



SUOMI – FINLAND  
(FI)

PATENTTI- JA REKISTERIHALLITUS  
PATENT- OCH REGISTERSTYRELSEN

(12) **PATENTTIJULKAISU**  
**PATENTSKRIFT**

(10) **FI 120495 B**

(45) Patenti myönnetty - Patent beviljats

13.11.2009

(51) Kv.lk. - Int.kl.

**C07K 14/155** (2006.01)  
**C07K 14/025** (2006.01)  
**C07K 19/00** (2006.01)  
**C12N 15/49** (2006.01)  
**C12N 15/62** (2006.01)  
**C12N 5/10** (2006.01)  
**C12P 21/02** (2006.01)

(21) Patentihakemus - Patentansökning

945248

(22) Tekemispäivä - Ingivningsdag

08.11.1994

(24) Alkupäivä - Löpdag

19.08.1993

(41) Tullut julkiseksi - Blivit offentlig

05.01.1995

(86) Kv. hakemus - Int. ansökan

PCT/US1993/007833

(32) (33) (31) Etuoikeus - Prioritet

21.08.1992 US 934375 P

(73) Haltija - Innehavare

**1 •Biogen Idec MA Inc.**, 15 Cambridge Center, CAMBRIDGE, MA 02142, AMERIKAN YHDYSVALLAT, (US)

(72) Keksijä - Uppfinnare

**1 •Barsoum, James G.**, Nine Marlboro Road, Lexington, MA 02173, AMERIKAN YHDYSVALLAT, (US)  
**2 •Fawell, Stephen E.**, One Black Horse Terrace, Winchester, MA 01890, AMERIKAN YHDYSVALLAT, (US)  
**3 •Pepinsky, R. Blake**, 30 Falmouth Road, Arlington, MA 02174, AMERIKAN YHDYSVALLAT, (US)

(74) Asiamies - Ombud

**Kolster Oy Ab**, Iso Roobertinkatu 23, 00120 Helsinki

(54) Keksinnön nimitys - Uppfinningens benämning

**Tat-peräisiä kuljetuspolypeptidejä**  
**Tat-härledda transportpolypeptider**

(83) Mikro-organismitalletus - Deposition av mikroorganism

69368 ATCC

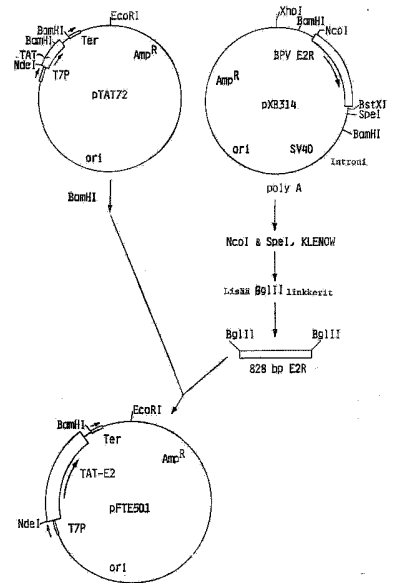
(56) Viitejulkaisut - Anförda publikationer

WO 91/09958 A2

(57) Tiivistelmä - Sammandrag

Tämä keksintö koskee biologisesti aktiivisten lastimolekyylien, kuten polypeptidien ja nukleiinihappojen, viemistä solujen sytoplasmaan ja tumiin in vitro ja in vivo käyttämällä uusia kuljetuspolypeptidejä, jotka käsittävät yhden tai useamman osan HIV:n tat-proteiinia ja jotka on liitetty kovalenttisesti lastimolekyyliin. Tämän keksinnön mukaisille kuljetuspolypeptideille on tunnusmerkillistä, että luonnossa ilmenevästä tat-proteiinista niissä on läsnä tat:n emäksinen alue (aminohapot 49 - 57), niistä puuttuu tat:n kysteiniipitoinen alue (aminohapot 22 - 36), ja niistä puuttuu tat:n eksoni 2:n koodittama karboksiterminaalinen alue (aminohapot 73 - 86). Tavanomaisista tat-proteiineista löytyvän kysteiniipitoisen alueen poissaolo ratkaisee väärän transaktivaation ja disulfidiaggregaation ongelmat.

Uppfinningen avser transporten av biologiskt aktiva lastmolekyler, såsom polypeptider och nukleinsyror, in till cellers cytoplasma och kärna in vitro och in vivo med hjälp av nya transport polypeptider, vilka innehåller en eller flera delar av HIV tat-protein och vilka är kovalent bundna till lastmolekylerna. Denna upfinningens transport polypeptider består av närvaron av en basisk region (aminosyror 49 - 57), frånvaron av en tat cystein-rik region (aminosyror 22 - 36) och frånvaron av tat exon 2-kodande karboxiterminal domenen (aminosyror 73 - 86) av den i naturen förekommande tat-proteinen. Frånvaron av cystein-rika regionen, vilken existerar i vanliga tat-proteiner, löser problemet, vilket är förknippat till fel transaktivaation och disulfid aggregaation.



## Tat-peräisiä kuljetuspolypeptidejä

Tämä keksintö on jatkohakemus samaan aikaan vireillä olevalle hakemukselle, jonka sarjanumero on 07/934 375 ja joka on jätetty 21. elokuuta 1992.

### Keksinnön tekninen ala

Tämä keksintö koskee biologisesti aktiivisten lastimolekyylien, kuten polypeptidien ja nukleiinihappojen, viemistä solujen sytoplasmaan ja tumiin in vitro. Tämän keksinnön mukainen lastimolekyylien solunsisäinen vieminen saadaan aikaan käyttämällä uusia kuljetuspolypeptidejä, jotka käsittävät yhden tai useamman osan HIV:n tat-proteiinista, ja jotka on kiinnitetty kovalenttisesti lastimolekyyliin. Tämän keksinnön mukaisille kuljetuspolypeptideille on tunnusmerkillistä se, että luonnossa ilmenevän tat-proteiinin tat:n emäksinen alue (aminohapot 49 - 57) on läsnä, tat:n kysteiniipitoinen alue (aminohapot 22 - 36) on poissa, ja tat:n eksoni 2:n koodittama karboksiterminaalinen alue (aminohapot 73 - 86) on poissa. Tavanomaisista tat-proteiineista löytyvän kysteiniipitoisen alueen poissaolon ansiosta tämän keksinnön mukaiset kuljetuspolypeptidit ratkaisevat väärän transaktivaation ja disulfi-diaggregaation ongelmat. Tämän keksinnön mukaisten kuljetuspolypeptidien pienentynyt koko minimoi myös lastimolekyylin biologisen aktiivisuuden häiritsemisen.

### Keksinnön tausta

Biologiset solut ovat yleensä läpäisemättömiä makromolekyylien suhteen, mukaan lukien proteiinit ja nukleiinihapot. Jotkin pienet molekyylit menevät solujen sisään hyvin alhaisilla nopeuksilla. Se, että ei ole ollut keinoja viedä makromolekyyliä soluihin in vivo, on ollut este sille, että käytettäisiin terapeuttisesti, profylaktisesti ja diagnostisesti mahdollisesti suurta määrää proteiineja ja

nukleiinihappoja, joiden toimintapaikka on solunsisäinen. Niinpä useimmat terapeuttiset, profylaktiset ja diagnostiset ehdokkaat, jotka tähän saakka on tuotettu yhdistelmä-DNA-teknologiaa käyttämällä, ovat polypeptidejä, jotka toimivat solunulkoisessa ympäristössä tai kohdesolun pinnalla.

5 On kehitetty useita menetelmiä makromolekyylien viemiseksi soluihin in vitro. Luetteloon tällaisista menetelmistä kuuluvat elektroporaatio, membraanifuusio liposomien kanssa, suurinopeuksinen pommitus DNA-päälysteisillä mikroammuksilla, inkubointi kalsiumfosfaatti-DNA-sakan 10 kanssa, DEAE-dekstraani-välitteinen transfektio, infektiomuokatuilla virusnukleiinihappoilla ja suora mikroinjektio yksittäisiin soluihin. Nämä in vitro -menetelmät vievät tyypillisesti nukleiinihappomolekyyliä vain osaan kokoluopopulaatiosta, ja niillä on taipumus vahingoittaa suurta 15 joukkoa soluja. Makromolekyylien kokeellinen vieminen soluihin in vivo on saatu aikaan raaputuslataamisella ("scrape loading"), kalsiumfosfaattisakoilla ja liposomeilla. Nämä tekniikat ovat kuitenkin osoittautuneet tähän mennessä rajoitusti käyttökelpoisiksi soluihin viemisessä in vivo. 20 Lisäksi jopa soluissa in vitro tällaisten menetelmien käyttökelpoisuus on äärimmäisen rajoittunut proteiinien viemisen kannalta.

Tarvitaan yleisiä menetelmiä biologisesti aktiivisten proteiinien viemiseksi ehjiin soluihin in vitro ja in vivo. (L. A. Sternson, "Obstacles to Polypeptide Delivery", Ann. N. Y. Acad. Sci., 57, s. 19 - 21 (1987)). Lipopeptidin (P. Hoffmann ym., "Stimulation of Human and Murine Adherent Cells by Bacterial Lipoprotein and Synthetic Lipopeptide Analogues", Immunobiol., 177, s. 158 - 170 (1988)) tai emäksisen polymeerin, kuten polylysiinin tai polyarginiinin, kemiallinen lisäys (W-C. Chen ym., "Conjugation of Poly-L-Lysine Albumin and Horseradish Peroxidase: A Novel 30

Method of Enhancing the Cellular Uptake of Proteins", Proc. Natl. Acad. Sci. USA, 75, s. 1872 - 1876 (1978)) eivät ole osoittautuneet kovin luotettaviksi tai yleisesti käyttökelpoiksi (katso esimerkki 4 jäljempänä). Foolihappoa on käytetty kuljetusosana (C. P. Leamon ja Low, Delivery of Macromolecules into Living Cells: A Method That Exploits Folate Receptor Endocytosis", Proc. Natl. Acad. Sci. USA, 88, s. 5572 - 5576 (1991)). Folaattikonjugaattien internalisaatiosta esitettiin todisteita, mutta ei sytoplasmaan viemisestä. Kun otetaan huomioon kiertävän folaatin korkeat tasot in vivo, tämän järjestelmän hyödyllisyyttä ei ole täysin osoitettu. Pseudomonas-eksotoksiinia on myös käytetty kuljetusosana (T. I. Prior ym. "Barnase Toxin: A New Chimeric Toxin Composed of Pseudomonas Exotoxin A and Barnase", Cell, 64, s. 1017 - 1023 (1991)). Tämän järjestelmän tehokkuus ja yleinen soveltuvuus ei kuitenkaan ole selvä julkaistun työn perusteella.

Ihmisen immuunikatoviruksen 1-tyyppin ("HIV") tat-proteiini on osoittanut potentiaalia lastiproteiinien soluihin viemiseen (julkaistu PCT-hakemus WO 91/09 958). Kun kuitenkin otetaan huomioon täyspitkän tat-proteiinin kemialliset ominaisuudet, alalla ei ole esitetty yleisesti käyttökelpoisia menetelmiä sen tehokkaaksi käyttämiseksi biologisesti aktiivisen lastin viemisessä.

Tat on HIV:n koodittama proteiini, joka transaktivoi tiettyjä HIV-geenejä ja on oleellinen viruksen replikaatiolle. HIV-1:n täyspitkässä tat-proteiinissa on 86 aminohappotähdettä. HIV:n tat-geenissä on kaksi eksonia. Tat:n aminohappoja 1 - 72 koodittaa eksoni 1, ja aminohappoja 73 - 86 koodittaa eksoni 2. Täyspitkälle tat-proteiinille ominainen on emäksinen alue, joka käsittää kaksi lysiniä ja kuusi arginiinia (aminohapot 49 - 57), ja kysteiniipitoinen alue, joka käsittää seitsemän kysteinitähdettä

(aminohapot 22 - 37). Viljelmässä kasvavat ihmissolut ottavat sisäänsä puhdistettua tat-proteiinia ympäröivästä elatusaineesta (A. D. Frankel ja C. O. Pabo, "Cellular Uptake of the Tat Protein from Human Immunodeficiency Virus", Cell, 55, s. 1189 - 1193 (1988)). Alalla ei ole esitetty sitä, vaaditaanko tat-proteiinin kysteiinipitoista aluetta (joka saa aikaan aggregaatiota ja huonoliukoisuutta) tat-proteiinin ottamiseen solun sisään.

PCT-patenttihakemus WO 91/09 958 ("958-hakemus") paljastaa, että heterologinen proteiini, joka käsittää HIV:n tat-proteiinin aminohapot 1 - 67, jotka on fuusioitu geneettisesti papilloomaviruksen E2-transaktivaatiorepressoripolypeptidiin, otetaan sisään soluihin. Lastipolypeptidin biologisen aktiivisuuden (E2-transaktivaation repressio) säilymistä ei kuitenkaan tässä osoiteta.

Tat-proteiinin käyttö, kuten '958-hakemuksessa osoitetaan, käsittää mahdollisesti käytännön vaikeuksia, kun sitä käytetään lastiproteiinien viemiseen soluun. Nämä käytännön vaikeudet käsittävät proteiinin aggregoitumisen ja liukenemattomuuden, joihin liittyy tat-proteiinin kysteiinipitoinen alue. Lisäksi '958-hakemus ei tarjoa esimerkkejä tat-proteiinin kemiallisesta silloittamisesta lastiproteiinien kanssa, mikä voi olla kriittisen tärkeää tilanteissa, joissa tat:n geneettinen fuusio lastiproteiiniin häiritsee tat-proteiinin, lastiproteiinin tai molempien oikeaa laskostumista. Lisäksi sekä '958-hakemus että Frankel ja Pabo (edellä) esittävät tat-kuljetusproteiinien käyttöä yhdessä klorokiinin kanssa, joka on sytotoksinen. Niinpä on olemassa tarve yleensä sovellettavissa olevalle keinolle, jolla voitaisiin viedä turvallisesti ja tehokkaasti biologisesti aktiivisia lastimolekyylejä elävien solujen sytoplasmaan ja tumiin.

### Keksinnön yhteenveto

Tämä keksintö ratkaisee edellä esitetyt ongelmat tuomalla käytettäviksi tuotteita, joilla sytoplasmaan ja tumaan voidaan viedä tehokkaasti biologisesti aktiivisia ei-tat-proteiineja, nukleiinihappoja ja muita molekyyliä, jotka eivät 1) sinänsä pysty menemään sisään kohdesoluihin tai solujen tumiin, tai 2) eivät sinänsä pysty menemään sisään kohdesoluihin hyödyllisellä nopeudella. Kohdemolekyylien solunsisäinen antaminen saadaan aikaan käyttämällä uusia kuljetusproteiineja, jotka käsittävät yhden tai useamman osan HIV:n tat-proteiinista ja jotka on kiinnitetty kovalenttisesti lastimolekyyliin. Erityisemmin tämä keksintö koskee uusia kuljetuspolypeptidejä, menetelmiä, joilla voidaan valmistaa näitä kuljetuspolypeptidejä, kuljetuspolypeptidi-lastikonjugaatteja, farmaseuttisia, profylaktisia ja diagnostisia koostumuksia, jotka käsittävät kuljetuspolypeptidi-lastikonjugaatteja.

Tämän keksinnön kuljetuspolypeptideille on tunnusomaista tat:n emäksisen alueen aminohapposekvenssin läsnäolo (luonnossa esiintyvän tat-proteiinin aminohapot 49 - 57), tat:n kysteiinipitoisen alueen aminohapposekvenssin poissaolo (luonnossa esiintyvän tat-proteiinin aminohapot 22 - 36) ja tat:n eksoni 2:n koodittaman karboksiterminaalisen osan poissaolo (luonnossa ilmenevän tat-proteiinin aminohapot 73 - 86). Tällaisten kuljetuspolypeptidien edullisia suoritusmuotoja ovat: tat37-72 (SEQ ID nro 2), tat37-58 (SEQ ID nro 3), tat38-58GGC (SEQ ID nro 4), tatCGG47-58 (SEQ ID nro 5), tat47-58GGC (SEQ ID nro 6), ja tat $\Delta$ cys (SEQ ID nro 7). Ammattimiehet ymmärtävät, että kun kuljetuspolypeptidi fuusioidaan geneettisesti lastiosaan, täytyy lisätä aminotermiinalinen metioniini, mutta väliosian aminohappoja (esim. CysGlyGly tai GlyGlyCys) ei tarvitse lisätä. Sen ansiosta, että tavanomaisissa tat-proteiineissa läsnä

oleva kysteiinipitoinen alue on poissa, tämän keksinnön mukaiset kuljetuspolyeptidit ratkaisevat disulfidiaggregaation ongelman, joka voi johtaa lastin biologisen aktiivisuuden katoamiseen, kuljetuspolyeptidi-lastikonjugaatin liukenemattomuuteen tai molempiin. Tämän keksinnön mukaisen kuljetuspolyeptidien pienentynyt koko minimoi myös edullisella tavalla lastin biologisen aktiivisuuden häirintää. Pienentyneen kuljetuspolyeptidin koon etuna on lisäksi viemisen lisääntynyt teho niissä tämän keksinnön mukaisissa suoritusmuodoissa, joihin kuuluu useampien kuljetuspolyeptidien kiinnittäminen lastimolekyyliä kohden.

Tämän keksinnön mukaiset kuljetuspolyeptidit voidaan kiinnittää edullisesti lastimolekyyliin kemiallisesti silloittamalla tai geneettisellä fuusiolla. Tämän keksinnön edullisten suoritusmuotojen mukaisesti kuljetuspolyeptidi ja lastimolekyyli ovat kemiallisesti silloitettuja. Ainutkertainen terminaalinen kysteiinitähde on edullinen keino kemialliseen silloitukseen. Muiden tämän keksinnön edullisten suoritusmuotojen mukaisesti kuljetusosan karboksipää fuusioidaan geneettisesti lastiosan aminopäähän. Eriytyisen edullinen tämän keksinnön mukainen suoritusmuoto on JB106, joka koostuu aminotermiinisestä metioniinista, josta seuraavat tat-tähteet 47-58, joita seuraavat HPV-16:n E2-tähteet 245-365.

Monissa tapauksissa tämän keksinnön mukaiset uudet kuljetuspolyeptidit välttävät edullisella tavalla klorokiiniin liittyvän sytotoksisuuden. Edellä mainittujen piirteiden ansiosta tämä keksintö avaa tien sellaiselle biologiselle tutkimukselle ja taudin hoidolle, joihin liittyvät proteiinit, nukleiinihapot ja muut molekyylit, joiden vaikutuspaikat ovat sytoplasmassa tai tumassa.



### Kuvoiden lyhyt kuvaus

Kuvio 1 kuvaa HIV-1:n tat-proteiinin aminohapposekvenssiä (SEQ ID nro 1).

5 Kuvio 2 esittää yhteenvetona tulokset solujen sisäänottokokeista, joissa käytettiin kuljetuspolypeptidin ja Pseudomonas-eksotoksiinin ribosylaatioalueen muodostamia konjugaatteja (varjostetut pylvää: konjugoimattomat, vino- viivoitetut pylvää, konjugoidut).

10 Kuvio 3 esittää yhteenvetona tulokset solujen sisäänottokokeista, joissa käytettiin kuljetuspolypeptidin ja ribonukleaasin muodostamia konjugaatteja (täytetyt neliöt, ribonukleaasi-SMCC ilman kuljetusosaa, täytetyt ympyrät, tat37-72-ribonukleaasi, täytetyt kolmiot, tat38-58GGC-ribonukleaasi, täytetyt vinoneliöt, tatCGG38-58-ribonukleaasi, 15 avoimet neliöt, tatCGG47-58-ribonukleaasi).

Kuvio 4 esittää kaavamaisesti plasmidin pAHE2 muodostamisen.

Kuvio 5 esittää kaavamaisesti plasmidin pET8c123 muodostamisen.

20 Kuvio 6 esittää kaavamaisesti plasmidin pET8c123CCSS muodostamisen.

Kuvio 7 esittää yhteenvetona tulokset solujen sisäänottokokeista, joissa käytettiin kuljetuspolypeptidin ja E2-repressorin muodostamia konjugaatteja (avoimet vinoneliöt, E2.123 silloitettuna tat37-72:een ilman klorokiiniä, 25 täytetyt vinoneliöt, E2.123 silloitettuna tat37-72:een klorokiinin kanssa, avoimet ympyrät, E2.123CCSS silloitettuna tat37-72:een ilman klorokiiniä, suljetut ympyrät, E2.123CCSS silloitettuna tat37-72:een klorokiinin kanssa).

30 Kuvio 8 esittää kaavamaisesti plasmidin pTATAcys muodostamisen.

Kuvio 9 esittää kaavamaisesti plasmidin pFTE501 muodostamisen.

Kuvio 10 esittää kaavamaisesti plasmidin pTATAcys-249 muodostamisen.

Kuvio 11 esittää kaavamaisesti plasmidin pJB106 muodostamisen.

5 Kuvio 12 esittää proteiinin JB106 täydellisen aminohapposekvenssin.

10 Kuvio 13 esittää yhteenvetona E2-repressiomääritysten tulokset, joissa oli mukana JB106 (neliöt), TxHE2CCSS (vinoneliöt) ja HE2.123 (ympyrät). Määritykset suoritettiin COS7-soluissa ilman klorokiiniä, kuten esimerkissä 14 kuvataan.

#### **Keksinnön yksityiskohtainen kuvaus**

Jotta tässä kuvattu keksintö voidaan ymmärtää täydellisemmin, esitetään seuraava yksityiskohtainen kuvaus.

15 Kuvauksessa käytetään seuraavia termejä:

Aminohappo - Peptidin, polypeptidin tai proteiinin monomeerinen yksikkö. Kaksikymmentä proteiiniaminohappoa (L-isomeerit) ovat: alaniini ("Ala" tai "A"), arginiini ("Arg" tai "R"), asparagiini ("Asn" tai "N"), asparagiinihappo ("Asp" tai "D"), kysteiini ("Cys" tai "C"), glutamiini ("Gln" tai "Q"), glutamiinihappo ("Glu" tai "E"), glysiini ("Gly" tai "G"), histidiini ("His" tai "H"), isoleusiini ("Ile" tai "I"), leusiini ("Leu" tai "L"), lysiini ("Lys" tai "K"), metioniini ("Met" tai "M"), fenyyialaniini ("Phe" tai "F"), proliini ("Pro" tai "P"), seriini ("Ser" tai "S"), treoniini ("Thr" tai "T"), tryptofaani ("Trp" tai "W"), tyrosiini ("Tyr" tai "Y"), ja valiini ("Val" tai "V"). Termi aminohappo tässä käytettynä käsittää myös proteiiniaminohappojen analogit ja proteiiniaminohappojen ja niiden analogien D-isomeerit.

20  
25  
30

Lasti - Molekyyli, joka ei ole tat-proteiini tai sen fragmentti ja joka joko 1) ei itsessään kykene menemään sisään kohdesoluihin tai 2) ei itsessään kykene menemään

sisään kohdesoluihin hyödyllisellä nopeudella. ("Lasti" tässä hakemuksessa käytettynä viittaa joko molekyyliin siinänsä, se on: ennen konjugointia, tai kuljetuspolypeptidilastikonjugaatin lastiosaan.) Esimerkkeihin "lastista" kuuluvat, näihin rajoittumatta, pienet molekyylit ja makromolekyylit, kuten polypeptidit, nukleiinihapot ja polysakkaridit.

Kemiallinen silloittaminen - Kahden tai useamman ennalta muodostetun molekyylin kovalenttinen sitominen.

10 Lastikonjugaatti - Molekyyli, joka käsittää ainakin yhden kuljetuspolypeptidiosan ja ainakin yhden lastiosan ja joka on muodostettu joko geneettisen fuusion tai kuljetuspolypeptidin tai lastimolekyylin kemiallisen silloittamisen avulla.

15 Geneettinen fuusio - Kahden tai useamman proteiinin kolineaarinen, kovalenttinen sidos niiden polypeptidirunkojen välillä, proteiineja koodittavien peräkkäisten sekvenssien geneettisen ekspression välityksellä.

20 Makromolekyyli - Molekyyli, kuten peptidi, polypeptidi, proteiini tai nukleiinihappo.

25 Polypeptidi - Mikä tahansa polymeeri, joka koostuu oleellisesti mistä tahansa 20 proteiiniamiinohaposta (edellä) sen koosta riippumatta. Vaikka usein käytetään "proteiinia" viittaamaan suhteellisen suuriin polypeptideihin, ja "peptidiä" käytetään usein viittaamaan pieniin polypeptideihin, näiden termien käyttö alalla limittyy ja vaihtelee. Termi "polypeptidi" tässä käytettynä viittaa peptideihin, polypeptideihin ja proteiineihin, ellei toisin ilmaista.

30 Reportterigeeni - Geeni, jonka ekspressio riippuu halutun solutason tapahtuman ilmenemisestä, ja jonka ekspressio voidaan havaita mukavasti geneettisesti transformoidussa isäntäsolussa.

Reportteriplasmidi - Plasmidivektori, joka käsittää yhden tai useamman reportterigeenin.

Pieni molekyyli - Molekyyli, joka on muu kuin makromolekyyli.

5 Väliosaa-aminohappo - Aminohappo (jolla on edullisesti pieni sivuketju), joka sisällytetään kuljetusosan ja kemiallisessa silloituksessa käytetyn aminohappotähteen väliin (esimerkiksi molekyylytason taipuisuuden tarjoamiseksi ja steerisen estämisen välttämiseksi).

