



(12) PATENT

(19) NO

(11) 331826

(13) B1

NORGE

(51) Int Cl.

C12N 15/48 (2006.01)

A61P 31/18 (2006.01)

C07K 14/16 (2006.01)

C12N 15/09 (2006.01)

A61K 48/00 (2006.01)

C12N 15/49 (2006.01)

Patentstyret

(21)	Søknadsnr	20041157	(86)	Int.inng.dag og søknadsnr	2002.09.18 PCT/EP2002/10592
(22)	Inng.dag	2004.03.19	(85)	Videreføringsdag	2004.03.19
(24)	Løpedag	2002.09.18	(30)	Prioritet	2001.12.11, GB, 0129604 2002.03.19, GB, 0206462 2001.09.20, WO, PCT/GB01/04207
(41)	Alm.tilgj	2004.05.19			
(45)	Meddelt	2012.04.16			
(73)	Innehaver	Glaxo Group Limited, Glaxo Wellcome House, Berkeley Avenue, GB-UB60NN GREENFORD, MIDDLESEX, Storbritannia			
(72)	Oppfinner	Peter Franz Ertl, c/o GlaxoSmithKline, Gunnels Wood Road, GB-SG12NY STEVENAGE, HERTFORDSHIRE , Storbritannia Gerald Wayne Gough, c/o GlaxoSmithKline, Gunnels Wood Road, GB-SG12NY STEVENAGE, HERTFORDSHIRE , Storbritannia John Philip Tite, c/o GlaxoSmithKline, Gunnels Wood Road, GB-SG12NY STEVENAGE, HERTFORDSHIRE , Storbritannia Andrew Beaton, c/o GlaxoSmithKline, Gunnels Wood Road, GB-SG12NY STEVENAGE, HERTFORDSHIRE, Storbritannia Andrew Lear, c/o GlaxoSmithKline, Gunnels Wood Road, GB-SG12NY STEVENAGE, HERTFORDSHIRE , Storbritannia			
(74)	Fullmektig	Bryn Aarflot AS, Postboks 449 Sentrum, 0104 OSLO, Norge			
(54)	Benevnelse	Nukleotidsekvens, vektor, protein, farmasøytisk preparat, intradermal avleveringsanordning, samt fremgangsmåte for fremstilling av en nukleotidsekvens.			
(56)	Anførte publikasjoner	Kotsopoulou E. et al. A rev-independent human immunodeficiency virus type 1 (HIV-1)-based vector that exploits a codon-optimized HIV-1 gag-pol gene. Journal of virology. 2000, Vol. 74, No. 10, side 4839-4852., WO 0039302 A			
(57)	Sammendrag				

Det beskrives en nukleotidsekvens som koder for et HIV-1 gag-protein eller fragment derav inneholdende en gag-epitop og et andre HIV-antigen eller et fragment som koder for en epitop av nevnte andre HIV-antigen, operativt koblet til en heterolog promoter. Foretrukne polynukleotidsekvenser koder dessuten for nef eller et fragment derav, og RT eller et fragment derav.

Foreliggende oppfinnelse vedrører en nukleotidsekvens, en vektor som omfatter nevnte nukleotidsekvens, et protein som kodes for av nevnte nukleotidsekvens, et farmasøytisk preparat som omfatter nevnte nukleotidsekvens, en intradermal avleveringsanordning som omfatter nevnte farmasøytiske preparat, samt en fremgangsmåte for fremstilling av en nukleotidsekvens ifølge oppfinnelsen.

Vaksineformuleringer som omfatter slike nukleinsyrekonstruksjoner og anvendelsen av slike formuleringer innen medisinen er beskrevet, særlig DNA-vaksiner som er anvendelige ved forebyggelse og behandling av HIV-infeksjoner, nærmere bestemt når slike administreres ved partikkel-mediert avlevering.

HIV-1 er hovedårsaken til AIDS (tilegnet immunsviktsyndrom) som ansees som ett av verdens hovedhelseproblem. Selv om det har vært gjennomført omfattende forskning over hele verden for å frembringe en vaksine, har slike anstrengelser hittil ikke vært vellykkede.

Ikke-kappe-proteiner av HIV-1 har vært beskrevet og inkluderer for eksempel interne strukturproteiner, så som produktene av *gag*- og *pol*-genene og andre ikke-strukturelle proteiner, så som Rev, Nef, Vif og Tat (Green *et al.*, New England J. Med. **324**, 5, 308 *et seq* (1991) og Bryant *et al.*, (red. Pizzo), *Pediatr. Infect. Dis. J.*, **11**, 5, 390 *et seq* (1992).

Gag-genet translateres fra full-lengde RNA under dannelse av et forløper-polyprotein som deretter spaltes i 3-5 kapsidproteiner; matriksproteinet, kapsidproteinet og nukleinsyre-bindingsprotein og protease. (1. Fundamental Virology, Fields BN, Knipe DM & Howley M, 1996 2. Fields Virology Bd. 2, 1996).

Kotsopoulou E. Et al. A rev-independent human immunodeficiency virus type 1 (HIV-1)-based vector that exploits a codon-optimized HIV-1 gag-pl gene. *Journal of virology*. 2000, Vol. 74, No. 10, page 4839-4852 lærer kodon-optimaliserte sekvenser for HIV polyprotein (gag-pol).

WO 0039302 A (CHIRON CORP)2000.07.06 lærer forskjellige gag-protease fusjoner og muterte env eller tat gener.

Genet *gag* forårsaker det 55-kilodalton (kD) Gag-forløperprotein, også kalt p55, som uttrykkes fra den uspleisede virale mRNA. Under translasjon myristoyleres N-terminalen av p55, som utløser dets assosiasjon med det cytoplasmatiske aspekt ved cellemembraner. Det membran-assosierte Gag-polyprotein rekrutterer to kopier av den virale genomiske RNA sammen med andre virale og cellulære proteiner som

utløser avsnøringen av viruspartikkelen fra overflaten av en infisert celle. Etter avsnøring spaltes p55 av den viralt kodede protease (et produkt av pol-genet) under den virale modningsprosess, i fire mindre proteiner betegnet MA (matriks [p17], CA (kapsid [p24], NC (nukleokapsid [p9] og p6. (4).

5 I tillegg til de tre Gag-hovedproteinene, inneholder alle Gag-forløpere diverse andre regioner som spaltes av, og forblir i, virionet som peptider av ulike størrelser. Disse proteinene har forskjellige roller, f.eks. har p2-proteinet vært angitt å ha en rolle innen den regulerende aktivitet av proteasen og til å bidra til den riktige tids-tilpasning av proteolytisk prosessering.

10 MA-polypeptidet avledes fra den N-terminale myristoylerte ende av p55. De fleste MA-molekyler forblir forbundet med den indre overflate av virionets lipid-dobbeltlag, hvilket stabiliserer partikkelen. Et subsett av MA rekrutteres inne i de dypere lag av virionet, hvor det blir del av det kompleks som ledsager den virale DNA til kjernen (5). Disse MA-molekylene letter nukleærtransporten av det virale genom
15 fordi et karyofilt signal på MA gjenkjennes av det cellulære nukleære importmaskineri. Dette fenomen lar HIV infisere ikke-delende celler, en uvanlig egenskap ved et retrovirus.

Proteinet p24 (CA) danner den koniske kjerne av viruspartikler. Cyklofilin A har vært vist å interagere med p24-regionen av p55, hvilket fører til dets

20 inkorporering i HIV-partikler. Interaksjonen mellom Gag og cyklofilin A er essensiell fordi forstyrrelse av denne interaksjon med cyklosporin A inhiberer viral replikasjon.

NC-regionen av Gag er ansvarlig for spesifikk gjenkjenning av det såkalte HIV-pakkingssignal. Pakkingssignalet består av fire stem loop strukturer lokalisert nær 5'-enden av den virale RNA, og er tilstrekkelig til å mediere inkorporeringen av en
25 heterolog RNA i HIV-1 virioner. NC binder til pakkingssignalet gjennom interaksjoner mediert av to Zinkfinger-motiver. NC letter også revers transkripsjon.

Polypeptidregionen p6 medierer interaksjoner mellom p55 Gag og hjelpe-proteinet Vpr, som fører til inkorporeringen av Vpr i sammensettende virioner. Regionen p6 inneholder også et såkalt sent domene som fordres for den effektive
30 frigjøring av avsnørende virioner fra en infisert celle.

Pol-genet koder for to proteiner som inneholder de to aktiviteter som viruset trenger ved tidlig infeksjon, nemlig RT og integraseproteinet som trengs for integrering av viral DNA i celle-DNA. Primærproduktet av Pol avspaltes av virion-proteasen for å gi det aminoternale RT-peptid som inneholder nødvendige

aktiviteter for DNA-syntese (RNA- og DNA-rettet DNA-polymerase, ribonuklease H) og karboksyterminalt integraseprotein.

HIV RT er en heterodimer av full-lengde RT (p66) og et spaltningsprodukt (p51) som mangler det karboksyterminale Rnase-integrase-domenet.

5 RT er en av de sterkest konserverte proteiner som kodes av det retrovirale genom. To hovedaktiviteter av RT er DNA Pol og ribonuklease H. DNA Pol-aktiviteten av RT benytter RNA og DNA som templater om hverandre, og er som alle kjente DNA-polymeraser ikke i stand til å initiere DNA-syntese *de novo*, men fordrer et på forhånd eksisterende molekyl for å tjene som en primer (RNA).

10 Den iboende Rnase H aktivitet i alle RT-proteiner spiller den vesentlige rolle tidlig i replikasjonen ved å fjerne RNA-genomet etter som DNA-syntesen skrider frem. Den nedbryter selektivt RNA fra alle RNA-DNA-hybridmolekyler. Strukturelt opptar polymerasen og ribo H separate, ikke-overlappende domener hvor Pol dekker to tredjedeler av aminoenden av Pol.

15 Den p66-katalytiske subenhet foldes i 5 distinkte subdomener. De 23 amino-terminale av disse har delen med RT-aktivitet. Karboksyenden av disse er Rnase H domenet.

Etter infeksjon av vertscellen kopieres det retrovirale RNA-genom inn i lineær ds DNA med den reverse transkriptase som er tilstede i den infiserende partikkel.

20 Integrasen (omtalt i Skalka AM '99 Adv in Virus Res. **52** 271-273) gjenkjenner endene av den virale DNA, trimmer disse og ledsager den virale DNA til et vertskromosomalt sete for å katalysere integrasjon. Mange seter i vertens DNA kan målsøkes for integrasjon. Selv om integrasen er tilstrekkelig til å katalysere integrasjon *in vitro*, er det ikke det eneste protein assosiert med den virale DNA *in vivo*. Det store protein, viralt DNA-kompleks, isolert fra de infiserte cellene har vært betegnet pre-integrasjonskomplekset. Dette letter tilegnelsen av vertscellegenene med avkommets virale genomer.

30 Integrasen dannes av 3 distinkte domener, det N-terminale domene, den katalytiske kjerne og det C-terminale domene. Det katalytiske kjernedomenet inneholder alt som fordres for polynukleotidyl-overføringens kjemi.

Nef-proteinet fjerner som kjent CD4, HIV-reseptoren, fra celleoverflaten, men den biologiske betydning av denne funksjon er under diskusjon. Dessuten interagerer Nef med signalveien til T-celler og induserer en aktiv tilstand, som i sin tur kan fremme mer effektiv genekspresjon. Enkelte HIV-isolater har mutasjoner i dette

området som bevirker at de ikke koder for funksjonelt protein og er alvorlig hemmet i deres replikasjon og patogenese *in vivo*.

DNA-vaksiner består vanligvis av en bakteriell plasmidvektor som det er innført en sterk promoter i, genet av interesse som koder for et antigen peptid, og polyadenylerings/transkripsjonelle termineringssekvenser. Genet av interesse kan kode for et fullt protein eller ganske enkelt en antigen peptidsekvens relatert til patogenet, tumoren eller annet agens som det er ment å beskyttes mot. Plasmidet kan dyrkes i bakterier som for eksempel *E. coli*, og deretter isoleres og tilberedes i et passende medium, avhengig av den tiltenkte administrasjonsmåte, før det administreres til verten. Etter administrering tas plasmidet opp av vertens celler, hvor det kodede peptid produseres. Plasmidvektoren vil fortrinnsvis være fremstillet uten et replikasjonsorigo som er funksjonelt i eukaryote celler, for å forhindre plasmid-replikasjon i pattedyrverten og integrasjon i kromosomalt DNA til vedkommende dyr.

Det finnes en rekke fordeler ved DNA-vaksinasjon i forhold til tradisjonelle vaksinasjonsteknikker. For det første forventes det siden de proteiner som kodes av DNA-sekvensen syntetiseres i verten, at strukturen eller konformasjonen av proteinet vil tilsvare det naturlige protein som er assosiert med sykdomstilstanden. Det er også sannsynlig at DNA-vaksinasjon vil gi beskyttelse mot forskjellige stammer av et virus ved å frembringe cytotoksiske T-lymfocytteresponser som gjenkjenner epitoper fra konserverte proteiner. Siden plasmidene tas opp av de vertsceller hvor det kan produseres antigen protein, vil det dessuten utløses en langvarig immunrespons. Teknologien byr også på den mulighet å kombinere diverse immunogener i et enkelt preparat for å lette samtidig immunisering med relasjon til en rekke sykdomstilstander.

Nyttig bakgrunnsinformasjon med relasjon til DNA-vaksinasjon er gitt i Donnelly *et al.* "DNA vaccines" Ann. Rev. Immunol. 1997 15:617-648.

Foreliggende oppfinnelse tilveiebringer en nukleotidsekvens som omfatter en sekvens som koder for et immunogenisk fusjonsprotein som omfatter et kodon optimalisert HIV-1 gag-protein, et kodon optimalisert RT protein og et HIV-1 Nef protein, operativt koblet til en heterolog promoter, hvori gag proteinet omfatter p17 og p24, RT proteinet inneholder en mutasjon for i alt vesentlig å inaktivere en hvilken som helst revers transkriptase aktivitet, og Nef proteinet inneholder en N-terminal trunkering.

Oppfinnelsen tilveiebringer også en vektor som omfatter en nukleotidsekvens ifølge oppfinnelsen.

Oppfinnelsen tilveiebringer også et protein som kodes for av en nukleotidsekvens ifølge oppfinnelsen.

5 Oppfinnelsen tilveiebringer også et farmasøytisk preparat som omfatter en nukleotidsekvens eller en vektor ifølge oppfinnelsen og et farmasøytisk akseptabelt hjelpestoff, fortynningsmiddel, bærer eller adjuvans.

Oppfinnelsen tilveiebringer også en intradermal avleveringsanordning, som omfatter et farmasøytisk preparat ifølge oppfinnelsen.

10 Endelig tilveiebringer oppfinnelsen en fremgangsmåte for fremstilling av en nukleotidsekvens ifølge oppfinnelsen som omfatter operativ kobling av en nukleotidsekvens som koder for HIV-1 gag-protein eller fragment derav og et HIV-1 Nef-protein eller fragment derav, til en heterolog promotersekvens.

Det er beskrevet et nukleinsyremolekyl som omfatter en nukleotidsekvens
15 som koder for HIV gag-protein, eller fragment derav, koblet til en nukleotidsekvens som koder for et ytterligere HIV-antigen, eller fragment derav, og som operativt er koblet til en heterolog promoter. Fragmentet av nevnte nukleotidsekvens vil kode for en HIV-epitop og koder typisk for et peptid på minst 8 aminosyrer. Nukleotid-
sekvensen er fortrinnsvis en DNA-sekvens og inngår fortrinnsvis i et plasmid uten et
20 replikasjonsorigo. Slike nukleinsyremolekyler formuleres med farmasøytisk
akseptable hjelpestoffer, bærere, fortynningsmidler eller adjuvanser for fremstilling
av en farmasøytisk blanding egnet for behandlingen og/eller forebyggelsen av HIV-
infeksjon og AIDS.

Foretrukket formuleres DNA-sekvensen på overflaten av inerte partikler eller
25 kuler egnet for partikkelmediert medikamentavlevering. Fortrinnsvis er kulene av gull.

Det tilveiebringes en DNA-sekvens som høyt uttrykt koder for gag-protein, en sekvens som optimaliseres for å ligne kodonanvendelsen av gener i pattedyrceller. Særlig optimaliseres gag-proteinet for å gi likhet med det til høyt uttrykte humane gener.

30 DNA-koden har 4 bokstaver (A, T, C og G) og benytter disse til å stave tre-bokstavs "kodoner" som representerer aminosyrene i proteiner som kodes av en organismes gener. Den lineære sekvens av kodoner langs DNA-molekylet translateres til den lineære sekvens av aminosyrer i de proteiner som kodes av disse genene. Koden er sterkt degenerert, med 61 kodoner som koder for de 20 naturlige

forekommende aminosyrer og 3 kodoner som representerer "stop"-signaler. De fleste aminosyrer kodes således av mer en ett kodon – faktisk kodes flere av fire eller flere forskjellige kodoner.

Når mer enn ett kodon er tilgjengelig til koding av en gitt aminosyre, har det vært observert at organismers mønster for kodonanvendelse i høy grad ikke er vilkårlig. Ulike arter oppviser en forskjellig forfordeling i deres kodonvalg, og kodonanvendelsen kan i en enkelt art være klart forskjellig mellom gener som uttrykkes i høye og lave nivåer. Denne forfordeling er forskjellig i virus-, plante-, bakterie- og pattedyrceller, og enkelte arter oppviser en sterkere forfordeling bort fra tilfeldig kodonvalg enn andre. For eksempel skjer mindre sterk forfordeling hos mennesker og andre pattedyr enn hos enkelte bakterier eller virus. Det er av disse grunner en signifikant sannsynlighet for at et pattedyrgen uttrykt i *E.coli* eller et fremmed eller rekombinant gen uttrykt i pattedyrceller, vil ha en upassende fordeling av kodoner for effektiv ekspresjon. I en heterolog DNA-sekvens antas det at nærværet av ansamlinger (clusters) av kodoner eller en rikelighet av kodoner som sjeldent observeres i den vert hvor ekspresjon skal skje, er prediktiv for lave heterologe ekspresjonsnivåer i vedkommende vert.

Det er beskrevet en gag polynukleotidsekvens som koder for en aminosyresekvens, hvor kodonanvendelsesmønsteret for polynukleotidsekvensen ligner det for høyt uttrykte pattedyrgener. Fortrinnsvis er polynukleotidsekvensen en DNA-sekvens. Det er ønskelig at kodonanvendelsesmønsteret av polynukleotidsekvensen typisk er av høyt uttrykte humangener.

I polynukleotidene ifølge foreliggende oppfinnelse er kodonanvendelsesmønsteret endret fra det typiske for humane immunsviktvirus til nærmere å representere kodon-forfordelingen av målorganismen, f.eks. et pattedyr, spesielt et menneske. "Kodonanvendelseskoeffisienten" er et mål på hvor sterkt kodonmønsteret av en gitt polynukleotidsekvens ligner det for en tilsiktet art. Kodonfrekvenser kan avledes fra litteraturkilder for de høyt uttrykte genene til mange arter (se f.eks. Nakamura *et al.* Nucleic Acids Research 1996, 24:214-215). Kodonfrekvensene for hver av de 61 kodonene (uttrykt som det antall forekomster som forekommer per 1000 kodoner av den valgte klasse av gener) er normalisert for hver av de tyve naturlige aminosyrene, slik at verdien for det hyppigst anvendte kodon for hver aminosyre er satt til 1, og frekvensene for de mindre vanlige kodonene er anbragt på en skala mellom null og 1. Hver av de 61 kodonene tilordnes således en

verdi på 1 eller mindre for de høyt uttrykte gener til de tilsiktede arter. For å beregne en kodonanvendelseskoeffisient for et bestemt polynukleotid, i forhold til de høyt uttrykte gener av vedkommende art, registreres den skalerte verdi for hvert kodon av det bestemte polynukleotid, hvorpå det geometriske middel av alle disse verdiene tas (ved å dividere summen av de naturlige logaritmer av disse verdier med det totale antall kodoner og benytte anti-log-verdien). Koeffisienten vil ha en verdi mellom null og 1 og jo høyere koeffisient dess flere kodoner i polynukleotidet er hyppig benyttede kodoner. Dersom en polynukleotidsekvens har en kodonanvendelseskoeffisient på 1, er alle kodonene "mest hyppige" kodoner for høyt uttrykte gener av de tilsiktede arter.

Kodonanvendelsesmønsteret for polynukleotidet vil fortrinnsvis utelukke kodoner med en RSCU-verdi på mindre enn 0,2 i høyt uttrykte gener i målorganismen. Alternativt vil kodonanvendelsesmønsteret utelukke kodoner som representerer <10% av de kodoner som benyttes for en bestemt aminosyre. En RSCU- (relative synonymous codon usage) verdi er det observerte antall kodoner dividert med det forventede antall dersom alle kodoner for vedkommende aminosyre ble benyttet like hyppig. Et polynukleotid ifølge foreliggende oppfinnelse vil i alminnelighet ha en kodonanvendelseskoeffisient (eller RSCU) for høyt uttrykte humangener på mer enn 0,3, fortrinnsvis mer enn 0,4, mest foretrukket mer enn 0,5. Kodonanvendelsestabeller for mennesker finnes også i Genebank.

Til sammenligning har et høyt uttrykt beta-aktingen en RSCU på 0,747. Kodonanvendelsestabellen for *homo sapiens* er angitt nedenfor:

Kodonanvendelsestabell 1:

Homo sapiens [gbpril]: 27143 CDS (12816923 kodoner)

felt: [triplett] [frekvens: per tusen] ([antall])

UUU 17,0 (217684)	UCU 14,8 (189419)	UAU 12,1 (155645)	UGU 10,0 (127719)
UUC 20,5 (262753)	UCC 17,5 (224470)	UAC 15,8 (202481)	UGC 12,3 (157257)
UUA 7,3 (93924)	UCA 11,9 (152074)	UAA 0,7 (9195)	UGA 1,3 (16025)
UUG 12,5 (159611)	UCG 4,5 (57572)	UAG 0,5 (6789)	UGG 12,9 (165930)
CUU 12,8 (163707)	CCU 17,3 (222146)	CAU 10,5 (134186)	CGU 4,6 (59454)
CUC 19,3 (247391)	CCC 20,0 (256235)	CAC 14,9 (190928)	CGC 10,8 (137865)
CUA 7,0 (89078)	CCA 16,7 (214583)	CAA 12,0 (153590)	CGA 6,3 (80709)

	CUG 39,7 (509096)	CCG 7,0 (89619)	CAG 34,5 (441727)	CGG 11,6 (148666)
	AUU 15,8 (202844)	ACU 12,9 (165392)	AAU 17,0 (218508)	AGU 12,0 (154442)
	AUC 21,6 (277066)	ACC 19,3 (247805)	AAC 19,8 (253475)	AGC 19,3 (247583)
5	AUA 7,2 (92133)	ACA 14,9 (191518)	AAA 24,0 (308123)	AGA 11,5 (147264)
	AUG 22,3 (285776)	ACG 6,3 (80369)	AAG 32,6 (418141)	AGG 11,3 (145276)
	GUU 10,9 (139611)	GCU 18,5 (236639)	GAU 22,4 (286742)	GGU 10,8 (138606)
	GUC 14,6 (187333)	GCC 28,3 (362086)	GAC 26,1 (334158)	GGC 22,7 (290904)
10	GUA 7,0 (89644)	GCA 15,9 (203310)	GAA 29,1 (373151)	GGA 16,4 (210643)
	GUG 28,8 (369006)	GCG 7,5 (96455)	GAG 40,2 (515485)	GGG 16,4 (209907)

Kodende GC 52,51% 1.bokstav GC 56,04% 2.bokstav, GC 42,35% 3.bokstav GC 59,13%.

15

Kodonanvendelsestabell 2 (foretrukket):

Kodonanvendelse for humangener (sterkt uttrykt) 1/24/91 (human_high.cod)

20	Aminosyre	Kodon	Antall	/1000	Fraksjon
	Gly	GGG	905,00	18,76	0,24
	Gly	GGA	525,00	10,88	0,14
	Gly	GGT	441,00	9,14	0,12
25	Gly	GGC	1867,00	38,70	0,50
	Glu	GAG	2420,00	50,16	0,75
	Glu	GAA	792,00	16,42	0,25
	Asp	GAT	592,00	12,27	0,25
30	Asp	GAC	1821,00	37,75	0,75
	Val	GTG	1866,00	38,68	0,64
	Val	GTA	134,00	2,78	0,05
	Val	GTT	198,00	4,10	0,07
35	Val	GTC	728,00	15,09	0,25
	Ala	GCG	652,00	13,51	0,17
	Ala	GCA	488,00	10,12	0,13
	Ala	GCT	654,00	13,56	0,17
40	Ala	GCC	2057,00	42,64	0,53
	Arg	AGG	512,00	10,61	0,18
	Arg	AGA	298,00	6,18	0,10
	Ser	AGT	354,00	7,34	0,10
45	Ser	AGC	1171,00	24,27	0,34
	Lys	AAG	2117,00	43,88	0,82
	Lys	AAA	471,00	9,76	0,18
	Asn	AAT	314,00	6,51	0,22

	Asn	AAC	1120,00	23,22	0,78
	Met	ATG	1077,00	22,32	1,00
	Ile	ATA	88,00	1,82	0,05
5	Ile	ATT	315,00	6,53	0,18
	Ile	ATC	1369,00	28,38	0,77
	Thr	ACG	405,00	8,40	0,15
	Thr	ACA	373,00	7,73	0,14
10	Thr	ACT	358,00	7,42	0,14
	Thr	ACC	1502,00	31,13	0,57
	Trp	TGG	652,00	13,51	1,00
	End	TGA	109,00	2,26	0,55
15	Cys	TGT	325,00	6,74	0,32
	Cys	TGC	706,00	14,63	0,68
	End	TAG	42,00	0,87	0,21
	End	TAA	46,00	0,95	0,23
20	Tyr	TAT	360,00	7,46	0,26
	Tyr	TAC	1042,00	21,60	0,74
	Leu	TTG	313,00	6,49	0,06
	Leu	TTA	76,00	1,58	0,02
25	Phe	TTT	336,00	6,96	0,20
	Phe	TTC	1377,00	28,54	0,80
	Ser	TCG	325,00	6,74	0,09
	Ser	TCA	165,00	3,42	0,05
30	Ser	TCT	450,00	9,33	0,13
	Ser	TCC	958,00	19,86	0,28
	Arg	CGG	611,00	12,67	0,21
	Arg	CGA	183,00	3,79	0,06
35	Arg	CGT	210,00	4,35	0,07
	Arg	CGC	1086,00	22,51	0,37
	Gln	CAG	2020,00	41,87	0,88
	Gln	CAA	283,00	5,87	0,12
40	His	CAT	234,00	4,85	0,21
	His	CAC	870,00	18,03	0,79
	Leu	CTG	2884,00	59,78	0,58
	Leu	CTA	166,00	3,44	0,03
45	Leu	CTT	238,00	4,93	0,05
	Leu	CTC	1276,00	26,45	0,26
	Pro	CCG	482,00	9,99	0,17
	Pro	CCA	456,00	9,45	0,16
50	Pro	CCT	568,00	11,77	0,19

Pro	CCC	1410,00	29,23	0,48
-----	-----	---------	-------	------

I henhold til et ytterligere aspekt ved oppfinnelsen tilveiebringes en ekspresjonsvektor som omfatter, og er i stand til å styre ekspresjonen av en polynukleotidsekvens i henhold til det første aspekt ved oppfinnelsen, særlig er kodon-anvendelsesmønsteret av gag polynukleotidsekvensen typisk for høyt uttrykte pattedyrgener, fortrinnsvis høyt uttrykte humangener. Vektoren kan være egnet for å styre ekspresjon av heterolog DNA i bakterie-, insekt- eller pattedyrceller, særlig humanceller. I én utførelsesform er ekspresjonsvektoren p7313 (se Figur 1).

Foretrukket koder ikke gag-genet for gag p6-peptidet. NEF-genet er trunkert for å fjerne sekvensen som koder for den N-terminale region, dvs. fjerning av 30-85, fortrinnsvis 60-85, typisk ca. 81, fortrinnsvis de N-terminale 65 aminosyrene.

RT-genet er også optimalisert for å ligne et høyt uttrykt humangen. RT koder fortrinnsvis for en mutasjon for i det vesentlige å inaktivere enhver revers transkriptaseaktivitet. En foretrukket inaktiveringsmutasjon involverer bruk av W-tryptofan 229 i stedet for K (lysin).

Det kan tilveiebringes en vertscelle som omfatter en polynukleotidsekvens i henhold til oppfinnelsen eller en ekspresjonsvektor i henhold til oppfinnelsen. Vertscellen kan være bakteriell, f.eks. *E.coli*, pattedyr, f.eks. menneske, eller være en insektcelle. Pattedyrceller som omfatter en vektor i henhold til foreliggende oppfinnelse kan være dyrkede celler transfektert *in vitro* eller transfektert *in vivo* ved administrering av vektoren til pattedyret.

Foreliggende oppfinnelse tilveiebringer et farmasøytisk preparat som omfatter en polynukleotidsekvens i henhold til oppfinnelsen. Fortrinnsvis omfatter preparatet en DNA-vektor. Foretrukket omfatter preparatet en pluralitet av partikler, fortrinnsvis gullpartikler, belagt med DNA som omfatter en vektor som koder for en polynukleotidsekvens ifølge oppfinnelsen. Fortrinnsvis koder sekvensen for en HIV gag aminosyresekvens, hvor kodonanvendelsesmønsteret av polynukleotidsekvensen er typisk for høyt uttrykte pattedyrgener, særlig humangener. I alternative utførelsesformer omfatter blandingen et farmasøytisk akseptabelt hjelpestoff og en DNA-vektor i henhold til det andre aspekt ved foreliggende oppfinnelse. Blanding kan også innbefatte en adjuvans.

Vektorene kan benyttes med immunstimulerende midler. Fortrinnsvis administreres det immunstimulerende middel samtidig med nukleinsyrevektoren

ifølge oppfinnelsen, og formuleres foretrukket sammen. Slike immunstimulerende midler innbefatter, men denne liste er ikke på noen måte uttømmende og utelukker ikke andre midler: syntetiske imidazokinoliner som imiquimod [S-26308, R-837], (Harrison *et al.* "Reduction of recurrent HSV disease using imiquimod alone or
 5 combined with a glycoprotein vaccine", Vaccine 19:1820-1826, (2001)); and resiquimod [S-28463, R-848] (Vasilakos *et al.* «Adjuvant activities of immune response modifier R-848: Comparison with CpG ODN», Cellular immunology 204:64-74 (2000).), Schiffske baser av karbonylgrupper og aminer som konstitutivt uttrykkes på antigenpresenterende celler og T-celleoverflater, så som tucarecol (Rhodes, J. *et al.* "Therapeutic potentiation of the immune system by costimulatory Schiff-base-
 10 forming drugs", Nature 377:71-75 (1995)), cytokin, kjemokin og kostimulerende molekyler som enten protein eller peptid, hvilket ville inkludere proinflammatoriske cytokiner så som GM-CSF, IL-1 alfa, IL-1 beta, TGF-alfa og TGF-beta, Th1 indusere så som interferon gamma, IL-2, IL-12, IL-15 og IL-18, Th2 indusere så som IL-4, IL-
 15 5, IL-6, IL-10 og IL-13 og andre kjemokin- og kostimulerende gener så som MCP-1, MIP-1 alfa, MIP-1 beta, RANTES, TCA-3, CD80, CD86 og CD40L, andre immunstimulerende målsøkende ligander så som CTLA-4 og L-selektin, apoptosestimulerende proteiner og peptider så som Fas (49), syntetiske lipidbaserte adjuvaner så som vaxfectin, (Reyes *et al.*, "Vaxfectin enhances antigen specific
 20 antibody titres and maintains Th1 type immune responses to plasmid DNA immunization", Vaccine 19:3778-3786), skvalen, alfa-tokoferol, polysorbat 80, DOPC og kolesterol, endotoksin, [LPS], Beutler, B., "Endotoxin, "Toll-like receptor 4, and the afferent limb of innate immunity", Current Opinion in Microbiology 3:23-30 (2000)); CpG oligo- og dinukleotider, Sato, Y. *et al.*, "Immunostimulatory DNA sequences
 25 necessary for effective intradermal gene immunization", Science 273 (5273):352-354 (1996). Hemmi, H. *et al.*, "Atoll-like receptor recognizes bacterial DNA", Nature 408:740-745, (2000) og andre potensielle ligander som utløser Toll-reseptorer for å frembringe Th1-induserende cytokiner, så som syntetiske mykobakterielle lipo-
 30 proteiner, mykobakterielt protein p19, peptidoglykan, teikoinesyre og lipid A.

Enkelte foretrukne adjuvaner for utløsning av en hovedsakelig Th1-type respons innbefatter for eksempel et lipid A-derivat, så som monofosforyl lipid A, eller fortrinnsvis 3-de-O-acylert monofosforyl lipid A. MPL® adjuvaner er tilgjengelige fra Corixa Corporation (Seattle, WA; se for eksempel US-patent 4.436.727; 4.877.611;

4.866.034 og 4.912.094). CpG-holdige oligonukleotider (hvor CpG-dinukleotidet er umetylert) induserer også en hovedsakelig Th1-respons. Slike oligonukleotider er velkjente og er beskrevet for eksempel i WO 96/02555, WO 99/33488 og US-patent 6.008.200 og 5.856.462. Immunstimulerende DNA-sekvenser er også beskrevet, for eksempel av Sato *et al.*, Science 273:352, 1996. En annen foretrukket adjuvans omfatter et saponin som f.eks. Quil A, eller derivater derav, inklusivt QS21 og QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA); Escin; Digitonin; eller Gypsophila eller Chenopodium quinoa saponiner.

Ved oppfinnelsen muliggjøres også anvendelsen av et polynukleotid ifølge oppfinnelsen eller en vektor ifølge oppfinnelsen, ved behandlingen eller forebyggelse av en HIV-infeksjon.

Fremgangsmåter for behandling eller forhindring av HIV-infeksjoner, eventuelle symptomer eller sykdommer assosiert med slike er beskrevet, som omfatter administrering av en effektiv mengde av et polynukleotid, en vektor eller et farmasøytisk preparat ifølge oppfinnelsen. Administrering av et farmasøytisk preparat kan skje i form av én eller flere individuelle doser, for eksempel som gjentatte doser av det samme DNA-plasmid eller i et "prime-boost" terapeutisk vaksinasjonsregime. I enkelte tilfeller kan "prime" vaksinasjonen være via partikkel-mediert DNA-avlevering av et polynukleotid i henhold til foreliggende oppfinnelse, fortrinnsvis inkorporert i en plasmid-avledet vektor, og "boost" ved administrering av en rekombinant viral vektor som omfatter den samme polynukleotidsekvens, eller ved å gi booster-dosen med proteinet i adjuvans. Omvendt kan primingen være med den virale vektor eller med en proteinformulering, i alminnelighet et protein formulert i adjuvans, og booster-dosen en DNA-vaksine ifølge foreliggende oppfinnelse. Det kan benyttes flere doser av primer og/eller booster.

Fragmenter av gag, nef eller RT-proteiner kan være aktuelle. For eksempel kan et polynukleotid ifølge oppfinnelsen kode for et fragment av HIV gag, nef eller RT protein. Et polynukleotid som koder for et fragment av minst 8, for eksempel 8-10 aminosyrer eller opp til 20, 50, 60, 70, 80, 100, 150 eller 200 aminosyrers lengde ansees å falle innenfor rammen av oppfinnelsen så lenge det kodede oligo- eller polypeptid oppviser HIV-antigenitet. Særlig, men ikke utelukkende, omfatter dette den situasjon hvor polynukleotidet koder for et fragment av en komplett HIV-proteinsekvens, og kan representere én eller flere diskrete epitoper av

vedkommende protein. Slike fragmenter kan kodonoptimaliseres slik at fragmentet har et kodonanvendelsesmønster som ligner det for et høyt uttrykt pattedyrgen.

Foretrukne konstruksjoner inkluderer:

- 5 1. p17, p24, fusjonert til trunkert NEF (som mangler nukleotider som koder for terminale aminosyrer 1-65)
2. p17, p24, RT, trunkert NEF (som mangler nukleotider som koder for terminale aminosyrer 1-65)
- 10 3. p17, p24 (optimalisert gag) trunkert NEF (som mangler nukleotider som koder for terminale aminosyrer 1-65)
4. p17, p24 (optimalisert gag) RT (optimalisert) trunkert NEF (som mangler nukleotider som koder for terminale aminosyrer 1-85)
- 15 5. p17, p24, RT (optimalisert) trunkert NEF (som mangler nukleotider som koder for terminale aminosyrer 1-65)
6. Trunkert NEF – (som mangler nukleotid 1-65) fusjonert til optimalisert p17, p24 gag.
- 20 7. Særlig foretrukne konstruksjoner ifølge oppfinnelsen inkluderer trippel-fusjonene RT-NEF-Gag og RT-Gag-Nef, særlig:
- 25 8. Optimalisert RT, trunkert NEF og optimalisert p17, p24 (gag) (RNG) og
9. Optimalisert RT, optimalisert p17, 24 (gag), Nef trunkert (mangler aminosyrene 1-65) RGN
- 30

Det er å foretrekke at HIV-konstruksjonene er avledet fra en HIV "Clade B" eller "Clade C", særlig "Clade B".

Som omtalt ovenfor inkluderer foreliggende oppfinnelse ekspresjonsvektorer som omfatter nukleotidsekvensene ifølge oppfinnelsen. Slike ekspresjonsvektorer

konstrueres rutinemessig på feltet molekylær biologi, og kan for eksempel involvere anvendelsen av plasmid DNA og passende initiatorer, promotere, enhancere og andre elementer, som for eksempel polyadenyleringssignaler som kan være nødvendige og som er anbrakt i den riktige orientering, for å muliggjøre protein-ekspressjon. Andre egnede vektorer vil være åpenbare for fagmannen. I henhold til et ytterligere eksempel i så henseende viser vi til Sambrook *et al.* Molecular Cloning: a Laboratory Manual. 2. utg. CSH Laboratory Press, (1989).

Fortrinnsvis kobles et polynukleotid ifølge oppfinnelsen eller for bruk i oppfinnelsen, i en vektor, operativt til en kontrollsekvens som kan sørge for ekspresjonen av den kodende sekvens i vertscellen, dvs. vektoren er en ekspresjonsvektor. Betegnelsen "operativt koblet" refererer seg til en posisjonering ved siden av hverandre, hvor de beskrevne komponenter står i et forhold som lar dem virke på tiltenkt måte. En regulatorsekvens, som f.eks. en promoter, "operativt koblet" til en kodende sekvens er posisjonert på en slik måte at ekspresjon av den kodende sekvens oppnås under betingelser som er forenelige med regulatorsekvensen.

Vektorene kan for eksempel være plasmider, kunstige kromosomer (f.eks. BAC, PAC, YAC), virus eller fagvektorer forsynt med et replikasjonsorigo, eventuelt en promoter for ekspresjonen av polynukleotidet og eventuelt en regulator av promoteren. Vektorene kan inneholde ett eller flere selekterbare markørgener, for eksempel et ampicillin- eller kanamycinresistens-gen når det er tale om et bakterielt plasmid, eller et resistensgen for en soppvektor. Vektorer kan benyttes *in vitro*, for eksempel for fremstillingen av DNA eller RNA eller benyttes til å transfekte eller transformere en vertscelle, for eksempel en pattedyr-vertscelle, f.eks. for fremstillingen av protein som kodes av vektoren. Vektorene kan også tilpasses anvendelse *in vivo*, for eksempel i en fremgangsmåte for DNA-vaksinasjon eller for genterapi.

Promotere og andre ekspresjonsreguleringssignaler kan velges slik at de er fortlidelige med den vertscelle som ekspresjonen er beregnet for.

For eksempel kan pattedyr-promotere inkludere metallotionein-promoteren som kan induseres som respons på tungmetaller, som f.eks. kadmium, og β -aktin promoteren. Virale promotere så som den SV40 store T-antigen-promoteren, humant cytomegalovirus (CMV) umiddelbart tidlige (IE) promoteren, Rous' sarcoma virus

LTR-promoter, adenovirus promoter eller en HPV-promoter, særlig den HPV oppstrøms regulerende region (URR) kan også benyttes. Alle disse promoterene er grundig beskrevet og lett tilgjengelige på området.

Et foretrukket promoterelement er den CMV umiddelbart tidlige promoter, uten
 5 intron A, men inklusivt ekson 1. Promoterelementet kan være det minimale promoter-
 element eller den forsterkede promoter, hvorav den forsterkede promoter fore-
 trekkes. Det tilveiebringes følgelig en vektor som omfatter et polynukleotid ifølge opp-
 finnelsen under kontroll av HCMV 1E tidlige promoter.

Eksempler på egnede virale vektorer er herpes simplex virale vektorer,
 10 vaccinia eller alfavirus-vektorer og retrovirus, inklusivt lentivirus, adenovirus og
 adenoassosiert virus. Genoverføringsteknikker som gjør bruk av disse virus er kjent
 for fagmannen. Retrovirus-vektorer kan for eksempel benyttes til stabilt å integrere
 polynukleotidet ifølge oppfinnelsen i vertsgenomet, selv om slik rekombinasjon ikke
 er å foretrekke. Replikasjonsdefekte adenovirus-vektorer forblir derimot episomale
 15 og vil derfor tillate transient ekspresjon. Vektorer som er i stand til å styre ekspresjon
 i insektceller (for eksempel baculovirus-vektorer), i humanceller, gjær eller i bakterier,
 kan benyttes for å produsere mengder av det HIV-protein som kodes av poly-
 nukleotidene ifølge foreliggende oppfinnelse, for eksempel for bruk som subunit-
 vaksiner eller i immunoassays.

Polynukleotidene i henhold til oppfinnelsen kan benyttes i produksjonen ved
 20 ekspresjon av de kodede proteiner, hvor ekspresjon kan skje *in vitro*, *in vivo* eller *ex vivo*.
 Nukleotidene kan derfor være involvert i rekombinant proteinsyntese, for
 eksempel for å øke utbytter, eller kan faktisk finne anvendelse som terapeutiske
 midler i egen kraft, benyttet i DNA-vaksinasjonsteknikker. Når polynukleotidene ifølge
 25 foreliggende oppfinnelse benyttes ved fremstillingen av de kodede proteiner *in vitro*
 eller *ex vivo*, vil celler, for eksempel i cellekultur, bli modifisert til å inkludere det poly-
 nukleotid som skal uttrykkes. Slike celler inkluderer transiente eller fortrinnsvis
 stabile, pattedyrcellelinjer. Spesielle eksempler på celler som kan modifiseres ved
 innføring av vektorer som koder for et polypeptid i henhold til oppfinnelsen, innbe-
 30 fatter pattedyr HEK293T-, CHO-, HeLa-, 293- og COS-celler. Fortrinnsvis vil celle-
 linjen som velges være en som ikke bare er stabil, men som også tillater moden
 glykosylering og celleoverflateekspresjon av et polypeptid. Ekspresjon kan oppnås i
 transformerte oocytter. Et polypeptid kan uttrykkes fra et polynukleotid ifølge fore-
 liggende oppfinnelse, i celler fra et transgent dyr utenom mennesket, fortrinnsvis en

mus. Et transgent dyr utenom mennesket som uttrykker et polypeptid fra et polynukleotid ifølge oppfinnelsen, faller inn under rammen for oppfinnelsen.

