

PCT

WELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationales Büro



INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation ⁷ : A61K 39/00		A2	(11) Internationale Veröffentlichungsnummer: WO 00/47223 (43) Internationales Veröffentlichungsdatum: 17. August 2000 (17.08.00)																								
(21) Internationales Aktenzeichen: PCT/EP99/09759 (22) Internationales Anmeldedatum: 3. Dezember 1999 (03.12.99) (30) Prioritätsdaten: 199 07 485.2 12. Februar 1999 (12.02.99) DE (71) Anmelder (für alle Bestimmungsstaaten ausser US): STRATHMANN AG & CO. [DE/DE]; Sellhopsweg 1, D-22459 Hamburg (DE). (72) Erfinder; und (75) Erfinder/Anmelder (nur für US): SCHREIBER, Michael [DE/DE]; Heidberg 35, D-22301 Hamburg (DE). (74) Anwälte: WEBER-QUITZAU, Martin usw.; Uexküll & Stolberg, Beselerstrasse 4, D-22607 Hamburg (DE).		(81) Bestimmungsstaaten: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Veröffentlicht <i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i>																									
(54) Title: VIRAL VACCINE (54) Bezeichnung: VIRUS-VAKZINE (57) Abstract <p>The invention relates to a pharmaceutical composition or vaccine containing a mixture of viral protein molecules which are sequence variants of a single viral protein or of part of such a protein. The invention also relates to a DNA vaccine coding for a mixture of structurally different virus proteins. Said vaccine contains a mixture of sequence variants of a viral DNA molecule or of a part thereof which code for sequence variants of a viral protein or of a part thereof. According to a preferred embodiment of the invention the viral proteins are sequence variants of the GP120 protein of the human immunodeficiency virus (HIV) which differ from each other in terms of the amino acid sequence in the area of the V2-loop and/or the V3-loop, preferably both the V2- and V3-loop. The invention also relates to the production of said viral vaccines, including the related intermediate stages and constructs, as well as to their methods of production and their uses.</p>																											
<p style="text-align: center;">V3 LOOP SEQUENCE DATA OF HIV-1 PATIENT ISOLATES (PI) V3-Loop Sequenzdaten von HIV-1 Patientenisolaten (PI).</p> <table><tr><td></td><td>CTRPNNNTRKSI.HIGPGRAFYTATGDIIGDIRQAHC</td></tr><tr><td>PI-903</td><td>-----G--.-----STN-----A-----S-----</td></tr><tr><td>PI-951</td><td>----H-----N-----W-T-----</td></tr><tr><td>PI-918</td><td>---S-----E-----</td></tr><tr><td>PI-970</td><td>---S-----E-----</td></tr><tr><td>PI-990</td><td>-----R--.P-----T---V-----</td></tr><tr><td>PI-991</td><td>-----P-A-----E---N-----</td></tr><tr><td>PI-952</td><td>-I-----R-.TL---VL-T-E-----K---</td></tr><tr><td>PI-932</td><td>-I--H-TVTDR-.-----S-HT-RKIK-----</td></tr><tr><td>PI-910</td><td>----SIQK-R-V.R-----S-I-ARAAT----K-Q-</td></tr><tr><td>PI-911</td><td>----SIQK-R-V.R-----S-I--RAAT----K-Q-</td></tr><tr><td>PI-930</td><td>----YR-AKHR-M-----NVKGN1K;----</td></tr></table>					CTRPNNNTRKSI.HIGPGRAFYTATGDIIGDIRQAHC	PI-903	-----G--.-----STN-----A-----S-----	PI-951	----H-----N-----W-T-----	PI-918	---S-----E-----	PI-970	---S-----E-----	PI-990	-----R--.P-----T---V-----	PI-991	-----P-A-----E---N-----	PI-952	-I-----R-.TL---VL-T-E-----K---	PI-932	-I--H-TVTDR-.-----S-HT-RKIK-----	PI-910	----SIQK-R-V.R-----S-I-ARAAT----K-Q-	PI-911	----SIQK-R-V.R-----S-I--RAAT----K-Q-	PI-930	----YR-AKHR-M-----NVKGN1K;----
	CTRPNNNTRKSI.HIGPGRAFYTATGDIIGDIRQAHC																										
PI-903	-----G--.-----STN-----A-----S-----																										
PI-951	----H-----N-----W-T-----																										
PI-918	---S-----E-----																										
PI-970	---S-----E-----																										
PI-990	-----R--.P-----T---V-----																										
PI-991	-----P-A-----E---N-----																										
PI-952	-I-----R-.TL---VL-T-E-----K---																										
PI-932	-I--H-TVTDR-.-----S-HT-RKIK-----																										
PI-910	----SIQK-R-V.R-----S-I-ARAAT----K-Q-																										
PI-911	----SIQK-R-V.R-----S-I--RAAT----K-Q-																										
PI-930	----YR-AKHR-M-----NVKGN1K;----																										

(57) Zusammenfassung

Die vorliegende Erfindung betrifft eine pharmazeutische Zusammensetzung bzw. eine Vakzine, die eine Mischung viraler Protein-Moleküle umfaßt, die Sequenzvarianten eines einzigen viralen Proteins oder eines Teils desselben sind. Die Erfindung betrifft ferner u.a. eine DNA-Vakzine, die für eine Mischung strukturellunterschiedlicher Virus-Proteine kodiert, wobei die Vakzine eine Mischung von Sequenzvarianten eines viralen DNA-Moleküls oder eines Teils desselben enthält, die Sequenzvarianten eines viralen Proteins oder eines Teils kodieren. Gemäß einer bevorzugten Ausführungsform der Erfindung handelt es sich bei den viralen Proteinen um Sequenzvarianten des GP120-Proteins des Humanen Immundefizienzvirus (HIV), die sich jeweils in ihrer Aminosäuresequenz im Bereich der V2-Schleife und/oder der V3-Schleife, vorzugsweise sowohl der V2- als auch der V3-Schleife, voneinander unterscheiden. Die Erfindung betrifft ferner die Herstellung der Virus-Vakzinen einschließlich der damit im Zusammenhang stehenden Zwischenstufen bzw. Konstrukte, Herstellungsverfahren und Verwendungen.

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland			TR	Türkei
BG	Bulgarien	HU	Ungarn	ML	Mali	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MN	Mongolei	UA	Ukraine
BR	Brasilien	IL	Israel	MR	Mauretanien	UG	Uganda
BY	Belarus	IS	Island	MW	Malawi	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	MX	Mexiko		
CF	Zentralafrikanische Republik	JP	Japan	NE	Niger	UZ	Usbekistan
CG	Kongo	KE	Kenia	NL	Niederlande	VN	Vietnam
CH	Schweiz	KG	Kirgisistan	NO	Norwegen	YU	Jugoslawien
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	NZ	Neuseeland	ZW	Zimbabwe
CM	Kamerun			PL	Polen		
CN	China	KR	Republik Korea	PT	Portugal		
CU	Kuba	KZ	Kasachstan	RO	Rumänien		
CZ	Tschechische Republik	LC	St. Lucia	RU	Russische Föderation		
DE	Deutschland	LI	Liechtenstein	SD	Sudan		
DK	Dänemark	LK	Sri Lanka	SE	Schweden		
EE	Estland	LR	Liberia	SG	Singapur		

VIRUS-VAKZINE

Die vorliegende Erfindung betrifft eine pharmazeutische Zusammensetzung bzw. eine Vakzine, die eine Mischung viraler Protein-Moleküle umfaßt, die Sequenzvarianten eines einzigen viralen Proteins oder eines Teils desselben sind. Die Erfindung betrifft ferner u.a. eine DNA-Vakzine, die für eine Mischung strukturell unterschiedlicher Virus-Proteine kodiert, wobei die Vakzine eine Mischung von Sequenzvarianten eines viralen DNA-Moleküls oder eines Teils desselben enthält, die Sequenzvarianten eines viralen Proteins oder eines Teils kodieren. Gemäß einer bevorzugten Ausführungsform der Erfindung handelt es sich bei den viralen Proteinen um Sequenzvarianten des GP120-Proteins des Humanen Immundefizienzvirus (HIV), die sich jeweils in ihrer Aminosäuresequenz im Bereich der V2-Scheife und/oder der V3-Schleife, vorzugsweise sowohl der V2- als auch der V3-Schleife, voneinander unterscheiden. Die Erfindung betrifft ferner die Herstellung der Virus-Vakzinen einschließlich der damit im Zusammenhang stehenden Zwischenstufen bzw. Konstrukte, Herstellungsverfahren und Verwendungen.

Bei vielen viralen Infektionen, insbesondere bei der HIV-1-Infektion, beobachtet man eine starke Immunabwehr, die in der

- 2 -

Lage ist, das Virus über einen Zeitraum von mehreren Jahren zu kontrollieren. Den Zeitraum, in dem das Virus kontrolliert wird und keine Krankheitssymptome beobachtet werden, bezeichnet man als die asymptomatische Phase der (HIV-)Erkrankung. Im Verlauf
5 der Erkrankung entstehen immer wieder neue Virusvarianten. Dadurch ist es dem Virus möglich, der Immunabwehr des Menschen zu entkommen und immer wieder neue Abwehrzellen des Immunsystems zu infizieren (vgl. M. Schreiber et al., J. Virol. 68 Nr. 6 (1994) 3908-3916; J. Gen. Virol. 77 (1996) 2403-2414; Clin. Exp.
10 Immunol. 107 (1997) 15-20; J. Virol. 71 Nr. 12 (1997) 9198-9205).

Im Stand der Technik sind zur Behandlung viraler Erkrankungen, wie z. B. der Poliovirus- (Heraud F et al., Biologicals, 1993, 21:311-316), Hantavirus- (Ulrich R et al., 1998 Vaccine 16:272-
15 280; Schmaljohn CS et al., 1992 Vaccine 10:10-13), Lassavirus- (Morrison HG et al., 1989 Virology 171:179-188; Clegg JC et al., 1987 Lancet 2:186-188), Hepatitis-A-Virus- (Clemens et al., 1995 J. Infect. Dis. 171:Suppl1:S44-S49; Andre et al., 1990 Prog. Med. Virol. 37:72-95) und Hepatitis-B-Virus-Infektion (McAleer et al.,
20 1984 Nature 307:178-180) aber auch der HIV-Infektion (Egan et al., 1995 J. Infect. Dis. 171:1623-1627, Kovac et al., 1993 J. Clin. Invest. 92:919-928) nur einzelne, genetisch unveränderte, spezifische Virus-Antigene oder einzelne inaktivierte Virusstämme zu Untersuchungen geeigneter Vakzinierungsstrategien eingesetzt
25 worden. Im Falle von HIV wurden bisher beispielsweise die externen Hüllproteine von zwei Virusstämmen für Impfstoff-experimente am Menschen hergestellt (MN und SF2) (Zolla-Pazner et al., J. Infect. Dis. 178 (1998) 1502-1506). Beide GP120-Moleküle unterscheiden sich in der Aminosäuresequenz, besonders
30 aber in der Aminosäuresequenz der variablen Bereiche, wie z.B. der V3-Schleife (V3-Loop). Die beiden Virusstämme haben aufgrund der verschiedenen V3-Loop-Sequenzen unterschiedliche phänotypische Eigenschaften. Der HIV-1 MN-Stamm ist eine Virusvariante, die bevorzugt Makrophagen und Zellen, die den viralen Korezeptor
35 CCR5 besitzen, infiziert. Viren des SF2-Typs dagegen vermehren sich bevorzugt in T-Zellen und benutzen den viralen Korezeptor

- 3 -

CXCR4. Man nennt solche Viren daher auch T-Zelltrophe (z.B. HIV-Stamm SF2) oder Makrophagotrophe (z.B. HIV-Stamm MN) Virusvarianten. Im Schimpansen sind mit diesen beiden GP120-Varianten Impfexperimente durchgeführt worden. In diesen Experimenten ist
5 gezeigt worden, daß man eine Immunantwort induzieren kann, die nicht nur gegen die beiden HIV-Stämme MN und SF2 schützt, sondern auch in der Lage ist, die Infektion mit anderen Virusstämmen, aber auch mit HIV-1-Patientenisolaten zu verhindern. Im Unterschied zum Schimpansen hat man beim Menschen bisher nur eine der
10 beiden GP120-Varianten für Impfexperimente eingesetzt. Sowohl MN GP120 als auch SF2 GP120 schützen nicht mit Sicherheit vor einer Infektion durch eine Virusvariante wie sie in einem HIV-1-Infizierten vorkommt (Patienten-Isolat oder Wildtyp-Virus). Einen Überblick über den Stand der Forschung im Zusammenhang mit der
15 Vakzinierung mit GP120-Hüllprotein bietet J.A. Levy in "*HIV and the Pathogenesis of AIDS*", Herausgeber: Jay A. Levy, Kapitel 15, 2. Auflage, ASM Press Washington, D.C., 1998).

Der Nachteil der im Stand der Technik bislang diskutierten bzw.
20 untersuchten Vakzinierungsstrategien besteht unter anderem darin, daß man mit den eingesetzten Vakzinen nicht in der Lage ist, die Entstehung immer neuer Virusvarianten im Verlauf der Virus-Erkrankung zu verhindern.

25 Aufgabe der vorliegenden Erfindung ist es daher, eine Vakzine zur Verfügung zu stellen, die insbesondere in der Lage ist, die Entstehung neuer Virusvarianten im Verlauf einer Virus-Erkrankung zu verhindern bzw. deutlich zu vermindern und die Möglichkeiten einzuschränken bzw. zu verhindern, daß das Virus der Immunabwehr
30 des Menschen entkommen und immer wieder neue Abwehrzellen des Immunsystems infizieren kann.

Zur Lösung der Aufgabe werden erfindungsgemäß die in den nachfolgenden Ansprüchen zum Ausdruck kommenden Gegenstände
35 vorgeschlagen.

- 4 -

Gegenstand der vorliegenden Erfindung ist somit unter anderem eine Protein-Vakzine, die eine Mischung viraler Protein-Moleküle umfaßt, wobei die Moleküle Sequenzvarianten eines einzigen viralen Proteins oder eines Teils desselben sind.

5

Unter Sequenzvarianten eines Proteins werden erfindungsgemäß solche Moleküle verstanden, die eine gegenüber einem nativen viralen Protein oder einem Teil (Fragment) desselben abgeleitete Aminosäuresequenzen aufweisen, wobei sich die Varianten dadurch voneinander unterscheiden, daß mindestens eine Aminosäure an beliebigen Stellen der Sequenz bzw. Teilen derselben ausgetauscht sein kann. Vorzugsweise weisen die Sequenzvarianten mehrere Aminosäureaustausche an verschiedenen Stellen der Sequenz auf, die für die Produktion, aber auch die Bindung virusneutralisierender Antikörper verantwortlich sind. Die Zahl und Lage der ausgetauschten Aminosäuren hängt dabei von der Aminosäurevariabilität der Regionen des gp120 ab, die bei Zellkultur adaptierten und Wildtyp HIV-Isolaten beobachtet wird. Erfindungsgemäß weisen die Sequenzvarianten eine Heterogenität an mindestens zwei Aminosäurepositionen der Sequenz oder Teilen derselben auf. Besonders bevorzugt ist eine Heterogenität an drei bis acht und vorzugsweise an mehr als acht Aminosäurepositionen, wobei alle vorkommenden Aminosäuren an diesen Positionen möglich sind. Aus der Kombination der verschiedenen Aminosäuren, die je Position möglich sind, ergibt sich die Anzahl der möglichen Sequenzvarianten, die in ihrer Gesamtheit die Vakzine darstellen.

Insbesondere betrifft die Erfindung eine Vakzine, die eine Mischung von $\geq 10^2$ Sequenzvarianten umfaßt, d.h. eine Mischung von mehr als 10^2 Molekülen unterschiedlicher Aminosäuresequenz (Homologen). Besonders bevorzugt ist eine Vakzine, die $\geq 10^3$ und vorzugsweise $\geq 10^4$ Sequenzvarianten umfaßt.

Im Rahmen der vorliegenden Erfindung wird davon ausgegangen, daß die Vakzine neben den Sequenzvarianten auch das Protein als

- 5 -

solches umfassen kann, von dem die Sequenzvarianten abgeleitet sind.

Erfindungsgemäß wird somit ein Wirkstoff gegen Viren bereitgestellt, der virusspezifische Proteine umfaßt, die sich alle in ihren Sequenzen bzw. Teilen derselben unterscheiden. Um dies zu erreichen, wird das Protein-kodierende Gen neu synthetisiert, um neue Schnittstellen für DNA-schneidende Enzyme zu erzeugen, um den Austausch spezifischer Regionen zu erlauben. Durch chemische
10 Synthese von DNA-Fragmenten, die für den betreffenden Proteinabschnitt kodieren, wird dann eine Genbank für die Protein-kodierende Sequenz hergestellt. Die Expressionsvektoren, die das Protein kodieren, werden dann als Mischung in Zellen transfiziert. Aus diesen Protein-produzierenden Zellen kann dann die
15 Mischung der verschiedenen viralen Proteine isoliert werden, die erfindungsgemäß den bevorzugten Wirkstoff darstellt.

Im Falle von HIV als viraler Erkrankung kommt es durch die chronische Infektion und die damit verbundene kontinuierliche
20 Schädigung des Immunsystems im Verlauf der Erkrankung zum vollständigen Verlust der spezifischen Virusabwehr. Zur spezifischen Virusabwehr gehören neutralisierende Antikörper, die die Eigenschaft besitzen, mit bestimmten antigenen Strukturen eine spezifische, sehr feste Bindung einzugehen. Dadurch werden fremde
25 Antigene markiert und deren Interaktion mit z.B. Virus-Rezeptoren blockiert. Eine solche spezifische Virusabwehr sind neutralisierende Antikörper gegen den V3-Loop des externen GP120-Hüllproteins. Solche anti-V3-Loop-Antikörper sind in der Lage, die Infektion von Zellen zu verhindern. Im Tiermodell ist gezeigt
30 worden, daß sich durch die Verabreichung eines bestimmten monoklonalen V3-Loop-Antikörpers eine Infektion mit HIV-1 verhindern lässt. Durch die Gabe des gleichen monoklonalen Antikörpers war es ebenfalls möglich, eine bestehende Infektion zu heilen.

35 Im Unterschied zu experimentellen Systemen beobachtet man jedoch im natürlichen Verlauf der HIV-Infektion die Entwicklung immer

- 6 -

neuer Virusvarianten. So finden sich zu einem Zeitpunkt in einem einzigen Patienten hunderte von verschiedenen V3-Loop-Varianten, da die Variation des HIV-1 gerade im V3-Loop besonders hoch ist. Der V3-Loop ist eine wichtige dominante antigene Domäne des GP120. Daher wird gegen jeden V3-Loop eine sehr spezifische humorale Immunantwort gebildet. Resultat der Induktion einer hoch spezifischen Immunantwort gegen den V3-Loop ist, daß neutralisierende Antikörper gegen den V3-Loop der HIV-1-Variante A nicht in der Lage sind, die Variante B zu neutralisieren und umgekehrt (Schreiber et al., J. Virol. 68 (1994) 3908-16, Abrahamsson et al., 4 (1990) 107-12).

In der asymptomatischen Phase der HIV-Erkrankung werden die zellfreien Virusvarianten im Serum durch Antikörper neutralisiert. Erst zu einem späteren Zeitpunkt können Virusvarianten im Serum beobachtet werden, die sich der autologen Neutralisation entziehen. Diese V3-Loop-Escape-Varianten werden nicht mehr durch V3-Loop-Antikörper erkannt und daher auch nicht mehr neutralisiert. Alle anderen Virusvarianten, die im Serum der Patienten gefunden werden, werden aber durch autologe V3-Loop-Antikörper erkannt. Dies deutet darauf hin, daß im Verlauf der Erkrankung Virusvarianten auftreten, gegen die keine V3-Loop-Antikörper existieren. Überprüft man den Antikörpergehalt von Serumproben eines Patienten über einen Zeitraum von mehreren Jahren, so kann man in den Serumproben, die am Anfang der asymptomatischen Phase entnommen wurden, die V3-Loop-Antikörper gegen die im späteren Stadium auftretende V3-Loop-Escape-Variante nachweisen. Es kommt daher nicht wie bei anderen Infektionserkrankungen zum klassischen Escape des Erregers durch immer neue antigene Variation, sondern zum Abschalten der neutralisierenden Immunantwort gegen eines der im Patienten replizierenden Viren. Das beobachtete Fehlen der neutralisierenden V3-Loop-Antikörper ist daher das Resultat eines kontinuierlichen Verlusts der typenspezifischen V3-Loop-Antikörper. Durch den Verlust der neutralisierenden Antikörper kann sich das entsprechende Virus vermehren, was zum Anstieg der Viruslast im Serum der Patienten führt. Im Verlauf

- 7 -

der Erkrankung beobachtet man ebenfalls den Anstieg der Viruslast im Lymphknoten (Chun et al., Nature 387 (1997) 183-188).

Ein Virus, welches sich im Patienten gegen das Immunsystem durchsetzt und sich zur dominierenden Virusvariante entwickelt, muß zwangsläufig alle antiviralen Barrieren überwinden, die das Immunsystem bereitstellt. Neben neutralisierenden Antikörpern sind cytotoxische T-Zellen ebenfalls in der Lage die Virusvermehrung zu unterdrücken. Cytotoxische T-Zellen bewirken den Tod HIV infizierter CD4+ T-Zellen. Neben dem Verlust der neutralisierenden Antikörper wird auch der Verlust solcher cytotoxischer T-Zellen gegen HIV infizierte Zellen beobachtet (Zerhoui et al., Thymus 24 (1997) 203-219; Gould et al., Semin Cell Dev Biol 9 (1998) 321-328; Wagner et al., J Gen Virol 74 (1993) 1261-1269; O'Toole et al., AIDS Res Hum Retroviruses 8 (1992) 1361-1368; Shearer et al., 137 (1986) 2514-2521). Daher ist die Aufgabe einer heterologen Vakzine, der Mischung von vielen verschiedenen GP120 Hüllproteinen des HIV, sowohl neutralisierende Antikörper als auch cytotoxische T-Zellen gegen viele verschiedenen Virusvarianten zu induzieren bzw. zu aktivieren.

Wie schon ausgeführt, zeichnet sich die HIV-Erkrankung dadurch aus, daß sich das Virus ständig verändert. Dies wird durch die Fehlerrate des viralen Enzyms Reverse-Transkriptase bewirkt, welches bei der Vermehrung der viralen Erbinformation Fehler macht. Dadurch werden Mutanten erzeugt, die einerseits aufgrund ihrer biologischen Eigenschaften und der antiviralen Wirkung des humanen Immunsystems selektiert werden. Wie bei HIV so gibt es auch weitere Krankheitserreger, die ihre Erbinformation ohne Fehlerkorrektur umschreiben, was zur Ausbildung von vielen Varianten führen kann. Dazu zählen neben Hepatitis A-, B- und C-Viren auch Hanta- und Lassa-Viren. So beobachtet man bei Hepatitis B geimpften Personen, daß es bei ca 2% zum Impfversagen kommt. Bei diesen 2% werden Hepatitis B "escape" Varianten gefunden, die sich nicht durch die durch den Impfstoff induzierte Immunantwort unterdrücken lassen.