10 Kohdesolu - Solu, johon lasti viedään kuljetuspoly-peptidillä. "Kohdesolu" voi olla mikä tahansa solu, mukaan lukien ihmissolut, joko in vivo tai in vitro.

Kuljetusosa tai kuljetuspolyeptidi - Polypeptidi, joka kykenee viemään kovalenttisesti kiinnittyneen lastin  
15 kohdesoluun.

Tämän keksinnön mukainen fuusioproteiini on yleisesti sovellettavissa sellaisten pienten molekyylien ja makromolekyylien terapeuttiseen, profylaktiseen tai diagnostiseen viemiseen solun sisälle, kuten proteiinit, nukleiinihapot ja polysakkaridit, jotka eivät itsessään kykene menemään kohdesolujen sisään hyödyllisellä nopeudella. Pitäisi kuitenkin ymmärtää, että tämän keksinnön mukaiset vaihtoehtoiset suoritusmuodot eivät rajoitu kliinisiin sovellutuksiin. Tätä keksintöä voidaan soveltaa edullisesti lääketieteellisessä ja biologisessa tutkimuksessa. Tämän keksinnön tutkimussovellutuksissa lasti voi olla lääkeaine tai reportterimolekyyli. Tämän keksinnön mukaisia kuljetuspolyeptidejä voidaan käyttää tutkimuslaboratorioreagensseina joko yksinään tai osana kuljetuspolyeptidin konjugaatiotarvikesarjaa.  
20  
25  
30

Kohdesolut voivat olla in vivo soluja, se on: soluja, jotka muodostavat elävien eläinten tai ihmisten elimiä tai kudoksia, tai mikro-organismeja, joita löytyy elävistä

eläimistä tai ihmisistä. Kohdesolut voivat myös olla in vitro soluja, se on: viljeltyjä eläinsoluja, ihmissoluja tai mikro-organismeja.

5 Tämän keksinnön mukaisen fuusioproteiinin eri käyttötarkoituksissa käytettävien lääkeaineiden ja reportterimolekyylien valitsemisessa on laaja liikkumavapaus. Tekijöihin, jotka on otettava huomioon reportterimolekyyliä valittaessa, kuuluvat, näihin rajoittumatta, etsityn korkeellisen informaation tyyppi, myrkyttömyys, detektion mukavuus, detektion kvantitoitavuus, ja saatavuus. Ammattimies tuntee monia tällaisia reportterimolekyyliä.

10 Kuten ymmärretään jäljempänä esitettyjen esimerkkien perusteella, olemme käyttäneet sellaisia entsyymejä, joille on olemassa kolorimetrisiä määrittämiä, mallilastina osoittaaksemme tämän keksinnön mukaisten kuljetuspolyeptidien toimivuuden ja hyödylliset piirteet. Nämä entsyymilastit tarjoavat solun sisään ottamisen herkän, mukavan, visuaalisen detektion. Lisäksi koska visuaalinen tulos ilmenee vain jos lastin entsyymaattinen aktiivisuus säilyy, nämä 15 entsyymit tarjoavat herkän ja luotettavan testin lastiosan biologisen aktiivisuuden säilymiselle tämän keksinnön mukaisissa kuljetuspolyeptidi-lastikonjugaateissa. Tämän keksinnön mukainen edullinen suoritusmuoto käsittää piparjuuriperoksidaasin ("HRP") kuljetuspolyeptidi-lastikonjugaatin lastiosana. Eriyisen edullinen lastiosamalli tämän 20 keksinnön suorittamista varten on  $\beta$ -galaktosidaasi.

Lastiproteiinimalleja voidaan valita myös niiden solunsisäisen vaikutuspaikan mukaan. Kuten esimerkeissä 6 ja 7 jäljempänä kuvataan, olemme käyttäneet Pseudomonas-eksotoksiinista ("PE") peräisin olevaa ADP-ribosylaatio- 30 osaa ja haiman ribonukleaasia vahvistaaksemme sen, että tämän keksinnön mukaiset kuljetuspolyeptidit ovat vieneet sytoplasmaan oikein laskostuneita lastiproteiineja.

Täyspitkä Pseudomonas-eksotoksiini on itsessään kykenevä menemään sisään soluihin, missä se inaktivoi ribosomeja ADP-ribosylaatioreaktion välityksellä ja näin tappaa solut. Osa Pseudomonas-eksotoksiiniproteiinista, joka tunnetaan ADP-ribosylaatio-osana, ei kykene menemään sisään soluihin, mutta siinä on jäljellä kyky inaktivoida ribosomit, jos se saatetaan yhteyteen niiden kanssa. Niinpä solukuolema, jonka kuljetuspolyeptidin ja PE-ADP-ribosylaatio-osan muodostamat konjugaatit saavat aikaan, on testi kuljetuspolyeptidin lastin viemiselle sytoplasmaan.

Olemme käyttäneet myös ribonukleaasia vahvistaaksemme sen, että tämän keksinnön mukaiset kuljetuspolyeptidit ovat vieneet sytoplasmaan oikein laskostuneen lastiproteiinin. Proteiinisynteesi, RNA:sta riippuvainen prosessi, on hyvin herkkä ribonukleaasille, joka digestoi RNA:ta. Ribonukleaasi ei itsessään kykene kuitenkaan menemään solujen sisään. Niinpä kuljetuspolyeptidin ja ribonukleaasin muodostaman konjugaatin aikaan saama proteiinisynteesin inhibitio on testi biologisesti aktiivisen ribonukleaasin viemiselle solun sisään.

Tietenkin tietyn lastimolekyylin viemistä sytoplasmaan voi seurata saman lastimolekyylin vieminen edelleen tumaan. Tumaan vieminen käsittää välttämättä sen, että läpäistään jokin osa sytoplasmosta.

Papilloomaviruksen E2-repressoriproteiinit ovat esimerkkejä makromolekyylilääkkeistä, jotka voidaan viedä kohdesolujen tumiin tämän keksinnön mukaisilla kuljetuspolyeptideillä. Papilloomaviruksen E2-proteiini, joka esiintyy normaalisti homodimeerinä, säätelee sekä papilloomavirusgenomin transkriptiota että replikaatiota. E2-proteiinin karboksiterminaalinen osa sisältää DNA:han sitoutumis- ja dimerisaatio-aktiivisuudet. Sellaisten DNA-sekvenssien lyhytkestoinen ekspressio transfektoiduissa nisäkässoluissa,

jotka koodittavat erilaisia E2-analogeja tai E2:n karboksi-terminaalaisia fragmentteja, inhiboi täyspitkän E2-proteiinin aikaansaamaa transaktivaatiota (J. Barsoum ym., "Mechanism of Action of the Papillomavirus E2 Repressor: 5 Repression in the Absence of DNA Binding", J. Virol., 66, s. 3941 - 3945 (1992)). E2-repressorit, jotka on lisätty viljeltyjen nisäkässolujen kasvatuselatusaineeseen, eivät mene solujen sisään, ja niinpä ne eivät inhiboi E2:n transaktivaatiota näissä soluissa. Tämän keksinnön mukaisten 10 kuljetuspolypeptidien konjugointi E2-repressoreihin saa kuitenkin aikaan E2-repressorien translokaation kasvatuselatusaineesta viljeltyihin soluihin, missä ne ilmentävät biologista aktiivisuutta ja repressoivat reportterigeenin E2-riippuvaista ekspressiota.

15 Nopeutta, jolla yksi- ja kaksijuosteiset nukleiinihapot menevät solujen sisään in vitro ja in vivo, voidaan tehostaa edullisesti käyttämällä tämän keksinnön mukaisia kuljetuspolypeptidejä. Kuten esimerkissä 11 (jäljempänä) osoitetaan, menetelmät polypeptidien silloittamiseksi kemiallisesti nukleiinihappoihin ovat sinänsä tunnettuja. Tämän 20 keksinnön edullisessa suoritusmuodossa lasti on yksijuosteinen "antisense"-nukleiinihappo. "Antisense"-nukleiinihapot ovat hyödyllisiä inhiboimaan sellaisten sekvenssien solunsisäistä ekspressiota, joille ne ovat komplementaarisia. 25 Eräässä toisessa tämän keksinnön suoritusmuodossa lasti on kaksijuosteinen nukleiinihappo, joka käsittää sitoutumiskohdan, jonka nukleiinihappoon sitoutuva proteiini tunnistaa. Esimerkki tällaisesta nukleiinihappoon sitoutuvasta proteiinista on virusperäinen transaktivaattori.

30 Luonnossa ilmenevällä HIV-1:n tat-proteiinilla (kuvio 1) on alue (aminohapot 22 - 37), jossa 7 16:sta aminohaposta on kysteiini. Nämä kysteinitähteet kykenevät muodostamaan disulfididisidoksia keskenään, toisten tat-proteiini-

nimolekyylien kysteiniipitoisella alueella olevien kyste-  
 iinitähteiden kanssa, ja lastiproteiinissa tai konjugaatin  
 lastiosassa olevien kysteiniitähteiden kanssa. Sellainen  
 disulfididisidoksen muodostuminen voi saada aikaan lastin  
 biologisen aktiivisuuden katoamisen. Lisäksi vaikka ei oli-  
 sikaan mahdollisuutta disulfididisidoksen muodostumiseen las-  
 tiosan kanssa (esimerkiksi kun lastiproteiinissa ei ole  
 kysteiniitähteitä), disulfididisidoksen muodostuminen kulje-  
 tuspolypeptidien välillä johtaa kuljetuspolypeptidin, kul-  
 jetuspolypeptidi-lastikonjugaatin tai molempien aggregoitum-  
 miseen ja liukenemattomuuteen. Tat:n kysteiniipitoinen alue  
 on potentiaalisesti vakavien ongelmien lähde käytettäessä  
 luonnossa ilmenevää tat-proteiinia lastimolekyylien viemi-  
 seen soluun.

Kysteiniipitoista aluetta tarvitaan tat:n dimeri-  
 soitumiseen in vitro, ja se vaaditaan HIV:n DNA-sekvenssien  
 transaktivaatioon. Niinpä tat:n kysteiniipitoisen alueen  
 poistamisella on se lisäetu, että tat:n luonnollinen aktii-  
 visuus poistuu, se on: HIV:n transkription ja replikaation  
 induktio. Alalla ei kuitenkaan esitetä tarvitaanko tat-pro-  
 teiinin kysteiniipitoista aluetta solun sisään ottamiseen.

Tämä keksintö käsittää suoritusmuotoja, joissa on  
 ratkaistu tat:n kysteiniipitoiseen alueeseen liittyvät on-  
 gelmat, koska tämä alue ei ole läsnä tässä kuvatuissa kul-  
 jetuspolypeptideissä. Näissä suoritusmuodoissa kuljetuspo-  
 lypeptidin tai kuljetuspolypeptidi-lastikonjugaatin otto  
 solun sisään tapahtuu yhä. Yhdessä tämän keksinnön edullis-  
 ten suoritusmuotojen ryhmässä aminohapposekvenssi, joka  
 edeltää kysteiniipitoista aluetta, fuusioidaan suoraan sii-  
 hen aminohapposekvenssiin, joka seuraa kysteiniipitoista  
 aluetta. Sellaisille kuljetuspolypeptideille on annettu ni-  
 meksi tatAcys ja niillä on yleinen kaava (tat1-21)-(tat38-  
 n), jossa n on karboksiterminaalisen tähteen numero, se on:



49 - 86. Edullisesti n on 58 - 72. Kuten jäljempänä olevis-  
ta esimerkeistä käy ilmi, tat-proteiinin kysteiniipitoista  
aluetta edeltävää aminohapposekvenssiä ei tarvita ottoon  
solun sisään. Edullinen kuljetuspolyeptidi (tai kuljetus-  
5 osa) koostuu tat-proteiinin aminohapoista 37 - 72, ja sille  
on annettu nimeksi tat37-72 (SEQ ID nro 2). Tat-tähteen 37,  
kysteiniin, säilyttäminen kuljetuspolyeptidin aminopäässä  
on edullista, koska se on käyttökelpoinen kemiallisessa  
silloittamisessa.

10 Tat $\Delta$ cys-polyeptidien, tat37-72:n ja muiden tämän  
keksinnön suoritusmuotojen etuihin kuuluvat seuraavat:

a) tat-proteiinin luonnollinen aktiivisuus, se on:  
HIV-transkription induktio, eliminoituu,

b) dimeerit ja muut kuljetuspolyeptidin korkeammat  
15 multimeerit vältetään,

c) tat $\Delta$ cys-geneettisten fuusioiden ekspressiotaso E.  
colissa voi parantua,

d) joillakin kuljetuspolyeptidikonjugaateilla il-  
menee lisääntynyttä liukoisuutta ja ylivertaista käsittelyn  
20 helppoutta ja

e) joillakin fuusioproteiineilla ilmenee lisäänty-  
nyttä aktiivisuutta lastiosassa, verrattuna fuusioihin,  
jotka sisältävät kysteiniirikkaan osan.

25 Monia kemiallisia silloitusmenetelmiä tunnetaan ja  
ne ovat mahdollisesti käyttökelpoisia, kun halutaan konju-  
goida tämän keksinnön mukaisia kuljetuspolyeptidejä lasti-  
makromolekyyleihin. Monet tunnetut kemialliset silloitusme-  
netelmät ovat epäspesifisiä, se on: ne eivät suuntaa lii-  
toskohtaa mihinkään tiettyyn kohtaan kuljetuspolyeptidissä  
30 tai lastimakromolekyylissä. Tuloksena tästä epäspesifisten  
silloitusaineiden käyttö voi hyökätä funktionaalsiin koh-  
tiin tai tukkia steerisesti aktiivisia kohtia, mikä tekee  
konjugoiduista proteiineista biologisesti inaktiivisia.

Edullinen lähestymistapa, jolla liittämisen spesifisyyttä voidaan lisätä tätä keksintöä suoritettaessa, on suunnata kemiallinen liittäminen funktionaaliseen ryhmään, joka havaitaan vain kerran tai muutaman kerran yhdessä tai  
5 molemmissa silloitettavista polypeptideistä. Esimerkiksi monissa proteiineissa kysteini, joka on ainoa tioliryhmän sisältävä proteiiniaminohappo, esiintyy vain joitakin kertoja. Samoin esimerkiksi jos polypeptidi ei sisällä lysiinitähteitä, silloitusreagenssi, joka on spesifinen primäärisille amiineille, on selektiivinen kyseisen polypeptidin aminopäälle. Tämän lähestymistavan menestykseks hyödyntäminen liittämsspesifisyyden parantamiseksi vaatii sen, että polypeptideillä on sopivan harvinaiset ja reaktiiviset aminohappotähteet sellaisilla molekyylin alueilla, joita  
10 voidaan muuttaa ilman että molekyylin biologinen aktiivisuus katoaa.

Kuten jäljempänä olevissa esimerkeissä osoitetaan, kysteinitähteet voidaan korvata, kun ne ilmenevät sellaisissa polypeptidisekvenssin osissa, joissa niiden osallistuminen silloitusreaktioon todennäköisesti häiritsisi biologista aktiivisuutta. Kun kysteinitähde korvataan, on  
20 tyypillisesti suotavaa minimoida tuloksena olevat muutokset polypeptidin laskostumisessa. Muutokset polypeptidin laskostumisessa minimoituvat, kun korvike on kemiallisesti ja steerisesti samanlainen kysteiinin kanssa. Näistä syistä seriini on edullinen kysteiinin korvike. Kuten jäljempänä olevissa esimerkeissä osoitetaan, kysteinitähde voidaan tuoda polypeptidin aminohapposekvenssiin silloitustarkoituksissa. Kun tuodaan kysteinitähde, tuominen amino- tai  
25 karboksipäähän tai sen lähelle on edullista. Sellaisia aminohapposekvenssimuutoksia varten on saatavilla tavanomaisia menetelmiä, tuotettiinpa haluttu polypeptidi kemiallisella synteesillä tai yhdistelmä-DNA:n ekspressiolla.  
30

Silloitusreagenssit voivat olla homobifunktionaalisia, se on: joilla on kaksi funktionaalista ryhmää, jotka käyvät läpi saman reaktion. Edullinen homobifunktionaalinen silloitusreagenssi on bismaleimidoheksaani ("BMH"). BMH sisältää kaksi funktionaalista maleimidiryhmää, jotka reagoivat spesifisesti sulfhydryylin sisältävien yhdisteiden kanssa hellävaraisissa olosuhteissa (pH 6,5 - 7,7). Nuo kaksi maleimidiryhmää on yhdistetty toisiinsa hiilivetyketjulla. Niinpä BMH on hyödyllinen sellaisten polypeptidien irreversiibeliä silloittamista varten, jotka sisältävät kysteiniinitähteitä.

Silloitusreagenssit voivat olla myös heterobifunktionaalisia. Heterobifunktionaalisilla silloitusaineilla on kaksi erilaista funktionaalista ryhmää, esimerkiksi amiini-reaktiivinen ryhmä ja tiolireaktiivinen ryhmä, jotka silloittavat kaksi proteiinia, joilla on vapaita amiineja ja vastaavasti tioleja. Esimerkkejä heterobifunktionaalisista silloitusaineista ovat sukkinimidyyli-4-(N-maleimidometyyli)sykloheksaani-1-karboksylaatti ("SMCC"), m-maleimido-bentsoyyli-N-hydroksisukkinimidiesteri ("MBS"), ja sukkinimidi-4-(p-maleimidofenyyli)butyraatti ("SMPB"), MBS:n pitempiketjuinen analogi. Näiden silloitusaineiden sukkinimidyyli-ryhmä reagoi primäärisen amiinin kanssa ja tiolireaktiivinen maleimidi muodostaa kovalenttisen sidoksen kysteiniinitähteen tiolin kanssa.

Silloitusreagensseilla on usein huono liukoisuus veteen. Silloitusreagenssiin voidaan lisätä hydrofiilinen osa, kuten sulfonaattiryhmä, sen vesiliukoisuuden parantamiseksi. Sulfo-MBS ja sulfo-SMCC ovat esimerkkejä silloitusreagensseista, joita on modifioitu vesiliukoisuutta silloittamalla pitäen.

Monet silloitusreagenssit saavat aikaan konjugaa-  
 tin, joka on oleellisesti katkeamaton solun olosuhteissa.  
 Jotkin silloitusreagenssit sisältävät kuitenkin kovalentti-  
 sen sidoksen, kuten disulfidin, joka voidaan katkaista so-  
 5 lun olosuhteissa. Esimerkiksi ditiobis(sukkinimidyylipro-  
 pionaatti) ("DSP"), Trautin reagenssi ja N-sukkinimidyyli-  
 3-(2-pyridyyliditio)propionaatti ("SPDP") ovat hyvin tun-  
 nettuja katkaistavissa olevia silloitusaineita. Katkaista-  
 vissa olevan silloitusreagenssin käyttö mahdollistaa sen,  
 10 että lastiosa erotetaan kuljetuspolypeptidistä kohdesoluun  
 viemisen jälkeen. Suora disulfidiliitos voi olla myös hyö-  
 dyllinen.

Joillakin uusilla silloitusreagensseilla, kuten n-  
 y-maleimidobutyryylioksi-sukkinimidiesteri ("GMBS") ja sul-  
 15 fo-GMBS, on vähentynyt immunogeenisyys. Joissakin tämän  
 keksinnön mukaisissa suoritusmuodoissa tällainen vähentynyt  
 immunogeenisyys voi olla edullinen.

Monia silloitusreagensseja, mukaan lukien edellä  
 käsitellyt, on kaupallisesti saatavilla. Yksityiskohtaiset  
 20 ohjeet niiden käyttöä varten ovat helposti saatavilla kau-  
 pallisilta toimittajilta. Yleinen lähde teos proteiinien  
 silloittamisesta ja konjugaattien valmistamisesta on S. S.  
 Wong, Chemistry of Protein Conjugation and Cross-Linking,  
 CRC Press (1991).

Kemialliseen silloittamiseen voi sisältyä väliosien  
 käyttö. Väliosat tarjoavat molekyylin sisäistä joustoa tai  
 ne säätävät molekyylin sisäisiä etäisyyksiä konjugoitujen  
 osien välillä ja näin ollen voivat auttaa biologisen aktii-  
 25 visuuden säilyttämisessä. Väliosia voi olla polypeptidiosan  
 muodossa, joka käsittää väliosia-aminohappoja. Väli-ryhmä voi  
 30 vaihtoehtoisesti olla osa silloitusreagenssia, kuten "pit-  
 kaketjuisessa SPDP:ssä" (Pierce Chem. Co., Rockford, IL,  
 luettelonro 21651 H).

Kuten seuraavissa esimerkeissä osoitetaan käyttämällä tässä tarjottua aminohappo- ja DNA-sekvenssi-informaatiota, tämän keksinnön mukaiset kuljetuspolypeptidit voidaan syntetisoida kemiallisesti tai tuottaa yhdistelmä-DNA-menetelmillä. Menetelmät, joilla aminohapposekvenssiltään tunnettuja polypeptidejä voidaan tuottaa kemiallisella synteisillä tai yhdistelmä-DNA-menetelmillä, ovat hyvin tunnettuja. Automaattilaitteet polypeptidi- tai DNA-synteesiä varten ovat kaupallisesti saatavilla. Isäntäsolut, kloonausvektorit, DNA-ekspressiokontrollisekvenssit ja oligonukleotidilinkkerit ovat myös kaupallisesti saatavilla.

Tässä esitettyihin edullisiin kuljetuspolypeptidien aminohapposekvensseihin voidaan helposti tehdä pieniä lisäyksiä, deletioita tai substituutioita sinänsä tunnetulla tavalla. Pitäisi kuitenkin ymmärtää, että sellaiset variaatiot ovat tämän keksinnön suojapiirin sisällä.

Lisäksi tunnetaan tat-proteiineja muista viruksista, kuten HIV-2 (M. Guyader ym., "Genome Organization and Transactivation of the Human Immunodeficiency Virus Type 2", Nature, 326, s. 662 - 669 (1987)), hevosen tarttuva anemia virus (R. Carroll ym., "Identification of Lentivirus Tat Functional Domains Through Generation of Equine Infectious Anemia Virus/Human Immunodeficiency Virus Type 1 tat Gene Chimeras", J. Virol., 65, s. 3460 - 3467 (1991)), ja apinan immuunipuutosvirus (L. Chakrabarti ym., "Sequence of Simian Immunodeficiency Virus from Macaque and Its Relationship to Other Human and Simian Retroviruses", Nature, 328, s. 543 - 547 (1987); S. K. Arya ym., "New Human and Simian HIV-Related Retroviruses Possess Functional Transactivator (tat) Gene", Nature 328, s. 548 - 550 (1987)). Pitäisi ymmärtää, että polypeptidit, jotka ovat peräisin näistä tat-proteiineista ja joiden piirteenä on tat:n emäk-

sisen alueen läsnäolo ja tat:n kysteiinipitoisen alueen poissaolo, kuuluvat tämän keksinnön suojapiiriin.

Jotta tässä kuvattu keksintö olisi mahdollista ymmärtää täydellisemmin, esitetään seuraavat esimerkit. Pitäisi ymmärtää, että nämä esimerkit ovat pelkästään havainnollistamistarkoituksia varten ja niiden ei ole tarkoitus rajoittaa tätä keksintöä millään tavalla. Kaikkien näiden esimerkkien yhteydessä kaikki molekyylikloonausreaktiot suoritettiin J. Sambrookin ym. teoksen mukaisesti, Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory (1989), paitsi missä toisin mainitaan.

#### **Esimerkki 1**

#### **Kuljetuspolypeptidien tuotanto ja puhdistus**

#### **Yhdistelmä-DNA**

15 Plasmidi pTat72 oli lähtökloonin tat-peräisten kuljetuspolypeptidien tuottamiseksi bakteereissa ja sellaisten geenien muodostamiseksi, jotka koodittavat kuljetuspolypeptidin ja lastiproteiinin fuusioita. Hankimme plasmidin pTat72 (kuvattu Frankelin ja Pabon artikkelissa edellä) Alan Frankelilta (The Whitehead Institute for Biomedical Research, Cambridge, MA). Plasmidi pTat72 oli peräisin F. W. Studierin ym. pET-3a-ekspressiovektorista ("Use of T7 RNA Polymerase to Direct Expression of Cloned Genes", Methods Enzymol., 185, s. 60 - 90 (1990)) sijoittamalla syntetttinen geeni, joka kooditti HIV-1:n tat:n aminohappoja 1 - 72. Tat:n koodittava alue käyttää E. colin kodoninkäyttöä ja sitä kontrolloi bakteriofaagi T7:n polymeraasin promoottori, joka on indusoitavissa isopropyli-beta-D-tiogalaktopyranosidilla ("IPTG"). Tat-proteiini käsitti 5 % E. colin kokonaisproteiinista IPTG-induktion jälkeen.