En fremgangsmåte for vaksinerings av et pattedyr er beskrevet, som omfatter administrering av en effektiv mengde av en slik vaksine eller vaksineblanding. Mest foretrukket vil ekspresjonsvektorer for bruk i DNA-vaksiner, vaksineblandinger og immunterapeutiske midler være plasmidvektorer.

DNA-vaksiner kan administreres i form av "naken DNA", for eksempel i en flytende formulering, administrert ved å benytte en sprøyte eller høytrykksstråle, eller DNA-formulert med liposomer eller en irriterende transfeksjons-forsterker eller ved partikkelmediert DNA-mediert avlevering (PMDD). Alle disse avleveringssystemene er velkjente på området. Vektoren kan føres inn i et pattedyr, for eksempel ved hjelp av et viralt vektoravleveringssystem.

Preparatet ifølge foreliggende oppfinnelse kan avgis ved hjelp av en rekke administrasjonsmåter, så som intramuskulært, subkutan, intraperitonealt, intravenøst eller mukosalt.

Foretrukket administreres preparatet intradermalt. Spesielt avleveres preparatet ved hjelp av en genkanon, særlig ved partikkelbombardement som innebærer at vektoren belegges på en kule (f.eks. gull) som deretter administreres under høyt trykk inn i epidermis; som for eksempel beskrevet i Haynes *et al.*, J. Biotechnology 44:37-42 (1996).

I ett illustrerende eksempel kan det oppnås gassdrevet partikkelakselerasjon med anordninger fremstillet av Powderject Pharmaceuticals PLC (Oxford, UK) og Powderject Vaccines Inc. (Madison, WI), hvorav enkelte eksempler er beskrevet i US-patent 5.846.796; 6.010.478; 5.865.796; 5.584.807 og i EP patent nr. 0500 799. Denne fremgangsmåte byr på en nålfri avleveringsmåte, hvor en tørrpulverformulering av mikroskopiske partikler, så som polynukleotid, akselereres til høy hastighet i en helium gass-stråle utviklet av en manuell anordning, hvor partiklene drives inn i et målvev av interesse, i alminnelighet huden. Partiklene er fortrinnsvis gullkuler med en 0,4-4,0 μm , mer foretrukket 0,6-2,0 μm diameter som DNA-konjugatet er belagt på og som så er innkapslet i en patron eller kassett for anbringelse i "genkanonen".

Andre anordninger og fremgangsmåter som kan være egnet for gassdrevet nålfri injeksjon av preparatet ifølge foreliggende oppfinnelse kan omfatte dem som leveres av Bioject, Inc. (Portland, OR), hvorav enkelte eksempler er beskrevet i US-

patent 4.790.824; 5.064.413; 5.312.335; 5.383.851; 5.399.163; 5.520.639 og 5.993.412.

Vektorene som omfatter de nukleotidsekvenser som koder for antigene peptider, administreres i en mengde som vil være profylaktisk eller terapeutisk effektiv. Mengden som administreres ligger i alminnelighet i området fra 1 pikogram til 1 milligram, fortrinnsvis 1 pikogram til 10 mikrogram for partikkelmediert avlevering, og 100 nanogram til 1 milligram, fortrinnsvis 10 mikrogram til 1 milligram for andre administrasjonsmåter, nukleotid per dose. Den nøyaktige mengde kan variere betydelig avhengig av vekten av pasienten som immuniseres og av administrasjonsmåten.

Det er mulig å administrere den immunogenkomponent som omfatter nukleotidsekvensen som koder for det antigene peptid, som en engangsoperasjon eller å foreta administreringen gjentatt, for eksempel mellom 1 og 7 ganger, fortrinnsvis mellom 1 og 4 ganger, med mellomrom på ca. 1 dag og ca. 18 måneder. Dette behandlingsregime vil imidlertid variere betydelig avhengig av størrelsen av angjeldende pasient, mengden av administrert nukleotidsekvens, administrasjonsmåten og andre faktorer som vil være åpenbare for en veterinær eller lege. Pasienten kan få én eller flere andre anti-HIV retrovirale medikamenter som del av deres totale behandlingsregime. Nukleinsyre-immunogenet kan dessuten administreres med en adjuvans.

Den her angitte adjuvanskomponent kan administreres tilsvarende via en rekke forskjellige administrasjonsmåter, som for eksempel peroralt, nasalt, pulmonalt, intramuskulært, subkutant, intradermalt eller topisk. Fortrinnsvis administreres adjuvanskomponenten intradermalt eller topisk, mest foretrukket topisk. Administreringen kan foretas mellom ca. 14 dager før og ca. 14 dager etter administrering av nukleotidsekvensen, fortrinnsvis mellom ca. 1 dag før og ca. 3 dager etter administrering av nukleotidsekvensen. I en utførelsesform administreres adjuvanskomponenten i det vesentlige samtidig med administreringen av nukleotidsekvensen. Med "i det vesentlige samtidig med" menes at administreringen av adjuvanskomponenten fortrinnsvis skjer samtidig som administreringen av nukleotidsekvensen, eller om ikke, minst i løpet av få timer før eller etter nukleotidsekvensadministrering. Etter den mest foretrukne behandlingsprotokoll vil adjuvanskomponenten bli administrert i det vesentlige samtidig med administrering av nukleotidsekvensen. Denne protokoll kan selvsagt varieres etter behov, overens-

stemmende med typen av variabler omtalt ovenfor. Det er å foretrekke at adjuvansen er et 1H-imidazo[4,5-c]-kinolin-4-aminderivat, så som imiquimod. I alminnelighet vil imiquimod bli tilberedt som en topisk kremformulering og bli administrert i henhold til ovennevnte protokoll.

5 Igjen vil den administrerte dose av derivatet også variere med slike variabler, men kan for eksempel variere mellom ca. 0,1 mg per kg til ca. 100 mg per kg, hvor "per kg" refererer seg til kroppsvekten av angjeldende pattedyr. Denne administrering av 1H-imidazo[4,5-c]kinolin-4-aminderivatet vil fortrinnsvis bli gjentatt ved hver påfølgende administrering eller booster-administrering av nukleotidsekvensen. Mest foretrukket vil den administrerte dose være mellom ca. 1 mg per kg til ca. 50 mg per 10 kg. Når det gjelder et "prime-boost"-skjema som her beskrevet, kan imiquimod eller annet 1H-imidazo[4,5-c]kinolin-4-aminderivatet administreres enten med primer- eller booster-dosen eller med både primer- og booster-dosen.

Selv om adjuvanskomponenten kan omfatte kun 1H-imidazo[4,5-c]kinolin-4- 15 aminderivater for administrering i rå kjemisk tilstand, er det å foretrekke at administreringen skjer i form av en farmasøytisk formulering. Det vil si at adjuvanskomponenten fortrinnsvis vil omfatte 1H-imidazo[4,5-c]-kinolin-4-aminet kombinert med én eller flere farmasøytisk akseptable bærere, og eventuelt med andre terapeutiske ingredienser. Bærerne må være "akseptable" i den forstand at de er 20 forlikelige med andre ingredienser i formuleringen og ikke skadelige for mottakeren av denne. Formuleringenes natur vil selvsagt variere med den tiltenkte administrasjonsmåte, og de kan fremstilles etter fremgangsmåter velkjent på det farmasøytiske området. Alle metodene inkluderer det trinn å bringe et 1H-imidazo[4,5-c]-kinolin-4-aminderivat sammen med én eller flere passende bærere. I 25 alminnelighet fremstilles formuleringene ved jevnt og intimt å bringe sammen derivatet med flytende bærere eller findelte faste bærere, eller begge, og deretter om nødvendig, forme produktet til den ønskede formulering. Formuleringer ifølge foreliggende oppfinnelse egnet for peroral administrering kan tilberedes som diskrete enheter, som f.eks. kapsler, oblatkapsler eller tabletter som hver inneholder en fastsatt 30 mengde virkestoff; som et pulver eller granuler; en løsning eller en suspensjon i en vandig væske eller en ikke-vandig væske; eller som en flytende olje-i-vann emulsjon eller en vann-i-olje emulsjon. Virkestoffet kan også tilberedes som bolus, klyster eller pasta.

En tablett kan fremstilles ved pressing eller støping, eventuelt med én eller flere hjelpeingredienser. Pressede tabletter kan fremstilles ved i en egnet maskin og presse virkestoffet i en frittstrømmende form så som pulver eller granuler, eventuelt blandet med et bindemiddel, smøremiddel, inert fortynningsmiddel, smørende,
5 overflateaktivt eller dispergerende middel. Støpte tabletter kan fremstilles i en passende maskin og støpe en blanding av den pulveriserte forbindelse fuktet med et inert flytende fortynningsmiddel.

Tablettene kan eventuelt være drasjert eller forsynt med deleskår og kan formuleres slik at de gir langsom eller kontrollert frigjøring av virkestoffet.

10 Formuleringer for injeksjon, for eksempel ved intramuskulær, intraperitoneal eller subkutan administrering, innbefatter vandige og ikke-vandige sterile injeksjonsløsninger som kan inneholde antioksydanter, buffere, bakteriestatika og oppløst stoff som gjør formuleringen isotonisk med blodet til den tiltenkte mottaker; og vandige og ikke-vandige sterile suspensjoner som kan inkludere suspenderingsmidler og for-
15 tykningsmidler. Formuleringene kan tilberedes i enhetsdoseform eller som fler-dosebeholdere, for eksempel forseglede ampuller og glass, og kan oppbevares i frysetørket (lyofilisert) tilstand som bare fordrer tilsetning av den sterile flytende bærer, for eksempel vann for injeksjon, umiddelbart før bruk. Injeksjonsløsninger og suspensjoner for tilberedning på stedet kan fremstilles fra sterile pulvere, granuler og
20 tabletter av den ovenfor beskrevne type. Formuleringer egnet for pulmonal administrasjon via kinn- eller nesehulen tilberedes slik at partiklene som inneholder virkestoffet og hensiktsmessig har en diameter i området fra 0,5 til 7 mikron, avleveres til mottakerens bronkialtre. Slike formuleringer har den mulighet at de er i form av findelte pulvere som lett kan fremstilles enten som en perforerbar kapsel, for
25 eksempel av gelatin, for anvendelse i en inhalasjonsanordning, eller alternativt, som en selvdrevet formulering som omfatter virkestoff, et egnet flytende drivmiddel og eventuelt andre ingredienser, så som surfaktant og/eller fast fortynningsmiddel. Det kan også benyttes selvdrevne formuleringer hvor virkestoffet avgis i form av dråper av en løsning eller suspensjon. Slike selvdrevne formuleringer tilsvarer de som er
30 kjent på området og kan fremstilles etter etablerte fremgangsmåter. Disse er hensiktsmessig forsynt med en ventil som kan betjenes manuelt eller automatisk og som har de ønskede forstøvningsegenskaper. Ventilene kan med fordel være av avmålingstype som avgir et fast volum, for eksempel 50 til 100 μL ved hver betjening av denne.

Som en ytterligere mulighet kan adjuvanskomponenten ha form av en løsning for bruk i en atomisør eller nebulisator, hvorved en akselerert luftstrøm eller ultralyd-agitasjon benyttes for å produsere en tåke av fine dråper for inhalasjon.

Formuleringer egnet for intranasal administrasjon innbefatter i alminnelighet
5 lignende former som de som ovenfor er beskrevet for pulmonal administrasjon, selv om det er å foretrekke at slike formuleringer har en partikkeldiameter i området fra ca. 10 til ca. 200 mikron, slik at de kan holdes igjen i nesehulen. Dette kan oppnås ved å benytte et pulver av egnet partikkelstørrelse eller ved valg av en passende ventil. Andre egnede formuleringer er grove pulvere som har en partikkeldiameter i området fra ca. 20 til ca. 500 mikron, for administrasjon ved hurtig
10 inhalasjon gjennom nesepassasjen, fra en beholder som holdes tett opp til nesen, og nese-dråper som omfatter ca. 0,2 til 5 vekt % virkestoff i vann- eller oljebaserte løsninger. I én utførelsesform av oppfinnelsen kan vektoren som omfatter den nukleotidsekvens som koder for det antigene peptid, administreres i den samme
15 formulering som 1H-imidazo[4,5-c]kinolin-4-aminderivatet. I denne utførelsesform finnes således immunogen- og adjuvanskomponenten i den samme formulering.

I én utførelsesform fremstilles adjuvanskomponenten i en form egnet for administrering med genkanon og administreres på denne måte i det vesentlige samtidig med administrering av nukleotidsekvensen. For fremstilling av formuleringer
20 egnet for bruk på denne måte kan det være nødvendig at 1H-imidazo[4,5-c]-kinolin-4-aminderivatet lyofiliseres og klebes til for eksempel gullkuler som er egnet for administrering med genkanon.

I en alternativ utførelsesform kan adjuvanskomponenten administreres som et tørt pulver via gassfremdrift under høyt trykk.

25 Selv om de ikke formuleres sammen kan det være hensiktsmessig at adjuvanskomponenten administreres til eller omtrent det samme administrasjonssete som nukleotidsekvensen.

Andre detaljer angående farmasøytiske preparater finnes i Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pennsylvania
30 (1985), som her med hele sitt innhold inkorporeres ved referanse.

Passende teknikker for innføring av det nakne polynukleotid eller vektoren i en pasient inkluderer også topisk administrering med en passende bærer. Nukleinsyren kan administreres topisk til huden eller til slimhinner, for eksempel ved intranasal, peroral, intravaginal eller intrarektal administrering. Det nakne polynukleotid eller

vektoren kan foreligge sammen med et farmasøytisk akseptabelt hjelpestoff, som f.eks. fosfatbuffret saltvann (PBS). DNA-opptak kan lettes ytterligere ved å benytte midler som f.eks. bupivakain, enten separat eller inkludert i DNA-formuleringen. Andre fremgangsmåter for administrering av nukleinsyren direkte til en mottaker
5 omfatter ultralyd, elektrisk stimulering, elektroporering og "microseeding" som er beskrevet i US-patent 5.697.901.

Opptak av nukleinsyrekonstruksjoner kan forbedres ved hjelp av flere kjente transfeksjonsteknikker, for eksempel de som inkluderer anvendelsen av trans-feksjonsmidler. Eksempler på slike midler er kationiske midler, for eksempel kalsium-
10 fosfat og DEAE-dekstran og lipofektanter, for eksempel lipofectam og transfectam. Doseringen av den nukleinsyre som skal administreres kan endres.

En nukleinsyresekvens kan også administreres ved hjelp av spesielle avleveringsvektorer egnet innen genterapi. Genterapimetoder er omtalt for eksempel av Verme *et al.*, Nature 1997, **389**:239-242. Både virale og ikke-virale
15 vektorsystemer kan benyttes. Virusbaserte systemer omfatter retrovirale, lentivirale, adenovirale, adenoassosiert virale, herpes-virale, Canarypox- og vaccinia-virale systemer. Systemer som ikke er basert på virus, inkluderer direkte administrering av nukleinsyrer, mikrokuleinnkapslings-teknologi (poly(laktid-ko-glykolid) og liposombaserte systemer. Virale og ikke-virale avleveringssystemer kan kombineres
20 når det er ønskelig å gi boosterinjeksjoner etter en initial vaksinasjon, for eksempel en initial "prime" DNA-vaksinasjon, ved å benytte en ikke-viral vektor som f.eks. et plasmid, etterfulgt av én eller flere "booster" vaksinasjoner ved bruk av en viral vektor eller et ikke-viralt system. Tilsvarende kan det i forbindelse med oppfinnelsen vurderes "prime boot" systemer med polynukleotidet ifølge oppfinnelsen, etterfulgt av
25 booster-injeksjoner med protein i adjuvans eller *vice versa*.

En nukleinsyresekvens kan også administreres ved hjelp av transformerte celler. Slike celler inkluderer celler tatt fra et individ. Det nakne polynukleotid eller vektoren ifølge foreliggende oppfinnelse kan innføres i slike celler *in vitro*, og de transformerte cellene kan senere returneres til individet. Polynukleotidene ifølge
30 oppfinnelsen kan integreres i nukleinsyre som allerede foreligger i en celle, ved homolog rekombinasjon. Om ønskes kan en transformert celle dyrkes *in vitro* og én eller flere av de resulterende celler kan benyttes. Celler kan tilføres til et passende sted i en pasient ved hjelp av kjent kirurgisk eller mikrokirurgisk teknikk (f.eks. grafting, mikro-injeksjon, etc.).

De farmasøytiske preparatene ifølge foreliggende oppfinnelse kan inkludere adjuvansforbindelser slik som angitt detaljert ovenfor, eller andre substanser som kan tjene til å øke den immunrespons som induseres av proteinet som kodes av denne DNA. Disse kan kodes av DNA, enten separat fra antigenet eller som en fusjon med dette, eller kan inkluderes som ikke-DNA-elementer av formuleringen. Eksempler på substanser av adjuvans-type som kan inkluderes i formuleringen ifølge foreliggende oppfinnelse er ubikvitin, lysosomalt assosiert membranprotein (LAMP), hepatitt B, viruskjerneantigen, FLT3-ligand (et cytokin som er viktig ved utviklingen av profesjonelle antigenpresenterende celler, særlig dendritiske celler) og andre cytokiner som f.eks. IFN- γ og GM-CSF. Andre foretrukne adjuvanser omfatter imiquimod og Resiquimod og Tucarasol. Imiquimod er særlig å foretrekke.

Ved den foreliggende oppfinnelse muliggjøres foretrukket anvendelsen av et nukleinsyremolekyl som her beskrevet, for behandling eller forebyggelse av HIV-infeksjon. Nukleinsyremolekylet administreres fortrinnsvis med imiquimod. Imiquimod administreres fortrinnsvis topisk, mens nukleinsyremolekylet fortrinnsvis administreres ved hjelp av partikkelmediert avlevering.

En fremgangsmåte for behandling av et individ som lider av eller er mottagelig for HIV-infeksjon er beskrevet, som omfatter administrering av et nukleinsyremolekyl som her beskrevet, og imiquimod.

Foreliggende oppfinnelse vil nedenfor bli beskrevet med henvisning til de følgende eksempler:

EKSEMPLER

Eksempel 1: Optimalisering av p55 gag (p17, p24, p13) for etterligning av kodonanvendelse av høyt uttrykte humane gener.

Gen av interesse

Et syntetisk gen som koder for p55 gag-antigenet av HIV-1 clade B stamme HXB2 (GenBank tilgang K03455), optimalisert for ekspresjon i pattedyrceller, ble sammensatt av overlappende oligonukleotider ved PCR.

Optimalisering involverte endring av kodonanvendelsesmønsteret for det virale gen for å gi en kodonfrekvens som var nærmere den som finnes i høyt uttrykte

humangener. Kodoner ble tilordnet ved å benytte et statistisk Visual Basic program betegnet Syngene (en oppdatert versjon av Calcgene, skrevet av R.S. Hale og G. Thompson, Protein Expression and Purification bd. 12 s. 185-188, 1998).

5 Kloning:

1528bp gag PCR-produktet ble gelrenset, kuttet med restriksjonsendo-nukleasene Not og Bam HI og ligert inn i NotI/BamHI kuttet vektor WRG7077. Dette plasserer genet mellom CMV promoteren/intron A og polyadenyleringssignalet for bovin veksthormon.

10 Kloner ble sekvensert og kontrollert med henblikk på feil. Ingen enkeltklon var 100% korrekt. Regioner med korrekt sekvens fra to kloner ble derfor kombinert ved overlappende PCR under bruk av passende kombinasjoner av det optimaliserte oligosett for å gi et full-lengde kodonoptimalisert gag-gen. Denne endelige klon ble senere funnet å inneholde en enkelt nukleotiddelesjon som resulterte i en ramme-
15 forskyvning og for tidlig terminering av translasjon. Delesjonen ble reparert ved å kutte ut det området av genet som inneholdt den feilaktige sekvens og kloner inn den riktige sekvens fra den tilsvarende region av en annen klon. Dette ga den endelige kodonoptimaliserte p55 gag-klon:
Gagoprpr2. (Se Figur 2)

20

Eksempel 2: Fremstilling av et p17/p24 trunkert Nef fusjonsgen

Gen av interesse

p17- og p24-delene av p55gag-genet avledet fra HIV-1 clade B-stamme HXB2 ble PCR-amplifisert fra plasmidet pHXB?Pr (B.Maschera, E. Furfine og E.D. Blair,
25 1995 J. Virol 69, 5431-5436). pHXB?Pr. 426bp fra 3' enden av HXB2 nef-genet ble amplifisert fra det samme plasmid. Siden HXB2 nef-genet inneholder en prematur termineringskodon ble to overlappende PCR brukt til å reparere kodonet (TGA [stop] til TGG [Trp]).

30 PCR-produktene p17/p24-linker og trNEF-linker ble forbundet for å danne p17p24trNEF-fusjonsgenet (Figur 3) i en PCR-reaksjon (**antisense**).

1542bp-produktet ble gelrenset, kuttet med restriksjonsendonukleasene NotI og BamHI og klonet inn i NotI BamHI setene til vektor WRG7077. Dette anbringer

genet mellom CMV promoter/intron A og polyadenyleringssignalet for bovin veksthormon.

Eksempel 3: Produksjon av et Gag p17/24opt/trNef1 ("Gagopt/Nef") fusjonsgen

Gen av interesse

p17/p24-delen av det kodonoptimaliserte p55gag-gen avledet fra HIV-1 clade B stamme HXB2 ble PCR-amplifisert fra plasmidet pGagOPTpr2. Det trunkerte HXB2 Nef-gen med den premature termineringskodon reparert (TGA [stop] til TGG [Trp]) ble amplifisert ved PCR fra plasmidet 7077trNef20. De to PCR-produktene ble konstruert med overlappende ender slik at de to genene kunne forbindes ved en andre PCR.

1544bp produktet ble gelrenset, kuttet med restriksjonsendonukleasene NotI og BamHI og klonet (se figurer) inn i NotI BamHI setene av vektor WRG7077. Dette anbringer genet mellom CMV promoter/intron A og polyadenyleringssignalet for det bovine veksthormon.

Eksempel 4: Plasmid: p7077-RT3 klon #A

Gen av interesse

Et syntetisk gen som koder for RT-delen av pol-genet av HIV-1 clade B stamme HXB2, optimalisert for ekspresjon i pattedyrceller, ble sammensatt av overlappende oligonukleotider ved PCR. Den klonede sekvens tilsvarer posisjonene 2550-4222 av HXB2-referansesekvensen (GenBank tilgang K03455). For å sikre ekspresjon har den klonede sekvens to ytterligere kodoner i 5' enden som ikke forekom i det opprinnelige gen – AUG GGC (Met Gly). Optimalisering involverte endring av kodonanvendelsesmønsteret av det virale gen for å gi en kodonfrekvens nærmere den som finnes i høyt uttrykte humangener, men uten sjeldent benyttede kodoner. Kodoner ble tilordnet ved å benytte et statistisk Visual Basic program betegnet Syngene (en oppdatert versjon av Calcgene, skrevet av R.S. Hale & G. Thompson, Protein Expression and Purification bd. 12 s. 185-188, 1998).

Den endelige klon ble konstruert fra to intermediære kloner, #16 og #21.

Kloning:

De 1,7 kb PCR-produktene ble gelrenset, kuttet med NotI og BamHI og PCR-renset før de ble ligert med NotI/BamHI-kuttet pWRG7077. Dette anbringer genet mellom CMV-promoter og polyadenyleringssignalet for bovint veksthormon. Kloner ble sekvensert. Ingen klon var 100% korrekt, men klon #16 ble korrigert ved å erstatte 403 bp KpnI-BamHI-fragmentet som inneholdt 3 feil med et korrekt KpnI-BamHI-fragment fra klon #21. Den endelige klon ble verifisert ved sekvensering (se Figur 5).

Eksempel 5: Optimalisert RT**Gen av interesse**

Det syntetiske gen som koder for RT-delen av pol-genet av HIV-1 clade B stamme HXB2, optimalisert for ekspresjon i pattedyrceller, ble kuttet ut fra plasmid p7077-RT3 som et 1697 bp NotI/BamHI-fragment, gelrenset og klonet inn i NotI & BamHI-setene av p7313-ie (avledet fra pspC31) for å anbringe genet nedstrøms for en Iowa lengde HCMV-promoter + ekson 1, og oppstrøms for et kanin globin polyadenyleringssignal. (R7004 p27) (Figur 6).

Eksempel 6**Plasmid: 7077trNef20****Gen av interesse**

Innskuddet omfatter del av Nef-genet fra HIV-1 clade B stamme HXB2. 195 bp er deletert fra 5' enden av genet som fjerner kodonene for de første 65 aminosyrene av Nef. Dessuten er det premature termineringskodon i den publiserte HXB2 nef-sekvens blitt reparert (TAG til TGG [Trp]), som beskrevet for plasmid p17/24trNEF1. Den trunkerte nef-sekvens ble PCR-amplifisert fra plasmidet p17/24trNef1. Den klonede sekvens er ekvivalent med posisjonene 8992-9417 av HXB2 referansesekvensen (GenBank tilgang K03455). For å sikre ekspresjon har den klonede sekvens et ytterligere kodon ved 5' enden som ikke er tilstede i det opprinnelige gen - AUG (Met).

Primere:

StrNef (sense) ATAAGAATGCGGCCGCCATGGTGGGTTTTCCAGTCACACCTT
[SEKV. ID. Nr:1]

AstrNef (antisense)

CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [SEKV. ID. Nr:2]

PCR: 94°C 2 min, deretter 25 sykluser: 94°C 30 sek, 50°C 30 sek, 72°C 2min, avslutning 72°C 5 min.

5

Kloning:

455 bp RT PCR-produktet ble gelrenset, kuttet med restriksjons-
endonukleasene NotI og BamHI og ligert inn i NotI/BamHI-kuttet vektor WRG7077.
Dette anbringer genet mellom CMV promoteren/intron A og polyadenyleringssignalet
av bovin veksthormon.

10

Eksempel 7

Plasmid: 7077RT 8

Gen av interesse

RT-delen av pol-genet ble avledet fra HIV-1 clade B stamme HXB2. Den ble
PCR-amplifisert fra plasmidet p7077Pol14.

Den klonede sekvens er ekvivalent med posisjonene 2550-4234 av HXB2
referansesekvensen (GenBank tilgang K03455). For å sikre ekspresjon har den
klonede sekvens to ytterligere kodoner ved 5' enden som ikke er tilstede i det opp-
rinnelige gen - AUG GGC (Met Gly).

20

Primere:

SRT (sense) ATAAGAATGCGGCCGCCATGGGCCCCATTAGCCCTATTGAGACT
[SEKV. ID. Nr:3]

ASRT (antisense)

CGCGGATCCTTAATCTAAAAATAGTACTTTCCTGATT [SEKV. ID. Nr:4]

PCR: 94°C 2 min, deretter 25 sykluser: 94°C 30 sek, 50°C 30 sek, 72°C 4 min,
avslutning 72°C 5 min.

25

Kloning:

1720 bp RT PCR-produktet ble gelrenset, kuttet med restriksjons-
endonukleasene NotI og BamHI og ligert inn i NotI/BamHI-kuttet vektor WRG7077. Dette anbringer

30

genet mellom CMV promoter/intron A og polyadenyleringssignalet av bovin veksthormon.

Eksempel 8

5

p17/24opt/RT/trNef13 ("Gagopt/RT/Nef")

Denne konstruksjon inneholder en PCR som bevirker en R- til H-aminosyreendring.

10 **Gen av interesse**

p17/p24-delen av det kodonoptimaliserte p55 gag-gen avledet fra HIV-1 clade B stamme HXB2 ble PCR-amplifisert fra plasmidet pGagOPT rpr2. Den RT-kodende sekvens ble PCR-amplifisert fra plasmidet 7077RT 8. Det trunkerte HXB2 Nef-gen med det premature termineringskodon reparert (TGA [stop] til TGG [Trp]) ble
15 amplifisert ved PCR fra plasmidet 7077trNef20. De tre PCR-produktene ble konstruert med overlappende ender slik at de tre genene kunne forbindes ved en andre PCR.

Primere:

20 (P17/24)

Sp17p24opt (sense)

ATAAGAATGCGGCCGCCATGGGTGCCCGAGCTTCGGT [SEKV. ID. Nr:5]

ASp17p24optRTlinker (antisense)

TGGGGCCCATCAAACTCTGGCTTTGTGTC [SEKV. ID. Nr:6]

25 PCR: 94°C 1 min, deretter 20 sykluser: 94°C 30 sek, 50°C 30 sek, 72°C 2 min, avslutning 72°C 4 min.

1114 bp p17/24opt-produkter ble gelrenset
(RT)

Sp17p24optRTlinker (sense)

30 CAGAGTGTTGATGGGCCCCATTAGCCCTAT [SEKV. ID. Nr:7]

ASRTtrNeflinker (antisense)

AACCCACCATATCTAAAAATAGTACTTTCC [SEKV. ID. Nr:8]

PCR som ovenfor

1711 bp RT PCR-produktet ble gelrenset

(5' trunkert nef)

SRTtrNeflinker (sense)

CTATTTTATAGATATGGTGGGTTTTCCAGTCAC [SEKV. ID. Nr:9]

AstrNef (antisense)

5 CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [SEKV. ID. Nr:10]

PCR som ovenfor

448 bp-produktet ble gelrenset.

De tre PCR-produktene ble deretter heftet sammen i en andre PCR med primerne Sp17/24opt og AstrNef.

10 PCR: 94°C 1 min, deretter 30 sykluser: 94°C 30 sek, 50°C 30 sek, 72°C 4 min, avslutning 72°C 4 min.

3253 bp-produktet ble gelrenset, kuttet med restriksjons-endonukleasene NotI og BamHI og klonet inn i NotI BamHI-setene av vektor WRG7077. Dette bringer genet mellom CMV promoteren/intron A og polyadenyleringssignalet av bovin veksthormon.

15

Eksempel 9

Plasmid: pGRN#16 (p17/p24opt corr /RT/trNef.)

Gen av interesse

20 Polyproteinet utviklet med p17/24opt/RT/trNef13 ("Gagopt/RT/Nef") ble observert å uttrykke et trunkert produkt av ~30 kDa som følge av en ansamling av ugunstige kodoner innen p24 omkring aminosyre 270. Denne ble erstattet med optimale kodoner ved PCR sammenhektende mutagenese.

p17/24opt/RT/trNef13 ble benyttet som templat for å amplifisere delen av Gag 5' til

25 mutasjonen med primerne Sp17/p24opt og GTR-A, og delen av Gag 3' til mutasjonen med primerne GTR-S og Asp17/p24optRTlinker. Overlappingen av de produkter som inneholdt kodonendringene og de gelrensede produktene ble heftet sammen ved å benytte Sp17/p24opt- og Asp17/p24optRTlinker-primerne. Produktet ble kuttet med NotI og AgeI og innskutt i et tilsvarende kuttet p17/24opt/RT/trNef13,

30 for å frembringe pGRN.

Klon #16 ble verifisert og progressert.

Primere:5' PCR:

Sp17p24opt (sense)

ATAAGAATGCGGCCGCCATGGGTGCCCCGAGCTTCGGT [SEKV. ID. Nr:11]

5 GTR-A (Antisense)

GCGCACGATCTTGTTTCAGGCCAGGATGATCCACCGTTTATAGATTTCTCC
[SEKV. ID. Nr:12]3' PCR

Sense: GTR-S (Sense)

10 ATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCC
[SEKV. ID. Nr:13]

ASp17p24optRTlinker (antisense)

TGGGGCCCATCAACACTCTGGCTTTGTGTC [SEKV. ID. Nr:14]

PCR-betingelser for individuelle produkter og sammenheking ved å benytte

15 PWO DNA-polymerase (Roche):

95°C 1 min, deretter 20 sykluser: 95°C 30 sek, 55°C 30 sek, 72°C 180 sek,
avslutning 72°C 120 sek og 4°C hold.1114 bp-produktet ble gelrenset og kuttet med NotI og AgeI for å frigjøre et
6647 bp-fragment som ble gelrenset og ligert inn i NotI/AgeI-kuttet gelrenset

20 p17/24opt/RT/trNef13 for å frembringe pGRN#16.

Eksempel 10:**Plasmid: p73i-GRN2 Klon #19****(p17/p24(opt) /RT (opt) trNef) - reparert**25 **Gen av interesse**p17/p24-delen av det kodonoptimaliserte gag, kodonoptimalisert RT og
trunkerte Nef-gen fra HIV-1 clade B stamme HXB2 nedstrøms for en lowa lengde
HCMV-promoter + ekson1, og oppstrøms for et kanin β -globin
polyadenyleringssignal.30 Plasmider inneholdende trNef-genet avledet fra plasmid p17/24trNef1 inne-
holder en PCR-feil som gir en R- til H-aminosyreendring 19 aminosyrer fra enden av
nef. Denne ble korrigert ved PCR-mutagenese, den korrigerte nef PCR heftet til

kodonoptimalisert RT fra p7077-RT3 og det tilhekte fragment kuttet med Apal og BamHI og klonet inn i Apal/BamHI kuttet p73i-GRN.

Primere:

- 5 PCR coRT fra p7077-RT3 under bruk av primere:
(Polymerase = PWO (Roche) gjennom det hele.
Sense: U1
GAATTCGCGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGA
ACTGAAACCCGGGAT [SEKV. ID. Nr:15]
- 10 AScoRT-Nef
GGTGTGACTGGAAAACCCACCATCAGCACCTTTCTAATCCCCGC [SEKV. ID.
Nr:16]
Syklus: 95°C (30 sek) deretter 20 sykluser 95°C (30 sek), 55°C (30 sek), 72°C
(180 sek), deretter 72°C (120 sek) og hold ved 4°C.
- 15 Det 1,7 kb PCR-produktet ble gelrenset.
PCR 5' Nef fra p17/24trNef1 ved å benyttet primere:
Sense: S-Nef
ATGGTGGGTTTTCCAGTCACACC [SEKV. ID. Nr:17]
Antisense: ASNef-G:
20 GATGAAATGCTAGGCGGCTGTCAAACCTC [SEKV. ID. Nr:18]
Syklus: 95°C (30 sek) deretter 15 sykluser 95°C (30 sek), 55°C (30 sek), 72°C
(60 sek), deretter 72°C (120 sek) og hold ved 4°C.
PCR 3' Nef fra p17/24trNef1 ved å benytte primere:
Sense SNEF-G
25 GAGGTTTGACAGCCGCCTAGCATTTCATC [SEKV. ID. Nr:19]
Antisense:
AStrNef (antisense)
CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [SEKV. ID. Nr:20]
Syklus: 95°C (30 sek) deretter 15 sykluser 95°C (30 sek), 55°C (30 sek), 72°C
30 (60 sek), deretter 72°C (120 sek) og hold ved 4°C.

PCR-produktene ble gelrenset. I første omgang ble de to Nef-produktene heftet sammen ved å benytte 5' (S-Nef) og 3' (AstrNef) primerne.

Syklus: 95°C (30 sek) deretter 15 sykluser 95°C (30 sek), 55°C (30 sek), 72°C (60 sek), deretter 72°C (180 sek) og hold ved 4°C.

PCR-produktet ble PCR-renset og heftet til RT-produktet ved å benytte U1 og AstrNef-primere.

- 5 Syklus: 95°C (30 sek) deretter 20 sykluser 95°C (30 sek), 55°C (30 sek), 72°C (180 sek), deretter 72°C (180 sek) og hold ved 4°C.

2,1 kb-produktet ble gelrenset og kuttet med Apal og BamHI. Plasmidet p73I-GRN ble også kuttet med Apal og BamHI gelrenset og ligert med Apal-Bam RT3trNef for å regenerere p17/p24 (opt)/RT (opt) trNef-genet.

10

Eksempel 11

p73i-GN2 Klon #2 (p17/p24opt/trNef) - reparert

Gen av interesse

- 15 p17/p24-delen av de kodonoptimaliserte gag- og trunkerte Nef-genene fra HIV-1 clade B stamme HXB2 nedstrøms for en lowa lengde HCMV-promoter + ekson1 og oppstrøms for et kanin β -globin polyadenyleringssignal.

- 20 Plasmider inneholdende trNef-genet avledet fra plasmid p17/24trNef1 inneholder en PCR-feil som gir en R- til H-aminosyreendring 19 aminosyrer fra enden av Nef. Denne ble korrigert ved PCR-mutagenese og det korrigerte fragment kuttet med BglII og BamHI og klonet inn i BglII/BamHI-kuttet p73I-GN. (Figur 12) regenererer det korrigerte p17/p24opt/trNef-fusjonsgen nedstrøms for lowa lengde HCMV-promoter + ekson1, og oppstrøms for kanin β -globin polyadenyleringssignalet.

PCR 5' Nef fra p17/24trNef1 ved å benytte primere:

- 25 Polymerase = PWO (Roche) gjennom det hele.

Sense: S-Nef

ATGGTGGGTTTTCCAGTCACACC [SEKV. ID. Nr:21]

Antisense: ASNef-G:

- 30 GATGAAATGCTAGGC^uGGCTGTCAAACCTC [SEKV. ID. Nr:22]

Syklus: 95°C (30 sek) deretter 15 sykluser 95°C (30 sek), 55°C (30 sek), 72°C (60 sek), deretter 72°C (120 sek) og hold ved 4°C.

PCR 3' Nef fra p17/24trNef1 ved å benytte primere:

Sense: SNEF-G

GAGGTTTGACAGCCGCTAGCATTTTCATC [SEKV. ID. Nr:23]

Antisense: AStrNef

CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [SEKV. ID. Nr:24]

- 5 Syklus: 95°C (30 sek) deretter 15 sykluser 95°C (30 sek), 55°C (30 sek), 72°C (60 sek), deretter 72°C (120 sek) og hold ved 4°C.

PCR-produktene ble gelrenset og heftet sammen ved å benytte 5' (S-Nef) og 3' (AstrNef) primere.

Syklus: 95°C (30 sek) deretter 15 sykluser 95°C (30 sek), 55°C (30 sek), 72°C (60 sek), deretter 72°C (180 sek) og hold ved 4°C.

10

PCR-produktet ble PCR-renset, kuttet med BglII/BamHI og det 367 bp-fragmentet gelrenset og klonet inn i BglII/BamHI-kuttet gelrenset p73i-GN.

Eksempel 12

- 15 **Plasmid: p73i-RT w229k (inaktivert RT)**

Gen av interesse

Utvikling av et inaktivert RT-gen nedstrøms for en lowa lengde HCMV promoter + ekson 1 og oppstrøms for et kanin β -globin polyadenyleringssignal.

- 20 Som følge av betenkeligheter ved anvendelsen av en aktiv HIV RT art i en terapeutisk vaksine var inaktivering av genet ønskelig. Dette ble oppnådd ved PCR-mutagenese av RT (avledet fra P73I-GRN2) aminosyreposisjon 229 fra Trp til Lys (R7271 p1-28).

Primere:

PCR 5' RT + mutasjon ved bruk av primere:

- 25 (polymerase = PWO (Roche) gjennom det hele)

Sense: RT3-u:1

GAATTCGCGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGA
AGCTGAAACCCGGGAT [SEKV. ID. Nr:25]

Antisense: AScoRT-Trp229Lys

- 30 GGAGCTCGTAGCCCATCTTCAGGAATGGCGGCTCCTTCT [SEKV. ID. Nr:26]

Syklus:

1 x [94°C (30 sek)]

15 x [94°C (30 sek) / 55°C (30 sek) / 72°C (60 sek)]

1 x [72°C (180 sek)]

PCR-gelrens

PCR 3' RT + mutasjon ved bruk av primere:

Antisense: RT3-1:1

5 GAATTCGGATCCTTACAGCACCTTTCTAATCCCCGCACTCACCAGCTTGTCTGAC
CTGCTCGTTGCCGC [SEKV. ID. Nr:27]

Sense: ScoRT-Trp229Lys

CCTGAAGATGGGCTACGAGCTCCATG [SEKV. ID. Nr:28]

10

Syklus:

1 x [94°C (30 sek)]

15 x [94°C (30 sek) / 55°C (30 sek) /72°C (60 sek)]

1 x [72°C (180 sek)]

15 PCR- gelrens

PCR-produktene ble gelrenset og 5'- og 3'- enden av RT ble hektet sammen ved å benytte 5' (RT3-U1)- og 3' (RT3-L1)-primerne.

Syklus:

1 x [94°C (30 sek)]

20 15 x [94°C (30 sek) / 55°C (30 sek) /72°C (120 sek)]

1 x [72°C (180 sek)]

PCR-produktet ble gelrenset og klonet inn i p7313ie, ved å benytte NotI og BamHI restriksjons seter for å utvikle p73I-RT w229k. (Se Figur 13).

25 **Eksempel 13:**

Plasmid: p73i-Tgrn (#3)

Gen av interesse

p17/p24-delen av det kodonoptimaliserte gag-, kodonoptimalisert RT- og trunkerte Nef-gen fra HIV-1 clade B stamme HXB2 nedstrøms for en lowa lengde
30 HCMV promoter + ekson1 og oppstrøms for et kanin β -globin polyadenylerings-signal.

Trippel fusjonskonstruksjoner som inneholder en aktiv form av RT kan kanskje ikke aksepteres for human anvendelse av regulerende myndigheter, og inaktivering

av RT ble deretter foretatt ved insersjon av et NheI- og ApaI-kuttet fragment fra p73i-RT w229k, i en NheI/ApaI-kuttet p73i-GRN2#19 (Figur 14). Dette resulterer i en W→K endring i posisjon 229 i RT.

5 Eksempel 14

p73I-Tnrg (#16)

Gen av interesse

Det trunkerte Nef, inaktiverte kodonoptimaliserte RT og p17/p24-delen av det kodonoptimaliserte gag-gen fra HIV-1 clade B stamme HXB2 nedstrøms for en Iowa
10 lengde HCMV promoter + ekson1 og oppstrøms for et kanin β-globin poly-adenyleringssignal.

Rekkefølgen av genet i polyprotein som kodes av p73I-Tgrn ble omleiret ved PCR og PCR-sammenheking for å frembringe p73I-Tnrg (Figur 15). Hvert gen ble PCR-amplifisert og gelrenset før PCR-sammenheking av genene for å danne et
15 enkelt polyprotein. Produktet ble gelrenset, NotI/BamHI-kuttet og ligert inn i NotI/BamHI-kuttet p7313ie.

Primere:

trNef PCR

S-Nef (Not I)

20 CATTAGAGCGGCCGCGATGGTGGGTTTTCCAC [SEKV. ID. Nr:29]

AS-Nef-coRT linker

GATGGGACTGATGGGGCCCATGCAGTTCTTGAAGTACTCCGG [SEKV. ID. Nr:30]

RTw229k PCR

S-coRT

25 ATGGGCCCCATCAGTCCCATCGAG [SEKV. ID. Nr:31]

AS-coRT-p17p24 linker

CAGTACCGAAGCTCGGGCACCCATCAGCACCTTTCTAATCCCCGC [SEKV. ID. Nr:32]

p17p24opt PCR

30 S-p17p24opt

ATGGGTGCCCCGAGCTTCGGTACTG [SEKV. ID. Nr:33]

AS-p17p24opt (BamHI)

GATGGGGGATCCTCACAACACTCTGGCTTTGTGTCC [SEKV. ID. Nr:34]

PCR-betingelser for individuelle produkter og sammenheking ved å benytte VENT DNA-polymerase (NEB):

1 x [94°C (30 sek)]

25 x [94°C (30 sek) / 55°C (30 sek) /72°C (120 sek) [p17p24 eller RT] eller
60 sek [trNef]]]

1 x [72°C (240 sek)]

PCR-produktene ble gelrenset og benyttet i en PCR-sammenheking ved å benytte primerne S-trNef (NotI) og AS-p17p24opt (BamHI):

1 x [94°C (30 sek)]

25 x [94°C (30 sek) / 55°C (30 sek) /72°C (210 sek)]

1 x [72°C (240 sek)]

Produktet på 3000 bp ble gelrenset og kuttet med NotI og BamHI som ble PCR-renset og ligert inn i NotI/BamHI-kuttet gelrenset p7313ie for å frembringe p73i-Tnrg.