Das Entstehen von Impfstoff-, Therapie- und Immun-Escape-Varianten beruht auf der genetischen Variabilität dieser Viren (Blum, Int J Clin Lab Res 27 (1997) 213-214; Jongerius et al., Transfusion 38 (1998) 56-59).

5

Im Stand der Technik ist bislang nicht bekannt, wie einem solchen Verlust Virus-neutralisierender Antikörper entgegengewirkt werden kann. Die bisher eingeschlagenen Strategien waren jeweils nur punktuell erfolgreich, d.h. in Bezug auf die Bildung und den Erhalt einzelner variantenspezifischer Antikörper gegen bestimmte Viren, dem Verlust typenspezifischer Antikörper gegen die Vielfalt der vom Virus im Verlauf der Erkrankung gebildeten Proteinvarianten können die bislang vorgeschlagenen Behandlungsmethoden in Form von antiviralen Medikamenten oder Impfstoffen auf der Basis einer Memo-Substanz jedoch nicht entgegenwirken.

10
15

Durch die vorliegende Erfindung einer heterogenen Vakzine ist es nunmehr erstmals möglich, den Verlust neutralisierender V3-Antikörper zu verhindern oder diesem Verlust zumindest deutlich entgegenzuwirken.

20

Auf der Basis der vorliegenden Erfindung wird erstmals eine immunrekonstitutive Behandlung von Virus-infizierten Patienten, insbesondere von HIV-1-Infizierten, zur Verfügung gestellt, bei der man das Immunsystem in einer Art und Weise aktiviert, daß die natürlich erworbene Immunabwehr gegen die Virus-Population der verschiedenen Virusvarianten regeneriert und neu stimuliert wird, um so die im Patienten vorhandenen Virusvarianten an ihrer Vermehrung zu hindern.

25
30

Die Erfindung basiert kurz gesagt auf dem Konzept, für einen Wirkstoff gegen HIV-1 ein Gemisch von GP120-Proteinen herzustellen, wobei sich diese GP120-Proteine z.B. entweder in der V2- oder der V3-Schleife oder gleichzeitig in den beiden variablen Domänen, d.h. dem V2-Loop und dem V3-Loop, unterscheiden. Um dies zu erreichen, muß das GP120-kodierende Gen neu synthetisiert

35

werden unter Bildung neuer (monovalenter) Schnittstellen für DNA-schneidende Enzyme, die den spezifischen Austausch der z.B. V2- und V3-Region erlauben. Durch chemische Synthese von DNA-Fragmenten, die für den V2-Loop und den V3-Loop kodieren, wird
5 dann eine V2/V3-Genbank für GP120 hergestellt. Die GP120-Expressionsvektoren werden schließlich als Mischung in Zellen transfiziert, und aus diesen GP120-produzierenden Zellen kann dann die Mischung der verschiedenen GP120-Proteine isoliert werden, die als Wirkstoff zur Prophylaxe und/oder Therapie der HIV-Erkrankung
10 bzw. AIDS eingesetzt werden kann.

Die der vorliegenden Erfindung zugrundeliegende Überlegung besteht somit darin, für eine Immuntherapie oder einen Impfstoff gegen HIV-1 viele verschiedene GP120-Moleküle herzustellen, die
15 V3-Loop-Sequenzen (und/oder V2-Loop-Sequenzen) tragen, wie sie auch bei Virusvarianten in Patienten identifiziert werden können.

Die Produktion einer Mischung von natürlichen Virusvarianten im Zellkultursystem ist sehr aufwendig. Beonders zu beachten ist,
20 daß sich die Zusammensetzung einer Virus-Mischung verändert, da man Zellen in der Zellkultur einsetzen muß, die von HIV-1 infiziert werden können. Einige Virusvarianten, die selektive Vorteile besitzen, z.B. eine schnellere Wachstumskinetik, werden daher im Zellkultursystem schon nach wenigen Tagen dominieren und
25 die anderen langsameren Virusvarianten verdrängen. Dies ist vor allen Dingen dann gegeben, wenn die verschiedenen Virusvarianten, die in den peripheren mononukleären Blutzellen (PBMC) oder im Serum von Patienten vorkommen, isoliert werden sollen. Da es nicht möglich ist, eine gleichbleibende Mischung von HIV-Varianten zu kultivieren, kann auch auf diese Weise keine entsprechende
30 Mischung von GP120-Proteinen aus Viruskulturen produziert bzw. isoliert werden.

Daher wurde im Rahmen der vorliegenden Erfindung ein gentechnischer Ansatz für die Herstellung von HIV-1-Varianten und GP120-Varianten gewählt. Zur Herstellung von Virusvarianten und rekomb-

- 10 -

binantem GP120 wurde ein Klonierungssystem konstruiert, das aus zwei Vektoren besteht. Zentraler Bestandteil beider Vektoren ist ein neues, chemisch synthetisiertes HIV-1 Hüllprotein-Gen (HIV-1 *env*-Gen), welches für das GP120-Protein kodiert. Ein Vektor
5 enthält das gesamte Genom des HIV-1. Nach Transfektion dieses Vektors in Zellen produzieren die Zellen infektiöse Viruspartikel. Dies ermöglicht die Produktion einer Mischung von Virusvarianten, da Zellen verwendet werden können, die resistent gegen eine HIV-1-Infektion sind. Daher kann sich in einem solchen
10 Zellkultursystem die Zusammensetzung der Virusmischung nicht aufgrund der biologischen Eigenschaften der Virusvarianten ändern. Der zweite Vektor ermöglicht die Expression von GP120-Varianten in eukaryontischen Zellen. Dieser Vektor dient direkt zur Herstellung des Impfstoffs (der Vakzine).

15 Für die gentechnische Herstellung und Expression der V3-Loop-GP120-Varianten in eukaryotischen Zellen wurde ein spezielles Genkonstrukt für HIV-1-*env* hergestellt, welches im nachfolgenden als Genkassette bezeichnet wird. Zur Herstellung der Genkassette
20 wurde erfindungsgemäß die gesamte kodierende Sequenz des *env*-Gens chemisch synthetisiert. Mit Hilfe dieser Methode ist es möglich, die kodierende Sequenz des Gens beliebig zu verändern, wobei neue DNA-Erkennungssequenzen für DNA-schneidende Restriktionsenzyme in die *env*-Gensequenz eingebaut werden können. Dabei muß darauf
25 geachtet werden, daß sich die Aminosäuresequenz des ursprünglichen GP120 (vorzugsweise der Stämme NL4-3 und PI-932) nicht ändert, während in der DNA-Sequenz des *env*-Gens hingegen neue Restriktionsschnittstellen entstehen. Gleichzeitig werden erfindungsgemäß Schnittstellen für Enzyme, die in der *env*-Sequenz
30 mehrfach vorkommen, entfernt, so daß jeweils nur eine Schnittstelle für ein bestimmtes Restriktionsenzym, wie z.B. BglII, vorhanden ist. Gemäß einer bevorzugten Ausführungsform der Erfindung werden im *env*-Gen zehn neue, einmal-vorkommende (monovalente) Schnittstellen in Abständen von ca. 150 Basenpaaren
35 eingefügt. Das so erzeugte neue *env*-Gen wird im Rahmen der

- 11 -

vorliegenden Erfindung auch als Genkassette bezeichnet. Im Prinzip ist das neue *env*-Fragment einem Polylinker ähnlich.

Die Schnittstellen für die Restriktionsenzyme BstEII und BamHI begrenzen die erfindungsgemäß bevorzugte Genkassette (SEQ ID NO: 9). Um den Bereich des *env*-Gens austauschen zu können, sind beide Schnittstellen nur einmal im pBSCenvATG Expressionsvektor (SEQ ID NO: 10) für gp120 und im retroviralen Vektor pNL4-3 vorhanden. Der Vektor mit dem vom Hinterleger zugeteilten Bezugszeichen pBSCenv-V3 wurde am 6. Januar 1999 bei der DSMZ - Deutschen Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, 38124 Braunschweig, Deutschland, unter der Zugriffsnummer DSM 12612 nach dem Budapester Vertrag hinterlegt.

Durch den Austausch des durch BstEII und BamHI begrenzten Genabschnitts lassen sich die neu hergestellten Abschnitte des *env*-Gens, die ebenfalls durch BstEII und BamHI begrenzt sind, in beide Vektoren klonieren. Das PstI/BclI V2-Loop und das BglII/XbaI V3-Loop *env*-Fragment werden chemisch bzw. enzymatisch synthetisiert und anschließend in einen Standardvektor wie z.B. pUC 18 oder 19 kloniert. In diesem Vektor wird das Fragment durch Sequenzierung überprüft und anschließend durch einen Pst/BclI oder BglII/XbaI Verdau aus dem Vektor geschnitten und in den pBSCenvATG und dann in den pNL4-3 Vektor überführt. Anstelle der Zwischenklonierung in einen Standardvektor (pUC 18, pUC 19 etc.) können die PstI/BclI V2-Loop- und BglII/XbaI V3-Loop-Fragmente auch direkt in pBSCenvATG kloniert werden.

Da sich in der Genkassette alle ca. 100 Basenpaare eine Schnittstelle für ein Restriktionsenzym befindet, können alle Bereiche des *env*-Gens, insbesondere der V3-Loop oder V2-Loop, gegen beliebige DNA-Fragmente ausgetauscht werden. Der für den V3-Loop kodierende Abschnitt läßt sich durch ein BglII-XbaI-Fragment, das eine Größe von 244 Basenpaaren besitzt (vgl. SEQ ID NO: 9; Nukleotide 708 bis 955), austauschen. Ziel ist es, die Genabschnitte, die für die variablen Loops des GP120-Proteins

- 12 -

5 kodieren, gegen neue DNA-Fragmente auszutauschen. Die Fragmente werden synthetisch hergestellt und anschliessend in die *env*-Genkassette kloniert. Wenn man bei der Synthese der V2-Loop- und V3-Loop-DNA-Fragmente an bestimmten Positionen die Sequenz variiert, d.h. alle vier Nukleotide gleichberechtigt einbaut oder an dieser Position das Nukleotid Inosin verwendet, lassen sich *env*-Genvarianten mit einer vorgegebenen Variation der dadurch kodierten Aminosäuresequenz herstellen. Die V3-Loop-Sequenzen von Patientenisolaten sind Ausgangsbasis für die Herstellung der GP120-Mischung. Werden diese Varianten in die *gp120*-Genkassette kloniert, so können die entsprechenden GP120-Proteine mit den variablen antigenen Domänen in eukaryontischen Zellen exprimiert werden. Für die Variation der V3-Loop Region des *env*-Gens stehen Sequenzdaten von HIV-1 Patientenisolaten zur Verfügung. Patienten-
15 isolate sind Virusvarianten, die direkt aus Patientenmaterial, Serum oder infizierten Zellen aus dem Blut im Labor angezüchtet worden sind. Diese Viren zeichnen sich, im Unterschied zu HIV-1 Viren, die lange im Labor gezüchtet wurden, durch besondere Eigenschaften aus. Diese Viren sind resistenter gegen neutralisierende Antikörper und Chemokine. Im Unterschied zu Zellkultur-
20 adaptierten Viren benutzen Patientenisolate verschiedene Corezeptoren für die Infektion von Zellen. Da sich Patientenisolate sehr von Labor-adaptierten Viren unterscheiden, sollen die gesammelten V2-Loop- bzw. V3-Loop-Sequenzdaten der *env*-Gene solcher Viren für die Herstellung der GP120-Mischung benutzt werden. Im Rahmen der Erfindung sind zusätzliche Basenaustausche in Bereichen außerhalb der für V2 und V3 kodierenden Sequenzabschnitte möglich.

30 Gemäß einer besonderen Ausführungsform der Erfindung wird eine Protein-Vakzine zur Verfügung gestellt, die eine Mischung von GP120-Proteinen des Humanen Immundefizienzvirus (HIV) umfaßt, die sich jeweils in ihrer Aminosäuresequenz im Bereich der V3-Scheife und/oder der V2-Schleife voneinander unterscheiden.

35 Im Rahmen der vorliegenden Erfindung werden bei der gentechnischen Herstellung der Protein-Mischung bzw. der Protein-Vakzine

- 13 -

Sequenzvariationen vorzugsweise in Bereichen des V3-Loops eingeführt, die außerhalb von Konsensussequenzen liegen (vgl. z.B. M. Schreiber et al., J. Virol. 71 Nr. 12 (1997) 9198-9205). Bei den Konsensussequenzen handelt es sich um Sequenzabschnitte, die sowohl zwischen verschiedenen Virusstämmen als auch bei der Bildung von Virusvarianten bei fortschreitender Virus-Erkrankung (HIV-Erkrankung) im Körper im wesentlichen erhalten bleiben. Ein solcher Abschnitt ist im Falle des HIV-1 vom B Subtyp die Sequenz Gly-Pro-Gly-Arg-Ala-Phe (GPGRAPH). Links und rechts von diesem Sequenzmotiv liegen kurze 4-10 Aminosäure lange Bereiche, die stark variieren können. In Fig. 1 ist dies am Beispiel des V3-Loop Sequenzvariationen von Patientenisolaten dargestellt.

Zusätzlich zu Variationen im Bereich des V3-Loops können die Moleküle der erfindungsgemäßen Protein-Vakzine auch im Bereich des V2-Loops Sequenzvariationen aufweisen. Auch in diesem Fall sind Änderungen in der Aminosäuresequenz außerhalb von Konsensussequenz-Bereichen bevorzugt (siehe Fig. 2). Ferner sind zusätzliche Aminosäureaustausche in Bereichen außerhalb der V2- und V3-Loops möglich.

In Zusammenhang mit der vorliegenden Erfindung werden stets die üblichen Ein- bzw. Dreibuchstabencodes zur Bezeichnung der Aminosäuren verwendet. Bei der Variation der Aminosäuresequenzen innerhalb der erfindungsgemäßen Protein-Vakzine sind beliebige Aminosäureaustausche möglich. In jedem Fall werden aber Aminosäuren der viralen GP120 Sequenz bzw. der Sequenzen des V2-Loop und V3-Loop vorzugsweise durch Aminosäuren ausgetauscht, die auch bei anderen Virusvarianten an den entsprechenden Sequenzpositionen vorkommen (siehe Fig. 2, Stand der Sequenzdaten 1997 in: Human Retroviruses and AIDS, A compilation and analysis of nucleic acid and amino acid sequences, Los Alamos National Laboratory, Los Alamos, NM 87545, U.S.A., Editors: B. Korber, B. Hahn, B. Foley, J.W. Mellors, T. Leitner, G. Myers, F. McCutchan, C. Kuiken).

- 14 -

Im Rahmen der vorliegenden Erfindung wird erstmals eine Protein-Vakzine zur Verfügung gestellt, die auf gentechnischem Wege hergestellt wird und Sequenzvarianten spezifischer viraler Antigene (Proteine) umfaßt. Die Herstellung der Vakzine wird nachfolgend am Beispiel des V3-Loops von HIV beschrieben. Für den Fachmann versteht es sich, daß das nachfolgend beschriebene Syntheseprinzip auf andere Viren bzw. virale Proteine oder Teile von deren Sequenzen übertragen werden kann.

10 Kurz gesagt wird als Basis für die Vakzine eine Mischung aus GP120-Proteinen mit variablen V3-Sequenzen hergestellt, die der Variabilität des V3-Loops des GP120-Proteins von HIV nachempfunden sind.

15 Dafür wird zunächst die für das GP120-Protein kodierende Sequenz stückweise in das Plasmid pUC 18/19 hineinkloniert. Dies erfolgt durch einen Polylinkeraustausch, so daß mit den möglichen Restriktionssites, die vorher in der *gp120*-Sequenz durch stille Mutationen eingefügt wurden, alle interessanten Abschnitte mit
20 monovalenten Restriktionssites flankiert sind und somit später herausgeschnitten und ausgetauscht werden können.

Anschließend werden zwei homologe Nukleinsäureoligomere mit Hilfe eines DNA-Synthesizers hergestellt (Länge ca. 300 Basenpaare).

25 Nach erfolgter Hybridisierung bildet sich das doppelsträngige DNA-Fragment mit der DNA-Sequenz des V3-Loops. Für die Herstellung des doppelsträngigen V3-Loop DNA-Fragments können verschiedene Standardmethoden aus der Molekularbiologie verwendet werden. Bei einer Methode werden die Variablen durch Inosine
30 eingeführt und bei dem entsprechenden komplementären Strang werden die Variablen durch einen Nukleotid-Mix (AGCT, AGC, AG, ...) eingeführt. Bei dieser Methode wird das DNA-Fragment ausschließlich chemisch hergestellt. Bei einer anderen Methode, einer Kombination aus chemischer und enzymatischer DNA-Synthese,
35 werden die Variablen durch Nukleotid-Mischungen eingeführt. Die Hybridisierung der Oligonukleotide erfolgt dann an komplementären

- 15 -

Enden einer Länge von 20-30 Basen, die nicht variabel sind. Die Synthese zum vollständig doppelsträngigen Molekül erfolgt enzymatisch mit Hilfe einer DNA-Polymerase. Dabei können sowohl isotherme DNA-Polymerasen (Klenow-Fragment) oder thermostabile DNA-Polymerasen verwendet werden (Taq-Polymerase). Verwendet man Taq-Polymerase, dann können größere Mengen für die Klonierung des V3-Loop DNA-Fragments auch mit Hilfe der Polymerase-Ketten-Reaktion hergestellt werden. Mit Hilfe dieser Methoden wird ein doppelsträngiges DNA-Fragment hergestellt, welches an bestimmten Positionen variabel ist. Diese degenerierte DNA-Sequenz kodiert für die entsprechende Vielzahl von V3-Loop Aminosäure-Sequenzen der herzustellenden GP120-Proteinmischung, der Protein-Vakzine.

Die Mischung der synthetisierten V2-Fragmente besitzt am 5'-Ende eine PstI- und am 3'-Ende eine BclI-Schnittstelle. Die Mischung der synthetisierten V3-Fragmente besitzt am 5'-Ende eine BglII- und am 3'-Ende eine XbaI-Schnittstelle. Mit Hilfe der jeweiligen zwei Schnittstellen wird die Fragment-Mischung in einen Vektor, z.B. pUC18 delta-env oder pUC 18 BstEII-BamHI (Fig. 3), kloniert. Es entsteht nach dieser Klonierung eine Mischung oder ein Pool von Plasmid-DNAs des Vektors pUC 18, die alle das vollständige BstEII-BamHI env Fragment besitzen. Alle Plasmid-DNAs unterscheiden sich ausschließlich in der Sequenz des env V2-Loop oder V3-Loop. Durch Insertion der V2 bzw. der V3 Fragmente wird die Deletion (delta env) des env-Genabschnitts im pUC18 Vektor aufgehoben.

Diese so hergestellte Mischung der Plasmid-DNAs mit variablen env-Fragmenten wird in *E. coli* transformiert und fermentiert. Dabei amplifiziert und repliziert *E. coli* diese Mischung an Plasmid-DNA. Dabei werden alle möglichen Variablen dieser V2-/V3-Loop-Plasmide erzeugt.

Anschließend wird dieser Plasmid Pool isoliert. Aus der Mischung der Plasmid-DNAs wird dann das env-Fragment durch BstEII- und BamHI-Verdau ausgeschnitten und direkt in den Vektor für die

- 16 -

Expression des viralen *gp120* kloniert. Dieser Vektor hat den Vorteil, daß das GP120 auch im eukaryotischen Zellsystem exprimiert wird.

- 5 Dieser Pool an BSCenvATG-Vektor-DNA mit variablem *gp120*-Konstrukt wird nun in Cos- bzw. Chinese Hamster Ovary-Zellen (CHO-Zellen) transfiziert. Diese Eukaryonten exprimieren dann diesen Pool an Plasmiden, so daß statistisch das entsprechende Protein zu jeder Variablen translatiert und anschließend - je nachdem, welche
- 10 Eukaryonten eingesetzt werden - entsprechend glykosyliert wird. Es schließt sich die Ernte der Proteine mit anschließender Aufreinigung bis zum Fertigprodukt (Protein-Mischung bzw. GP120-Mischung mit variabler Aminosäuresequenz) an.
- 15 Zur Veranschaulichung wird das Verfahren zur Erzeugung der erfindungsgemäßen Protein-Vakzine in Fig. 4 übersichtlich dargestellt.

Wie bereits dargelegt, können Variationen im V2-Loop und/oder im

20 V3-Loop vorgenommen werden. Die Variation des V2-Loops kann dabei nach demselben, für V3 beschriebenen Schema erfolgen.

Die Herstellung der erfindungsgemäßen Protein-Vakzine erfolgt, wie oben beschrieben, über verschiedene Schlüssel-Konstrukte,

25 d.h. Nukleinsäure-Zwischenstufen und DNA-Konstrukte, die zur Ausführung der Erfindung wesentlich sind.

Wie oben bereits ausgeführt, wird die das GP120-Protein kodierende Nukleinsäuresequenz stückweise in pUC 18/19 hineinkloniert,

30 wobei erfindungsgemäß von einer *gp120*-Sequenz ausgegangen wird, bei der die mehrfach vorhandenen Restriktionsspaltorte zunächst durch stille Mutationen so modifiziert werden, daß von den verbleibenden Restriktionsspaltorten jeder nur noch einmal in der Sequenz vorkommt. Diese Sequenzmodifikation ist notwendig, um die

35 Expressionskassette, die zur Erzeugung der Sequenzvarianten erforderlich ist, gezielt an einer ganz bestimmten Stelle, die

- 17 -

durch zwei, nur je einmal in der Sequenz vorkommende Restriktionsspaltorte begrenzt ist, einfügen zu können. Dem Fachmann sind Verfahren zur Einführung stiller Mutationen in einer Nukleinsäuresequenz wohlbekannt.

5

Gegenstand der vorliegenden Erfindung ist daher auch eine Nukleinsäuresequenz, die von der *env*-Sequenz in SEQ ID NO: 1 oder einem Fragment derselben abgeleitet ist, wobei sie derart modifiziert ist, daß sie ausschließlich monovalente Restriktionsspaltorte enthält. Vorzugsweise erfolgt die Modifikation durch Einführen stiller Mutationen. Gemäß einer bevorzugten Ausführungsform der Erfindung weist die Nukleinsäuresequenz die in SEQ ID NO: 9 dargestellte Sequenz auf. Die Sequenz der Genkassette kann derart modifiziert sein, daß sie das gesamte *env*-Gen oder Teile des *env*-Gens von Virusisolaten aus Patienten enthält. Vorzugsweise soll eine solche Genkassette die *env*-Sequenz des Patientenisolats PI-932 (SEQ ID NO: 11) enthalten. Eine erfindungsgemäß bevorzugte Genkassette basierend auf der Sequenz PI-932 ist in SEQ ID NO: 12 dargestellt.