### Tat1-72:n puhdistus bakteereista

Suspensioimme E. colin, joka ekspressoi tat1-72-proteiinia, 10-kertaiseen tilavuuteen 25-millimolaarista Tris-HCl:ää (pH 7,5), joka oli 1 mM EDTA:n suhteen. Hajotimme solut "French press"-laitteessa ja poistimme liukenemattomat roskat sentrifugoimalla 10 000 g:n kiihtyvyydessä 1 tunnin ajan. Latasimme supernatantin Q Sepharose Fast Flow -ioninvaihtopylvääseen (Pharmacia LKB, Piscataway, NJ) (20 ml hartsia/60 ml lyaattia). Käsittelimme läpi tulleen fraktion 0,5 M NaCl:lla, mikä aiheutti tat-proteiinin saostumisen. Keräsimme suolalla saostetun proteiinin sentrifugoimalla 35 000 g:n kiihtyvyydellä 50.2-roottorissa 1 tunnin ajan. Liuotimme pohjaan ajetun sakan 6-molaariseen guanidiini-HCl:ään ja kirkastimme liuoksen sentrifugoimalla 35 000 rpm:ssä, 50.2-roottorissa 1 tunnin ajan. Latasimme kirkastetun näytteen A.5-agarosigeelisuodatuspylvääseen, joka oli tasapainotettu 6-molaarisella guanidiini-HCl:llä, joka oli 50-millimolaarinen natriumfosfaatin suhteen (pH 5,4) ja 10-millimolaarinen DTT:n suhteen, ja eluimme sitten näytteen samalla puskurilla. Latasimme tat-proteiinin sisältävät geelisuodatusfraktiot C<sub>4</sub>-käänteisfaasi-HPLC-pylvääseen ja eluimme gradientilla, jossa oli 0,1-%:isessa trifluorietikkahapossa asetonitriiliä 0 - 75 %. Tällä menetelmällä tuotimme noin 20 mg tat1-72-proteiinia per litra E. coli -viljelmää (olettaen että soluja oli 6 g per litra). Tämä edusti noin 50-%:ista kokonaissaantoa.

SDS-PAGE-analyysissä tat1-72-polyeptidi liikkui noin 10 kDa:n yksittäisenä juovana. Puhdistettu tat1-72-polyeptidi oli aktiivinen sisäänotto/transaktivaatiomäärityksessä. Lisäsimme polypeptidin sellaisten ihmisen hepatomasolujen viljelmäelatusaineeseen, jotka sisälsivät tat-responsiivisen kudospasminogeeniaktivaattori ("tPA") -reportterigeenin. Klorokiinin ollessa läsnä (0,1 mM), puhdis-

tettu tat1-72-proteiini (100 ng/ml) indusoi tPA:n ekspres-  
sion noin 150-kertaiseksi.

### **Kuljetuspolypeptidien kemiallinen synteesi**

5 Monien erilaisten kuljetuspolypeptidien kemiallista  
synteesiä varten käytimme kemiallisesti saatavilla olevaa  
automatisoitua järjestelmää (Applied Biosystems Model 430A  
-syntetisaattoria) ja noudatimme järjestelmän valmistajan  
suosittelemia menetelmiä. Poistimme suojaryhmät HF-käsit-  
telyllä ja eristimme synteettiset polypeptidit tavanomai-  
10 silla käänteisfaasi-HPLC-menetelmillä. Kaikkien synteettis-  
ten polypeptidien eheys varmistettiin massaspektrometriana-  
lyysillä.

#### **Esimerkki 2**

#### **$\beta$ -galaktosidaasikonjugaatit**

#### **15 Kemiallinen silloitus SMCC:llä**

$\beta$ -galaktosidaasin asetylointia varten (kysteiini-  
sulfhydryyliryhmien suojaamiseksi) liuotimme 6,4 mg kauppal-  
lisesti saatavilla olevaa  $\beta$ -galaktosidaasia (Pierce Chem.  
Co., luettelonro 32101G) 200  $\mu$ l:aan 50-millimolaarista fos-  
20 faattipuskuria (pH 7,5). Lisäsimme 10  $\mu$ l jodietikkahappoa,  
joka valmistettiin liuottamalla 30 mg jodietikkahappoa 4  
ml:aan 50-millimolaarista fosfaattipuskuria (pH 7,5), 200  
 $\mu$ l:aan  $\beta$ -galaktosidaasiluosta. (Seuraavissa kokeissa ha-  
vaitsimme jodiasetamidin olevan edullinen korvike jodietik-  
25 kahapolle.) Annoimme reaktion edetä 60 minuutin ajan hu-  
oneenlämpötilassa. Sitten erotimme asetyloidun  $\beta$ -galaktosi-  
daasin reagoimattomasta jodietikkahaposta lataamalla reak-  
tioseoksen (Pharmacia) pieneen G-25-geelisuodatuspylvääseen  
(Pharmacia LKB, Piscataway, NJ) ja keräämällä välisijatila-  
30 vuuden.

Ennen asetyloidun  $\beta$ -galaktosidaasin amiiniryhmien  
SMCC-aktivaatiota konsentroimme 2 ml G-25-pylväästä kerät-  
tyä entsyymiä 0,3 ml:ksi Centricon 10 -ultrasuodatuslait-



teessa (Amicon, Danvers, MA). Lisäsimme konsentroituu ase-  
tyloituun  $\beta$ -galaktosidaasiin 19  $\mu$ g sulfo-SMCC:tä (Pierce  
Chem. Co., luettelonro 22322G), joka oli liuotettu 15  $\mu$ l:aan  
5 dimetyyliformamidia ("DMF"). Annoimme reaktion edetä 30 mi-  
nuuttia huoneenlämpötilassa. Erotimme sitten  $\beta$ -galaktosi-  
daasi-SMCC:n DMF:stä ja reagoimattomasta SMCC:stä viemällä  
sen pienen G-25-geelisuodatuspylvään läpi.

Jotta saisimme silloitettua kuljetuspolypeptidit  
kemiallisesti  $\beta$ -galaktosidaasiin, sekoitimme  $\beta$ -galaktosi-  
10 daasi-SMCC-liuoksen 100  $\mu$ g:aan kuljetuspolypeptidiä (tat1-  
72, tat37-72, tat38-58GGC, tat37-58, tat47-58GGC tai  
tatCGG47-58), joka oli liuotettu 200  $\mu$ l:aan 50-millimolaa-  
rista fosfaattipuskuria (pH 7,5). Annoimme reaktion edetä  
60 minuuttia huoneenlämpötilassa. Eristimme sitten kulje-  
15 tuspolypeptidi- $\beta$ -galaktosidaasikonjugaatin lataamalla reak-  
tioseoksen S-200HR-geelisuodatuspylvääseen ja keräämällä  
välisijatilavuuden.

Näin hankittu kuljetuspolypeptidi- $\beta$ -galaktosidaasi-  
konjugaatti antoi positiivisia tuloksia, kun siitä määri-  
20 tettiin tat tavanomaisissa Western blot- ja ELISA-ana-  
lyyseissä, jotka suoritettiin tat:ia vastaan suunnatuilla  
kanin polyklonaalisilla vasta-aineilla. Western blot- ja  
ELISA-analyysistä yleisen esityksen voi löytää E. Harlowin  
ja D. Lanen teoksesta Antibodies: A Laboratory Manual, Cold  
25 Spring Harbor Laboratory (1988). Geelisuodatusanalyysi Su-  
perose 6:lla (Pharmacia LKB, Piscataway, NJ) osoitti, että  
kuljetuspolypeptidi- $\beta$ -galaktosidaasikonjugaatin molekyyli-  
paino oli noin 540 000 daltonia. Kuljetuspolypeptidi- $\beta$ -  
galaktosidaasikonjugaatin spesifinen aktiivisuus oli 52 %  
30  $\beta$ -galaktosidaasilähtömateriaalin spesifisestä aktiivisuu-  
desta, kun se määritettiin o-nitrofenyyli- $\beta$ -D-galaktopyra-  
nosidilla ("ONPG"). ONPG-määritysmenetelmä kuvataan yksi-

tyiskohtaisesti Sambrookin ym. teoksessa (edellä) sivuilla 16.66 - 16.67.

### **$\beta$ -galaktosidaasikonjugaattien otto solujen sisään**

Lisäsimme konjugaatit HeLa-solujen (ATCC nro CCL2) elatusaineeseen 100-mikromolaarisen klorokiinin ollessa 5 elänsä tai poissa. Inkuboimme soluja 4 - 18 tuntia 37 °C:n lämpötilassa/5,5 % CO<sub>2</sub>. Kestävöimme solut fosfaattipuskuroidussa suolaliuoksessa ("PBS"), joka oli 2-%:inen formaldehydin ja 0,2-%:inen glutaraldehydin suhteen, 5 minuutin 10 ajan 4 °C:n lämpötilassa. Pesimme sitten solut kolme kertaa 2-millimolaarisella MgCl<sub>2</sub>:lla PBS:ssä ja värjäsimme ne X-Gal:lla 37 °C:n lämpötilassa. X-gal on väritön  $\beta$ -galaktosidaasin substraatti (5-bromi-4-kloro-3-indolyyli-D-galaktosidi), joka tuottaa sinisen tuotteen, kun  $\beta$ -galaktosidaasi 15 katkaisee sen. Meidän X-gal-värjäysliuoksemme sisälsi 1 mg:n X-gal:ia (Bio-Rad, Richmond, CA, luettelonumero 170-3455) per ml PBS:ää, joka oli 5-millimolaarinen kaliumferrisyanidin, 5-millimolaarinen kaliumferrosyanidin ja 2-millimolaarinen MgCl<sub>2</sub>:n suhteen.

20 Tutkimme värjätyt solut mikroskoopilla jopa 400-kertaisilla suurennoksilla. Sellainen mikroskooppinen tutkiminen paljasti tuman värjäytymisen, samoin kuin sytoplasman värjäytymisen.

25 Solut, joihin tat37-72- $\beta$ -galaktosidaasikonjugaatti tai tat1-72- $\beta$ -galaktosidaasikonjugaatti lisättiin, värjäytyivät tummansinisiksi.  $\beta$ -galaktosidaasiaktiivisuutta voitiin havaita niinkin lyhyen kehitysajan kuin 15 minuutin jälkeen. Vertailun vuoksi pitäisi huomata, että ainakin 6 tunnin pituinen värinkehitysaika vaaditaan normaalisti, kun 30  $\beta$ -galaktosidaasiaktiivisuus viedään soluihin  $\beta$ -galaktosidaasigeenin transfektion avulla. Tuman värjäytyminen oli näkyvissä klorokiinin poissa ollessa, vaikkakin tuman värjäytymisen intensiteetti oli hieman suurempi klorokiinillä

käsitellyissä soluissa. Kontrollisoluissa, jotka oli käsitelty konjugoitumattomalla  $\beta$ -galaktosidaasilla, ei ollut havaittavissa värjäytymistä.

**Katkaistavissa oleva konjugointi suoralla disulfidilla**

5

Kullakin  $\beta$ -galaktosidaasitetrameerilla on 12 kysteiinitähdettä, joita voidaan käyttää suoraan disulfididisidokseen kuljetuspolyeptidin kysteiinitähteeseen. Tat37-72:n sulfhydryylin pelkistämiseksi ja sitten suojaamiseksi liuotimme 1,8 mg (411 nanomoolia) tat37-72:ta 1 ml:aan 50-millimolaarista natriumfosfaattia (pH 8,0), joka oli 150-millimolaarinen NaCl:n, 2-millimolaarinen EDTA:n suhteen ja panimme liuoksen Reduce-Imm-pylvääseen (Pierce Chem. Co., Rockford, IL). 30 minuutin kuluttua huoneenlämmössä eluimme tat37-72:n pylväästä 1 ml:n erillä samaa puskuria, putkiin jotka sisälsivät 0,1 ml 10-millimolaarista 5,5'-ditio-bis(2-nitrobensoehappoa) ("DTNB"). Jätimme pelkistyneen tat37-72-polyeptidin DTNB:n luo 3 tunniksi. Poistimme sitten reagoimattoman DTNB:n tat37-72-TNB:stä geelisuodatuksella 9 ml:n Sephadex G 10 -pylväässä (Pharmacia LKB, Piscataway, NJ). Liuotimme 5 mg  $\beta$ -galaktosidaasia 0,5 ml:aan puskuria ja poistimme siitä suolat 9 ml:n Sephadex G-25 -pylväässä (Pharmacia LKB, Piscataway, NJ), jolloin saimme 3,8 mg  $\beta$ -galaktosidaasia/ml puskuria. Sekoitimme 0,5 ml:n erät  $\beta$ -galaktosidaasiliuosta, josta suolat oli poistettu, 0,25 tai 0,5 ml:aan tat37-72-TNB-valmistetta, ja annoimme suoran disulfidisilloitusreaktion jatkua huoneenlämpötilassa 30 minuutin ajan. Poistimme reagoimattoman tat37-72-TNB:n  $\beta$ -galaktosidaasikonjugaatista geelisuodatuksella 9 ml:n Sephacryl S-200 -pylväässä. Monitoroimme silloitusreaktion laajuutta epäsuorasti mittaamalla absorbanssia 412 nm:n aallonpituudella, joka johtui vapautuneesta TNB:stä.

10

15

20

25

30

Näin tuotetut suorat disulfidikonjugaatit otettiin soluihin (tuloksia ei esitetä).

**Katkaistavissa oleva konjugointi SPDP:llä**

Käytimme heterobifunktionaalista silloitusreagenssia ("SPDP"), joka sisältää katkaistavissa olevan disulfidisidoksen, muodostamaan silloituksen seuraavien välille: 5  
1)  $\beta$ -galaktosidaasin primääriset amiiniryhmät ja tat1-72:n kysteiinisulfhydryylit (metabolisesti leimattu  $^{35}\text{S}$ :llä), tai 2) rodamiinileimatun  $\beta$ -galaktosidaasin primääriset 10  
amiiniryhmät ja tat37-72:n aminoterminaalinen kysteiinisulfhydryyli.

Tat1-72-konjugaatiota varten liuotimme 5 mg  $\beta$ -galaktosidaasia 0,5 ml:aan 50-millimolaarista natriumfosfaattia (pH 7,5), joka oli 150 mM NaCl:n ja 2 mM  $\text{MgCl}_2$ :n suhteen, ja poistimme suolat  $\beta$ -galaktosidaasista 9 ml:n Sephadex G-25-pylväessä (Pharmacia LKB, Piscataway, NJ). Käsittelimme  $\beta$ -galaktosidaasin, josta suolat oli poistettu, 88-kertaisella molaarisella ylimäärällä jodiasetamidia huoneenlämmössä 2 tunnin ajan vapaiden sulfhydryyliryhmien tukkimiseksi. Sen jälkeen kun reagoimaton jodiasetamidi oli poistettu geelisuodatuksella, käsittelimme suojatun  $\beta$ -galaktosidaasin 10-kertaisella molaarisella ylimäärällä SPDP:tä huoneenlämpötilassa. 2 tunnin kuluttua vaihdoimme puskurin ultrasuodatuksella (Ultrafree 30, Millipore, Bedford, MA). Lisäsimme sitten 4-kertaisen molaarisen ylimäärän leimattua tat1-72:ta ja annoimme silloitusreaktion edetä yli yön huoneenlämpötilassa. Poistimme reagoimattoman tat1-72:n geelisuodatuksella 9 ml:n Sephacryl S-200-pylväessä. Leimatun tat1-72:n tunnettua spesifistä aktiivisuutta käyttämällä laskimme, että yhtä  $\beta$ -galaktosidaasitetrameeriä kohti oli 1,1 silloitettua tat1-72-polypeptidiä. ONPG-määrittystä käyttämällä havaitsimme, että konjugoitu  $\beta$ -galaktosidaasi oli säilyttänyt 100 % entsyymaatti-

sesta aktiivisuudestaan. Käyttämällä soluun inkorporoituneen radioaktiivisuuden mittausta ja X-gal-värjäystä, osoitimme että konjugaatti oli otettu sisään viljeltyihin HeLa-soluihin.

5           Tat37-72-konjugaatiota varten menettelimme aikaisemmassa kappaleessa kuvatulla tavalla, paitsi että leimasimme  $\beta$ -galaktosidaasin rodamiinimaleimidilla (molaarinen suhde 5:1) huoneenlämpötilassa 1 tunnin ajan, ennen jodiasetamidikäsittelyä (100:1 jodiasetamidin molaarinen ylimäärä). Silloitusreaktiossa käytimme SPDP-suhdetta 20:1 ja  
10           tat37-72-suhdetta 10:1. Arvioimme että konjugoidulla tuotteella olisi noin 5 rodamiinimolekyyliä (UV-absorbanssin perusteella) ja noin 2 tat37-72-osaa (geelifiltraation perusteella) per  $\beta$ -galaktosidaasitetrameeri. Tästä menettelystä saadulla konjugaatilla oli säilynyt noin 35 %  $\beta$ -  
15           galaktosidaasin alkuperäisestä entsyymaattisesta aktiivisuudesta. Osoitimme X-gal-värjäystä ja rodamiinifluoresenssia käyttämällä, että SPDP-konjugaatti otettiin sisään viljeltyihin HeLa-soluihin.

20

### **Esimerkki 3**

#### **Eläintutkimukset $\beta$ -galaktosidaasikonjugaateilla**

Konjugaatin puoliintumisajan määrittämistä ja biologisen jakautumisen analysointia varten injektoimme joko 200  $\mu$ g SMCC- $\beta$ -galaktosidaasia (kontrolli) tai tat1-72- $\beta$ -  
25           galaktosidaasia suonensisäisesti ("IV") Balb/c-hiirten (Jackson Laboratories) häntälaskimoihin, joko klorokiinin kanssa tai ilman. Keräsimme verinäytteitä tietyin välein 30 minuutin ajan. 30 minuutin kuluttua uhrasimme eläimet ja poistimme elimet ja kudokset histokemiallista analyysiä  
30           varten.

Mittasimme  $\beta$ -galaktosidaasin aktiivisuuden verinäytteissä ONPG-määrityksellä. ONPG-määritysmenetelmä kuvataan yksityiskohtaisesti Sambrookin ym. teoksessa (edellä)

sivuilla 16.66 - 16.67.  $\beta$ -galaktosidaasi ja tat1-72- $\beta$ -galaktosidaasi poistuivat nopeasti verenkierrosta. Arvioimme että niiden puoliintumisajat olivat 3 - 6 minuuttia. Nämä kokeelliset vertailut viittasivat siihen, että tat1-72-  
 5 kuljetuspolyeptidin kiinnittymisellä on vähän tai ei lainkaan vaikutusta  $\beta$ -galaktosidaasin poistumisnopeuteen verestä.

Jotta voisimme havaita kuljetuspolyeptidi- $\beta$ -galaktosidaasikonjugaattien oton solun sisään, valmistimme ohuita kudossjääleikkeitä uhratuista eläimistä (edellä), suoritimme kestäväöinnin esimerkissä 2 (edellä) kuvatulla tavalla ja alistimme ne tavanomaiseen X-gal-värjäysmenettelyyn. Maksa, perna ja sydän värjäytyivät voimakkaasti. Keuhko ja luurankolihas värjäytyivät vähemmän voimakkaasti. Aivoissa,  
 10 haimassa ja munuaisessa ei näkynyt havaittavissa olevaa värjäytymistä. Mikroskooppinen tutkimus suurella suurennoksella paljasti vahvaa värjäytymistä solussa ja joissakin tapauksissa tumassa, ilmeisesti endoteelisoluisissa, jotka ympäröivät kudoksiin menevää verivarastoa.  
 15

#### 20 **Esimerkki 4**

**Soluihin sisään menon testauksia  $\beta$ -galaktosidaasin ja polyarginiinin sekä  $\beta$ -galaktosidaasin ja polylysiinin muodostamilla konjugaateilla**

Verrataksemme yksinkertaisten, emäksisistä aminohapoista koostuvien polymeerien tehoa meidän tat-peräisten kuljetuspolyeptidiemme tehoon, konjugoimme kaupallisesti saatavilla olevan polyarginiinin (Sigma Chem Co., St. Louis, MO, luettelonro P-4663) ja polylysiinin (Sigma, luettelonro P-2658)  $\beta$ -galaktosidaasiin esimerkissä 2 edellä kuvatulla tavalla. Lisäsimme konjugaatit HeLa-solujen elatusaineeseen konsentraationa 1 - 30  $\mu\text{g/ml}$ , klorokiinin kanssa tai ilman. Sen jälkeen kun oli inkuboitu konjugaattien  
 25  
 30

kanssa, kestäväimme ja värjäsimme solut ja tutkimme ne mikroskoopilla, kuten esimerkissä 2 edellä on kuvattu.

Polylyysiini- $\beta$ -galaktosidaasikonjugaatti antoi matalat pintavärjäytymistasot eikä tuman värjäytymistä. Polyarginiini- $\beta$ -galaktosidaasikonjugaatti antoi intensiivisen yleisen värjäytymisen, mutta tällöin näkyi vähemmän tuman värjäytymistä kuin tat1-72- $\beta$ -galaktosidaasi- ja tat37-72- $\beta$ -galaktosidaasikonjugaateilla. Jotta voisimme tehdä eron polyarginiini- $\beta$ -galaktosidaasikonjugaatin solun pinnalle sitoutumisen ja varsinaisen internalisaation välillä, käsitelimme solut trypsiinillä, proteaasilla, ennen kestäväinti- ja värjäysmenettelyjä. Trypsiinikäsittely eliminoi suurimman osan polyarginiini- $\beta$ -galaktosidaasilla värjättyjen solujen X-gal-värjäytymisestä, mikä viittasi siihen että polyarginiini- $\beta$ -galaktosidaasikonjugaatti oli sitoutunut solujen ulkopuolisiin rakenteisiin pikemminkin kuin todella internalisoitunut. Sitä vastoin solut, jotka oli altistettu tat1-72- tai 37-72- $\beta$ -galaktosidaasikonjugaateille, värjäytyivät trypsiinikäsittelystä huolimatta, mikä viittasi siihen että  $\beta$ -galaktosidaasilasti oli solujen sisällä ja näin suojassa trypsiinidigestiolta. Kontrollisoluuissa, jotka oli käsitelty konjugoimattomalla  $\beta$ -galaktosidaasilla, ei näkynyt havaittavissa olevaa värjäytymistä.

## Esimerkki 5

## Piparjuuriperoksidaasikonjugaatit

## Kemiallinen silloitus

Jotta saisimme tuotettua tat1-72-HRP- ja tat37-72-HRP-konjugaatit, käytimme kaupallisesti saatavilla olevaa HRP-liittämistarvikesarjaa (maleimidilla aktivoitu Immuno-pure-HRP, Pierce Chem. Co., luettelonro 31498G). Tarvikesarjan mukana tuleva HRP on muodossa, joka on selektiivisesti reaktiivinen vapaiden -SH-ryhmien suhteen. (Kysteini on 20 proteiiniaminohaposta ainoa, jolla on vapaa -SH-ryhmä.) Kuljetuspolyeptidi-HRP-konjugaatiokokeessa, joka käsitti tat1-72:n, tuotimme tat1-72-aloitusmateriaalin E. colissa ja puhdistimme sen HPLC:llä kuten esimerkissä 1 edellä kerrotaan. Lyofilisoimme 200 µg puhdistettua tat1-72:ta (joka liuotettiin TFA/asetonitriiliin) ja liuotimme sen uudestaan 100 µl:aan 100-millimolaarista HEPES-puskuria (pH 7,5), joka oli 0,5-millimolaarinen EDTA:n suhteen. Lisäsimme 50 µl tat1-72- tai tat37-72-liuosta 50 µl:aan Immunopure HRP:tä (750 µg entsyymiä) 250-millimolaarisessa trietanoliamiinissa (pH 8,2). Annoimme reaktion jatkua 80 minuuttia huoneenlämpötilassa. Näissä olosuhteissa noin 70 % HRP:stä oli liittynyt kemiallisesti tat1-72-molekyyleihin. Monitoroimme liittämisreaktion laajuutta SDS-PAGE-analyysillä.

## 25 HRP-konjugaattien otto solun sisään

Lisäsimme konjugaatit HeLa-solujen elatusaineeseen pitoisuutena 20 µg/ml, klorokiinin (100 µM) ollessa läsnä tai poissa. Inkuboimme soluja 4 - 18 tuntia 37 °C:n lämpötilassa, 5,5 % CO<sub>2</sub>. Kehitimme HRP-värjäyksen käyttämällä 4-kloro-1-naftolia (Bio-Rad, Richmond, CA, luettelonro 170-6431) ja vetyperoksidia HRP:n substraattina. Seuraavissa kokeissa korvasimme diaminobentsidiinin (Sigma Chem. Co., St. Louis, MO) 4-kloro-1-naftolilla.

35 Soluissa, joihin lisäsimme kuljetuspolyeptidi-HRP-konjugaatit, näkyi soluun liittynyttä HRP-aktiivisuutta.



Lyhyet konjugaatille altistamisen ajanjaksot saivat aikaan värjäytymiskuvioita, jotka vaikuttivat pistemäisiltä, mikä mahdollisesti heijastaa HRP:tä endosyyttisissä vehikkeleissä. Pitempien inkubaatioiden jälkeen havaitsimme dif-  
 5 fuusia tuman ja sytoplasman värjäytymistä. Kontrollisoluissa, jotka oli käsitelty konjugoimattomalla HRP:llä, ei näkynyt havaittavissa olevaa värjäytymistä.

#### Esimerkki 6

#### PE-ADP-ribosylaatio-osa-konjugaatit

10 Kloonasimme ja ekspressoimme *E. coli*ssa Pseudomonas-eksotoksiinin ("PE") sekä täyspitkinä muotona että sen ADP-ribosylaatio-osan muodossa. Tuotimme kuljetuspolypeptidi-PE-konjugaatteja sekä geneettisellä fuusiolla että kemiallisella silloituksella.