Eksempel 15:

1. Plasmid: P73i-Tnrg (#3)

Gen av interesse

Den trunkerte Nef, p17/p24-del av det kodonoptimaliserte gag- og inaktiverte kodonoptimaliserte RT-gen fra HIV-1 clade B stamme HXB2 nedstrøms for en Iowa lengde HCMV promoter + ekson1, og oppstrøms for et kanin β -globin polyadenyleringssignal.

Rekkefølgen av genene i det polyprotein som kodes av p73i-Tgrn ble omleiret ved PCR for å frembringe p73i-Tnrg (Figur 16). Kodonoptimalisert p17/p24 og RT ble utviklet som et enkelt produkt og PCR-hektet til amplifisert trNef. Produktet ble gelrenset, NotI/BamHI-kuttet og ligert inn i NotI/BamHI-kuttet p7313ie.

Primere:

p17/p24 - RT 3'PCR:

Sp17p24opt (sense)

ATGGGTGCCCGAGCTTCGGTACTG [SEKV. ID. Nr:35]

RT3 1:1 (antisense)

GAATTCGGATCCTTACAGCACCTTTCTAATCCCCGCACTCACCAGCTTGTGAC
CTGCTCGTTGCCGC [SEKV. ID. Nr:36]

TrNef 5'PCR
S-Nef (NotI)

CATTAGAGCGGCCGCGATGGTGGGTTTTCCAC [SEKV. ID. Nr:37]

AS-Nef-p17p24

5 CAGTACCGAAGCTCGGGCACCCATGCAGTTCTTGAAGTACTCCGG [SEKV. ID. Nr:38]

PCR-betingelser for individuelle produkter og sammenheking ved bruk av VENT DNA-polymerase (NEB):

1 x [94°C (30 sek)]

10 25 x [94°C (30 sek) / 55°C (30 sek) /72°C (180 sek [p17p24+RT] eller 60 sek [trNef] eller 210 sek [sammenheking]]]

1 x [72°C (240 sek)]

Produktet på 3000 bp ble gelrenset og kuttet med NotI og BamHI som ble PCR-renset og ligert inn i NotI/BamHI-kuttet gelrenset p7313ie for å frembringe p73i-
15 Tngr.

Eksempel 16:

Plasmid: p73I-Trgn (#6)

Gen av interesse

20 Det inaktiverte kodonoptimaliserte RT, p17/p24-delen av det kodon-optimaliserte gag- og trunkerte Nef-gen fra HIV-1 clade B stamme HXB2 nedstrøms for en lowa lengde HCMV promoter + ekson1 og oppstrøms for et kanin β-globin polyadenyleringssignal.

Rekkefølgen av genene i konstruksjonen ble oppnådd ved PCR-amplifikasjon
25 av p17p24-trNef og RTw229k fra plasmidene henholdsvis p73I-GN2 og p73I-RTw229k. PCR-sammenheking ble foretatt og produktet gelrenset og NotI/BamHI-kuttet før ligering med NotI/BamHI-kuttet p7313ie. Sekvensering viste at p17p24 ikke var fullstendig optimalisert og et 700 bp-fragment ble deretter AgeI/MunI-kuttet fra den kodende region og erstattet med MunI/Age-fragment fra p73i-Tgrn#3 inne-
30 holdende den korrekte kodende sekvens. (Se Figur 17).

Primere:

p17p24-trNef PCR

S-p17p24opt

ATGGGTGCCCGAGCTTCGGTACTG [SEKV. ID. Nr:39]

AstrNef (BamHI)

RTw229k

RT3-U:1

GAATTCGCGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGA

5 AGCTGAAACCCGGGAT [SEKV. ID. Nr:40]

AS-coRT-p17p24opt linker

CAGTACCGAAGCTCGGGCACCCATCAGCACCTTTCTAATCCCCGC [SEKV. ID. NR:41]

PCR-betingelser for individuelle produkter og sammenheking ved bruk av
10 VENT DNA-polymerase (NEB):

1 x [94°C (30 sek)]

25 x [94°C (30 sek) / 55°C (30 sek) / 72°C (120 sek (PCR) eller 180 sek
(sammenheking)

1 x [72°C (240 sek)]

15 Produktet på 3000 bp fra PCR-sammenhekingen ble gelrenset og kuttet med
NotI og BamHI, som ble PCR-renset og ligert inn i NotI/BamHI-kuttet gelrenset
p7313ie for å frembringe p73i-Tngr. Sekvensanalyse viste at p17p24-sekvensen
oppnådd fra p73I-GN2 ikke var fullstendig kodonoptimalisert og at denne var blitt
overført til det nye plasmid. Dette ble rettet ved kutting av et 700 bp-fragment fra
20 p73i-Tngr kuttet med MunI og AgeI, som ble erstattet ved ligering med et 700 bp
MunI/AgeI-oppsluttet produkt fra p73i-Tgrn for å frembringe konstruksjonen p73I-
Tngr#6.

Eksempel 17:

25 **Plasmid: p73i-Trng (#11)**

Gen av interesse

Den inaktiverte kodonoptimaliserte RT, trunkerte Nef- og p17/p24-del av det
kodonoptimaliserte gag-gen fra HIV-1 clade B stamme HXB2 nedstrøms for en lowa
lengde HCMV promoter + ekson1 og oppstrøms for et kanin β -globin poly-
30 adenyleringssignal.

Rekkefølgen av genene i konstruksjonen ble oppnådd ved PCR-amplifikasjon
av RT-trNef og p17p24-genene fra p73i-Tgrn. PCR-sammenheking av de to DNA-
fragmentene ble foretatt og produktet på 3 kb gelrenset og NotI/BamHI-kuttet før
ligering med NotI/BamHI-oppsluttet p7313ie, hvilket førte til p73I Trng (#11).

Primere:RTw229k-trNef

RT3-u:1

GAATTGCGGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGA

5 AGCTGAAACCCGGGAT [SEKV. ID. Nr:42]

AS-Nef-p17p24opt linker

CAGTACCGAAGCTCGGGCACCCATGCAGTTCTTGA ACTACTCCGG [SEKV. ID.
Nr:43]P17p24

10 S-p17p24opt

ATGGGTGCCCCGAGCTTCGGTACTG [SEKV. ID. Nr:44]

AS-p17p24opt (BamHI)

GATGGGGGATCCTCACAACACTCTGGCTTTGTGTCC [SEKV. ID. Nr:45]

PCR-betingelser for individuelle produkter og sammenheking ved å benytte

15 VENT DNA-polymerase (NEB):

1 x [94°C (30 sek)]

25 x [94°C (30 sek)/55°C (30 sek)/72°C (120 sek (PCR av gener) eller 180 sek
(sammenheking)]

1 x [72°C (240 sek)]

20 Produktet på 3000 bp fra PCR-sammenhekingen ble gelrenset og kuttet med
NotI og BamHI, som ble PCR-renset og ligert inn i NotI/BamHI-kuttet gelrenset
p7313ie for å frembringe p73i-Tngr.**Eksempel 18:**25 **p73i-Tngr (#f1)****Gen av interesse**p17/p24-delen av det kodonoptimaliserte gag-, trunkerte Nef- og
kodonoptimaliserte inaktivererte RT-gen fra HIV-1 clade B stamme HXB2 nedstrøms
for en lowa lengde HCMV promoter + ekson1, og oppstrøms for et kanin β -globin
polyadenyleringssignal.
30Rekkefølgen av genene i konstruksjonen ble oppnådd ved PCR-amplifikasjon
av p17p24-trNef og RTw229k fra plasmidene henholdsvis p73I-GN2 og
p73I-RTw229k. PCR-sammenheking ble foretatt og produktet gelrenset og

NotI/BamHI-kuttet før ligering med NotI/BamHI-oppsluttet p7313ie. To sekvensfeil ble registrert i sekvensen (p17p24 og RT) som senere ble reparert ved erstatning med korrekte deler av genene ved å benytte restriksjons seter innen polyproteinet. (Se Figur 19).

5 **Primere:**

p17p24-trNef PCR

S-p17p24opt

ATGGGTGCCCCGAGCTTCGGTACTG [SEKV. ID. Nr:46]

AS-Nef-coRTlinker

10 GATGGGACTGATGGGGCCCATGCAGTTCTTGA ACTACTCCGG [SEKV. ID. Nr:47]

RTw229k

S-coRT

ATGGGCCCCATCAGTCCCATCGAG [SEKV. ID. Nr:48]

RT3-1:1

15 GAATTCGGATCCTTACAGCACCTTTCTAATCCCCGCACTCACCAGCTTGTGCGAC
CTGCTCGTTGCCGC [SEKV. ID. Nr:49]

PCR-betingelser for individuelle produkter og sammenheking ved å benytte VENT DNA-polymerase (NEB):

1 x [94°C (30 sek)]

20 25 x [94°C (30 sek)/55°C (30 sek)/72°C (120 sek (PCR) eller 180 sek (sammen-
heking)

1 x [72°C (240 sek)]

Produktet på 3000 bp ble gelrenset og kuttet med NotI og BamHI som ble PCR-renset og ligert inn i NotI/BamHI-oppsluttet gelrenset p7313ie for å frembringe
25 p73i-Tngr. Sekvensering viste at p17p24 ikke var fullstendig optimalisert, og et
700 bp-fragment ble deretter AgeI/MunI-kuttet fra den kodende region og erstattet
med MunI/Age-fragment fra p73i-Tgrn#3 inneholdende den korrekte kodende
sekvens. Polyproteinet inneholdt også en enkelt punktmutasjon (G2609A) som
resulterte i en aminosyresubstitusjon av Thr til Ala i RT-delen av polyproteinet.

30 Mutasjonen ble korrigert ved ApaI/BamHI-oppslutning av konstruksjonen og PCR-
opprensning for å fjerne den muterte sekvens, som ble erstattet ved ligering med en
ApaI/BamHI-oppsluttet del av RT fra p73i-Tgrn.

Eksempel 19:**Fremstilling av plasmid-belagt «gull-oppslemming» for «genkanon» DNA-patroner**

Plasmid DNA (ca. 1 µg/µL), f.eks. 100 µg, og 2 µm gullpartikler, f.eks. 50 mg (PowderJect), ble suspendert i 0,05M spermidin, f.eks. 100 µL (Sigma). DNA ble utfelt på gullpartiklene ved tilsetning av 1M CaCl₂, f.eks. 100 µL (American Pharmaceutical Partners, Inc., USA). DNA/gull-komplekser ble inkubert i 10 minutter ved romtemperatur, vasket tre ganger i absolutt etanol, f.eks. 3 x 1 mL (på forhånd tørket på molekylsikt 3Å (BDH)). Prøver ble resuspendert i absolutt etanol inne-

holdende 0,05 mg/mL polyvinylpyrrolidon (PVP, Sigma) og delt i tre like alikvoter i 1,5 mL mikrosentrifugerør (Eppendorf). Alikvotene var for analyse av (a) «gull-oppslemming», (b) eluat-plasmid eluert fra (a), og (c) for fremstilling av gull/plasmid-belagte Tefzel-patroner for «genkanonen» (se Eksempel 3 nedenfor). For fremstilling av prøvene (a) og (b) ble rørene inneholdende plasmid DNA/»gull-oppslemming» i etanol/PVP sentrifugert i 2 minutter med maksimal hastighet i en Eppendorf 5418 mikrosentrifuge, hvorpå supernatanten ble fjernet og «gull-oppslemmingen» tørket i 10 minutter ved romtemperatur. Prøve (a) ble resuspendert til 0,5-1,0 µg/µL plasmid DNA i TE pH 8,0, anslått til ca. 50% belegging. For eluering ble prøve (b) resuspendert til 0,5-1,0 µg/µL av plasmid DNA i TE pH 8,0, og inkubert ved 37°C i 30 minutter, kraftig ristet og deretter sentrifugert i 2 minutter ved maksimal hastighet i en Eppendorf 5418 mikrosentrifuge og supernatanten, eluatet, tatt ut og oppbevart ved -20°C. Den eksakt eluerte DNA-konsentrasjon ble bestemt ved spektrofotometrisk kvantifisering ved å benytte Genequant II (Pharmacia Biotech).

Eksempel 20:**Fremstillinger av patroner for DNA-immunisering**

Fremstilling av patroner for Accell-genoverføringsanordningen var som tidligere beskrevet (Eisenbraun *et al.*, DNA and Cell Biology, 1993, Bd. 12, Nr. 9, s. 791-797; Pertner *et al.*). I korte trekk ble plasmid DNA belagt på 2 µm gullpartikler (DeGussa Corp., South Plainfield, N.J., USA) og fylt over i Tefzel-slange, som senere ble skåret i 1,27 cm lengde for å tjene som patroner og oppbevart i tørket tilstand ved 4°C inntil bruk. Ved en typisk vaksinasjon inneholdt hver patron 0,5 mg gull belagt med totalt 0,5 µg DNA/patron.

Eksempel 21:**Immunrespons på HIV-antigener etter DNA-vaksinasjon ved å benytte gen-kanonen**

- 5 Mus (n=3/gruppe) ble vaksinert med antigener kodet av nukleinsyre og lokalisert i to vektorer. P7077 som utnytter HCMV IE-promoter som inkluderer Intron A og ekson1 (fcmv-promoter). P73I avleverer det samme antigen, men inneholder den HCMV IE-promoter (icmv-promoter) som mangler Intron A, men inkluderer ekson 1.
- 10 Plasmidet ble avlevert til det barberte målsetet av abdominal hud av F1 (C3H x Balb/c) mus. Musene ble gitt en primær immunisering av 2 x 0,5 µg DNA på dag 0, booster doser på 2 x 0,5 µg DNA på dag 35, og cellulær respons ble påvist på dag 40 ved å benytte IFN-gamma Elispot.
- P73I - tom vektor
 - 15 - P7077 - tom vektor
 - P7077 GRN - (f CMV promoter) Gag, RT, Nef
 - P73I GRN - (i CMV promoter) Gag, RT, Nef
 - P73I GR3N - (CMV promoter) optimalisert Gag, optimalisert RT, Nef
 - 20 - P7077 GN - (f CMV promoter) Gag, Nef
 - P73I GN - (i CMV promoter) Gag, Nef

Cytotoksiske T-celleresponser

- Den cytotoksiske T-cellerespons ble bestemt ved CD8+ T-celle-begrenset
- 25 IFN-γ ELISPOT-analyse av splenocytter oppsamlet 5 dager senere. Mus ble avlivet ved cervical dislokasjon og milten oppsamlet i iskald PBS. Splenocytter ble «teased» ut i fosfat-buffret saltvann (PBS) etterfulgt av lyse av røde blodceller (1 minutt i buffer bestående av 155 mM NH₄Cl, 10 mM KHCO₃, 0,1 mM EDTA). Etter to vaskinger i PBS for å fjerne partikler, ble enkeltcellesuspensjonen delt i aliquoter i ELISPOT-
- 30 plater som på forhånd var belagt med oppfangende IFN-γ antistoff og stimulert med CD8-begrenset kognat-peptid (Gag, Nef eller RT). Etter dyrking over natten ble IFN-γ produserende celler visualisert ved anvendelse av anti-murint IFN-γ biotin-merket

antistoff (Pharmingen) etterfulgt av streptavidin-konjugert alkalisk fosfatase og kvantifisert ved å benytte avbildningsanalyse.

Resultatene av dette forsøk er vist i Figur 20, 21 og 22.

5 Eksempel 22

Immunogenitet av vaksinekonstruksjoner

1. Cellulære tester

Den cellulære immunrespons omfatter cytotoksiske CD8-celler og CD4-hjelperceller. En følsom metode for å påvise spesifikke CD8- og CD4-celler er ELISpot-testen som kan benyttes for å kvantifisere antall celler som kan utskille interferon- γ eller IL-2. ELISpot-testen baserer seg på oppfangingen av cytokiner utskilt fra individuelle celler. I korte trekk belegges spesialiserte mikrotiterplater med anti-cytokin antistoffer. Splenocytter isolert fra immuniserte dyr inkuberes over natten i nærvær av spesifikke peptider som representerer kjente epitoper (CD8) eller proteiner (CD4). Dersom celler stimuleres til frigjøring av cytokiner, vil de binde til antistoffene på den plateoverflate som omgir lokaliteten av de individuelle produserende celler. Cytokiner forblir festet til antistoff-belegget etter at cellene er lysert og platene vasket. Testen er utviklet på lignende måte som en ELISA-test ved å benytte et biotin/avidin amplifikasjonssystem. Antallet av flekker er likt antallet av cytokinproduserende celler.

CD8-responser på følgende K2^d-begrensede murine epitoper: Gag (AMQMLKETI), Nef (MTYKAAVDL) og RT (YYDPSKDLI) og CD4-responser på Gag- og RT-proteiner ble registrert for alle seks konstruksjoner. Resultatene av disse testene ble analysert statistisk og konstruksjoner rangert i henhold til deres immunogenitet. Resultatene er vist i Figur 23 av figurene.

2. Humorale tester

Blodprøver fra to forsøk ble oppsamlet for antistoffanalyse 7 og 14 dager etter booster-dosen. Serum ble fraskilt og oppbevart nedfrosset inntil antistoff-titeret kunne måles ved å benytte spesifikke ELISA-tester. Alle prøver ble testet med henblikk på antistoffer mot Gag, Nef og RT. I korte trekk ble ELISA-plater belagt med det relevante protein. Overskytende protein ble vasket av før fortynnede serumprøver ble inkubert i brønnene. Serumprøvene ble vasket av, og anti-mus antiserum konjugert

til en passende merking, ble tilsatt. Platen ble fremkalt og lest på en plateleser. Resultatene er vist i Figur 24.

3. Antistoff-data

Antistoff-titere ble målt for alle seks konstruksjoner i fire forsøk. Konstruksjon p73i-GNR ga konsekvent ingen antistoffrespons på Gag og begrensede antistoff-responser på Nef. Grunnen til dette er uklar, idet T-celleresponser ble observert fra splenocytter isolert fra den samme mus, hvilket tyder på at Gag-proteinet ble uttrykt *in vivo*.

Rangeringen for utviklingen av Gag-spesifikke antistoffer var:
RNG>GRN>NRG>RGN>NGR>GNR

Analyse av cellulære immunologi-data

Formålet var å rangere de seks konstruksjonene på grunnlag av data fra telling av flekker fra 3 immunologiforsøk. Tre sett av responser ble målt: CD8-responser på Gag, Nef og RT på Dag 7 (7 dager etter primærdose), CD4-responser på Gag og RT på Dag 35 (7 dager etter boosterdose), CD8-responser på Gag, Nef og RT på Dag 35 (7 dager etter boosterdose).

Hver respons (f.eks. CD8-respons på Gag) ble forsøkt i modell ved å benytte en modell med lineær blandet effekt i SAS-versjon 8. Modellen inkluderte faste effekter av konstruksjon, enten angjeldende antigen (Gag, Nef eller RT) var tilstede eller fraværende, og om IL-2 var tilstede eller fraværende. For CD8-responser, hvor data var tilgjengelige fra hver enkelt mus, ble dessuten individ inkludert som en tilfeldig effekt i modellen. Modellen inkluderte interaksjonsbetingelser for å ta hensyn til en ulik konstruksjonseffekt for hver kombinasjon av antigenet (tilstede/fraværende) og IL-2 (tilstede/fraværende).

Ut fra modellen ble forskjellen i justert gjennomsnittlig respons mellom hver konstruksjon og p7313 (kontrollgruppen) anslått separat for hver kombinasjon av antigen (tilstede/fraværende) og IL-2 (tilstede/fraværende), sammen med en p-verdi som indikerer hvorvidt forskjellen var statistisk signifikant. Basert på forskjellene og p-verdiene i nærvær av antigenet og i fravær av IL-2, ble konstruksjoner rangert ved å tilordne en skåre på 6 til konstruksjonen med størst forskjell, 5 til den nest største, etc., men 0 til eventuelle konstruksjoner hvor forskjellen ikke var statistisk signifikant på 5% nivået.

Forutsetningene for modellen - at restene var normalfordelt med konstant varians, ble bestemt ved å benytte grafiske metoder og sensitivitetsanalyser, hvor først en log og dernest en kvadratrots-transformasjon av responsen ble benyttet i modellen. Rangeringen av konstruksjonene var ikke følsom for avvik fra for-

5 utsetningene for modellen.

Etter beregning av rangeringene for hver respons i hvert forsøk for seg, ble total rangering for de tre settene av responser beregnet over alle tre forsøk. Den følgende tabell viser de totale rangeringer over de tre forsøkene.

10 **Totale rangeringer av konstruksjoner for hvert av tre sett av responser, kombinert over 3 immunologiske forsøk**

15	Konstruksjon	Dag 7 (7 dager etter primærdose)	Dag 35 (7 dager etter boosterdose)	
		CD8	CD4	CD8
	GRN	5	18	3
	GNR	17	24	28
	RGN	28	23	33
20	RNG	25	27	37
	NRG	25	19	0
	NGR	4	14	10

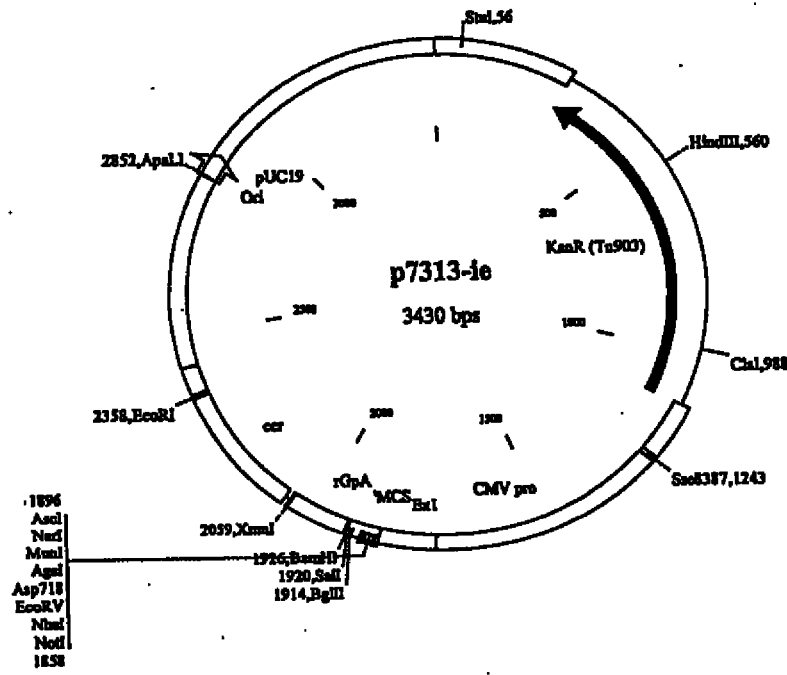
25 RNG har den høyeste rangering for begge sett av responser på Dag 35 og den nest høyeste rangering, etter RGN, på Dag 7. RGN får også høye rangeringer for begge sett av responser på Dag 35.

Patentkrav

1. Nukleotidsekvens, karakterisert ved at den omfatter en sekvens som koder for et immunogenisk fusjonsprotein som omfatter et kodon optimalisert HIV-1 gag-protein, et kodon optimalisert RT protein og et HIV-1 Nef protein, operativt koblet til en heterolog promoter, hvori gag proteinet omfatter p17 og p24, RT proteinet inneholder en mutasjon for i alt vesentlig å inaktivere en hvilken som helst revers transkriptase aktivitet, og Nef proteinet inneholder en N-terminal trunkering.
2. Nukleotidsekvens ifølge krav 1, karakterisert ved at den kodonoptimaliserte sekvens etterligner kodonanvendelsen i et høyt uttrykt humant gen som har en RSCU-verdi på 0,5.
3. Nukleotidsekvens ifølge krav 1, karakterisert ved at rekkefølgen av sekvensene er RT, gag, Nef eller RT, Nef, gag.
4. Nukleotidsekvens ifølge krav 1, karakterisert ved at den heterologe promoter er promoteren fra HCMV IE-gen.
5. Nukleotidsekvens ifølge krav 4, karakterisert ved at 5' av promoteren omfatter ekson 1.
6. Nukleotidsekvens ifølge krav 1, karakterisert ved at RT er mutert ved å erstatte lysin med tryptofan 229.
7. Vektor, karakterisert ved at den omfatter en nukleotidsekvens ifølge krav 1.
8. Vektor ifølge krav 7, karakterisert ved at den er en viral vektor.
9. Vektor ifølge krav 8, karakterisert ved at den er et replikasjonsdefekt adenovirus.

10. Vektor ifølge krav 9, k a r a k t e r i s e r t v e d a t den er et dobbeltkjedet DNA-plasmid.
11. Protein, k a r a k t e r i s e r t v e d a t det kodes for av en nukleotidsekvens ifølge krav 1.
12. Farmasøytisk preparat, k a r a k t e r i s e r t v e d a t det omfatter en nukleotidsekvens ifølge krav 1 eller en vektor ifølge krav 7 og et farmasøytisk akseptabelt hjelpestoff, fortynningsmiddel, bærer eller adjuvans.
13. Farmasøytisk preparat ifølge krav 12, k a r a k t e r i s e r t v e d a t det er tilpasset intramuskulær eller intradermal administrering.
14. Farmasøytisk preparat ifølge krav 12, k a r a k t e r i s e r t v e d a t bæreren er en gullkule.
15. Intradermal avleveringsanordning, k a r a k t e r i s e r t v e d a t den omfatter et farmasøytisk preparat ifølge krav 12.
16. Fremgangsmåte for fremstilling av en nukleotidsekvens ifølge krav 1, k a r a k t e r i s e r t v e d a t den omfatter operativ kobling av en nukleotidsekvens som koder for et HIV-1 gag-protein eller fragment derav og et HIV-1 Nef-protein eller fragment derav, til en heterolog promotersekvens.

Figur 1



2/50

Figur 2

Sekvens av p55 gag-innskudd i pGagOptpr2

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
 AGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
 AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
 CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
 TCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
 ATTGTCCAAAACATTCAGGGCCAGATGGTTCATCAGGCCATCAGCCCCCGGACGCTC
 AATGCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
 TTCTCCGCTTTGAGTGAGGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
 GTGGGCGGCCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
 GCCGAGTGGGACAGAGTGCATCCCGTCCAGCTGGCCCAATCGCGCCCGGACAGATG
 CGGGAGCCTCGCGGCTCTGACATTGCCCGCACCACTCTACACTGCAAGAGCAAATC
 GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAAATCTATAACGGTGGATC
 ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTtagggattacgtcgaccggTTTTATAAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAAGTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCCTGTGAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
 GTGTTGGCCGAAGCCATGAGCCAGGTGACGAACTCCGCAACCATCATGATGCAGAGA
 GGGAACTTCCGCAATCAGCGGAAGATCGTGAAGTGTTCATTGCGGCAAGGAGGGT
 CATACCGCCCCGCAACTGTGCGGGCCCCCTAGGAAGAAAGGGTGTGGAAGTGCGGCAAG
 GAGGGACACCAGATGAAAGACTGTACAGAACGACAGGCCAATTTTCTTGAAAGATT
 TGGCCGAGCTACAAGGGGAGACCTGGTAATTTCTGCAAAGCAGGCCCGAGCCCACC
 GCCCCCCTGAGGAATCCTTCAGGTCCGGAGTGGAGACCACAACGCCTCCCCAAAAA
 CAGGAACCAATCGACAAGGAGCTGTACCCTTTAACTTCTCTGCGTTCTCTCTTTGGC
 AACGACCCGTCGTCTCAATAA [SEKV. ID. NR:50]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNITMLNT VGGHQAAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPK EPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQNANPDC KTILKALGPA ATLEEMMTAC
 QGVGGPGHKA RVLAEAMSQV TNSATIMMQR GNFRNQRKIV KCFNCGKEGH
 TARNCRAPRK KGCWKCGKEG HQMKDCTERQ ANFLGKIWPS YKGRPGNFLQ
 SRPEPTAPPE ESFRSGVETT TPPQKQEPID KELYPLTSLR SLFGNDPSSQ

*

[SEKV. ID. NR:51]

3/50

Figur 3**Sekvens av p17/24trNEF-innskuddet i p17/24trNEF1**

ATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGATGGGAAAAAATT
 CGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGC
 AGGGAGCTAGAACGATTTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGT
 AGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGA
 TCATTATATAATACAGTAGCAACCTCTATTGTGTGCATCAAAGGATAGAGATAAAA
 GACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAA
 GCACAGCAAGCAGCAGCTGACACAGGACACAGCAATCAGGTCAGCCAAAATTACCCT
 ATAGTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTA
 AATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTTCAGCCCAGAAGTGATACCCATG
 TTTTCAGCATTATCAGAAGGAGCCACCCACAAGATTTAAACACCATGCTAAACACA
 GTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCT
 GCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATG
 AGAGAACCAAGGGGAAGTGACATAGCAGGAACACTAGTACCCTTCAGGAACAAATA
 GGATGGATGACAAATAATCCACCTATCCCAGTAGGAGAAATTTATAAAAGATGGATA
 ATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATA
 AGACAAGGACCAAAAGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAACTCTA
 AGAGCCGAGCAAGCTTCACAGGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTC
 CAAAATGCGAACCCAGATTGTAAGACTATTTTAAAAGCATTTGGGACCAGCGGCTACA
 CTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGA
 GTTTTGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAG
 GCAGCTGTAGATCTTAGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATT
 CACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTAC
 TTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTT
 GGATGGTGTCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAA
 GGAGAGAACACCAGCTTGTTACACCTGTGAGCCTGCATGGGATGGATGACCCGGAG
 AGAGAAGTGTTAGAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCCCGA
 GAGCTGCATCCGGAGTACTTCAAGAACTGCTGA [SEKV. ID. NR:52]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQA AA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTMLNT VGGHQAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPKEPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQNANPDC KTILKALGPA ATLEEMMTAC
 QGVGGPGHKA RVLVGFPVTP QVPLRPMYK AAVDLSHFLK EKGGLLEGLIH
 SQRRQDILD L WIYHTQGYFP DWQNYTPGPG VRYPLTFGWC YKLVPVEPDK
 VEEANKGENT SLLHPVSLHG MDDPEREVLE WRFDShLAFH HVARELHPEY
 FKNC*
 [SEKV. ID. NR:53]

4/50

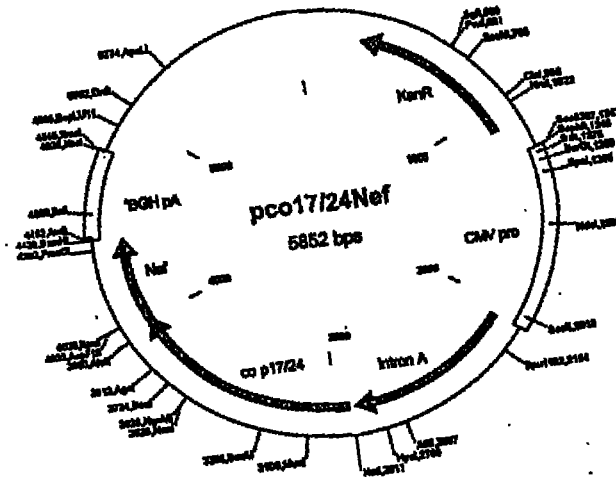
Figur 4

**Sekvens av p17/24opt/trNef-innskuddet i
p17/24opt/trNef1**

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
 AGGCTGCGCCCCGGGAGGCCAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
 AGGGAGCTTGAACGGTTTGGCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
 CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
 TCCTTGTATAACACAGTGGCTACCCCTCTACTGCGTACACCAGAGGATCGAGATTAAAG
 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
 ATTGTCCAAAACATTCAAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
 AATGCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
 TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
 GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
 CGGGAGCCTCGCGGCTCTGACATTGCCCGGCACCACCTCTACACTGCAAGAGCAAATC
 GGATGGATGACCAACAATCCTCCCATCCAGTTGGAGAAATCTATAAACGGTGGATC
 ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCCGGTTTTATAAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
 GTGTTGATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTAC
 AAGGCAGCTGTAGATCTTAGCCACTTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTA
 ATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACAAGGC
 TACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACC
 TTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAAT
 AAAGGAGAGAACACCAGCTTGTTACACCTGTGAGCCTGCATGGGATGGATGACCCG
 GAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCC
 CGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGA [SEKV. ID. NR:54]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTLNT VGGHQAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMR PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPK EPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQNPDC KTILKALGPA ATLEEMMTAC
 QGVGGPGHKA RVL MVGFPVT PQVPLRPMTY KAAVDLSHFL KEKGGLEGLI
 HSQRRQDILD LWIYHTQGYF PDWQNYTPGP GVMRYPLTFGW CYKLPVPEPD
 KVEEANKGEN TSL LHPVSLH GMDDPEREVL EWRFD SHLAF HHVARELHPE
 YFKNC*

[SEKV. ID. NR:55]



6/50

Figur 5

Sekvens av RT-innskudd av p7077-RT3:

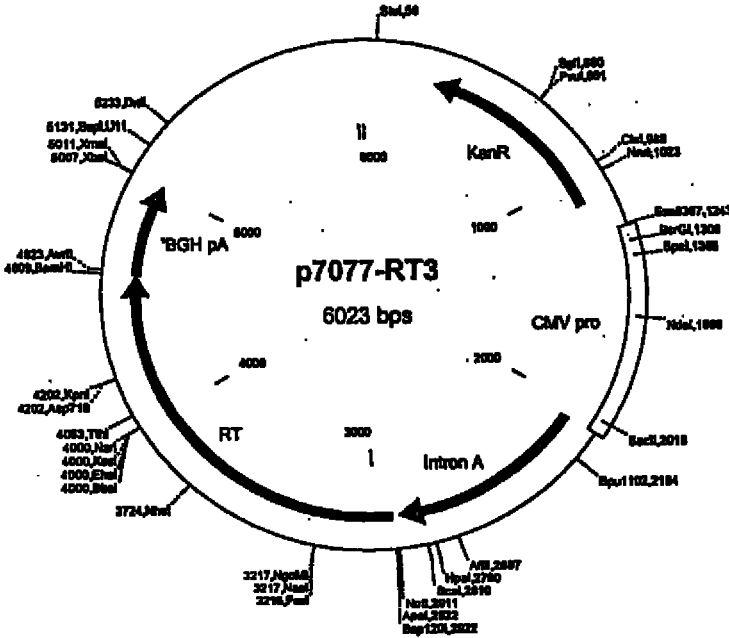
ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC
 GGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAG
 ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCCA
 TACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG
 GTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC
 ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC
 GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC
 ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCC
 CAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAG
 CCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGACCTGTAC
 GTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTGAGGCAG
 CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA
 TTCTGTGGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGACGCTATCGTC
 CTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTC
 AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG
 CGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCTCGAG
 CTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGTGTACTATGACCCC
 TCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATAACAG
 ATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGC
 GCCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAG
 TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC
 TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTC
 GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTG
 GGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGGGGAAG
 GCCGGGTACGTACCAACCGGGGCCGCCAGAAGGTCGTACCCCTGACCGACACCACC
 AACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAG
 GTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTCAGGCCCCAGCCGGAC
 CAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAG
 GTCTACCTCGCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAGGTGAC
 AAGCTGGTGAAGTGGGGGATTAGAAAGGTGCTGTAA [SEKV. ID. NR:56]

MGPISPIETV	SVKCLKPGMDG	PKVKQWPLTE	EKIKALVEIC	TEMEKEGKIS
KIGPENPYNT	PVFAIKKKDS	TKWRKLVDFR	ELNKRTQDFW	EVQLGIPHPA
GLKKKKSFTV	LDVGDAYFSV	PLDEDFRKYT	AFTIPSINNE	TPGIRYQYNV
LPQGWKGSFA	IFQSSMTKIL	EPFRKQNPDI	VIYQYMDDLY	VGSLEIGQH
RTKIEELRQH	LLRWGLTTPD	KKHQKEPPFL	WMGYELHPDK	WTVQPIVLPE
KDSWTVNDIQ	KLVGKLNWAS	QIYPGIKVRQ	LCKLLRGTKA	LTEVIPLTEE
AELELAENRE	ILKEPVHGVY	YDPSKDLIAE	IQKQGQGWWT	YQIYQEPFKN
LKTGKYARMR	GAHTNDVKQL	TEAVQKITTE	SIVIWGKTPK	FKLPIQKETW
ETWWTEYWQA	TWIPEWBFVN	TPPLVKLWYQ	LEKEPIVGAE	TFYVDGAANR
ETKLKAGYV	TNRGRQKVVT	LTDTTNQKTE	LQAIYLALQD	SGLEVNIVTD
SOYALGIIQA	QPDQSESELV	NQIIEQLIKK	EKVYLAWVPA	HKGIGGNEQV

7/50

DKLVSAGIRK VL*

[SEKV. ID. NR:57]



8/50

Figur 6

Sekvens av det kodende innskudd i p73i-RT3:

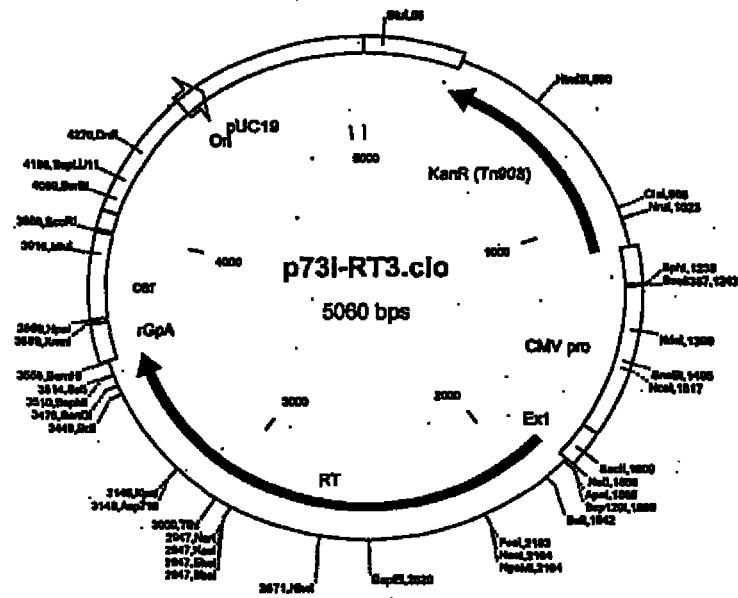
ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC
 GGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAG
 ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCCA
 TACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG
 GTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC
 ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC
 GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC
 ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCC
 CAGGGCTGGAAGGGCTCTCCCGCCATTTTCAGAGCTCCATGACCAAGATCCTGGAG
 CCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGACCTGTAC
 GTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTGAGGCAG
 CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA
 TTCCTGTGGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGCAGCCTATCGTC
 CTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTC
 AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG
 CGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCTCGAG
 CTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGTGTACTATGACCCC
 TCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAG
 ATTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGC
 GCCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAG
 TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC
 TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTC
 GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTG
 GGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGGGAA
 GGCCGGGTACGTCACCAACCGGGGCCCGCAGAAGGTGCTCACCTGACCGACACCAC
 CAACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGA
 GGTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTAGGCCCAGCCGGA
 CCAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAA
 GGTCTACCTCGCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAGGTGCA
 CAAGCTGGTGAAGTGGGGGATTAGAAAGGTGCTGTAA [SEKV. ID. NR:58]

MGPISPIETV SVKLKPGMDG PKVKQWPLTE EKIKALVEIC TEMEKEGKIS
 KIGPENPYNT PVFAIKKKDS TKWRKLVDFR ELNKRTOQDFW EVQLGIPHPA
 GLKKKKSVM LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV
 LPQGWKGSPA IFQSSMTKIL EPFRKQNPDI VIYQYMDDLY VGSLEIGQH
 RTKIEELRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLPE
 KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEE
 AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGQGWY YQIYQEPFKN
 LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW
 ETWWTEYWQA TWIPEWEFVN TPPLVKLWYQ LEKEPIVGAE TFIYVDAANR
 ETKLGKAGYV TNRGRQKVV LTDTTNQKTE LQAIYLLALQD SGLEVNIVTD
 SQYALGIIQA QPDQSESELV NQIIEQLIKK EKVYLAWVPA HKGIGGNEQV

9/50

DKLVSAGIRK VL*

[SEKV. ID. NR:59]



10/50

Figur 7**Sekvens av Nef-innskudd i 7077trNef20**

ATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCA
GCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCAC
TCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTC
CCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGA
TGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGA
GAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGA
GAAGTGTTAGAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCCCCGAGAG
CTGCATCCGGAGTACTTCAAGAACTGCTGA [SEKV. ID. NR:60]

MVGFPVTPQV PLRPMTYKAA VDLSHFLKEK GGLEGLIHSQ RRQDILDLWI
YHTQGYFPDW QNYTPGPGVR YPLTFGWYK LVPVEPKVE EANKGENTSL
LHPVSLHGMD DPEREVLEWR FDSVLAFHHV ARELHPEYFK NC*
[SEKV. ID. NR:61]

Figur 8

11/50

Sekvens av RT-innskudd i 7077RT8

ATGGGCCCCATTAGCCCTATTGAGACTGTGTCAGTAAAATTAAAGCCAGGAATGGAT
 GGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAATAAAAGCATTAGTAGAA
 ATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCCTGAAAATCCA
 TACAATACTCCAGTATTTGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAATTA
 GTAGATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGA
 ATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGT
 GATGCATATTTTTTCAGTTCCTTAGATGAAGACTTCAGGAAATATACTGCATTTACC
 ATACCTAGTATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCA
 CAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAG
 CCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTAT
 GTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTGAGACAA
 CATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAACATCAGAAAGAACCTCCA
 TTCCTTTGGATGGGTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTG
 CTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGGAAATTG
 AATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAACTCCTT
 AGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAA
 CTGGCAGAAAAACAGAGAGATTCTAAAAGAACCACTACATGGAGTGTATTATGACCCA
 TCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAA
 ATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAAGAATGAGGGGT
 GCCCACACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAAAAATAACCACAGAA
 AGCATAGTAATATGGGGAAAGACTCCTAAATTTAAACTGCCCATACAAAAGGAAACA
 TGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTTGAGTGGGAGTTT
 GTTAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCCTAGTA
 GGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAA
 GCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAACTGACACAACA
 AATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAA
 GTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGAT
 CAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAGGAAAAG
 GTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGAT
 AAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTTTAGATTAA

[SEKV. ID. NR:62]

MGPISPIETV SVKLKPGMDG PKVKQWPLTE EKIKALVEIC TEMEKEGKIS
 KIGPENFYNT PVFAIKKKDS TKWRKLVDFR ELNKRTQDFW EVQLGIPHPA
 GLKKKKS MTV LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV
 LPQGWKGSPA IFQSSMTKIL EPFRKQNPDI VIYQYMDDLV VGS DLEIGQH
 RTKIEELRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLPE
 KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEE
 AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGQGWY YQIYQEPFKN
 LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW
 ETWWTEYWQA TWIPEWEFVN TPPLVKLWYQ LEKEPIVGAE TIFYVDGAANR
 ETKLGKAGYV TNRGRQKVVV LTDTTNQKTE LQAIYLLALQD SGLEVNIVTD
 SQYALGIIQA QPDQSESELV NQIIEQLIKK EKVYLAWVPA HKGIGGNEQV
 DKLVSA GIRK VLFLD*

12/50

[SEKV. ID. NR:63]