20

Gegenstand der vorliegenden Erfindung ist daher auch eine einzelsträngige Nukleinsäuresequenz, die den für den V3-Loop und/oder den für den V2-Loop kodierenden Bereich oder Fragmente bzw. Teile derselben enthält, wobei im Falle des V3-Loop ein BglII-XbaI (247 bp) oder auch ein BglII-NheI (283 bp) Fragment gegen ein verändertes Fragment welches an mindestens 6, vorzugsweise an 9 bis 20 Positionen, Nukleinsäureaustausche oder Mutationen trägt, und im Falle des V2-Loop ein PstI-BclI (139 bp) oder auch ein PstI-EcoRI (339 bp) Fragment gegen ein verändertes Fragment welches an mindestens 6, vorzugsweise an 9 bis 20 Positionen, Nukleinsäureaustausche oder Mutationen trägt, ausgetauscht sein kann. Dabei sind die Nukleotide innerhalb einer Nukleinsäuresequenz entweder jeweils durch Inosin oder jeweils durch eine Mischung von 2-4 Nukleotiden ersetzt. Am Beispiel in Fig. 5 ist dies erläutert. Sollen an 7 Aminosäurepositionen des V3-Loop 21 verschiedene Aminosäuren zu 1152 Varianten kombiniert werden,

- 18 -

dann müssen an 11 Nukleinsäurepositionen jeweils zwei Nukleotide durch chemische Synthese in die Sequenz der einzelsträngigen Nukleinsäuren (Oligonukleotide) eingeführt werden (Fig. 5).

- 5 Die einzelsträngigen Nukleinsäuresequenzen werden - wie oben bereits erwähnt - zum Doppelstrang hybridisiert bzw. synthetisiert.

Die vorliegende Erfindung betrifft daher ferner doppelsträngige
10 DNA, die Hybride der o.g. einzelsträngigen Nukleinsäuresequenzen umfaßt, wobei jeweils eine Nukleinsäuresequenz (5'-3'-Oligomer oder 3'-5'-Oligomer) an ein oder mehreren ausgewählten Positionen Inosine enthält und die andere Nukleinsäuresequenz (3'-5'-
15 Oligomer oder 5'-3'-Oligomer) an den bei der späteren Hybridisierung entsprechenden komplementären Positionen jeweils zwei, drei oder vier der möglichen Nukleotide (Adenin, A; Thymin, T; Guanin, G; Cytosin, C) enthält. Dadurch werden Sequenzen erzeugt, bei denen die vier Basen zufallsverteilt an den jeweiligen Positionen der einzelsträngigen Nukleinsäuresequenz (DNA) vorhanden
20 sind, so daß die gewünschten Kombinationen von A, T, G oder C für die entsprechenden gewünschten Aminosäure Codone (ein Codon wird aus einer Abfolge von drei Nukleotiden gebildet) erzeugt werden. Aufgrund der Variation von Nukleotidpositionen kommt es zu zufallsverteilten Sequenzkombinationen. Daher entstehen z.B. für
25 die Kombination aus (A,C) (ACT) (ACGT) insgesamt $2 \times 3 \times 4 = 24$ mögliche DNA-Sequenzen, die für verschiedene Aminosäuren kodieren. Die Anzahl der variablen Positionen bestimmt auch gleichzeitig die Anzahl und Position der eingeführten Inosine in der komplementären DNA-Sequenz. Die Berechnung der Heterogenität
30 eines solchen Konstruktes ist in Fig. 5 dargestellt.

Da zur Bildung der doppelsträngigen DNA jeweils Inosine enthaltende einzelsträngige DNA mit Oligomeren hybridisiert wird, das an den entsprechenden Positionen jeweils zufallverteilt A, T, G
35 oder C enthält, wird eine Mischung aus doppelsträngiger *gp120*-Sequenz oder eines Teils (Fragments) derselben erhalten, wobei

die Nukleinsäuresequenzen jeweils von der *env*-Sequenz (SEQ ID NO: 1 oder 12) abgeleitet sind und wobei sich die Nukleinsäuresequenzen in dem für die V2-Schleife kodierenden Bereich und/oder in dem für die V3-Schleife kodierenden Bereich jeweils voneinander unterscheiden. Das heißt, daß sich die in der Mischung enthaltenen Nukleinsäuresequenzen derart voneinander unterscheiden, daß sie für eine Mischung von Proteinen kodieren, die in der V2-Schleife und/oder in der V3-Schleife jeweils voneinander verschiedene Aminosäuresequenzen aufweisen.

10

Gemäß einer bevorzugten Ausführungsform der Erfindung können auf diese Weise mindestens 10^2 , vorzugsweise mindestens 10^3 und gemäß einer besonders bevorzugten Ausführungsform der Erfindung mindestens 10^4 Sequenzvarianten auf Nukleinsäure-Ebene (DNA-Ebene) erhalten werden. Diese DNA-Mischung wird nachfolgend als DNA-Pool bezeichnet, bezogen auf die Varianten der *gp120*-Sequenz wird dieser nachfolgend als Pool mit variablem *gp120*-Konstrukt (*gp120*-Pool) bezeichnet.

15

Unter Sequenzvarianten eines DNA-Sequenz werden im Rahmen der vorliegenden Erfindung solche Moleküle verstanden, die eine gegenüber einer nativen viralen DNA-Molekül oder einem Teil (Fragment) desselben abgeleitete Nukleinsäuresequenz aufweisen, wobei sich die Varianten dadurch voneinander unterscheiden, daß mindestens ein Nukleotid an beliebigen Stellen der Sequenz ausgetauscht sein kann. Vorzugsweise weisen die Sequenzvarianten mehrere Nukleotidaustausche an verschiedenen Stellen der Sequenz auf, wobei die Zahl und die Lage der ausgetauschten Nukleinsäuren im wesentlichen von der Länge der Nukleinsäuresequenz abhängt.

25

30

Bei der Expression der *env*-Gen Varianten ausgehend von der hergestellten Plasmid-DNA-Mischung ist zu erwarten, daß aufgrund der Zufallsverteilung alle möglichen DNA Sequenzen exprimiert werden. Daher werden dann auch alle aufgrund der vorgegebenen DNA-Mischung möglichen Sequenzvarianten des *gp120*-Moleküls oder eines Teils desselben gebildet. Die Heterogenität des *gp120*-

35

- 20 -

Gemisches berechnet sich einerseits anhand der Heterogenität der DNA-Sequenz und der Degeneration des genetischen Codes für die Aminosäuren (siehe auch Fig. 5).

5 Der Nachweis der Heterogenität der Plasmid-DNA Mischung erfolgt durch die DNA-Sequenzierung von einzelnen Klonen. Dazu wird die Mischung der Plasmid-DNA in *E. coli* transformiert und einzelne Klone werden zufällig ausgewählt und deren V2- bzw. V3-Loop-Sequenz bestimmt. Insgesamt sollen ca. 100-200 verschiedene Klone
10 sequenziert werden. Aus der statistischen Verteilung der DNA-Sequenzen kann auf die Verteilung der gesamten Mischung zurückgerechnet werden. Die Methode entspricht weitestgehend der Methode der Stichproben-Entnahme, wie sie standardmäßig für die Qualitätskontrolle von verschiedensten Produkten eingesetzt wird.

15

Die direkte molekulare Analyse der Heterogenität der GP120-Protein Mischung gestaltet sich etwas schwieriger, da sich einzelne GP120-Moleküle aus der Mischung nicht abtrennen und nachweisen lassen. Dies liegt in der Natur der Sache, weil sich
20 die einzelnen GP120-Varianten nur in wenigen Aminosäuren voneinander unterscheiden. Durch Gelelektrophoretische Auftrennung oder eine Massenspektroskopische Analyse kann festgestellt werden, ob es sich um eine Mischung oder eine einzelne Form des GP120-Moleküls handelt. Wenn die GP120-Mischung z.B. im
25 Tier als Vakzine eingesetzt wird, ist die im Tier induzierte Immunantwort direkt von der Anzahl und der Zusammensetzung der GP120-Mischung abhängig. Die Immunantwort des Tieres, die neutralisierenden Antikörper kann dann in Virus-Neutralisations-Tests untersucht werden. Zur Kontrolle einer entsprechend
30 Varianten-übergreifenden Immunantwort können verschiedene Patienten isolate des HIV-1 im Neutralisationstest eingesetzt werden. Bevorzugt werden Patienten isolate gewählt, die sich in der Fähigkeit unterscheiden, verschiedene Corezeptoren für die Infektion der Zielzellen benutzen zu können. Das Neutralisations-
35 potential der GP120-Mischung dient dann als Qualitätsmaßstab.

Gegenstand der vorliegenden Erfindung ist daher ferner eine Protein-Mischung, die GP120-Proteine umfaßt, die in der V2-Scheife und/oder in der V3-Schleife jeweils voneinander verschiedene Aminosäuresequenzen aufweisen, wobei die Mischung gemäß
5 einer Ausführungsform der Erfindung mindestens 10^2 , vorzugsweise mindestens 10^3 und gemäß einer besonders bevorzugten Ausführungsform der Erfindung mindestens 10^4 Sequenzvarianten enthält.

Wie bereits zuvor erwähnt, wird die Mischung doppelsträngiger
10 DNA, die durch Hybridisierung von Inosine-enthaltender einzelsträngiger DNA mit zufallsverteilt A-, T-, G- und/oder C-enthaltender einzelsträngiger DNA erhalten werden kann, d.h. der Pool mit variablem GP120-Konstrukt, in prokaryontischen oder eukaryontischen Wirtszellen, vorzugsweise *E. coli*, transformiert und fer-
15 mentiert.

Gegenstand der vorliegenden Erfindung sind daher Plasmide, die doppelsträngige DNA insertiert enthalten, die jeweils Hybride der einzelsträngigen, Inosin-enthaltenden Nukleinsäuresequenz (s.o.)
20 mit der einzelsträngigen, eine Mischung aller vier Nukleotidvarianten (A, T, G, C) enthaltende Nukleinsäuresequenz (s.o.) umfassen. Ferner betrifft die Erfindung eine Vektor-Mischung, die eine Mischung dieser Plasmide, wobei sich die Nukleinsäuresequenzen der Plasmide in dem für die V3-Schleife kodierenden Bereich und/oder in dem für die V2-Schleife kodierenden Bereich
25 jeweils voneinander unterscheiden. Gemäß einer Ausführungsform der Erfindung umfaßt die Vektor-Mischung mindestens 10^2 , vorzugsweise mindestens 10^3 und gemäß einer besonders bevorzugten Ausführungsform der Erfindung mindestens 10^4 der genannten Plasmide. Je nachdem, welche Wirtszellen zur Expression der Vektoren (Plasmide) verwendet werden sollen, kommen verschiedene, dem Fachmann wohlbekannte Expressionssysteme und Basisvektoren in Frage (vgl. Methods in Enzymology, Vol. 185, Gene Expression Technology, 1991, Herausgeber D.V. Goeddel, Academic Press,
30 Inc.). So ist das Plasmid pUC 18/19 für die Expression in *E. coli* als Basisplasmid bevorzugt, in das die Nukleinsäuresequenz-

- 22 -

Varianten hineinkloniert werden. Im Rahmen der vorliegenden Erfindung lassen sich die Vektor-Mischungen somit entweder in bakteriellen Wirtszellen, wie *E. coli*, oder in eukaryontischen Wirtszellen, vorzugsweise aus der Gruppe bestehend aus Cos-,
5 CHO-, oder Baby Hamster Kidney-Zellen (BHK-Zellen), oder in anderen Wirtszellen exprimieren.

Die vorliegende Erfindung betrifft somit ferner *E. coli*-Wirtszellen oder eukaryontische Wirtszellen, die mit einer erfindungsgemäßen Vektor-Mischung transfiziert ist.
10

Da Wirtszellen mit dem DNA-Pool (der Vektor-Mischung) transfiziert werden, ist zu erwarten, daß eine ebenso große oder zumindest annähernd ebenso große Anzahl unterschiedlich transformierter Wirtszellen entsteht, wie Sequenzvarianten im DNA-Pool vorhanden sind. Bei der Herstellung der Vakzine ist besonderes Augenmerk auf die Ausbeuten bei den einzelnen Klonierungsschritten zu legen. Die Ausbeuten der Herstellung der doppelsträngigen DNA-Fragmente, der Ligation der V3-Loop und V2-Loop
15 DNA-Fragmente in das *env*-Gen und die Transformation bzw. die Herstellung der Bakterien und Zellklone muß so gestaltet sein, daß die zu erzielende Heterogenität des Gemisches nicht eingeschränkt wird. An einem Beispiel soll dies verdeutlicht werden. Wenn z.B. aus einer Menge von Kugeln, die sich aufgrund zweier
20 Farben unterscheiden, eine Anzahl Kugeln entnommen werden soll, aber beide Farben mit 99,9 % Wahrscheinlichkeit in dieser Auswahl vorhanden sein sollen, müßten ca. 13 Kugeln entnommen werden (G. Schreiber, Ein kombinatorisches Problem aus der Genetik; Bioengineering 1988, 2:32-35). Übertragen auf die Klonierungsschritte der HIV-Vakzine bedeutet dies: Wenn eine Heterogenität von ca. 6000 Virusvarianten erzeugt werden soll, müßten in jedem einzelnen Klonierungsschritt dreizehnmal so viele Klone erzeugt werden. Es muß also eine Genbank von ca. 80000 Klonen für den V3-Loop hergestellt werden. Ausgehend von dieser Genbank werden dann
30 die Expressionsvektoren für die Transfektion der CHO-Zellen hergestellt. Bei der Transfektion der CHO-Zellen müssen dann

demzufolge ca. 80000 Transfektionsergebnisse erzielt werden. Eine solche Mischung von transfizierten CHO-Zellen erzeugt dann die gewünschte Menge von 6000 verschiedenen gp120 Varianten.

- 5 Die vorliegende Erfindung betrifft auch ein Verfahren zur Herstellung einer für ein virales Protein kodierenden Nukleinsäuresequenz (Expressionskassette), bei dem man die Sequenz derart modifiziert oder vorzugsweise so viele stille Mutationen einführt, daß sie anschließend mindestens zwei und vorzugsweise
10 ausschließlich monovalente Restriktionsspaltorte enthält. Besonders bevorzugt enthält die Sequenz nur noch monovalente Restriktionsspaltorte. Vorzugsweise handelt es sich bei dem von der Nukleinsäuresequenz kodierten Protein um GP120, wobei man durch stille Mutationen die für das GP120 kodierende *env*-Wildtyp-
15 Sequenz variiert.

- Ferner ist Gegenstand der vorliegenden Erfindung ein Verfahren zur Herstellung der erfindungsgemäßen Vektor-Mischung, die vorzugsweise eine Mischung von Plasmiden enthält, deren Nukleinsäuresequenzen sich in dem für die V2-Schleife kodierenden Bereich und/oder in dem für die V3-Schleife kodierenden Bereich jeweils voneinander unterscheiden, wobei man die erfindungsgemäßen Plasmide in einen in Wirtszellen (vorzugsweise *E. coli*-,
20 *Cos*-, CHO- oder BHK-Zellen) exprimierbaren (Basis-)Vektor hineinligiert. Bei dem (Basis-)Vektor handelt es sich vorzugsweise um den pUC 18-, den pUC 19- oder den BSCenvATG-Vektor.

- Im Rahmen der vorliegenden Erfindung wird ferner ein Verfahren zur Herstellung/Erzeugung von Wirtszellen, vorzugsweise ausgewählt aus der Gruppe bestehend aus *E. coli*-, *Cos*-, BHK- oder CHO-Zellen, bei dem man die Wirtszellen mit einer erfindungsgemäßen Vektor-Mischung transformiert, die eine Mischung von Plasmiden enthält, deren Nukleinsäuresequenzen sich in dem für die V2-Schleife kodierenden Bereich und/oder in dem für die V3-Schleife
35 kodierenden Bereich jeweils voneinander unterscheiden.

Schließlich wird erfindungsgemäß erstmals ein Verfahren zur Herstellung einer DNA-Vakzine zur Verfügung gestellt, bei dem man das erfindungsgemäße Verfahren zur Herstellung der Vektor-Mischung durchführt, wobei die erfindungsgemäßen Plasmide nach
5 Applikation in Wirtszellen (z.B. menschliche und tierische Monocyten) die Mischung der *gp120*-Proteine exprimieren. Die DNA-Vakzine kann zur Applikation gegebenenfalls mit pharmazeutisch verträglichen Hilfs- und/oder Trägerstoffen formuliert sein und nach Verabreichung im Organismus die Produktion der Sequenzva-
10 rianten viraler Proteine ermöglichen.

Die vorliegende Erfindung betrifft somit auch eine pharmazeuti-
sche Zusammensetzung bzw. eine DNA-Vakzine, die für eine Mischung
strukturell unterschiedlicher Virus-Proteine kodiert, wobei die
15 Vakzine eine Mischung von Sequenzvarianten eines viralen DNA-Moleküls oder eines Teils desselben enthält, d.h. von DNA-Molekülen, deren Nukleinsäuresequenzen sich in dem für das Protein kodierenden Bereich, einem Teil oder einem Fragment desselben voneinander unterscheiden. Unter dem Begriff "struktu-
20 rell unterschiedlichen Virus-Proteinen" werden erfindungsgemäß solche Proteine verstanden, deren Aminosäuresequenzen sich von der Wildtyp-Sequenz des entsprechenden Virus-Proteins ableiten, wobei die Aminosäuresequenzen insofern voneinander abweichen, als sie sich untereinander und im Vergleich zur Wildtyp-Sequenz durch
25 ein oder mehrere ausgetauschte Aminosäuren an gleichen oder unterschiedlichen Positionen der Sequenz unterscheiden. Wie bereits ausgeführt, kodieren die Nukleinsäuresequenzen in der Vakzine erfindungsgemäß bevorzugt für eine Mischung strukturell unterschiedlicher GP120-Proteine von HIV (Sequenzvarianten von
30 GP120), wobei eine Vakzine besonders bevorzugt ist, die eine Mischung von DNA-Molekülen enthält, deren Nukleinsäuresequenzen sich in dem für die V2-Schleife kodierenden Bereich und/oder in dem für die V3-Schleife kodierenden Bereich von *gp120* des HIV-1 voneinander unterscheiden. In diesem Zusammenhang wird auch auf
35 Fig. 3 verwiesen, in der im V3-Loop von GP120 bekannte strukturelle Unterschiede bzw. Variationen dargestellt sind.

Die erfindungsgemäße DNA-Vakzine besitzt besondere Bedeutung im Rahmen einer Gentherapie.

Schließlich wird erfindungsgemäß erstmals ein Verfahren zur Herstellung einer pharmazeutischen Zusammensetzung bzw. einer Protein-Vakzine zur Verfügung gestellt, bei dem man die erfindungsgemäßen Wirtszellen, d.h. Wirtszellen, die mit einer erfindungsgemäßen Vektor-Mischung transformiert sind, wobei die Vektoren jeweils Plasmide enthalten, deren Nukleinsäuresequenzen sich in dem für die V2-Schleife kodierenden Bereich und/oder in dem für die V3-Schleife kodierenden Bereich jeweils voneinander unterscheiden, unter Bedingungen kultiviert, die die Expression der Mischung viraler Protein-Sequenzvarianten gestatten. Bei den Wirtszellen handelt es sich vorzugsweise um bakterielle Wirtszellen, wie *E. coli*, oder um eukaryontische Wirtszellen, vorzugsweise aus der Gruppe bestehend aus Cos-, Chinese Hamster Ovary-(CHO-), oder Baby Hamster Kidney-Zellen (BHK-Zellen).

Durch die vorliegende Erfindung ist es somit möglich, eine Vakzine aus variablen GP120-Proteinen zur Verfügung zu stellen, die bei der prophylaktischen und/oder therapeutischen Behandlung einer HIV-Infektion oder AIDS

1. das Immunsystem aktiviert,
2. die Bildung HIV-1-neutralisierender Antikörper induziert,
3. den Verlust von neutralisierenden Antikörpern verhindert und
4. die Kontrolle über die Stimulation der GP120-spezifischen Immunantwort durch den Wirkstoff übernommen wird.

Dies erfolgt vorzugsweise dadurch, daß bereits vor dem mit dem Ausbruch von AIDS im Zusammenhang stehenden Verlust HIV-neutralisierender Antikörper ausreichend Antigene zur Verfügung stehen, die aufgrund ihrer Vielfalt unterschiedlicher Aminosäuresequenzen (d.h. ihrer Sequenzvariationen) in der Lage sind, die Bildung

derjenigen HIV-Antikörper zu induzieren, die normalerweise, d.h. ohne entsprechende Prophylaxe gemäß der vorliegenden Erfindung, bei Progression der Erkrankung verloren gehen oder in ihrer Konzentration abnehmen. Neben einer vorbeugenden Behandlung

5 betrifft die vorliegende Erfindung auch die therapeutische Behandlung, wobei die GP120-Mischung (Protein-Vakzine) eingesetzt wird, um einer Abnahme bzw. einem Verlust HIV-neutralisierender Antikörper entgegenzuwirken, indem das Immunsystem aktiviert und die Bildung neuer bzw. zusätzlicher HIV-1-neutralisierender

10 Antikörper induziert wird.

Erfindungsgemäß wird daher die Mischung strukturell unterschiedlicher viraler Proteine, die Sequenzvarianten eines viralen Proteins (vorzugsweise von GP120) oder eines Teils desselben sind

15 zur Herstellung einer Vakzine zur Prävention und/oder Therapie einer Virus-Infektion beim Menschen verwendet. Ferner betrifft die vorliegende Erfindung die Verwendung einer Mischung von DNA-Molekülen, die für Sequenzvarianten eines viralen Proteins oder eines Teils desselben kodieren, zur Herstellung einer Vakzine zur

20 Prävention und/oder Therapie einer Virus-Infektion beim Menschen. Gemäß einer bevorzugten Ausführungsform betrifft die Erfindung die Verwendung zur Herstellung einer Vakzine zur Prävention und/oder Therapie einer HIV-Infektion beim Menschen.