#### 15 Plasmidin muodostaminen

Plasmidin pTat70(ApaI) muodostamiseksi sijoitimme ainutkertaisen ApaI-kohdan tat:n avoimeen lukukehykseen digestoimalla pTat72:n BamHI:llä ja EcoRI:llä, ja sijoitimme kaksijuosteisen linkkerin, joka koostui seuraavista  
 20 synteettisistä oligonukleotideista:

GATCCCAGAC CCACCAGGTT TCTCTGTCGG GCCCTTAAG (SEQ ID nro 8)

AATTCTTAAG GGCCCGACAG AGAAACCTGG TGGGTCTGG (SEQ ID nro 9)

25 Linkkeri korvasi tat:n C-pään, LysGlnStop, GlyPro-Stop:lla. Linkkeri lisäsi myös ainutkertaisen ApaI-kohdan, joka oli sopiva tat-sekvenssin kehyksensijaista fuusiota varten PE-ADP:n ribosylaatio-osaa koodittavien sekvenssien kanssa, käyttämällä PE-sekvenssissä luonnostaan olevaa  
 30 ApaI-kohtaa. Plasmidin pTat70PE (SEQ ID nro 10) muodostamiseksi poistimme ApaI-EcoRI-fragmentin, joka koodittaa PE-ADP:n ribosylaatio-osaa, plasmidista CD4(181)-PE(392). CD4(181)-PE(392):n muodostaminen on kuvattu G. Winklerin ym. artikkelissa ("CD4-Pseudomonas Exotoxin Hybrid  
 35 Proteins: Modulation of Potency and Therapeutic Window

Through Structural Design and Characterization of Cell Internalization", AIDS Research and Human Retroviruses, 7, s. 393 - 401 (1991)). Sijoitimme ApaI-EcoRI-fragmentin pTat70(ApaI):een, joka oli digestoitu ApaI:llä ja EcoRI:llä.

Plasmidin pTat8PE (SEQ ID nro 11) muodostamiseksi poistimme 214 emäsparin kokoisen NdeI-ApaI-fragmentin pTat70PE:stä ja korvasimme sen kaksijuosteisella linkkerillä, jossa on NdeI- ja ApaI-kohesiiviset päät, ja joka koodittaa tat-tähteitä 1 - 4 ja 67 - 70, ja joka koostuu seuraavista synteettisistä oligonukleotideista:

TATGGAACCG GTCGTTTCTC TGTCGGGCC (SEQ ID nro 12)

CGACAGAGAA ACGACCGGTT CCA (SEQ ID nro 13).

#### TAT8-PE:n puhdistus

pTat8-PE-konstruktion puhdistus tuotti PE-ADP:n ribosylaatio-osa-polypeptidin, joka oli fuusioitu tat-proteiinin aminohappoihin 1 - 4 ja 67 - 70. pTat8-PE-ekspressiotuote ("tat8-PE") toimi PE-ADP:n ribosylaatio-osa-osana (ja konjugoimattomana kontrollina) jäljempänä kuvatuissa kemiallisissa silloituskokeissa. 8:n tat-aminohapon koodit olivat artefaktoja kloonausmenetelmästä, joka valittiin mukavuussyistä. PE-ADP:n ribosylaatio-osaan fuusioidulla 8 aminohapolla ei ole kuljetusaktiivisuutta (kuvio 2).

Tat8-PE:n puhdistusta varten suspensioimme 4,5 g pTAT8-PE-transformoitua *E. colia* 20 ml:aan 50-millimolaarista Tris-HCl:ää (pH 8,0), joka oli 2-millimolaarinen EDTA:n suhteen. Hajotimme solut "French press" -laitteessa ja poistimme liukenemattoman roskan sentrifugoimalla 10 000 rpm:n nopeudella 1 tunnin ajan SA600-roottorissa. Suurin osa tat8-PE:stä oli supernatantissa. Latasimme supernatantin 3 ml:n Q-Sepharose Fast Flow -ioninvaihtopylväeseen (Pharmacia LKB, Piscataway, NJ). Näytteen latauksen jälkeen pesimme pylvään 50-millimolaarisella Tris-HCl:llä (pH 8,0), joka oli 2-millimolaarinen EDTA:n suh-

teen. Pylvään pesemisen jälkeen suoritimme vaiheittaisen gradienttieluution käyttämällä samaa puskuria, joka oli 100-, 200- ja 400-millimolaarinen NaCl:n suhteen. Tat8-PE eluoitui 200-millimolaarisella NaCl:llä. Ioninvaihtokromatografian jälkeen puhdistimme edelleen tat8-PE:n geelisuodatuksella Superdex 75 FPLC-pylväässä (Pharmacia LKB, Piscataway, NJ). Tasapainotimme geelisuodatuspylvään 50-millimolaarisella HEPES-puskurilla (pH 7,5). Sitten la-  
 5 tasimme näytteen ja suoritimme eluution tasapainotuspuskurilla nopeudella 0,34 ml/min. Keräsimme 1,5 minuutin fraktioita ja varastoimme tat8-PE-fraktiot -70 °C:n lämpötilaan.  
 10

#### TAT8-PE:n silloitus

Koska PE-ADP-ribosylaatio-osassa ei ole kysteiniini-  
 15 tähteitä, käytimme sulfo-SMCC:tä (Pierce Chem. Co., Rockford, IL, luettelonro 22322 G) kuljetuspolypeptidin ja tat8-PE:n konjugointiin. Suoritimme konjugoinnin 2-vaiheisella reaktiomenetelmällä. Ensimmäisessä reaktiovaiheessa käsittelimme tat8-PE:n (3 mg/ml) 50-millimolaarisella  
 20 HEPES-puskurilla (pH 7,5), joka oli 10-millimolaarinen sulfo-SMCC:n suhteen, huoneenlämpötilassa 40 minuutin ajan. Sulfo-SMCC lisättiin reaktioon 100-millimolaarisena kantaliuoksena 1-molaarisessa HEPES-puskurissa, pH 7,5). Erotimme tat8-PE-sulfo-SMCC:n reagoimattomasta sulfo-SMCC:stä geelisuodatuksella P6DG-pylväässä (Bio-Rad,  
 25 Richmond, CA), joka oli tasapainotettu 25-millimolaarisessa HEPES-puskurissa (pH 6,0), joka oli 25-millimolaarinen NaCl:n suhteen. Toisessa reaktiovaiheessa annoimme tat8-PE-sulfo-SMCC:n (1,5 mg/ml, 100 mM HEPES (pH 7,5), 1 mM  
 30 EDTA) reagoida puhdistetun tat37-72:n kanssa (lopullinen konsentraatio 600 µM) huoneenlämpötilassa 1 tunnin ajan. Silloitusreaktion lopettamiseksi lisäsimme kysteiniiniä. Analysoimme silloitusreaktiotuotteet SDS-PAGE:ssa. Noin 90 % tat8-PE:stä silloittui tat37-72-kuljetuspolypeptidiin  
 35 näissä oloissa. Noin puolella konjugoituneesta tuotteesta

oli yksi kuljetuspolypeptidiosa, ja puolella oli kaksi kuljetuspolypeptidiosaa.

#### Soluton määrittäminen PE-ADP:n ribosylaatiolle

Sen varmistamiseksi, että PE:n ribosylaatio-osassa  
 5 säilyi sen biologinen aktiivisuus (se on: tuhoava ribosomien muokkaus) kuljetuspolypeptideihin konjugoimisen jälkeen, testasimme kuljetuspolypeptidi-PE-ADP-ribosylaatio-konjugaattien vaikutusta in vitro (se on: soluttomasti) translaatioon. Kutakin in vitro translaatiokoetta varten  
 10 teimme tuoreen translaatioseoksen ja pidimme sen jäällä. In vitro translaatioseos sisälsi 200 µl kanin retikuloosyyttilysaattia (Promega, Madison, WI), jossa oli 2 µl 10-millimolaarista ZnCl<sub>2</sub>:ta (valinnainen), 4 µl 20:n proteiiniaminohapon muodostamaa seosta, josta puuttui metioniini, ja 20 µl <sup>35</sup>S-metioniinia. Lisäsimme 9 µl:aan translaatioseosta 1 - 1000 ng kuljetuspolypeptidi-PE-konjugaattia (edullisesti 1 µl:n tilavuudessa) tai kontrollia, ja esi-  
 15 inkuboimme seosta 60 minuuttia 30 °C:n lämpötilassa. Lisäsimme sitten 0,5 µl BMV-RNA:ta kuhunkin näytteeseen ja  
 20 inkuboimme vielä 60 minuuttia 30 °C:n lämpötilassa. Varastoimme näytteet -70 °C:n lämpötilaan sen jälkeen kun olimme lisänneet 5 µl 50-prosenttista glyserolia per näyte. Analysoimme in vitro translaatioreaktion tuotteet SDS-PAGE-menetelmin. Latasimme 2 µl kutakin translaatioreak-  
 25 tioseosta (sekä sopivan määrän SDS-PAGE-näytepuskuria) kaistaa kohden SDS-geeleihin. Elektroforeesin jälkeen visualisoimme <sup>35</sup>S:n sisältävät in vitro -translaatiotuotteet fluorografialla.

Havaitsimme edeltävässä kappaleessa kuvattua menetelmää käyttämällä, että PE-ADP:n ribosylaatio-osalla, joka on fuusioitu geneettisesti tat1-70-kuljetuspolypeptidiin, ei ollut biologista aktiivisuutta, se on: se ei estänyt in vitro translaatiota. Sitä vastoin samaa menetelmää käyttämällä havaitsimme, että PE-ADP:n ribosylaatio-  
 30 osa, joka on silloitetTU kemiallisesti tat37-72-kuljetus-

polypeptidiin, oli säilyttänyt täyden biologisen aktiivisuuden, se on: inhiboinut in vitro translaation yhtä hyvin kuin kontrollit, joissa oli konjugoimattomia PE-ADP:n ribosylaatio-osia (kuvio 2).

5                    **Sytotoksisuusmääritys PE-ADP:n ribosylaatiolle**

Toisessa testissä, jossa oli mukana tat37-72:n ja PE-ADP:n ribosylaatio-osan konjugaatti, lisäsimme sen viljellyille HeLa-soluille klorokiinin (100 µM) ollessa läsnä tai poissa. Määritimme sitten sytotoksisuuden mittaamalla in vitro proteiinisynteesiä, jonka osoittaa trikloorietikkahapolla ("TCA") saostettavissa oleva radioaktiivisuus soluekstrakteissa.

Suoritimme sytotoksisuusmäärityksen seuraavasti. Hajotimme HeLa-solukerrokset, sentrifugoimme solut ja resuspendoimme ne elatusaineeseen tiheydellä  $2,5 \times 10^4$ /ml. Käytimme 0,5 ml suspensiota/kuoppa, kun käytimme 24-kuoppaisia levyjä, tai 0,25 ml suspensiota/kuoppa, kun käytimme 48-kuoppaisia levyjä. Lisäsimme konjugaatteja tai konjugoimattomia kontrolleja, jotka oli liuotettu 100 µl:aan PBS:ää, kuoppiin sen jälkeen kun solujen oli annettu asetua ainakin 4 tunnin ajan. Inkuboimme soluja konjugaattien tai kontrollien läsnäollessa 60 minuutin ajan 37 °C:n lämpötilassa, sitten lisäsimme 0,5 ml tuoretta elatusainetta kullekin solulle ja inkuboimme soluja vielä 5 - 24 tuntia. Tämän inkubaation jälkeen poistimme elatusaineen kustomakin kuopasta ja pesimme solut kerran noin 0,5 ml:lla PBS:ää. Sitten lisäsimme 1 µCi <sup>35</sup>S-metioniinia (Amersham) per 100 µl per kuoppa, in vivo soluleimauslaatu (SJ.1015) ja inkuboimme soluja 2 tunnin ajan. Kahden tunnin kuluttua poistimme radioaktiivisen elatusaineen ja pesimme solut 3 kertaa kylmällä 5-prosenttisellä TCA:lla ja sitten kerran PBS:llä. Lisäsimme 100 µl 0,5-molaarista NaOH:ta kuhunkin kuoppaan ja annoimme solujen hajoamisen ja proteiinien liukenemisen tapahtua ainakin 45 minuutin ajan. Lisäsimme sitten kuhunkin kuoppaan 50 µl 1-molaarista HCl:ää ja

siirsimme kunkin kuopan koko sisällön tuikenesteeseen radioaktiivisuuden nestetuikemittausta varten.

Klorokiinin poissa ollessa oli olemassa selvä annosriippuvainen soluproteiinisynteesin inhibitio vasteena käsittelylle kuljetuspolypeptidin ja PE-ADP:n ribosylaatio-osan konjugaatilla, mutta ei vasteena käsittelylle konjugoimattomalla PE-ADP:n ribosylaatio-osalla. Tulokset esitetään yhteenvedona kuviossa 2. Tat37-72:een konjugoituna PE-ADP:n ribosylaatio-osa vaikutti olevan kuljetettu 3 - 10 kertaa tehokkaammin kuin tat1-72:een konjugoituna. Konjugoimme myös kuljetuspolypeptidit tat38-58GGC, tat37-58, tat47-58GGC ja tatCGG-47-58 PE-ADP:n ribosylaatio-osaan. Kaikki nämä konjugaatit saivat aikaan biologisesti aktiivisen PE-ADP:n ribosylaatio-osan oton solun sisään (tuloksia ei esitetä).

#### Esimerkki 7

#### Ribonukleaasikonjugaatit

#### Kemiallinen silloitus

Liuotimme 7,2 mg naudan haiman ribonukleaasi A:ta, tyyppi 12A (Sigma Chem. Co., St. Louis, MO, luettelonro R5500) 200 µl:aan PBS:ää (pH 7,5). Ribonukleaasiliuokseen lisäsimme 1,4 mg sulfo-SMCC:tä (Pierce Chem. Co., Rockford, IL, luettelonro 22322H). Vortex-laitteella sekoituksen jälkeen annoimme reaktion edetä huoneenlämpötilassa 1 tunnin ajan. Poistimme reagoimattoman SMCC:n ribonukleaasi-SMCC:stä viemällä reaktiotuotteen 9 ml:n P6DG-pylvään läpi (Bio-Rad, Richmond, CA) ja keräämällä 0,5 ml:n fraktioita. Tunnistimme välisijatilavuuden huippufraktiot (jotka sisälsivät ribonukleaasi-SMCC-konjugaatin) monitoroimalla UV-absorbanssia 280 nm:n aallonpituudella. Jaoimme yhdistetyt, ribonukleaasi-SMCC:n sisältävät fraktiot 5 yhtäsuureen erään. Kuhunkin 4:ään ribonukleaasi-SMCC-erään lisäsimme kemiallisesti syntetisoidun kuljetuspolypeptidin, joka vastasi tat-tähteitä: 37-72 ("37-72"), 38-58 sekä GGC karboksipäässä ("38-58GGC"), 37-58 ("CGG37-

58"), tai 47-58 sekä CGG aminopäässä ("CGG47-58"). Annoim-  
mekuljetuspolyeptidi-ribonukleaasikonjugaatioreaktioiden  
edettä 2 tunnin ajan huoneenlämmössä, ja sitten yli yön  
4 °C:n lämpötilassa. Analysoimme reaktiotuotteet SDS-  
5 PAGE:ssa 10 - 20-prosenttisessa gradienttigelissä. Sil-  
loitustehokkuus oli noin 60-prosenttinen kuljetuspolyep-  
ptideille tat38-58GGC, tat37-58 ja tatCGG47-58 ja 40-pro-  
senttinen tat37-72:lle. Modifioiduista luokista 72 % si-  
sälsi yhden ja 25 % sisälsi kaksi kuljetuspolyeptidisubs-  
10 titutiota.

**Tat37-72-ribonukleaasikonjugaattien otto solun si-  
sään**

Pidimme soluja yllä 37 °C:n lämpötilassa soluvilje-  
lyinkubaattorissa Dulbeccon muokatussa Eaglen elatusai-  
15 neessa, jossa oli lisänä 10-prosenttinen fetaalivasikan  
seerumi ja penisilliini/streptomysiini. Solun sisäänoton  
määrittäjä varten jaoimme  $10^5$  solua 24-kuoppaiselle levyl-  
le ja viljelimme niitä yli yön. Pesimme solut Dulbeccon  
PBS:llä ja lisäsimme ribonukleaasikonjugaatin, joka oli  
20 liuotettu 300 µl:aan PBS:ää, joka oli 80 µM klorokiinin  
suhteen, konsentraatioina 0, 10, 20, 40 ja 80 µg/ml. Kun  
oli inkuboitu 1,25 tuntia 37 °C:n lämpötilassa, lisäsimme  
750 µl kasvatuselatusainetta ja inkuboimme solunäytteitä  
vielä yli yön. Yli yön inkuboinnin jälkeen pesimme solut  
25 kerran PBS:llä ja inkuboimme niitä 1 tunnin ajan Minimal  
Essential Medium -elatusaineessa ilman metioniinia (Flow  
Labs) (250 µl/kuoppa), joka sisälsi <sup>35</sup>S-metioniinia  
(1 µCi/kuoppa). Kun oli inkuboitu 1 tunti radioaktiivisen  
metioniinin kanssa, poistimme elatusaineen ja pesimme so-  
30 lut kolme kertaa 5-prosenttisellä TCA:lla (1 ml/kuoppa/pe-  
su). Sitten lisäsimme 250 µl 0,5-molaarista NaOH:ta per  
kuoppa. Kun oli kulunut 1 tunti huoneenlämpötilassa, pipe-  
toimme 200 µl kunkin kuopan sisältöä tuikepulloon, li-  
säsimme 100 µl 1-molaarista HCl:ää ja 4 ml tuikenestettä.  
35 Sen jälkeen kun kunkin pullon sisältöä oli sekoitettu pe-

rusteellisesti, mittasimme kussakin näytteessä olevan radioaktiivisuuden nestetuikelaskennalla.

Solun sisään ottoa koskevat tulokset esitetään yhteenvedona kuviossa 3. Kuljetuspolyeptidi tat38-58GGC  
 5 toimi yhtä hyvin tai hieman paremmin kuin tat37-72. Kuljetuspolyeptidillä tatCGG47-58 oli vähentynyt aktiivisuus (tuloksia ei esitetä). Emme tiedä, oliko tällä polyeptidillä vähentynyt sisäänmenoaktiivisuus vai oliko emäksisen alueen läheisyys ribonukleaasiin häirinnyt entsyymiaktiivisuutta.  
 10

Olemme käyttäneet kationinvaihtokromatografiaa (BioCAD-perfuusiokromatografijärjestelmä, PerSeptive Biosystems) sellaista ribonukleaasikonjugaattien puhdistamiseen, joilla on yksi tai kaksi kuljetuspolyeptidiosaa.

15 **Esimerkki 8**

**Proteiinikinaasi A:n inhibiittorikonjugaatit**

**Kemiallinen silloitus**

Ostimme proteiinikinaasi A:n inhibiittori ("PKAI")  
 -peptidin (20 aminohappoa) Bachem California -yhtiöltä  
 20 (Torrence, CA). PKAI:n kemialliseksi silloittamiseksi kuljetuspolyeptideihin käytimme joko sulfo-MBS:ää (konsentraationa 10 mM) tai sulfo-SMPB:tä (konsentraationa 15 mM). Molemmat näistä silloitusaineista ovat heterobifunktionaalisia tioliryhmien ja primääristen amiiniryhmien  
 25 suhteen. Koska PKAI:sta puuttuvat lysiini- ja kysteiini-tähteet, sekä sulfo-MBS että sulfo-SMPB kohdistavat selektiivisesti silloituksen PKAI:n aminopäähän. Saatoimme PKAI:n reagoimaan konsentraatiossa 2 mg/ml, niin että läsnä oli 50-millimolaarinen HEPES-puskuri (pH 7,5) joka oli  
 30 25-millimolaarinen NaCl:n suhteen, huoneenlämpötilassa 50 minuutin ajan kumman tahansa silloitusaineen kanssa. Sulfo-MBS-reaktioseos sisälsi 10-millimolaarisen sulfo-MBS:n ja 20-prosenttisen DMF:n. Sulfo-SMPB-reaktioseos oli 15-millimolaarinen sulfo-SMPB:n suhteen ja 20-prosenttinen  
 35 dimetyylisulfoksidin suhteen ("DMSO"). Puhdistimme PKAI:n



ja silloitusaineiden aduktit käänteisfaasi-HPLC:llä käyttäen C<sub>4</sub>-pylvästä. Eluimme näytteet C<sub>4</sub>-pylvästä 20 - 75 -prosenttisella asetonitriiligradientilla, joka oli 0,1-prosenttinen trifluorietikkahapon suhteen. Poistimme aduk-  
 5 teista asetonitriilin ja trifluorietikkahapon lyofilisoi-  
 malla ja liuotimme ne uudelleen 25-millimolaariseen HEPES-  
 puskuriin (pH 6,0), joka oli 25-millimolaarinen NaCl:n  
 suhteen. Lisäsimme tat1-72:n tai tat37-72:n ja säädimme  
 10 reaktioseoksen pH:n 7,5:ksi lisäämällä 1-molaarista HEPES-  
 puskuria (pH 7,5) niin että loppukonsentraatio oli 100 mM.  
 Annoimme sitten silloitusreaktion edetä huoneenlämpötilas-  
 sa 60 minuutin ajan.

Säätelimme silloittumisen laajuutta muuttamalla  
 kuljetuspolyeptidin ja PKAI:n suhdetta. Analysoimme sil-  
 15 loitusreaktion tuotteita SDS-PAGE:ssa. Tat37-72:lla sil-  
 loitusreaktioissa muodostui yksittäinen uusi elektrofo-  
 reettinen juova. Tämä tulos oli yhdenmukainen sen kanssa  
 että yksittäiseen PKAI-molekyyliin oli lisätty yksittäinen  
 tat37-72-molekyyli. Tat1-72:n tapauksessa silloitusreak-  
 20 tioissa muodostui kuusi uutta tuotetta. Tämä tulos on yh-  
 täpitävä sen kanssa, että monia PKAI-molekyylejä lisättiin  
 yhtä tat1-72-polyeptidiä kohti, mikä johtuu tat1-72:ssa  
 olevista monista kysteiniinihteistä. Kun lisäsimme PKAI:n  
 silloitusreaktioon suurena molaarisena ylimääränä, saimme  
 25 aikaan vain konjugaatteja, jotka sisälsivät 5 tai 6 PKAI-  
 osaa tat1-72:ta kohden.

#### PKAI-aktiivisuuden fosforylaatiomääritys *in vitro*

Jotta saisimme testattua sulfo-MBS:llä silloitet-  
 tujen konjugaattien biologisen PKAI-aktiivisuuden säilymi-  
 30 sen, käytimme in vitro -fosforylaatiomääritystä. Tässä  
 määrittäksessä histoni V toimi substraattina proteiiniki-  
 naasi A:n fosforylaatiossa PKAI:n ollessa läsnä tai poissa  
 (PKAI-konjugaattia silmällä pitäen). Sitten käytimme SDS-  
 PAGEa monitoroidaksemme PKAI-riippuvaisia eroja fosfory-  
 35 laatioasteessa. Kussakin reaktiossa inkuboimme 5:tä yksik-

köä proteiinikinaasi A:n katalyyttistä alayksikköä (Sigma) ja vaihtelevia määriä PKAI:tä tai PKAI-konjugaattia 37 °C:n lämpötilassa 30 minuutin ajan. Määritysreaktioseos oli 24-millimolaarinen natriumasetaatin suhteen (pH 6,0) ja 25-millimolaarinen MgCl<sub>2</sub>:n suhteen, 100-millimolaarinen DTT:n suhteen, ja sisälsi 50 µCi [γ-<sup>32</sup>P]ATP:tä ja 2 µg histoni V:tä, 40 µl:n kokonaisreaktiotilavuudessa. Tätä määrittystä käyttämällä havaitsimme, että tat1-72:een tai tat37-72:een konjugoitu PKAI inhiboi fosforylaatiota yhtä hyvin kuin konjugoimaton PKAI (tuloksia ei esitetä).

#### Solumääritys

PKAI:n ja kuljetuspolypeptidi-PKAI-konjugaattien solun sisäänoton testaamiseksi käytimme viljeltyjä soluja, jotka sisälsivät kloramfenikoliasetyylitransferaasi ("CAT") -reportterigeenin cAMP-responsiivisen ekspressio-kontrollisekvenssin alaisuudessa. Niinpä kvantitoimme proteiinikinaasi A:n aktiivisuutta epäsuorasti mittaamalla CAT-aktiivisuutta. Tämän määrittäksen ovat kuvanneet yksityiskohtaisesti J. R. Grove ym. ("Probind cAMP-Related Gene Expression with a Recombinant Protein Kinase Inhibitor", Molecular Aspects of Cellular Regulation, Vol. 6, toim. P. Cohen ja J. G. Folkes, Elsevier Scientific, Amsterdam, s. 173 - 195 (1991)).

Tätä määrittystä käyttämällä emme havainneet PKAI:n tai minkään kuljetuspolypeptidi-PKAI-konjugaatin aktiivisuutta. Tätä tulos antoi meille kuvan, että PKAI-osa voisi hajota nopeasti mentyään solun sisään.

#### PKAI:n silloitus tat37-72-β-galaktosidaasiin

Olimme havainneet aikaisemmin tat37-72-β-galaktosidaasin solun sisään menon olevan klorokiinistä riippumattonta (esimerkki 2 edellä). Niinpä silloitimme PKAI:n tat37-72-β-galaktosidaasiin, jotta PKAI olisi mahdollisesti suojattu nopealta hajoamiselta.