Figur 9

Sekvens av p17/24opt/RT7trNef-innskuddet i p17/24opt/RT/trNef13

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
 AGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
 AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
 CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
 TCCTTGTATAACACAGTGGCTACCTCTACTGCGTACACCAGAGGATCGAGATTAAG
 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
 ATTGTCCAAAACATTGAGGGCCAGATGGTTCATCAGGCCATCAGCCCCGGACGCTC
 AATGCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
 TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
 GTGGGCGGCCATCAGGCCGCCATGCAATGTTGAAGGAGACTATCAACGAGGAGGCA
 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
 CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
 GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
 ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCGGTTTTATAAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCCGGACCCGGACACAAAGCCAGA
 GTGTTGATGGGCCCCATTAGCCCTATTGAGACTGTGTGAGTAAATTAAGCCAGGA
 ATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAGCATTA
 GTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAA
 AATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGA
 AAATTAGTAGATTTTCAAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAA
 TTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGAT
 GTGGGTGATGCATATTTTTAGTTCCCTTAGATGAAGACTTCAGGAAATATACTGCA
 TTTACCATACCTAGTATAAACAAATGAGACACCAGGGATTAGATATCAGTACAATGTG
 CTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATC
 TTAGAGCCTTTTAGAAAAACAAATCCAGACATAGTTATCTATCAATACATGGATGAT
 TTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTG
 AGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAA
 CCTCCATTCTTTGGATGGGTATGAACTCCATCCTGATAAATGGACAGTACAGCCT
 ATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGG
 AAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAA
 CTCCTTAGAGGAACCAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAG
 CTAGAACTGGCAGAAAACAGAGAGATTCTAAAAGAACAGTACATGGAGTGTATTAT
 GACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACA
 TATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAAGAATG
 AGGGGTGCCACACTAATGATGTAAACAATTAACAGAGGCAGTGCAAAAAATAACC
 ACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAAACTGCCCATACAAAAG
 GAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTGAGTGG

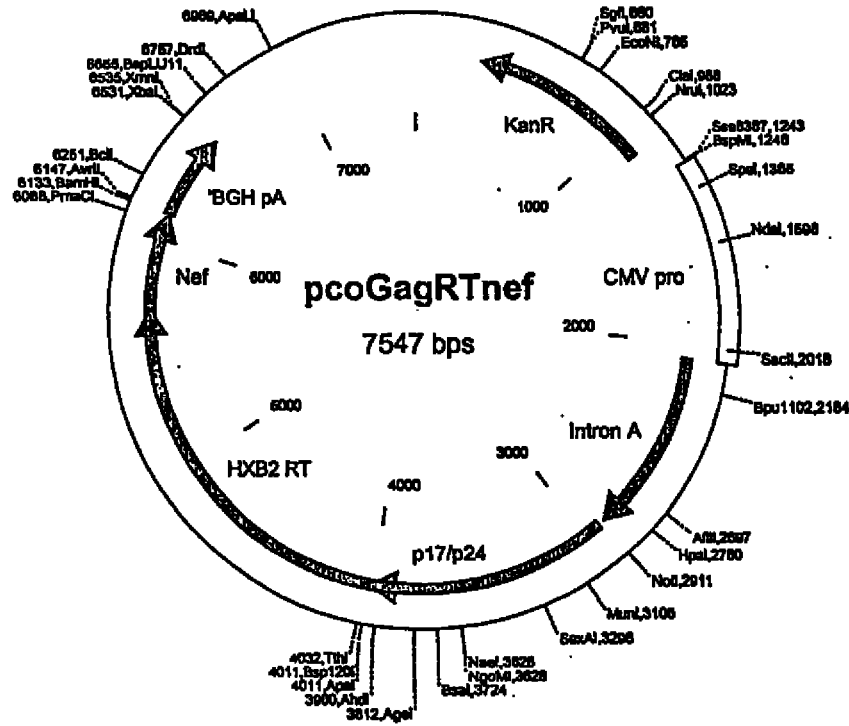
14/50

GAGTTTGTTAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCC
 ATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTA
 GGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCCTAACTGAC
 ACAACAAATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGA
 TTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAA
 CCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAG
 GAAAAGGTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAA
 GTAGATAAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTTTAGATATGGTGGGT
 TTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGAT
 CTTAGCCACTTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGA
 AGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGG
 CAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTAC
 AAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACC
 AGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTA
 GAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCG
 GAGTACTTCAAGAACTGCTGA [SEKV. ID. NR:64]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LXHIVWASRE LERFAVNPGL
 LETSEGCQRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTLNT VGGHQAAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMR PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSP SILDIRQGP EPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQANPDC KTIKALGPA ATLEEMMTAC
 QGVGGPGHKA RVLMPISPI ETVSVKLKP MDGPKVKQWP LEEKIKALV
 BICTEMEKEG KISKIGPENP YNTPVFAIK KDSTKWRKLV DFRELNKRTQ
 DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY FSVPLDEDFR KYTAFTIPSI
 NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQN PDIVYQYMD
 DLYVGSLEI GQHRTKIEEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH
 PDKWTVQPIV LPEKDSWTVN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG
 TKALTEVIPL TEEAELELAE NREILKEPVH GVYYDPSKDL IAEIQKQGQG
 QWTYQIYQBP FKNLKTGKYA RMRGAHTNDV KQLTEAVQKI TTESIWIWGK
 TPKFKLPIQK ETWETWTEY WQATWIPWE FVNTPLVLK WYQLEKEPIV
 GAETFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDITNQ KTELQAIYLA
 LQDSGLEVNI VTDSQYALGI IQAQPQSES ELVNQIEQL IKKEKVYLAW
 VPAHKGIGGN EQVDKLVSAG IRKVLFLDMV GFPVTPQVPL RPTYKAAVD
 LSHFLKEKGG LEGLIHSQRR QDILDWYH TQGYFPDWQN YTPGPGVRYP
 LTFGWCYKLV PVEPDKVEEA NKGENTSLLH PVSLHGMDDP EREVLEWRFD
 SHLAFHHVAR ELHPEYFKNC *

[SEKV. ID. NR:65]

15/50



16/50

Figur 10

Sekvens av det p17/p24opt(cor)/RT/trNef-kodende innskudd i WRG7077:

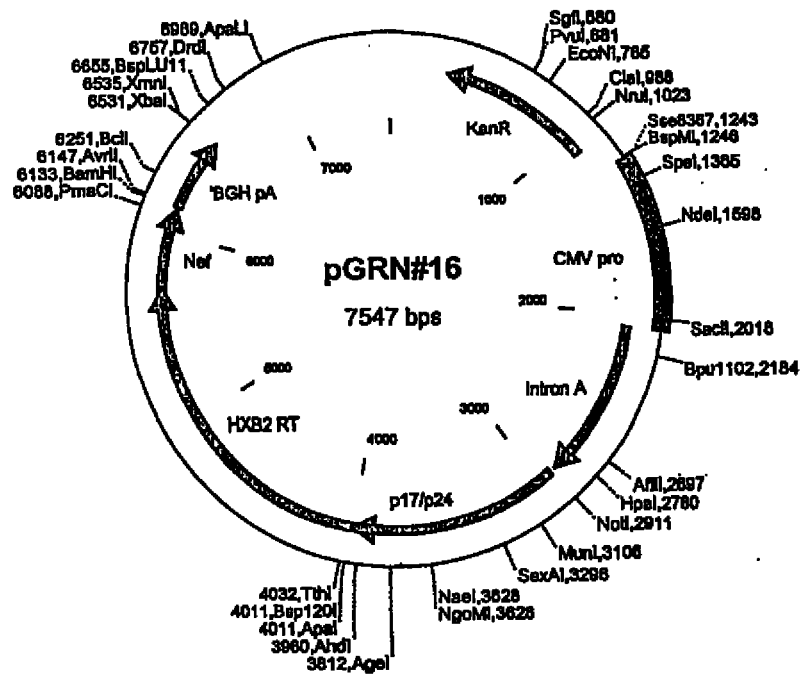
ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
 AGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
 AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
 CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
 TCCTTGATATAACACAGTGGCTACCTCTACTGCGTACACCAGAGGATCGAGATTAAG
 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
 ATTGTCCAAAACATTGAGGGCCAGATGGTTCATCAGGCCATCAGCCCCCGGACGCTC
 AATGCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCATG
 TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
 GTGGGCGGCCATCAGGCCGCCATGCAATGTTGAAGGAGACTATCAACGAGGAGGCA
 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
 CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
 GGATGGATGACCAACAATCCTCCCATCCAGTTGGAGAAATCTATAAACGGTGGATC
 ATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCGGTTTTATAAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCTGTGAGGAGTAGGCGGACCCGGACACAAAGCCAGA
 GTGTTGATGGGCCCCATTAGCCCTATTGAGACTGTGTGAGTAAATTAAGCCAGGA
 ATGGATGGCCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAATAAAAGCATTAA
 GTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAA
 AATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGA
 AAATTAGTAGATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAA
 TTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGAT
 GTGGGTGATGCATATTTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAATATACTGCA
 TTTACCATACCTAGTATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTG
 CTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAATC
 TTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGAT
 TTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTG
 AGACAACATCTGTTGAGGTGGGGACTTACCACACAGACAAAAACATCAGAAAGAA
 CCTCCATTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCT
 ATAGTGCTGCCAGAAAAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGG
 AAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAA
 CTCCTTAGAGGAACCAAGCACTAACAGAAGTAATACCACCTAACAGAAGAAGCAGAG
 CTAGAACTGGCAGAAAACAGAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTAT
 GACCCATCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACA
 TATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAAGAATG
 AGGGGTGCCCACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAAAAATAACC
 ACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAAACTGCCCATACAAAAG
 GAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTGAGTGG
 GAGTTTGTTAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCC

17/50

ATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTA
 GGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTACCCTAACTGAC
 ACAACAAATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTGGGA
 TTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAGCACAA
 CCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAG
 GAAAAGGTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAA
 GTAGATAAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTATAGATATGGTGGGT
 TTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGAT
 CTTAGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACCTCCCAAAGA
 AGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGG
 CAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACCTTTGGATGGTGCTAC
 AAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACC
 AGCTTGTTACACCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTA
 GAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCG
 GAGTACTTCAAGAACTGCTGA [SEKV. ID. NR:66]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCROI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEBQNK SKKKAQQA AA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTLNT VGGHQAAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMR PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSP SILDIRQGPB EPFRDYVDRF
 YKTLRABQAS QEVKNWMTET LLVQANANPDC KTIKALGPA ATLEEMMTAC
 QGVGGPGHKA RVLMPISPI ETVSVKLKP MDGPKVKQWP LTEEKIKALV
 EICTEMEKEG KISKIGPENP YNTPVFAIKK KDSTKWRKLV DFRELNKRTQ
 DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY PSVPLDEDFR KYTAFTIPSI
 NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQN PDIVYQYMD
 DLYVGSLEI GQHRTKIEEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH
 PDKWTVQPIV LPEKDSWTVN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG
 TKALTEVIPL TEEAELELAE NREILKEPVH GVEYDPSKDL IAEIQKQGQG
 QWTYQIYQEP FKNLKTGKYA RMRGAHTNDV KQLTEAVQKI TTESIWIWK
 TPKFKLPIQK ETWETWWTEY WQATWIPEWE FVNTPLVLK WYQLEKEPIV
 GAETFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDTTNQ KTELQAIYLA
 LQDSGLEVNI VTDSQYALGI IQAQPDQSES ELVNQIIEQL IKKEKVYLA
 VPAHKGIGGN EQVDKLVSAG IRKVLFLDMV GFPVTPQVPL RPMTYKAAVD
 LSHFLKEKGG LEGLIHSQRR QDILDWYH TQGYFPDWQN YTPGPGVRY
 LTFGWCYKLV PVEPDKVEEA NKGENTSLLH PVSLHGMDDP EREVLEWRFD
 SHLAFHHVAR ELHPEYFKNC *

[SEKV. ID. NR:67]



19/50

Figur 11**Sekvens av det p17/p24(opt)/RT(opt)trNef-innskudd i p73i-GRN2:**

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGA
 GAAAATTAGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCAT
 ATCGTGTGGGCCTCGAGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCT
 GCTGGAAACATCTGAGGGATGTCGCCAGATCCTGGGGCAATTGCAGCCAT
 CCCTCCAGACCGGGAGTGAAGAGCTGAGGTCCTTGTATAACACAGTGGCT
 ACCCTCTACTGCGTACACCAGAGGATCGAGATTAAAGGATACCAAGGAGGC
 CTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAGGCCCAG
 CAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCC
 TATTGTCCAAAACATTCAGGGCCAGATGGTTCATCAGGCCATCAGCCCCC
 GGACGCTCAATGCCTGGGTGAAGGTTGTCGAAGAGAAGGCCTTTTCTCCT
 GAGGTTATCCCCATGTTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGAC
 CTCAATACAATGCTTAATACCGTGGGCGGCCATCAGGCCGCCATGCAAAT
 GTTGAAGGAGACTATCAACGAGGAGGCAGCCGAGTGGGACAGAGTGCAT
 CCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATGCGGGAGCCTCGCGG
 CTCTGACATTGCCGGCACCACTCTACACTGCAAGAGCAAATCGGATGGA
 TGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
 ATCCTGGGCCTGAACAAGATCGTGCGCATGTA CTCTCCGACATCCATCCTT
 GACATTAGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCGGTT
 TTATAAGACCCTGCGAGCAGAGCAGGCCTCTCAGGAGGTCAAAA ACTGGA
 TGACGGAGACACTCCTGGTACAGAACGCTAACCCCGACTGCAAAACAATC
 TTGAAGGCACTAGGCCCGGCTGCCACCCTGGAAGAGATGATGACCGCCTG
 TCAGGGAGTAGGCGGACCCGGACACAAAGCCAGAGTGTTGATGGGCCCC
 ATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGACGG
 CCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTG
 GTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCG
 GGCCTGAGAACCCATACAACACCCCCGTGTTTGCCATCAAGAAGAAGGAC
 AGCACCAAGTGGCGCAAGCTGGTGGATTTCCGGGAGCTGAATAAGCGGAC
 CCAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCATCCGGCCGGCCTGA
 AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCTTACTTCAGC
 GTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACCATCCCATCT

20/50

ATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCCCA
GGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCC
TGGAGCCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATG
GACGACCTGTACGTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAA
GATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCGG
ACAAGAAGCATCAGAAGGAGCCGCCATTCTGTGGATGGGCTACGAGCTC
CATCCCGACAAGTGGACCGTGCAGCCTATCGTCCTCCCCGAGAAGGACAG
CTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGCTA
GCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTGCGC
GGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCT
CGAGCTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGCACGGCGTGT
ACTATGACCCCTCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAG
GGGCAGTGGACATAACCAGATTTACCAGGAGCCTTTCAAGAACCTCAAGAC
CGGCAAGTACGCCCCGATGAGGGGCGCCACACCAACGATGTCAAGCAG
CTGACCGAGGCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGG
GAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACCTGGGAGACGT
GGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTCGTG
AATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCAT
CGTGGGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACA
AAGCTCGGGAAGGCCGGGTACGTCACCAACCGGGGCCCGCAGAAGGTCG
TCACCCTGACCGACACCACCAACCAGAAGACGGAGCTGCAGGCCATCTAT
CTCGCTCTCCAGGACTCCGGCCTGGAGGTGAACATCGTGACGGACAGCCA
GTACGCGCTGGGCATTATTCAGGCCAGCCGGACCAAGTCCGAGAGCGAAC
TGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTC
GCCTGGGTCCCGGCCCATAGGGCATTGGCGGCAACGAGCAGGTGACAA
GCTGGTGAGTGCGGGGATTAGAAAGGTGCTGATGGTGGGTTTTCCAGTCA
CACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTT
AGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATCACTCCCA
AAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACT
TCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTG
ACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGA
AGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGC

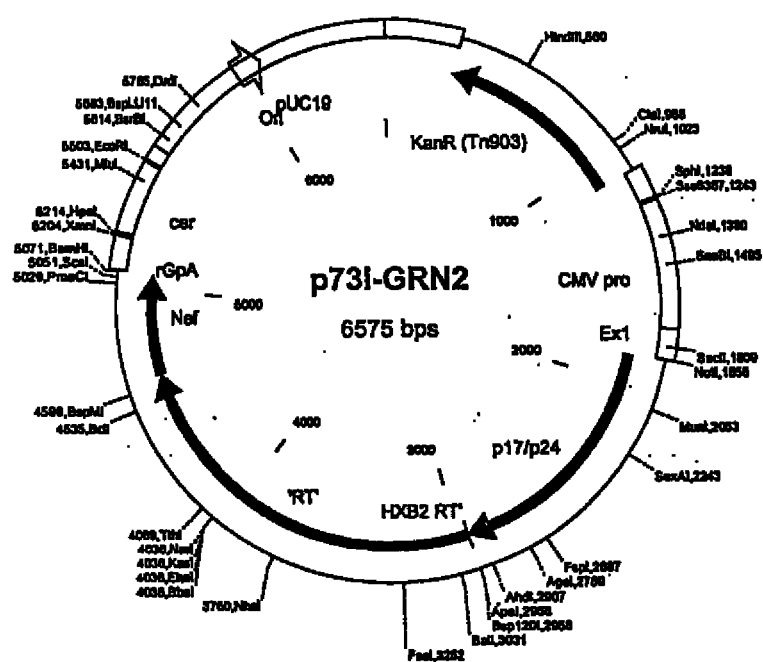
21/50

ATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAG
 CCGCCTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCA
 AGAACTGCTGA [SEKV. ID. NR:68]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTMLNT VGGHQAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPKEPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQANANPDC KTILKALGPA ATLEEMMTAC
 QGVGGPGHKA RVLMPISPPI ETVSVKLPKG MDGPKVKQWP LTEEKIKALV
 EICTEMEKEG KISKIGPENP YNTPVFAIKK KDSTKWRKLV DFRBLNKRTQ
 DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY FSVPLDEDFR KYTAFTIPSI
 NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQN PDIVIQYMD
 DLYVGSLEI GQHRTKIEEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH
 PDKWTVQPIV LPEKDSWTVN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG
 TKALTEVIPL TEEAELELAE NREILKEPVH GVYYDPSKDL IAEIQKQGGQ
 QWTYQIYQEP FKNLKTGKYA RMGAHTNDV KQLTEAVQKI TTESIVIWGK
 TPKFKLPIQK ETWETWTEY WQATWIPEWE FVNTPLVLK WYQLEKEPIV
 GAETFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDTTNQ KTELQAIYLA
 LQDSGLEVNI VTDSQYALGI IQAQPDSSES ELVNQIIEQL IKKEKVYLAW
 VPAHKGIGGN EQVDKLVSAG IRKVLVVGFP VTPQVPLRPM TYKAAVDLSH
 FLKEKGGLEG LIHSQRRQDI LDLWIYHTQG YFPDWQNYTP GPGVRYPLTF
 GWCYKLVPE PDKVEEANKG ENTSLLHPVS LHGMDDPERE VLEWRFDSRL
 AFHHVARELH PEYFKNC*

[SEKV. ID. NR:69]

22/50



23/50

Figur 12

Sekvens av p17/p24opt/trNef-innskuddet i p73i-GN2:

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
 AGGCTGCGCCCGGGAGGCCAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
 AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
 CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
 TCCTTGTATAACACAGTGGCTACCCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
 ATTGTCCAAAACATTTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
 AATGCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
 TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
 GTGGGCGGCCATCAGGCCGCCATGCAAAATGTTGAAGGAGACTATCAACGAGGAGGCA
 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
 CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
 GGATGGATGACCAACAATCCTCCCATCCAGTTGGAGAAATCTATAAACGGTGGATC
 ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTTATAAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCCTGTGAGGAGTAGGCGGACCCGGACACAAAGCCAGA
 GTGTTGATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTAC
 AAGGCAGCTGTAGATCTTAGCCACTTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTA
 ATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGC
 TACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACC
 TTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAAT
 AAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCG
 GAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCTAGCATTTCATCACGTGGCC
 CGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGA [SEKV. ID. NR:70]

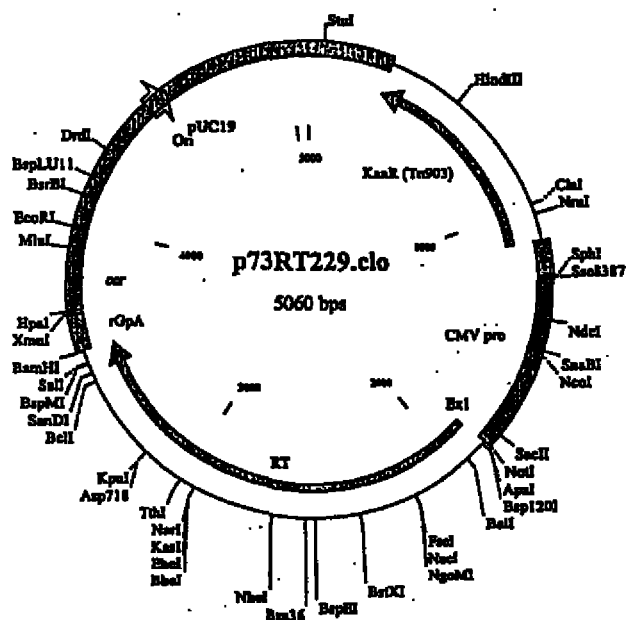
MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNTMLNT VGGHQAAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMR PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPK EPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQANANPDC KTIKALGPA ATLEEMMTAC
 QVGGGPGHKA RVL MVGFPVT PQVPLRPMTY KAAVDLSHFL KEKGGLEGLI
 HSQRRQDILD LWIYHTQGYF PDWQNYTPGP GVRYP LTFGW CYKLVPVEPD
 XVEEANKGEN TSL LHPVSLH GMDDPEREVL EWRFD SRLAF HHVARELHPE
 YFKNC* [SEKV. ID. NR:71]

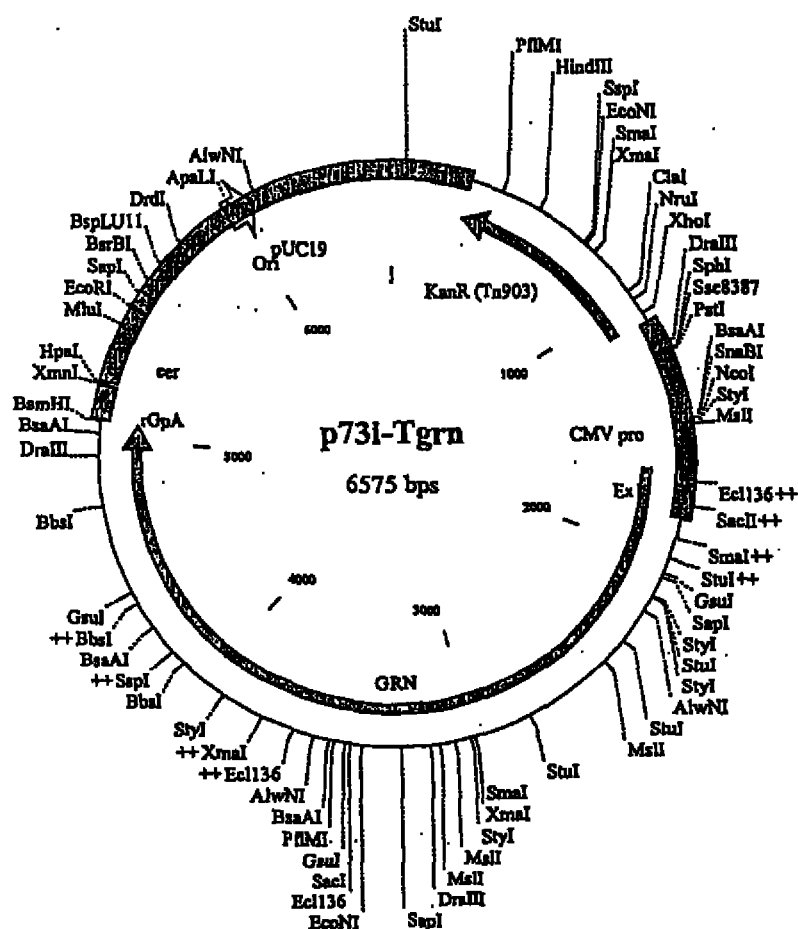
25/50

Figur 13

Sekvens av RT w229k:

ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGACGGCCC
 CAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCG
 AGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCCATACAACACCCCCGTG
 TTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGATTTCCGGGAGCTGAA
 TAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCATCCGGCCGGCCTGAAGA
 AGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGTTACTTCAGCGTCCCTCTGGACGAG
 GACTTTAGAAAGTACACCGCCTTTACCATCCCATCTATCAACAACGAGACCCCTGGCATCAG
 ATATCAGTACAACGTCTCTCCCCAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCA
 TGACCAAGATCCTGGAGCCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATG
 GACGACCTGTACGTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCT
 GAGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGC
 CATTCTCTgaaGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGACGCTATCGTCCTC
 CCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGC
 TAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTGCGCGGCACCAAGG
 CCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCTCGAGCTGGCTGAGAACCGGGAG
 ATCCTGAAGGAGCCCGTGACGGCGTGTACTATGACCCCTCCAAGGACCTGATCGCCGAAAT
 CCAGAAGCAGGGCCAGGGGCAGTGGACATACCAGATTTACCAGGAGCCTTTCAAGAACCTCA
 AGACCGGCAAGTACGCCCCGATGAGGGGCGCCACACCAACGATGTCAAGCAGCTGACCGAG
 GCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCT
 GCCTATCCAGAAGGAGACCTGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTTC
 CCGAGTGGGAGTTTCGTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAG
 CCCATCGTGGGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGG
 GAAGGCCGGGTACGTACCAACCGGGGCGCCAGAAGGTCGTCAACCTGACCGACACCACCA
 ACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAGGTGAAC
 ATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTTCAGGCCCAGCCGGACCAAGTCCGAGAG
 CGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTCGCCTGGG



Figur 14

Sekvens:

ATGGGTGCCCGAGCTTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAAATT
AGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
AGGGAGCTTGAACGGTTTGCCTGTAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
TCCTTGATATAACACAGTGGCTACCTCTACTGCGTACACCAGAGGATCGAGATTAAAG
GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCAAGAACTATCCT
ATTGTCCAAAACATTTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
AATGCCTGGGTGAAGGTTGTCGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
TTCTCCGCTTTGAGTGAGGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
CGGGAGCCTCGCGGCTCTGACATTGCCGGGCAACACCTCTACACTGCAAGAGCAAATC
GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC

28/50

ATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCGGTTTTATAAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAACTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
 GTGTTGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGG
 ATGGACGGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTG
 GTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAG
 AACCATAACAACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGC
 AAGCTGGTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAG
 CTGGGCATCCCCATCCGGCCGGCCTGAAGAAGAAGAGCGTGACCGTGCTGGAC
 GTGGGCGACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCC
 TTTACCATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTC
 CTCCCCAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATC
 CTGGAGCCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGAC
 CTGTACGTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTG
 AGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAG
 CCGCCATTCTCTgaaGATGGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGACGCCT
 ATCGTCCTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGC
 AAGCTCAACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAG
 CTGCTGCGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAG
 CTCGAGCTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGTGACTAT
 GACCCCTCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACA
 TACCAGATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCCGATG
 AGGGGCGCCCAACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACG
 ACCGAGTCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAG
 GAGACCTGGGAGACGTGGTGGACCGAATATGGGCAGGCCACCTGGATTCCCGAGTGG
 GAGTTCGTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCC
 ATCGTGGGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTC
 GGGAAGGCCGGGTACGTACCAACCGGGGCCCGCAGAAGGTGTCACCCTGACCGAC
 ACCACCAACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGC
 CTGGAGGTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTAGGCCAG
 CCGGACCAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAA
 GAGAAGGTCTACCTCGCCTGGGTCCCGGCCCATAGGGCATTTGGCGGCAACGAGCAG
 GTCGACAAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGATGGTGGGTTTTCCAGTC
 ACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCAC
 TTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGAT
 ATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTAC
 ACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTA
 CCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTA
 CACCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGG
 TTTGACAGCCGCCTAGCATTTTCATCACGTGGCCCCGAGAGCTGCATCCGGAGTACTTC
 AAGAACTGCTGA [SEKV. ID. NR:73]

29/50

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQAAAA DTGHSNOVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTLNT VGGHQAA MQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGP K EPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQNANPDC KTI LKALGPA ATLEEMMTAC
 QGVGGPGHKA RVLMPISPI ETVSVKLKPG MDGPKVKQWP LTEEKIKALV
 EICTEMEKEG KISKIGPENP YNTPVFAIKK KDSTKWRKLV DFRELNKRTQ
 DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY FSVPLDEDFR KYTAFTIPSI
 NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQ PDIVIQYMD
 DLYVGSLEI GQHRTKIEEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH
 PDKWTVQPIV LPEKDSWTVN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG
 TKALTEVIPL TEEAELELAE NREILKEPVH GVYYDPSKDL IAEIQKQGG
 QWTYQIQEP FKNLKTGKYA RMRGAHTNDV KQLTEAVQKI TTESIWIWGK
 TPKFKLPQK ETWETWWTEY WQATWIPEWE FVNTPLVLK WYQLEKEPIV
 GAETFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDTTNQ KTELQAIYLA
 LQDSGLEVNI VTDSQYALGI IQAQPDQSES ELVNQIIIEQL IKKEKVYLAW
 VPAHKGIGGN EQVDKLVSAG IRKVLVVGFP VTPQVPLRPM TYKAAVDLSH
 FLKEKGGLG LIHSQRRQDI LDLWIYHTQG YFPDWQNYTP GPGVRYPLTF
 GWCYKLVPE PDKVEEANKG ENTSL LHPVS LHGMDDPERE VLEWRFD SRL
 AFHHVARELH PEYFKNC*

[SEKV. ID. NR:74]

30/50

Figur 15

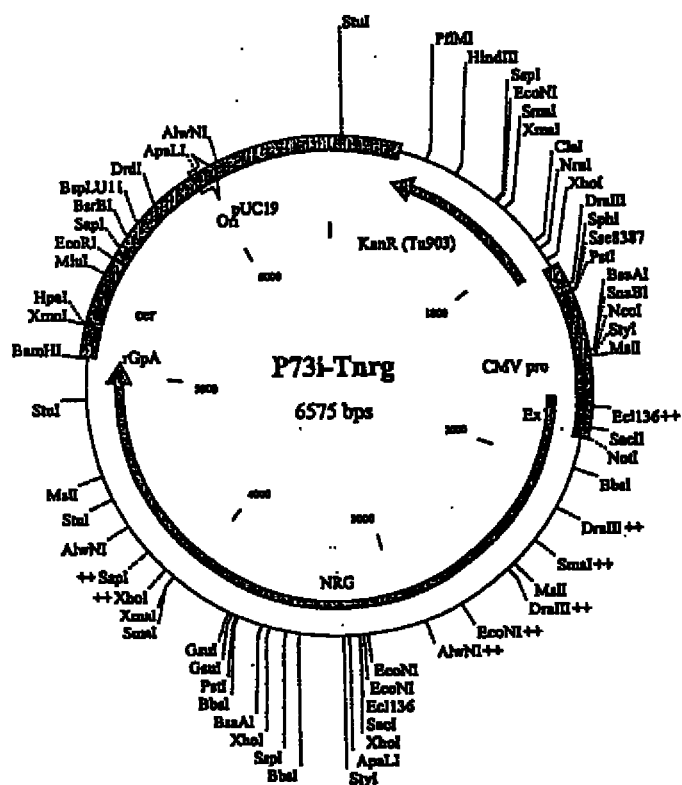
Sekvens av Tnrg:

ATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCA
 GCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCAC
 TCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTC
 CCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGA
 TGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGA
 GAGAACACCAGCTTGTTACACCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGA
 GAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTTCATCACGTGGCCCCGAGAG
 CTGCATCCGGAGTACTTCAAGAACTGCATGGGCCCCATCAGTCCCATCGAGACCGTG
 CCGGTGAAGCTGAAACCCGGGATGGACGGCCCCAAGGTCAAGCAGTGGCCACTCACC
 GAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAG
 ATCAGCAAGATCGGGCCTGAGAACCATAACAACCCCCGTGTTTGCCATCAAGAAG
 AAGGACAGCACCAAGTGGCGCAAGCTGGTGGATTTCGGGGAGCTGAATAAGCGGACC
 CAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCCATCCGGCCGGCCTGAAGAAGAAG
 AAGAGCGTGACCGTGCTGGACGTGGGCGACGCTTACTTCAGCGTCCCTCTGGACGAG
 GACTTTAGAAAGTACACCGCCTTTACCATCCCATCTATCAACAACGAGACCCCTGGC
 ATCAGATATCAGTACAACGTCTCTCCCCAGGGCTGGAAGGGCTCTCCCGCCATTTTC
 CAGAGCTCCATGACCAAGATCCTGGAGCCGTTTTCGGAAGCAGAACCCCGATATCGTC
 ATCTACCAGTACATGGACGACCTGTACGTGGGCTCTGACCTGGAATCGGGCAGCAT
 CGCACGAAGATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCG
 GACAAGAAGCATCAGAAGGAGCCGCCATTCCTgaaGATGGGCTACGAGCTCCATCCC
 GACAAGTGGACCGTGACGCCTATCGTCTCTCCCCGAGAAGGACAGCTGGACCGTGAAC
 GACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGCTAGCCAGATCTATCCCGGGATC
 AAGGTGCGCCAGCTCTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGAGGTGATT
 CCCCTCACGGAGGAAGCCGAGCTCGAGCTGGCTGAGAACCGGGAGATCCTGAAGGAG
 CCCGTGCACGGCGTGTAATGACCCCTCCAAGGACCTGATCGCCGAAATCCAGAAG
 CAGGGCCAGGGGCAGTGGACATACCAGATTTACCAGGAGCCTTTCAAGAACCTCAAG
 ACCGGCAAGTACGCCCCGATGAGGGGCGCCACACCAACGATGTCAAGCAGCTGACC
 GAGGCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGGGAAGACACCCAAG
 TTCAAGCTGCCTATCCAGAAGGAGACCTGGGAGACGTGGTGGACCGAATATTGGCAG
 GCCACCTGGATTCCCGAGTGGGAGTTCTGTAATACACCTCCTCTGGTGAAGCTGTGG
 TACCAGCTCGAGAAGGAGCCCATCGTGGGCGCGGAGACATTCTACGTGGACGGCGCG
 GCCAACCGCGAAACAAAGCTCGGGAAGGCCGGGTACGTACCAACCGGGGGCCGCCAG
 AAGGTCGTACCCCTGACCGACACCAACCAAGACGAGCTGCAGGCCATCTAT
 CTCGCTCTCCAGGACTCCGGCCTGGAGGTGAACATCGTGACGGACAGCCAGTACGCG
 CTGGGCATTATTCAGGCCCAGCCGGACCAAGTCCGAGAGCGAACTGGTGAACCAGATT
 ATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTCGCCTGGGTCCCGGCCATAAG
 GGCATTGGCCGCAACGAGCAGGTGCAAGCTGGTGGAGTGCGGGGATTAGAAAGGTG
 CTGATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAA
 ATTAGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCC
 TCGAGGGAGCTTGAACGGTTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGGA
 TGTCGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTG
 AGGTCCTTGATAACACAGTGGCTACCTCTACTGCGTACACCAGAGGATCGAGATT

31/50

AAGGATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAG
 AAGGCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTAT
 CCTATTGTCCAAAACATTACAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACG
 CTCAATGCCTGGGTGAAGGTTGTCTGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCC
 ATGTTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAAT
 ACCGTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAG
 GCAGCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAG
 ATGCGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAA
 ATCGGATGGATGACCAACAATCCTCCCATCCAGTTGGAGAAATCTATAAACGGTGG
 ATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTTGAC
 ATTAGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCGGTTTTATAAGACC
 CTGCGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTG
 GTACAGAACGCTAACCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCC
 ACCCTGGAAGAGATGATGACCGCCTGTCTAGGGAGTAGGCGGACCCGACACAAAGCC
 AGAGTGTGTGA [SEKV. ID. NR:75]

MVGFVPTPQV PLRPMYKAA VDLSHFLKEK GGLEGLIHSQ RRQDILDLWI
 YHTQGYFPDW QNYTPGPGVR YPLTFGWCYK LVPVEPDKVE EANKGENTS
 LHPVSLHGMD DPEREVLEWR FDSRLAFHHV ARELHPEYFK NCMGPISPIE
 TVSVKLKPGM DGPVKVQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
 NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKS
 TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
 PAIFQSSMTK ILEPFRKQNP DIVIYQYMD LYGSDLEIG QHRTKIEELR
 QHLRLWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
 IQKLVGKLNW ASQIYPGIKV RQLCKLLRGT KALTEVIPLT EEAELELAEN
 REILKEPVHG VYYDPSKDLI AEIQKQGQWQ WTYQIYQEPF KNLKTGKYAR
 MRGAHTNDVK QLTEAVQKIT TESIVIWGKT PKFKLPIQKE TWETWTEYW
 QATWIPEWEF VNTPLVLKLW YQLEKEPIVG AETFYVDGAA NRETKLGKAG
 YVTNRGRQKV VTLTDTTNQK TELQAIYLAL QDSGLEVNIV TDSQYALGII
 QAQPDQSESE LVNQIIEQLI KKEKVYLAWV PAHKGIGGNE QVDKLVSAGI
 RKVLMGARAS VLSGGELDRW EKIRLRPGGK KKYKLKHIVW ASRELERFAV
 NPGLLETSEG CRQILGQLQP SLQTGSEELR SLYNTVATLY CVHQRIEIKD
 TKEALDKIEE EQNKSKKKAQ QAAADTGHSN QVSQNYPIVQ NIQQQMVHQA
 ISPRTLNAWV KVVEEKAFSP EVIPMFSAIS EGATPQDLNT MLNTVGGHQA
 AMQMLKETIN EEAIEWDRVH PVHAGPIAPG QMREPRGSDI AGTTSTLQEQ
 IGWMTNNPPI PVGEIYKRWI ILGLNKIVRM YSPTSILDIR QGPKEPFRDY
 VDRFYKTLRA EQASQEVKNW MTETLLVQNA NPDCKTILKA LGPAATLEEM



Figur 16

Sekvens av Tngr-innskuddet i p7313ie:

ATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGT
 AGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACCTCCCAAAGAA
 GACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAAC
 TACACACCAGGGCCAGGGGT CAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACC
 AGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTG
 TGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGC
 CTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCATGGGTGC
 CCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATTAGGCTGCGCCCGG
 GAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCGAGGGAGCTTGAACGGTTT
 GCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGTGCGCAGATCCTGGGGCAATTGCA
 GCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGGTCTTGTATAACAAGTGGCTACCCCTCT
 ACTGCGTACACCAGAGGATCGAGATTAAGGATACCAAGGAGGCCTTGGACAAAATTGAGGAG
 GAGCAAAACAAGAGCAAGAAGAAGGCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCA
 GGTATCACAGAACTATCCTATTGTCCAAAACATT CAGGGCCAGATGGTT CATCAGGCCATCA
 GCCCCCGGACGCTCAATGCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTT
 ATCCCCATGTTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAA
 TACCGTGGGCGGCCATCAGGCCGCCATGCAATGTTGAAGGAGACTATCAACGAGGAGGCAG
 CCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATGCGGGAG
 CCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATCGGATGGATGAC
 CAACAATCCTCCCATCCAGTTGGAGAAATCTATAAACGGTGGATCATCCTGGGCCTGAACA
 AGATCGTGCGCATGTACTCTCCGACATCCATCCTTGACATTAGACAGGGACCCAAAGAGCCT
 TTTAGGGATTACGTGACCCGGTTTTATAAGACCCTGCGAGCAGAGCAGGCCTCTCAGGAGGT
 CAAAAACTGGATGACGGAGACTCCTGGTACAGAACGCTAACCCCGACTGCAAAACAATCT
 TGAAGGCACTAGGCCCGGCTGCCACCCTGGAAGAGATGATGACCGCTGT CAGGGAGTAGGC
 GGACCCGGACACAAAGCCAGAGTGTGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGT
 GAAGCTGAAACCCGGGATGGACGGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGA
 TCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGG
 CCTGAGAACCATAACAACACCCCGTGTGTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCG
 CAAGCTGGTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGG
 GCATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGAC
 GCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACCATCCCATC
 TATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCCCAGGGCTGGAAGG
 GCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAGCCGTTTCGGAAGCAGAAC
 CCGATATCGTCATCTACCAGTACATGGACGACCTGTACGTGGGCTCTGACCTGGAAATCGG
 GCAGCATCGCACGAAGATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACCACTC
 CGGACAAGAAGCATCAGAAGGAGCCGCCATTCTGaaGATGGGCTACGAGCTCCATCCCGAC
 AAGTGGACCGTGACGCTATCGTCTCCCGAGAAGGACAGCTGGACCGTGAACGACATCCA
 GAAGCTGGTGGGCAAGCTCAACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGC
 TCTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCC
 GAGCTCGAGCTGGCTGAGAACCGGGAGATCTGAAGGAGCCCGTGACCGCGTGTACTATGA
 CCCCCTCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAGA
 TTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGCGCCAC

34/50

ACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAGTCCATCGTGAT
 CTGGGGGAAGACACCCCAAGTTCAAGCTGCCTATCCAGAAGGAGACCTGGGAGACGTGGTGA
 CCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTCGTGAATACACCTCCTCTGGTG
 AAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTGGGCGCGGAGACATTCTACGTGGACGG
 CGCGGCCAACCGCGAAACAAAGCTCGGGAAGGCCGGGTACGTACCAACCGGGGCCGCCAGA
 AGGTCGTCAACCTGACCGACACCACCAACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCT
 CTCCAGGACTCCGGCCTGGAGGTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTAT
 TCAGGCCCAGCCGACAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCA
 AGAAAGAGAAGGTCTACCTCGCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAG
 GTCGACAAGCTGGTGAAGTGCGGGGATTAGAAAGGTGCTGTAA [SEKV. ID. NR:77]

MVGFPVTPQV PLRPMTYKAA VDLSHFLKEK GGLEGLIHSQ RRQDILDWI
 YHTQGYFPDW QNYTPGPGVR YPLTFGWYK LVPVEPKVE EANKGENTSL
 LHPVSLHGMD DPEREVLEWR FDSRLAFHHV ARELHPEYFK NCMGARASVL
 SGGELDRWEK IRLRPGGKKK YKLKHIVWAS RELERFAYNP GLLETSEGCR
 QILGQLQPSL QTGSEELRSL YNTVATLYCV HQRIEIKDTK EALDKIEEEQ
 NKSKKKAQQA AADTGHSNOV SQNYPIVQNI QGQMVHQAI S PRTLNAWVKV
 VEEKAFSPEV IPMFSALSEG ATPQDLNML NTVGGHQAAM QMLKETINEE
 AAEDRVHPV HAGPIAPGOM REPRGSDIAG TTSTLQEQIG WMTNNPPIPV
 GEIYKRWIIL GLNKIVRMYS PTSILDIRQG PKEPFRDYVD RFYKTLRAEQ
 ASQEVKNWMT ETLLVQANANP DCKTILKALG PAATLEEMMT ACQGVGGPGH
 KARVLMGPIS PIETVSVKLK PGMDGPKVKQ WPLTEEKIKA LVEICTEMEK
 EGKISKIGPE NPYNTFVFAI KKKDSTKWRK LVDFRELNRK TQDFWEVQLG
 IPHPAGLKKK KSVTVLDVGD AYFSVPLDED FRKYTAFTIP SINNETPGIR
 YQYNVLPQGW KGSPAIFQSS MTKILEPFRK QNPDIYIYQY MDDLYVGS DL
 EIGQHRTKIE ELRQHLLRWG LTPDKKHQK EPPFLWMGYE LHPDKWTVQP
 IVLPEKDSWT VNDIQKLVGK LNWASQIYPG IKVRQLCKLL RGTKALTEVI
 PLTEEAEL EL AENREILKEP VHGVYDPSK DLIAEIQKQG QGQWYQIYQ
 EPFKNLKTGK YARMRGAHTN DVKQLTEAVQ KITTESIWIW GKTPKFKLPI
 QKETWETWWT EYWQATWIPE WEFVNTPLV KLWYQLEKEP IVGAETFYVD
 GAANRETKLG KAGYVTNRGR QKVVTLTDTT NQKTELQAIY LALQDSGLEV
 NIVTDSQYAL GIIQAQPDQS ESELVNIIE QLIKKEKVYL AWPAPKGGIG
 GNEQVDKLVS AGIRKVL*