25 Bei der Virus-Infektion kann es sich im Rahmen der vorliegenden Erfindung um jede beliebige Infektion handeln, bei der im Verlauf der Erkrankung sich replizierende Virusvarianten entstehen, wobei als ausgewählte Protein-Sequenzen oder proteinkodierende Nukleinsäuresequenzen diejenigen Aminosäuresequenzabschnitte oder

30 für diese kodierende Nukleinsäureabschnitte erfindungsgemäß in Frage kommen, bei denen im Verlauf viraler Erkrankungen stets bzw. häufig Varianten, d.h. Sequenzvariationen, beobachtet werden. Vorzugsweise handelt es sich vorliegend um Sequenzvarianten des GP120-Moleküls (auf Protein-Ebene) bzw. des für GP120 kodierenden

35 Bereichs (auf DNA-Ebene) oder Teilen derselben, insbesondere um

- 27 -

Sequenzvarianten im V3-Loop oder im V2-Loop, vorzugsweise sowohl im V3- als auch im V2-Loop.

Erfindungsgemäß werden folglich pharmazeutische Zusammensetzungen
5 bzw. Vakzinen zur Verfügung gestellt, die bei der immunre-
konstitutiven Behandlung von Virus-Infizierten das Immunsystem
in einer Art und Weise aktivieren, daß die natürlich erworbene
Immunabwehr gegen das Virus regeneriert und neu stimuliert wird,
um so die im Patienten replizierenden Virusvarianten an ihrer
10 Vermehrung zu hindern. Gemäß einer bevorzugten Ausführungsform
der Erfindung werden erstmals Vakzinen zur immunrekonstitutiven
Behandlung von HIV-1-Infizierten bereitgestellt.

Im Rahmen der vorliegenden Erfindung können die Vakzinen entweder
15 als solche, d.h. als DNA- und/oder Protein-Mischungen ohne
weitere Zusätze, oder zusammen mit weiteren Wirkstoffen, wie z.B.
Immunstimulantien wie Interleukin-2, CC- und CXC-Chemokine,
und/oder pharmazeutisch verträglichen Hilfs- und Trägerstoffen
formuliert bzw. verabreicht werden. In der Regel sind die
20 erfindungsgemäßen Vakzinen zur intravenösen Applikation formu-
liert, wie z.B. intravenös, intramuskulär, subcutan. Außerdem
können die Vakzinen zur oralen, mucosalen (intravaginalen, in-
trarektalen) und transdermalen Anwendung formuliert sein.

25 Die vorliegende Erfindung weist folgende Vorteile auf:

Durch eine gezielte gentechnische Manipulation der Nukleinsäure-
sequenz, mit Hilfe einer Genkassette, lassen sich beliebige
Varianten eines Krankheitserregers oder eines Gens des Erregers
30 herstellen. Die Variationen betreffen bevorzugt die Bereiche,
gegen die neutralisierende Antikörper oder cytotoxische T-Zellen
gebildet werden. Dadurch läßt sich ein Impfstoff in Form einer
Mischung von Varianten des Krankheitserregers oder von Varianten
der entsprechenden Antigene eines Krankheitserregers herstellen.
35 Die Vorteile dieses Konzeptes in Bezug auf HIV als Krankheits-
erreger liegen in der Herstellung einer Mischung des HIV-GP120,

dem äußeren Membranprotein des HIV, gegen welches sich die virusneutralisierende Immunantwort des Menschen richtet. Da gerade im Falle der HIV-Erkrankung jeder Patient einer Vielzahl von HIV-Varianten ausgesetzt ist, die sich alle in der Sequenz des GP120 Proteins unterscheiden, ist der Einsatz einer Vielzahl von GP120-Varianten als Impfstoff von besonderem Vorteil, wobei diese GP120-Mischung sowohl für die Immunisierung normaler gesunder Personen aber auch für die Therapie von bereits mit HIV infizierten Personen eingesetzt werden kann. Bei der Immunisierung mit der GP120-Mischung wird eine möglichst breite Immunantwort gegen möglichst viele Virusvarianten im Menschen induziert, die im idealen Falle gegen alle HIV-Varianten schützt. Bei der Therapie mit der GP120-Mischung kann der Verlust von neutralisierenden Antikörpern und der Verlust der HIV-spezifischen zellulären Immunantwort bekämpft werden.

Erfindungsgemäß ist es ferner erstmals möglich, *gp120*-Klone herzustellen, die in der Art bzw. Vielfalt der beobachteten Sequenzvariationen Plasmaisolaten von HIV-Infizierten bzw. an AIDS Erkrankten entsprechen (vgl. M. Schreiber et al., J. Virol. 68 No. 6 (1994) 3908-3916). In diesem Zusammenhang wird im Gegensatz zu verschiedenen bisher verfolgten Ansätzen kein isoliertes GP120-Molekül oder ein antigenes Fragment desselben zur Immunisierung verwendet, sondern eine Protein-Mischung, die aus einer Vielzahl von Sequenzvarianten besteht, die durch eine vollständige GP120-Sequenz charakterisiert sind, die darüber hinaus die korrekte, der natürlichen Faltung des GP120-Moleküls entsprechende Tertiärstruktur aufweisen. Die Konformation der in der Vakzine zur Verfügung gestellten Virusvarianten ist insofern von Bedeutung, als dadurch Sequenz- und Strukturvarianten bereitgestellt werden, die mit den im Verlauf der Viruserkrankung nachweisbaren GP120-Varianten eine größtmögliche Übereinstimmung hinsichtlich der beobachteten Sequenzvariationen als auch der für die Bindung an CD4 und den Corezeptor erforderlichen Konformation erzielt wird, wodurch eine effektive Immunstimulation erzielt werden kann.

Durch die Erfindung wird eine Mischung von GP120-Proteinen bereitgestellt, die die Protein-Vakzine darstellt. Gleichzeitig wird durch die Erfindung eine Mischung der für diese Protein-Mischung kodierenden Gene bereitgestellt. Diese Gene können in
5 einen Vektor überführt werden der sich für die direkte Anwendung am Menschen eignet (DNA-Vakzine) (Ulmer et al., 1995 Ann NY Acad Sci 772:117-125; Donnelly et al., 1995 Ann NY Acad Sci 1995 772:40-46). Eine DNA-Vakzine hat den Vorteil, daß die Heterogenität der Mischung der DNA-Vektoren höher sein kann als die
10 Mischung der rekombinanten GP120-Proteine. Es ist leichter eine hohe Heterogenität eines DNA-Vektorgemisches herzustellen. Eine solche DNA-Vakzine würde ein viel breiteres Spektrum an HIV-Varianten abdecken. Von der technischen Seite betrachtet ist die Herstellung der DNA-Mischung einfacher. Voraussetzung für die
15 DNA-Vakzine ist ein *gp120*-Expressionsvektor, der die beiden Schnittstellen BstEII und BamHI trägt. Dies ermöglicht die Umsetzung der V3-Loop und V2-Loop Variablen *env*-Genfragmente in einen solchen Vektor.

20 Im Hinblick auf die GP120-Proteine ist zu beachten, daß sich im Bereich des V3-Loop 5 potentielle Stellen für die N-Glykosylierung befinden. Die Zuckerreste schützen den V3-Loop vor der Erkennung durch neutralisierende Antikörper. Virusvarianten, die einen glykosylierten V3-Loop aufweisen sind schlechter neutrali-
25 sierbar als Virusvarianten bei denen die Glykosylierung unvollständig ist. Da sich Virusvarianten mit einer vollständigen V3-Loop-Glykosylierung der neutralisierenden Immunantwort entziehen können, sind sie für die Übertragung und für das Etablieren der Infektion von Bedeutung. Im Verlauf der Erkrankung, wenn ein Teil
30 der neutralisierenden Antikörperantwort verlorengegangen ist, setzen sich Virusvarianten durch, deren V3-Loop nicht vollständig glykosyliert ist. Da der V3-Loop nicht mehr durch Zuckerreste verdeckt ist, replizieren diese Virusvarianten schneller als die vollständig glykosylierten V3-Loop Mutanten. Erfindungsgemäß ist
35 es daher von Vorteil, bei der Konstruktion der GP120-Vakzine die unterschiedliche Glykosylierung des GP120 V3-Loop zu berücksichtigen.

- 30 -

Im Rahmen der vorliegenden Erfindung kann das an HIV dargestellte Prinzip auf andere Viren, die im Verlauf der viralen Erkrankung ebenfalls Varianten ausbilden, übertragen werden. Insbesondere können Vakzinen gegen eine Vielzahl von Viren bereitgestellt werden, die eine breite Immunantwort gegen viele Virusvarianten im Menschen induzieren, die im Idealfall gegen alle Varianten schützt. Bei der Therapie mit solchen Vakzinen kann der Verlust von neutralisierenden Antikörpern und der Verlust der HIV-spezifischen zellulären Immunantwort wirksam bekämpft werden. In diesem Zusammenhang ist es ferner möglich oder sogar ratsam, bei Erregertypen (Viren), die im Verlauf einer Erkrankung Sequenzvarianten mehrerer Proteine bilden, mehrere erfindungsgemäße Genkassetten bereitzustellen, die jeweils die für Varianten jedes dieser Proteine oder Fragmente derselben kodierende Nukleinsäuresequenz enthalten, um dem Verlust neutralisierender Antikörper gegen alle denkbaren Virusvarianten entgegenzuwirken.

Im Rahmen der vorliegenden Erfindung können die Vorteile einer Protein-Vakzine und einer DNA-Vakzine in vorteilhafter Weise kombiniert werden, um den Erfolg einer präventiven und/oder therapeutischen Behandlung einer viralen Erkrankung zu erhöhen. Es wird somit ferner eine pharmazeutische Zusammensetzung zur Prävention und/oder Therapie einer Virus-Infektion zur Verfügung gestellt, die eine Protein-Mischung und eine Nukleinsäuremischung umfaßt, wobei die Protein-Mischung Sequenzvarianten eines viralen Proteins oder eines Teils desselben umfaßt, und die Nukleinsäuremischung DNA-Moleküle umfaßt, die für Sequenzvarianten eines viralen Proteins oder eines Teils desselben kodieren. Insbesondere handelt es sich bei der pharmazeutischen Zusammensetzung um ein Kombinationspräparat, das sowohl eine der oben genannten Sequenzvarianten des GP120-Proteins umfassende Protein-Mischung als auch eine der oben genannten, von der *env*-Sequenz in SEQ ID NO: 1 oder SEQ ID NO: 11 oder einem Fragment derselben abgeleitete Nukleinsäuremischung umfaßt.

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

BEISPIELE

- 5 1. Allgemeine Beschreibung des Herstellungsverfahrens
- 1.1 Klonierung von V3-loop kodierenden Oligonukleotiden
 - 1.2 Klonierung der V3-env-Varianten
 - 1.3 Analyse der Variabilität
- 10 2. Material und Methoden
- 2.1 Abkürzungen
 - 2.1.1 Allgemeine Abkürzungen
 - 2.1.2 Nukleinsäuren
 - 15 2.2 Bakterienstämme
 - 2.3 Plasmide
 - 2.4 Enzyme
 - 2.5 Chemikalien
 - 2.6 Oligonukleotide
 - 20 2.7 Molekulargewichtsstandards
 - 2.8 Verwendete Reagenzien-Kits
 - 2.9 Medien
 - 2.10 Sterilisieren von Lösungen
 - 2.11 Anzucht und Lagerung von Bakterien
 - 25 2.12 Herstellung kompetenter Zellen
 - 2.13 Transformation in E.coli
 - 2.14 Plasmid-DNA Präparation
 - 2.15 Auftrennung von DNA in Agarosegelen
 - 2.16 Reinigung von DNA aus Agarosegelen
 - 30 2.17 Auftrennung von DNA in Polyacrylamidgelen
 - 2.18 Schneiden von DNA
 - 2.19 Ligation von DNA
 - 2.20 DNA-Sequenzierung
 - 35 2.21 Herstellung doppelsträngiger DNA mit Hilfe von Oligonukleotiden
 - 2.22 Transfektion von COS-Zellen und CHO-Zellen
 - 2.23 Chromatographische Reinigung der GP120-Mischung
 - 2.24 Herstellung der DNA-Vakzine

40

1. Allgemeine Beschreibung des Herstellungsverfahrens

1.1 Klonierung einer Mischung von V3-loop *gp120*-Varianten

5 Die Oligonukleotide werden wie unter 2.6 beschrieben synthetisiert. Als V3-loop-Sequenz ist beispielhaft eine mutierte Sequenz des HIV-1 Patientenisolates F1-01 (M. Schreiber et al., J. Virol. 68 Nr. 6 (1994) 3908-3916) angegeben. Jede andere Sequenz oder Mischung von Sequenzen ist ebenfalls möglich. Um viele verschiedene Varianten einer V3-loop-Sequenz zu klonieren, werden
10 an bestimmten Positionen der Sequenz anstelle von reinen Nukleotidbausteinen Gemische verwendet. Auf diese Weise erhält man ein Gemisch von Oligonukleotiden, die sich alle in der Sequenz unterscheiden und somit für unterschiedliche V3-loops
15 kodieren. Man bezeichnet solche Oligonukleotid-Mischungen auch als Oligonukleotide mit degenerierten Sequenzen. Ausgehend von chemisch synthetisierten Oligonukleotiden mit degenerierten Sequenzen wird der für die verschiedenen V3-loops kodierende Bereich dargestellt.

20

Ein Oligonukleotid, in Leseraster-Orientierung (forward) wird für den env-Sequenzabschnitt von der BglII Schnittstelle bis zur ersten degenerierten Position synthetisiert. Ein zweites Oligonukleotid, in komplementärer Orientierung (reverse) wird entsprechend der Sequenz ab einer Position, die 15 Basen vor dem
25 Beginn des variablen Bereiches liegt synthetisiert (Methode 2.6). Durch Überlappung der 3'-Region auf einer Länge von insgesamt 15 Basen erfolgt die Hybridisierung beider Oligonukleotide. In einer anschließenden Reaktion mit DNA-Polymerase (z.B. Taq DNA Polymerase oder Klenow-Fragment) entsteht aus den beiden hybridisierten
30 Oligonukleotide ein vollständig doppelsträngiges DNA-Molekül (Methode 2.21).

- 33 -

Das so erhaltene DNA-Gemisch wird mit den Restriktionsendonucleasen BglII und XbaI verdaut. Die Klonierung des DNA-Gemisches erfolgt in den mit BglII und XbaI geschnittenen Expressionsvektor Δ V3-pBSCenvV3 (Methoden 2.15 und 2.16).

5

Dieser Vektor enthält die kodierende Sequenz des *gp160* des HIV-1 Stammes NL4-3 (IIIB). Das NL4-3 env Gen wurde so manipuliert, daß es die Restriktionschnittstellen BglII und XbaI sowie ApaLI, PstI und BclI nur jeweils einmal besitzt. Der für den V3 loop kodierende Bereich, der zwischen den Schnittstellen BglII und XbaI liegt, wurde entfernt und durch eine 15 Basenpaar Sequenz ersetzt, wodurch eine analytische Schnittstelle für das Enzym AscI eingeführt wird. In diesen Vektor wird das BglII und XbaI geschnittene V3-loop DNA-Gemisch kloniert (Methode 2.19). Der Verlust der AscI Schnittstelle im fertigen V3-loop pBSCenvV3 Vektor kann für die Selektion der V3-loop kodierenden Klone benutzt werden (Methode 2.18).

Das so entstandene Plasmidgemisch wird in DH5 α -Bakterien transformiert (Methode 2.12), wobei eine Transformationsrate von $>10^5$ erreicht werden soll. Dadurch erhält man eine Genbibliothek von V3-loop-Fragmenten mit einer Größe von ca. 10^5 Klonen. Das Gemisch soll durch DNA-Sequenzierung analysiert werden (siehe 1.3). Mit Hilfe dieses Verfahrens erhält man ein Gemisch von Klonen, die alle für verschiedene V3-loop Varianten des GP120 Proteins kodieren. Dieses Gemisch von Vektoren dient als Ausgangsprodukt für die Herstellung des Proteingemisches für die Verwendung als Impfstoff und als Ausgangsprodukt für die Anwendung als DNA Vaccine.

30

Für die Herstellung der GP120-Proteinmischung werden die pBSCenvV3 Expressionvektoren in COS-Zellen und CHO-Zellen transfiziert (Methode 2.22). Die Reinigung der verschiedenen GP120-Proteine, der Wirkstoffmischung, wird, wie in der Literatur beschrieben, nach Methode 2.23 durchgeführt.

35

1.2 Klonierung einer Mischung von V2-loop *gp120*-Varianten

Varianten des V1-loop und V2-loop werden in gleicher Weise hergestellt. Der Expressionsvektor Δ V3-pBSCenvV3 besitzt dafür
5 drei zusätzliche Restriktionsschnittstellen. Die Variation des V1-loop erfolgt durch Klonierung in die ApaI und PstI Schnittstellen. Die Variation des V2-loop erfolgt durch Klonierung in die PstI und BclI Schnittstellen.

10

1.3 Analyse der Variabilität

Die Analyse der V3-loop-Gemische erfolgt durch DNA-Sequenzierung (Methode 2.20). Geplant ist die statistische Auswahl von ca.
15 100-200 Klonen, deren V1-, V2- und V3-loop Sequenzen bestimmt werden sollen. Sind alle diese Klone unterschiedlich, kann mit Hilfe einer statistischen Berechnung die Heterogenität der Genbibliothek und somit auch die Heterogenität der Proteinmischung bestimmt werden.

20

2. Material und Methoden

2.1 Abkürzungen

25

2.1.1 Allgemeine Abkürzungen

	μ g	Mikrogramm
	μ l	Mikroliter
	μ mol	Mikromol
30	APS	Ammoniumpersulfat
	bp	Basenpaare
	dNTP	Desoxynukleosidtriphosphat
	DNA	Desoxyribonukleinsäure
	DTT	Dithiothreitol
35	EDTA	Ethylendiamintetraessigsäure
	g	Gramm
	GP	Glykoprotein

- 35 -

	h	Stunde
	HIV	Human Immunodeficiency Virus
	IPTG	Isopropyl- β -D-thiogalaktopyranosid
	MOPS	3-Morpholinopropansulfonsäure
5	OD	Optische Dichte
	PAGE	Poly-Acrylamid-Gel-Elektrophorese
	PCR	Polymerase-Ketten-Reaktion
	RT	Raumtemperatur
10	TEMED	N,N,N',N'-Tetramethylethylendiamin

2.1.2 Nukleinsäuren

	A	Adenin
15	C	Cytosin
	G	Guanin
	T	Thymidin

20 2.2 Bakterienstämme

Escherichia coli DH5 α : F-, endA1, hsdR17, (rk-mk+), supE44, recA1, λ -, gyrA96, relA1, Φ 80d lac z Δ M15

25 2.3 Plasmide

pBSCenvATG Eigenkonstruktion, die Sequenz ist in SEQ ID NO: 10 angegeben.

30 2.4 Enzyme

Restriktionsenzyme	MBI-Fermentas, Gibco-BRL, Biolabs
DNA-Polymerasen	MBI-Fermentas, Gibco-BRL
35 T4-DNA-Ligase	MBI-Fermentas

2.5 Chemikalien

	[α -35S]dATP	Amersham Life Science
	Agarose, ultra pure	GIBCO BRL
5	Ammoniumpersulfat	Merck
	Ampicillin	US Biochemical
	Bacto-Agar	Becton Dickinson
	Bacto-Trypton	Becton Dickinson
	Borsäure	Merck
10	Bromphenolblau	Merck
	Calciumchlorid	Merck
	Desoxyribonukleotide	MBI-Fermentas
	Dithiothreitol	Biotechnik, St. Leon-Rot
	Eisessig	Merck
15	Ethidiumbromid	Sigma
	Ethylendiamintetraessigsäure	Merck
	Glycerin	Merck
	Harnstoff	ICN Biomedicals
	Hefeextrakt	Becton Dickinson
20	Kaliumchlorid	Merck
	Kaliumdihydrogenphosphat	Merck
	Magnesiumchlorid	Merck
	Acrylamid-Mix	Roth
	Natriumacetat	Merck
25	Natriumchlorid	Merck
	di-Natriumhydrogenphosphat	Merck
	Natriumdihydrogenphosphat	Merck
	2-Propanol	Merck
	Sigmacote (Chloriertes Polysiloxan)	Sigma
30	N,N,N',N'-Tetramethylethyldiamin	Merck
	Tris(hydroxymethyl)-aminomethan	GIBCO BRL

2.6 Oligonukleotide

Alle aufgeführten Oligonukleotide wurden mit dem ExpediteTM Nucleic Acid Synthesis System der Firma PE Biosystems (Weiterstadt) hergestellt. Für die Klonierung von env Genen, die sich nur in der Sequenz für den V3 loop unterscheiden, werden Oligonukleotide verwendet. Am Beispiel der Klonierung der V3 loop Region für das HIV-1-Patienteisolat F1-01 sind die Sequenzen der Oligonukleotide aufgeführt.

10

V3-Loop: für die Klonierung in pNL4-3/BglIII-NheI, F1-01, forward:

5'-AAG ATG TAG TAA TTA GAT CTG CCA ATT TCA CAG ACA ATG CTA AAA
CCA TAA TAG TAC AGC TGA ACA CAT CGT TAG AAA TTA ATT GTA CAA GAC

15 CCA ACA ACA AT ACA-3' (SEQ ID NO: 3)

V3-Loop: für die Klonierung in pNL4-3/BglIII-NheI, F1-01:

5'-TTT TGC TCT AGA AAT GTT ACA ATG TGC TTG TCT TAT GTC TCC TGT
20 TGC AGC TTC TGT TGC ATG AAA TGC TCT CCC TGG TCC GAT ATG GAT ACT
ATG-(GA)(AT)(GATC) TTT TCT TGT ATT GTT GTT GGG-3' (SEQ ID NO: 4)

Für die Sequenzierung von V3 loop-Klonen wurden die Oligonukleotide 7010, 7011 und die M13-Standard-Primer verwendet (M13, 25 M13r).

Primer 7010: 5'-CCA TGT ACA AAT GTC AG-3' (SEQ ID NO: 5)

Primer 7011: 5'-AAA ACT GTG CGT TAC AA-3' (SEQ ID NO: 6)

M13-Primer: 5'-GTA AAA CGA CGG CCA GT-3' (SEQ ID NO: 7)

M13r-Primer: 5'-CAG GAA ACA GCT ATG AC-3' (SEQ ID NO: 8)

30

2.7 Molekulargewichtsstandards

1 kb-Leiter MBI-Fermentas

100 bp Leiter MBI Fermentas

35

2.8 Verwendete Reagenzien-Kits

	T7-Sequenzier-Kit	Pharmacia
	Qiaquick PCR Purification Kit	Qiagen
5	Qiagen Plasmid Kit	Qiagen

2.9 Medien

10	YT-Medium:
	10 g Trypton
	5 g Hefeextrakt
	5 g NaCl
15	dYT-Medium:
	16 g Trypton
	10 g Hefeextrakt
	5 g NaCl
20	dYT-Agarplatten:
	10 g Trypton
	5 g Hefeextrakt
	5 g NaCl
	15 g Agar

25

Die angegebenen Mengen beziehen sich auf jeweils 1000 ml deionisiertes Wasser. Autoklaviert wurden die Ansätze bei 121°C und 1,5 bar für 20 min. Sollten die Medien Antibiotika oder andere hitzeempfindliche Reagenzien enthalten, wurden die entsprechenden

30 Mengen sterilfiltriert und nach dem Abkühlen des Mediums zugegeben.

	Ampicillin	3,3 ml/l	(60 mg/ml)
	IPTG	3 ml/l	(100 mM)
35	xgal	3 ml /l	(2% in DMF, nicht filtrieren)

2.10 Sterilisieren von Lösungen und Geräten

Lösungen und Medien sowie Pipettenspitzen, Eppendorfgefäße und Glasgeräte wurden 20-40 min bei 121°C und 1,5 bar autoklaviert.