Käsittelimme β-galaktosidaasin 20-millimolaarisella DTT:llä (pelkistävä aine) huoneenlämpötilassa 30 minuutin

ajan ja poistimme sitten DTT:n geelisuodatuksella G50-pylväässä MES-puskurissa (pH 5). Annoimme pelkistyneen  $\beta$ -galaktosidaasin reagoida SMPB-aktivoidun PKAI:n kanssa (edellä) pH:ssa 6,5 60 minuutin ajan. Jäljellä olevien vapaiden sulfhydryyliryhmien suojaamiseksi lisäsimme N-etyylimaleimidia tai jodiasetamidia. SDS-PAGE-analyysi osoitti, että ainakin 95 %  $\beta$ -galaktosidaasista oli konjugoitunut. Noin 90 % konjugoidusta beta-galaktosidaasituotteesta sisälsi yhden PKAI-osan per alayksikkö, ja noin 10 % sisälsi 2 PKAI-osaa. Käsittelimme PKAI- $\beta$ -galaktosidaasikonjugaatin 10-kertaisella sulfo-SMCC:n molaarisella ylimäärällä. Saatoimme sitten PKAI- $\beta$ -galaktosidaasi-SMCC:n reagoimaan tat1-72:n kanssa. SDS-PAGE-analyysin mukaan PKAI- $\beta$ -galaktosidaasin ja tat1-72:n suhde vaikutti olevan 1:0,5. Olemme tuottaneet noin 100  $\mu$ g lopputuotetta. Saostumisongelmien vuoksi lopputuotteen konsentraatio liuoksessa on rajattu 100  $\mu$ g/ml:ksi.

#### Esimerkki 9

##### E2-repressorikonjugaatit

Kuljetuspolyeptidi-E2-repressorikonjugaattien solun sisään oton ja E2-repressoriaktiivisuuden testaamiseksi transfektoimme samanaikaisesti E2-riippuvaisen reporteriplasmidin ja E2-ekspressioplasmidin SV40-transformoituneeseen "African green monkey" -munuaissoluihin ("COS7"). Sitten altistimme transfektoidut solut kuljetuspolyeptidi-E2-repressorikonjugaateille (tehty geneettisellä fuusiolla tai kemiallisella silloituksella) tai soveliaille kontroleille. Repressiomääritys, joka kuvataan jäljempänä, oli oleellisesti sama kuin Barsoumin ym. (edellä).

##### Repressiomäärityssolut

Hankimme COS7-solut American Type Culture Collection -talletuslaitoksesta, Rockville, MD (ATCC nro CRL 1651). Kasvatimme COS7-soluja Dulbeccon modifioidussa Eaglen elatusaineessa (Gibco, Grand Island, NY), jossa oli

10 % fetaalivasikan seerumia (JRH Biosciences, Lenexa, KS) ja joka oli 4-millimolaarinen glutamiinin suhteen ("kasvatuselatusaine"). Solujen inkubaatio-olosuhteet olivat 5,5 % CO<sub>2</sub>, 37 °C.

5                    **Repressiomäärittelyplasmidit**

Meidän E2-riippuvainen reportteriplasmidimme pXB322hGH sisälsi ihmisen kasvuhormoni-reportterigeenin, joka oli lyhennetyn, SV40:n varhaisen promoottorin alainen, joka sisälsi 3 E2-sitomiskohtaa ylävirrassa. Muodostimme hGH-reportteriplasmidin pXB322gGH, kuten Barsoum ym. ovat kuvanneet (edellä).

Täyspitkän HPV-E2-geenin ekspressiota varten muodostimme plasmidin pAGE2 (kuviot 4). Plasmidi pAHE2 sisältää E2-geenin HPV-kannasta 16, joka on liitetty toimivasti adenoviruksen myöhäiseen pääpromoottoriin, jota vahvistaa promoottorista ylävirtaan sijaitseva SV40-tehostaja. Eristimme HPV-E2-geenin plasmidista pHPV16 (täyspitkä HPV16-genomi, joka on kloonattu pBR322:een), jonka M. Durst ym. ovat kuvanneet artikkelissa "A Papillomavirus DNA from Cervical Carcinoma and Its Prevalence in Cancer Biopsy Samples from Different Geographic Regions", Proc. Natl. Acad. Sci. USA, 80, s. 3812 - 3815 (1983) Tth111I-AseI-fragmenttina. HPV-genomissa Tth111I katkaisee nukleotidin 2711 ja AseI katkaisee nukleotidin 3929 kohdalta. Teimme Tth111I-AseI-fragmentista tasapäisen DNA-polymerasi I Klenow -reaktiolla ja ligoimme BamHI-linkkerit (New England Biolabs, luettelonro 1021). Sijoitimme tämän linkkerin sisältävän fragmentin BamHI-katkaistuun pBG331-plasmidiin saadaksemme aikaan plasmidin pAHE2.

30                    Plasmidi pBG331 on sama kuin pBG312 (R. L. Cate ym., "Isolation of the Bovine and Human Genes for Mullerian Inhibiting Substance and Expression of the Human Gene in Animal Cells", Cell, 45, s. 685 - 698 (1986)), paitsi että siitä puuttuu BamHI-kohta SV40-polyadenylaatio-signaalista alavirtaan, mikä tekee BamHI-kohdan pro-

moottorin ja SV40-intronin välillä ainutkertaiseksi. Pois-  
timme epäsuotavan BamHI-kohdan pBG312:n osittaisella  
BamHI-digestiolla, linearisoidun plasmidin geelipuhdistuk-  
sella, tasapäiden muodostamisella DNA-polymeraasi I Klenow  
5 -käsittelyllä, ligaatiolla itsen kanssa ja seulomalla  
plasmideja, joissa olisi haluttu BamHI-kohdan deletio.

#### E2-repressoriproteiinien tuotanto bakteereissa

Yksi meidän E2-repressoriproteiineistamme, E2.123,  
koostui HPV E2:n karboksiterminaalisisista 121 aminohaposta,  
10 jossa aminopäähän oli lisätty MetVal. Käytimme myös va-  
rianttia E2.123, jolle annettiin nimeksi E2.123CCSS.  
E2.123:ssa on kysteiinitähteet HVP16 E2:n aminohappokoh-  
dissa 251, 281, 300 ja 309. E2.123CCSS:ssä kysteiinitäh-  
teet kohdissa 300 ja 309 vaihdettiin seriiniksi ja ly-  
15 siinitähde kohdassa 299 vaihdettiin arginiiniksi. Poistim-  
me kysteiinitähteet kohdissa 300 ja 309, niin että kyste-  
iiniriippuvainen kemiallinen silloitus voisi tapahtua E2-  
repressorin aminoterminalisessa osassa, mutta ei E2:n  
minimaalisessa DNA:han sitoutuvassa/dimerisaatio-osassa.  
20 Ajattelimme, että minimaalisessa DNA:han sitoutuvassa  
osassa silloitukset todennäköisesti häiritisivät repres-  
sorin biologista aktiivisuutta.

Plasmidin pET8c-123:n muodostamista varten (kuvio  
5, SEQ ID nro 14) tuotimme välttämättömän DNA-fragmentin  
25 standardinomaisella polymeraasiketjureaktio ("PCR") -tek-  
niikoilla, joissa plasmidi pHPV16 oli templaatti. (PCR-  
tekniikoista voi löytää yleisen esityksen Sambrookin ym.  
teoksen kappaleesta 14, edellä. Automatisoitu PCR-laite ja  
kemikaalit ovat kaupallisesti saatavilla.) EA52:n, PCR-  
30 oligonukleotidialukkeen 374 emäsparin E2-123-fragmentin  
5'-päästä varten, nukleotidisekvenssi esitetään sekvenssi-  
listassa nimellä SEQ ID nro 15. EA54:n, PCR-oligonukleoti-  
dialukkeen, jota käytettiin E2-123-fragmentin 3'-päästä  
varten, nukleotidisekvenssi esitetään sekvenssiluettelossa  
35 nimellä SEQ ID nro 16. Digestoimme PCR-tuotteet NcoI:llä

ja BamHI:llä ja kloonasimme syntyneen fragmentin NcoI/BamHI-digestoituun pET8c-ekspressioplasmidiin (Studier ym., edellä), jolloin saatiin aikaan plasmidi pET8c-123.

5 Käyttämällä samaa menetelmää, jossa oli eri 5'-pään oligonukleotidi-PCR-alue, hankimme 260 emäsparin fragmen-  
 10 tin ("E2-85"), joka sisälsi metioniinikodonin ja alaniini-  
 kodonin, jota seurasivat välittömästi kodonit HPV16-E2:n  
 83:lle karboksiterminaaliseen aminohapolle. EA57:n, 5'-  
 15 pään PCR-alueen E2-85:n tuottamiseksi, nukleotidisek-  
 20 venssi esitetään sekvenssiluettelossa nimellä SEQ ID nro  
 34.

Plasmidin pET8c-123CCSS:n muodostamiseksi (kuvio 6,  
 SEQ ID nro 17), E2.123CCSS:n bakteerissa tuottoa varten,  
 15 syntetisoimme 882 emäsparin PstI-EagI-DNA-fragmentin PCR-  
 tekniikoilla. PCR-templaatti oli pET8c-123. Yksi PCR-alue-  
 keista, jolle oli annettu nimeksi 374.140, kooditti kaik-  
 kua kolmea aminohappomuutosta: CGACACTGCA GTATACAATG  
 TAGAATGCTT TTTAAATCTA TATCTTAAAG ATCTTAAAG (SEQ ID nro  
 20 18). Toisella PCR-alueella, 374.18, oli seuraava sek-  
 venssi: GCGTCGGCCG CCATGCCGGC GATAAT (SEQ ID nro 19). Di-  
 gestoimme PCR-reaktiotuotteet PstI:llä ja EagI:llä ja  
 eristimme 882 emäsparin fragmentin standardimenetelmällä.  
 Lopullinen vaihe oli pET8c-123CCSS:n tuottaminen kolmen  
 25 kappaleen ligaatiossa, jossa pET8c-123:sta peräisin oleva  
 3424 bp:n EcoRI-EagI-fragmentti liitettiin 882 emäsparin  
 PCR-fragmenttiin ja 674 emäsparin PstI-EcoRI-pET8c-123-  
 fragmenttiin, kuten kuviossa 6 esitetään. Varmistimme  
 konstruktion DNA-sekvenssianalyysillä. E2.123:n ja  
 30 E2.123CCSS-proteiinien tuottamista varten ekspressoimme  
 plasmidit pET8c-123 ja pET8c-123CCSS *E. coli* -kannassa  
 BL21(DE3)pLysS, kuten Studier ym. ovat kuvanneet (edellä).

#### E2-repressoriproteiinien puhdistus

35 Sulatimme 3,6 grammaa pakastettuja, pET8c-123-  
 transformoituja *E. coli* -soluja ja suspensoimme ne 35

ml:aan 25-millimolaarista Tris-HCl:ää (pH 7,5), joka oli 0,5-millimolaarinen EDTA:n, 2,5-millimolaarinen DTT:n suhteen ja sisälsi proteaasi-inhibiittoreita (1 mM PMSF, 3 mM bentsamidiini, 50 µg/ml pepstatiini A, 10 µg/ml aprotiniini). Hajotimme solut viemällä ne kahdesti "French press"-laitteen läpi paineella 10 000 psi (68 880 kPa). Sentrifugoimme lyaattia 12 000 rpm:n nopeudella SA600-roottorissa 1 tunnin ajan. E2.123-proteiini oli supernatantissa. Lisäsimme supernatanttiin MES-puskuria (pH 6) 25 mM:ksi, MES-puskuria (pH 5) 10 mM:ksi ja NaCl:a 125 mM:ksi. Panimme sitten supernatantin 2 ml:n S Sepharose Fast Flow -pylvääseen nopeudella 6 ml/h. Lataamisen jälkeen pesimme pylvään 50-millimolaarisella Tris-HCl:llä (pH 7,5), joka oli 1-millimolaarinen DTT:n suhteen. Sitten suoritimme vaiheittaisen gradienttieluution (2 ml/vaihe), jossa oli 200-, 300-, 400-, 500-, 700- ja 1000-millimolaarinen NaCl 50-millimolaarisessa Tris-HCl:ssä (pH 7,5), jotka olivat 1-millimolaarisia DTT:n suhteen. E2.123-repressoriproteiini eluoitui 500- ja 700-millimolaarisissa NaCl-fraktioissa. SDS-PAGE-analyysi osoitti, että E2.123-repressorin puhtaus oli yli 95-prosenttinen.

Sulattimme 3,0 grammaa pakastettua, pET8c-123CCSS-transformoitua *E. colia* ja suspensoimme solut 30 ml:aan samaa puskuria, jota käytettiin pET8c-123-transformoiduille soluille (edellä). Hajotus, liukenemattoman soluroskan ja MES-puskurin ja NaCl:n lisäys oli myös samanlainen kuin E2-123:n puhdistuksen tapauksessa kuvattiin. E2.123CCSS:n puhdistusmenetelmä erosi MES-puskurin ja NaCl:n lisäyksen jälkeen, koska E2.123CCSS:n tapauksessa tässä menetelmän vaiheessa muodostui sakka. Poistimme sakan sentrifugoimalla, ja havaitsimme että se ja supernatantti sisälsivät molemmat oleellista E2-repressoraktiivisuutta. Niinpä alistimme molemmat puhdistusvaiheisiin. Panimme supernatantin 2 ml:n S Sepharose Fast Flow -pylvääseen (Pharmacia LKB, Piscataway, NJ) nopeudella 6 ml/h. Lataamisen jälkeen

pesimme pylvään 50-millimolaarisella Tris-HCl:llä (pH 7,5), joka oli 1-millimolaarinen DTT:n suhteen. Pylvään pesun jälkeen suoritimme vaiheittaisen gradienttieluution (2 ml/vaihe) käyttämällä puskuria, joka oli 300-, 400-, 500-, 700- ja 1000-millimolaarinen NaCl:n suhteen ja 50-millimolaarinen Tris-HCl:n suhteen (pH 7,5) sekä 1-millimolaarinen DTT:n suhteen. E2.123CCSS-proteiini eluoitui 700-millimolaarisella NaCl:llä. SDS-PAGE-analyysi osoitti, että sen puhtaus oli yli 95 %. Liuotimme E2.123CCSS-sakan 7,5 ml:aan puskuria, joka oli 25-millimolaarinen Tris-HCl:n (pH 7,5), 125-millimolaarinen NaCl:n, 1-millimolaarinen DTT:n ja 0,5-millimolaarinen EDTA:n suhteen. Laatasimme liuotetun materiaalin 2 ml:n S Sepharose Fast Flow -pylvääseen ja pesimme pylvään kuten on kuvattu E2.123:n ja ei-sakkautuneen E2.123CCSS:n tapauksessa. Suoritimme vaiheittaisen gradienttieluution (2 ml/vaihe) käyttämällä 300-, 500-, 700- ja 1000-millimolaarista NaCl:ää. E2-repressori eluoitui 500 - 700-millimolaarisen NaCl:n fraktioissa. SDS-PAGE-analyysi osoitti, että sen puhtaus oli yli 98 %. Heti E2.123- ja E2.123CCSS-proteiinien puhdistuksen jälkeen lisäsimme glyserolia, niin että sen loppukonsentraatio oli 15 % (v/v) ja varastoimme pikapakastetuja (nestetyppi) eriä -70 °C:n lämpötilaan. Kvantitoimme puhdistetut E2-repressoriproteiinit UV-absorbanssilla 280 nm:n aallonpituudella käyttämällä ekstinktiokerrointa 1,8 pitoisuudessa 1 mg/ml.

#### Kemiallinen silloitus

Suoritimme kuljetuspolypeptidin, joka koostui tatar-aminohapoista 37 - 72, kemiallisen synteessin esimerkissä 1 kuvatulla tavalla. Liuotimme polypeptidin (5 mg/ml) 10-millimolaariseen MES-puskuriin (pH 5,0), joka oli 50-millimolaarinen NaCl:n, 0,5-millimolaarinen EDTA:n suhteen (ekstinktiokerroin 0,2 pitoisuudessa 1 ml/ml). Kuljetuspolypeptidiliuokseen lisäsimme bismaleimidoheksaanin ("BMH") (Pierce Chemical Co., Rockford, IL, luettelonro 22319G)



kantaliuosta (6,25 mg/ml DMF) niin, että loppukonsentraatio oli 1,25 mg/ml, ja HEPES-puskurin kantaliuosta (1 M, pH 7,5) niin että loppukonsentraatio oli 100 mM. Annoimme BMH:n reagoida proteiinin kanssa 30 minuutin ajan huoneen-  
5 lämpötilassa. Sitten erotimme proteiini-BMH:n reagoimattomasta BMH:sta geelisuodatuksella G-10-pylväässä, joka oli tasapainotettu puskurilla, joka oli 10-millimolaarinen MES:n (pH 5), 50-millimolaarinen NaCl:n ja 0,5-millimolaarinen EDTA:n suhteen. Varastoimme kuljetuspolyeptidi-BMH-  
10 konjugaatin erät -70 °C:n lämpötilaan.

Kuljetuspolyeptidi-BMH-konjugaatinsilloittamiseksi E2-repressorin poistimme E2-repressoriproteiinin sen säilytyspuskurista. Laimensimme E2-repressoriproteiinin kolmella tilavuudella 25-millimolaarista MES:iä (pH 6,0),  
15 joka oli 0,5-millimolaarinen EDTA:n suhteen, ja latasimme sen eränä S Sepharose Fast Flow -hartsiin (Pharmacia LKB, Piscataway, NJ), 5 mg proteiinia per ml hartsia. Sen jälkeen kun proteiinilla ladattu hartsiliete oli kaadettu pylvääseen, pesimme pylvään 25-millimolaarisella MES:llä  
20 (pH 6,0), joka oli 0,5-millimolaarinen EDTA:n ja 250-millimolaarinen NaCl:n suhteen. Sitten eluimme sitoutuneen E2-repressoriproteiinin pitoisuuteen 1 mg/ml 25-millimolaarisella MES:llä (pH 6,0), joka oli 0,5-millimolaarinen EDTA:n suhteen. Kokeiluluontoisista silloitustutkimuksista,  
25 jotka oli suoritettu kullakin E2-repressoriproteiinerällä ja BMH-aktivoitulla kuljetuspolyeptidillä, määritimme että E2-repressoriproteiinin (1 mg) käsittely 0,6 mg:lla BMH-aktivoitua kuljetuspolyeptidiä tuottaa halutun inkorporaation, 1 kuljetusmolekyylä per E2-repressorihomodimeeri. Tyypillisesti sekoitimme 2 ml E2-repressorin (1 mg/ml) 300 µl:aan tat37-72-BMH:ta (4 mg/ml) ja 200 µl:aan 1-molaarista HEPES:iä (pH 7,5). Annoimme silloitusreaktion edetä 30 minuuttia huoneenlämmössä. Lopetimme silloitusreaktion lisäämällä 2-merkaptetanolia niin, että loppukonsentraatio oli 14 mM. Määritimme silloittumisen määrän  
35

SDS-PAGE-analyysissä. Varastoimme tat37-72-E2-repressorikonjugaatin eriä -70 °C:n lämpötilaan. Käytimme identtisiä menetelmiä tat37-72-kuljetuspolypeptidin silloittamiseksi kemiallisesti HPVE2 123 -repressoriproteiiniin ja HPVE2  
 5 CCSS -repressoriproteiiniin.

#### E2-repressorikonjugaattien otto solun sisään

Meidän E2-repressorimäärityksiämme varten käytimme COS7-soluihin transfektoitujen plasmidien lyhytaikaista ekspressiota. Meidän E2-repressiomääritysmenetelmämme oli  
 10 samanlainen kuin Barsoumin ym. artikkelissaan kuvaama (edellä). Transfektoimme  $4 \times 10^6$  COS7-solua (noin 50-prosenttisen konfluentti keräysaikana) elektroporaatiolla kahdessa eri transfektiossa ("EP1" ja "EP2"). Transfektiossa EP1 käytimme 20 µg pXB322hGH:ta (reportteriplasmi-  
 15 di) sekä 380 µg sonikoitua taimenen sperman kantaja-DNA:ta (Pharmacia LKB, Piscataway, NJ). Transfektiossa EP2 käytimme 20 µg pXB322hGH:ta sekä 30 µg pAHE2:ta (E2-transaktivaattori) ja 350 µg taimenen sperman kantaja-DNA:ta. Suoritimme elektroporaatiot Bio-Rad Gene Pulser -laitteel-  
 20 la, 270 voltia, 960 µFD, noin 11 millisekunnin pulssi-ajalla. Elektroporaatioiden jälkeen jaoimme solut 6-kuoppalevyille,  $2 \times 10^5$  solua per kuoppa. Viisi tuntia elektroporaatioiden jälkeen imimme pois kasvatuselatusaineen, huuhtelimme solut kasvatuselatusaineella ja lisäsimme  
 25 1,5 ml tuoretta kasvatuselatusainetta kuhunkin kuoppaan. Tässä vaiheessa lisäsimme klorokiinia ("CQ") niin että loppukonsentraatio oli 80 µM (tai kontrolleihin tyhjä liuos). Sitten lisäsimme silloitetun tat37-72-E2.123:n ("TxHE2") tai silloitetun tat37-72-E2.123CCSS:n  
 30 ("TxHE2CCSS"). Näiden kuljetuspolypeptidi-lastikonjugaattien lopullinen konsentraatio oli 6, 20 tai 60 µg/ml solukasvatuselatusainetta (taulukko I).

Taulukko I  
Näytteiden nimeäminen

	Kuoppa	CO ( $\mu\text{M}$ )	Proteiini ( $\mu\text{g/ml}$ )
5	EP1.1	0	0
	EP1.2	80	0
10	EP2.1	0	0
	EP2.2	0	6 TxHE2
	EP2.3	0	20 TxHE2
	EP2.4	0	60 TxHE2
	EP2.5	0	6 TxHE2CCSS
	EP2.6	0	20 TxHE2CCSS
	EP2.7	0	60 TxHE2CCSS
	EP2.8	80	0
	EP2.9	80	6 TxHE2
	EP2.10	80	20 TxHE2
	EP2.11	80	60 TxHE2
	EP2.12	80	6 TxHE2CCSS
	EP2.13	80	20 TxHE2CCSS
	EP2.14	80	60 TxHE2CCSS

18 tunnin inkuboinnin jälkeen poistimme elatusai-  
neen, huuhtelimme solut tuoreella elatusaineella ja lisä-  
simme 1,5 ml tuoretta elatusainetta, joka sisälsi samat  
20 klorokiinin ja kuljetuspolypeptidi-lastikonjugaattien kon-  
sentraatiot kuin edeltävä 18-tuntinen inkubaatio. Tämä  
elatusaineen vaihto tehtiin mahdollisen hGH:n poistamiseksi,  
joka on mahdollisesti ollut läsnä ennen kuin represso-  
ri meni sisään soluihin. 24 tuntia elatusaineen vaihdon  
25 jälkeen keräsimme solut ja suoritimme solulaskennat viabi-  
liteetin tarkistamiseksi. Sitten määritimme hGH:n laimen-  
tamattomista kasvuelatusainenäytteistä Seldonin menetelmän  
mukaisesti, joka on kuvattu teoksessa Protocols in Molecu-  
lar Biology, Green Publishing Associates, New York, s.  
30 9.7.1 - 9.7.2 (1987), käyttämällä Allegro Human Growth  
Hormone -tarvikesarjaa lyhytkestoista geeniekspressiojär-  
jestelmää varten (Nichols Institute, San Juan Capistrano,  
CA). Vähensimme määrityksen taustan (se on: määrityskompo-  
nentit, joihin oli lisätty muokkaamatonta elatusainetta)  
35 hGH:n cpm:stä kaikkien näytteiden tapauksessa. Suoritimme

erilliset repressiolaskut prosenttiosuuksina kullekin proteiinikäsittelylle sen mukaan oliko klorokiiniä läsnä ("(+)"CQ") vai poissa ("(-)"CQ") proteiinin sisäänottotestissä. Laskimme repression prosenttiosuuden seuraavan kaavan mukaisesti:

$$\text{Repressio} = \frac{(\text{ACT} - \text{BKG}) - (\text{REP} - \text{BKG})}{\text{ACT} - \text{BKG}} \times 100$$

10

jossa:

BKG = hGH:n cpm pelkän reportterin transfektioissa (esim: EP1.1 (-)CQ:lle ja EP1.2 (+)CQ:lle);

15 ACT = hGH:n cpm reportterin ja transaktivaattorin transfektiossa, mutta missä ei lisätty repressorikonjugaattia (esim: EP2.1 (-)CQ:lle ja EP2.8 (+)CQ:lle);

REP = hGH:n cpm reportterin ja transaktivaattorin transfektiossa, missä lisättiin repressorikonjugaatti (esim: EP2.2 - 2.7 (-)CQ:lle ja EP2.9 - 2.14 (+)CQ:lle.

20

Tulokset edustavasta E2-repressiomäärityksestä esitetään taulukossa II. Taulukko I nimeää taulukossa II esitetyt eri näytteet. Kuvio 7 kuvaa graafisesti taulukossa II esitetyt tulokset.