[SEKV. ID. NR:78]

36/50

Figur 17

Sekvens av innskudd Trgn#6

ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC
 GGCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAG
 ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCA
 TACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG
 GTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC
 ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC
 GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC
 ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCTCTCCC
 CAGGGCTGGAAGGGCTCTCCCGCCATTTTCAGAGCTCCATGACCAAGATCCTGGAG
 CCGTTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGACCTGTAC
 GTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTGAGGCAG
 CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA
 TTCCTgaaGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGCAGCCTATCGTC
 CTCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTC
 AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG
 CGCGGCACCAAGGCCCTGACCGAGGTGATTTCCCTCACGGAGGAAGCCGAGCTCGAG
 CTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGTGTAATGACCCC
 TCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAG
 ATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCCGATGAGGGGC
 GCCCACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCAGGACCGAG
 TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC
 TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTC
 GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCATCGTG
 GGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGGGGAAG
 GCCGGGTACGTCACCAACCGGGGCCGCCAGAAGTTCGTCACCCTGACCGACACCACC
 AACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAG
 GTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTAGGCCCAGCCGGAC
 CAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAG
 GTCTACCTCGCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAGGTCGAC
 AAGCTGGTGAGTGCGGGGATTAGAAAGTGCTGATGGGTGCCCGAGCTTCGGTACTG
 TCTGGTGGAGAGCTGGACAGATGGGAGAAAATTAGGCTGCGCCCGGAGGCAAAAAG
 AAATACAAGCTCAAGCATATCGTGTGGGCCTCGAGGGAGCTTGAACGGTTTGCCGTG
 AACCCAGGCCTGCTGGAAACATCTGAGGGATGTCGCCAGATCCTGGGGCAATTGCAG
 CCATCCCTCCAGACCGGGAGTGAAGAGCTGAGGTCCTTGTATAACACAGTGGCTACC
 CTCTACTGCGTACACCAGAGGATCGAGATTAAGGATACCAAGGAGGCCTTGACAAA
 ATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAGGCCCAGCAGGCAGCTGCTGACACT
 GGGCATAGCAACCAGGTATCACAGAACTATCCTATTGTCCAAACATTTCAGGGCCAG
 ATGGTTTCATCAGGCCATCAGCCCCCGGACGCTCAATGCCTGGGTGAAGGTTGTCGAA
 GAGAAGGCCCTTTCTCCTGAGGTTATCCCCATGTTCTCCGCTTTGAGTGAGGGGGCC
 ACTCCTCAGGACCTCAATACAATGCTTAATACCGTGGGCGGCCATCAGGCCGCCATG

37/50

CAAATGTTGAAGGAGACTATCAACGAGGAGGCAGCCGAGTGGGACAGAGTGCATCCC
 GTCCACGCTGGCCCAATCGCGCCCGGACAGATGCGGGAGCCTCGCGGCTCTGACATT
 GCCGGCACCACCTCTACACTGCAAGAGCAAATCGGATGGATGACCAACAATCCTCCC
 ATCCCAGTTGGAGAAATCTATAAACGGTGGATCATCCTGGGCCTGAACAAGATCGTG
 CGCATGTACTCTCCGACATCCATCCTTGACATTAGACAGGGACCCAAAGAGCCTTTT
 AGGGATTACGTGACCGGTTTTATAAGACCCTGCGAGCAGAGCAGGCCTCTCAGGAG
 GTCAAAAACCTGGATGACGGAGACACTCTGGTACAGAACGCTAACCCCGACTGCAAA
 ACAATCTTGAAGGCACTAGGCCCGGCTGCCACCCTGGAAGAGATGATGACCGCCTGT
 CAGGGAGTAGGCGGACCCGGACACAAAGCCAGAGTGTGATGGTGGGTTTTCCAGTC
 ACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCAC
 TTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGAAGACAAGAT
 ATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTAC
 ACACCAGGGCCAGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTA
 CCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTA
 CACCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGG
 TTTGACAGCCGCTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTC
 AAGAACTGCTGA [SEKV. ID. NR:79]

MGPISPIETV SVKCLKPGMDG PKVKQWPLTE EKIKALVEIC TEMEKEGKIS
 KIGPENPYNT PVFAIKKKDS TKWRKLVDFR ELNKRTQDFW EVQLGIPHPA
 GLKKKKSVMV LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV
 LPQGWKQSPA IFQSSMTKIL EPFRKQNPDI VIYQYMDDL YVSDLEIGQH
 RTKIEELRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLPE
 KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEE
 AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGGQWT YQIYQEPFKN
 LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW
 ETWWTEYWQA TWIPEWEFVN TPPLVKLWYQ LEKEPIVGAE TFYVDGAANR
 ETKLGKAGYV TNRRGRQVVT LTDTTNQKTE LQAIYLALQD SGLEVNIVTD
 SQYALGIIQA QPDQSESELV NQIIEQLIKK EKVYLAWVPA HKGIGGNEQV
 DKLVSAGIRK VLMGARASVL SGGELDRWEK IRLRPGGKKK YKLKHIVWAS
 RELERFAVNP GLEETSEGCR QILGQLQPSL QTGSEELRSL YNTVATLYCV
 HQRIEIKDTK EALDKIEEEQ NKSCKKAQQA AADTGHSNQV SQNYPIVQNI
 QGQMVHQAIS PRTLNAWVKV VEEKAFSPEV IPMFSALSEG ATPQDLNMTL
 NTVGGHQAAM QMLKETINEE AAEWDRVHPV HAGPIAPGQM REPRGSDIAG
 TTSTLQEQIG WMTNNPPIPV GEIYKRWIL GLNKIVRMYS PTSILDIRQG
 PKEPFRDYVD RFYKTLRAEQ ASQEVKNWMT ETLLVQNANP DCKTILKALG
 PAATLEEMMT ACQGVGGPGH KARVLMVGFP VTPQVPLRPM TYKAAVDLSH
 FLKEKGLEG LIHSQRRQDI LDLWIYHTQG YFPDWQNYTP GPGVRYPLTF

Figur 18

Sekvens av innskudd av Trng#11

ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC
 GGCCCCAAGGTCAAGCAGTGGCCACTCACCAGGAGAAGATCAAGGCCCTGGTGGAG
 ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCGGAGAACCCA
 TACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG
 GTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC
 ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC
 GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC
 ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCTCTCCC
 CAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAG
 CCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGACCTGTAC
 GTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTGAGGCAG
 CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA
 TTCCTgaaGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGCAGCCTATCGTC
 CTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTC
 AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG
 CGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCTCGAG
 CTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGTGTACTATGACCCC
 TCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAG
 ATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGC
 GCCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCAGACCGAG
 TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC
 TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTC
 GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTG
 GCGCGGAGACATTCTACGTGGACGGCGCGGCCAACC CGGAAACAAAGCTCGGGAAG
 GCCGGGTACGTCAACCAACCGGGGCCCGCAGAAGGTCGTACCCTGACCGACACCACC
 AACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAG
 GTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTCAGGCCCAGCCGGAC
 CAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAG
 GTCTACCTCGCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAGGTGAC
 AAGCTGGTGAAGTGCGGGGATTAGAAAGGTGCTGATGGTGGGTTTTCCAGTCACACCT
 CAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTAA
 AAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGAAGACAAGATATCCTT
 GATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTACACACCA
 GGGCCAGGGGTGAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTT
 GAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCT
 GTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGAC
 AGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAAC
 TGCTGAATGGGTGCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAG
 AAAATTAGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGG
 GCCTCGAGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAG
 GGATGTGCGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAG
 CTGAGGTCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAG

40/50

ATTAAGGATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAG
 AAGAAGGCCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAAC
 TATCCTATTGTCCAAAACATTGAGGGCCAGATGGTTCATCAGGCCATCAGCCCCGG
 ACGCTCAATGCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTTATC
 CCCATGTCTCTCGCTTTGAGTGAGGGGGGCCACTCCTCAGGACCTCAATACAATGCTT
 AATACCGTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAG
 GAGGCAGCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGA
 CAGATGCGGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACTCTACACTGCAAGAG
 CAAATCGGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGG
 TGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTT
 GACATTAGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTATAAG
 ACCCTGCGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTC
 CTGGTACAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCT
 GCCACCCTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCGGACCCGGACACAAA
 GCCAGAGTGTTG [SEKV. ID. NR:81]

MGPISPIETV SVKLKPGMDG PKVKQWPLTE EKIKALVEIC TEMEKEGKIS
 KIGPENPYNT PVFAIKKIDS TKWRKLVDFR ELNKRTOQDFW EVQLGIPHPA
 GLKKKKSVTV LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV
 LPQGWKGSPA IFQSSMTKIL EPFRKQNPDI VIYQYMDLY VGSDLEIGQH
 RTKIEELRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLPE
 KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEE
 AELELAENRE ILKEPVHGVY YDPSKDLIAR IQKQGGQWT YQIYQEPFKN
 LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW
 ETWWTEYWQA TWIPEWEFVN TPPLVKLWYQ LEKEPIVGAE TFYVDGAANR
 ETKLGKAGYV TNRGRQKVT LTDTTNQKTE LQAIYLALQD SGLEVNIVTD
 SQYALGIIQA QPDQSESELV NQIEQLIKK EKVYLAWVPA HKGIGGNEQV
 DKLVSAIRK VLMVGFPVTP QVPLRPMYK AAVDLSHFLK EKGGLLEGLIH
 SQRRODILD LWIYHTQGYFP DWQNYTPGPG VRYPLTFGWC YKLVPVEPDK
 VEEANKGENT SLLHPVSLHG MDDPEREVLE WRFDSRLAFH HVARELHPEY
 FKNCMGARAS VLSGGELDRW EKIRLRPGGK KKYKLKHIVW ASRELERFAV
 NPGLLETSEG CRQILGQLQP SLQTGSEELR SLYNTVATLY CVHQRIEIKD
 TKEALDKIEE EQNKSKKKAQ QAAADTGHNS QVSQNYPIVQ NIQQQMVHQA
 ISPRTLNAWV KVVEEKAFSP EVIPMFSAIS EGATPQDLNT MLNTVGGHQA
 AMQMLKETIN EEAAEWDRVH PVHAGPIAPG QMREPRGSDI AGTTSTLQEQ
 IGWMTNNPPI PVGEIYKRWI ILGLNKIVRM YSPTSILDIR QGPKEPFRDY
 VDRFYKTLRA EQASQEVKNW MTETLLVQNA NPDCKTILKA LGPAATLEEM
 MTACQGVGGP GHKARVL [SEKV. ID. NR:82]

42/50

Figur 19

Sekvens av TgnR (F1):

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
 AGGCTGCGCCCGGGAGGCCAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
 AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
 CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
 TCCTTGTATAACACAGTGGCTACCCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
 ATTGTCCAAAACATTTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
 AATGCCTGGGTGAAGGTTGTCTGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
 TTCTCCGCTTTGAGTGAGGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
 GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCCA
 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
 CGGGAGECTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
 GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
 ATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCCGGTTTATAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCCTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCCGGACCCGGACACAAAGCCAGA
 GTGTTGATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTAC
 AAGGCAGCTGTAGATCTTAGCCACTTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTA
 ATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGC
 TACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACC
 TTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAAT
 AAAGGAGAGAAACACCAGCTTGTTACACCTGTGAGCCTGCATGGGATGGATGACCCG
 GAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTTCATCACGTGGCC
 CGAGAGCTGCATCCGGAGTACTTCAAGAACTGCATGGGCCCCATCAGTCCCATCGAG
 ACCGTGCCGGTGAAGCTGAAACCCGGGATGGACGGCCCCAAGGTCAAGCAGTGGCCA
 CTCACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAAGAG
 GGCAAGATCAGCAAGATCGGGCCTGAGAACCCATACAACACCCCCCGTGTTTGCCATC
 AAGAAGAAGGACAGCACCAAGTGCGCAAGCTGGTGGATTTCCGGGAGCTGAATAAG
 CGGACCCAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCATCCGGCCGGCCTGAAG
 AAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCTTACTTCAGCGTCCCTCTG
 GACGAGGACTTTAGAAAGTACACCGCCTTTACCATCCCATCTATCAACAACGAGACC
 CCTGGCATCAGATATCAGTACAACGTCTCCCCCAGGGCTGGAAGGGCTCTCCCGCC
 ATTTTCCAGAGCTCCATGACCAAGATCCTGGAGCCGTTTTCGGAAGCAGAACCCCGAT
 ATCGTCATCTACCAGTACATGGACGACCTGTACGTGGGCTCTGACCTGGAAATCGGG
 CAGCATCGCACGAAGATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACC
 ACTCCGGACAAGAAGCATCAGAAGGAGCCGCCATTCTTgaaGATGGGCTACGAGCTC
 CATCCCGACAAGTGGACCGTGCAGCCTATCGTCTCCCCGAGAAGGACAGCTGGACC
 GTGAACGACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGCTAGCCAGATCTATCCC
 GGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGAG

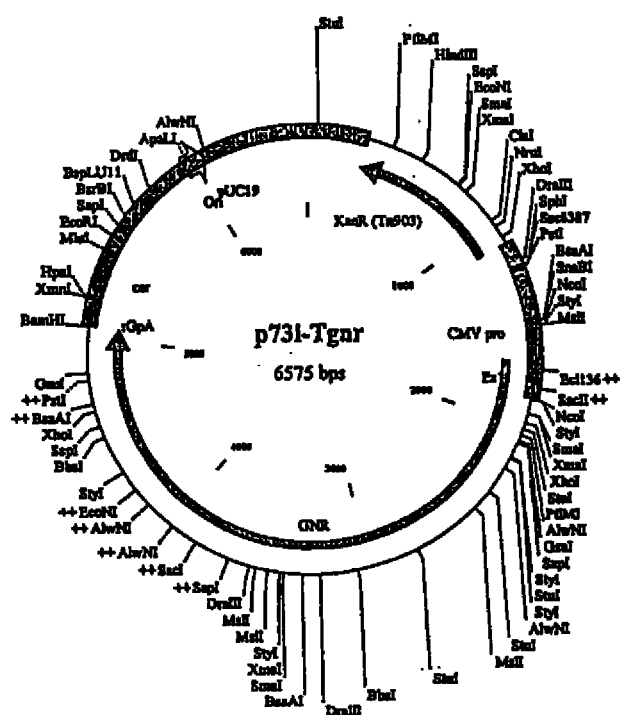
43/50

GTGATTCCCCTCACGGAGGAAGCCGAGCTCGAGCTGGCTGAGAACCGGGAGATCCTG
AAGGAGCCCCGTGCACGGCGTGTACTATGACCCCTCCAAGGACCTGATCGCCGAAATC
CAGAAGCAGGGCCAGGGGCAGTGGACATACCAGATTTACCAGGAGCCTTTCAAGAAC
CTCAAGACCGGCAAGTACGCCCCGATGAGGGGGCGCCACACCAACGATGTCAAGCAG
CTGACCGAGGCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGGGAAGACA
CCCAAGTTCAAGCTGCTTATCCAGAAGGAGACCTGGGAGACGTGGTGGACCGAATAT
TGGCAGGCCACCTGGATTCCCGAGTGGGAGTTCGTGAATACACCTCCTCTGGTGAAG
CTGTGGTACCAGCTCGAGAAGGAGCCCATCGTGGGCGCGGAGACATTCTACGTGGAC
GGCGCGGCCAACC CGCAAACAAAGCTCGGGAAGGCCGGGTACGTACCAACCGGGGC
CGCCAGAAGGTGCTCACCCCTGACCGACACCACCAACCAGAAGACGGAGCTGCAGGCC
ATCTATCTCGCTCTCCAGGACTCCGGCCTGGAGGTGAACATCGTGACCGACAGCCAG
TACGCGCTGGGCATTATTACGGCCAGCCGGACCAGTCCGAGAGCGAACTGGTGAAC
CAGATTATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTCGCCTGGGTCCCGGCC
CATAAGGGCATTGGCGGCAACGAGCAGGTCGACAAGCTGGTGAGTGCGGGGATTAGA
AAGGTGCTGTAA [SEKV. ID. NR:83]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
LETSEGCRI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
LDKIEEBEQNK SKKKAQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAISPR
TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTLNT VGGHQAAMQM
LKETINEEAA EWDRVHPVHA GPIAPGQMR PRGSDIAGTT STLQEQIGWM
TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGP KEFRDYVDRF
YKTLRAEQAS QEVKNWMTET LLVQANANPDC KTILKALGPA ATLEEMMTAC
QGVGGPGHKA RVLVVGFPVT PQVPLRPMTY KAAVDLSHFL KEKGGLEGLI
HSQRRQDILD LWIYHTQGYF PDWQNYTPGP GVRYPPLTFGW CYKLVPVEPD
KVEEANKGEN TSLHHPVSLH GMDDPEREVL EWRFD SRLAF HHVARELHPE
YFKNMGPIS PIETVSVKLK PGMDGPKVKQ WPLTEEKIKA LVEICTEMEK
EGKISKIGPE NPYNTPVFAI KKKDSTKWRK LVDFRELNRK TQDFWEVQLG
IPHPAGLKKK KSVTVLDVGD AYFSVPLDED FRKYTAFTIP SINNETPGIR
YQYNVLPQGW KGSPAIFQSS MTKILEPFRK QNPDIYIYQY MDDLIVGSDL
EIGQHRTKIE ELRQHLLRWG LTPDKKHQK EPPFLWMGYE LHPDKWTVQP
IVLPEKDSWT VNDIQKLVGK LNWASQIYPG IKVRQLCKLL RGTKALTEVI
PLTEEALEL AENREILKEP VHGVYDPSK DLIAEIQKQG QGQWYQIYQ
EPFKNLKTGK YARMRGAHTN DVKQLTEAVQ KITTESIVIW GKTPKPKLPI
QKETWETWWT EYWQATWIPE WEFVNTPLV KLWYQLEKEP IVGAETFYVD
GAANRETKLG KAGYVTNRGR QKVVTLTDTT NQKTELQAIY LALQDSGLEV
NIVTDSQYAL GIIQAQPDQS ESELVNIIE QLIKKEKVYL AWVPAHKGIG
GNEQVDKLV AGIRKVL*

[SEKV. ID. NR:84]

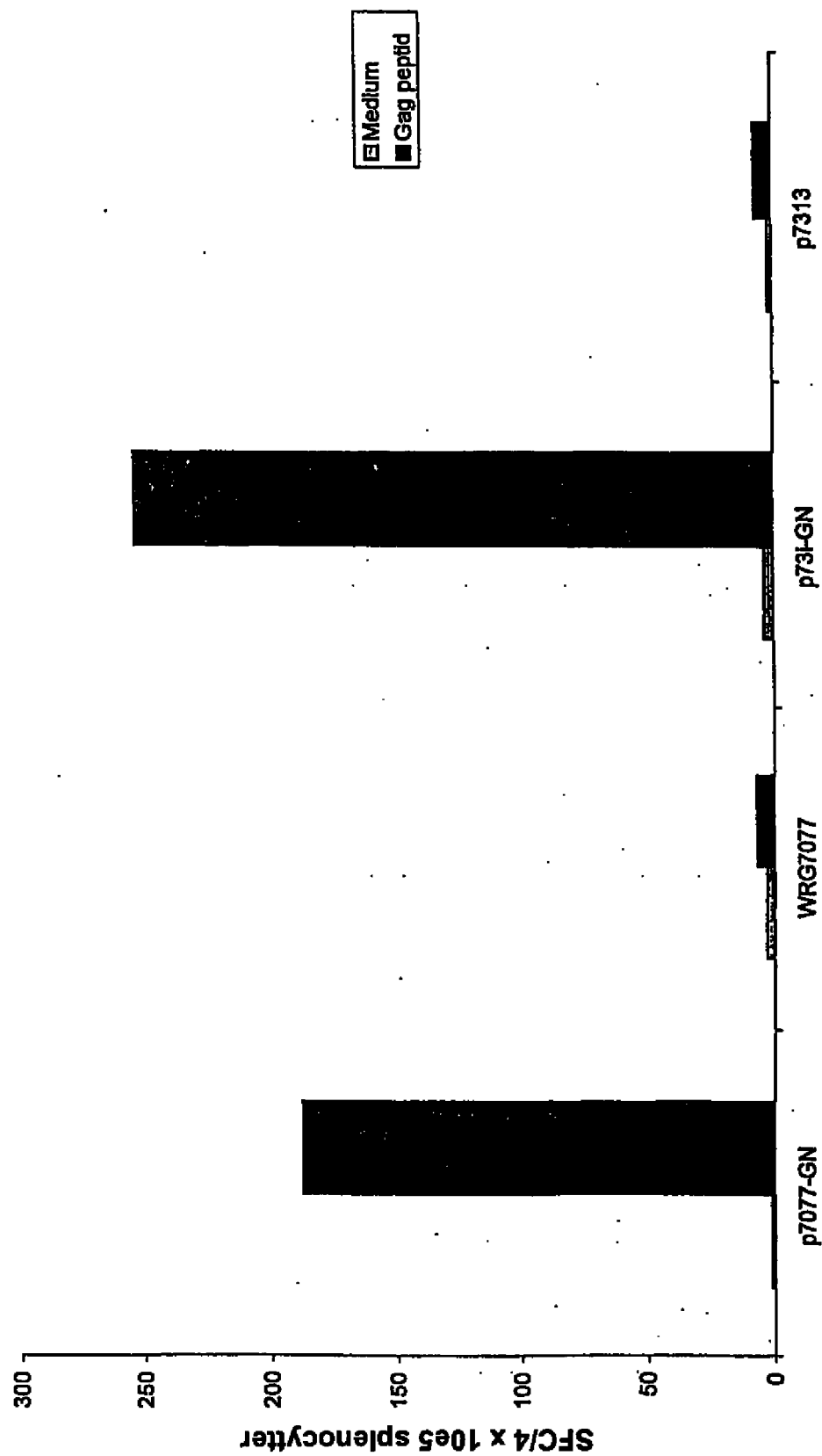
44/50



45/50

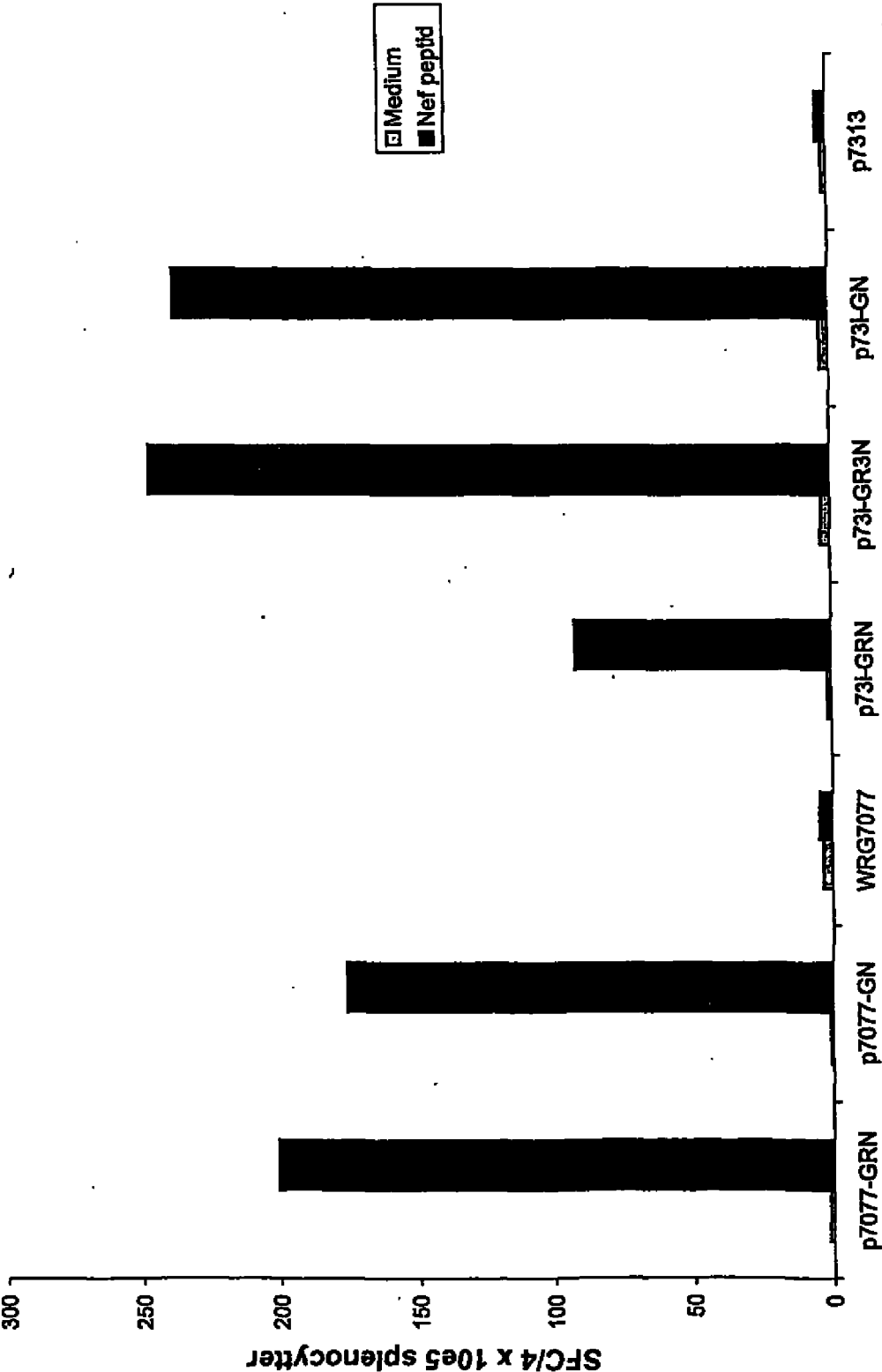
Figur 20

Responser på Gag-peptid målt ved bruk av IFN-gamma ELIsport 5 dager etter booster

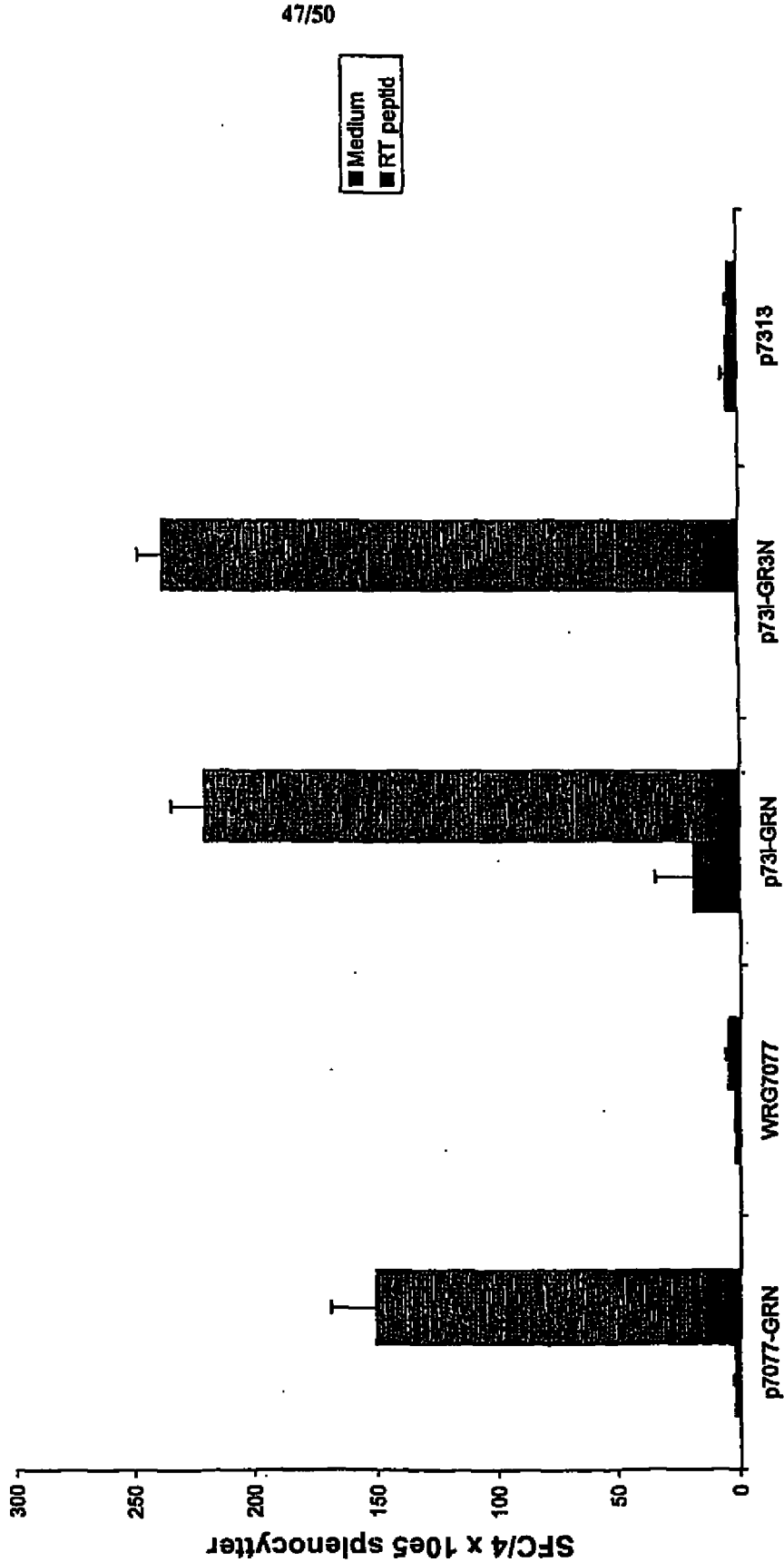


Figur 21

Responser på Nef-peptid ved bruk av IFN-gamma ELIsot 5 dager etter boosterdose



Figur 22
Responser på Rt-peptid ved IFN-gamma ELIspot 5 dager etter boosterdose

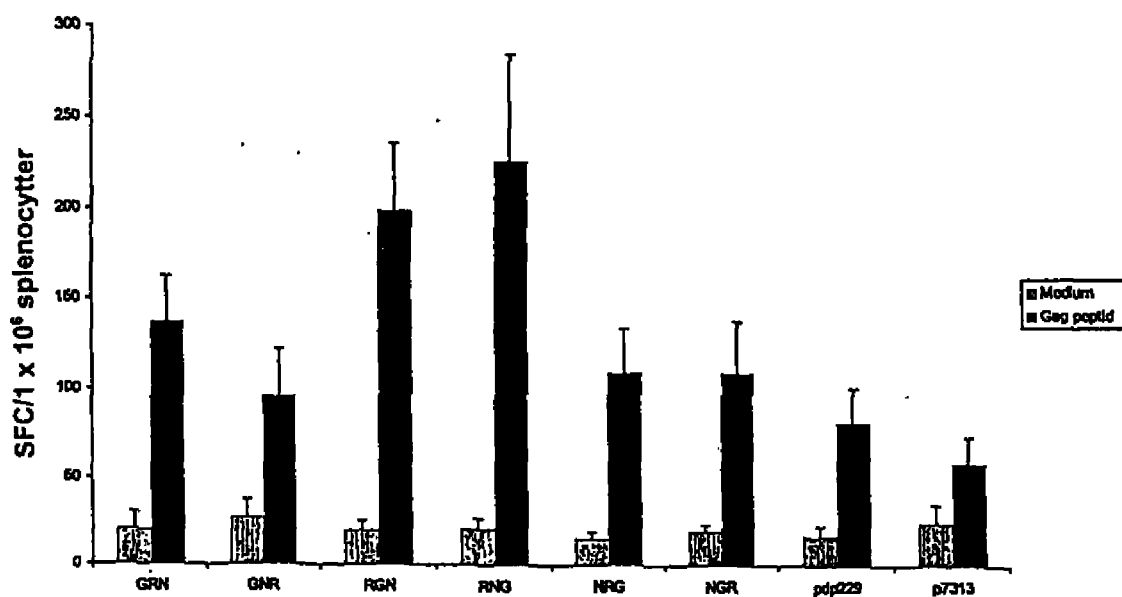


48/50

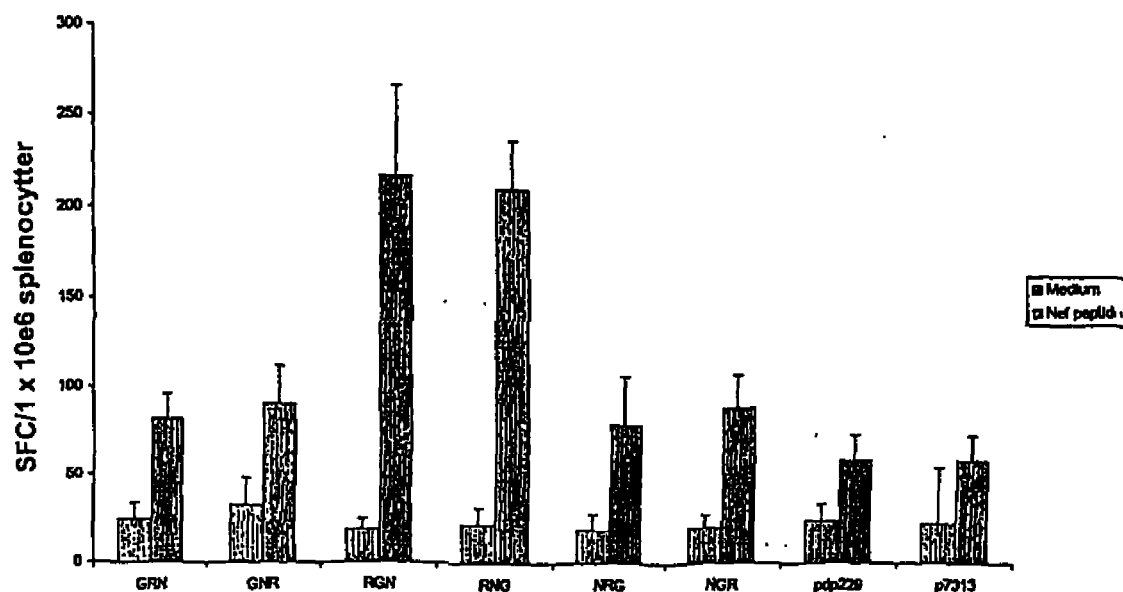
Figur 23

Eksempel på den CD8 immunrespons som påvises i musemodellen etter immunisering med de reformaterte konstruksjoner. Konstruksjonene omfatter Gag, RT og Nef som en trippelfusjon. G er Gag, R er RT og N er Nef.

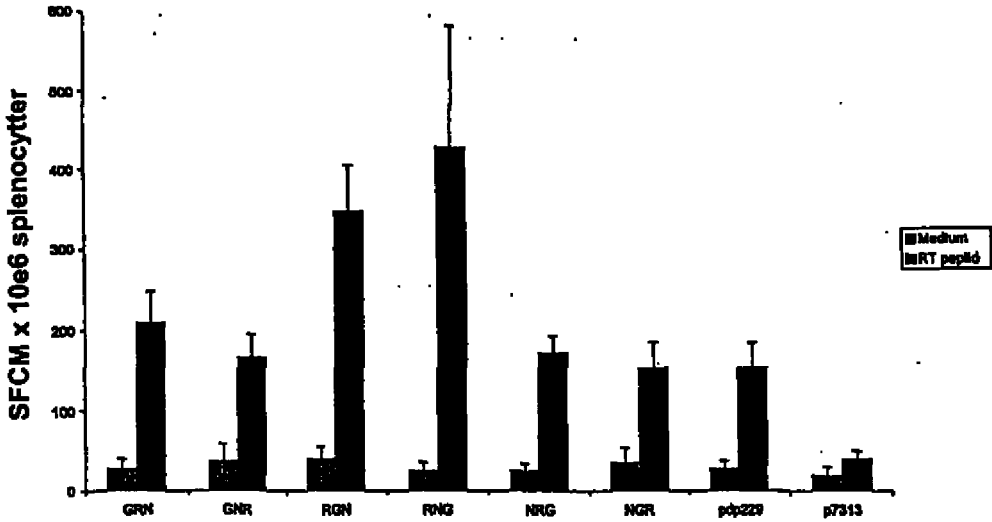
Responser på Gag-peptid ved IFN-gamma ELIspot på Dag 35 (7 dager etter boosterdose)



Responser på Nef-peptid ved IFN-gamma ELIspot på Dag 35 (7 dager etter boosterdose)



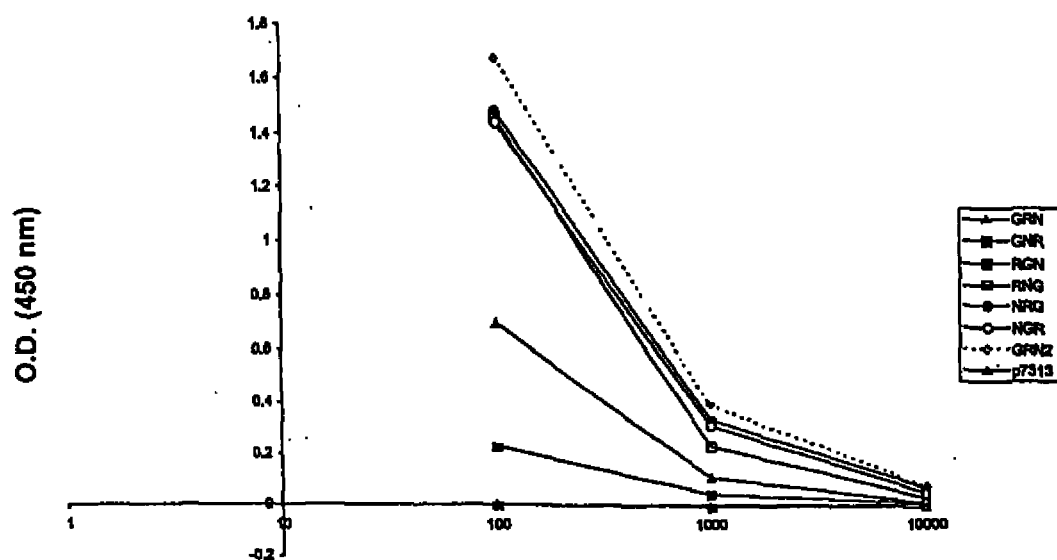
Responser på RT-peptid ved IFN-gamma ELIsot på Dag 35 (7 dager etter boosterdose)



50/50

Figur 24

Gag-antistoffer ved ELISA på Dag 35 (7 dager etter booster dose)



SEKVENSLISTE

<110> Beaton, Andrew
 Ertl, Peter Franz
 Gough, Gerald Wayne
 Lear, Andrew
 Tite, John Philip
 Van Wely, Catherine Ann

<120> HIV-GAG CODON-OPTIMISED DNA VACCINES

<130> PG4653C1

<140> US 12/015,756

<141> 2008-01-17

<150> US 10/490,011

<151> 2004-10-25

<150> PCT/EP02/10592

<151> 2002-09-18

<150> PCT/GB01/04207

<151> 2001-09-20

<150> GB 0129604.5

<151> 2001-12-11

<150> GB 0206462.4

<151> 2002-03-19

<160> 87

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Nef primer

<400> 1

ataagaatgc ggccgccatg gtgggttttc cagtcacacc tt

42

<210> 2

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> AStrNef primer

<400> 2

cgcggtatcct cagcagttct tgaagtactc c

31

<210> 3

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> srt primer

<400> 3

ataagaatgc ggccgccatg ggccccatta gccctattga gact

44

<210> 4

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Asrt primer

<400> 4

ataagaatgc ggccgccatg ggccccatta gccctattga gact

44

<210> 5

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> sp17p24 primer

<400> 5

ataagaatgc ggccgccatg ggtgcccag cttcggt

37

<210> 6

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> sp17p24 primer

<400> 6

tggggcccat caacactctg gctttgtgtc

30

<210> 7

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> linker

<400> 7

cagagtgttg atgggcccga ttagccctat

30

<210> 8

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> linker

<400> 8

aaccacccat atctaataat agtactttcc

30

<210> 9

<211> 32

<212> DNA

<213> Artificial Sequence

<220>
 <223> linker

 <400> 9
 ctatttttag atatggtggg ttttccagtc ac 32

 <210> 10
 <211> 31
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> linker

 <400> 10
 cgcggtatcct cagcagttct tgaagtactc c 31

 <210> 11
 <211> 37
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 11
 ataagaatgc ggccgcatg ggtgcccag cttcggt 37

 <210> 12
 <211> 51
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 12
 gcgcacgatac ttgttcaggc ccaggatgat ccaccgttta tagatttctc c 51

 <210> 13
 <211> 49
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 13
 atcctggggcc tgaacaagat cgtgcgcagc tactctccga catccatcc 49

 <210> 14
 <211> 30
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 14
 tgggggcccac caacactctg gctttgtgtc 30

 <210> 15
 <211> 68

<212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 15
 gaattcgcgg ccgcgatggg ccccatcagt cccatcgaga ccgtgccggt gaagctgaaa 60
 cccgggat 68

 <210> 16
 <211> 44
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 16
 ggtgtgactg gaaaaccac catcagcacc tttctaatacc ccgc 44

 <210> 17
 <211> 23
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 17
 atggtgggtt ttccagtcac acc 23

 <210> 18
 <211> 29
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 18
 gatgaaatgc taggcggctg tcaaacctc 29

 <210> 19
 <211> 29
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 19
 gaggtttgac agccgcctag catttcac 29

 <210> 20
 <211> 31
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 20
 cgcggatcct cagcagttct tgaagtactc c 31

<210> 21
 <211> 23
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 21
 atggtgggtt ttccagtcac acc 23

 <210> 22
 <211> 29
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 22
 gatgaaatgc taggcggctg tcaaacctc 29

 <210> 23
 <211> 29
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 23
 gaggtttgac agccgcctag catttcac 29

 <210> 24
 <211> 31
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 24
 cgcggatcct cagcagttct tgaagtactc c 31

 <210> 25
 <211> 68
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 25
 gaattcgcgg ccgcgatggg ccccatcagt cccatcgaga ccgtgccggt gaagctgaaa 60
 cccgggat 68

 <210> 26
 <211> 39
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

<400> 26
 ggagctcgta gcccatcttc aggaatggcg gctccttct 39

<210> 27
 <211> 68
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 27
 gaattcggat ccttacagca cttttctaata ccccgcactc accagcttgt cgacctgctc 60
 gttgccgc 68

<210> 28
 <211> 26
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 28
 cctgaagatg ggctacgagc tccatg 26

<210> 29
 <211> 32
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 29
 cattagagcg gccgcgatgg tgggttttcc ac 32

<210> 30
 <211> 42
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 30
 gatgggactg atggggccca tgcagttctt gaactactcc gg 42

<210> 31
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 31
 atggggccca tcagtcccat cgag 24

