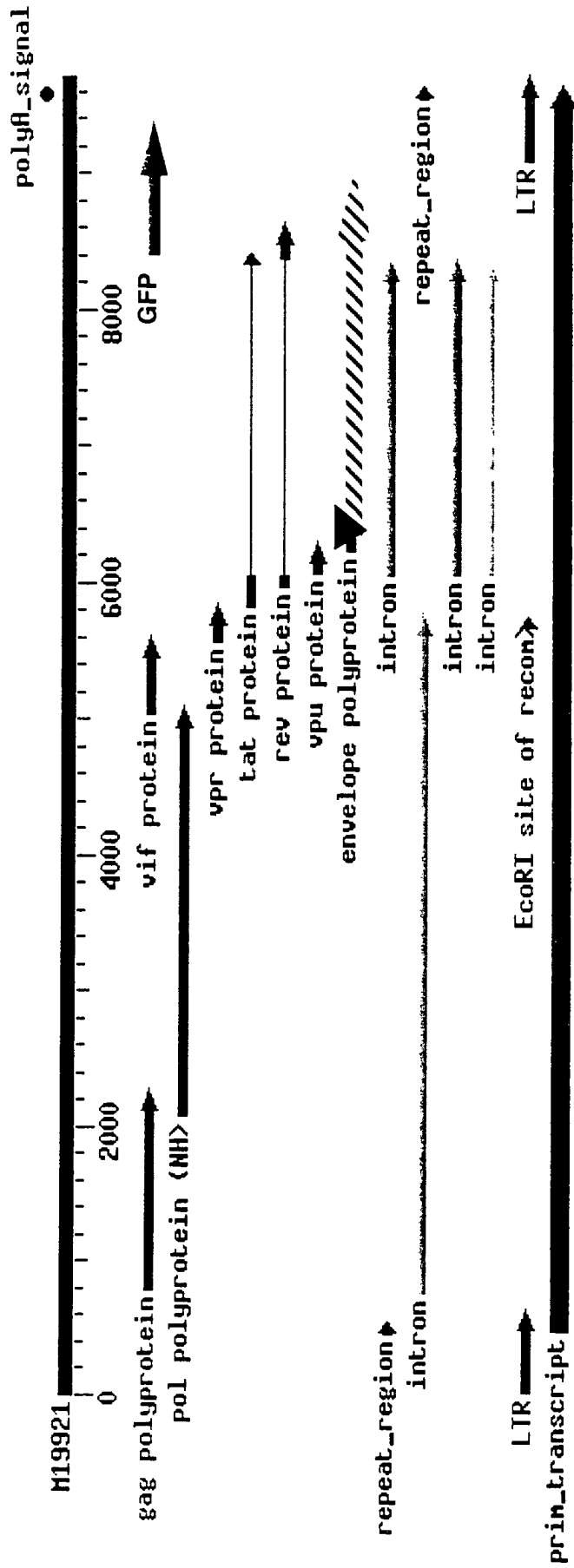


Fig. 5



▼ Verschiebung des Leserahmens

Fig. 6

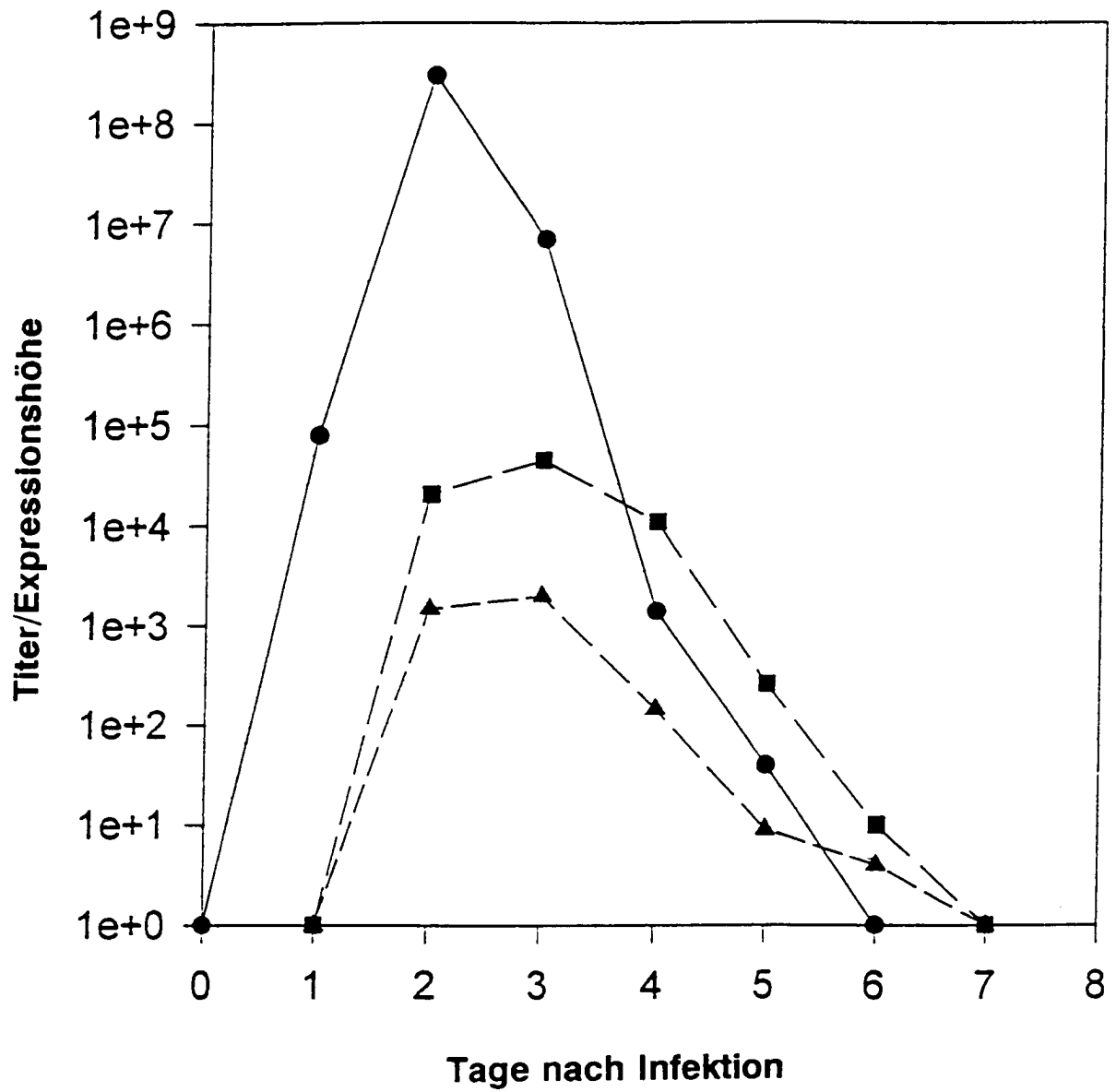


Fig. 7