Taulukko II  
E2-repressimääritys

	Näyte	hGH:n cpm	cpm - mää- rityksen tausta	cpm - PKG	Repressio- prosentti
5	EP1.1	3958	3808	--	--
	EP1.2	5401	5251	--	--
	EP2.1	15 161	15 011	11,203	--
	EP2.2	12 821	12 671	8863	20,9
10	EP2.3	10 268	10 118	6310	43,7
	EP2.4	8496	8346	4538	59,5
	EP2.5	11 934	11 784	7976	28,8
	EP2.6	9240	9090	5282	52,9
	EP2.7	7926	7776	3968	64,6
15	EP2.8	15 120	14 970	9719	--
	EP2.9	12 729	12 579	7328	24,6
	EP2.10	9590	9440	4189	56,9
	EP2.11	8440	8290	3039	68,7
	EP2.12	11 845	11 695	6444	33,7
	EP2.13	8175	8025	2774	71,5
20	EP2.14	6697	6547	1296	86,7

Kuljetuspolypeptidi tat37-72, joka oli silloitettu joko E2-repressoriin (E2.123 tai E2.123CCSS) johti E2-riippuvaisen geeniekspression annosriippuvaiseen inhibi-  
 25 tioon viljellyissä nisäkässoluissa (taulukko II; kuvio 7). Olemme toistaneet tämän kokeen neljä kertaa samanlaisilla tuloksilla. Vaikutus oli E2-spesifinen sikäli että tat37-72-konjugaateilla ei ollut vaikutusta pXB332hGH:n E2-in-  
 30 duktioon (tuloksia ei esitetä). Tat37-72xHE2-konjugaateilla ei myös ollut vaikutusta sellaisen reportterin hGH-ekspressiotasoon, jossa hGH-geenin ekspressiota ohjasi kons-  
 titutiivinen promoottori, jossa ei ollut vastetta E2:n suhteen. E2-repressori, jossa oli CCSS-mutaatio, represso-  
 35 suuremmassa määrin kuin repressori, jossa oli villityyppi-  
 nen aminohapposekvenssi. Tämä oli odotusten mukaista, kos-

5 ka kuljetuspolypeptidin silloitus kumpaan tahansa viimeisest<sup>ä</sup> kahdesta kysteinistä villityyppisessä repressorissa todennäköisesti vähentäisi tai eliminoisi repressorin aktiivisuutta. Klorokiiniä ei vaadittu repressioaktiivisuuteen. Klorokiini kuitenkin todella tehosti repressiota kaikissa testeissä. Nämä tulokset esitetään yhteenvetona taulukossa II ja kuviossa 7.

#### Esimerkki 10

##### TATΔCYS-konjugaatit

##### 10 TATΔcys:n muodostaminen

Sellaisen kuljetuspolypeptidin tuotantoa varten bakteerissa, joka sisältäisi tat-aminohapot 1 - 21, jotka olisi fuusioitu suoraan tat-aminohappoihin 38 - 72, muodostimme ekspressioplasmidin pTATΔcys (kuvio 8; SEQ ID nro 15 20). Plasmidin pTATΔcys muodostamiseksi käytimme tavanomaisia PCR-tekniikoita, joissa plasmidi pTAT72 oli PCR-templaattina. Yksi PCR:ään käytetyistä oligonukleotidialukkeista oli 374.18 (SEQ ID nro 19), joka kattaa EagI-kohdan tat:n koodittavasta sekvenssistä ylävirtaan. (Käytimme myös oligonukleotidia 374.18 plasmidin pET8c-123CCSS 20 muodostamiseen. Katso esimerkki 9.) Toinen oligonukleotidialuke PCR:lle, 374.28, kattaa EagI-kohdan tat:n koodittavan sekvenssin sisällä, ja sillä on deletio tat-DNA-sekvenssissä, joka koodittaa aminohappoja 22 - 37. 25 374.28:n nukleotidisekvenssi on: TTTACGGCCG TAAGAGATAC CTAGGGCTTT GGTGATGAAC GCGGT (SEQ ID nro 21). Digestoimme PCR-tuotteet EagI:llä ja eristimme tuloksena olevan 762 emäsparin fragmentin. Sijoitimme tämän EagI-fragmentin 4057 emäsparin vektoriin, joka tuotettiin katkaisemalla 30 pTAT72 EagI:llä. Varmistimme konstruktion DNA-sekvenssi-analyysillä ja ekspressoimme tatΔcys-polypeptidin Studierin ym. menetelmällä (edellä). SDS-PAGE-analyysi osoitti, että tatΔcys-polypeptidillä on oikea koko.

35 TatΔcys-proteiinin puhdistusta varten sulatimme 4,5 grammaa pTATΔcys-transformoituja *E. coli* -soluja, resusp-

endoimme solut 35 ml:aan 20-millimolaarista MES:iä (pH 6,2), joka oli 0,5-millimolaarinen EDTA:n suhteen. Hajotimme solut viemällä ne kahdesti "French press"-laitteen läpi 10 000 psi:n (68 000 kPa) paineessa. Poistimme liu-

5 kenemattoman roskan sentrifugoimalla 10 000 rpm:n nopeudella SA600-roottorissa 1 tunnin ajan. Panimme supernatantin 5 ml:n S Sepharose Fast Flow -pylvääseen nopeudella 15 ml/tunti. Pesimme pylvään 50-millimolaarisella Tris-HCl:llä (pH 7,5), joka oli 0,3-millimolaarinen DTT:n

10 suhteen. Sitten suoritimme vaiheittaisen gradienttiseluutuksen (2 ml/vaihe) samalla puskurilla, joka oli 300-, 400-, 500-, 700- ja 950-millimolaarinen NaCl:n suhteen. Tat $\Delta$ cys-proteiini eluoitui 950-millimolaarisessa NaCl-fraktiossa.

Konjugoimme tat $\Delta$ cys-kuljetuspolypeptidin rodamiiniisotiosyanaattiin ja testasimme sen määrittämällä suoraan otto solun sisään. Tulokset olivat positiivisia (samanlaisia kuin tulokset vastaavissa kokeissa tat1-72:lla).

#### Geneettinen TAT $\Delta$ cys-249-fuusio

Tat $\Delta$ cys-kuljetuspolypeptidin, joka olisi fuusioitu geneettisesti natiivin E2-repressoriproteiinin aminopäähän (se on: BPV-1 E2:n 249 karboksiterminaaliseen aminohappoon), bakteeriekspressiota varten muodostimme plasmidin pTAT $\Delta$ cys-249 seuraavasti. Muodostimme plasmidin pFTE501 (kuvio 9) plasmideista pTAT72 (Frankel ja Labo, edellä) ja

20 pXB314 (Barsoum ym. edellä). Plasmidista pXB314 eristimme NcoI-SpeI-DNA-fragmentin, joka koodittaa 249 aminohapon BPV-1 E2-repressorin. (NcoI katkaisee nukleotidista 296 ja SpeI katkaisee pXB314:n nukleotidista 1118.) Teimme tämän fragmentin päät tasaisiksi DNA polymeraasi I Klenow -kä-

25 sittelyllä ja lisäsimme kaupallisesti saatavilla olevan BglII-linkkerin (New England Biolabs, luettelonro 1090). Sijoitimme tämän linkkerin sisältävän fragmentin BamHI-katkaistuun (täydellinen digestio) pTAT72-plasmidiin. pTAT72:ssa on BamHI-katkaisukohta tat:ia koodittavan alueen sisällä, lähellä sen 3'-päättä, ja toinen BamHI-kat-

30

35

kaisukohta hieman alavirtaan tat-geenistä. BgIII-linkkeri liitti tat:ia ja E2:ta koodittavat sekvenssit samaan luku-kehykseen, jolloin ne koodittivat fuusiota, jossa oli tat-proteiinin 62 ensimmäistä aminohappoa, joita seurasi seriinitähde ja BPV-1 E2-proteiinin 249 viimeistä aminohappoa. Annoimme tälle bakteeriekspressioplasmidille nimeksi pFTE501 (kuvio 9). Plasmidin pTAT $\Delta$ cys-249 muodostamiseksi (kuvio 10; SEQ ID nro 22) sijoitimme 762 emäsparin EagI-fragmentin plasmidista pTATcys, joka sisältää osan tat:ia, joka sisältää kysteiinidelektion, plasmidin pFTE501 4812 emäsparin EagI-fragmenttiin.

#### Tat $\Delta$ cys-249:n puhdistus

Sulatimme 5 g *E. colia*, joka ekspressoi tat $\Delta$ cys-249:tä ja suspendoimme solut 40 ml:aan puskuria, joka oli 25-millimolaarinen Tris-HCl:n (pH 7,5), 25-millimolaarinen NaCl:n, 0,5-millimolaarinen EDTA:n, 5-millimolaarinen DTT:n suhteen ja joka sisälsi proteaasin inhibiittoreita (1,25 mM PMSF, 3 mM bentsamiidini, 50  $\mu$ g/ml pepstatiini A, 50  $\mu$ g/ml aprotiniini, 4  $\mu$ g/ml E64). Hajotimme solut viemällä ne kahdesti "French"-paine-kammion läpi 10 000 psi:n (68 000 kPa) paineessa. Poistimme liukenemattoman roskan lysaatista sentrifugoimalla 12 000 rpm:n nopeudella SA600-roottorilla 1 tunnin ajan. Puhdistimme tat $\Delta$ cys-249:n liukoisesta fraktiosta. Supernatantti ladattiin 2 ml:n S Sepharose Fast Flow -pylvääseen (Pharmacia LKB, Piscataway, NJ) virtausnopeudella 6 ml/h. Pylväs pestiin puskurilla, joka oli 25-millimolaarinen Tris-HCl:n (pH 7,5), 25-millimolaarinen NaCl:n, 0,5-millimolaarinen EDTA:n ja 1-millimolaarinen DTT:n suhteen, ja käsiteltiin peräkkäisillä suolavaiheilla samassa puskurissa, joka oli 100-, 200-, 300-, 400-, 500-, 600- ja 800-millimolaarinen NaCl:n suhteen. Otimme talteen tat $\Delta$ cys-249:n 600 - 800 mM suolafraktioista. Yhdistimme huippufraktiot, teimme sen 15-prosenttiseksi glyserolin suhteen ja varastoimme erät -70 °C:n lämpötilaan.



### Immunofluoresenssimääritys

Jotta saisimme analysoitua tat $\Delta$ cys-E2-repressorifuusioproteiinin oton solun sisään, käytimme epäsuoria immunofluoresenssitekniikoita. Jaoimme HeLa-soluja "cover slip" -lasien päälle 6-kuoppaisissa kudosisviljelylevyissä, 5  
niin että ne olivat 50-prosenttisen konfluentteja. Yli yön inkuboinnin jälkeen lisäsimme tat $\Delta$ cys-E2-repressorifuusioproteiinin (loppukonsentraatio 1  $\mu$ g/ml) ja klorokiinin (loppukonsentraatio 0,1-millimolaarinen). Kuuden tunnin 10  
kuluttua poistimme fuusioproteiinin/klorokiinin sisältävän kasvatuselatusaineen ja pesimme solut kahdesti PBS:llä. Kestävöimme pestyt solut 3,5-prosenttisellä formaldehydillä huoneenlämmössä. Permeabilisoimme kestävöidyt solut PBS:llä, joka oli 0,2-prosenttinen Triton-X-100:n, 2-prosenttinen 15  
naudan seerumialbumiinin ("BSA") suhteen ja 1-millimolaarinen MgCl<sub>2</sub>:n ja 0,1-millimolaarinen CaCl<sub>2</sub>:n suhteen ("PBS+") 5 minuutin ajan huoneenlämpötilassa. Permeabilisoitujen solujen suojaamiseksi käsittelimme ne PBS:llä, joka oli 2-prosenttinen BSA:n suhteen, 1 tunnin 20  
ajan 4 °C:n lämpötilassa.

Inkuboimme "cover slip" -laseja 20  $\mu$ l:lla primääri-vasta-aine-liuosta kussakin kuopassa, 1:100-laimennoksena PBS+:ssa, joka oli 2-prosenttinen BSA:n suhteen, 1 tunnin ajan 4 °C:n lämpötilassa. Primäärinen vasta-aine oli joko 25  
kanin polyklonaalinen vasta-aine BPV-1 E2-repressorille (saatu aikaan injektoimalla puhdistetut E2:n karboksiterminaaliset 85 aminohappoa) tai kanin polyklonaalinen vasta-aine tat:ia vastaan (saatu aikaan injektoimalla tat:n puhdistetut aminotermiinaaliset 72 aminohappoa). Lisäsimme 30  
sekundäärisen vasta-aineen 1:100-laimennoksena PBS+:ssa, joka oli 0,2-prosenttinen Tween-20:n ja 2-prosenttinen BSA:n suhteen, 30 minuutiksi 4 °C:n lämpötilassa.

Sekundäärinen vasta-aine oli rodamiinikonjugoitu, kanin IgG:tä vastaan konjugoitu vuohen vasta-aine (Cappel, 35  
nro 2212-0081). Sen jälkeen kun soluja oli inkuboitu se-

kundäärisen vasta-aineen kanssa, pesimme solut PBS+:lla, joka oli 0,2-prosenttinen Tween 20:n ja 2-prosenttinen BSA:n suhteen, ja sijoitimme "cover slip"-lasit liuoksen päälle, joka oli 90-prosenttinen glyserolin, 25-millimolaarinen natriumfosfaatin (pH 7,2) ja 150-millimolaarinen NaCl:n suhteen. Tutkimme solut fluoresenssimikroskoopilla, jossa oli rodamiinisuodatin.

#### Tat $\Delta$ cys-fuusioiden otto soluun

Havaitsimme, että tat $\Delta$ cys-E2-repressorifuusioproteiinia otettiin merkittävästi solun sisään, kun käytimme joko tat-vasta-ainetta tai E2-vasta-ainetta. Kontrollisoluissa, jotka oli altistettu konjugoimattomalle tat-proteiinille, havaitsimme solunsisäistä fluoresenssia kun käytimme tat-vasta-ainetta, mutta ei E2-vasta-ainetta. Kontrollisoluissa, jotka oli altistettu seokselle, jossa oli konjugoimaton E2-repressorin ja tat-proteiini tai tat $\Delta$ cys, havaitsimme fluoresenssia kun käytimme tat-vasta-ainetta, mutta ei E2-vasta-ainetta. Tämä vahvisti, että tat välittää E2-repressorin sisäänottoa vain silloin, kun tämä on liitetty tat-proteiiniin. Kuten konjugoimattomalla tat-proteiinilläkin, havaitsimme tat $\Delta$ cys-E2-repressorifuusioproteiinia kaikkialla soluissa, mutta se oli konsentroitunut solunsisäisiin vehikkeleihin. Nämä tulokset osoittavat, että tat-peräinen polypeptidi, josta täysin puuttuvat kysteinitähteet, voi viedä heterologisen proteiinin (se on: kuljetuspolyeptidin ja lastiproteiinin geneettisen fuusion) eläinsoluihin.

Samanlaisessa menetelmässä kuin edellä kuvattiin, tuotimme tat $\Delta$ cys:n geneettisen fuusion HPV E2:n 123 C-terminaaliseen aminohappoon. Kasvatuselatusaineeseen lisätynä tämä fuusiopolypeptidi osoittautui repressoivan E2-riippuvaista geeniekspressiota COS7-soluissa (tuloksia ei näytetä).

**Esimerkki 11****"Antisense"-oligodeoksinukleotidikonjugaatit**

Syntetisoimme DNA-fosforitionaattianalogeja (pituusdeltaan 4 - 18 nukleotidia), jotka kukin sisälsivät 5'-päässä vapaan aminoryhmän, käyttämällä automatisoitua DNA/RNA-syntetisaattoria (Applied Biosystems model 394). Amiiniryhmä inkorporoitiin oligonukleotideihin käyttämällä kaupallisesti modifioituja nukleotideja ("aminolink 2", Applied Biosystems). Oligonukleotidit vastasivat "sense"- ja "antisense"-juosteita ihmisen kasvuhormoni-RNA:n ja CAT-lähetti-RNA:n alueilta.

Kutakin silloitusreaktiota varten liuotimme 200 µg oligonukleotidia 100 µl:aan 25-millimolaarista natriumfosfaattipuskuria (pH 7,0). Sitten lisäsimme 10 µl sulfo-SMCC:n 50-millimolaarista kantaliuosta ja annoimme reaktion edetä huoneenlämpötilassa 1 tunnin ajan. Poistimme reagoimattoman sulfo-SMCC:n geelisuodattamalla reaktioseoksen P6DG-pylväässä (Bio-Rad) 25-millimolaarisessa HEPES-puskurissa (pH 6,0). Kuivasimme oligonukleotidi-sulfo-SMCC-aduktin alipaineessa. Oligonukleotidien saalis tässä menetelmässä oli välillä 58 - 95 %. Kuljetuspolypeptidin kanssa tapahtuvaa reaktiota varten liuotimme kukin oligonukleotidi-sulfo-SMCC-aduktin 50 µl:aan 0,5-millimolaarista EDTA:ta, siirsimme liuoksen koeputkeen, joka sisälsi 50 µg lyofilisoitua kuljetuspolypeptidiä, ja annoimme reaktion edetä huoneenlämpötilassa 2 tunnin ajan. Analysoimme reaktiotuotteet SDS-PAGE:ssa.

**Esimerkki 12****Vasta-aine-konjugaatit****30 Anti-tubuliinikonjugaatti 1**

Hankimme tubuliinia vastaan suunnattua, hiiren kaupallista monoklonaalista IgG1:tä (Amersham) ja puhdistimme sen askitesnesteestä käyttämällä proteiini A:ta tavanomaisin menetelmin. Leimasimme puhdistetun vasta-aineen rodamiini-isotiosyanaatilla käyttämällä 1,2 moolia rodamiinia

per mooli vasta-ainetta. Kun altistimme kestäväidyt, permeabilisoidut HeLa-solut leimatulle vasta-aineelle, tutkiminen mikroskoopilla paljasti kirkkaasti värjäytyneitä mikrotubuleita. Vaikka rodamiinileimaus oli riittävä, tehostimme vasta-aineen signaalia hiiren immunoglobuliinia vastaan suunnatulla, FITC-leimatulla vasta-aineella.

Menetelmässä, joka oli oleellisesti esimerkissä 2 kuvatun kaltainen (edellä), annoimme vasta-aineen (250 µg) reagoida sulfo-SMCC:n kanssa (10:1 molaarinen ylimäärä). Sitten lisäsimme 48 µg (<sup>35</sup>S-leimattua) tat1-72:ta. Tat1-71:n ja vasta-aineen molaarinen suhde oli 2,7:1. Radioaktiivisuuden inkorporaation perusteella tat1:72 oli silloittunut vasta-aineeseen suhteessa 0,6:1.

Tat1-72:n ja vasta-aineen muodostaman konjugaatin sisäänmenon analysoimiseksi lisäsimme konjugaatin elatusaineeseen (10 µg/ml), joka huuhtoi "coverslip"-laseilla kasvatettuja soluja. Havaitimme solussa pistemäisen fluoresenssikuvion. Pistemäinen kuvio viittasi konjugaatin sijaintiin vehikkeleissä, ja siitä ei näin ollen voitu tehdä sytoplasmaan joutumista koskevia johtopäätöksiä.

Jotta saisimme osoitettua konjugoidun vasta-aineen immunologisen reaktiivisuuden, testasimme sen kykyä sitoa tubuliinia. Liitimme puhdistetun tubuliinin syano-geenibromidilla aktivoituun Sepharose 4B:hen (Sigma Chem. Co., St. Louis, MO). Panimme radioaktiivisen konjugaatin sisältävät näytteet tubuliinipylvääseen (ja Sepharose 4B-kontrollipylvääseen) ja mittasimme sitoutuneen konjugaatin määrän. Affiniteettimatriisiin sitoutui enemmän radioaktiivisuutta kuin kontrollipylvääseen, mikä viittasi tubuliinia sitovaan aktiivisuuteen.

#### Tubuliinia vastaan suunnattu konjugaatti 2

Erillisessä silloituskokeessa hankimme tubuliinia vastaan suunnatun rotan monoklonaalisen IgG2a-vasta-aineen (Serotec) ja puhdistimme sen askitesnesteestä tavanomaisin menetelmin käyttämällä proteiini G:tä. Eluoimme vasta-ai-

neen Caps-puskurilla (pH 10). Puhdistettu vasta-aine oli positiivinen tubuliininsitoutumismäärityksessä. Annoimme tat1-72:n reagoita rodamiini-isotiosyanaatin kanssa molaarisessa suhteessa 1:1. Reaktiotuotteen  $A_{550}/A_{280}$ -suhde oli 0,63, mikä viittasi siihen että noin 0,75 moolia väriä oli substituotunut per mooli tat1-72:ta. Kun reagoimaton väri erotettiin tat1-72-rodamiinista G-25-geelisuodatuksella (Pharmacia LKB, Piscataway, NJ), havaitsimme että vain 52 µg 150 µg:sta tat1-72:ta käytettiin reaktiossa.

Säästimme erän tat1-72-rodamiinista käytettäväksi (kontrollina) soluunottokokeissa, ja lisäsimme loput 0,4 mg:aan vasta-ainetta, joka oli reagoinut SMCC:n kanssa (20:1). Reaktioseos sisälsi tat1-72:vasta-aine-suhteen noin 1:1, pikemminkin kuin odotettu 5:1. (Seuraavassa kokeessa 5:1-suhde osoittautui olevan epätyytyttävä, ja sai aikaan sakan). Annoimme silloitusreaktion edetä yli yön 4 °C:n lämpötilassa. Sitten lisäsimme molaarisen ylimäärän kysteiiniä jäljellä olevien maleimidiryhmien suojaamiseksi ja näin silloitusreaktion pysäyttämiseksi. Sentrifugoimme reaktioseokset poistaaksemme mahdollisen läsnä olevan sakan.

Suoritimme elektroforeesin käyttäen 4 - 20-prosenttista polyakryyliamidigradianttigeeliä analysoidaksemme supernatantin pelkistävässä ja ei-pelkistävässä olosuhteissa. Analysoimme myös sakan tällä menetelmällä. Vasta-aineella tat-1-72 (ilman rodamiinia) tehdyistä konjugaatiokokeista peräisin olevissa supernatanteissa havaitsimme hyvin vähän materiaalia 4 - 20-prosenttisessä geelissä. Supernatanteissa, jotka olivat peräisin vasta-aine-tat-1-72-rodamiinin konjugaatiokokeista, havaitsimme kuitenkin suhteellisen raskaita juovia vasta-aineen yläpuolella pelkistyneen näytteen tapauksessa. Vasta-aine vaikutti olevan konjugoitunut tat-1-72:een noin 1:1 suhteessa.

Soluunottokokeissa, jotka suoritettiin konjugaatilla 2 (menetelmä kuten edellä kuvattiin konjugaatille 1),

saimme tuloksia, jotka olivat samanlaisia kuin jotka saatiin konjugaatille 1. Kun konjugaatti visualisoitiin rodaminiifluoresenssilla tai toiseen vasta-aineeseen assosioituneella fluoresiinilla, havaitsimme konjugaatin vehikkeleissä.

### Esimerkki 13

#### Muita tat-E2-konjugaatteja

#### Kemiallisesti silloitetut tat-E2-konjugaatit

Silloitimme kuljetuspolyeptidin tat37-72 E2:n neljään eri repressorimuotoon. Näissä kokeissa käytetyt E2-repressorin neljä osaa olivat BPV-1:n 103 karboksiterminaalista tähdettä (se on: 308 - 410) ("E2.103"), BPV-1:n 249 karboksiterminaalista tähdettä (se on: 162 - 410) ("E2.249"), HPV-16:n 121 karboksiterminaalista tähdettä (se on: 245 - 365) ("HE2"), ja HPV-16:n 121 karboksiterminaalista tähdettä, jossa kysteiinitähteet kohdissa 300 ja 309 vaihdettiin seriiniksi, ja lysiinitähde kohdassa 299 vaihdettiin arginiiniksi ("HE2CCSS"). HE2:n ja HE2CCSS:n rekombinanttituotanto ja puhdistus, jota seurasi HE2:n ja HE2CCSS:n kemiallinen silloitus tat37-72:een, jolloin muodostui TxHE2 ja vastaavasti TxHE2CCSS, kuvataan esimerkissä 9 (edellä). E2.103:n ja E2.249:n kemialliseksi silloittamiseksi tat37-72:een (jolloin saatiin aikaan konjugaatit, joille annettiin nimeksi TxHE2.103 ja TxHE2.249), käytimme samaa menetelmää, jota käytettiin tekemään TxHE2 ja TxHE2CCSS (esimerkki 9 edellä).

Ekspressoimme proteiinin E2.103 *E. coli*ssa plasmidista pET-E2.103. Hankimme pET-E2.103:n PCR-kloonausmenetelmällä, joka oli analoginen sen kanssa, jota käytettiin pET8c-123:n tuottamiseksi, joka on kuvattu esimerkissä 9 (edellä) ja kuviossa 5. Kuten pET8c-123:n muodostamisessakin, ligoimme PCR:llä tuotetun NcoI-BamHI-E2-fragmentin NcoI-BamHI-katkaistuun pET8c:hen. Meidän PCR-templaattimme E2-fragmentille oli plasmidi pCO-E2 (Hawley-Nelson ym., *EMBO J.*, vol. 7, s. 525 - 31 (1988), US-pa-

tentti nro 5 219 990). Oligonukleotidialukkeet, joita käytettiin E2-fragmentin tuottamiseksi pCO-E2:sta, olivat EA21 (SEQ ID nro 36) ja EA22 (SEQ ID nro 37). Aluke EA21 toi NcoI-kohdan, joka lisäsi metioniinikodonin, jota seurasi alaniinikodoni, joka oli 5'-suunnassa sellaisen alueen vieressä, joka kooditti BPV-1 E2:n 101:tä karboksiterminaalista tähdettä.

Ekspressoimme proteiinin E2.249 *E. coli*ssa plasmidista pET8c-249. Muodostimme pET8c-249:n sijoittamalla 1362 emäsparin NcoI-BamHI-fragmentin plasmidista pXB314 (kuvio 9) NcoI-BamHI-katkaistuun pET8c:hen (kuvio 5).