<210> 32
 <211> 45
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

 <400> 32
 cagtaccgaa gctcgggcac ccatcagcac ctttctaatac cccgc 45

 <210> 33
 <211> 24
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 33
 atgggtgccc gagcttcggt actg 24

 <210> 34
 <211> 36
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 34
 gatgggggat cctcacaaca ctctggcttt gtgtcc 36

 <210> 35
 <211> 24
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 35
 atgggtgccc gagcttcggt actg 24

 <210> 36
 <211> 68
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 36
 gaattcggat ccttacagca cttttctaata ccccgcaactc accagcttgt cgacctgctc 60
 gttgccgc 68

 <210> 37
 <211> 32
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 37
 cattagagcg gccgcgatgg tgggttttcc ac 32

 <210> 38

<211> 45
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 38
 cagtaccgaa gctcgggcac ccatgcagtt cttgaactac tccgg 45

 <210> 39
 <211> 24
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 39
 atgggtgccc gagcttcggt actg 24

 <210> 40
 <211> 68
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 40
 gaattcgcg cgcgatggg ccccatcagt cccatcgaga ccgtgccggt gaagctgaaa 60
 cccgggat 68

 <210> 41
 <211> 45
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 41
 cagtaccgaa gctcgggcac ccatcagcac ctttctaata cccgc 45

 <210> 42
 <211> 68
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 42
 gaattcgcg cgcgatggg ccccatcagt cccatcgaga ccgtgccggt gaagctgaaa 60
 cccgggat 68

 <210> 43
 <211> 45
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

<400> 43
 cagtaccgaa gctcgggcac ccatgcagtt cttgaactac tccgg 45

 <210> 44
 <211> 24
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 44
 atgggtgccc gagcttcggt actg 24

 <210> 45
 <211> 36
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 45
 gatgggggat cctcacaaca ctctggcttt gtgtcc 36

 <210> 46
 <211> 24
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 46
 atgggtgccc gagcttcggt actg 24

 <210> 47
 <211> 42
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 47
 gatgggactg atggggccca tgcagttctt gaactactcc gg 42

 <210> 48
 <211> 24
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR primer

 <400> 48
 atgggccccca tcagtcccat cgag 24

 <210> 49
 <211> 68
 <212> DNA
 <213> Artificial Sequence

 <220>

<223> PCR primer

<400> 49

```
gaattcggat cttacagca cttttctaat ccccgcactc accagcttgt cgacctgctc 60
gttgccgc                                          68
```

<210> 50

<211> 1503

<212> DNA

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> HIV

<400> 50

atgggtgcc	gagcttcggt	actgtctggt	ggagagctgg	acagatggga	gaaaattagg	60
ctgcgcccgg	gaggcaaaaa	gaaatacaag	ctcaagcata	tcgtgtgggc	ctcgagggag	120
cttgaacggt	ttgccgtgaa	cccaggcctg	ctggaaacat	ctgagggatg	tcgccagatc	180
ctggggcaat	tgcagccatc	cctccagacc	gggagtgaag	agctgaggtc	cttgataaac	240
acagtggcta	ccctctactg	cgtaaccag	aggatcgaga	ttaaggatac	caaggaggcc	300
ttggacaaaa	ttgaggagga	gcaaaacaag	agcaagaaga	aggcccagca	ggcagctgct	360
gacactgggc	atagcaacca	ggtatcacag	aactatccta	ttgtccaaaa	cattcagggc	420
cagatggttc	atcaggccat	cagcccccg	acgtcfaatg	cctgggtgaa	ggttgtcgaa	480
gagaaggcct	tttctcctga	ggttatcccc	atgttctccg	ctttgagtga	gggggccact	540
cctcaggacc	tcaatacaat	gcttaataac	gtgggcggcc	atcaggccgc	catgcaaatg	600
ttgaaggaga	ctctcaacga	ggaggcagcc	gagtgggaca	gagtgcattc	cgtccacgct	660
ggcccaatcg	gcctccgaca	gatgcgggag	cctgcggcct	ctgacattgc	cggcaccacc	720
tctacactgc	aagagcaaat	cggatggatg	accaacaatc	ctccatccc	agttggagaa	780
atctataaac	ggtggatcat	tctcgggtct	aataaaattg	ttagaatgta	ctctccgaca	840
tccatccttg	acattagaca	gggacccaaa	gagcctttta	gggattacgt	cgaccggttt	900
tataagacc	tgcgagcaga	gcaggcctct	caggaggtca	aaaactggat	gacggagaca	960
ctcctggtac	agaacgctaa	ccccgactgc	aaaacaatct	tgaaggcact	aggcccggtc	1020
gccaccctgg	aagagatgat	gaccgcctgt	cagggagtag	gcggaccggg	acacaaagcc	1080
agagtgttg	ccgaagccat	gagccagggt	acgaactccg	caaccatcat	gatgcagaga	1140
gggaacttcc	gcaatcagcg	gaagatcgtg	aagtgtttca	attgcggcaa	ggagggtcac	1200
accgcccgca	actgtcgggc	ccctaggaag	aaagggtgtt	ggaagtgcgg	caaggaggga	1260
caccagatga	aagactgtac	agaacgcacg	gccaattttc	ttggaaaagat	ttggccgagc	1320
tacaagggga	gacctggtaa	tttcttgcaa	agcaggcccg	agccccaccg	cccccttgag	1380
gaatccttca	ggtccggagt	ggagaccaca	acgcctcccc	aaaaacagga	accaatcgac	1440
aagagctgt	accctttaac	ttctctgcgt	tctctctttg	gcaacgacc	gtcgtctcaa	1500
ttaa						1503

<210> 51

<211> 500

<212> PRT

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> HIV

<400> 51

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

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

<210> 52
 <211> 1515
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 52
 atgggtgcga gagcgtcagt attaagcggg ggagaattag atcgatggga aaaaattcgg 60
 ttaaggccag ggggaagaa aaaatataaa ttaaacata tagtatgggc aagcaggag 120

```

ctagaacgat tgcagtttaa tcctggcctg ttagaaacat cagaaggctg tagacaaata 180
ctgggacagc tacaaccatc ccttcagaca ggatcagaag aacttagatc attatataat 240
acagtagcaa ccctctattg tgtgcatcaa aggatagaga taaaagacac caagggaagct 300
ttagacaaga tagaggaaga gcaaaacaaa agtaagaaaa aagcacagca agcagcagct 360
gacacaggac acagcaatca ggctcagccaa aattacccta tagtgcagaa catccagggg 420
caaatggtac atcaggccat atcacctaga actttaaatg catgggtaaa agtagtagaa 480
gagaaggctt tcagcccaga agtgataccc atgttttcag cattatcaga aggagccacc 540
ccacaagatt taaacaccat gctaaacaca gtggggggac atcaagcagc catgcaaagt 600
ttaaagaga ccatcaatga ggaagctgca gaatgggata gagtgcattc agtgcattgc 660
gggcctattg caccaggcca gatgagagaa ccaaggggaa gtgacatagc aggaactact 720
agtacccttc aggaacaaat aggatggatg acaataatc cacctatccc agtaggagaa 780
atttataaaa gatggataat cctgggatta aataaaatag taagaatgta tagccctacc 840
agcattcttg acataagaca aggacaaaaa gaacccttta gagactatgt agaccgggtc 900
tataaaactc taagagccga gcaagcttca caggaggtaa aaaattggat gacagaaacc 960
ttgttggtcc aaaatgcgaa ccagattgt aagactatct taaaagcatt gggaccagcg 1020
gctacactag aagaaatgat gacagcatgt cagggagtag gaggacccgg ccataaggca 1080
agagtttttg tgggttttcc agtcacacct caggtagctt taagaccaat gacttacaag 1140
gcagctgtag atcttagcca ctttttaaaa gaaaaggggg gactggaagg gctaattcac 1200
tcccaagaaa gacaagatat ccttgatctg tggatctacc acacacaagg ctacttccct 1260
gattggcaga actacacacc agggccaggg gtcagatatc cactgacctt tggatggtgc 1320
tacaagctag taccagttga gccagataag gtagaagagg ccaataaagg agagaacacc 1380
agcttggttac accctgtgag cctgcatggg atggatgacc cggagagaga agtggttagag 1440
tggaggtttg acagccacct agcatttcat cacgtggccc gagagctgca tccggagtac 1500
ttcaagaact gctga 1515

```

<210> 53

<211> 504

<212> PRT

<213> Artificial Sequence

<220>

<223> HIV

<400> 53

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

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

<210> 54
<211> 1518
<212> DNA
<213> Artificial Sequence

<220>
<223> HIV

<400> 54
atgggtgccc gagcttcggt actgtctggt ggagagctgg acagatggga gaaaattagg 60
ctgcgcccgg gaggcaaaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgagggag 120
cttgaacggt ttgccgtgaa cccaggcctg ctggaaacat ctgagggatg tcgccagatc 180
ctggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggtc cttgtataac 240
acagtggcta ccctctactg cgtaaccag aggatcgaga ttaaggatag caaggaggcc 300
ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga aggccagca ggcagctgct 360
gacactgggc atagcaacca ggtatcacag aactatccta ttgtccaaaa cattcagggc 420
cagatgggtc atcaggccat cagcccccg acgctcaatg cctgggtgaa ggttgtcgaa 480
gagaaggcct tttctcctga ggttatcccc atgttctcgc ctttgagtga gggggccact 540
cctcaggacc tcaatacaat gcttaatacc gtgggaggcc atcaggccgc catgcaaatg 600
ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gagtgcattcc cgtccacgct 660
ggcccaatcg cgcccgga gaatgcgggag cctcgcggtc ctgacattgc cggcaccacc 720
tctacactgc aagagcaaat cggatggatg accaacaatc ctcccatccc agttggagaa 780
atctataaac ggtggatcat tctcgtctc aataaaattg ttagaatgta ctctccgaca 840
tccatccttg acattagaca gggacccaaa gagcctttta gggattacgt cgaccgggtt 900
tataagaccc tgcgagcaga gcaggcctct caggaggtca aaaactggat gacggagaca 960
ctcctggtac agaacgctaa ccccgactgc aaaacaatct tgaaggcact agggccggtc 1020
gccaccctgg aagagatgat gaccgcctgt cagggagtag gcggacccgg acacaaagcc 1080
agagtgttga tgggtgggtt tccagtcaca cctcagggtac ctttaagacc aatgacttac 1140

```

aaggcagctg tagatcttag ccacttttta aaagaaaagg ggggactgga agggctaatt 1200
cactcccaaa gaagacaaga tacccttgat ctgtggatct accacacaca aggctacttc 1260
cctgattggc agaactacac accagggcca ggggtcagat atccactgac ctttggatgg 1320
tgctacaagc tagtaccagt tgagccagat aaggtagaag aggccaataa aggagagaac 1380
accagcttgt tacaccctgt gagcctgcat gggatggatg acccggagag agaagtgtta 1440
gagtggagggt ttgacagcca cctagcattt catcacgtgg cccgagagct gcatccggag 1500
tacttcaaga actgctga                                     1518

```

<210> 55

<211> 505

<212> PRT

<213> Artificial Sequence

<220>

<223> HIV

<400> 55

```

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

```

370 375 380
Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile
385 390 395 400
His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu Trp Ile Tyr His Thr
405 410 415
Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Val
420 425 430
Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys Leu Val Pro Val Glu
435 440 445
Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu Asn Thr Ser Leu Leu
450 455 460
His Pro Val Ser Leu His Gly Met Asp Asp Pro Glu Arg Glu Val Leu
465 470 475 480
Glu Trp Arg Phe Asp Ser His Leu Ala Phe His His Val Ala Arg Glu
485 490 495
Leu His Pro Glu Tyr Phe Lys Asn Cys
500 505

<210> 56
<211> 1689
<212> DNA
<213> Artificial Sequence

<220>
<223> HIV

<400> 56
atggggcccca tcagtcccat cgagaccgtg ccggtgaagc tgaaacccgg gatggacggc 60
cccaagggtca agcagtggcc actcaccgag gagaagatca aggccctggt ggagatctgc 120
accgagatgg agaaagaggg caagatcagc aagatcgggc ctgagaaccc atacaacacc 180
cccggtgtttg ccatcaagaa gaaggacagc accaagtggc gcaagctggt ggatttccgg 240
gagctgaata agcggaccca ggatttctgg gaggtccagc tgggcatccc ccatccggcc 300
ggcctgaaga agaagaagag cgtgaccgtg ctggacgtgg gcgacgctta cttcagcgtc 360
cctctggacg aggacttttag aaagtacacc gcctttacca tcccatctat caacaacgag 420
acccctggga tcagatatca gtacaacgtc ctcccccagg gctggaaggg ctctcccggc 480
attttccaga gctccatgac caagatcctg gagccgtttc ggaagcagaa ccccgatatc 540
gtcatctacc agtacatgga cgacctgtac gtgggctctg acctggaaat cgggcagcat 600
cgcacgaaga ttgaggagct gaggcagcat ctgctgagat ggggcctgac cactccggac 660
aagaagcatc agaaggagcc gccattcctg tggatgggct acgagctcca tcccagacaag 720
tggaccgtgc agcctatcgt cctccccgag aaggacagct ggaccgtgaa cgacatccag 780
aagctggtgg gcaagctcaa ctgggctagc cagatctatc ccgggatcaa ggtgcgccag 840
ctctgcaagc tgctgcgcgg caccaaggcc ctgaccgagg tgattcccct cacggaggaa 900
gccgagctcg agctggctga gaaccgggag atcctgaagg agcccgtgca cggcgtgtac 960
tatgacccct ccaaggacct gatcgccgaa atccagaagc agggccaggg gcagtggaca 1020
taccagattt accaggagcc tttcaagaac ctcaagaccg gcaagtacgc ccgcatgagg 1080
ggcgcccaca ccaacgatgt caagcagctg accgaggccg tccagaagat caccgaccgag 1140
tccatcgtga tctgggggaa gacacccaag ttcaagctgc ctatccagaa ggagacctgg 1200
gagacgtggt ggaccgaata ttggcaggcc acctggattc ccgagtggga gttcgtgaat 1260
acacctctc tggtgaagct gtggtaccag ctcgagaagg agcccatcgt gggcgcgagg 1320
acattctacg tggacggcgc ggccaaccgc gaaacaaagc tcgggaaggc cgggtacgtc 1380
accaaccggg gccgccagaa ggtcgtcacc ctgaccgaca ccaccaacca gaagacggag 1440
ctgcaggcca tctatctcgc tctccaggac tccggcctgg aggtgaacat cgtgacggac 1500
agccagtagc cgctgggcat tattcaggcc cagccggacc agtccgagag cgaactggtg 1560
aaccagatta tcgagcagct gatcaagaaa gagaagggtc acctcgctg ggtcccggcc 1620
cataagggca ttggcggcaa cgagcaggtc gacaagctgg tgagtgcggg gattagaaag 1680
gtgctgtaa 1689

<210> 57
<211> 562
<212> PRT
<213> Artificial Sequence

<220>

<223> HIV

<400> 57

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

[illegible]

```
<210> 58
<211> 1689
<212> DNA
<213> Artificial Sequence
```

<220>
<223> HIV

<400> 58						
atggggcccca	tcagtcccat	cgagaccgtg	cgggtgaagc	tgaaaccccg	gatggacggc	60
cccaaggtca	agcagtggcc	actcaccgag	gagaagatca	aggccctggt	ggagatctgc	120
accgagatgg	agaaagaggg	caagatcagc	aagatcgggc	ctgagaaccc	atacaacacc	180
cccgtgtttg	ccatcaagaa	gaaggacagc	accaagtggc	gcaagctggt	ggatttccgg	240
gagctgaata	agcggaccca	ggattttctg	gaggtccagc	tgggcatccc	ccatccggcc	300
ggcctgaaga	agaagaagag	cgtgaccgtg	ctggacgtgg	gcgacgctta	cttcagcgtc	360
cctctggacg	aggacttttag	aaagtacacc	gcctttacca	tcccatctat	caacaacgag	420
acccctggca	tcagatatca	gtacaacgtc	ctccccagg	gctggaaggg	ctctcccgcc	480
attttccaga	gctccatgac	caagatcctg	gagccgtttc	ggaagcagaa	ccccgatatc	540
gtcatctacc	agtacatgga	cgacctgtac	gtgggctctg	acctggaaat	cgggcagcat	600
cgcacgaaga	ttgaggagct	gaggcagcat	ctgctgagat	ggggcctgac	cactccggac	660
aagaagcatc	agaaggagcc	gccattcctg	tgggtgggct	acgagctcca	tcccgacaag	720
tggaccgtgc	agcctatcgt	ctctcccag	aaggacagct	ggaccgtgaa	cgacatccag	780
aagctggtgg	gcaagctcaa	ctgggctagc	cagatctatc	ccgggatcaa	ggtgcgccag	840
ctctgcaagc	tgctgcgcgg	caccaaggcc	ctgaccgagg	tgattccccct	cacggaggaa	900
gccgagctcg	agctggctga	gaaccgggag	atcctgaagg	agcccgtgca	cggcgtgtac	960
tatgacccct	ccaaggacct	gatcgccgaa	atccagaagc	agggccaggg	gcagtggaca	1020
taccagattt	accaggagcc	tttcaagaac	ctcaagaccg	gcaagtacgc	ccgcatgagg	1080
ggcgcccaca	ccaacgatgt	caagcagctg	accgaggccg	tccagaagat	cacgaccgag	1140
tccatcgtga	tctgggggaa	gacacccaag	ttcaagctgc	ctatccagaa	ggagacctgg	1200
gagacgtggt	ggaccgaata	ttggcaggcc	acctggattc	ccgagtggga	gttcgtgaat	1260
acacctcctc	tggtgaagct	gtggtaccag	ctcgagaagg	agcccatcgt	gggcgcggag	1320
acattctacg	tggacggcgc	ggccaaccgc	gaaacaaagc	tcgggaaggc	cgggtacgtc	1380
accaaccggg	gccgccagaa	ggtcgtcacc	ctgaccgaca	ccaccaacca	gaagacggag	1440
ctgcaggcca	tctatctcgc	tctccaggac	tccggcctgg	aggtgaacat	cgtgacggac	1500
agccagtagc	cgctgggcat	tattcaggcc	cagccggacc	agtccgagag	cgaactgggtg	1560
aaccagatta	tcgagcagct	gatcaagaaa	gagaaggtct	acctgcctg	ggtcccggcc	1620
cataagggca	ttggcggcaa	cgagcaggtc	gacaagctgg	tgagtgcggg	gattagaaag	1680
gtgctgtaa						1689

```
<210> 59
<211> 562
<212> PRT
<213> Artificial Sequence
```

<220>
<223> HIV

<400> 59
Met Gly Pro Ile Ser Pro Ile Glu Thr Val Ser Val Lys Leu Lys Pro
1 5 10 15
Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys

			20					25				30			
Ile	Lys	Ala	Leu	Val	Glu	Ile	Cys	Thr	Glu	Met	Glu	Lys	Glu	Gly	Lys
		35					40					45			
Ile	Ser	Lys	Ile	Gly	Pro	Glu	Asn	Pro	Tyr	Asn	Thr	Pro	Val	Phe	Ala
	50					55					60				
Ile	Lys	Lys	Lys	Asp	Ser	Thr	Lys	Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg
65				70					75						80
Glu	Leu	Asn	Lys	Arg	Thr	Gln	Asp	Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile
			85					90						95	
Pro	His	Pro	Ala	Gly	Leu	Lys	Lys	Lys	Lys	Ser	Val	Thr	Val	Leu	Asp
			100					105					110		
Val	Gly	Asp	Ala	Tyr	Phe	Ser	Val	Pro	Leu	Asp	Glu	Asp	Phe	Arg	Lys
		115						120				125			
Tyr	Thr	Ala	Phe	Thr	Ile	Pro	Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile
	130					135					140				
Arg	Tyr	Gln	Tyr	Asn	Val	Leu	Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala
145					150					155					160
Ile	Phe	Gln	Ser	Ser	Met	Thr	Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln
			165						170					175	
Asn	Pro	Asp	Ile	Val	Ile	Tyr	Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly
		180						185					190		
Ser	Asp	Leu	Glu	Ile	Gly	Gln	His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg
	195						200					205			
Gln	His	Leu	Leu	Arg	Trp	Gly	Leu	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln
	210					215					220				
Lys	Glu	Pro	Pro	Phe	Leu	Trp	Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys
225					230					235					240
Trp	Thr	Val	Gln	Pro	Ile	Val	Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val
			245						250					255	
Asn	Asp	Ile	Gln	Lys	Leu	Val	Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile
		260						265					270		
Tyr	Pro	Gly	Ile	Lys	Val	Arg	Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr
	275						280					285			
Lys	Ala	Leu	Thr	Glu	Val	Ile	Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu
	290					295					300				
Leu	Ala	Glu	Asn	Arg	Glu	Ile	Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr
305					310					315					320
Tyr	Asp	Pro	Ser	Lys	Asp	Leu	Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln
			325						330					335	
Gly	Gln	Trp	Thr	Tyr	Gln	Ile	Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys
		340						345					350		
Thr	Gly	Lys	Tyr	Ala	Arg	Met	Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys
	355						360					365			
Gln	Leu	Thr	Glu	Ala	Val	Gln	Lys	Ile	Thr	Thr	Glu	Ser	Ile	Val	Ile
	370					375					380				
Trp	Gly	Lys	Thr	Pro	Lys	Phe	Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp
385					390					395					400
Glu	Thr	Trp	Trp	Thr	Glu	Tyr	Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp
			405						410					415	
Glu	Phe	Val	Asn	Thr	Pro	Pro	Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu
		420						425					430		
Lys	Glu	Pro	Ile	Val	Gly	Ala	Glu	Thr	Phe	Tyr	Val	Asp	Gly	Ala	Ala
	435						440					445			
Asn	Arg	Glu	Thr	Lys	Leu	Gly	Lys	Ala	Gly	Tyr	Val	Thr	Asn	Arg	Gly
	450					455					460				
Arg	Gln	Lys	Val	Val	Thr	Leu	Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu
465					470					475					480
Leu	Gln	Ala	Ile	Tyr	Leu	Ala	Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn
			485						490					495	
Ile	Val	Thr	Asp	Ser	Gln	Tyr	Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro
		500						505					510		
Asp	Gln	Ser	Glu	Ser	Glu	Leu	Val	Asn	Gln	Ile	Ile	Glu	Gln	Leu	Ile
	515						520					525			
Lys	Lys	Glu	Lys	Val	Tyr	Leu	Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile

530 535 540
 Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys
 545 550 555 560
 Val Leu

<210> 60
 <211> 429
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 60
 atggtgggtt ttccagtcac acctcaggtta cctttaagac caatgactta caaggcagct 60
 gtagatctta gccacttttt aaaagaaaag gggggactgg aagggtctaat tcactcccaa 120
 agaagacaag atatccttga tctgtggatc taccacacac aaggctactt ccctgattgg 180
 cagaactaca caccagggcc aggggtcaga tatccactga cctttggatg gtgctacaag 240
 ctagtaccag ttgagccaga taaggtagaa gagggccaata aaggagagaa caccagcttg 300
 ttacaccctg tgagcctgca tgggatggat gacccggaga gagaagtgtt agagtggagg 360
 tttgacagcc acctagcatt tcatcacgtg gcccgagagc tgcattccgga gtacttcaag 420
 aactgctga 429

<210> 61
 <211> 142
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 61
 Met Val Gly Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr
 1 5 10 15
 Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly
 20 25 30
 Leu Glu Gly Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu
 35 40 45
 Trp Ile Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
 50 55 60
 Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys
 65 70 75 80
 Leu Val Pro Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu
 85 90 95
 Asn Thr Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro
 100 105 110
 Glu Arg Glu Val Leu Glu Trp Arg Phe Asp Ser Val Leu Ala Phe His
 115 120 125
 His Val Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys
 130 135 140

<210> 62
 <211> 1698
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 62
 atggggcccca ttagccctat tgagactgtg tcagtaaaat taaagccagg aatggatggc 60

ccaaaagtta aacaatggcc attgacagaa gaaaaaataa aagcattagt agaaatttgt 120
acagagatgg aaaaggaagg gaaaatttca aaaattgggc ctgaaaatcc atacaatact 180
ccagtatttg ccataaagaa aaaagacagt actaaatgga gaaaattagt agatttcaga 240
gaacttaata agagaactca agacttctgg gaagttcaat taggaatacc acatcccgca 300
gggttaaaaa agaaaaaatc agtaacagta ctggatgtgg gtgatgcata tttttcagtt 360
cccttagatg aagacttcag gaaatatact gcattttacca tacctagtat aaacaatgag 420
acaccagggg ttagatatca gtacaatgtg cttccacagg gatggaaagg atcaccagca 480
atattccaaa gtagcatgac aaaaatctta gagcctttta gaaaacaaaa tccagacata 540
gttatctatc aatacatgga tgatttgtat gtaggatctg acttagaaat agggcagcat 600
agaacaaaaa tagaggagct gagacaacat ctggtgaggt ggggacttac cacaccagac 660
aaaaaacatc agaaagaacc tccattcctt tggatgggtt atgaactcca tcctgataaa 720
tggacagtac agcctatagt gctgccagaa aaagacagct ggactgtcaa tgacatacag 780
aagttagtgg ggaaattgaa ttgggcaagt cagatttacc cagggattaa agtaaggcaa 840
ttatgtaaac tccttagagg aaccaaagca ctaacagaag taataccact aacagaagaa 900
gcagagctag aactggcaga aaacagagag attctaaaag aaccagtaca tggagtgtat 960
tatgacccat caaaagactt aatagcagaa atacagaagc aggggcaagg ccaatggaca 1020
tatcaaattt atcaagagcc atttaaaaat ctgaaaacag gaaaatatgc aagaatgagg 1080
ggtgccaca ctaatgatgt aaaacaatta acagaggcag tgcaaaaaat aaccacagaa 1140
agcatagtaa tatggggaaa gactcctaaa tttaaactgc ccatacaaaa ggaaacatgg 1200
gaaacatggt ggacagagta ttggcaagcc acctggattc ctgagtggga gtttgtaaat 1260
accctccct tagtgaaatt atgggtaccag ttagagaaag aacctatagt agggcagaa 1320
accttctatg tagatggggc agctaacagg gagactaaat taggaaaagc aggatatgtt 1380
actaatagag gaagacaaaa agttgtcacc ctaactgaca caacaaatca gaagactgag 1440
ttacaagcaa tttatctagc tttgcaggat tccggattag aagtaaacat agtaacagac 1500
tcacaatatg cattaggaat cattcaagca caaccagatc aaagtgaatc agagttagtc 1560
aatcaaataa tagagcagtt aataaaaaag gaaaaggtct atctggcatg ggtaccagca 1620
cacaagggaa ttggaggaaa tgaacaagta gataaattag tcagtgtctg aatcaggaaa 1680
gtactatttt tagattaa 1698

<210> 63
<211> 565
<212> PRT
<213> Artificial Sequence

<220>
<223> HIV

<400> 63
Met Gly Pro Ile Ser Pro Ile Glu Thr Val Ser Val Lys Leu Lys Pro
1 5 10 15
Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys
20 25 30
Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys
35 40 45
Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala
50 55 60
Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg
65 70 75 80
Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile
85 90 95
Pro His Pro Ala Gly Leu Lys Lys Lys Ser Val Thr Val Leu Asp
100 105 110
Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys
115 120 125
Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile
130 135 140
Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala
145 150 155 160
Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln
165 170 175
Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly
180 185 190
Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg
195 200 205

Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln
210 215 220
Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys
225 230 235 240
Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val
245 250 255
Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile
260 265 270
Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr
275 280 285
Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu
290 295 300
Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr
305 310 315 320
Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln
325 330 335
Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys
340 345 350
Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys
355 360 365
Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile
370 375 380
Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp
385 390 395 400
Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp
405 410 415
Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu
420 425 430
Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Asp Gly Ala Ala
435 440 445
Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly
450 455 460
Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Glu
465 470 475 480
Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn
485 490 495
Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro
500 505 510
Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile
515 520 525
Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile
530 535 540
Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys
545 550 555 560
Val Leu Phe Leu Asp
565

<210> 64
<211> 3213
<212> DNA
<213> Artificial Sequence

<220>
<223> HIV

<400> 64
atgggtgccc gagcttcggt actgtctggt ggagagctgg acagatggga gaaaattagg 60
ctgcgcccgg gaggcaaaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgaggagg 120
cttgaacggt ttgccgtgaa cccaggcctg ctggaaacat ctgagggatg tcgccagatc 180
ctggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggtc cttgtataac 240
acagtggcta ccctctactg cgtacaccag aggatcgaga ttaaggatac caaggaggcc 300
ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga aggcccagca ggcagctgct 360
gacactgggc atagcaacca ggtatcacag aactatccta ttgtccaaaa cattcagggc 420

cagatgggttc atcaggccat cagcccccg acgctcaatg cctgggtgaa ggttgctgaa 480
gagaaggcct tttctcctga gggtatcccc atgttctccg ctttgagtga gggggccact 540
cctcaggacc tcaatacaat gcttaatacc gtgggcggcc atcaggccgc catgcaaagt 600
ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gagtgcattcc cgtccacgct 660
ggcccaatcg cgcccgga ca gatgcgggag cctcgcggct ctgacattgc cggcaccacc 720
tctacactgc aagagcaaat cggatggatg accaacaatc ctcccatccc agttggagaa 780
atctataaac ggtggatcat tctcgtctc aataaaattg ttagaatgta ctctccgaca 840
tccatccttg acattagaca gggacccaaa gagcctttta gggattacgt cgaccgggtt 900
tataagaccc tgcgagcaga gcaggcctct caggagggtca aaaactggat gacggagaca 960
ctcctggtac agaacgctaa ccccgactgc aaaacaatct tgaaggcact agggccggct 1020
gccaccctgg aagagatgat gaccgcctgt cagggagtag gcggacccgg acacaaagcc 1080
agagtgttga tgggccccat tagccctatt gagactgtgt cagtaaaatt aaagccagga 1140
atggatggcc caaaagttaa acaatggcca ttgacagaag aaaaaataaa agcattagta 1200
gaaatttgta cagagatgga aaaggaaggg aaaatttcaa aaattggggc tgaaaatcca 1260
tacaatactc cagtatttgc cataaagaaa aaagacagta ctaaattggag aaaattagta 1320
gatttcagag aacttaataa gagaactcaa gacttctggg aagttcaatt aggaatacca 1380
catcccgca ggttaaaaaa gaaaaaatca gtaacagtac tggatgtggg tgatgcata 1440
ttttcagttc ccttagatga agacttcagg aatatactg catttaccat acctagtata 1500
aacaatgaga caccagga tagatatcag tacaatgtgc ttccacaggg atggaaagga 1560
tcaccagcaa tattccaaag tagcatgaca aaaatcttag agccttttag aaaacaaaat 1620
ccagacatag ttatctatca atacatggat gatttgtatg taggatctga cttagaaata 1680
gggcagcata gaacaaaaat agaggagctg agacaacatc tgttgagggtg gggacttacc 1740
acaccagaca aaaaacatca gaaagaacct ccattccttt ggatgggtta tgaactccat 1800
cctgataaat ggacagtaca gcctatagtg ctgccagaaa aagacagctg gactgtcaat 1860
gacatacaga agttagtggg gaaattgaat tgggcaagtc agatttacc agggattaaa 1920
gtaaggcaat tatgtaaact ccttagagga accaaagcac taacagaagt aataccacta 1980
acagaagaag cagagctaga actggcagaa aacagagaga ttctaaaaga accagtacat 2040
ggagtgtatt atgacccatc aaaagactta atagcagaaa tacagaagca ggggcaaggc 2100
caatggacat atcaaattta tcaagagcca tttaaaaatc tgaaaacagg aaaatagtga 2160
agaatgaggg gtgcccacac taatgatgta aaacaattaa cagaggcagt gcaaaaaata 2220
accacagaaa gcatagtaat atggggaaag actcctaaat ttaaactgcc catacaaaag 2280
gaaacatggg aaacatggtg gacagagtat tggcaagcca cctggattcc tgagtgggag 2340
tttgtttaata cccctccctt agtgaaatta tggtagcagt tagagaaaga acccatagta 2400
ggagcagaaa ccttctatgt agatggggca gctaacaggg agactaaatt aggaaaagca 2460
ggatatgtta ctaatagagg aagacaaaaa gttgtcaccc taactgacac aacaaatcag 2520
aagactgagt tacaagcaat ttatctagct ttgcaggatt cgggattaga agtaaacata 2580
gtaacagact cacaatatgc attaggaatc attcaagcac aaccagatca aagtgaatca 2640
gagttagtca atcaataat agagcagtta ataaaaaagg aaaaggctta tctggcatgg 2700
gtaccagcac acaaaggaat tggaggaaat gaacaagtag ataaattagt cagtgcctga 2760
atcaggaaag tactattttt agatatggtg gggtttccag tcacacctca ggtaccttta 2820
agaccaatga cttacaaggc agctgtagat cttagccact ttttaaaaga aaagggggga 2880
ctggaagggc taattcactc ccaaagaaga caagatatcc ttgatctgtg gatctaccac 2940
acacaaggct acttccctga ttggcagaac tacacaccag ggccaggggt cagatatcca 3000
ctgacctttg gatggtgcta caagctagta ccagttgagc cagataaggt agaagaggcc 3060
aataaaggag agaaccacag cttgttacac cctgtgagcc tgcattgggt ggatgacctg 3120
gagagagaag tgtagagtg gaggtttgac agccacctag catttcatca cgtggcccca 3180
gagctgcac cggagtactt caagaactgc tga 3213

<210> 65
<211> 1070
<212> PRT
<213> Artificial Sequence

<220>
<223> HIV

<400> 65
Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp
1 5 10 15
Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys
20 25 30
His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro
35 40 45
Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu

50	55	60
Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn		
65	70	75
Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp		80
	85	90
Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys		95
	100	105
Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn Gln Val		110
	115	120
Ser Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His		125
	130	135
Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu		140
	145	150
Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser		155
	165	170
Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly		175
	180	185
Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu		190
	195	200
Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala		205
	210	215
Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr		220
	225	230
Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile		235
	245	250
Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys		255
	260	265
Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly		270
	275	280
Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu		285
	290	295
Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr		300
	305	310
Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala		315
	325	330
Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly		335
	340	345
Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Met Gly Pro Ile Ser		350
	355	360
Pro Ile Glu Thr Val Ser Val Lys Leu Lys Pro Gly Met Asp Gly Pro		365
	370	375
Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val		380
	385	390
Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly		395
	405	410
Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp		415
	420	425
Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg		430
	435	440
Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly		445
	450	455
Leu Lys Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr		460
	465	470
Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr		475
	485	490
Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn		495
	500	505
Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser		510
	515	520
Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val		525
	530	535
Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile		540
	545	550
Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg		555
	560	

				565					570					575			
Trp	Gly	Leu	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe		
			580					585					590				
Leu	Trp	Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro		
		595					600					605					
Ile	Val	Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys		
	610					615					620						
Leu	Val	Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro	Gly	Ile	Lys		
625					630					635					640		
Val	Arg	Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala	Leu	Thr	Glu		
			645						650					655			
Val	Ile	Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg		
		660						665					670				
Glu	Ile	Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys		
	675						680					685					
Asp	Leu	Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr		
	690					695					700						
Gln	Ile	Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala		
705					710					715					720		
Arg	Met	Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala		
			725						730					735			
Val	Gln	Lys	Ile	Thr	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro		
		740						745					750				
Lys	Phe	Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr	Trp	Trp	Thr		
	755					760					765						
Glu	Tyr	Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe	Val	Asn	Thr		
	770					775					780						
Pro	Pro	Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu	Lys	Glu	Pro	Ile	Val		
785					790					795					800		
Gly	Ala	Glu	Thr	Phe	Tyr	Val	Asp	Gly	Ala	Ala	Asn	Arg	Glu	Thr	Lys		
			805					810						815			
Leu	Gly	Lys	Ala	Gly	Tyr	Val	Thr	Asn	Arg	Gly	Arg	Gln	Lys	Val	Val		
		820						825					830				
Thr	Leu	Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu	Leu	Gln	Ala	Ile	Tyr		
	835					840						845					
Leu	Ala	Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn	Ile	Val	Thr	Asp	Ser		
	850					855					860						
Gln	Tyr	Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro	Asp	Gln	Ser	Glu	Ser		
865					870					875					880		
Glu	Leu	Val	Asn	Gln	Ile	Ile	Glu	Gln	Leu	Ile	Lys	Lys	Glu	Lys	Val		
			885					890						895			
Tyr	Leu	Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	Gly	Gly	Asn	Glu	Gln		
	900							905					910				
Val	Asp	Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys	Val	Leu	Phe	Leu	Asp		
	915						920					925					
Met	Val	Gly	Phe	Pro	Val	Thr	Pro	Gln	Val	Pro	Leu	Arg	Pro	Met	Thr		
	930					935					940						
Tyr	Lys	Ala	Ala	Val	Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly		
945					950					955					960		
Leu	Glu	Gly	Leu	Ile	His	Ser	Gln	Arg	Arg	Gln	Asp	Ile	Leu	Asp	Leu		
			965						970					975			
Trp	Ile	Tyr	His	Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr		
		980						985						990			
Pro	Gly	Pro	Gly	Val	Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Tyr	Lys		
		995					1000						1005				
Leu	Val	Pro	Val	Glu	Pro	Asp	Lys	Val	Glu	Glu	Ala	Asn	Lys	Gly	Glu		
	1010					1015					1020						
Asn	Thr	Ser	Leu	Leu	His	Pro	Val	Ser	Leu	His	Gly	Met	Asp	Asp	Pro		
1025					1030					1035					1040		
Glu	Arg	Glu	Val	Leu	Glu	Trp	Arg	Phe	Asp	Ser	His	Leu	Ala	Phe	His		
			1045						1050					1055			
His	Val	Ala	Arg	Glu	Leu	His	Pro	Glu	Tyr	Phe	Lys	Asn	Cys				
			1060					1065					1070				

<210> 66
 <211> 3213
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 66

```

atgggtgccc gagcttcggt actgtctggt ggagagctgg acagatggga gaaaattagg 60
ctgcgcccgg gaggcaaaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgagggag 120
cttgaacggt ttgcogtgaa cccaggcctg ctggaaacat ctgagggatg tcgccagatc 180
ctggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggtc cttgtataac 240
acagtggcta ccctctactg cgtaaccag aggatcgaga ttaaggatac caaggaggcc 300
ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga aggccagca ggcagctgct 360
gacactgggc atagcaacca ggtatcacag aactatccta ttgtccaaaa cattcagggc 420
cagatggttc atcaggccat cagccccgg acgctcaatg cctgggtgaa ggttgtcgaa 480
gagaaggcct tttctcctga ggttatcccc atgttctcgg ctttgagtga gggggccact 540
cctcaggacc tcaatacaat gcttaatacc gtgggcggcc atcaggccgc catgcaaagt 600
ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gagtgcattc cgtccacgct 660
ggcccaatcg cgcccgga ca gatgcgggag cctcgcggct ctgacattgc cggcaccacc 720
tctacactgc aagagcaaat cggatggatg accaacaatc ctcccatccc agttggagaa 780
atctataaac ggtggatcat cctgggcctg aacaagatcg tgcgcatgta ctctccgaca 840
tccatccttg acattagaca gggacccaaa gagcctttta gggattacgt cgaccggttt 900
tataagaccc tgcgagcaga gcaggcctct caggaggtca aaaactggat gacggagaca 960
ctcctggtac agaacgctaa ccccgactgc aaaacaatct tgaaggcact aggcccggt 1020
gccaccctgg aagagatgat gaccgcctgt cagggagtag gcggaccgg acacaaagcc 1080
agagtgttga tgggccccat tagccctatt gagactgtgt cagtaaaatt aaagccagga 1140
atggatggcc caaaagttaa acaatggcca ttgacagaag aaaaaataaa agcattagta 1200
gaaatttgta cagagatgga aaaggaaggg aaaatttcaa aaattgggcc tgaaaatcca 1260
tacaatactc cagtatttgc cataaagaaa aaagacagta ctaaattggag aaaattagta 1320
gatttcagag aacttaataa gagaactcaa gacttctggg aagttcaatt aggaatacca 1380
catcccgcat ggttaaaaaa gaaaaaatca gtaacagtac tggatgtggg tgatgcatat 1440
ttttcagttc ccttagatga agacttcagg aaataactg catttaccat acctagtata 1500
aacaatgaga caccaggat tagatatcag tacaatgtgc ttccacaggg atggaaagga 1560
tcaccagcaa tattccaaag tagcatgaca aaactcttag agccttttag aaaacaaaa 1620
ccagacatag ttatctatca atacatggat gatttgtatg taggatctga cttagaaata 1680
gggcagcata gaacaaaaat agaggagctg agacaacatc tgttgagggtg gggacttacc 1740
acaccagaca aaaaacatca gaaagaacct ccattccttt ggatgggtta tgaactccat 1800
cctgataaat ggacagtaca gcctatagtg ctgccagaaa aagacagctg gactgtcaat 1860
gacatacaga agttagtggg gaaattgaat tgggcaagtc agatttacc agggattaaa 1920
gtaaggcaat tatgtaaact ccttagagga accaaagcac taacagaagt aataccacta 1980
acagaagaag cagagctaga actggcagaa aacagagaga ttctaaaaga accagtacat 2040
ggagtgtatt atgacccatc aaaagactta atagcagaaa tacagaagca ggggcaaggc 2100
caatggacat atcaaattta tcaagagcca tttaaaaatc tgaaaacagg aaaatatgca 2160
agaatgaggg gtgcccacac taatgatgta aaacaattaa cagaggcagt gcaaaaaata 2220
accacagaaa gcatagtaat atggggaaaag actcctaaat ttaaactgcc catacaaaag 2280
gaaacatggg aaacatggtg gacagagtat tggcaagcca cctggattcc tgagtgggag 2340
tttgtttaata cccctccctt agtgaaatta tggtaaccagt tagagaaaga acccatagta 2400
ggagcagaaa ccttctatgt agatggggca gctaacaggg agactaaatt aggaaaagca 2460
ggatatgtta ctaatagagg aagacaaaaa gttgtcacc taactgacac aacaaatcag 2520
aagactgagt tacaagcaat ttatctagct ttgcaggatt cgggattaga agtaaacata 2580
gtaacagact cacaataatg attaggaatc attcaagcac aaccagatca aagtgaatca 2640
gagttagtca atcaaataat agagcagtta ataaaaagg aaaaggtcta tctggcatgg 2700
gtaccagcac acaaaggaat tggaggaaat gaacaagtag ataaattagt cagtgcctgga 2760
atcaggaaaag tactattttt agatatggtg ggttttccag tcacacctca ggtaccttta 2820
agaccaatga cttacaaggc agctgtagat cttagccact ttttaaaaga aaagggggga 2880
ctggaagggc taattcactc ccaaagaaga caagatatcc ttgatctgtg gatctaccac 2940
acacaaggct acttcctga ttggcagaac tacacaccag ggccaggggt cagatatcca 3000
ctgacctttg gatggtgcta caagctagta ccagttgagc cagataaggt agaagaggcc 3060
aataaaggag agaaccacag cttgttacac cctgtgagcc tgcattgggt ggatgacctg 3120
gagagagaag tgttagagtg gaggtttgac agccacctag catttcatca cgtggcccg 3180
gagctgcatc cggagtactt caagaactgc tga
3213