5 Hitzeempfindliche Lösungen wie z.B. Antibiotika- und IPTG-Lösungen wurden sterilfiltriert.

2.11 Anzucht und Lagerung von Bakterien

10

Bakterienkulturen wurden jeweils mit einer einzelnen Kolonie angeimpft. Zur Isolierung einer Einzelkolonie wurden Zellen einer Flüssigkultur auf einer Agarplatte ausgestrichen. Nach Inkubation bei 37°C über Nacht sind vereinzelt Kolonien gewachsen. Zur
15 kurzfristigen Aufbewahrung der Bakterien wurden die Agar-Platten mit Parafilm versiegelt und bei 4°C gelagert. Für eine dauerhafte Lagerung von Bakterienstämmen wurden 0,75 ml einer YT-Übernachtskultur mit 0,25 ml sterilem Glycerin vermischt, in flüssigem Stickstoff schockgefroren und bei -70°C aufbewahrt.

20

2.12 Herstellung kompetenter Zellen

100 ml YT-Medium wurden mit 100 µl einer Übernachtskultur ange-
25 impft. Die Bakterien wurden bis zum Erreichen einer $OD_{560}=0,4$ bei 37°C im Schüttelinkubator inkubiert und anschließend 8 min bei 4°C und 1000 g zentrifugiert. Alle anschließenden Arbeiten erfolgten auf Eis unter Verwendung vorgekühlter Gefäße und 4°C kalter Lösungen. Die Zellen wurden in 50 ml 50 mM $CaCl_2$ resus-
30 pendiert und 30 min auf Eis inkubiert. Nach erneutem Zentrifugieren wurden die Bakterien in 2,5 ml sterilem TFBII-Puffer aufgenommen und in Portionen von jeweils 100 µl aufgeteilt. Die 100 µl Portionen der kompetenten Zellen konnten, soweit sie nicht für eine sofortige Transformation benötigt wurden, nach Schockge-
35 frieren in flüssigem Stickstoff bei -70°C gelagert werden.

- 40 -

TFBII-Puffer :

	10 mM	MOPS pH 7,0
	75 mM	CaCl ₂
	10 mM	KCl
5	15%	Glycerin

2.13 Transformation von *E. coli*

10 (Hanahan, J. Mol. Biol. 166 (1983) 557-580)

Eine 100 µl Portion kompetenter Zellen wurde im Eisbad aufgetaut und mit 1-100 ng Plasmid-DNA für 30 min bei 0°C inkubiert. Anschließend wurde der Transformationsansatz 1 min bei 42°C inkubiert. Danach wurde 1 ml YT-Medium zugegeben und bei 37°C für
15 eine Stunde geschüttelt. Um eine Selektion der transformierten Zellen zu ermöglichen, wurden 100-500 µl des Ansatzes auf antibiotikahaltigen YT-Agarplatten ausgestrichen. Nach Inkubation über Nacht bei 37°C sind nur Kolonien gewachsen, die das plasmid-kodierte Resistenzgen tragen. Bei einer Transformation von
20 Zellen, die keine funktionsfähige β-Galaktosidase enthalten (lacZΔM15 Mutation, z.B. DH5α), ermöglichen Vektoren wie der verwendete pUC über eine plasmid-kodierte β-Galaktosidase eine direkte Selektion (*blue white screening*) rekombinanter Bakterien. Für die Blau-Weiß-Selektion wurden die Bakterien auf Amp/
25 IPTG/Xgal-YT-Agarplatten ausgestrichen. Das Substrat Xgal wird durch die β-Galaktosidase in einen blauen Farbstoff gespalten. Aufgrund der Zerstörung des Leserasters des lacZ-Gens durch die Insertion eines fremden DNA-Fragments in das lacZ-Gen werden weiße Kolonien erzeugt. Dagegen bedeuten blaue Kolonien, daß das
30 lacZ-Gen bei der Klonierung und Ligation funktionsfähig geblieben ist und keine DNA inseriert wurde.

2.14 Plasmid-DNA Präparation

(Quiagen, Hilden)

Um Plasmid-DNA aus *E. coli* zu isolieren, wurde der QIAprep Spin
5 Miniprep Kit der Firma Qiagen verwendet. Die DNA-Präparation
erfolgte nach Anleitung des Herstellers (QIAprep Plasmid Handbook
03/95). Plasmidhaltige *E. coli*-Bakterien wurden in dYT-amp-Medium
angeimpft und bei 37°C über Nacht geschüttelt. Drei ml der *E.*
coli-Übernachtskultur wurden für die Plasmid Isolierung einge-
10 setzt. Die Bakterien wurden geerntet (1min, 14.000 U/min, Heraeus
Zentrifuge) und in 250 µl P1-Puffer aufgenommen. Durch alkalische
Lyse wurden die Bakterien aufgeschlossen (250 µl, 0,2N NaOH/1%-
SDS). Durch Zugabe von 3M Kaliumacetat, (350 µl, pH 5,5) wurde
die Mischung neutralisiert. Die Aufreinigung der Plasmid-DNA
15 basiert auf der selektiven Bindung von Plasmid-DNA an DEAE-
Anionenaustauschersäulen. Bei den gewählten Salzkonzentrationen
und pH-Bedingungen bindet die Plasmid-DNA an der DEAE-Matrix. Die
DEAE-Matrix wurde gewaschen (750 µl, PE-Puffer, Qiagen). Die
Plasmid-DNA wurde anschließend mit 50 µl H₂O eluiert und bei
20 -20°C gelagert.

Um Plasmid-DNA in größerer Menge (ca. 100 µg) zu erhalten, wurden
25 ml Übernachtskultur eingesetzt. Für die Aufreinigung der DNA
wurde der QIAGEN Plasmid Midi Kit verwendet. Die Aufreinigung
25 beruht hier auf dem unter dem "QIAprep Spin Miniprep Kit"
beschriebenen Prinzip und wurde nach Herstellerangaben (QIAGEN
Plasmid Purification Handbook 01/97) durchgeführt.

2.15 Auftrennung von DNA in Agarosegelen

30

Die Trennung von DNA erfolgte durch Gelelektrophorese in Agarose-
segelen. In Abhängigkeit von der Größe der untersuchten Fragmente
wurden 0,8-2% Agarose in TBE-Puffer (für analytische Agarosegele)
oder TAE-Puffer (für präparative Agarosegele) durch Aufkochen
35 gelöst. Nach dem Abkühlen auf ca. 60°C wurde die Gelflüssigkeit
mit 1 µl Ethidiumbromid (10 mg/ml) versetzt und in ein vor-

bereitetes Gelbett mit eingesetztem Kamm gegossen. Das vollständig abgekühlte Gel wurde in einer Elektrophoresekammer mit TBE-, bzw. TAE-Puffer überschichtet und der Kamm entfernt. Die DNA-Proben wurden mit 1/5 Vol. Auftragspuffer gemischt und in die Probestaschen pipettiert. Die Fragmente wurden bei 5-10 Volt/cm Gellänge 0,5-2 h aufgetrennt. Zur Detektion wurde das Gel nach der Elektrophorese im UV-Durchlicht photographiert. Die Molekulargewichte der DNA-Banden wurden im Vergleich zu mitlaufenden DNA-Molekulargewichtsmarkern bestimmt.

10

TBE-Puffer :

100 mM	Tris
100 mM	Borsäure
3 mM	EDTA

15

TAE-Puffer :

40 mM	Tris
2mM	EDTA
0,114%	Eisessig

20

Auftragspuffer :

0,25%	Bromphenolblau
0,25%	Xylencyanol
30%	Glycerin
50 mM	EDTA

25

2.16 Reinigung von DNA aus Agarosegelen

Für die Extraktion der DNA aus Agarosegelen wurde der "QIAquick Gel Extraction Kit" (Qiagen, Hilden) verwendet. Die Pufferbedingungen wurden so gewählt, daß die Nukleinsäuren an die Silica-Membran der Säulen binden, während niedermolekulare Bakterienbestandteile die Membran passieren (Methode 2.15). Es wurde nach dem "QIAquick Gel Extraction Kit Protokoll" des QIAquick Spin

35

Handbooks 07/97 vorgegangen. Die gereinigte DNA wurde jeweils mit 50 µl H₂O von der Silica-Membran eluiert.

5 2.17 Auftrennung von DNA in Polyacrylamidgelen

Die Trennung von radioaktiv markierten DNA-Fragmenten zur DNA-Sequenzierung wurde unter Verwendung von denaturierenden Polyacrylamidgelen durchgeführt. Zwei gereinigte Glasplatten wurden mit Ethanol von Fettresten befreit und mit 1 ml Sigmacote pro Platte beschichtet, um eine hydrophobe Oberfläche zu erhalten. Durch die Beschichtung sollte ein späteres Zerreißen des Gels beim Abziehen von der Glasplatte vermieden werden. Zwischen die beiden Glasplatten wurden Abstandshalter (*Spacer*) gelegt, welche gleichzeitig zum seitlichen Abdichten dienten. Die ganze Apparatur wurde mit mehreren Klammern fixiert. Zur Herstellung des Sequenzgels wurden 21 g Harnstoff in etwa 20 ml Wasser gelöst. Nach Zugabe von 7,5 ml Acrylamid-Mix und 5 ml 10x TBE-Puffer (siehe 2.15) wurde der Ansatz mit Wasser auf 50 ml aufgefüllt. Die Polymerisation des Gels wurde durch Zugabe von 300 µl APS und 100 µl TEMED gestartet. Sofort danach wurde die Gellösung in die vorbereitete Gelapparatur gegossen, und durch Einsetzen des flachen Rückens des Sägezahnkamms der Taschenboden geformt. Bis zur vollständigen Polymerisation wurde das Gel waagerecht bei Raumtemperatur gelagert. Nach dem Einsetzen des Gels in die Gelkammer wurde diese mit TBE-Puffer gefüllt. Durch Umdrehen des Sägezahnkammes wurden die Auftragstaschen gebildet. Um das Gel auf die optimale Betriebstemperatur zu erwärmen, wurde ein Vorlauf von 20 min durchgeführt. Die Taschen wurden gründlich mit TBE-Puffer gespült und anschließend mit 4-5 µl der Proben gefüllt. Die Elektrophorese erfolgte bei 2 kV (150 Watt) für etwa 2-3 h. Das Gel wurde nach Beenden der Elektrophorese aus der Gelkammer genommen und eine der Glasplatten vorsichtig abgehoben. Durch Auflegen und leichtes Andrücken wurde das Gel auf Schleicher & Schuell-Papier (Whatman, England) übertragen. Nach dem Trocknen bei 80°C unter Vakuum wurde das Gel 1-3 Tage bei RT auf

- 44 -

einem Röntgen-Film (Kodak BioMax MR-1, Sigma Deisenhofen) exponiert.

Acrylamid-Mix:

- 5 40% Polyacrylamid
 0,8% Bisacrylamid

2.18 Schneiden von DNA

10

Restriktionsendonukleasen erkennen und hydrolysieren enzymespezifische palindrome DNA-Sequenzen von meist 4-8 Nukleotiden Länge. Bei der Hydrolyse entstehen je nach Art des Enzyms stumpfe (*blunt ends*) oder überhängende Enden einzelsträngiger DNA (*sticky ends*).

- 15 Für einen Restriktionsverdau wurden 0,1-60 µg DNA mit 2-120 Einheiten (*Units*) eines Restriktionsenzym in dem vom Hersteller angegebenen Puffer für 2-5 h bei 37°C inkubiert. Das Gesamtvolumen betrug zwischen 10 µl für analytische und 300 µl für präparative Ansätze. Für Restriktionen mit zwei verschiedenen
- 20 Enzymen wurde ein Puffer gewählt, in dem beide Enzyme eine ausreichende Aktivität besitzen, z.B. den EcoRI-Puffer für einen EcoRI/BamHI-Verdau, oder es wurden nach Inkubation mit dem ersten Enzym die Bedingungen für das zweite Enzym eingestellt.

25

2.19 Ligation von DNA

DNA-Ligasen katalysieren die Verknüpfung von DNA-Molekülen unter Verbrauch von NAD⁺ oder ATP durch Bildung einer Phosphodiesterbindung zwischen einer freien 5'-Phosphatgruppe und einer 3'-Hydroxylgruppe. Ein drei- bis fünffacher Überschuß an DNA-Fragment wurde mit 100-500 ng geschnittenen Plasmids und 5 Einheiten T4-DNA-Ligase sowie 10 nmol ATP über Nacht bei 12°C in Ligationspuffer inkubiert. Es wurden nur Ligationen kohäsiver Enden

35 durchgeführt.

Ligationspuffer:

	40 mM	Tris-HCl pH 7,8
	10 mM	MgCl ₂
	10 mM	DTT
5	0,5 mM	ATP

2.20 DNA-Sequenzierung

10 (Sanger et al., Proc. Natl. Acad. Sci. USA 74 (1977) 5463-5467)
Die Sequenzierung von DNA wurde nach der Kettenabbruchmethode
durchgeführt. Als Substrat diente doppelsträngige Plasmid-DNA.
Aus Einzelstrang-DNA kann ausgehend von einem Oligonukleotid-
15 primer in Anwesenheit von dNTPs durch eine DNA-Polymerase dop-
pelsträngige DNA synthetisiert werden. Ist im Nukleotidgemisch
ein geringer Anteil Didesoxynukleotide (ddNTPs) enthalten, führt
dies zu einem zufällig verteiltem Einbau des ddNTPs, da die DNA-
Polymerase nicht zwischen dNTPs und ddNTPs unterscheiden kann.
Der Einbau führt aufgrund der fehlenden 3'-Hydroxylgruppe der
20 ddNTPs zu einem Abbruch der Doppelstrangsynthese und, da er
statistisch verteilt erfolgt, zu unterschiedlich langen DNA-
Einzelsträngen. Da der Syntheseansatz viergeteilt ist und in
jeder Probe nur eines der vier ddNTPs vorhanden ist, kommt es in
den jeweiligen Ansätzen zu basenspezifischen Kettenabbrüchen. Zur
25 Markierung des synthetisierten Einzelstranges wird der Reaktion
 α -³⁵S-dATP zugegeben. Nach Denaturierung und Trennung der DNA
über Polyacrylamidgelelektrophorese lassen sich die Banden
autoradiographisch detektieren und die Sequenz des DNA-Stranges
direkt ablesen.

30

Durchgeführt wurden die Sequenzierungen mit dem T7-Sequenzier-Kit
(Pharmacia) nach der Anleitung des Herstellers. Als Oligonu-
kleotidprimer wurde dabei entweder der M13-Universal-Primer
(Pharmacia) oder Primer M13r verwendet. 32 µl (ca. 2 µg) gerei-
35 nigter doppelsträngiger Plasmid-DNA wurden mit 8 µl 2 M NaOH für
10 min bei RT denaturiert. Nach Zugabe von 7 µl 3 M Natriumacetat

- 46 -

pH 4,8, 4 µl H₂O und 120 µl -20°C Ethanol wurde die DNA 20 min bei -70°C gefällt. Durch Zentrifugieren (15 min, 4°C) wurde die DNA isoliert, zweimal mit -20°C kaltem 70% Ethanol gewaschen und in einer Vakuumzentrifuge getrocknet. Das DNA-Sediment wurde
5 anschließend in 10 µl H₂O resuspendiert und nach Zugabe von 2 µl *Annealing*-Puffer und 2 µl Primerlösung (10 pmol in Wasser) bei 65°C für 5 min, 37°C für 10 min und 5min bei RT mit dem Oligonukleotidprimer hybridisiert. Direkt darauf wurden zwecks Primere-
10 elongation und radioaktiver Markierung 3 µl *Labelling*-Mix, 1 µl α-³⁵S-dATP (1000 Ci/mmol, 10 µCi/µl) und 2 µl T7-Polymeraselösung (mit *Enzyme Dilution* Puffer 1:5 verdünnt) zugegeben und 5 min bei RT inkubiert. Jeweils 4,5 µl dieses Ansatzes wurden auf eine, auf
15 37°C vorgewärmte *MicroSample Plate* (Greiner) gegeben, in der je 2,5 µl der vier verschiedenen dNTP/ddNTP-Mixe vorlagen. Nach fünfminütiger Inkubation bei 37°C, die der weiteren Elongation und basenspezifischen Termination diente, wurden die Reaktionen durch Zugabe von 5 µl Stopplösung beendet. Vor dem Auftrag auf das Polyacrylamidgel wurden die Proben für 2 min bei 80°C denaturiert. Gelagert wurden die Proben bei -20°C.

20

Annealing-Puffer:

1 M	Tris-HCl pH 7,6
100 mM	MgCl ₂
160 mM	DTT

25

Labelling-Mix:

1,375 µM	dCTP
1,375 µM	dGTP
1,375 µM	dTTP
333,5 mM	NaCl

30

Enzyme Dilution Puffer:

20 mM	Tris-HCl pH 7,5
5 mM	DTT
100 µg/ml	BSA
5%	Glycerin

35

- 47 -

Stopplösung:

	10 mM	EDTA
	97,5%	Formamid
	0,3%	Bromphenolblau
5	0,3%	Xylencyanol

	A-Mix:	840 μ M dCTP	C-Mix:	840 μ M dATP
		840 μ M dGTP		840 μ M dGTP
		840 μ M dTTP		840 μ M dTTP
10		93,5 μ M dATP		93,5 μ M dCTP
		14 μ M ddATP		14 μ M ddCTP
		40 mM Tris-HCl pH 7,6		40 mM Tris-HCl pH 7,6
		50 mM NaCl		50 mM NaCl
15	G-Mix:	840 μ M dATP	T-Mix:	840 μ M dATP
		840 μ M dCTP		840 μ M dCTP
		840 μ M dTTP		840 μ M dGTP
		93,5 μ M dGTP		93,5 μ M dTTP
		14 μ M ddGTP		14 μ M ddTTP
20		40 mM Tris-HCl pH 7,6		40 mM Tris-HCl pH 7,6
		50 mM NaCl		50 mM NaCl

2.21 Herstellung doppelsträngiger DNA mit Hilfe von Oligonukleotiden

Für die Herstellung des DNA-Gemisches wurde die Oligonukleotide bei 60°C hybridisiert. Je Oligonukleotid wurden 100 pmol eingesetzt. Von der hybridisierten Probe wurden 1 pmol für die PCR oder 100 pmol für die Klenow-Reaktion eingesetzt.

- 48 -

Für die PCR wurden folgende Zyklen durchgeführt:

	Am Anfang:	10 min	94°C
	30 Zyklen:	1 min	94°C
		1 min	45°C
5		1 min	72°C
	Am Schluß:	10 min	72°C

Als DNA-*template* für die PCR wurden 1 pmol hybridisierte Oligonukleotide eingesetzt. PCR-Primer wurden bei allen durchgeführten
10 PCR-Reaktionen in einer Endkonzentration von 0,1 pmol/ μ l eingesetzt. Von der Taq-Polymerase wurden 1 Unit pro 50 μ l des Reaktionsansatzes benutzt. Die Konzentration der dNTPs im PCR-Ansatz wurde auf 0,1 mM eingestellt.

15 Für die Herstellung des DNA-Gemisches mit Hilfe der Klenow-Reaktion wurden 100 pmol der hybridisierten Oligonukleotide in Klenow-Puffer (50 mM Tris-HCl, pH 8,0; 50 mM MgCl₂, 10 mM DTT, 0,05 mM dNTP) gelöst. Die Reaktion wurde durch Zugabe von 5 Units DNA-Polymerase I (Klenow-Fragment) gestartet. Nach 30 min bei 37°C
20 wurde die Reaktion durch Erhitzen auf 75°C (10 min) gestoppt.

2.22 Transfektion von COS-Zellen und CHO-Zellen

25 5-10 μ g linearisierte Plasmid-DNA werden in 150 μ l Zellkulturmedium (ohne FKS und Antibiotika) gelöst. Der Lösung werden 30 μ l des Transfektionsmittels zugegeben (SuperFect, Qiagen, Hilden). Die Mischung wird 10 min bei RT inkubiert und nach Zugabe von 1 ml Medium auf die Zellen gegeben. Die COS-Zellen (80% konfluent
30 gewachsen) wurden in 60 mm-Platten gezüchtet und kurz vor der Transfektion mit PBS gewaschen. Nach einer Inkubation von 3 h bei 37°C und 5% CO₂ wurde das Transfektionsmedium entfernt. Anschließend wurden die Zellen 4x mit PBS gewaschen und in Selektionsmedium (DMEM, 5% FKS, 600 μ g/ml Geneticin) aufgenommen.

2.23 Chromatographische Reinigung der GP120-Mischung

Die chromatographische Reinigung erfolgt nach Standardmethoden (Techniques in HIV Research, Aldovini & Walker, Stockton Press, 1990; S.W. Pyle et al. Purification of 120,00 dalton envelope glycoprotein from culture fluids of human immunodeficiency virus (HIV)-infected H9 cells. AIDS Res Hum Retroviruses 1987, 3:387-400). Zellextrakte und Zellkulturüberstände von GP120-exprimierenden COS-Zellen werden für die Isolierung der GP120-Mischung eingesetzt. Nach Zentrifugation (25 000 g) wird das Lysat wie folgt aufgereinigt:

1. Gelfiltration (Sephadex G-20)
2. Immunoaffinitätschromatographie
(anti-GP120-Antikörper gebunden an CH-Sepharose 4B)
3. Anreicherung durch Proteinfällung
4. Anreicherung durch Dialyse

2.24 Herstellung der DNA-Vakzine


Ausgangsmaterial für die Herstellung der DNA-Vakzine ist die Mischung der BstEII-BamHI DNA-Fragmente des *env*-Gens, die für die Herstellung der GP120-Protein Mischung verwendet wurden. Ein eukaryotischer Expressionsvektor für das HIV-1 GP120 der für DNA-Vakzinierungen zugelassen ist (vgl. z.B. J.D. Boyer et al., J Infect Dis 1997, 176:1501-1509; J.D. Boyer et al., Nat Med 1997, 3:526-532; M.L. Bagarazzi et al., J Med Primatol 1997, 26:27-33) wird so verändert, daß sich im *env*-Gen die Schnittstellen für BstEII und BamHI an identischen Stellen des *gp120*-Leserasters befinden. In einen solchen Vektor werden die mutierten, variablen *env*-Genfragmente (BstEII-BamHI) kloniert. Die Veränderung der *env*-Gene im V2-Loop- und V3-Loop-Bereich und deren Klonierung in den DNA-Vakzine-Vektor erfolgt analog der Klonierungsschritte die für die Herstellung der *gp120*-Mischung durchgeführt werden.