#### Geneettiset TATΔcys-BPV E2-fuusiot

TATΔcys-249:n lisäksi testasimme useita muita TATΔcys-BPV-1 E2 -repressorifuusioita. Plasmidi pTATΔcys-105 kooditti tat-tähteitä 1 - 21 ja 38 - 67, joita seurasivat BPV-1:n karboksiterminaaliset 105 tähdettä. Plasmidi pTATΔcys-161 kooditti tat-tähteitä 1 - 21 ja 38 - 62, joita seurasivat BPV-1:n 161 karboksiterminaalista tähdettä. Muodostimme plasmidit pTATΔcys-105 ja pTATΔcys-161 väli-  
 20 tuoteplasmideista pFTE103 ja vastaavasti pFTE403.

Tuotimme pFTE103:n ja pFTE403:n (samoin kuin pFTE501:n) ligoimalla eri insertit BamHI-katkaistuun (täydellinen digestio) vektoriin pTAT72.

Jotta saisimme insertiofragmentin pFTE103:lle, eristimme 929 emäsparin PleI-BamHI-fragmentin pXB314:stä ja ligoimme sen kaksijuosteiseen linkkeriin, joka koostui synteettisestä oligonukleotidista FTE.3 (SEQ ID nro 23) ja synteettisestä oligonukleotidista FTE.4 (SEQ ID nro 24). Linkkeri kooditti tat-tähteitä 61 - 67 ja siinä oli yksijuosteisena yli tuleva BamHI-osa 5'-päässä ja yksijuosteisena yli tuleva PleI-osa 3'-päässä. Ligoimme linkkerin sisältävän fragmentin pXB314:stä BamHI-katkaistuun pTAT72:een, jolloin saatiin pFTE103. Jotta saisimme insertiofragmentin pFTE403:een, digestoimme pXB314:n NcoI:llä ja SpeI:llä, teimme tasapäätt Klenow-käsittelyllä ja li-  
 35

5 goimme BglIII-linkkerin, joka koostui GAAGATCTTC:stä (New England Biolabs, Beverly, MA, luettelonro 1090) (SEQ ID nro 35), joka oli muodostanut dupleksin itsensä kanssa. Puhdistimme tuloksena olevan 822:n emäsparin fragmentin elektroforeesilla ja sitten ligoimme sen BamHI-digestoi-

10 tuun pTAT72-vektoriin, jolloin saimme PFTE403:n.  
Jotta saisimme deletoitua tat-tähteet 22 - 37, jolloin saatiin plasmidi pTATΔcys-105 pFTE103:sta ja pTATΔcys-161 pFTE403:sta, käytimme samaa menetelmää (kuvattu edellä), jota käytettiin hankkimaan plasmidi pTATΔcys-249 pFTE501:stä.

#### Geneettiset TATΔcys-HPV E2-fuusiot

15 Muodostimme plasmidit pTATΔcys-HE2.85 ja pTATΔcys-HE2.121, koodittaaksemme fuusioproteiinin, joka koostuisi tatΔcys-kuljetusosasta (tat-tähteet 1 - 21, 38 - 72), jota seuraisi HPV-16:n 85 tai vastaavasti 121 karboksiterminaalista tähdettä.

20 Meidän lähtöplasmidimme pTATΔcys-HE2.85:n ja pTATΔcys-HE2.121:n muodostamisessa olivat pTET8c-85 ja vastavasti pET8c-123 (molemmat kuvattu edellä). Digestoimme pET8c-85:n ja pET8c-123:n BglIII:lla ja NcoI:llä, ja eristimme ison fragmentin molemmissa tapauksissa (4769 emäsparia pET8c-85:stä tai 4880 emäsparia pET8c-123:sta) käytettäväksi vektorina. Molemmissa vektoreissa E2:ta koodittavat alueet alkavat NcoI-kohdasta. Sijoitimme molempiin vektoreihin 220 emäsparin BglIII-AatII-fragmentin plasmidista pTATΔcys, ja synteettisen fragmentin. BglIII-AatII-fragmentin 5'-pää on ylävirtaan T7-promoottorista ja koodittaa tatΔcys:n ensimmäistä 40 tähdettä (se on: tähteet 1 - 21, 38 - 56). Synteettinen fragmentti, joka koostui hybridisoituneista oligonukleotideista 374.67 (SEQ ID nro 25) ja 374.68 (SEQ ID nro 26), kooditti tat-tähteitä 57 - 72, jossa oli yksijuosteisena yli tuleva AatII-osa 5'-päässä ja yksijuosteisena yli tuleva NcoI-osa 3-päässä.



### Sarja geneettisiä JB-fuusioita

Plasmidi pJB106 koodittaa fuusioproteiinia (kuvio 12) (SEQ ID nro 38), jossa aminoterminaalista metioniini-tähdettä seuraavat tat-tähteet 47 - 58 ja sitten HPV-16:n E2-tähteet 245 - 365. Jotta saisimme aikaan pJB106:n, suoritimme kolmivaiheisen ligaation, joka kuvataan kaavamaisesti kuviossa 11. Saimme aikaan 4602 emäsparin vektorifragmentin digestoimalla plasmidin pET8c NcoI:llä ja BamHI:llä. Yksi insertti oli 359 emäsparin MspI-BamHI-fragmentti pET8c-123:sta, joka kooditti HPV-16 E2-tähteitä 248 - 365. Toinen insertti oli synteettinen fragmentti, joka koostui hybridisoituneesta oligonukleotidiparista 374.185 (SEQ ID nro 27) ja 374.186 (SEQ ID nro 28). Synteettinen fragmentti kooditti aminoterminaalista metioniinia ja tat-tähteitä 47 - 58, sekä HPV16-tähteitä 245 - 247 (se on: ProAspThr). Synteettisellä fragmentilla oli yksijuosteisena yli tuleva NcoI-osa 5'-päässä ja yksijuosteisena yli tuleva MspI-osa 3'-päässä.

Hankimme plasmidit pJB117 (SEQ ID nro 59), pJB118 (SEQ ID nro 60), pJB119 (SEQ ID nro 61), pJB120 (SEQ ID nro 62) ja pJB122 (SEQ ID nro 63) PCR-deleetiokloonauksella, samanlaisella tavalla kuin käytettiin pTATΔcys:n tapauksessa (kuvattu edellä ja kuviossa 8). Muodostimme plasmidit pJB117 ja pJB118 delectoimalla pTATΔcys-HE2.121:n segmenttejä. Muodostimme plasmidit pJB119 ja pJB120 delectoimalla pTATΔcys-161:n segmenttejä. Kaikissa neljässä kloonauksessa käytimme PCR-aluketta 374.122 (SEQ ID nro 29) kattaaksemme HindIII-kohdan, joka sijaitsee alavirtaan tat-E2:n koodittavasta alueesta. Kussakin tapauksessa toinen aluke ulottui NdeI-kohtaan tatΔcys:n koodittavan sekvenssin alussa ja poisti kodonit tähteille tatΔcys:n alussa (se on: tähteet 2 - 21 ja 38 - 46 pJB117:lle ja pJB119:lle, ja tähteet 2 - 21 pJB118:lle ja pJB120:lle). Tähteiden 2 - 21 delectointia varten käytimme aluketta 379.11 (SEQ ID nro 30). Tähteiden 2 - 21 ja 38 - 46 dele-

tointia varten käytimme aluketta 379.12 (SEQ ID nro 31). PCR-reaktion jälkeen digestoimme PCR-tuotteet NdeI:llä ja HindIII:lla. Sitten kloonasimme tuloksena olevat restriktiofragmentit vektoriin pTATAcys-HE2.121, joka oli diges-

5 toitu aikaisemmin NdeI:llä ja HindIII:llä, jolloin saimme 4057 emäsparin reseptorifragmentin. Näin muodostimme ekspressioplasmideja, jotka koodittivat fuusioproteiineja, jotka koostuivat aminohappotähteistä seuraavasti:

10 JB117 = Tat47-72-HPV16 E2 245 - 365,  
 JB118 = Tat38-72-HPV16 E2 245 - 365,  
 JB119 = Tat47-62-BPV1 E2 250 - 410 ja  
 JB120 = Tat38-62-BPV1 E2 250 - 410.

Muodostimme PJB122:n, joka koodittaa tat-tähteitä 38 - 58, jota seuraa HPV16:n E2-tähteet 245 - 365 (se on: E2:n karboksiterminaaliset 121 aminohappoa), deletoimalla pJB118-kodoneista tat-tähteet 59 - 72. Suoritimme tämän deleetion PCR:llä käyttämällä alukkeita 374.13 (SEQ ID nro 32), joka kattaa AatII-kohdan tat:n koodittavan alueen sisällä, ja alukkeella 374.14 (SEQ ID nro 33), joka kattaa AatII-kohdan, joka sijaitsee hieman alavirtaan Tat-E2-geenin ainutkertaisesta HindIII-kohdasta. Digestoimme PCR-tuotteen AatII:lla ja eristimme tuloksena olevan restriktiofragmentin. Lopullisessa pJB122-muodostusvaiheessa sijoitimme eristetyn AatII-fragmentin AatII-digestoituun pJB118-vektoriin.

15  
20  
25

Pitäisi panna merkille, että meidän kaikissa viidessä edellä kuvatussa pJB-konstruktiossamme tat:n koodittavaa sekvenssiä edelsi metioniinikodoni translaation initiaatiota varten.

### 30 Tat-E2-fuusioproteiinien puhdistus

Käytimme kaikissa tapauksissa E. colia ekspresoidaksemme geneettiset tat-E2-fuusiomme. Meidän yleinen menetelmämme tat-E2-proteiinin puhdistamiseen käsitti seuraavat alkuvaiheet: solujen ajamisen sakaksi, niiden resuspendoinnin 8 - 10 tilavuuteen lyysipuskuria (25-milli-

35

molaarinen Tris (pH 7,5), 25-millimolaarinen NaCl, 1-millimolaarinen DTT, 0,5-millimolaarinen EDTA), joka sisälsi proteaasin inhibiittoreita - yleensä 1 mM PMSF:n, 4 µg/ml E64:ää, 50 µg/ml aprotiniinia, 50 µg/ml pepstatiini A:ta, ja 3 mM bentsamidiinin), solujen hajotuksen "French press"-laitteessa (2 kertaa 12 000 psi (82 656 kPa)), ja lysaattien sentrifugoinnin 10 000 - 12 000 g:n kiihtyvyydessä 1 tunnin ajan (paitsi FTE-proteiinit), 4 °C:n lämpötilassa. Lisävaiheet, joita käytettiin tiettyjen tat-E2-fuusioproteiinien puhdistukseen, kuvataan jäljempänä.

E2.103 ja E2.249 - Lysaatin sentrifugoinnin jälkeen latasimme supernatantin Fast S Sepharose -pylvääseen ja eluimme E2.103- tai E2.249-proteiinin 1-molaarisella NaCl:llä. Sitten puhdistimme lisäksi E2.103:n tai E2.249:n kromatografiassa P60-geelisuodatuspylväessä, joka oli tasapainotettu 100-millimolaarisella HEPES-puskurilla (pH 7,5), joka oli 0,1-millimolaarinen EDTA:n ja 1-millimolaarinen DTT:n suhteen.

FTE103 - Sen jälkeen kun lysaattia oli sentrifugoitu 10 000 g:n kiihtyvyydessä 10 minuutin ajan 4 °C:n lämpötilassa, otimme talteen FTE103-proteiinin (joka sakkautui) resuspendoimalla sakan 6-molaarisella urealla ja lisäämällä kiinteää guanidiini-HCl:ää niin että loppukonsentraatio oli 7-molaarinen. Suspension sentrifugoinnin jälkeen puhdistimme FTE103-proteiinin supernatantista kromatografialla A.5M-geelisuodatuspylväessä 6-molaarisella guanidiinilla, joka oli 50-millimolaarinen natriumfosfaatin (pH 5,4) ja 10-millimolaarinen DTT:n suhteen. Keräsimme FTE103:n sisältävät fraktiot geelisuodatuspylvästä sen perusteella, että Coomassie-värjättyihin SDS-polyakryyliamidielektroforeesigeeleihin ilmestyi juova, jonka näennäinen molekyylipaino oli 19 kDa.

FTE403 - Meidän puhdistusmenetelmämme FTE403:lle oli oleellisesti sama kuin FTE103:lle, paitsi että FTE403

liikkui geelisuodatuspylväässä 25 kDa:n näennäisen molekyyli­painon mukaisesti.

FTE501 - Sen jälkeen kun lysaattia oli sentrifugoi-  
tu 10 000 g:n kiihtyvyydessä 30 minuutin ajan, resuspen-  
soimme sakan 6-molaarisessa ureassa, lisäsimme kiinteää  
5 guanidiini-HCl:ää niin että loppukonsentraatio oli 6 M, ja  
DTT:tä 10-millimolaariseksi konsentraatioksi. 30 minuutin  
jälkeen 37 °C:n lämpötilassa, kirkastimme liuoksen senti-  
fugoimalla 10 000 g:n lämpötilassa 30 minuutin ajan. Sit-  
10 ten latasimme näytteen A.5-agarosigeelisuodatuspylvääseen  
6-molaarisessa guanidiini-HCl:ssä, joka oli 50-millimolaa-  
rinen natriumfosfaatin (pH 5,4) ja 10-millimolaarinen  
DTT:n suhteen, ja keräsimme FTE501:n sisältävät fraktiot  
geelisuodatuspylväästä, sen perusteella, että Coomassie-  
15 värjättyihin SDS-polyakryyliamidielektroforeesigeeleihin  
ilmestyi juova, jonka näennäinen molekyyli­paino oli 40  
kDa. Latasimme geelisuodatuksella puhdistetun FTE501:n C<sub>18</sub>-  
käänteisfaasi-HPLC-pylvääseen ja eluimme gradientilla,  
jossa oli 0 - 75 % asetonitriili 0,1-prosenttisessa tri-  
20 fluorietikkahapossa. Keräsimme FTE501-proteiinin yksittäi-  
senä piikkinä, jonka näennäinen molekyyli­paino oli 40 kDa.

TatΔcys-105 - Lysaatin sentrifugoinnin jälkeen la-  
tasimme supernatantin Q-Sepharose-pylvääseen, joka oli  
tasapainotettu 25-millimolaarisella Tris:llä (pH 7,5),  
25 joka oli 0,5-millimolaarinen EDTA:n suhteen. Latasimme Q-  
Sepharose-pylväästä läpi tulleen nesteen S-Sepharose-pyl-  
vääseen, joka oli tasapainotettu 25-millimolaarisella  
MES:llä (pH 6,0), sen jälkeen kun Q-Sepharose-pylväästä  
läpi tullut nesteen pH oli säädetty noin 6,0:ksi lisäämäl-  
30 lä MES:iä (pH 6,0) niin että loppukonsentraatio oli 30 mM.  
Otimme talteen tatΔcys-105-proteiinin S-Sepharose-pylvääs-  
tä käyttämällä peräkkäisiä NaCl-konsentroidivaiheita 25-  
millimolaarisessa MES:ssä (pH 6,0). TatΔcys-105 eluoitui  
pH 6,0 -puskurissa 800 - 1 000 mM NaCl:lla.

Tat $\Delta$ cys-161 - Lysaatin sentrifugoinnin jälkeen la-  
 5 tasimme supernatantin S-Sepharose-pylvääseen, joka oli  
 tasapainotettu 25-millimolaarisella Tris:lla (pH 7,5),  
 joka oli 0,5-millimolaarinen EDTA:n suhteen. Otimme tal-  
 teen tat $\Delta$ cys-161:n S-Sepharose-pylväästä käyttämällä vai-  
 heittaista NaCl-gradienttia 25-millimolaarisessa Tris:ssa  
 (pH 7,5). Tat $\Delta$ cys-161 eluoitui pH 7,5 -puskurissa 500 -  
 700-millimolaarisella NaCl:lla.

Tat $\Delta$ cys-249 - Lysaatin sentrifugoinnin jälkeen la-  
 10 tasimme supernatantin Q-Sepharose-pylvääseen, joka oli  
 tasapainotettu 25-millimolaarisella Tris:llä (pH 7,5),  
 joka oli 0,5-millimolaarinen EDTA:n suhteen. Otimme tal-  
 teen tat $\Delta$ cys-249:n S-Sepharose-pylväästä käyttämällä vai-  
 heittaista NaCl-gradienttia 25-millimolaarisessa Tris:ssa  
 15 (pH 7,5). Tat $\Delta$ cys-249 eluoitui vaiheittaisessa NaCl-gra-  
 dientissa osassa, joka oli 600 - 800-millimolaarinen  
 NaCl:n suhteen.

Tat $\Delta$ cys-HE2.85 ja Tat $\Delta$ cys.He2.121 - Lysaatin sent-  
 rifugoinnin jälkeen latasimme supernatantin Q-Sepharose-  
 20 pylvääseen. Latasimme läpi tulleen nesteen S-Sepharose-  
 pylvääseen. Otimme talteen tat $\Delta$ cys-HE2.85:n tai tat $\Delta$ cys-  
 HE2.121:n käyttämällä vaiheittaista NaCl-gradienttia. Mo-  
 lemmat proteiinit eluoituivat 1-molaarisella NaCl:lla.

HPV E2 ja HPV E2CCSS - Katso esimerkki 9 (edellä).  
 25 JB106 - Lysaatin sentrifugoinnin ja supernatantin  
 keräämisen jälkeen llisäsimme NaCl:a niin että loppukon-  
 sentraatio oli 300 mM. Latasimme supernatantin, johon oli  
 lisätty NaCl:a, S-Sepharose-pylvääseen joka oli tasapai-  
 notettu 25 mM HEPES-puskurilla (pH 7,5). Käsittelimme pyl-  
 30 vään peräkkäisillä suolakonsentraatiovaiheilla puskurissa,  
 joka oli 25-millimolaarinen HEPES:n (pH 7,5), 1,5-millimo-  
 laarinen EDTA:n ja 1-millimolaarinen DTT:n suhteen. Elu-  
 oimme JB106-proteiinin S-Sepharose-pylväästä 1-molaarisel-  
 la NaCl:lla.

JB117 - Lysaatin sentrifugoinnin ja supernatantin keräämisen jälkeen lisäsimme NaCl:a niin että loppukonsentraatio oli 300 mM. Koska JB117 sakkautui 300-millimolaarisessa NaCl:ssa, laimensimme JB117-supernatantin 100-mil-

5 limolaariseen NaCl:iin ja latasimme proteiinin eränä S-Sepharose-pylvääseen. Eluimme JB117:n S-Sepharose-pylväästä 1-molaarisella NaCl:lla, joka oli 25-millimolaarinen Tris:n (pH 7,5) ja 0,3-millimolaarinen DTT:n suhteen.

JB118 - Lysaatin sentrifugoinnin ja supernatantin keräämisen jälkeen lisäsimme NaCl:a niin että loppukonsentraatio oli 300 mM. Latasimme supernatantin, johon oli lisätty NaCl:a, S-Sepharose-pylvääseen, joka oli tasapainotettu 25 mM HEPES-puskurilla (pH 7,5). Eluimme JB118-

10 proteiinin S-Sepharose-pylväästä 1-molaarisella NaCl:lla, joka oli 25-millimolaarinen Tris:n (pH 7,5) ja 0,3-millimolaarinen DTT:n suhteen.

JB119, JB120, JB121 ja JB122 - Lysaatin sentrifugoinnin ja supernatantin keräämisen jälkeen lisäsimme JB119:n ja JB121:n tapauksessa NaCl:a niin että sen konsentraatio oli 150 mM, ja JB120:n ja JB122:n tapauksessa

20 niin että sen konsentraatio oli 200 mM. Latasimme supernatantin, johon oli lisätty NaCl:a, S-Sepharose-pylvääseen, joka oli tasapainotettu 25 mM HEPES-puskurilla (pH 7,5). Eluimme proteiinit JB119, JB120, JB121, ja JB122 S-Sepha-

25 rose-pylväästä 1-molaarisella NaCl:lla, joka oli 25-millimolaarinen Tris:n (pH 7,5) ja 0,3-millimolaarinen DTT:n suhteen.

#### Esimerkki 14

##### E2-repressiomääritykset - muut konjugaatit

30 Testasimme tat-E2-fuusioproteiininimme silmällä pitäen papilloomaviruksen täyspitkän E2-proteiinin aikaansaamaa transkription aktivaatiota ("repressiota"). Mittasimme E2-repression lyhytkestoisella kotransfektiomäärityksellä COS7-soluissa. Tässä määrityksessä käytetyt COS7-

35 solut pidettiin viljelmässä vain lyhyitä ajanjaksoja. Su-

latimme COS7-solut pasaasissa 13 ja käytimme niitä vain  
 pasaasiin 25 asti. Pitkät kasvatusajat johtivat alhaisiin  
 E2:n transkription aktivaatiotasoihin ja alentuneeseen  
 repressioon ja toistettavuuteen. Meidän repressiomäärityk-  
 5 semme ja repressioaktiivisuuden laskentamenetelmämme kuva-  
 taan esimerkissä 9 (edellä). Konjugaattien TxE2.103,  
 TxE22.249, FTE103, FTE202, FTE403 ja FTE501 tapauksessa  
 korvasimme HPV-16 E2-transaktivaattorin yhtä suurella mää-  
 10 rällä BPV-1:n E2-transaktivaattoria. Vastaavasti, sen si-  
 jaan että olisi transfektoitu HPV-16 E2 -ekspressioplasmii-  
 dilla pAHE2, transfektoimme BPV-1 E2 -ekspressioplasmidil-  
 la pXB323, joka kuvataan täysin US-patentissa 5 219 990.

Geneettinen fuusioproteiini JB106 on ollut jatku-  
 vasti meidän tehokkain tat-E2-repressorikonjugaattimme.  
 15 Tulokset JB106:ta ja TxHE2CCSS:ää vertaavasta repres-  
 siomäärityksestä esitetään taulukossa III. Kuvio 13 kuvaa  
 graafisesti taulukossa III esitettyjä tuloksia.

JB106:n lisäksi useat muut tat-E2-repressorikonju-  
 gaatit ovat saaneet aikaan merkittävää repressiota. Kuten  
 20 taulukossa IV esitetään, TxHE2:lla, TxHE2CCSS:llä,  
 JB117:llä, JB118:lla, JB119:llä, JB120:lla ja JB122:lla  
 havaittiin repressiotasoja +-alueella.

Taulukko III

	Lisätty proteiini (µg/ml)	cpm- tausta*	Rinnakkais- parien keskiarvo	cpm- taustan keskiarvo	Repressio- prosentti
5	0	3 872			
	0	3 694	3783	--	--
	0	17 896			
	0	18 891	18 393	14 610	--
10	1 JB106	16 384			
	1 JB106	17 249	16 816	13 033	10,8
	3 JB106	11 456			
	3 JB106	10 550	11 003	7 220	50,6
	10 JB106	6 170			
	10 JB106	7 006	6 588	2 805	81,0
	30 JB106	4 733			
30 JB106	4 504	4 618	835	94,3	
15	1 TxHE2CCSS	17 478			
	1 TxHE2CCSS	18 047	17 762	13 979	4,3
	3 TxHE2CCSS	14 687			
	3 TxHE2CCSS	15 643	15 165	11 382	22,1
	10 TxHE2CCSS	12 914			
	10 TxHE2CCSS	12 669	12 791	9 008	38,3
	30 TxHE2CCSS	7 956			
20	30 TxHE2CCSS	8 558	8 257	4 474	69,4
25	1 HE2.123	18 290			
	1 HE2.123	18 744	18 517	14 734	0
	3 HE2.123	17 666			
	3 HE2.123	18 976	18 321	14 538	1,3
	10 HE2.123	18 413			
	10 HE2.123	17 862	18 137	14 354	2,6
	30 HE2.123	18 255			
30	30 HE2.123	18 680	18 467	14 684	0,3

\*Tausta = 158 cpm

30 Taulukossa IV esitetään yhteenvedona meidän tat-E2-repressiomääritystuloksemme. Vaikka testasimmekin kaikki tat-E2-repressorikonjugaattimme samanlaisissa määrittelyksissä, kaikkia konjugaatteja ei testattu yhtäaikaa samassa määrittelyksessä. Olemme vastaavasti ilmaisseet repressioaktiivisuuden tason semikvantitatiivisesti seuraavasti: +++,

35



++, +, +/- tai -, jossa +++ on vahva repressio ja - tarkoittaa ettei repressiota ole havaittavissa. Kuvio 13 esittää havainnollisesti repressioaktiivisuutta luokittelevan järjestelmän, jota käytetään taulukossa IV. JB106 on 5 esimerkki aktiivisuustasosta +++. TxHE2CCSS on esimerkki aktiivisuustasosta ++. Negatiivinen kontrolli, HE2.123, on esimerkki aktiivisuustasosta -. Aktiivisuustaso + on väli-  
muoto TxHE2CCSS:llä ja HE2.123:lla havaitusta aktiivisuudesta. Niillä kahdella konjugaatilla, joiden aktiivisuus 10 esitetään +/-:na, oli heikko (mutta havaittavissa oleva) aktiivisuus joissakin määrityksissä eikä havaittavissa olevaa aktiivisuutta muissa määrityksissä.