```

<210> 67
 <211> 1070
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 67

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

Ser	Thr	Lys	Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg	435	440	445
Thr	Gln	Asp	Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	450	455	460
Leu	Lys	Lys	Lys	Lys	Ser	Val	Thr	Val	Leu	Asp	Val	Gly	Asp	Ala	Tyr	465	470	475
Phe	Ser	Val	Pro	Leu	Asp	Glu	Asp	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	485	490	495
Ile	Pro	Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	500	505	510
Val	Leu	Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Ser	Ser	515	520	525
Met	Thr	Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro	Asp	Ile	Val	530	535	540
Ile	Tyr	Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	545	550	555
Gly	Gln	His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Gln	His	Leu	Leu	Arg	565	570	575
Trp	Gly	Leu	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe		580	585	590
Leu	Trp	Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	595	600	605
Ile	Val	Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	610	615	620
Leu	Val	Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro	Gly	Ile	Lys	625	630	635
Val	Arg	Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala	Leu	Thr	Glu	645	650	655
Val	Ile	Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	660	665	670
Glu	Ile	Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	675	680	685
Asp	Leu	Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr	690	695	700
Gln	Ile	Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	705	710	715
Arg	Met	Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala	725	730	735
Val	Gln	Lys	Ile	Thr	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro	740	745	750
Lys	Phe	Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr	Trp	Trp	Thr	755	760	765
Glu	Tyr	Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe	Val	Asn	Thr	770	775	780
Pro	Pro	Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu	Lys	Glu	Pro	Ile	Val	785	790	795
Gly	Ala	Glu	Thr	Phe	Tyr	Val	Asp	Gly	Ala	Ala	Asn	Arg	Glu	Thr	Lys	805	810	815
Leu	Gly	Lys	Ala	Gly	Tyr	Val	Thr	Asn	Arg	Gly	Arg	Gln	Lys	Val	Val	820	825	830
Thr	Leu	Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu	Leu	Gln	Ala	Ile	Tyr	835	840	845
Leu	Ala	Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn	Ile	Val	Thr	Asp	Ser	850	855	860
Gln	Tyr	Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro	Asp	Gln	Ser	Glu	Ser	865	870	875
Glu	Leu	Val	Asn	Gln	Ile	Ile	Glu	Gln	Leu	Ile	Lys	Lys	Glu	Lys	Val	885	890	895
Tyr	Leu	Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	Gly	Gly	Asn	Glu	Gln	900	905	910
Val	Asp	Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys	Val	Leu	Phe	Leu	Asp	915	920	925
Met	Val	Gly	Phe	Pro	Val	Thr	Pro	Gln	Val	Pro	Leu	Arg	Pro	Met	Thr	930	935	940

Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly
945 950 955 960
Leu Glu Gly Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu
965 970 975
Trp Ile Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
980 985 990
Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys
995 1000 1005
Leu Val Pro Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu
1010 1015 1020
Asn Thr Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro
1025 1030 1035 1040
Glu Arg Glu Val Leu Glu Trp Arg Phe Asp Ser His Leu Ala Phe His
1045 1050 1055
His Val Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys
1060 1065 1070

<210> 68
<211> 3204
<212> DNA
<213> Artificial Sequence

<220>
<223> HIV

<400> 68
atgggtgccc gagcttcggt actgtctggt ggagagctgg acagatggga gaaaattagg 60
ctgcgcccgg gaggcaaaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgagggag 120
cttgaacggt ttgccgtgaa cccaggcctg ctggaaacat ctgagggatg tcgccagatc 180
ctggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggtc cttgtataac 240
acagtggcta ccctctactg cgtaaccag aggatcgaga ttaaggatac caaggaggcc 300
ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga aggcccagca ggcagctgct 360
gacactgggc atagcaacca ggtatcacag aactatccta ttgtccaaaa cattcagggc 420
cagatgggtc atcaggccat cagccccgg acgctcaatg cctgggtgaa ggttgtcgaa 480
gagaaggcct tttctctga ggttatcccc atgttctcgg ctttgagtga gggggccact 540
cctcaggacc tcaatacaat gcttaatacc gtgggcggcc atcaggccgc catgcaaatg 600
ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gactgcatcc cgtccacgtc 660
ggcccaatcg cgcccgga caatgacgag gatgcgggag cctcgcggct ctgacattgc cggcaccacc 720
tctacactgc aagagcaaat cggatggatg accaacaatc ctcccatccc agttggagaa 780
atctataaac ggtggatcat cctgggcctg aacaagatcg tgcgcatgta ctctccgaca 840
tccatccttg acattagaca gggacccaaa gagcctttta gggattacgt cgaccggttt 900
tataagacc tgcgagcaga gcaggcctct caggaggtca aaaactggat gacggagaca 960
ctcctggtac agaacgctaa ccccgactgc aaaacaatct tgaaggcact agggccggct 1020
gccaccctgg aagagatgat gaccgcctgt cagggagtag gcggacccgg acacaaagcc 1080
agagtgttga tgggccccat cagtcccatc gagaccgtgc cggatgaagct gaaacccggg 1140
atggacggcc ccaaggtcaa gcagtggcca ctaccggagg agaagatcaa ggccctggtg 1200
gagatctgca ccgagatgga gaaagagggc aagatcagca agatcggggc tgagaacca 1260
tacaacaccc ccgtgtttgc catcaagaag aaggacagca ccaagtggcg caagctggtg 1320
gatttccggg agctgaataa gcggacccag gatttctggg aggtccagct gggcatcccc 1380
catccggccg gcctgaagaa gaagaagagc gtgaccgtgc tggacgtggg cgacgcttac 1440
ttcagcgtcc ctctggacga ggactttaga aagtacaccg cctttaccat cccatctatc 1500
aacaacgaga cccctggcat cagatatcag tacaacgtcc tccccaggg ctggaagggc 1560
tctcccgcca tttctcagag ctccatgacc aagatcctgg agccgtttcg gaagcagaac 1620
cccgatatcg ttctctacca gtacatggac gacctgtacg tgggctctga cctggaaatc 1680
gggcagcatc gcacgaagat tgaggagctg aggcagcatc tgctgagatg gggcctgacc 1740
actccggaca agaagcatca gaaggagccg ccattcctgt ggatgggcta cgagctccat 1800
ccgcacaagt ggaccgtgca gcctatcgtc ctccccgaga aggacagctg gaccgtgaac 1860
gacatccaga agctggtggg caagctcaac tgggctagcc agatctatcc cgggatcaag 1920
gtgcgccagc tctgcaagct gctgcgcggc accaaggccc tgaccgaggt gattccccctc 1980
acggaggaag ccgagctcga gctggctgag aaccgggaga tcctgaagga gcccgtgcac 2040
ggcgtgtact atgacccctc caaggacctg atcgccgaaa tccagaagca gggccagggg 2100
cagtggacat accagattta ccaggagcct ttcaagaacc tcaagaccgg caagtacgcc 2160
cgcattgagg ggcgccacac caacgatgtc aagcagctga ccgaggccgt ccagaagatc 2220

acgaccgagt ccacgtgat ctgggggaag acacccaagt tcaagctgcc tatccagaag 2280
gagacctggg agacgtggtg gaccgaatat tggcaggcca cctggattcc cgagtgggag 2340
ttcgtgaata cacctcctct ggtgaagctg tggtagcagc tcgagaagga gcccatcgtg 2400
ggcgcgagga cattctacgt ggacggcgcg gccaaaccgc aaacaaagct cgggaaggcc 2460
gggtacgtca ccaaccgggg ccgccagaag gtcgtcaccg tgaccgacac caccaaccag 2520
aagacggagc tgcaggccat ctatctcgtc ctccaggact ccggcctgga ggtgaacatc 2580
gtgacggaca gccagtacgc gctgggcatt attcaggccc agccggacca gtccgagagc 2640
gaactggtga accagattat cgagcagctg atcaagaaag agaaggtcta cctcgcctgg 2700
gtcccggccc ataagggcat tggcggcaac gagcaggctg acaagctggt gagtgcgggg 2760
attagaaagg tgctgatggt gggttttcca gtcacacctc aggtaccttt aagaccaatg 2820
acttacaagg cagctgtaga tcttagccac tttttaaaag aaaagggggg actggaaggg 2880
ctaattcact cccaaagaag acaagatata cttgatctgt ggatctacca cacacaaggc 2940
tacttccctg attggcagaa ctacacacca gggccagggg tcagatatcc actgaccttt 3000
ggatggtgct acaagctagt accagttgag ccagataagg tagaagaggc caataaagga 3060
gagaacacca gcttgttaca ccctgtgagc ctgcatggga tggatgacct ggagagagaa 3120
gtgttagagt ggaggtttga cagccgccta gcatttcata acgtggcccc agagctgcat 3180
ccggagtact tcaagaactg ctga 3204

<210> 69

<211> 1067

<212> PRT

<213> Artificial Sequence

<220>

<223> HIV

<400> 69

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

290	295	300
Arg Ala Glu Gln Ala Ser	Gln Glu Val Lys Asn	Trp Met Thr Glu Thr
305	310	315
Leu Leu Val Gln Asn Ala	Asn Pro Asp Cys Lys	Thr Ile Leu Lys Ala
	325	330
Leu Gly Pro Ala Ala Thr	Leu Glu Glu Met Met	Thr Ala Cys Gln Gly
	340	345
Val Gly Gly Pro Gly His	Lys Ala Arg Val Leu	Met Gly Pro Ile Ser
	355	360
Pro Ile Glu Thr Val Ser	Val Lys Leu Lys Pro	Gly Met Asp Gly Pro
	370	375
Lys Val Lys Gln Trp Pro	Leu Thr Glu Glu Lys	Ile Lys Ala Leu Val
385	390	395
Glu Ile Cys Thr Glu Met	Glu Lys Glu Gly Lys	Ile Ser Lys Ile Gly
	405	410
Pro Glu Asn Pro Tyr Asn	Thr Pro Val Phe Ala	Ile Lys Lys Lys Asp
	420	425
Ser Thr Lys Trp Arg Lys	Leu Val Asp Phe Arg	Glu Leu Asn Lys Arg
	435	440
Thr Gln Asp Phe Trp Glu	Val Gln Leu Gly Ile	Pro His Pro Ala Gly
	450	455
Leu Lys Lys Lys Lys Ser	Val Thr Val Leu Asp	Val Gly Asp Ala Tyr
465	470	475
Phe Ser Val Pro Leu Asp	Glu Asp Phe Arg Lys	Tyr Thr Ala Phe Thr
	485	490
Ile Pro Ser Ile Asn Asn	Glu Thr Pro Gly Ile	Arg Tyr Gln Tyr Asn
	500	505
Val Leu Pro Gln Gly Trp	Lys Gly Ser Pro Ala	Ile Phe Gln Ser Ser
	515	520
Met Thr Lys Ile Leu Glu	Pro Phe Arg Lys Gln	Asn Pro Asp Ile Val
	530	535
Ile Tyr Gln Tyr Met Asp	Asp Leu Tyr Val Gly	Ser Asp Leu Glu Ile
545	550	555
Gly Gln His Arg Thr Lys	Ile Glu Glu Leu Arg	Gln His Leu Leu Arg
	565	570
Trp Gly Leu Thr Thr Pro	Asp Lys Lys His Gln	Lys Glu Pro Pro Phe
	580	585
Leu Trp Met Gly Tyr Glu	Leu His Pro Asp Lys	Trp Thr Val Gln Pro
	595	600
Ile Val Leu Pro Glu Lys	Asp Ser Trp Thr Val	Asn Asp Ile Gln Lys
	610	615
Leu Val Gly Lys Leu Asn	Trp Ala Ser Gln Ile	Tyr Pro Gly Ile Lys
625	630	635
Val Arg Gln Leu Cys Lys	Leu Leu Arg Gly Thr	Lys Ala Leu Thr Glu
	645	650
Val Ile Pro Leu Thr Glu	Glu Glu Ala Glu Leu	Glu Leu Ala Glu Asn Arg
	660	665
Glu Ile Leu Lys Glu Pro	Val His Gly Val Tyr	Tyr Tyr Asp Pro Ser Lys
	675	680
Asp Leu Ile Ala Glu Ile	Gln Lys Gln Gly Gln	Gly Gln Trp Thr Tyr
	690	695
Gln Ile Tyr Gln Glu Pro	Phe Lys Asn Leu Lys	Thr Gly Lys Tyr Ala
705	710	715
Arg Met Arg Gly Ala His	Thr Asn Asp Val Lys	Gln Leu Thr Glu Ala
	725	730
Val Gln Lys Ile Thr Thr	Glu Ser Ile Val Ile	Trp Gly Lys Thr Pro
	740	745
Lys Phe Lys Leu Pro Ile	Gln Lys Glu Thr Trp	Glu Thr Trp Trp Thr
	755	760
Glu Tyr Trp Gln Ala Thr	Trp Ile Pro Glu Trp	Glu Phe Val Asn Thr
	770	775
Pro Pro Leu Val Lys Leu	Trp Tyr Gln Leu Glu	Lys Glu Pro Ile Val
785	790	795
Gly Ala Glu Thr Phe Tyr	Val Asp Gly Ala Ala	Asn Arg Glu Thr Lys

				805					810					815			
Leu	Gly	Lys	Ala	Gly	Tyr	Val	Thr	Asn	Arg	Gly	Arg	Gln	Lys	Val	Val		
			820						825				830				
Thr	Leu	Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu	Leu	Gln	Ala	Ile	Tyr		
		835					840					845					
Leu	Ala	Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn	Ile	Val	Thr	Asp	Ser		
	850					855					860						
Gln	Tyr	Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro	Asp	Gln	Ser	Glu	Ser		
865				870					875					880			
Glu	Leu	Val	Asn	Gln	Ile	Ile	Glu	Gln	Leu	Ile	Lys	Lys	Glu	Lys	Val		
			885					890					895				
Tyr	Leu	Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	Gly	Gly	Asn	Glu	Gln		
	900							905				910					
Val	Asp	Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys	Val	Leu	Met	Val	Gly		
	915						920					925					
Phe	Pro	Val	Thr	Pro	Gln	Val	Pro	Leu	Arg	Pro	Met	Thr	Tyr	Lys	Ala		
	930				935						940						
Ala	Val	Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly	Leu	Glu	Gly		
945				950					955					960			
Leu	Ile	His	Ser	Gln	Arg	Arg	Gln	Asp	Ile	Leu	Asp	Leu	Trp	Ile	Tyr		
			965					970					975				
His	Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr	Pro	Gly	Pro		
	980						985					990					
Gly	Val	Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Tyr	Lys	Leu	Val	Pro		
	995					1000					1005						
Val	Glu	Pro	Asp	Lys	Val	Glu	Glu	Ala	Asn	Lys	Gly	Glu	Asn	Thr	Ser		
	1010					1015					1020						
Leu	Leu	His	Pro	Val	Ser	Leu	His	Gly	Met	Asp	Asp	Pro	Glu	Arg	Glu		
1025				1030					1035					1040			
Val	Leu	Glu	Trp	Arg	Phe	Asp	Ser	Arg	Leu	Ala	Phe	His	His	Val	Ala		
			1045					1050					1055				
Arg	Glu	Leu	His	Pro	Glu	Tyr	Phe	Lys	Asn	Cys							
		1060						1065									

<210> 70

<211> 1518

<212> DNA

<213> Artificial Sequence

<220>

<223> HIV

<400> 70

```

atgggtgccc gagcttcggt actgtctggt ggagagctgg acagatggga gaaaattagg 60
ctgcgcccgg gagggcaaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgagggag 120
cttgaacggt ttgccgtgaa cccaggcctg ctggaaacat ctgagggatg tgcgccagatc 180
ctggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggtc cttgtataac 240
acagtggcta ccctctactg cgtaaccag aggatcgaga ttaaggatac caaggaggcc 300
ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga aggcccagca ggcagctgct 360
gacactgggc atagcaacca ggtatcacag aactatccta ttgtccaaaa cattcagggc 420
cagatgggtc atcaggccat cagccccgg acgctcaatg cctgggtgaa ggttgctgaa 480
gagaaggcct tttctcctga ggttatcccc atgttctcgg ctttgagtga gggggccact 540
cctcaggacc tcaatacaat gcttaatacc gtgggcggcc atcaggccgc catgcaaatg 600
ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gagtgcattc cgtccacgct 660
ggcccaatcg cgcccgga ca gatgcgggag cctgcgggct ctgacattgc cggcaccacc 720
tctacactgc aagagcaaat cggtatggatg accaacaatc ctcccatccc agttggagaa 780
atctataaac ggtggatcat tctcgggtctc aataaaattg ttagaatgta ctctccgaca 840
tccatccttg acattagaca gggacccaaa gagcctttta gggattacgt cgaccggttt 900
tataagacc tgcgagcaga gcaggcctct caggaggtca aaaactggat gacggagaca 960
ctcctggtac agaacgctaa ccccgactgc aaaacaatct tgaaggcact agggccggct 1020
gccaccctgg aagagatgat gaccgcctgt cagggagtag gcggaccggg acacaaagcc 1080
agagtgttga tgggtgggtt tccagtcaca cctcaggtag ctttaagacc aatgacttac 1140
aaggcagctg tagatcttag ccacttttta aaagaaaagg ggggactgga agggctaatt 1200

```


cactcccaaa gaagacaaga tatccttgat ctgtggatct accacacaca aggctacttc 1260
 cctgattggc agaactacac accagggcca ggggtcagat atccactgac ctttggatgg 1320
 tgctacaagc tagtaccagt tgagccagat aaggtagaag aggccaataa aggagagaac 1380
 accagcttgt tacaccctgt gagcctgcat gggatggatg acccggagag agaagtgtta 1440
 gagtggaggt ttgacagccg cctagcattt catcacgtgg cccgagagct gcatccggag 1500
 tacttcaaga actgctga 1518

<210> 71
 <211> 505
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 71
 Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp
 1 5 10 15
 Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys
 20 25 30
 His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro
 35 40 45
 Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
 50 55 60
 Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn
 65 70 75 80
 Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp
 85 90 95
 Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
 100 105 110
 Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn Gln Val
 115 120 125
 Ser Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His
 130 135 140
 Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu
 145 150 155 160
 Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
 165 170 175
 Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly
 180 185 190
 Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu
 195 200 205
 Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala
 210 215 220
 Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 225 230 235 240
 Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile
 245 250 255
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270
 Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly
 275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
 290 295 300
 Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr
 305 310 315 320
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
 325 330 335
 Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
 340 345 350
 Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Met Val Gly Phe Pro
 355 360 365
 Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Ala Ala Val
 370 375 380

```

Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile
385                      390                      395                      400
His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu Trp Ile Tyr His Thr
                      405                      410                      415
Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Val
                      420                      425                      430
Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys Leu Val Pro Val Glu
                      435                      440                      445
Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu Asn Thr Ser Leu Leu
                      450                      455                      460
His Pro Val Ser Leu His Gly Met Asp Asp Pro Glu Arg Glu Val Leu
465                      470                      475                      480
Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His His Val Ala Arg Glu
                      485                      490                      495
Leu His Pro Glu Tyr Phe Lys Asn Cys
                      500                      505

```

<210> 72

<211> 1689

<212> DNA

<213> Artificial Sequence

<220>

<223> HIV

<400> 72

```

atggggcccca tcagtcccat cgagaccgtg ccggtgaagc tgaaaccccg gatggacggc 60
cccaaggtca agcagtggcc actcaccgag gagaagatca aggccctggt ggagatctgc 120
accgagatgg agaaagaggg caagatcagc aagatcgggc ctgagaaccc atacaacacc 180
cccggtgtttg ccatcaagaa gaaggacagc accaagtggc gcaagctggt ggatttccgg 240
gagctgaata agcggaccca ggattttctg gaggtccagc tgggcatccc ccatccggcc 300
ggcctgaaga agaagaagag cgtgaccgtg ctggacgtgg gcgacgctta cttcagcgtc 360
cctctggacg aggacttttag aaagtacacc gcctttacca tcccatctat caacaacgag 420
acccctggga tcagatatca gtacaacgtc ctcccccagg gctggaaggg ctctcccggc 480
attttccaga gctccatgac caagatcctg gagccgtttc ggaagcagaa ccccgatatc 540
gtcatctacc agtacatgga cgacctgtac gtgggctctg acctggaaat cgggcagcac 600
cgcacgaaga ttgaggagct gaggcagcat ctgctgagat ggggcctgac cactccggac 660
aagaagcatc agaaggagcc gccattcctg aagatgggct acgagctcca tcccgacaag 720
tggaaccgtg agcctatcgt cctccccgag aaggacagct ggaccgtgaa cgacatccag 780
aagctggtgg gcaagctcaa ctgggctagc cagatctatc ccgggatcaa ggtgcgccag 840
ctctgcaagc tgctgcgcgg caccaaggcc ctgaccgagg tgattcccct cacggaggaa 900
gccgagctcg agctggctga gaaccgggag atcctgaagg agcccgtgca cggcgtgtac 960
tatgaccctt ccaaggacct gatcgccgaa atccagaagc agggccaggg gcagtggaca 1020
taccagattt accaggagcc tttcaagaac ctcaagaccg gcaagtacgc ccgcatgagg 1080
ggcgcccaca ccaacgatgt caagcagctg accgaggccg tccagaagat cacgaccgag 1140
tccatcgtga tctgggggaa gacacccaag ttcaagctgc ctatccagaa ggagacctgg 1200
gagacgtggt ggaccgaata ttggcaggcc acctggattc ccgagtggga gttcgtgaat 1260
acacctcttc tgggtgaagct gtggtaccag ctcgagaagg agcccacgtt gggcgccggag 1320
acattctacg tggacggcgc ggccaaccgc gaaacaaagc tcgggaaggc cgggtacgtc 1380
accaaccggg gccgccagaa ggtcgtcacc ctgaccgaca ccaccaacca gaagacggag 1440
ctgcaggcca tctatctcgc tctccaggac tccggcctgg aggtgaacat cgtgacggac 1500
agccagtacg cgctgggcat tattcaggcc cagccggacc agtccgagag cgaactgggtg 1560
aaccagatta tcgagcagct gatcaagaaa gagaaggctc acctcgctg ggtcccggcc 1620
cataagggca ttggcggcaa cgagcaggtc gacaagctgg tgagtgcggg gattagaaag 1680
gtgctgtaa
1689

```

<210> 73

<211> 3204

<212> DNA

<213> Artificial Sequence

<220>

<223> HIV

<400> 73

```

atgggtgccc gagcttcggt actgtctggt ggagagctgg acagatggga gaaaattagg 60
ctgcgcccgg gaggcacaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgaggagg 120
cttgaacggt ttgcctgtaa cccaggcctg ctggaaacat ctgagggatg tcgccagatc 180
ctggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggtc cttgtataac 240
acagtggcta ccctctactg cgtacaccag aggatcgaga ttaaggatac caaggaggcc 300
ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga aggcccagca ggcagctgct 360
gacactgggc atagcaacca ggtatcacag aactatccta ttgtccaaaa cattcagggc 420
cagatggttc atcaggccat cagcccccg acgctcaatg cctgggtgaa ggttgtcgaa 480
gagaaggcct tttctcctga ggttatcccc atgttctccg ctttgagtga gggggccact 540
cctcaggacc tcaatacaat gcttaatacc gtgggcggcc atcaggccgc catgcaaagt 600
ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gagtgcattc cgtccacgct 660
ggcccaatcg cgcccggaca gatgcgggag cctcgcggct ctgacattgc cggcaccacc 720
tctacactgc aagagcaaat cggatggatg accaacaatc ctcccatccc agttggagaa 780
atctataaac ggtggatcat cctgggcctg aacaagatcg tgcgcatgta ctctccgaca 840
tccatccttg acattagaca gggacccaaa gagcctttta gggattacgt cgaccggttt 900
tataagacc tgcgagcaga gcaggcctct caggagggtca aaaactggat gacggagaca 960
tctcctggta agaacgctaa ccccgactgc aaaacaatct tgaaggcact aggcccggtc 1020
gccaccctgg aagagatgat gaccgcctgt cagggagtag gcggaccggg acacaaagcc 1080
agagtgttga tgggccccat cagtcccatc gagaccgtgc cgggtgaagct gaaaccggg 1140
atggacggcc ccaaggtcaa gcagtggcca ctaccggagg agaagatcaa ggccctggtg 1200
gagatctgca ccgagatgga gaaagagggc aagatcagca agatcggggc tgagaaccca 1260
tacaacaccc ccgtgtttgc catcaagaag aaggacagca ccaagtggcg caagctggtg 1320
gatttccggg agctgaataa gcggaccagc gatttctggg aggtccagct gggcatcccc 1380
catccggccg gcctgaagaa gaagaagagc gtgaccgtgc tggacgtggg cgacgcttac 1440
ttcagcgctc ctctggacga ggactttaga aagtacaccg cctttaccat cccatctatc 1500
aacaacgaga cccctggcat cagatatcag tacaacgtcc tccccaggg ctggaagggc 1560
tctcccgcca ttttccagag ctccatgacc aagatcctgg agccgtttcg gaagcagaac 1620
cccgatatcg tcatctacca gtacatggac gacctgtacg tgggctctga cctggaaatc 1680
gggcagcatc gcacgaagat tgaggagctg aggcagcatc tgctgagatg gggcctgacc 1740
actccggaca agaagcatca gaaggagccg ccattcctga agatgggcta cgagctccat 1800
cccgacaagt ggaccgtgca gcctatcgtc ctccccgaga aggacagctg gaccgtgaac 1860
gacatccaga agctggtggg caagctcaac tgggctagcc agatctatcc cgggatcaag 1920
gtgcgccagc tctgcaagct gctgcgcggc accaaggccc tgaccgaggt gattccccctc 1980
acggaggaag ccgagctgca gctggctgag aaccgggaga tctgaagga gcccggtgac 2040
ggcgtgtact atgacccctc caaggacctg atgcgcgaaa tccagaagca gggccagggg 2100
cagtggacat accagattta ccaggagcct ttcaagaacc tcaagaccgg caagtacgcc 2160
cgcattgagg ggcgccacac caacgatgtc aagcagctga ccgaggccgt ccagaagatc 2220
acgaccgagt ccattcgtgat ctgggggaag acacccaagt tcaagctgcc tatccagaag 2280
gagacctggg agacgtggtg gaccgaatat tggcaggcca cctggattcc cgagtgggag 2340
ttcgtgaata cacctcctct ggtgaagctg tggtagcagc tcgagaagga gcccattcgtg 2400
ggcgccggaga cattctacgt ggacggcgcg gccaaaccgc aaacaaagct cgggaaggcc 2460
gggtacgtca ccaaccgggg ccgccagaag gtcgtcacc tgaccgacac caccaaccag 2520
aagacggagc tgcaggccat ctatctcgct ctccaggact ccggcctgga ggtgaacatc 2580
gtgacggaca gccagtacgc gctgggcatt attcaggccc agccggacca gtccgagagc 2640
gaactggtga accagattat cgagcagctg atcaagaaa agaaaggtcta cctcgcctgg 2700
gtcccgggcc ataagggcatt tggcggaac gagcaggtcg acaagctggt gagtgcgggg 2760
attagaaagg tgctgatggt ggggttttcca gtcacacctc aggtaccttt aagaccaatg 2820
acttacaagg cagctgtaga tcttagccac tttttaaaa agaaagggggg actggaaggg 2880
ctaattcact cccaaagaag acaagatatc cttgatctgt ggatctacca cacacaaggc 2940
tacttccctg attggcagaa ctacacacca gggccagggg tcagatatcc actgaccttt 3000
ggatggtgct acaagctagt accagttgag ccagataagg tagaagaggc caataaagga 3060
gagaacacca gcttgttaca ccctgtgagc ctgcatggga tggatgacct ggagagagaa 3120
gtgttagagt ggaggtttga cagccgccta gcatttcac acgtggcccc agagctgcat 3180
ccggagtact tcaagaactg ctga

```

<210> 74

<211> 1067

<212> PRT

<213> Artificial Sequence

<220>

<223> HIV

<400> 74

```

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

```

Ile	Pro	Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	
			500					505					510			
Val	Leu	Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Ser	Ser	
		515					520					525				
Met	Thr	Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro	Asp	Ile	Val	
	530					535					540					
Ile	Tyr	Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	
545					550					555					560	
Gly	Gln	His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Gln	His	Leu	Leu	Arg	
			565						570					575		
Trp	Gly	Leu	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe	
		580						585					590			
Leu	Trp	Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	
	595						600					605				
Ile	Val	Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	
610						615					620					
Leu	Val	Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro	Gly	Ile	Lys	
625					630					635					640	
Val	Arg	Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala	Leu	Thr	Glu	
			645						650					655		
Val	Ile	Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	
		660						665					670			
Glu	Ile	Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	
	675						680					685				
Asp	Leu	Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr	
690						695					700					
Gln	Ile	Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	
705					710					715					720	
Arg	Met	Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala	
			725					730						735		
Val	Gln	Lys	Ile	Thr	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro	
		740						745					750			
Lys	Phe	Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr	Trp	Trp	Thr	
	755						760					765				
Glu	Tyr	Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe	Val	Asn	Thr	
770						775					780					
Pro	Pro	Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu	Lys	Glu	Pro	Ile	Val	
785					790					795					800	
Gly	Ala	Glu	Thr	Phe	Tyr	Val	Asp	Gly	Ala	Ala	Asn	Arg	Glu	Thr	Lys	
			805						810					815		
Leu	Gly	Lys	Ala	Gly	Tyr	Val	Thr	Asn	Arg	Gly	Arg	Gln	Lys	Val	Val	
		820						825					830			
Thr	Leu	Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu	Leu	Gln	Ala	Ile	Tyr	
	835						840					845				
Leu	Ala	Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn	Ile	Val	Thr	Asp	Ser	
850						855					860					
Gln	Tyr	Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro	Asp	Gln	Ser	Glu	Ser	
865					870					875					880	
Glu	Leu	Val	Asn	Gln	Ile	Ile	Glu	Gln	Leu	Ile	Lys	Lys	Glu	Lys	Val	
			885						890					895		
Tyr	Leu	Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	Gly	Gly	Asn	Glu	Gln	
		900						905					910			
Val	Asp	Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys	Val	Leu	Met	Val	Gly	
	915						920					925				
Phe	Pro	Val	Thr	Pro	Gln	Val	Pro	Leu	Arg	Pro	Met	Thr	Tyr	Lys	Ala	
	930					935					940					
Ala	Val	Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly	Leu	Glu	Gly	
945					950					955					960	
Leu	Ile	His	Ser	Gln	Arg	Arg	Gln	Asp	Ile	Leu	Asp	Leu	Trp	Ile	Tyr	
			965						970					975		
His	Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr	Pro	Gly	Pro	
		980						985					990			
Gly	Val	Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Tyr	Lys	Leu	Val	Pro	
		995					1000						1005			

Val	Glu	Pro	Asp	Lys	Val	Glu	Glu	Ala	Asn	Lys	Gly	Glu	Asn	Thr	Ser
1010					1015						1020				
Leu	Leu	His	Pro	Val	Ser	Leu	His	Gly	Met	Asp	Asp	Pro	Glu	Arg	Glu
1025					1030					1035					1040
Val	Leu	Glu	Trp	Arg	Phe	Asp	Ser	Arg	Leu	Ala	Phe	His	His	Val	Ala
					1045					1050					1055
Arg	Glu	Leu	His	Pro	Glu	Tyr	Phe	Lys	Asn	Cys					
					1060					1065					

<210> 75
 <211> 3204
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 75

atgggtgggtt	ttccagtcac	acctcaggta	cctttaagac	caatgactta	caaggcagct	60
gtagatctta	gccacttttt	aaaagaaaag	gggggactgg	aagggtctaat	tcactcccaa	120
agaagacaag	atataccttga	tctgtggatc	taccacacac	aaggctactt	ccctgatttg	180
cagaactaca	caccagggcc	aggggtcaga	tatccactga	cctttggatg	gtgctacaag	240
ctagtaccag	ttgagccaga	taaggtagaa	gaggccaata	aaggagagaa	caccagcttg	300
ttacaccctg	tgagcctgca	tgggatggat	gacccgagga	gagaagtgtt	agagtggagg	360
tttgacagcc	gcctagcatt	tcatacagtg	gcccagagagc	tgcataccgga	gtacttcaag	420
aactgcatgg	gccccatcag	tcccatcgag	accgtgccgg	tgaagctgaa	accggggatg	480
gacggcccca	aggtcaagca	gtggccactc	accgaggaga	agatcaaggc	cctggtggag	540
atctgcaccg	agatggagaa	agagggcaag	atcagcaaga	tcgggcctga	gaaccctac	600
aacacccccg	tgtttgccat	caagaagaag	gacagcacca	agtggcgcaa	gctggtggat	660
ttccggggagc	tgaataagcg	gacccaggat	ttctgggagg	tccagctggg	catcccccat	720
ccggccgggcc	tgaagaagaa	gaagagcggtg	accgtgctgg	acgtgggcga	cgcttacttc	780
agcgtccctc	tggacgagga	ctttagaaaag	tacaccgcct	ttaccatccc	atctatcaac	840
aacgagaccc	ctggcatcag	atatcagtag	aacgtcctcc	cccagggctg	gaagggctct	900
cccgccattt	tccagagctc	catgaccaag	atcctggagc	cgtttcggaa	gcagaacccc	960
gatatcgtca	tctaccagta	catggacgac	ctgtacgtgg	gctctgacct	ggaaatcggg	1020
cagcatcgca	cgaagattga	ggagctgagg	cagcatctgc	tgagatgggg	cctgaccact	1080
ccggacaaga	agcatcagaa	ggagccgcca	ttcctgaaga	tgggctacga	gctccatccc	1140
gacaagtggg	ccgtgcagcc	tatcgtcctc	cccagagaagg	acagctggac	cgtgaacgac	1200
atccagaagc	tgggtgggcaa	gctcaactgg	gctagccaga	tctatcccgg	gatcaagggtg	1260
cgccagctct	gcaagctgct	gcgcggcacc	aaggccctga	ccgaggtgat	tcccctcacg	1320
gaggaagccg	agctcgagct	ggctgagaac	cgggagatcc	tgaaggagcc	cgtgcacggc	1380
gtgtactatg	acccctccaa	ggacctgac	gccgaaatcc	agaagcaggg	ccaggggagc	1440
tggacatacc	agatttacca	ggagcctttc	aagaacctca	agaccggcaa	gtacgcccgc	1500
atgaggggag	cccacaccaa	cgatgtcaag	cagctgaccg	aggccgtcca	gaagatcacg	1560
accgagtcca	tctgtgatctg	ggggaagaca	cccaagttca	agctgcctat	ccagaaggag	1620
acctgggaga	cgtggtggac	cgaatattgg	caggccacct	ggattcccga	gtgggagttc	1680
gtgaatacac	ctcctctggt	gaagctgtgg	taccagctcg	agaaggagcc	catcgtgggc	1740
gcggagacat	tctacgtgga	cggcgcgggc	aaccgcgaaa	caaagctcgg	gaaggccggg	1800
tacgtcacca	accggggccg	ccagaagggtc	gtcacccctga	ccgacaccac	caaccagaag	1860
acggagctgc	aggccatcta	tctcgtctct	caggactccg	gcctggaggt	gaacatcgtg	1920
acggacagcc	agtacgcgct	gggcattatt	caggcccagc	cggaccagtc	cgagagcgaa	1980
ctggtgaacc	agattatcga	gcagctgac	aagaaagaga	aggtctacct	cgctgggtc	2040
ccggcccata	agggcatttg	cggcaacgag	caggtcgaca	agctggtgag	tgcggggatt	2100
agaaaggtgc	tgatgggtgc	ccgagcttcg	gtactgtctg	gtggagagct	ggacagatgg	2160
gagaaaatta	ggctgcgccc	gggaggcaaa	aagaaataca	agctcaagca	tatcgtgtgg	2220
gcctcgaggg	agcttgaacg	gtttgccgtg	aaccagggcc	tgctggaaac	atctgaggga	2280
tgctcgccaga	tcctggggca	attgcagcca	tccctccaga	ccgggagtg	agagctgagg	2340
tccttggtata	acacagtggc	taccctctac	tgctgacacc	agaggatcga	gattaaggat	2400
accaaggagg	ccttgagcaa	aattgaggag	gagcaaaaca	agagcaagaa	gaaggccag	2460
caggcagctg	ctgacactgg	gcatacgcaac	caggtatcac	agaactatcc	tattgtccaa	2520
aacattcagg	gccagatggg	tcatacaggcc	atcagcccc	ggacgctcaa	tgctgggtg	2580
aagggtgtcg	aagagaaggc	cttttctcct	gaggttatcc	ccatgttctc	cgctttgagt	2640
gagggggcca	ctcctcagga	cctcaataca	atgcttaata	ccgtgggcgg	ccatcaggcc	2700

gccatgcaaa tgttgaagga gactatcaac gaggaggcag ccgagtggga cagagtgcac 2760
 cccgtccacg ctggcccaat cgcgcccga cagatgcggg agcctcgcgg ctctgacatt 2820
 gccggcacca cctctacact gcaagagcaa atcggatgga tgaccaacaa tcctcccatc 2880
 ccagttggag aaatctataa acggtggatc atcctggggc tgaacaagat cgtgcgcatg 2940
 tactctccga catccatcct tgacattaga cagggaccca aagagccttt tagggattac 3000
 gtcgaccggt tttataagac cctgcgagca gagcaggcct ctcaggaggt caaaaactgg 3060
 atgacggaga cactcctggt acagaacgct aacccccgact gcaaaaacaat cttgaaggca 3120
 ctaggcccgg ctgccaccct ggaagagatg atgaccgcct gtcagggagt aggcggaccc 3180
 ggacacaaag ccagagtgtt gtga 3204

<210> 76
 <211> 1067
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 76
 Met Val Gly Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr
 1 5 10 15
 Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly
 20 25 30
 Leu Glu Gly Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu
 35 40 45
 Trp Ile Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
 50 55 60
 Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys
 65 70 75 80
 Leu Val Pro Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu
 85 90 95
 Asn Thr Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro
 100 105 110
 Glu Arg Glu Val Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His
 115 120 125
 His Val Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys Met Gly
 130 135 140
 Pro Ile Ser Pro Ile Glu Thr Val Ser Val Lys Leu Lys Pro Gly Met
 145 150 155 160
 Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys
 165 170 175
 Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser
 180 185 190
 Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys
 195 200 205
 Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu
 210 215 220
 Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His
 225 230 235 240
 Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly
 245 250 255
 Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr
 260 265 270
 Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr
 275 280 285
 Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe
 290 295 300
 Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro
 305 310 315 320
 Asp Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp
 325 330 335
 Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His
 340 345 350
 Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu

		355					360					365			
Pro	Pro	Phe	Leu	Trp	Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr
	370					375					380				
Val	Gln	Pro	Ile	Val	Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp
385					390					395					400
Ile	Gln	Lys	Leu	Val	Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro
				405					410					415	
Gly	Ile	Lys	Val	Arg	Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala
			420					425					430		
Leu	Thr	Glu	Val	Ile	Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala
			435					440				445			
Glu	Asn	Arg	Glu	Ile	Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp
	450					455					460				
Pro	Ser	Lys	Asp	Leu	Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln
465					470					475					480
Trp	Thr	Tyr	Gln	Ile	Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly
				485					490						495
Lys	Tyr	Ala	Arg	Met	Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu
			500					505					510		
Thr	Glu	Ala	Val	Gln	Lys	Ile	Thr	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly
			515				520					525			
Lys	Thr	Pro	Lys	Phe	Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr
	530					535					540				
Trp	Trp	Thr	Glu	Tyr	Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe
545					550					555					560
Val	Asn	Thr	Pro	Pro	Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu	Lys	Glu
				565					570					575	
Pro	Ile	Val	Gly	Ala	Glu	Thr	Phe	Tyr	Val	Asp	Gly	Ala	Ala	Asn	Arg
			580					585					590		
Glu	Thr	Lys	Leu	Gly	Lys	Ala	Gly	Tyr	Val	Thr	Asn	Arg	Gly	Arg	Gln
		595				600						605			
Lys	Val	Val	Thr	Leu	Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu	Leu	Gln
	610					615					620				
Ala	Ile	Tyr	Leu	Ala	Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn	Ile	Val
625					630					635					640
Thr	Asp	Ser	Gln	Tyr	Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro	Asp	Gln
				645					650					655	
Ser	Glu	Ser	Glu	Leu	Val	Asn	Gln	Ile	Ile	Glu	Gln	Leu	Ile	Lys	Lys
			660					665					670		
Glu	Lys	Val	Tyr	Leu	Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	Gly	Gly
			675				680					685			
Asn	Glu	Gln	Val	Asp	Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys	Val	Leu
	690					695					700				
Met	Gly	Ala	Arg	Ala	Ser	Val	Leu	Ser	Gly	Gly	Glu	Leu	Asp	Arg	Trp
705					710					715					720
Glu	Lys	Ile	Arg	Leu	Arg	Pro	Gly	Gly	Lys	Lys	Lys	Tyr	Lys	Leu	Lys
				725					730					735	
His	Ile	Val	Trp	Ala	Ser	Arg	Glu	Leu	Glu	Arg	Phe	Ala	Val	Asn	Pro
			740					745					750		
Gly	Leu	Leu	Glu	Thr	Ser	Glu	Gly	Cys	Arg	Gln	Ile	Leu	Gly	Gln	Leu
		755					760					765			
Gln	Pro	Ser	Leu	Gln	Thr	Gly	Ser	Glu	Glu	Leu	Arg	Ser	Leu	Tyr	Asn
	770					775					780				
Thr	Val	Ala	Thr	Leu	Tyr	Cys	Val	His	Gln	Arg	Ile	Glu	Ile	Lys	Asp
785					790					795					800
Thr	Lys	Glu	Ala	Leu	Asp	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Lys
				805					810					815	
Lys	Lys	Ala	Gln	Gln	Ala	Ala	Ala	Asp	Thr	Gly	His	Ser	Asn	Gln	Val
			820					825					830		
Ser	Gln	Asn	Tyr	Pro	Ile	Val	Gln	Asn	Ile	Gln	Gly	Gln	Met	Val	His
		835						840				845			
Gln	Ala	Ile	Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val	Val	Glu
	850					855					860				
Glu	Lys	Ala	Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Ser	Ala	Leu	Ser

865		870		875		880
Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly						
	885		890		895	
Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu						
	900		905		910	
Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala						
	915		920		925	
Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr						
	930		935		940	
Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile						
	945		950		955	
Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys						
	965		970		975	
Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly						
	980		985		990	
Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu						
	995		1000		1005	
Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr						
	1010		1015		1020	
Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala						
	1025		1030		1035	
Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly						
	1045		1050		1055	
Val Gly Gly Pro Gly His Lys Ala Arg Val Leu						
	1060		1065			