UDAPESTER VERTRAG ÜBER DIE INTERNATIONALE
ANERKENNUNG DER HINTERLEGUNG VON MIKROORGANISMEN
FÜR DIE ZWECKE VON PATENTVERFAHREN

INTERNATIONALES FORMBLATT

Strathmann AG & Co.
Sellhopsweg 1
22459 Hamburg

EMPFANGSBESTÄTIGUNG DER ERSTHINTERLEGUNG,
ausgestellt gemäß Regel 7.1 von der unten angegebenen
INTERNATIONALEN HINTERLEGUNGSSTELLE

I. KENNZEICHNUNG DES MIKROORGANISMUS	
Vom HINTERLEGER zugeteiltes Bezugszeichen: pBSCenv-V3	Von der INTERNATIONALEN HINTERLEGUNGSSTELLE zugeeilte EINGANGSNUMMER: DSM 12612
II. WISSENSCHAFTLICHE BESCHREIBUNG UND/ODER VORGESCHLAGENE TAXONOMISCHE BEZEICHNUNG	
<p>Mit dem unter I. bezeichneten Mikroorganismus wurde</p> <p><input checked="" type="checkbox"/> eine wissenschaftliche Beschreibung</p> <p><input checked="" type="checkbox"/> eine vorgeschlagene taxonomische Bezeichnung</p> <p>eingereicht. (Zutreffendes ankreuzen).</p>	
III. EINGANG UND ANNAHME	
Diese internationale Hinterlegungsstelle nimmt den unter I bezeichneten Mikroorganismus an, der bei ihr am 1999-01-06 (Datum der Ersthinterlegung) ¹ eingegangen ist.	
IV. EINGANG DES ANTRAGS AUF UMWANDLUNG	
Der unter I bezeichnete Mikroorganismus ist bei dieser internationalen Hinterlegungsstelle am 1999-01-06 eingegangen (Datum der Ersthinterlegung) und ein Antrag auf Umwandlung dieser Ersthinterlegung in eine Hinterlegung gemäß Budapest Vertrag ist am 1999-01-06 eingegangen (Datum des Eingangs des Antrags auf Umwandlung).	
V. INTERNATIONALE HINTERLEGUNGSSTELLE	
Name: DSMZ-DEUTSCHE SAMMLUNG VON MIKROORGANISMEN UND ZELLKULTUREN GmbH Anschrift: Mischeröder Weg 1b D-38124 Braunschweig	Unterschrift(en) der zur Vertretung der internationalen Hinterlegungsstelle befugten Person(en) oder des (der) von ihr ermächtigten Bediensteten:  Datum: 1999-01-11

¹ Falls Regel 6.4 Buchstabe d zutrifft, ist dies der Zeitpunkt, zu dem der Status einer internationalen Hinterlegungsstelle erworben worden ist.


BUDAPESTER VERTRAG ÜBER DIE INTERNATIONALE
ANERKENNUNG DER HINTERLEGUNG VON MIKROORGANISMEN
FÜR DIE ZWECKE VON PATENTVERFAHREN

INTERNATIONALES FORMBLATT

Strathmann AG & Co.
Sellhopsweg 1

22459 Hamburg

LEBENSFÄHIGKEITSDESCHENICUNG
ausgestellt gemäß Regel 10.2 von der unten angegebenen
INTERNATIONALEN HINTERLEGUNGSSTELLE

I. HINTERLEGER	II. KENNZEICHNUNG DES MIKROORGANISMUS
Name: Strathmann AG & Co. Sellhopsweg 1 Anschrift: 22459 Hamburg	Von der INTERNATIONALEN HINTERLEGUNGSSTELLE zugewiesene EINTRAGSNUMMER: DSM 12612 Datum der Hinterlegung oder Weiterleitung: 1999-01-06
III. LEBENSFÄHIGKEITSDESCHENICUNG	
Die Lebensfähigkeit des unter II genannten Mikroorganismus ist am 1999-01-07 ¹ geprüft worden. Zu diesem Zeitpunkt war der Mikroorganismus (X) ¹ lebensfähig () ¹ nicht mehr lebensfähig	
IV. BEDINGUNGEN, UNTER DENEN DIE LEBENSFÄHIGKEITSPRÜFUNG DURCHFÜHRT WORDEN IST ²	
V. INTERNATIONALE HINTERLEGUNGSSTELLE	
Name: DSMZ-DEUTSCHE SAMMLUNG VON MIKROORGANISMEN UND ZELLKULTUREN GmbH Anschrift: Mascheroder Weg 1b D-38124 Braunschweig	Unterschrift(en) der zur Vertretung der internationalen Hinterlegungsstelle befugten Person(en) oder des (der) von ihr ernittelten Bediensteten:  Datum: 1999-01-11

Angabe des Datums der Ershinterlegung. Wenn eine erneute Hinterlegung oder eine Weiterleitung vorgenommen worden ist, Angabe des Datums der jeweils letzten erneuten Hinterlegung oder Weiterleitung.

In den in Regel 10.2 Buchstabe a Ziffer ii und iii vorgesehenen Fällen Angabe der letzten Lebensfähigkeitsprüfung.
 Zutreffendes ankreuzen.

Ausfüllen, wenn die Angaben beantragt worden sind und wenn die Ergebnisse der Prüfung negativ waren.

Patentansprüche

1. Protein-Vakzine, die eine Mischung viraler Protein-Moleküle umfaßt, **dadurch gekennzeichnet**, daß die Moleküle Sequenzvarianten eines einzigen viralen Proteins oder eines Teils desselben sind.
2. Protein-Vakzine nach Anspruch 1, **dadurch gekennzeichnet**, daß die Mischung $\geq 10^2$ Sequenzvarianten enthält.
3. Protein-Vakzine nach Anspruch 2, **dadurch gekennzeichnet**, daß die Mischung $\geq 10^3$ und vorzugsweise $\geq 10^4$ Sequenzvarianten enthält.
4. Protein-Vakzine nach einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, daß sie eine Mischung von GP120-Proteinen von HIV umfaßt, die sich jeweils in ihrer Aminosäuresequenz im Bereich der V2-Scheife und/oder der V3-Schleife voneinander unterscheiden.
5. DNA-Vakzine, die für eine Mischung strukturell unterschiedlicher Virus-Proteine kodiert, **dadurch gekennzeichnet**, daß die Vakzine eine Mischung von Sequenzvarianten eines viralen DNA-Moleküls oder eines Teils desselben enthält.
6. DNA-Vakzine nach Anspruch 2, **dadurch gekennzeichnet**, daß sie eine Mischung von DNA-Molekülen enthält, die Sequenzvarianten eines viralen Proteins oder eines Teils kodieren.
7. DNA-Vakzine nach Anspruch 5 oder 6, **dadurch gekennzeichnet**, daß die Mischung $\geq 10^2$ DNA-Moleküle enthält, die sich in ihrer Nukleinsäuresequenz voneinander unterscheiden.
8. DNA-Vakzine nach Anspruch 7, **dadurch gekennzeichnet**, daß die Mischung $\geq 10^3$ und vorzugsweise $\geq 10^4$ DNA-Moleküle enthält,

die sich in ihrer Nukleinsäuresequenz voneinander unterscheiden.

9. DNA-Vakzine nach einem der Ansprüche 5 bis 8, **dadurch gekennzeichnet**, daß sie für eine Mischung strukturell unterschiedlicher GP120-Proteine von HIV kodiert, wobei die Vakzine eine Mischung von DNA Molekülen enthält, deren Nukleinsäuresequenzen sich in dem für die V2-Schleife kodierenden Bereich und/oder in dem für die V3-Schleife kodierenden Bereich voneinander unterscheiden.
10. DNA-Vakzine nach Anspruch 9, **dadurch gekennzeichnet**, daß sie eine Mischung von DNA-Molekülen enthält, die sich in ihrer Nukleinsäuresequenz derart voneinander unterscheiden, daß sie für eine Mischung von GP120-Proteinen kodieren, die in der V2-Schleife und/oder in der V3-Schleife jeweils voneinander verschiedene Aminosäuresequenzen aufweisen.
11. Nukleinsäuresequenz, die von der in SEQ ID NO: 1 dargestellten *env*-Sequenz oder einem Fragment derselben abgeleitet ist, **dadurch gekennzeichnet**, daß sie derart modifiziert ist, daß sie ausschließlich monovalente Restriktionsspaltorte enthält.
12. Nukleinsäuresequenz nach Anspruch 11, **dadurch gekennzeichnet**, daß die Sequenz durch Einführen stiller Mutationen modifiziert ist.
13. Nukleinsäuresequenz nach Anspruch 11 oder 12, **dadurch gekennzeichnet**, daß sie die in SEQ ID NO: 9 angegebene Sequenz aufweist.
14. Nukleinsäuresequenz, **dadurch gekennzeichnet**, daß sie die in SEQ ID NO: 11 angegebene Sequenz aufweist.

15. Nukleinsäuresequenz, **dadurch gekennzeichnet**, daß sie die in SEQ ID NO: 12 angegebene Sequenz aufweist.
16. Einzelsträngige Nukleinsäuresequenz, die den für die V3-Schleife und/oder den für die V2-Schleife von GP120 kodierenden Bereich enthält, **dadurch gekennzeichnet**, daß in der V3-Schleife ein 247 bp langes BglII-XbaI-Fragment oder ein 283 bp langes BglII-NheI-Fragment gegen ein verändertes Fragment ausgetauscht ist, und in der V2-Schleife ein 139 bp langes PstI-BclI-Fragment oder ein 339 bp langes PstI-EcoRI-Fragment gegen ein verändertes Fragment ausgetauscht ist, wobei das Fragment/die Fragmente jeweils an mindestens 6, vorzugsweise an 9 bis 20 Positionen, Inosin oder einen Nukleinsäureaustausch aufweist/aufweisen.
17. Einzelsträngige Nukleinsäuresequenz, die den für die V2-Schleife kodierenden Bereich und/oder den für die V3-Schleife kodierenden Bereich von GP120 enthält, **dadurch gekennzeichnet**, daß in der V3-Schleife ein 247 bp langes BglII-XbaI-Fragment oder ein 283 bp langes BglII-NheI-Fragment gegen ein verändertes Fragment ausgetauscht ist, und in der V2-Schleife ein 139 bp langes PstI-BclI-Fragment oder ein 339 bp langes PstI-EcoRI-Fragment gegen ein verändertes Fragment ausgetauscht ist, wobei das Fragment/die Fragmente jeweils an mindestens 6, vorzugsweise an 9 bis 20 Positionen, eine Mutation aufweist/aufweisen.
18. Doppelsträngige DNA, die Hybride der einzelsträngigen Nukleinsäuresequenz nach Anspruch 16 mit der einzelsträngigen Nukleinsäuresequenz nach Anspruch 17 umfaßt.
19. Nukleinsäure-Mischung, die doppelsträngige DNAs umfaßt, deren Nukleinsäuresequenzen von der *env*-Sequenz in SEQ ID NO: 1 oder SEQ ID NO: 11 oder einem Fragment derselben abgeleitet sind, **dadurch gekennzeichnet**, daß sich die Nukleinsäuresequenzen in dem für die V2-Schleife kodierenden

Bereich und/oder in dem für die V3-Schleife kodierenden Bereich jeweils voneinander unterscheiden.

20. Nukleinsäure-Mischung nach Anspruch 19, dadurch gekennzeichnet, daß sich die Nukleinsäuresequenz derart voneinander unterscheiden, daß sie für eine Mischung von Proteinen kodieren, die in der V2-Scheife und/oder in der V3-Schleife jeweils voneinander verschiedene Aminosäuresequenzen aufweisen.
21. Nukleinsäure-Mischung nach Anspruch 20, dadurch gekennzeichnet, daß die Mischung $\geq 10^2$ Sequenzvarianten enthält.
22. Nukleinsäure-Mischung nach Anspruch 21, dadurch gekennzeichnet, daß die Mischung $\geq 10^3$ und vorzugsweise $\geq 10^4$ Sequenzvarianten enthält.
23. Protein-Mischung, die Sequenzvarianten des GP120-Proteins umfaßt, dadurch gekennzeichnet, daß es eine Mischung von Proteinen ist, die in der V2-Scheife und/oder in der V3-Schleife jeweils voneinander verschiedene Aminosäuresequenzen aufweisen.
24. Protein-Mischung nach Anspruch 23, dadurch gekennzeichnet, daß die Mischung $\geq 10^2$ Sequenzvarianten enthält.
25. Protein-Mischung nach Anspruch 24, dadurch gekennzeichnet, daß die Mischung $\geq 10^3$ und vorzugsweise $\geq 10^4$ Sequenzvarianten enthält.
26. Plasmid, welches eine doppelsträngige DNA nach Anspruch 18 insertiert enthält.
27. Expressionsvektor, dadurch gekennzeichnet, daß er eine Nukleinsäuresequenz nach den Ansprüchen 11 bis 15 insertiert enthält.

28. Expressionsvektor nach Anspruch 27, **dadurch gekennzeichnet**, daß er die in SEQ ID NO: 10 angegebene Sequenz aufweist.
29. Expressionsvektor, **dadurch gekennzeichnet**, daß er DSM 12612 entspricht.
30. Vektor-Mischung, die eine Mischung von Plasmiden nach Anspruch 26 enthält, **dadurch gekennzeichnet**, daß sich die Nukleinsäuresequenzen der Plasmide in dem für die V2-Schleife kodierenden Bereich und/oder in dem für die V3-Schleife kodierenden Bereich jeweils voneinander unterscheiden.
31. Vektor-Mischung nach Anspruch 30, **dadurch gekennzeichnet**, daß die Mischung $\geq 10^2$ Plasmide enthält, die sich in ihrer Nukleinsäuresequenz voneinander unterscheiden.
32. Vektor-Mischung nach Anspruch 31, **dadurch gekennzeichnet**, daß die Mischung $\geq 10^3$ und vorzugsweise $\geq 10^4$ Plasmide enthält, die sich in ihrer Nukleinsäuresequenz voneinander unterscheiden.
33. Vektor-Mischung nach einem der Ansprüche 30 bis 32, **dadurch gekennzeichnet**, daß sich die Plasmide in *E. coli* als Wirtszelle exprimieren lassen.
34. Vektor-Mischung nach einem der Ansprüche 30 bis 32, **dadurch gekennzeichnet**, daß sich die Plasmide in Eukaryontenzellen, vorzugsweise in Cos-, CHO- oder BHK-Zellen, als Wirtszellen exprimieren lassen.
35. *E. coli*-Wirtszellen, die mit einer Vektor-Mischung nach Anspruch 33 transfiziert sind.
36. Eukaryontische Wirtszellen, die mit einer Vektor-Mischung nach Anspruch 34 transfiziert sind.

37. Eukaryontische Wirtszellen nach Anspruch 36, **dadurch gekennzeichnet**, daß sie eine Wirtszelle aus der Gruppe bestehend aus Cos-, BHK- oder CHO-Zellen ist.
38. Verfahren zur Herstellung der Nukleinsäuresequenz nach Anspruch 12, **dadurch gekennzeichnet**, daß man in eine für ein virales Protein kodierende Nukleinsäuresequenz so viele stille Mutationen einführt, daß die erhaltene Nukleinsäuresequenz nur noch monovalente Restriktionsspaltorte enthält.
39. Verfahren nach Anspruch 38, **dadurch gekennzeichnet**, daß die für ein virales Protein kodierende Nukleinsäuresequenz die Sequenz gemäß SEQ ID NO: 1 oder SEQ ID NO: 11 oder ein Fragment derselben ist, in die man so viele stille Mutationen einführt, daß sie nur noch monovalente Restriktionsspaltorte enthält.
40. Verfahren zur Herstellung der Vektor-Mischung nach den Ansprüchen 33 und 34, **dadurch gekennzeichnet**, daß man Plasmide, deren Nukleinsäuresequenzen sich in dem für die V2-Schleife kodierenden Bereich und/oder in dem für die V3-Schleife kodierenden Bereich jeweils voneinander unterscheiden, in einen in Wirtszellen, vorzugsweise in *E. coli*-, Cos-, CHO- oder BHK-Zellen, exprimierbaren Vektor hineinligiert.
41. Verfahren zur Herstellung der Wirtszellen nach Anspruch 35 oder 36, **dadurch gekennzeichnet**, daß man die Wirtszellen mit einer Vektor-Mischung nach den Ansprüchen 30 bis 32 transformiert.
42. Verfahren zur Herstellung einer Protein-Vakzine nach einem der Ansprüche 1 bis 4, **dadurch gekennzeichnet**, daß man die Wirtszellen nach einem der Ansprüche 35 bis 37 unter

Bedingungen kultiviert, die die Expression der Mischung viraler Protein-Sequenzvarianten gestatten.

43. Verfahren zur Herstellung einer DNA-Vakzine nach einem der Ansprüche 5 bis 10, **dadurch gekennzeichnet**, daß man das Verfahren nach Anspruch 40 durchführt, wobei man die erfindungsgemäßen Plasmide in einen Vektor hineinligiert, der in Wirtszellen des zu vakzinierenden Organismus exprimierbar ist.
44. Verwendung einer Mischung strukturell unterschiedlicher viraler Proteine, die Sequenzvarianten eines viralen Proteins oder eines Teils desselben sind, zur Herstellung einer Vakzine zur Prävention und/oder Therapie einer Virus-Infektion beim Menschen.
45. Verwendung einer Protein-Mischung nach einem der Ansprüche 23 bis 25 zur Herstellung einer Vakzine zur Prävention und/oder Therapie einer HIV-Infektion beim Menschen.
46. Verwendung einer Mischung von DNA-Molekülen, die für Sequenzvarianten eines viralen Proteins oder eines Teils desselben kodieren, zur Herstellung einer Vakzine zur Prävention und/oder Therapie einer Virus-Infektion beim Menschen.
47. Verwendung einer Nukleinsäure-Mischung nach den Ansprüchen 19 bis 22 zur Herstellung einer Vakzine zur Prävention und/oder Therapie einer Virus-Infektion beim Menschen.
48. Verwendung der Nukleinsäure-Mischung nach einem der Ansprüche 19 bis 22 zur Herstellung einer Vektor-Mischung nach einem der Ansprüche 30 bis 32, die in Wirtszellen exprimierbar ist, wobei die Wirtszellen aus der Gruppe bestehend aus *E. coli*-, Cos-, CHO- und BHK-Zellen ausgewählt ist.

49. Verwendung der Vektor-Mischung nach einem der Ansprüche 30 bis 32 zur Expression einer Protein-Mischung nach einem der Ansprüche 23 bis 25.
50. Verwendung der Wirtszelle nach einem der Ansprüche 35 bis 37 zur Herstellung einer Protein-Mischung nach einem der Ansprüche 23 bis 25.
51. Pharmazeutische Zusammensetzung zur Prävention und/oder Therapie einer Virus-Infektion, **dadurch gekennzeichnet**, daß sie eine Protein-Mischung und eine Nukleinsäuremischung umfaßt, wobei die Protein-Mischung Sequenzvarianten eines viralen Proteins oder eines Teils desselben umfaßt, und die Nukleinsäuremischung DNA-Moleküle umfaßt, die für Sequenzvarianten eines viralen Proteins oder eines Teils desselben kodieren.
52. Pharmazeutische Zusammensetzung nach Anspruch 51, **dadurch gekennzeichnet**, daß sie eine Protein-Mischung nach den Ansprüchen 23 bis 25 und eine Nukleinsäuremischung nach den Ansprüchen 19 bis 22 umfaßt.

Fig. 1

V3-Loop Sequenzdaten von HIV-1 Patientenisolaten (PI).

```
CTRPNNNTRKSI.HIGPGRAFYTGDIIIGDIRQAHC  
PI-903      -----G--.-----STN----A-----S-----  
PI-951      ----H-----N-----W-T-----  
PI-918      ----S-----E-----  
PI-970      ----S-----E-----  
PI-990      -----R--.P-----T---V-----  
PI-991      -----P-A-----E--N-----  
PI-952      -I-----R-.TL----VL-T--E-----K---  
PI-932      -I--H-TVTDR-.-----S-HT-RKIK-----  
PI-910      ----SIQK-R-V.R-----S-I-ARAAT'----K-Q-  
PI-911      ----SIQK-R-V.R-----S-I--RAAT----K-Q-  
PI-930      ----YR-AKHR-M-----NVKGNIK:-----
```

Fig. 2a

2/6

Aminosäuresequenzen des NL4-3 V2-Loop und NL4-3 V3-Loop.
Unterstrichen, die Bereiche die bevorzugt variiert werden
sollen.