<210> 77
 <211> 3204
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 77
 atgggtggggtt ttccagtcac acctcaggta cctttaagac caatgactta caaggcagct 60
 gtagatcttta gccactttttt aaaagaaaag gggggactgg aagggtctaat tcactcccaa 120
 agaagacaag atataccttga tctgtggatc taccacacac aagggtactt ccctgattgg 180
 cagaactaca caccagggcc aggggtcaga tatccactga cctttggatg gtgctacaag 240
 ctagtaccag ttgagccaga taaggtagaa gaggccaaata aaggagagaa caccagcttg 300
 ttacaccctg tgagcctgca tgggatggat gacccgagaga gagaagtgtt agagtggagg 360
 tttgacagcc gcctagcatt tcatcacgtg gcccgagagc tgcattccgga gtacttcaag 420
 aactgcatgg gtgcccagac ttcggtactg tctgggtggag agctggacag atggggagaaa 480
 attaggctgc gcccgggagg caaaaagaaa tacaagctca agcatatcgt gtgggcctcg 540
 agggagcttg aacggtttgc cgtgaaccca ggcctgctgg aaacatctga gggatgtcgc 600
 cagatccttg ggcaattgca gccatccctc cagaccggga gtgaagagct gaggtccttg 660
 tataacacag tggctaccct ctactgcgta caccagagga tgcagattaa ggataccaag 720
 gaggccttgg acaaaattga ggaggagcaa aacaagagca agaagaaggc ccagcaggca 780
 gctgctgaca ctgggcatag caaccaggta tcacagaact atcctattgt ccaaaacatt 840
 cagggccaga tggttcatca ggccatcagc ccccgagacg tcaatgcctg ggtgaagggt 900
 gtcgaagaga aggccttttc tctgagggtt atccccatgt tctccgcttt gagtgaaggg 960
 gccactcctc aggacctcaa tacaatgctt aataccgtgg gcggccatca ggccgcatg 1020
 caaatgttga aggagactat caacgaggag gcagccgagt gggacagagt gcatcccgtc 1080
 cacgctggcc caatcgcgcc cggacagatg cgggagcctc gcggctctga cattgcccgg 1140
 accacctcta cactgcaaga gcaaatcgga tggatgacca acaatcctcc catcccagtt 1200
 ggagaaatct ataaacggtg gatcatcctg ggcctgaaca agatcgtgcg catgtactct 1260
 ccgacatcca tccttgacat tagacaggga cccaaagagc ctttttaggga ttacgtcgac 1320
 cggttttata agaccctgcg agcagagcag gcctctcagg aggtcaaaaa ctggatgacg 1380
 gagacactcc tggtagagaa cgctaacccc gactgcaaaa caatcttgaa ggactaggc 1440
 ccggctgcca ccttgggaaga gatgatgacc gcctgtcagg gagtaggcgg acccgacac 1500
 aaagccagag tgttgatggg ccccatcagt cccatcgaga ccgtgccggt gaagctgaaa 1560
 cccgggatgg acggcccaa ggtcaagcag tggccactca ccgaggagaa gatcaaggcc 1620
 ctgggtggaga tctgcaccga gatggagaaa gagggcaaga tcagcaagat cgggcctgag 1680

```

aacccatata acacccccgt gtttgccatc aagaagaagg acagcaccaa gtggcgcaag 1740
ctgggtggatt tccgggagct gaataagcgg acccaggatt tctgggaggt ccagctgggc 1800
atcccccatc cggccggcct gaagaagaag aagagcgtga ccgtgctgga cgtgggcgac 1860
gcttacttca gcgtccctct ggacgaggac tttagaaagt acaccgcctt taccatccca 1920
tctatcaaca acgagacccc tggcatcaga tatcagtaca acgtcctccc ccagggctgg 1980
aagggctctc ccgccathtt ccagagctcc atgaccaaga tcctggagcc gtttcggaag 2040
cagaaccccc atatacgtcat ctaccagtac atggacgacc tgtacgtggg ctctgacctg 2100
gaaatcgggc agcatcgcac gaagattgag gagctgaggc agcatctgct gagatggggc 2160
ctgaccactc cggacaagaa gcatcagaag gagccgccat tcctgaagat gggctacgag 2220
ctccatcccc acaagtggac cgtgcagcct atcgtcctcc ccgagaagga cagctggacc 2280
gtgaacgaca tccagaagct ggtgggcaag ctcaactggg ctagccagat ctatcccggg 2340
atcaaggtgc gccagctctg caagctgctg cgcggcacca aggccctgac cgaggtgatt 2400
cccctcacgg aggaagccga gctcgagctg gctgagaacc gggagatcct gaaggagccc 2460
gtgcacggcg tgtactatga cccctccaag gacctgatcg ccgaaatcca gaagcagggc 2520
caggggcagt ggacatacca gatttaccag gagcctttca agaacctcaa gaccggcaag 2580
tacgcccgcg tgagggggcg ccacaccaac gatgtcaagc agctgaccga ggccgtccag 2640
aagatcacga ccgagtcctat cgtgatctgg gggaagacac ccaagttcaa gctgcctatc 2700
cagaaggaga cctgggagac gtggtggacc gaatattggc aggccacctg gattcccag 2760
tgggagttcg tgaatacacc tcctctggtg aagctgtggt accagctcga gaaggagccc 2820
atcgtggggc cggagacatt ctacgtggac ggcgcggcca accgcgaaac aaagctcggg 2880
aaggccgggt acgtcaccaa ccggggcgcg cagaaggctg tcaccctgac cgacaccacc 2940
aaccagaaga cggagctgca ggccatctat ctcgctctcc aggactccgg cctggaggtg 3000
aacatcgtga cggacagcca gtacgcgctg ggcattattc aggccagcc ggaccagtcc 3060
gagagcgaac tgggtgaacca gattatcgag cagctgatca agaaagagaa ggtctacctc 3120
gcctgggtcc cggcccataa gggcattggc ggcaacgagc aggtcgacaa gctggtgagt 3180
gcggggatta gaaaggtgct gtaa 3204

```

<210> 78
 <211> 1067
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 78

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

Ala	Thr	Leu	Tyr	Cys	Val	His	Gln	Arg	Ile	Glu	Ile	Lys	Asp	Thr	Lys	225	230	235	240
Glu	Ala	Leu	Asp	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Lys	Lys	Lys	245	250	255	
Ala	Gln	Gln	Ala	Ala	Ala	Asp	Thr	Gly	His	Ser	Asn	Gln	Val	Ser	Gln	260	265	270	
Asn	Tyr	Pro	Ile	Val	Gln	Asn	Ile	Gln	Gly	Gln	Met	Val	His	Gln	Ala	275	280	285	
Ile	Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val	Val	Glu	Glu	Lys	290	295	300	
Ala	Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Ser	Ala	Leu	Ser	Glu	Gly	305	310	315	320
Ala	Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly	Gly	His	325	330	335	
Gln	Ala	Ala	Met	Gln	Met	Leu	Lys	Glu	Thr	Ile	Asn	Glu	Glu	Ala	Ala	340	345	350	
Glu	Trp	Asp	Arg	Val	His	Pro	Val	His	Ala	Gly	Pro	Ile	Ala	Pro	Gly	355	360	365	
Gln	Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Thr	Thr	Ser	Thr	370	375	380	
Leu	Gln	Glu	Gln	Ile	Gly	Trp	Met	Thr	Asn	Asn	Pro	Pro	Ile	Pro	Val	385	390	395	400
Gly	Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys	Ile	Val	405	410	415	
Arg	Met	Tyr	Ser	Pro	Thr	Ser	Ile	Leu	Asp	Ile	Arg	Gln	Gly	Pro	Lys	420	425	430	
Glu	Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Tyr	Lys	Thr	Leu	Arg	Ala	435	440	445	
Glu	Gln	Ala	Ser	Gln	Glu	Val	Lys	Asn	Trp	Met	Thr	Glu	Thr	Leu	Leu	450	455	460	
Val	Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu	Lys	Ala	Leu	Gly	465	470	475	480
Pro	Ala	Ala	Thr	Leu	Glu	Glu	Met	Met	Thr	Ala	Cys	Gln	Gly	Val	Gly	485	490	495	
Gly	Pro	Gly	His	Lys	Ala	Arg	Val	Leu	Met	Gly	Pro	Ile	Ser	Pro	Ile	500	505	510	
Glu	Thr	Val	Ser	Val	Lys	Leu	Lys	Pro	Gly	Met	Asp	Gly	Pro	Lys	Val	515	520	525	
Lys	Gln	Trp	Pro	Leu	Thr	Glu	Glu	Lys	Ile	Lys	Ala	Leu	Val	Glu	Ile	530	535	540	
Cys	Thr	Glu	Met	Glu	Lys	Glu	Gly	Lys	Ile	Ser	Lys	Ile	Gly	Pro	Glu	545	550	555	560
Asn	Pro	Tyr	Asn	Thr	Pro	Val	Phe	Ala	Ile	Lys	Lys	Lys	Asp	Ser	Thr	565	570	575	
Lys	Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg	Thr	Gln	580	585	590	
Asp	Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	Leu	Lys	595	600	605	
Lys	Lys	Lys	Ser	Val	Thr	Val	Leu	Asp	Val	Gly	Asp	Ala	Tyr	Phe	Ser	610	615	620	
Val	Pro	Leu	Asp	Glu	Asp	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	Ile	Pro	625	630	635	640
Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	Val	Leu	645	650	655	
Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Ser	Ser	Met	Thr	660	665	670	
Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro	Asp	Ile	Val	Ile	Tyr	675	680	685	
Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	Gly	Gln	690	695	700	
His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Gln	His	Leu	Leu	Arg	Trp	Gly	705	710	715	720
Leu	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe	Leu	Trp	725	730	735	

Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	Ile	Val
			740					745					750		
Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	Leu	Val
		755					760					765			
Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro	Gly	Ile	Lys	Val	Arg
		770				775					780				
Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala	Leu	Thr	Glu	Val	Ile
785					790					795					800
Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	Glu	Ile
				805				810						815	
Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	Asp	Leu
			820					825					830		
Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr	Gln	Ile
		835					840					845			
Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	Arg	Met
		850				855					860				
Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala	Val	Gln
865					870					875					880
Lys	Ile	Thr	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro	Lys	Phe
			885						890					895	
Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr	Trp	Trp	Thr	Glu	Tyr
			900					905					910		
Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe	Val	Asn	Thr	Pro	Pro
		915					920					925			
Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu	Lys	Glu	Pro	Ile	Val	Gly	Ala
		930				935					940				
Glu	Thr	Phe	Tyr	Val	Asp	Gly	Ala	Ala	Asn	Arg	Glu	Thr	Lys	Leu	Gly
945					950					955					960
Lys	Ala	Gly	Tyr	Val	Thr	Asn	Arg	Gly	Arg	Gln	Lys	Val	Val	Thr	Leu
			965					970						975	
Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu	Leu	Gln	Ala	Ile	Tyr	Leu	Ala
			980					985					990		
Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn	Ile	Val	Thr	Asp	Ser	Gln	Tyr
		995					1000					1005			
Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro	Asp	Gln	Ser	Glu	Ser	Glu	Leu
		1010				1015					1020				
Val	Asn	Gln	Ile	Ile	Glu	Gln	Leu	Ile	Lys	Lys	Glu	Lys	Val	Tyr	Leu
1025					1030					1035					1040
Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	Gly	Gly	Asn	Glu	Gln	Val	Asp
			1045					1050						1055	
Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys	Val	Leu					
			1060					1065							

<210> 79
 <211> 3204
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 79
 atggggcccca tcagtcccat cgagaccgtg ccggtgaagc tgaaaccggt gatggacggc 60
 cccaaggtca agcagtggcc actcaccgag gagaagatca aggcctgtgt ggagatctgc 120
 accgagatgg agaaagaggg caagatcagc aagatcgggc ctgagaaccc atacaacacc 180
 cccgtgtttg ccatcaagaa gaaggacagc accaagtggc gcaagctggt ggatttccgg 240
 gagctgaata agcggaccca ggatttcttg gaggtccagc tgggcatccc ccatccggcc 300
 ggctgaaga agaagaagag cgtgaccgtg ctggacgtgg gcgacgctta cttcagcgtc 360
 cctctggacg aggacttttag aaagtacacc gcctttacca tcccatctat caacaacgag 420
 acccctggca tcagatatca gtacaacgtc ctccccagg gctggaaggg ctctcccgcc 480
 attttccaga gctccatgac caagatcctg gagccgtttc ggaagcagaa ccccgatata 540
 gtcacttacc agtacatgga cgacctgtac gtgggctctg acctggaaat cgggcagcat 600
 cgcacgaaga ttgaggagct gaggcagcat ctgctgagat ggggcctgac cactccggac 660

aagaagcatc agaaggagcc gccattcctg aagatgggct acgagctcca tcccgacaag 720
tggaccgtgc agcctatcgt cctccccgag aaggacagct ggaccgtgaa cgacatccag 780
aagctggtgg gcaagctcaa ctgggctagc cagatctatc ccgggatcaa ggtgcgccag 840
ctctgcaagc tgctgcgcgg caccaaggcc ctgaccgagg tgattcccct cacggaggaa 900
gccgagctcg agctggctga gaaccgggag atcctgaagg agcccgtgca cggcgtgtac 960
tatgacctct ccaaggacct gatcgccgaa atccagaagc agggccaggg gcagtggaca 1020
taccagattt accaggagcc tttcaagaac ctcaagaccg gcaagtacgc ccgcatgagg 1080
ggcgcccaca ccaacgatgt caagcagctg accgaggccg tccagaagat cacgaccgag 1140
tccatcgtga tctgggggaa gacacccaag ttcaagctgc ctatccagaa ggagacctgg 1200
gagacgtggt ggaccgaata ttggcaggcc acctggattc ccgagtggga gttcgtgaat 1260
acacctctc tggatgaagct gtggtaccag ctcgagaagg agcccatcgt gggcgcgagg 1320
acattctacg tggacggcgc ggccaaccgc gaaacaaagc tcgggaaggc cgggtacgtc 1380
accaaccggg gccgccagaa ggtcgtcacc ctgaccgaca ccaccaacca gaagacggag 1440
ctgcaggcca tctatctcgc tctccaggac tccggcctgg aggtgaacat cgtgacggac 1500
agccagtacg cgctgggcat tattcaggcc cagccggacc agtccgagag cgaactggtg 1560
aaccagatta tcgagcagct gatcaagaaa gagaaggctt acctcgcttg ggtcccggcc 1620
cataagggca ttggcggcaa cgagcaggct gacaagctgg tgagtgcggg gattagaaaag 1680
gtgctgatgg gtgcccagc ttcggtactg tctggtggag agctggacag atggggagaa 1740
attaggctgc gcccgggagg caaaaagaaa tacaagctca agcatatcgt gtgggcctcg 1800
agggagcttg aacggtttgc cgtgaaccca ggcctgctgg aaacatctga gggatgtcgc 1860
cagatcctgg ggcaattgca gccatccctc cagaccggga gtgaagagct gaggtccttg 1920
tataacacag tggctaccct ctactgcgta caccagagga tcgagattaa ggataccaag 1980
gaggccttgg acaaaattga ggaggagcaa aacaagagca agaagaaggc ccagcaggca 2040
gctgctgaca ctgggcatag caaccaggta tcacagaact atcctattgt ccaaaacatt 2100
cagggccaga tggttcatca ggccatcagc ccccgagcgc tcaatgcctg ggtgaagggt 2160
gtcgaagaga aggccttttc tctgagggtt atccccatgt tctccgcttt gagtgagggg 2220
gccactctc aggaacctca tacaatgctt aataccgtgg gcggccatca ggccgccatg 2280
caaatgttga aggagactat caacgaggag gcagccgagt gggacagagt gcatcccgtc 2340
cacgtgggcc caatcgcgcc cggacagatg cgggagcctc gcggctctga cattgccggc 2400
accacctcta cactgcaaga gcaaatcgga tggatgacca acaatcctcc catcccagtt 2460
ggagaaatct ataaacggtg gatcatcctg ggcctgaaca agatcgtgcg catgtactct 2520
ccgacatcca tccttgacat tagacaggga ccaaagagc cttttaggga ttacgtcgac 2580
cggttttata agacctgcg agcagagcag gcctctcagg aggtcaaaaa ctggatgacg 2640
gagacactcc tggtagagaa cgctaacccc gactgcaaaa caatcttgaa ggcactaggc 2700
ccggctgcca ccttgaaga gatgatgacc gcctgtcagg gagtaggcgg acccgacac 2760
aaagccagag tgttgatggt gggttttcca gtcacacctc aggtaccttt aagaccaatg 2820
acttacaagg cagctgtaga tcttagccac tttttaaaag aaaagggggg actggaaggg 2880
ctaattcact cccaagaag acaagatac cttgatctgt ggatctacca cacacaaggc 2940
tacttccctg attggcagaa ctacacacca gggccagggg tcagatatcc actgaccttt 3000
ggatggtgct acaagctagt accagttgag ccagataagg tagaagaggc caataaagga 3060
gagaacacca gcttgttaca ccctgtgagc ctgcatggga tggatgaccc ggagagagaa 3120
gtgttagagt ggaggtttga cagccgccta gcatttcac acgtggcccc agagctgcat 3180
ccggagtact tcaagaactg ctga 3204

<210> 80
<211> 1067
<212> PRT
<213> Artificial Sequence

<220>
<223> HIV

<400> 80
Met Gly Pro Ile Ser Pro Ile Glu Thr Val Ser Val Lys Leu Lys Pro
1 5 10 15
Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys
20 25 30
Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys
35 40 45
Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala
50 55 60
Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg
65 70 75 80
Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile

				85					90				95		
Pro	His	Pro	Ala	Gly	Leu	Lys	Lys	Lys	Lys	Ser	Val	Thr	Val	Leu	Asp
			100					105					110		
Val	Gly	Asp	Ala	Tyr	Phe	Ser	Val	Pro	Leu	Asp	Glu	Asp	Phe	Arg	Lys
		115					120					125			
Tyr	Thr	Ala	Phe	Thr	Ile	Pro	Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile
	130					135					140				
Arg	Tyr	Gln	Tyr	Asn	Val	Leu	Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala
145				150					155						160
Ile	Phe	Gln	Ser	Ser	Met	Thr	Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln
			165						170					175	
Asn	Pro	Asp	Ile	Val	Ile	Tyr	Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly
		180						185					190		
Ser	Asp	Leu	Glu	Ile	Gly	Gln	His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg
	195						200					205			
Gln	His	Leu	Leu	Arg	Trp	Gly	Leu	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln
	210					215					220				
Lys	Glu	Pro	Pro	Phe	Leu	Trp	Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys
225				230					235						240
Trp	Thr	Val	Gln	Pro	Ile	Val	Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val
			245						250					255	
Asn	Asp	Ile	Gln	Lys	Leu	Val	Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile
		260						265					270		
Tyr	Pro	Gly	Ile	Lys	Val	Arg	Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr
	275						280					285			
Lys	Ala	Leu	Thr	Glu	Val	Ile	Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu
	290					295					300				
Leu	Ala	Glu	Asn	Arg	Glu	Ile	Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr
305				310					315						320
Tyr	Asp	Pro	Ser	Lys	Asp	Leu	Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln
			325						330					335	
Gly	Gln	Trp	Thr	Tyr	Gln	Ile	Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys
		340						345					350		
Thr	Gly	Lys	Tyr	Ala	Arg	Met	Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys
	355						360					365			
Gln	Leu	Thr	Glu	Ala	Val	Gln	Lys	Ile	Thr	Thr	Glu	Ser	Ile	Val	Ile
	370					375					380				
Trp	Gly	Lys	Thr	Pro	Lys	Phe	Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp
385				390					395						400
Glu	Thr	Trp	Trp	Thr	Glu	Tyr	Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp
			405						410					415	
Glu	Phe	Val	Asn	Thr	Pro	Pro	Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu
		420						425					430		
Lys	Glu	Pro	Ile	Val	Gly	Ala	Glu	Thr	Phe	Tyr	Val	Asp	Gly	Ala	Ala
	435						440					445			
Asn	Arg	Glu	Thr	Lys	Leu	Gly	Lys	Ala	Gly	Tyr	Val	Thr	Asn	Arg	Gly
	450					455					460				
Arg	Gln	Lys	Val	Val	Thr	Leu	Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu
465					470					475					480
Leu	Gln	Ala	Ile	Tyr	Leu	Ala	Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn
			485						490					495	
Ile	Val	Thr	Asp	Ser	Gln	Tyr	Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro
		500						505					510		
Asp	Gln	Ser	Glu	Ser	Glu	Leu	Val	Asn	Gln	Ile	Ile	Glu	Gln	Leu	Ile
		515					520					525			
Lys	Lys	Glu	Lys	Val	Tyr	Leu	Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile
	530					535					540				
Gly	Gly	Asn	Glu	Gln	Val	Asp	Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys
545				550					555						560
Val	Leu	Met	Gly	Ala	Arg	Ala	Ser	Val	Leu	Ser	Gly	Gly	Glu	Leu	Asp
			565					570						575	
Arg	Trp	Glu	Lys	Ile	Arg	Leu	Arg	Pro	Gly	Gly	Lys	Lys	Lys	Tyr	Lys
		580						585					590		
Leu	Lys	His	Ile	Val	Trp	Ala	Ser	Arg	Glu	Leu	Glu	Arg	Phe	Ala	Val

		595					600				605				
Asn	Pro	Gly	Leu	Leu	Glu	Thr	Ser	Glu	Gly	Cys	Arg	Gln	Ile	Leu	Gly
	610					615					620				
Gln	Leu	Gln	Pro	Ser	Leu	Gln	Thr	Gly	Ser	Glu	Glu	Leu	Arg	Ser	Leu
625					630					635					640
Tyr	Asn	Thr	Val	Ala	Thr	Leu	Tyr	Cys	Val	His	Gln	Arg	Ile	Glu	Ile
				645					650					655	
Lys	Asp	Thr	Lys	Glu	Ala	Leu	Asp	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys
			660					665					670		
Ser	Lys	Lys	Lys	Ala	Gln	Gln	Ala	Ala	Ala	Asp	Thr	Gly	His	Ser	Asn
		675					680					685			
Gln	Val	Ser	Gln	Asn	Tyr	Pro	Ile	Val	Gln	Asn	Ile	Gln	Gly	Gln	Met
	690					695					700				
Val	His	Gln	Ala	Ile	Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val
705					710					715					720
Val	Glu	Glu	Lys	Ala	Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Ser	Ala
			725						730					735	
Leu	Ser	Glu	Gly	Ala	Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr
			740					745					750		
Val	Gly	Gly	His	Gln	Ala	Ala	Met	Gln	Met	Leu	Lys	Glu	Thr	Ile	Asn
	755						760					765			
Glu	Glu	Ala	Ala	Glu	Trp	Asp	Arg	Val	His	Pro	Val	His	Ala	Gly	Pro
	770					775					780				
Ile	Ala	Pro	Gly	Gln	Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly
785					790					795					800
Thr	Thr	Ser	Thr	Leu	Gln	Glu	Gln	Ile	Gly	Trp	Met	Thr	Asn	Asn	Pro
				805					810					815	
Pro	Ile	Pro	Val	Gly	Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu
			820					825					830		
Asn	Lys	Ile	Val	Arg	Met	Tyr	Ser	Pro	Thr	Ser	Ile	Leu	Asp	Ile	Arg
		835					840					845			
Gln	Gly	Pro	Lys	Glu	Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Tyr	Lys
	850					855					860				
Thr	Leu	Arg	Ala	Glu	Gln	Ala	Ser	Gln	Glu	Val	Lys	Asn	Trp	Met	Thr
865					870					875					880
Glu	Thr	Leu	Leu	Val	Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu
				885					890					895	
Lys	Ala	Leu	Gly	Pro	Ala	Ala	Thr	Leu	Glu	Glu	Met	Met	Thr	Ala	Cys
			900					905					910		
Gln	Gly	Val	Gly	Gly	Pro	Gly	His	Lys	Ala	Arg	Val	Leu	Met	Val	Gly
		915				920						925			
Phe	Pro	Val	Thr	Pro	Gln	Val	Pro	Leu	Arg	Pro	Met	Thr	Tyr	Lys	Ala
	930					935					940				
Ala	Val	Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly	Leu	Glu	Gly
945					950					955					960
Leu	Ile	His	Ser	Gln	Arg	Arg	Gln	Asp	Ile	Leu	Asp	Leu	Trp	Ile	Tyr
				965					970					975	
His	Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr	Pro	Gly	Pro
			980					985					990		
Gly	Val	Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Tyr	Lys	Leu	Val	Pro
		995					1000						1005		
Val	Glu	Pro	Asp	Lys	Val	Glu	Ala	Asn	Lys	Gly	Glu	Asn	Thr	Ser	
	1010					1015				1020					
Leu	Leu	His	Pro	Val	Ser	Leu	His	Gly	Met	Asp	Asp	Pro	Glu	Arg	Glu
1025					1030					1035					1040
Val	Leu	Glu	Trp	Arg	Phe	Asp	Ser	Arg	Leu	Ala	Phe	His	His	Val	Ala
				1045					1050						1055
Arg	Glu	Leu	His	Pro	Glu	Tyr	Phe	Lys	Asn	Cys					
			1060					1065							

<210> 81

<211> 3204

<212> DNA

<213> Artificial Sequence

<220>

<223> HIV

<400> 81

```

atggggcccca tcagtcccat cgagaccgtg ccggtgaagc tgaaacccgg gatggacggc 60
cccaagggtca agcagtggcc actcaccgag gagaagatca aggccctggt ggagatctgc 120
accgagatgg agaaagaggg caagatcagc aagatcgggc cggagaaccc atacaacacc 180
cccgtgtttg ccatcaagaa gaaggacagc accaagtggc gcaagctggt ggatttccgg 240
gagctgaata agcggaccca ggatttctgg gaggtccagc tgggcatccc ccatccggcc 300
ggcctgaaga agaagaagag cgtgaccgtg ctggacgtgg gcgacgctta cttcagcgtc 360
cctctggacg aggacttttag aaagtacacc gcctttacca tcccatctat caacaacgag 420
acccctggca tcagatatca gtacaacgtc ctccccagg gctggaaggg ctctcccgcc 480
atthttccaga gctccatgac caagatcctg gagccgtttc ggaagcagaa ccccgatata 540
gtcatctacc agtacatgga cgacctgtac gtgggctctg acctggaaat cgggcagcat 600
cgcacgaaga ttgaggagct gaggcagcat ctgctgagat ggggcctgac cactccggac 660
aagaagcatc agaaggagcg gccattcctg aagatgggct acgagctcca tcccgacaag 720
tggaccgtgc agcctatcgt cctccccgag aaggacagct ggaccgtgaa cgacatccag 780
aagctggttg gcaagctcaa ctgggctagc cagatctatc ccgggatcaa ggtgcgccag 840
ctctgcaagc tgctgcgcgg caccaaggcc ctgaccgagg tgattccctt caccgaggaa 900
gccgagctcg agctggctga gaaccgggag atcctgaagg agcccgtgca cggcgtgtac 960
tatgaccctt ccaaggacct gatcgccgaa atccagaagc agggccaggg gcagtggaca 1020
taccagattt accaggagcc tttcaagaac ctcaagaccg gcaagtacgc ccgcatgagg 1080
ggcgcccaca ccaacgatgt caagcagctg accgaggccg tccagaagat caccgaccgag 1140
tccatcgtga tctgggggaa gacacccaag ttcaagctgc ctatccagaa ggagacctgg 1200
gagacgtggt ggaccgaata ttggcaggcc acctggattc ccgagtggga gttcgtgaat 1260
acacctcttc tgggtgaagct gtggtaccag ctcgagaagg agcccacgtt gggcgccggag 1320
acattctacg tggacggcgc ggccaaccgc gaaacaaagc tcgggaaggc cgggtacgtc 1380
accaaccggg gccgcacagaa ggctcgtcac ctgaccgaca ccaccaacca gaagacggag 1440
ctgcaggcca tctatctcgc tctccaggac tccggcctgg aggtgaacat cgtgacggac 1500
agccagtacg cgtgggcat tattcaggcc cagccggacc agtccgagag cgaactggtg 1560
aaccagatta tcgagcagct gatcaagaaa gagaaggctt acctgcctg ggtcccggcc 1620
cataagggca ttggcggcaa cgagcaggtc gacaagctgg tgagtgcggg gattagaaag 1680
gtgctgatgg tgggttttcc agtcacacct caggtacctt taagaccaat gacttacaag 1740
gcagctgtag atcttagcca ctttttaaaa gaaaaggggg gactggaagg gctaattcac 1800
tcccaaagaa gacaagatat ccttgatctg tggatctacc acacacaagg ctacttccct 1860
gattggcaga actacacacc agggccaggg gtcagatata cactgacctt tggatggtgc 1920
tacaagctag taccagttga gccagataag gtagaagagg ccaataaagg agagaacacc 1980
agcttgttac acctgtgag cctgcatggg atggatgacc cggagagaga agtgtagtag 2040
tggaggtttg acagccgct agcatttcat cacgtggccc gagagctgca tccggagtac 2100
ttcaagaact gctgaatggg tgcccagact tcggtactgt ctggtggaga gctggacaga 2160
tgggagaaaa ttaggtgctg cccgggaggc aaaaagaaat acaagctcaa gcatactgtg 2220
tgggcctcga gggagcttga acggtttgcc gtgaaccagc gcctgctgga aacatctgag 2280
ggatgtcgcc agatcctggg gcaattgcag ccacccctcc agaccgggag tgaagagctg 2340
aggtccttgt ataacacagt ggctaccctc tactgcgtac accagaggat cgagattaag 2400
gataccaagg aggccttgga caaaattgag gaggagcaaa acaagagcaa gaagaaggcc 2460
cagcaggcag ctgctgacac tgggcatagc aaccagggtat cacagaacta tcctattgtc 2520
caaaacattc agggccagat ggttcatcag gccatcagcc cccggacgct caatgcctgg 2580
gtgaagggtg tcgaagagaa ggccctttct cctgaggtta tccccatgtt ctccgctttg 2640
agtgaggggg cactcctca ggacctcaat acaatgctta ataccgtggg cgccatcag 2700
gccgccatgc aaatggtgaa ggagactatc aacgaggagg cagccgagtg ggacagagtg 2760
catcccgtcc acgctggccc aatcgcgccc ggacagatgc gggagcctcg cggctctgac 2820
attccgggca ccacctctac actgcaagag caaatcggat ggatgaccaa caatcctccc 2880
atcccagttg gagaaatcta taaacggttg atcatcctgg gcctgaacaa gatcgtgcgc 2940
atgtactctc cgacatccat ccttgacatt agacagggac ccaaagagcc ttttagggat 3000
tacgtcgacc ggttttataa gacctgcga gcagagcagg cctctcagga ggtcaaaaac 3060
tggatgacgg agacactcct ggtacagaac gctaaccctg actgcaaaac aatcttgaag 3120
gcaactaggcc cggctgccac cctggaagag atgatgaccg cctgtcaggg agtaggcgga 3180
cccgacaca aagccagagt gttg 3204

```

<210> 82

<211> 1067

<212> PRT

<213> Artificial Sequence

<220>

<223> HIV

<400> 82

```

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

```

Arg	Gln	Lys	Val	Val	Thr	Leu	Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu	465	470	475	480
Leu	Gln	Ala	Ile	Tyr	Leu	Ala	Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn	485	490		495
Ile	Val	Thr	Asp	Ser	Gln	Tyr	Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro	500	505		510
Asp	Gln	Ser	Glu	Ser	Glu	Leu	Val	Asn	Gln	Ile	Ile	Glu	Gln	Leu	Ile	515	520		525
Lys	Lys	Glu	Lys	Val	Tyr	Leu	Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	530	535		540
Gly	Gly	Asn	Glu	Gln	Val	Asp	Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys	545	550		555
Val	Leu	Met	Val	Gly	Phe	Pro	Val	Thr	Pro	Gln	Val	Pro	Leu	Arg	Pro	565	570		575
Met	Thr	Tyr	Lys	Ala	Ala	Val	Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	580	585		590
Gly	Gly	Leu	Glu	Gly	Leu	Ile	His	Ser	Gln	Arg	Arg	Gln	Asp	Ile	Leu	595	600		605
Asp	Leu	Trp	Ile	Tyr	His	Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	610	615		620
Tyr	Thr	Pro	Gly	Pro	Gly	Val	Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	625	630		635
Tyr	Lys	Leu	Val	Pro	Val	Glu	Pro	Asp	Lys	Val	Glu	Glu	Ala	Asn	Lys	645	650		655
Gly	Glu	Asn	Thr	Ser	Leu	Leu	His	Pro	Val	Ser	Leu	His	Gly	Met	Asp	660	665		670
Asp	Pro	Glu	Arg	Glu	Val	Leu	Glu	Trp	Arg	Phe	Asp	Ser	Arg	Leu	Ala	675	680		685
Phe	His	His	Val	Ala	Arg	Glu	Leu	His	Pro	Glu	Tyr	Phe	Lys	Asn	Cys	690	695		700
Met	Gly	Ala	Arg	Ala	Ser	Val	Leu	Ser	Gly	Gly	Glu	Leu	Asp	Arg	Trp	705	710		715
Glu	Lys	Ile	Arg	Leu	Arg	Pro	Gly	Gly	Lys	Lys	Lys	Tyr	Lys	Leu	Lys	725	730		735
His	Ile	Val	Trp	Ala	Ser	Arg	Glu	Leu	Glu	Arg	Phe	Ala	Val	Asn	Pro	740	745		750
Gly	Leu	Leu	Glu	Thr	Ser	Glu	Gly	Cys	Arg	Gln	Ile	Leu	Gly	Gln	Leu	755	760		765
Gln	Pro	Ser	Leu	Gln	Thr	Gly	Ser	Glu	Glu	Leu	Arg	Ser	Leu	Tyr	Asn	770	775		780
Thr	Val	Ala	Thr	Leu	Tyr	Cys	Val	His	Gln	Arg	Ile	Glu	Ile	Lys	Asp	785	790		800
Thr	Lys	Glu	Ala	Leu	Asp	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Lys	805	810		815
Lys	Lys	Ala	Gln	Gln	Ala	Ala	Ala	Asp	Thr	Gly	His	Ser	Asn	Gln	Val	820	825		830
Ser	Gln	Asn	Tyr	Pro	Ile	Val	Gln	Asn	Ile	Gln	Gly	Gln	Met	Val	His	835	840		845
Gln	Ala	Ile	Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val	Val	Glu	850	855		860
Glu	Lys	Ala	Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Ser	Ala	Leu	Ser	865	870		880
Glu	Gly	Ala	Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly	885	890		895
Gly	His	Gln	Ala	Ala	Met	Gln	Met	Leu	Lys	Glu	Thr	Ile	Asn	Glu	Glu	900	905		910
Ala	Ala	Glu	Trp	Asp	Arg	Val	His	Pro	Val	His	Ala	Gly	Pro	Ile	Ala	915	920		925
Pro	Gly	Gln	Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Thr	Thr	930	935		940
Ser	Thr	Leu	Gln	Glu	Gln	Ile	Gly	Trp	Met	Thr	Asn	Asn	Pro	Pro	Ile	945	950		955
Pro	Val	Gly	Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys	965	970		975

```

Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly
      980                      985                      990
Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
      995                      1000                      1005
Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr
      1010                      1015                      1020
Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
      1025                      1030                      1035                      1040
Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
      1045                      1050                      1055
Val Gly Gly Pro Gly His Lys Ala Arg Val Leu
      1060                      1065

```

```

<210> 83
<211> 3204
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> HIV

```

```

<400> 83
atgggtgccc gagcttcggt actgtctggt ggagagctgg acagatggga gaaaattagg 60
ctgcgccccg gaggcacaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgagggag 120
cttgaacggg ttgccgtgaa cccaggcctg ctggaaacat ctgagggatg tgcgagatc 180
ctggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggtc cttgtataac 240
acagtggcta ccctctactg cgtacaccag aggatcgaga ttaaggatac caaggaggcc 300
ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga aggccagca ggcagctgct 360
gacactgggc atagcaacca ggtatcacag aactatccta ttgtccaaaa cattcagggc 420
cagatgggtt atcaggccat cagccccggg acgctcaatg cctgggtgaa ggttgtcgaa 480
gagaaggcct tttctcctga ggttatcccc atgttctccg ctttgagtga gggggccact 540
cctcaggacc tcaatacaat gcttaataac gtgggcggcc atcaggccgc catgcaaatg 600
ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gagtgcattc cgtccacgct 660
ggcccaatcg cgcccggaac gatgcgggag cctcgcggct ctgacattgc cggcaccacc 720
tctactctgc aagagcaaat cggatggatg accaacaatc ctcccatccc agttggagaa 780
atctataaac ggtggatcat cctgggcctg aacaagatcg tgcgcatgta ctctccgaca 840
tccatccttg acattagaca gggacccaaa gagcctttta gggattacgt cgaccggttt 900
tataagaccc tgcgagcaga gcaggcctct caggaggtca aaaactggat gacggagaca 960
ctcctggtac agaacgctaa ccccgactgc aaaacaatct tgaaggcact aggcccggt 1020
gccaccctgg aagagatgat gaccgcctgt cagggagtag gcggacccgg acacaaagcc 1080
agagtgttga tgggtgggtt tccagtcaca cctcaggtag ctttaagacc aatgacttac 1140
aaggcagctg tagatcttag ccacttttta aaagaaaagg ggggactgga agggctaatt 1200
cactcccaaa gaagacaaga tatccttgat ctgtggatct accacacaca aggtacttct 1260
cctgattggc agaactacac accagggcca ggggtcagat atccactgac ctttggatgg 1320
tgctacaagc tagtaccagt tgagccagat aaggtagaag aggccataaa aggagagaac 1380
accagcttgt tacaccctgt gagcctgcat gggatggatg acccgagag agaatgttta 1440
gagtggaggt ttgacagccg cctagcattt catcacgtgg cccgagagct gcatccggag 1500
tacttcaaga actgcatggg ccccatcagt cccatcgaga ccgtgccggt gaagctgaaa 1560
ccggggatgg acggcccaaa ggtcaagcag tggccactca ccgaggagaa gatcaaggcc 1620
ctggtggaga tctgcaccga gatggagaaa gagggcaaga tcagcaagat cgggcctgag 1680
aaccataaca acaccccgct gtttgccatc aagaagaagg acagcaccaa gtggcgcaag 1740
ctggtggatt tccgggagct gaataagcgg acccaggatt tctgggaggt ccagctgggc 1800
atccccatc cgcccgccct gaagaagaag aagagcgtga ccgtgctgga cgtgggcgac 1860
gcttacttca gcgtccctct ggacgaggac tttagaaagt acaccgcctt taccatccca 1920
tctatcaaca acgagacccc tggcatcaga tatcagtaca acgtcctccc ccagggctgg 1980
aagggtctct ccgccatttt ccagagctcc atgaccaaga tcctggagcc gtttcggaag 2040
cagaaccccg atatcgatc ctaccagtac atggacgacc tgtacgtggg ctctgacctg 2100
gaaatcgggc agcatcgcac gaagattgag gagctgaggc agcatctgct gagatggggc 2160
ctgaccactc cggacaagaa gcatcagaag gagccgcat tcctgaagat gggctacgag 2220
ctccatcccc acaagtggac cgtgcagcct atcgctctcc ccgagaagga cagctggacc 2280
gtgaacgaca tccagaagct ggtgggcaag ctcaactggg ctagccagat ctatcccggg 2340
atcaaggtgc gccagctctg caagctgctg cgcggcacca aggccctgac cgaggtgatt 2400
cccctcacgg aggaagccga gctcgagctg gctgagaacc gggagatcct gaaggagccc 2460

```

gtgcacggcg tgtactatga cccctccaag gacctgatcg ccgaaatcca gaagcagggc 2520
 caggggagcgt ggacatacca gatttaccag gagcctttca agaacctcaa gaccggcaag 2580
 tacgcccgcga tgagggggcgc ccacaccaac gatgtcaagc agctgaccga ggccgtccag 2640
 aagatcacga ccgagttccat cgtgatctgg gggaagacac ccaagttcaa gctgcctatc 2700
 cagaaggaga cctgggagac gtggtggacc gaatattggc aggccacctg gattcccag 2760
 tgggagttcg tgaatacacc tcctctgggtg aagctgtggt accagctcga gaaggagccc 2820
 atcgtgggagc cggagacatt ctacgtggac ggcgcgccca accgcgaaac aaagctcggg 2880
 aaggccgggt acgtcaccaa ccggggccgc cagaaggctcgc tcaccctgac cgacaccacc 2940
 aaccagaaga cggagctgca ggccatctat ctgcgtctcc aggactccgg cctggaggtg 3000
 aacatcgtga cggacagcca gtacgcgctg ggcattattc aggcccagcc ggaccagtcc 3060
 gagagcgaac tgggtgaacca gattatcgag cagctgatca agaaagagaa ggtctacctc 3120
 gcctgggtcc cggcccataa gggcattggc ggcaacgagc aggtcgacaa gctggtgagt 3180
 gcgggggatta gaaaggtgct gtaa 3204

<210> 84
 <211> 1067
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> HIV

<400> 84
 Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp
 1 5 10 15
 Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys
 20 25 30
 His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro
 35 40 45
 Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
 50 55 60
 Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn
 65 70 75 80
 Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp
 85 90 95
 Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
 100 105 110
 Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn Gln Val
 115 120 125
 Ser Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His
 130 135 140
 Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu
 145 150 155 160
 Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
 165 170 175
 Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly
 180 185 190
 Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu
 195 200 205
 Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala
 210 215 220
 Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 225 230 235 240
 Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile
 245 250 255
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270
 Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly
 275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
 290 295 300
 Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr
 305 310 315 320
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala

				325					330					335			
Leu	Gly	Pro	Ala	Ala	Thr	Leu	Glu	Glu	Met	Met	Thr	Ala	Cys	Gln	Gly		
			340					345					350				
Val	Gly	Gly	Pro	Gly	His	Lys	Ala	Arg	Val	Leu	Met	Val	Gly	Phe	Pro		
		355					360					365					
Val	Thr	Pro	Gln	Val	Pro	Leu	Arg	Pro	Met	Thr	Tyr	Lys	Ala	Ala	Val		
	370					375					380						
Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly	Leu	Glu	Gly	Leu	Ile		
385				390					395						400		
His	Ser	Gln	Arg	Arg	Gln	Asp	Ile	Leu	Asp	Leu	Trp	Ile	Tyr	His	Thr		
			405					410						415			
Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr	Pro	Gly	Pro	Gly	Val		
		420						425					430				
Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Tyr	Lys	Leu	Val	Pro	Val	Glu		
	435						440					445					
Pro	Asp	Lys	Val	Glu	Glu	Ala	Asn	Lys	Gly	Glu	Asn	Thr	Ser	Leu	Leu		
	450					455					460						
His	Pro	Val	Ser	Leu	His	Gly	Met	Asp	Asp	Pro	Glu	Arg	Glu	Val	Leu		
465				470					475						480		
Glu	Trp	Arg	Phe	Asp	Ser	Arg	Leu	Ala	Phe	His	His	Val	Ala	Arg	Glu		
			485					490						495			
Leu	His	Pro	Glu	Tyr	Phe	Lys	Asn	Cys	Met	Gly	Pro	Ile	Ser	Pro	Ile		
		500						505				510					
Glu	Thr	Val	Ser	Val	Lys	Leu	Lys	Pro	Gly	Met	Asp	Gly	Pro	Lys	Val		
	515						520					525					
Lys	Gln	Trp	Pro	Leu	Thr	Glu	Glu	Lys	Ile	Lys	Ala	Leu	Val	Glu	Ile		
	530					535					540						
Cys	Thr	Glu	Met	Glu	Lys	Glu	Gly	Lys	Ile	Ser	Lys	Ile	Gly	Pro	Glu		
545				550					555						560		
Asn	Pro	Tyr	Asn	Thr	Pro	Val	Phe	Ala	Ile	Lys	Lys	Lys	Asp	Ser	Thr		
			565					570						575			
Lys	Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg	Thr	Gln		
		580						585					590				
Asp	Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	Leu	Lys		
	595					600						605					
Lys	Lys	Lys	Ser	Val	Thr	Val	Leu	Asp	Val	Gly	Asp	Ala	Tyr	Phe	Ser		
	610					615				620							
Val	Pro	Leu	Asp	Glu	Asp	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	Ile	Pro		
625				630						635					640		
Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	Val	Leu		
			645					650						655			
Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Ser	Ser	Met	Thr		
		660						665					670				
Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro	Asp	Ile	Val	Ile	Tyr		
	675					680						685					
Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	Gly	Gln		
	690				695					700							
His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Gln	His	Leu	Leu	Arg	Trp	Gly		
705				710					715						720		
Leu	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe	Leu	Trp		
			725					730						735			
Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	Ile	Val		
		740						745					750				
Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	Leu	Val		
		755				760						765					
Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro	Gly	Ile	Lys	Val	Arg		
	770					775					780						
Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala	Leu	Thr	Glu	Val	Ile		
785				790					795						800		
Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	Glu	Ile		
			805					810						815			
Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	Asp	Leu		
			820					825					830				
Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr	Gln	Ile		

<400> 87
Tyr Tyr Asp Pro Ser Lys Asp Leu Ile
1 5