NL4-3 V2-Loop

GEIKNCFSNISTSIRDKVOKEYAFFYKLDIVPIDNTSYRLISCNTSVITQA

NL4-3 V3-Loop

VQLNTSVEINCTRPNNNTRKSIRIQRGPGRAFVTIGKIGNMRQAHCNISRAKWNATLKQ

Fig. 3

4/6

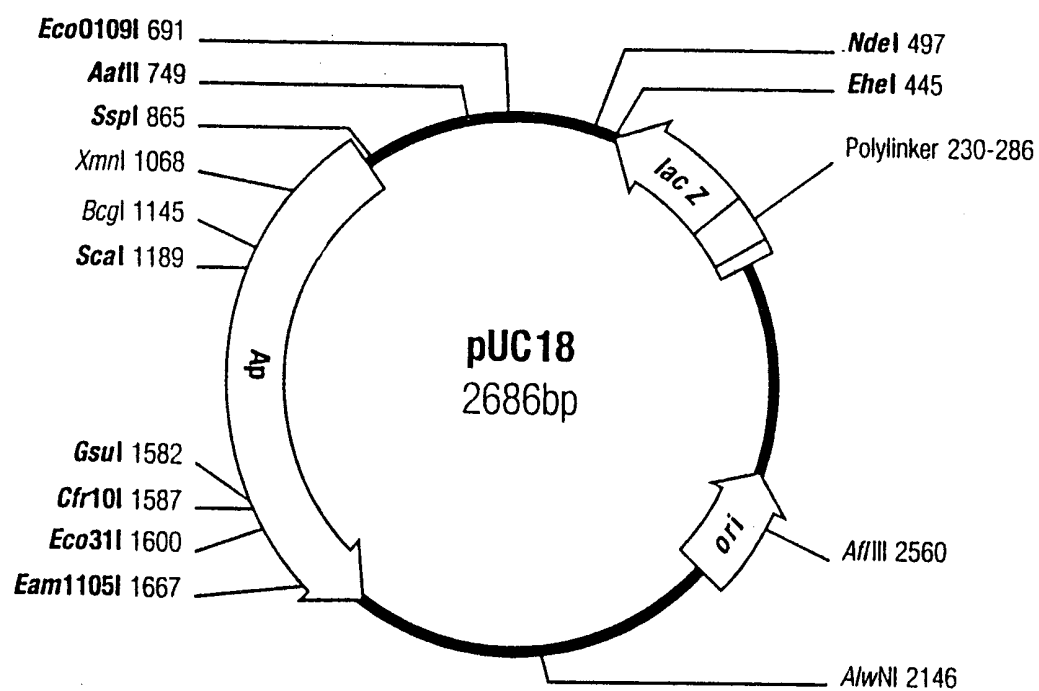


Fig. 4

5/6

Schematische Darstellung des Verfahrens zur Herstellung der Mischung von gp 120 exprimierenden Plasmid-Vektoren:

- 1.) Herstellung der degenerierten DNA-Fragmente für z.B. den V3-Loop
 - a.) Synthese von einzelsträngiger DNA
 - b.) Hybridisierung von zwei komplementären Oligonukleotiden.
- 2.) Klonierung der V3-Loop DNA-Fragmente in pUC18 delta env
- 3.) Klonierung der env Gene in den gp 120 Expressionsvektor pBSCenvATG

V3-Loop-Fragmente mit degenerierter Sequenz z.B.:

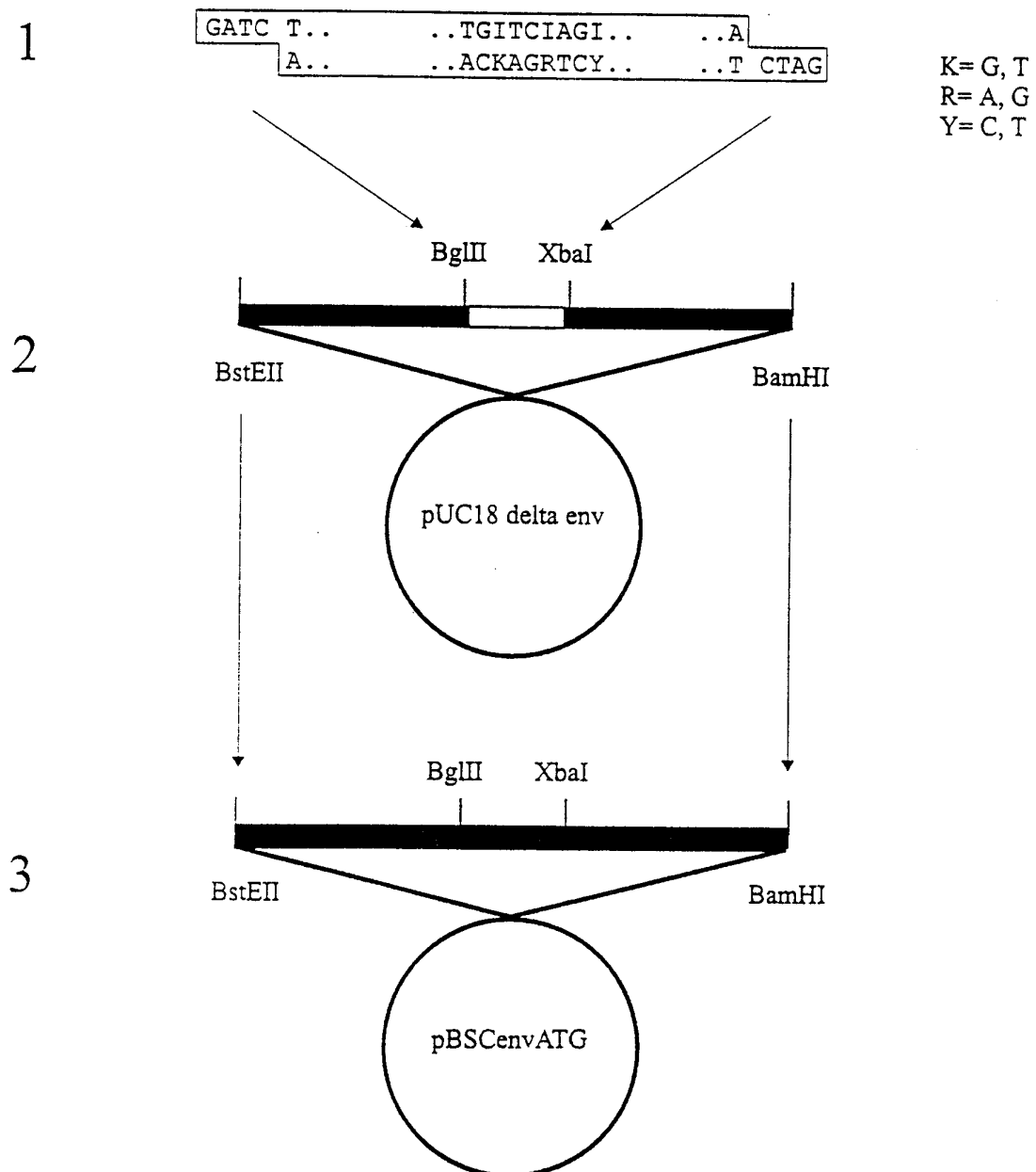


Fig. 5

6/6

Heterogenität der Vakzine am Beispiel des V3-Loop

V3-Loop Sequenz:

```

CTRPNNNTRKSIHIGPGRAFYTTGDIIGDIRQAHC
      G P      HA KK  E
      R      RN
      ET
      E
      A

```

Degeneration auf Protein-Ebene:

$$3 \times 2 \times 2 \times 2 \times 4 \times 6 \times 2 = 1152 \text{ Varianten}$$

```

Ser His Tyr Thr Gly Asp Asp
Gly Pro His Ala Lys Lys Glu
Arg      Arg Asn
      Glu Thr
      Glu
      Ala

```

Degenerierte-DNA-Sequenz der jeweiligen variablen Aminosäure Positionen:

```

AGA CCT TAT ACT GGG GAG GAC
G T  A  C   G  AA  ACC  G = 2048 Varianten

```

```

AGA CCT TAT ACT GGG GAG GAC
AGT CAT CAT GCT GAG GAC GAG
GGT      AGG GCG
GGA      AAG GCC
      AAG
      AAC
      ACG
      ACC

```

Aus der degenerierten DNA-Sequenz abgeleitete Variabilität auf Protein-Ebene:

```

Arg Pro Tyr Thr Gly Glu Asp
Ser His His Ala Glu Asp Glu
Gly      Arg Ala
Gly      Lys Ala
      Lys
      Asn
      Thr
      Thr

```

SEQUENZPROTOKOLL

<110> Strathmann AG & Co.

<120> Virus-Vakzine

<130> P052348

<140>

<141>

<150> 199 07 485.2

<151> 1999-02-12

<160> 12

<170> PatentIn Ver. 2.1

<210> 1

<211> 9709

<212> DNA

<213> Human immunodeficiency virus

<400> 1

```
tggaagggtct aattttggtcc caaaaaagac aagagatcct tgatctgtgg atctaccaca 60
cacaaggcta cttccctgat tggcagaact acacaccagg gccagggatc agatatccac 120
tgaccttttg atggtgcttc aagtttagtac cagttgaacc agagcaagta gaagaggcca 180
aataaggaga gaagaacagc ttgttacacc ctatgagcca gcatgggatg gaggaccgg 240
agggagaagt attagtgtgg aagtttgaca gcctcctagc atttcgtcac atggccccgag 300
agctgcatcc ggagtactac aaagactgct gacatcgagc tttctacaag ggactttccg 360
ctggggactt tccaggggagg tgtggcctgg gcgggactgg ggagtggcga gccctcagat 420
gctacatata agcagctgct ttttgctgt actgggtctc tctggttaga ccagatctga 480
gcctggggagc tctctggcta actagggaac ccactgctta agcctcaata aagcttgcc 540
tgagtgtcta aagtagtgtg tgcccgtctg ttgtgtgact ctggtaacta gagatccctc 600
agaccctttt agtcagtgtg gaaaatctct agcagtggcg cccgaacagg gacttgaaag 660
cgaaaagtaaa gccagaggag atctctcgac gcaggactcg gcttgctgaa gcgcgcacgg 720
caagaggcga ggggcggcga ctgggtgagta cgccaaaaat tttgactagc ggaggctaga 780
aggagagaga tgggtgagag agcgtcggtt ttaagcgggg gagaattaga taaatgggaa 840
aaaattcggt taaggccagg gggaaagaaa caatataaac taaaacatat agtatgggca 900
agcaggggagc tagaacgatt cgcagttaat cctggccttt tagagacatc agaaggctgt 960
agacaaatagc tgggacagct acaaccatcc cttcagacag gatcagaaga acttagatca 1020
ttatataata caatagcagt cctctattgt gtgcatacaa ggatagatgt aaaagacacc 1080
aaggaagcct tagataagat agaggaagag caaaacaaaa gtaagaaaaa ggcacagcaa 1140
gcagcagctg acacaggaaa caacagccag gtcagccaaa attaccctat agtgcagaac 1200
ctccagggggc aaatggtaca tcaggccata tcacctagaa ctttaaatgc atgggtaaaa 1260
gtagtagaag agaaggcttt cagcccagaa gtaataccca tgttttcagc attatcagaa 1320
ggagccaccc cacaagattt aaataccatg ctaaacacag tgggggggaca tcaagcagcc 1380
atgcaaatgt taaaagagac catcaatgag gaagctgcag aatgggatag attgcatcca 1440
gtgcatgcag ggcctattgc accaggccag atgagagaac caaggggaag tgacatagca 1500
ggaactacta gtacccttca ggaacaaata ggatggatga cacataatcc acctatccca 1560
gtaggagaaa tctataaaaag atggataatc ctgggattaa ataaaaatag aagaatgtat 1620
agccctacca gcattctgga cataagacaa ggaccaaaagg aaccctttag agactatgta 1680
gaccgattct ataaaaactc aagagccgag caagcttcac aagaggtaaa aaattggatg 1740
acagaaacct tgttggtcca aaatgcgaac ccagattgta agactatttt aaaagcattg 1800
ggaccaggag cgacactaga agaaatgatg acagcatgtc agggagtggg gggaccggc 1860
cataaagcaa gagttttggc tgaagcaatg agccaagtaa caatccagc taccataatg 1920
atacagaagg gcaatttttag gaaccaaaga aagactgtta agtgtttcaa ttgtggcaaa 1980
gaagggcaca tagccaaaaa ttgcagggcc cctaggaaaa agggctgttg gaaatgtgga 2040
aaggaaggac accaaatgaa agattgtact gagagacagg ctaatttttt aggggaagatc 2100
tggccttccc acaagggaag gccagggaat tttcttcaga gcagaccaga gccacagcc 2160
ccaccagaag agagcttcag gtttggggaa gagacaacaa ctccctctca gaagcaggag 2220
ccgatagaca aggaactgta tccttttagct tccctcagat cactctttgg cagcgacccc 2280
tcgtcacaaat aaagataggg gggcaattaa aggaagctct attagataga ggagcagatg 2340
```

atacagtatt	agaagaaatg	aatttgcag	gaagatggaa	accaaaaatg	ataggggggaa	2400
ttggagggtt	tatcaaagta	ggacagtatg	atcagatact	catagaaatc	tgcggacata	2460
aagctatagg	tacagtatta	gtaggacctt	cacctgtcaa	cataattgga	agaaatctgt	2520
tgactcagat	tggctgcact	ttaaattttc	ccattagtcc	tattgagact	gtaccagtaa	2580
aattaaagcc	aggaatggat	ggcccaaaag	ttaaacaatg	gccattgaca	gaagaaaaaa	2640
taaaagcatt	agtagaaatt	tgtacagaaa	tggaaaaggga	aggaaaaatt	tcaaaaattg	2700
ggcctgaaaa	tccatacaat	actccagtat	ttgccataaa	gaaaaaagac	agtactaaat	2760
ggagaaaaat	agtagatttc	agagaactta	ataagagaac	tcaagatttc	tgggaagttc	2820
aattaggaat	accacatcct	gcagggttaa	aacagaaaaa	atcagtaaca	gtactggatg	2880
tgggcgatgc	atatattttc	gttcccttag	ataaagactt	caggaagtat	actgcattta	2940
ccatacctag	tataaacaat	gagacaccag	ggattagata	tcagtacaat	gtgcttccac	3000
agggatggaa	aggatcacca	gcaatattcc	agtgtagcat	gacaaaaatc	ttagagcctt	3060
ttagaaaaaca	aatccagac	atagtcatct	atcaatacat	ggatgatttg	tatgtaggat	3120
ctgacttaga	aataggcgag	catagaacaa	aaatagagga	actgagacaa	catctgttga	3180
ggtggggggt	taccacacca	gacaaaaaac	atcagaaaaga	acctccattc	ctttggatgg	3240
gttatgaact	ccatcctgat	aaatggacag	tacagcctat	agtgtctgca	gaaaaggaca	3300
gctggactgt	caatgacata	cagaaattag	tgggaaaatt	gaattgggca	agtcagattt	3360
atgcagggat	taaagtaagg	caattatgta	aacttcttag	gggaaccaaa	gcactaacag	3420
aagtagtacc	actaacagaa	gaagcagagc	tagaactggc	agaaaacagg	gagattctaa	3480
aagaaccggg	acatggagtg	tattatgacc	catcaaaaaga	cttaatagca	gaaatacaga	3540
agcaggggca	aggccaatgg	acatatcaaa	tttatcaaga	gccattttaa	aatctgaaaa	3600
caggaaaaata	tgcaagaatg	aagggtgccc	acactaatga	tgtgaaacaa	ttaacagagg	3660
cagtacaaaa	aatagccaca	gaaagcatag	taatatgggg	aaagactcct	aaattttaat	3720
tacctacata	aaaggaaaca	tgggaagcat	ggtggacaga	gtattggcaa	gccacctgga	3780
ttcctgagtg	ggagtttgtc	aatacccttc	ccttagtgaa	gttatgggtac	cagtttagaga	3840
aagaacccat	aataggagca	gaaactttct	atgtagatgg	ggcagccaat	agggaaacta	3900
aattaggaaa	agcaggatat	gtaactgaca	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	gcacacaaag	gaattggagg	aaatgaacaa	gtagatgggt	4200
tggtcagtgc	tggaaatcagg	aaagtactat	ttttagatgg	aatagataag	gccaagaag	4260
aacatgagaa	atatcacagt	aattggagag	caatggctag	tgattttaac	ctaccacctg	4320
tagtagcaaa	agaaatagta	gccagctgtg	ataaatgtca	gctaaaaggg	gaagccatgc	4380
atggacaagt	agactgtagc	ccaggaatat	ggcagctaga	ttgtacacat	ttagaaggaa	4440
aagttatcct	ggtgacagtt	catgtagcca	gtggatataat	agaagcagaa	gtaattccag	4500
cagagacagg	gcaagaaaca	gcatacttcc	tcttaaaaatt	agcaggaaga	tggccagttaa	4560
aaacagtaca	tacagacaat	ggcagcaatt	tcaccagtac	tacagttaag	gccgcctgtt	4620
ggtggggcggg	gatcaagcag	gaatttggca	ttccctacaa	tccccaaagt	caaggagtaa	4680
tagaatctat	gaataaagaa	ttaaagaaaa	ttataggaca	ggtaagagat	caggctgaac	4740
atcttaagac	agcagtacaa	atggcagtat	tcattccaca	ttttaaaaga	aaagggggga	4800
ttgggggggta	catgtcaggg	gaaagaatat	tagacataat	agcaacagac	atacaacta	4860
aagaattaca	aaaacaaatt	acaaaaattc	aaaattttcg	ggttttattac	agggacagca	4920
gagatccagt	ttggaaaagga	ccagcaaagc	tcctctggaa	aggtgaaggg	gcagtagtaa	4980
tacaagataa	tagtgacata	aaagttagtg	caagaagaaa	agcaaagatc	atcagggatt	5040
atggaaaaaca	gatggcaggt	gatgattgtg	tggcaagtag	acaggatgag	gattaacaca	5100
tggaaaagat	tagtaaaaca	ccatatgtat	atltcaagga	aagctaagga	ctggttttat	5160
agacatcact	atgaaagtac	taatccaaaa	ataagttcag	aagtacacat	cccactaggg	5220
gatgctaaat	tagtaataac	aacatatttg	ggtctgcata	caggagaaaag	agactggcat	5280
ttgggtcagg	gagtctccat	agaatggagg	aaaaagagat	atagcacaca	agtagacctt	5340
gacctagcag	accaactaat	tcattctgcac	tatttttgatt	gtttttcaga	atctgctata	5400
agaaatacca	tattaggacg	tatagttagt	cctagggtgtg	aatatcaagc	aggacataac	5460
aaggtaggat	ctctacagta	cttggcacta	gcagcattaa	taaaaccaaa	acagataaag	5520
ccacctttgc	ctagtgttag	gaaactgaca	gaggacagat	ggaacaagcc	ccagaagacc	5580
aaggggccaca	gagggagcca	tacaatgaat	ggacactaga	gcttttagag	gaacttaaga	5640
gtgaagctgt	tagacatttt	cctaggatat	ggctccataa	cttaggacaa	catatctatg	5700
aaacttacgg	ggatacttgg	gcaggagtgg	aagccataat	aagaattctg	caacaactgc	5760
tgtttatcca	tttcagaatt	gggtgtcgag	atagcagaat	aggcggttact	cgacagagga	5820
gagcaagaaa	tggagccagt	agatccctaga	ctagagccct	ggaagcatcc	aggaagtcag	5880
cctaaaactg	cttgtaacca	ttgctattgt	aaaaagtgtt	gctttcattg	ccaagtttgt	5940
ttcatgacaa	aagccttagg	catctcctat	ggcaggaaga	agcggagaca	gcgacgaaga	6000
gctcatcaga	acagtacagc	tcattcaagct	tctctatcaa	agcagtaagt	agtacatgta	6060
atgcaacctt	taatagtagc	aatagtagca	ttagtagtag	caataataat	agcaatagtt	6120
gtgtgggtcca	tagtaatcat	agaatatagg	aaaatattaa	gacaaagaaa	aatagacagg	6180

ttaattgata	gactaataga	aagagcagaa	gacagtggca	atgagagtga	aggagaagta	6240
tcagcacttg	tggagatggg	ggtggaaatg	gggcaccatg	ctccttgga	tattgatgat	6300
ctgtagtgct	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
tagtttaaaag	tgcactgatt	tgaagaatga	tactaatacc	aatagtagta	gcgggagaat	6660
gataatggag	aaaggagaga	taaaaaactg	ctctttcaat	atcagcacia	gcataagaga	6720
taaggtgcag	aaagaatatg	cattctttta	taaacttgat	atagtaccaa	tagataatac	6780
cagctatagg	ttgataagtt	gtaacacctc	agtcattaca	caggcctgtc	caaagggtatc	6840
ctttgagcca	attcccatac	attattgtgc	cccggctggt	tttgcgattc	taaaatgtaa	6900
taataagacg	ttcaatggaa	caggaccatg	tacaaatgtc	agcacagtac	aatgtacaca	6960
tggaaatcagg	ccagtagtat	caactcaact	gctgttaaat	ggcagtcctag	cagaagaaga	7020
tgtagtaatt	agatctgcca	atttcacaga	caatgctaaa	accataatag	tacagctgaa	7080
cacatctgta	gaaattaatt	gtacaagacc	caacaacaat	acaagaaaaa	gtatccgtat	7140
ccagagggga	ccagggagag	catttggttac	aataggaaaa	ataggaaata	tgagacaagc	7200
acattgtaac	tttagtagag	caaaatggaa	tgccacttta	aaacagatag	ctagcaaatt	7260
aagagaacaa	atttgaaata	ataaaaacaat	aatctttaag	caatcctcag	gaggggaccc	7320
agaaattgta	ctgcacagtt	ttaattgtgg	aggggaattt	ttctactgta	attcaacaca	7380
actgtttaat	agtacttggt	ttaatagtac	ttggagtact	gaagggtcaa	ataaactga	7440
aggaagtgc	acaatcacac	tcccatgcag	aataaaacaa	tttataaaca	tgtggcagga	7500
agtaggaaaa	gcaatgtatg	cccctcccat	cagtggacaa	attagatggt	catcaaatat	7560
tactgggctg	ctattaacaa	gagatgggtg	taataacaac	aatgggtccg	agatcttcag	7620
acctggagga	ggcgatatga	gggacaattg	gagaagtga	ttatataaat	ataaagtagt	7680
aaaaattgaa	ccattaggag	tagcaccac	caaggcaaaag	agaagagtgg	tgacagagaga	7740
aaaaagagca	gtgggaatag	gagctttgtt	ccttgggttc	ttgggagcag	caggaagcac	7800
tatgggctgc	acgtcaatga	cgctgacggt	acaggccaga	caattattgt	ctgatatagt	7860
gcagcagcag	aacaatttgc	tgagggtcat	tgaggcgcaa	cagcatctgt	tgcaactcac	7920
agtctggggc	atcaaacagc	tccaggcaag	aatcctggct	gtggaaagat	acctaaagga	7980
tcaacagctc	ctggggattt	ggggttgtct	tggaaaactc	atttgcacca	ctgctgtgcc	8040
ttggaatgct	agttggagta	ataaatctct	ggaacagatt	tggataaaca	tgacctggat	8100
ggagtgggac	agagaaatta	acaattacac	aagcttaata	cactccttaa	ttgaagaatc	8160
gcaaaaccag	caagaaaaga	atgaacaaga	attattggaa	ttagataaat	gggcaagttt	8220
gtggaattgg	tttaacataa	caaattggct	gtggtatata	aaattattca	taatgatagt	8280
aggaggtctg	gtaggtttta	tctgtacttt	tctatagtga	tctatagtga	atagagttag	8340
gcagggatata	tcaccattat	cgtttcagac	ccacctccca	atccccaggg	gacccgacag	8400
gcccgaagga	atagaagaag	aagggtggaga	gagagacaga	gacagatcca	ttcgattagt	8460
gaacggatcc	ttagcactta	tctgggacga	tctgcggagc	ctgtgcctct	tcagctacca	8520
ccgcttgaga	gacttactct	tgattgtaac	gaggattgtg	gaacttcttg	gacgcagggg	8580
gtgggaagcc	ctcaaataat	ggtggaatct	cctacagtat	tggagtcagg	aactaaagaa	8640
tagtgctggt	aacttgctca	atgccacagc	catagcagta	gctgagggga	cagatagggt	8700
tatagaagta	ttacaagcag	cttatagagc	tattcgccac	atacctagaa	gaataagaca	8760
gggcttggaa	aggattttgc	tataagatgg	gtggcaagtg	gtcaaaaagt	agtgtgattg	8820
gatggcctgc	tgtaagggga	agaatgagac	gagctgagcc	agcagcagat	ggggtgggag	8880
cagtatctcg	agacctagaa	aaacatggag	caatcacaag	tagcaataca	gcagctaaca	8940
atgctgcttg	tgccgtgcta	gaagcacaag	aggaggaaga	ggtgggtttt	ccagtcacac	9000
ctcaggtacc	tttaagacca	atgacttaca	aggcagctgt	agatcttagc	cactttttaa	9060
aagaaaaggg	gggactggaa	gggctaattc	actcccaaag	aagacaagat	atccttgatc	9120
tgtggatcta	ccacacacaa	ggctacttcc	ctgattggca	gaactacaca	ccagggccag	9180
gggtcagata	tccactgacc	tttgatgggt	gctacaagct	agtaccagtt	gagccagata	9240
aggtagaaga	ggccaataaa	ggagagaaca	ccagcttggt	acaccctgtg	agcctgcatg	9300
gaatggatga	ccctgagaga	gaagtgttag	agtggaggtt	tgacagccgc	ctagcatttc	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	tgcttgagt	gcttcaagta	gtgtgtgcc	gtctgttggt	tgactctggt	9660
aactagagat	ccctcagacc	cttttagtca	gtgtggaaaa	tctctagca		9709

<210> 2
 <211> 854
 <212> PRT
 <213> Human immunodeficiency virus

<220>
 <223> Envelope Polyprotein

<400> 2

```

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 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

<210> 3
 <211> 107
 <212> DNA
 <213> Kuenstliche Sequenz

<220>
 <223> Beschreibung der kuenstlichen Sequenz:
 Oligonukleotid fuer Klonierung

<400> 3
 aagatgtagt aattagatct gccaatctca cagacaatgc taaaaccata atagtacagc 60
 tgaacacatc gttagaatt aattgtacaa gacccaacaa caataca 107

<210> 4
<211> 120
<212> DNA
<213> Kuenstliche Sequenz

<220>
<223> Beschreibung der kuenstlichen Sequenz:
Oligonukleotid fuer Klonierung

<220>
<221> misc_feature
<222> (97)..(99)
<223> Sequenz an dieser Position: (GA)(AT)(GATC), d.h.
Base an Position 97 kann G oder A sein, Base an
Position 98 kann A oder T sein, und Base an
Position 99 kann G, A, T oder C sein.

<400> 4
ttttgctcta gaaatgttac aatgtgcttg tcttatgtct cctgttcag cttctgttgc 60
atgaaatgct ctccctggtc cgatatggat actatgrwnt tttcttgtat tgttgttggg 120

<210> 5
<211> 17
<212> DNA
<213> Kuenstliche Sequenz

<220>
<223> Beschreibung der kuenstlichen Sequenz:
Sequenzierungsprimer

<400> 5
ccatgtacaa atgtcag

17

<210> 6
<211> 17
<212> DNA
<213> Kuenstliche Sequenz

<220>
<223> Beschreibung der kuenstlichen Sequenz:
Sequenzierungsprimer

<400> 6
aaaactgtgc gttacaa

17

<210> 7
<211> 17
<212> DNA
<213> Kuenstliche Sequenz

<220>
<223> Beschreibung der kuenstlichen Sequenz:
Sequenzierungsprimer

<400> 7
gtaaaacgac ggccagt

17

<210> 8
 <211> 17
 <212> DNA
 <213> Kuenstliche Sequenz

<220>
 <223> Beschreibung der kuenstlichen Sequenz:
 Sequenzierungsprimer

<400> 8
 caggaaacag ctatgac

17

<210> 9
 <211> 2148
 <212> DNA
 <213> Kuenstliche Sequenz

<220>
 <223> Beschreibung der kuenstlichen Sequenz: Synthetische
 DNA

<220>
 <221> misc_feature
 <222> (3)..(9)
 <223> BstEII-Schnittstelle

<220>
 <221> misc_feature
 <222> (2143)..(2148)
 <223> BamHI-Schnittstelle

<400> 9
 tgggtcaccg tctattatgg ggtgcctgtg tgggaaggaag caaccaccac tctattttgt 60
 gcatcagatg ctaaagcata tgatacagag gtacataatg tttggggccac acatgcctgt 120
 gtacccacag accccaaccc acaagaagta gtattggtaa atgtgacaga aaattttaac 180
 atgtggaaaa atgacatggg agaacagatg catgaggata taatcagttt atgggatcaa 240
 agccttaagc catgtgtaaa attaacccca ctctgtgtta gtttaaagtg cactgatttg 300
 aagaatgata ctaataccaa tagtagtagc gggagaatga taatggagaa aggagagata 360
 aaaaactgca gcttcaatat cagcacaagc ataagagata aggtgcagaa agaatatgca 420
 ttcttttata aacttgatat agtaccaata gataatacca gctatagggt gataagttgt 480
 aacacctcag tgatcacaca ggcctgtcca aaggatcctt ttgagccaat tcccatatcat 540
 tattgtgccc cggtgtggtt tgcgattcta aaatgtaata ataagacgtt caatggaaca 600
 ggaccatgta caaatgtcag cacagtacaa tgtacacatg gaattcgacc agtagtatca 660
 actcaactgc tgttaaatgg cagtctagca gaagaagatg tagtaattag atctgccaat 720
 ttcacagaca atgctaaaaa cataatagta cagctgaaca catctgtaga aattaattgt 780
 acaagaccca acaacaatac aagaaaaagt atccgtatcc agagggggacc agggagagca 840
 tttgttacia taggaaaaat aggaaatatg agacaagcac attgtaacat ttctagagca 900
 aaatgggaatg ccactttaaa acagatagct agcaaattaa gagaacaatt tggaaataat 960
 aaaacaataa tctttaagca gtcatccgga ggggacccag aaattgtaac gcacagtttt 1020
 aattgtggag gggaattttt ctactgtaat tcaacacaaac tgtttaatat tacttgggtt 1080
 aatagtactt ggagtactga aggggtcaaat aacactgaag gaagtgcacac aatcacactc 1140
 ccatgcagaa taaaacaatt tataaacatg tggcaggaag taggaaaagc aatgtatgcc 1200
 cctcccatca gtggccaaat tagatgttca tcaaatatta ctgggctgct attactcga 1260
 gatggtggta ataacaacaa tgggtccgag attttcagac ctggaggagg cgatatgagg 1320
 gataattgga gaagtgaatt atataaatat aaagtagtaa aaattgaacc attaggagta 1380
 gcaccacca aggcacaagag acgcgtggtg cacagagaaa agcgcgcagt gggaatagga 1440
 gctctgttcc ttgggttctt gggagcagca ggaagcacta tgggcgcagc gtcaatgacg 1500
 ctgacggtac aggccagaca attattgtct gatatagtgc agcagcagaa caatttgctg 1560
 agggcaattg aggcgcaaca gcatctgttg caactcacag tctggggcat caaacagctc 1620
 caggcaagaa tcctggctgt ggaaagatac ctaaaggatc aacagctcct ggggatttgg 1680
 ggttgctctg gaaaactcat ttgcaccact gctgtgcctt ggaatgctag ttggagtaat 1740
 aaatctcttg aacagatttg gaataacatg acctggatgg agtgggacag agaaattaac 1800
 aattacacaa gcttaataca ctcttaatt gaagaatcgc aaaaccagca agaaaaaat 1860

```

gaacaagaat tattggaatt agataaatgg gcaagtttgt ggaattgggt taacataaca 1920
aattggctgt ggtatataaa attattcata atgatatagtag gaggcttgggt aggtttaaga 1980
atagtttttg ctgtactttc tatagtgaat agagttaggc agggatattc accattatcg 2040
tttcagaccc acctcccaat cccgagggga cccgacaggc ccgaaggaat agaagaagaa 2100
ggtggagaga gagacagaga cagatccatt cgattagtga acggatcc 2148

```

<210> 10

<211> 6229

<212> DNA

<213> Kuenstliche Sequenz

<220>

<223> Beschreibung der kuenstlichen Sequenz: Synthetische DNA

<220>

<221> sig_peptide

<222> (1293)..(1295)

<223> env ATG

<220>

<221> misc feature

<222> (1377)..(1379)

<223> env AGT, gp120 Anfang

<220>

<221> misc feature

<222> (1397)..(1403)

<223> BstEII-Schnittstelle

<220>

<221> misc feature

<222> (3537)..(3542)

<223> BamHI-Schnittstelle

<220>

<221> misc feature

<222> (3855)..(3857)

<223> env TAA, Stop

<400> 10

```

ctgacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg cgacgcgtga 60
ccgctacact tgccagcgcc ctagcgcccc ctcctttcgc tttcttcctc tcctttctcg 120
ccacgttcgc cggctttccc cgtaagctc taaatcgggg gctcccttta gggttccgat 180
ttagtgcttt acggcacctc gaccccaaaa aacttgatta gggatgatgt tcacgtagt 240
ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg ttccttaata 300
gtggactctt gttccaaact ggaacaacac tcaaccctat ctcggtctat tcttttgatt 360
tataagggat tttgccgatt tcggcctatt ggttaaaaaa tgagctgatt taacaaaaaat 420
ttaacgcgaa ttttaacaaa atattaacgc ttacaatttc cattcgccat tcaggctgcg 480
caactgttgg gaaggcgcat cggcgcgggc ctcttcgcta ttacgccagc tggcgaaagg 540
gggatgtgct gcaaggcgat taagtgggt aacgccaggg ttttccagc cagcaggttg 600
taaaacgacg gccagtgcgc gtctagtatt taatagtaat caattacggg gtcattagtt 660
catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga 720
ccgccccacg acccccgcgc attgacgtca ataatgacgt atgttcccat agtaacgcca 780
atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaaactgc ccacttggca 840
gtacatcaag tgtatcatat gccaaagtac ccccttattg acgtcaatga cggtaaatgg 900
ccgccttggc attatgccca gtacatgacc ttatgggact ttcctacttg gcagtacatc 960
tacgtattag tcatcgctat taccatgggt atgcggttt ggcagtacat caatgggcgt 1020
ggatagcggg ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt 1080
ttgttttggc accaaaaatca acgggacttt ccaaaatgtc gtaacaactc cgccccattg 1140
acgcaaattg gcggtaggcg tgtacgggtg gaggtctata taagcagagc tcgttttagtg 1200
aaccgtcaga tcgcctggag acgccatcca cgctgttttg acctccatag aagacaccgg 1260
gacaattcga gctcgggtacc gtcgacgcca ccatgagagt gaaggagaag tatcagcact 1320

```

tgtggagatg	gggggtggaaa	tggggcacca	tgctccttgg	gatattgatg	atctgtagtg	1380
ctacagaaaa	attgtgggtc	accgtctatt	atgggggtacc	tgtgtggaag	gaagcaacca	1440
ccactctatt	ttgtgcatca	gatgctaaag	catatgatac	agaggtagat	aatgttttggg	1500
ccacacatgc	ctgtgtaccc	acagacccca	acccacaaga	agtagtattg	gtaaatgtga	1560
cagaaaaattt	taacatgtgg	aaaaatgaca	tggtagaaca	gatgcatgag	gatataatca	1620
gtttatggga	tcaaagccta	aagccatgtg	taaaattaac	cccactctgt	gttagtttaa	1680
agtgcactga	tttgaagaat	gatactaata	ccaatagtag	tagcgggaga	atgataatgg	1740
agaaaaggaga	gataaaaaac	tgctccttca	atatcagcac	aagcataaga	gataaggtgc	1800
agaaagaata	tgcatctctt	tataaacctg	atatagtacc	aatagataat	accagctata	1860
ggttgataag	ttgtaacacc	tcagtcatta	cacaggcctg	tccaaaggta	tcctttgagc	1920
caattcccat	acattattgt	gccccggctg	gtttttgcgat	tctaaaatgt	aataataaga	1980
cgttcaatgg	aacaggacca	tgtacaaatg	tcagcacagt	acaatgtaca	catggaaatca	2040
ggccagtagt	atcaactcaa	ctgctgttaa	atggcagctc	agcagaagaa	gatgtagtaa	2100
ttagatctgc	caatttcaca	gacaatgcta	aaaccataat	agtacagctg	aacacatctg	2160
tagaaattaa	ttgtacaaga	cccaacaaca	atacaagaaa	aagtatccgt	atccagaggg	2220
gaccagggag	agcattttgt	acaataggaa	aaataggaaa	tatgagacaa	gcacattgta	2280
acattagtag	agcaaaaatg	aatgccactt	taaaacagat	agctagcaaa	ttaagagaac	2340
aattttgaaa	taataaaaaca	ataatcttta	agcaatcctc	aggaggggac	ccagaaattg	2400
taacgcacag	ttttaattgt	ggaggggaat	ttttctactg	taattcaaca	caactgttta	2460
atagtacttg	gttttaatagt	acttggagta	ctgaagggtc	aaataacact	gaagggaagt	2520
acacaatcac	actcccatgc	agaataaaac	aattttataaa	catgtggcag	gaagtaggaa	2580
aagcaatgta	tgccccctcc	atcagtgagc	aaattagatg	ttcatcaaat	attactgggc	2640
tgctattaac	aagagatggg	ggtaataaca	acaatgggtc	cgagatcttc	agacctggag	2700
gaggcgatat	gagggacaat	tgagagaagt	aattatataa	atataaagta	gtaaaaattg	2760
aaccattagg	agtagcaccc	accaaggcaa	agtagaagag	ggtgcagaga	gaaaaaagag	2820
cagtgggaat	aggagctttg	ttccttgggt	tcttgggagc	agcagggaagc	actatgggct	2880
gcacgtcaat	gacgctgacg	gtacaggcca	gacaattatt	gtctgatata	gtgcagcagc	2940
agaacaattt	gctgagggct	attgagggcg	aacagcatct	gttgcaactc	acagtctggg	3000
gcatacaaca	gctccaggga	agaatccttg	ctgtggaaaag	atacctaaag	gatcaacagc	3060
tcctggggat	ttgggggtgc	tctggaaaac	tcatttgcac	cactgctgtg	ccttgggaatg	3120
ctagttaggag	taataaatct	ctggaaacaga	tttggaaata	catgacctgg	atggagtggg	3180
acagagaaat	taacaattac	acaagcttaa	tacactcctt	aattgaagaa	tcgcaaaacc	3240
agcaagaaaa	gaatgaacaa	gaattatttg	aattagataa	atgggcaagt	ttgtggaatt	3300
ggtttaacat	aacaaattgg	ctgtgggtata	taaaattatt	cataatgata	gtaggaggct	3360
tggtagggtt	aagaatagtt	tttgctgtac	tttctatagt	gaatagagtt	aggcagggat	3420
attcaccatt	atcgtttcag	accacctcc	caatcccag	gggacccgac	aggcccgaag	3480
gaatagaaga	agaagggtgga	gagagagaca	gagacagatc	cattcgatta	gtgaacggat	3540
ccttagcact	tatctgggac	gatctgcgga	gcctgtgcct	cttcagctac	caccgcttga	3600
gagacttact	cttgatttga	acgaggattg	tggaaacttct	gggacgcagg	gggtgggaag	3660
ccctcaaata	ttggtggaat	ctcctacagt	attggagtca	ggaactaaag	aatagtgtctg	3720
ttaaacttgct	tgactgccaca	gccatagcag	tagctgaggg	gacagatagg	gttatagaag	3780
tattacaagc	agcttataga	gctattcgcc	acatacctag	aagaataaga	caggccttgg	3840
aaaggatttt	gctataagat	gggtggcaag	tggtcaaaaa	gtagtgtgat	tggatggcct	3900
gctgtaaggg	aaagaatgag	acgagctgag	ccagcagcag	atgggggtggg	agcagtatct	3960
cgagatctag	actagaacta	gcttcgatcc	agacatgata	agatacattg	atgagtttgg	4020
acaaaccaca	actagaatgc	agtgaaaaaa	atgctttatt	tgtgaaattt	gtgatgctat	4080
tgctttattt	gtaaccatta	taagctgcaa	taaaacaagt	aacaacaaca	attgcattca	4140
ttttatgttt	cagggttcagg	gggaggtgtg	ggaggttttt	taaagcaagt	aaaacctcta	4200
caaattgtgg	atggctgatt	atgatcctgc	ctcgcgcgtt	tcggtgatga	cggtgaaaac	4260
ctctgacaca	tgcagctccc	ggagacgggc	acagcttgct	tgtaagcgga	tgccggggagc	4320
agacaagccc	gtcagggcgc	gtcagcgggt	gttggcgggt	gtcggggcgc	agccatgacc	4380
cagtcacgta	gcgatagcgg	agtgtatact	ggcttaacta	tgccggcatca	gagcagattg	4440
tactgagagt	gcaccatatg	tcgggcccgc	ttgctggcgt	ttttccatag	gctccgcccc	4500
cctgacgagc	atcacaaaaa	tcgacgctca	agtcagaggt	ggcgaacccc	gacaggacta	4560
taaagatacc	aggcggtttc	ccctgggaagc	tccctcgtgc	gctctcctgt	tcggaccctg	4620
ccgcttaccg	gatacctgtc	cgcctttctc	ccttcgggaa	gcgtggcgct	ttctcatagc	4680
tcacgctgta	ggtatctcag	ttcgggtgtg	ttcggttcgt	ccaagctggg	ctgtgtgcac	4740
gaaccccccg	ttcagcccg	ccgctgcgcc	ttatccggta	actatcgtct	tgagtcacac	4800
ccggtaagac	acgacttatc	gccactggca	gcagccactg	gtaacaggat	tagcagagcg	4860
agggtatgtag	gcggtgtctac	agagttcttg	aagtgggtggc	ctaactacgg	ctacactaga	4920
aggacagtat	ttggtatctg	cgctctgctg	aagccagtta	ccttcggaaa	aagagttggg	4980
agctcttgat	ccggcaaaaa	aaccaccgct	ggtagcgggt	gtttttttgt	ttgcaagcag	5040
cagattacgc	gcagaaaaaa	aggatctcaa	gaagatcctt	tgatcttttc	tacggggctc	5100
gacgctcagt	ggaacgaaaa	ctcacgttaa	gggattttgg	tcatgagatt	atcaaaaagg	5160

```

atcttcacct agatccctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat 5220
gagtaaacctt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc 5280
tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac tacgatacgg 5340
gagggccttac catctggccc cagtgtctgca atgataccgc gagaccacg ctcaccggct 5400
ccagattttat cagcaataaaa ccagccagcc ggaagggccg agcgcagaag tggtcctgca 5460
actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg 5520
ccagttaata gtttgcgcaa cgttggtgcc attgctacag gcatcgtggt gtcacgctcg 5580
tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt tacatgatcc 5640
cccatgttgt gcaaaaaaagc ggtagctcc ttcggtcctc cgatcgttgt cagaagtaag 5700
ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg 5760
ccatccgtaa gatgcttttc tgtgactggt gactactcaa ccaagtcatt ctgagaatag 5820
tgtatgcggc gaccgagttg ctcttgccc gcgtaatac gggataatac cgcgccacat 5880
agcagaactt taaaagtgtc catcattgga aaacgttctt cggggcgaaa actctcaagg 5940
atcttaccgc ttttgagatc cagttcgatg taaccactc gtgcaccaa ctgatcttca 6000
gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgccgca 6060
aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct ttttcaatat 6120
tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 6180
aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccac 6229

```

<210> 11
 <211> 860
 <212> DNA
 <213> Human immunodeficiency virus

<220>
 <221> misc_feature
 <222> (1)..(860)
 <223> PI-932 Originalsequenz V1-V2-V3-Loop

```

<400> 11
tgtgtaccca cagaccccaa cccacaaaag gtagtattgg aaaatgtgac agaaaatttt 60
aacatgtgga aaaatgacat ggtagaacag atgcatgagg atataatcaa tttatgggat 120
caaagcctaa agccatgtgt aaaactaacc ccactctgtg ttacttttaa ttgcaactgat 180
gctgatttaa attgcaataa tactgattta aattgcacta aagctaattt ggggaaaaaat 240
actcataaca atactatttag tgggaaaata atagagaaa tagaaataaa aaactgctct 300
ttcaagggtca ccacaggcat aagggaata atgcaaaaa aatatgcact tttgaataaa 360
cttgatatag taccaataga taatgataag aataatacta actttatatt gataagttgt 420
aacacctcga ccattacaca ggcctgtcca aaggatcctt ttgagccaat tcccatacat 480
ttttgtgcc cggctgggtt tgcgattcta aagtgtaatg aaaagagtta cagtggaaaa 540
ggaccatgta aaaatgtcag cacagtacaa ttagacatg gaattaggcc agtagtgtca 600
actcaactgc tgttgaaatg cagtctagca gaaaaagaag tagtaattag atctgagaat 660
ttcacagaca atgctaaaac cataatagta cagctgaagg aatctgtaaa cattacttgt 720
ataagacccc acaacactgt aacagacagg atacatatag ggccagggag atcatttcat 780
acaacaagaa aaataaaaag agatataaga caagcacatt gtagccttag gagaaaagat 840
tggaataaca ctttacaaga
860

```

<210> 12
 <211> 870
 <212> DNA
 <213> Kuenstliche Sequenz

<220>
 <223> Beschreibung der kuenstlichen Sequenz: PI-932
 Genkassette, beinhaltet die Schnittstellen fuer
 die Restriktionsenzyme BspT1, PstI, BclI, EcoRI,
 BglIII, PvuII, XbaI, NheI

```

<400> 12
tgtgtaccca cagaccccaa cccacaaaag gtagtattgg aaaatgtgac agaaaatttt 60
aacatgtgga aaaatgacat ggtagaacag atgcatgagg atataatcaa tttatgggat 120
caaagcctta agccatgtgt aaaactaacc ccactctgtg ttacttttaa ttgcaactgat 180
gctgatttaa attgcaataa tactgattta aattgcacta aagctaattt ggggaaaaaat 240

```

```
actcataact gcagtattag tgggaaaata atagagaaag tagaaataaa aaactgctct 300
ttcaaggtca ccacaggcat aagggataag atgcaaaaag aatatgcact tttgaataaa 360
cttgatatag taccaataga taatgataag aataatacta actttatatt gataagttgt 420
aacacctcgg tgatcacaca ggcctgtcca aaggatcct ttgagccaat tcccatacat 480
ttttgtgccc cggctgggtt tgcgattcta aagtgtaatg aaaagagtta cagtggaaaa 540
ggaccatgta aaaatgtcag cacagtacaa tgtacacatg gaattcggcc agtagtgtca 600
actcaactgc tgttgaatgg cagtctagca gaaaaagaag tagtaattag atctgagaat 660
ttcacagaca atgctaaaac cataatagta cagctgaagg aatctgtaaa cattacttgt 720
ataagacccc acaacactgt aacagacagg atacatatag ggccaggagg atcatttcat 780
acaacaagaa aaataaaaagg agatataaga caagcacatt gtagcctttc tagaaaagat 840
tggaataaca ctttacaaga gatagctagc                                     870
```