Taulukko IV

	Proteiini	Tat-tähteitä	E2-tähteitä	Repressio- taso
5	TxE2.103	37-72	BPV-1 308-410	+
	TxE2.249	37-72	BPV-1 162-410	-
	TxHE2	37-72	HPV-16 245-365	++
	TxHE2CCSS	37-72	HPV-16 245-365	++
10	FTE103	1-67	BPV-1 306-410	-
	FTE208	1-62	BPV-1 311-410	-
	FTE403	1-62	BPV-1 250-410	-
	FTE501	1-62	BPV-1 162-410	-
15	Tat $\Delta$ cys- 105	1-21,38-67	BPV-1 306-410	-
	Tat $\Delta$ cys- 161	1-21,38-62	BPV-1 250-410	+/-
20	Tat $\Delta$ cys- 249	1-21,38-62	BPV-1 162-410	+/-
	Tat $\Delta$ cys- HE2.85	1-21,38-72	HPV-16 281-365	+
	Tat $\Delta$ cys- HE2.121	1-21,38-72	HPV-16 245-365	+
25	JB106	47-58	HPV-16 245-365	+++
	JB117	47-72	HPV-16 245-365	++
	JB118	38-72	HPV-16 245-365	++
30	JB119	47-62	BPV-1 250-410	++
	JB120	38-62	BPV-1 250-410	++
	JB122	38-58	HPV-16 245-365	++

FTE103:lta, FTE403:lta, FTE208:lta ja FTE501:ltä, niiltä neljältä konjugaatilta, joilla oli tat:n aminoterminaalinen alue (se on: tähteet 1 - 21) ja kysteiinipitoinen alue (se on: tähteet 22 - 37), puuttui repressio kokonaan. Koska olemme osoittaneet epäsuoralla immunofluorensenssilla, että FTE501 menee sisään soluihin, pidämme todennäköisenä että E2-repressoriaktiivisuus on menetetty FTE-sarjassa johtuen liitoksesta tat-kuljetuspolyeptidiin. Meidän tuloksemme osoittavat, että kysteiinipitoisen alueen puuttuminen tat-osasta lisäsi yleisesti E2-repressoriaktiivisuutta. Lisäksi kysteiinipitoisen alueen poissaolo tat-E2-konjugaateissa vaikutti lisäävän proteiinin tuotantotasoa E. colissa, ja lisäävän proteiinin liukoisuutta, ilman että menetettiin kuljetusta kohdesoluihin. Tat:n aminoterminaalisen alueen deletio lisäsi myös E2-repressorin aktiivisuutta. Fuusioproteiini JB106, jossa oli vain tat-tähteet 47 - 58, oli tehokkain meidän tat-E2-repressorikonjugaateistamme. Tat:n kysteiinipitoisen alueen poissaolo ei kuitenkaan aina johda E2-repressoriaktiivisuuden säilymiseen konjugaatissa. Esimerkiksi kemiallinen konjugaatti TxE2.249 oli liukenematon ja myrkyllinen soluille. Niinpä jopa kysteiinittömän tat:n osan liittämisen voi saada aikaan toimimattoman E2-repressorikonjugatin.

25 **Esimerkki 15**

**Katkaistavissa olevat E2-konjugaatit**

Tat-osien kemiallinen konjugointi E2-proteiiniin sai aikaan ainakin 20-kertaisen vähentymisen E2-proteiinin sitoutumisessa E2-sitoutumiskohtiin DNA:ssa (tuloksia ei esitetä). Niinpä teimme kokeita arvioidaksemme katkaistavissa olevan silloituksen tat-kuljetusosan ja E2-repressoriosan välillä. Testasimme erilaisia menetelmiä, joilla voidaan tehdä katkaistavissa oleva silloitus.

35 Yhdessä koesarjassa aktivoimme HPV E2-CCSS-proteiinin kysteiinisulfydryyliryhmät aldritiolilla 100-millimo-

laarisessa HEPES-puskurissa (pH 7,5), joka oli 500-millimolaarinen NaCl:n suhteen. Eristimme aktivoitun E2-repressorin geelisuodatuskromatografialla ja käsittelimme sen tat37-72:lla. Saimme aikaan alhaisen silloitustehon, mikä  
 5 johtui nopeasta E2-CCSS-dimeerin muodostumisesta aldritiolikasittelyllä. Tämän ongelman välttämiseksi panimme E2-CCSS:n 8-molaariseen ureaan huoneenlämpötilassa ja käsittelimme sen aldritiolilla 23 °C:n lämpötilassa 60 minuutin ajan denaturoivissa olosuhteissa. Sitten laskostimme uudestaan E2CCSS-aldritioli-aduktiin, eristimme sen geelisuodatuskromatografilla ja annoimme sen sitten reagoida  
 10 tat37-72:n kanssa. Tällä menetelmällä saatiin aikaan erinomainen silloitus. Silloitimme myös E2CCSS:n ja E2CCSC:n tat37-72:een käyttämällä ureamenetelmän muunnelmaa, jossa  
 15 käytimme S-Sepharose-kromatografiaa geelisuodatuksen sijasta eristääksemme E2-aldritioli-adukteja. Tämä muunnos lisäsi aduktien saantoa ja sai aikaan reaktiossa käytetyssä E2-lähtömateriaalissa noin 90-prosenttisen silloittumisen.

20 Katkaistavissa olevilla tat-E2-konjugaateilla oli aktiivisuutta repressiomäärityksessä. Katkaistavissa olevien konjugaattien repressioaktiivisuus oli kuitenkin hieman alhaisempi kuin samanlaisilla konjugaateilla, jotka oli silloitettu irreversiibelisti. Katkaistavissa olevien  
 25 konjugaattien hieman alhaisempi aktiivisuus voi heijastaa proteiinin puoliintumisaikaa solussa. Tat on suhteellisen stabiili soluissa. E2-proteiineilla on yleensä lyhyet puoliintumisajat solussa. Niinpä irreversiibeli silloittuminen tat-osan ja E2-osan välillä voi stabiloida E2-osaa.

30 **Esimerkki 16**

**Herpes Simplex-viruksen repressorikonjugaatti**

Herpes Simplex -virus ("HSV") koodittaa transkription aktivaattoria VP16, joka indusoi hyvin varhaisten HSV-geenien ekspression. Friedman ym. ovat tuottaneet  
 35 HSV:n VP16-repressorin deletoimalla VP16:n karboksitermi-

naalisen transaktivaatio-osan ("Expression of a Truncated Viral Trans-Activator Selectively Impedes Lytic Infection by Its Cognate Virus", *Nature*, 335, s. 452 - 454 (1988)). Olemme tuottaneet HSV-2:n VP16-repressorin samanlaisella tavalla.

Jotta saisimme testattua kuljetuspolyeptidi-VP16-repressorikonjugaattien soluun oton ja VP16-repressorin aktiivisuuden, transfektoimme samanaikaisesti VP16-riippuvaisen reportteriplasmidin ja Vp16-repressoriplasmidin COS7-soluihin. Sitten altistimme transfektoidut solut kuljetuspolyeptidi-VP16-repressorikonjugaatille tai soveliaalle kontrollille. Jäljempänä kuvattu repressiomääritys oli analoginen edellä esimerkissä 9 kuvatun E2-repressiomäärityksen kanssa.

#### 15 VP16-repressiomääritysplasmidit

Meidän reportterikonstruktiomme VP16-repressiomääritystä varten oli plasmidi p175kCAT, joka oli hankittu G. Haywardilta (katso P. O'Hare ja G. S. Hayward, "Expression of Recombinant Genes Containing Herpes Simplex Virus Delayed-Early and Immediate-Early Regulatory Regions and Trans Activation by Herpes Virus Infection", *J. Virol.*, 52, s. 522 - 531 (1984)). Plasmidi p175kCAT sisältää HSV-1:n IE175-promoottorin, joka ohjaa CAT-reportterigeeniä.

Meidän HSV-2-transaktivaattorikonstruktiomme VP16-repressiomääritystä varten oli plasmidi pXB324, joka sisälsi villityyppisen HSV-2 VP16-geenin kanan  $\beta$ -aktiini-promoottorin ohjauksessa. Muodostimme pXB324:n sijoittamalla pXB100:aan (P. Han ym., "Transactivation of Heterologous Promoters by HIV-1 Tat", *Nuc. Acids. Res.*, 19, s. 7225 - 7229 (1991)), XhoI-kohdan ja BamHI-kohdan väliin, 280 emäsparin fragmentin joka sisälsi kanan  $\beta$ -aktiini-promoottorin, ja pCA5-plasmidista peräisin olevan 2318 emäsparin BamHI-EcoRI-fragmentin (O'Hare ja Hayward, edellä), joka koodittaa kokonaista villityyppistä HSV-2:n VP16-proteiinia.

### Tat-VP16-repressorifuusioproteiini

Tuotimme bakteereissa fuusioproteiinin tat-VP16R.GF (SEQ ID nro 58), joka koostui HIV:n tat-proteiinin aminohapoista 47 - 58, jota seurasivat HSV:n VP16-proteiinin aminohapot 43 - 412. Tat-VP16-repressorifuusioproteiinin bakteerituotantoa varten muodostimme plasmidin pET/tat-VP16R.GF kolmivaiheisessa ligaatiossa. Ensimmäinen fragmentti oli vektori pET-3d (kuvattu edellä vaihtoehtoisella nimellä "pET-8c"), joka oli digestoitu NcoI:llä ja BglIII:lla (noin 4600 emäsparia). Toinen fragmentti koostui synteettisistä oligonukleotideista 374.219 (SEQ ID nro 39) ja 374.220 (SEQ ID nro 40), jotka oli liitetty yhteen kaksijuosteisen DNA-molekyylin muodostamiseksi. Synteettisen fragmentin 5'-päässä oli yksijuosteisena yli tuleva NcoI-osa, joka sisälsi translaation aloituskodonin ATG. Aloituskodonin jälkeen olivat kodonit tat-tähteille 47 - 58. Heti tat-kodonien jälkeen lukukehyksessä olivat kodonit VP16:n tähteille 43 - 47. Synteettisen fragmentin 3'-pää oli tasapää kolmanteen fragmenttiin ligaatiota varten, joka oli 1134 emäsparin PvuII-BglIII-fragmentti pXB324R4:stä, joka sisälsi HSV-2 VP16:n kodonit 48 - 412. Johdimme pXB324R4:n pXB324:stä (kuvattu edellä). Plasmidi pXB324R2 oli välituote pXB324R4:n muodostamisessa.

Muodostimme pXB324R2:n sijoittamalla pXB100:aan 1342 emäsparin BamHI-AatII-fragmentin, joka oli peräisin pXB324:stä ja kooditti HSV-2 VP16:n 419 N-terminaalista aminohappoa. Jotta saisimme aikaan lukukehyksen sisällä olevan lopetuskodonin, käytimme 73 emäsparin AatII-EcoRI-fragmenttia, joka oli peräisin pSV2-CAT:sta (C.M. Gorman ym., Molecular and Cellular Biology, 2, s. 1044 - 1051 (1982)). Niinpä pXB324R2 kooditti HSV-2:n VP16:n 419 ensimmäistä aminohappoa ja lisäksi seitsemää ei-VP6-aminohappoa ennen lopetuskodonia. pXB324R4:n muodostamiseksi suoritimme kolmen palan ligaation, joka käsitti 5145 emäsparin MluI-EcoRI-fragmentin pXB324R2:sta ja kaksi insert-

tifragmenttia. Yksi insertti oli 115 emäsparin MluI-NspI-fragmentti pXB324R2:sta, joka kooditti VP16:n 198 ensimmäistä aminohappoa. Toinen inserttifragmentti oli kaksijuosteinen synteettinen DNA-molekyyli, joka koostui synteettisistä oligonukleotideista 374.32 (SEQ ID nro 41) ja 374.33 (SEQ ID nro 42). Yhteen hybridisoituina nämä oligonukleotidit muodostivat 5'-päähen yksijuosteisen NspI-pään ja 3'-päähen yksijuosteisen EcoRI-pään. Tämä synteettinen fragmentti kooditti VP16:n tähteitä 399-412, jota seurasi lopetuskodoni. Niinpä plasmidi pXB324R2 erosi pXB324R2:sta sikäli että siitä puuttuivat kodonit VP16:n aminohapoille 413 - 419 ja seitsemän ylimääräistä aminohappoa ennen lopetuskodonia.

#### Tat-VP16R.GF-fuusioproteiinin puhdistus

Ekspressoimme tat-VP16R.GF:ää varten tehdyn geneettisen konstruktiomme E. colissa. Keräsimme transformoidun E. colin sentrifugaatiolla, resuspendoimme solut 8 - 10 tilavuudessa hajotuspuskuria (25 mM Tris (pH 7,5), 25 mM NaCl, 1 mM DTT, 0,5 mM EDTA, 1 mM PMSF, 4 µg/ml E64:ää, 50 µg/ml aprotiniinia, 50 µg/ml pepstatiini A:ta, ja 3 mM bentsamidiini), hajotimme solut "French press"-laitteessa (2 kertaa, 12 000 psi:n paine (82 656 kPa), ja sentrifugoimme lyaatin 10 000 - 20 000 g:n kiihtyvyydessä 1 tunnin ajan 4 °C:n lämpötilassa. Lyaatin sentrifugoinnin jälkeen latasimme supernatantin Fast Q-Sepharose-pylvääseen, joka oli tasapainotettu 25-millimolaarisella Tris:llä (pH 7,5), joka oli 0,5-millimolaarinen EDTA:n suhteen. Latasimme Q-Sepharose:sta läpi tulleen nesteen Fast S-Sepharose-pylvääseen, joka oli tasapainotettu 25-millimolaarisella MES:llä (pH 6,0), joka oli 0,1-millimolaarinen EDTA:n ja 2-millimolaarinen DTT:n suhteen. Keräsimme tat-VP16-fuusioproteiinin S-Sepharose-pylvästä peräkkäisillä NaCl-konsentroitivaiheilla 25-millimolaarisessa MES:ssä (pH 6,0), joka oli 0,1-millimolaarinen EDTA:n ja 2-millimolaarinen DTT:n suhteen. Tat-VP16-fuu-

sioproteiini eluoitui 600 - 1000-millimolaarisissa NaCl-fraktioissa.

#### VP16:n repressiomääritys

Jaoimme HeLa-soluja 24-kuoppaisille viljelmälevyille,  $10^5$  solua/kuoppa. Seuraavana päivänä transfektoimme solut käyttämällä DEAE-dekstraanimenetelmää, kuten B. R. Cullen on kuvannut artikkelissa "Use of Eukaryotic Expression Technology in the Functional Analysis of Cloned Genes", Meth Enzymol., vol. 152, s. 684 (1987). Saostimme DNA:n transfektioita varten ja liuotimme sen uudelleen konsentraatioon noin 100  $\mu\text{g/ml}$  liuoksessa, joka oli 100-millimolaarinen NaCl:n ja 10-millimolaarinen Tris:n suhteen (pH 7,5). Kutakin transfektiota varten DNA-DEAE-seos koostui seuraavista: 200 ng p175kCAT (+/- 1 ng pXB324) tai 200 ng pSV-CAT (kontrolli), 1 mg/ml DEAE-dekstraania ja PBS:ää, lopputilavuus 100  $\mu\text{l}$ . Altistimme solut tälle seokselle 15 - 20 minuutin ajan 37 °C:n lämpötilassa, niin että viljelylevyjä heilutettiin ajoittain. Sitten lisäsimme kuhunkin kuoppaan 1 ml tuoretta DC-elatusainetta (DMEM + 10-prosenttinen seerumi), joka oli 80-mikromolaarinen klorokiinin suhteen. Sen jälkeen kun soluja oli inkuboitu 37 °C:n lämpötilassa 2,5 tuntia, imimme elatusaineen kussakin kuopasta ja korvasimme sen tuoreella DC:llä, joka oli 10-prosenttinen DMSO:n suhteen. Kun oli kulunut 2,5 tuntia huoneenlämpötilassa, imimme pois DMSO:n sisältävän elatusaineen ja korvasimme sen tuoreella DC:llä, joka sisälsi 0, 10 tai 50  $\mu\text{g/ml}$  puhdasta tat-VP16.GF:ää. Seuraavana päivänä korvasimme kussakin kuopassa elatusaineen tuoreella elatusaineella, jolla oli sama koostumus. 24 tuntia myöhemmin hajotimme HeLa-solut puskurilla, joka oli 0,65-prosenttinen NP-40:n (detergentti), 10-millimolaarinen Tris:n (pH 8,0), 1-millimolaarinen EDTA:n ja 150-millimolaarinen NaCl:n suhteen. Mittasimme proteiinikonsentraation kussakin ekstraktissa normalisoidaksemme näytteet kussakin määrityksessä.



Kun tat-VP16.GF:n konsentraatio oli 50 µg/ml, toksisuus soluille häiritsi määrittystä. Kun konsentraatio oli 10 µg/ml, tat-VP16.Gf-fuusioproteiini sai aikaan lähes täydellisen repression VP16-riippuvaisessa CAT-ekspressiossa, niin että havaittavaa solukuolemaa ei ollut ja 5 kontrolleissa oli noin 30-prosenttinen ei-VP16-riippuvaisen CAT-ekspression repression. Niinpä havaitsimme VP16-riippuvaisen transaktivaation spesifisen repression sen lisäksi että vähemmässä määrin esiintyi epäspesifistä repressionia. 10

#### Esimerkki 17

#### Kuljetuspolyeptidi-DNA-konjugaatit

DNA:han sitoutuvan transkriptiofaktorin aikaansaama transkription aktivaatio voidaan estää tuomalla soluihin 15 DNA, jossa on sitoutumiskohta kyseiselle transkriptiofaktorille. Tuotu DNA sitoo transkriptiofaktorin ja siitä tulee kyvytön sitoutumaan promoottorikohtaan, jossa se normaalisti toimii. Tätä strategiaa on käytetty inhiboimaan NF-κB:n aikaansaama transkription aktivoituminen 20 (Bielinska ym., "Regulation of Gene Expression with Double-Stranded Phosphorothioate Oligonucleotides", Science, vol. 250, s. 997 - 1000 (1990)). Bielinska ym. havaitsivat annosriippuvaisen inhibition, kun soluviljelyelatusaineeseen pantiin kaksijuosteista DNA:ta. Konjugoimme kuljetuspolyeptidi tat37-72:n kaksijuosteiseen DNA-molekyyliin, jotta saisimme määrittetyksi tehostaisiko sellainen 25 konjugaatio inhibitiota lisäämällä DNA:n ottoa solun sisään.

Ostimme neljä halutulla tavalla syntetisoitua 39- 30 meeristä fosforitioaattioligonukleotidia, joille oli annettu nimiksi NF1, NF2, NF3 ja NF4, ja joilla oli vastavasti nukleotidisekvenssit SEQ ID nro 43, SEQ ID nro 44, SEQ ID nro 45 ja SEQ ID nro 46. NF1 ja NF2 muodostavat dupleksin, joka vastaa villityyppistä NF-κB:n sitoutumis-

kohtaa. NF3 ja NF4 muodostavat dupleksin, joka vastaa mutatoitunutta NF- $\kappa$ B:n sitoutumiskohtaa.

5 Liuotimme NF1:n ja NF3:n veteen noin 4 mg/ml:n konsentraatioon. Sitten panimme 800  $\mu$ g NF1:tä ja NF3:a erikseen 400  $\mu$ l:aan puskuria, joka oli 50-millimolaarinen trietanoliamiinin (pH 8,2), 50-millimolaarinen NaCl:n ja 10-millimolaarinen Trautin reagenssin suhteen. Annoimme reaktion edetä 50 minuuttia huoneenlämpötilassa. Pysäytimme reaktion geelisuodatuksella P6DG-pylväässä (BioRad, 10 Richmond, CA), joka oli tasapainotettu 50-millimolaarisella HEPES-puskurilla (pH 6,0), joka oli 50-millimolaarinen NaCl:n suhteen, jotta saisimme poistettua ylimääräisen Trautin reagenssin. Monitoroimme 260 nm:n absorbanssia tunnistaksemme oligonukleotidin sisältävät fraktiot. Meidän 15 oligonukleotidisaantomme oli noin 75 %. Sitten hybridisoimme Trautilla muokatun NF1:n NF2:n kanssa (loppukonsentraatio 0,55 mg/ml) ja hybridisoimme Trautilla muokatun NF3:n NF4:n kanssa (loppukonsentraatio 0,50 mg/ml). Lopuksi annoimme 0,4 mg:n kutakin Trautilla muokattua DNA:ta 20 reagoida 0,6 mg:n kanssa tat-37-BMH:ta (valmistettu esimerkissä 9 edellä kuvatulla tavalla), 100-millimolaarisessa HEPES-puskurissa (1 ml, pH 7,5), 60 minuutin ajan huoneenlämpötilassa. Monitoroimme silloitusreaktion etenemistä polyakryyliamidigeelielektroforeesilla, jota seurasi 25 geelin värjäys etidiumbromidilla. Yleensä havaitsimme, että noin 50 % DNA:sta modifioituu näissä olosuhteissa.

Nämä kaksijuosteiset DNA-molekyylit testattiin oleellisesti Bielinskan ym. (edellä) menetelmien mukaisesti, tat-sidoksella tai ilman, silmällä pitäen NF- $\kappa$ B:n 30 transkription aktivaatiota. Tat-liitos tehosti merkittävästi NF- $\kappa$ B:n aikaansaamaa transaktivaatiota.

Tässä kuvatuilla menetelmillä valmistetuista rekombinantti-DNA-sekvensseistä on esimerkkinä viljelmä, joka on talletettu American Type Culture Collection -talletuslaitokseen, Rockville, Maryland. Escherichia coli -viljel-

mä, jolle on annettu nimeksi pJB106, talletettiin 28. heinäkuuta 1993 ja sille annettiin ATCC-talletusnumero 69368.

5 Vaikka olemme kuvanneet lukuisia tämän keksinnön suoritusmuotoja, on ilmeistä että meidän peruskonstruktioitamme voidaan muuttaa siten, että tarjotaan muita suoritusmuotoja, jotka käyttävät hyväksi tämän keksinnön mukaisia menetelmiä ja tuotteita. Niinpä ymmärretään, että tämän keksinnön suojapiirin määrittävät liitteinä olevat patenttivaatimukset pikemminkin kuin spesifiset suoritusmuodot, jotka on esitetty esimerkinomaisesti.

10

## Sekvenssilista:

## (1) GENERAL INFORMATION:

- (i) APPLICANT: BIOGEN, INC.
  - BARSOUM, James G. (US only)
  - FAWELL, Stephen E. (US only)
  - PEPINSKY, R. B. (US only)
- (ii) TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
- (iii) NUMBER OF SEQUENCES: 63
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: FISH & NEAVE
  - (B) STREET: 1251 Avenue of the Americas
  - (C) CITY: New York
  - (D) STATE: New York
  - (E) COUNTRY: USA
  - (F) ZIP: 10020
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 07/934,375
  - (B) FILING DATE: 21-AUG-1992
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Haley Jr., James F.
  - (B) REGISTRATION NUMBER: 27,794
  - (C) REFERENCE/DOCKET NUMBER: B170CIP
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (212) 596-9000
  - (B) TELEFAX: (212) 596-9090
  - (C) TELEX: 14-8367

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 86 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: human immunodeficiency virus  
 (B) STRAIN: type 1

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Glu Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser  
 1                                   5                                   10                                   15

Gln Pro Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe  
                                   20                                   25                                   30

His Cys Gln Val Cys Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly  
                                   35                                   40                                   45

Arg Lys Lys Arg Arg Gln Arg Arg Pro Pro Gln Gly Ser Gln Thr  
   50                                   55                                   60

His Gln Val Ser Leu Ser Lys Gln Pro Thr Ser Gln Ser Arg Gly Asp  
 65                                   70                                   75                                   80

Pro Thr Gly Pro Lys Glu  
                                   85

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Cys Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg  
 1                                   5                                   10                                   15

Arg Gln Arg Arg Arg Pro Pro Gln Gly Ser Gln Thr His Gln Val Ser  
                                   20                                   25                                   30

Leu Ser Lys Gln  
                                   35

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Cys Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg  
 1                   5                   10                   15

Arg Gln Arg Arg Arg Pro  
 20

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg Arg  
 1                   5                   10                   15

Gln Arg Arg Arg Pro Gly Gly Cys  
 20

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 15 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Cys Gly Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro  
 1                   5                   10                   15

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 15 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Gly Gly Cys  
 1                    5                                    10                                    15

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 56 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Glu Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser  
 1                    5                                    10                                    15

Gln Pro Lys Thr Ala Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly  
 20                                    25                                    30

Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln Gly Ser Gln Thr  
 35                                    40                                    45

His Gln Val Ser Leu Ser Lys Gln  
 50                                    55

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GATCCCAGAC CCACCAGGTT TCTCTGTCGG GCCCTTAAG

39

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AATTCTTAAG GGCCCGACAG AGAACCTGG TGGGTCTGG

39

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 5098 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTGAAGACGA AAGGGCCTCG TGATACGCCT ATTTTATAG GTTAATGTCA TGATAATAAT	60
GGTTTCTTAG ACGTCAGGTG GCACCTTTTCG GGGAAATGTG CGCGGAACCC CTATTTGTTT	120
ATTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCT GATAAATGCT	180
TCAATARTAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTGG CCCTTATTCC	240
CTTTTTGCG GCATTTTGGC TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA	300
AGATGCTGAA GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG	360
TAAGATCCTT GAGAGTTTTT GCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTTAAAGT	420
TCTGCTATGT GCGCGGTAT TATCCCGTGT TGACGCCGGG CAAGAGCAAC TCGGTGCGCG	480
CATACACTAT TCTCAGAATG ACTTGGTTGA GTACTACCA GTCACAGAAA AGCATCTTAC	540
GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGACTG ATAACACTGC	600
GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGACAA	660
CATGGGGGAT CATGTAATC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC	720



AAACGACGAG CGTGACACCA CGATGCCCTGC AGCAATGGCA ACAACGTTGC GCAAACCTATT	780
AACTGGCGAA CTACTIONTACTC TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA	840
TAAAGTTGCA GGACCACTTC TGGCCTCGGC CCTTCCGGCT GGCTGGTTTA TTGCTGATAA	900
ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATGCA GCACTGGGGC CAGATGGTAA	960
GCCCTCCCGT ATCGTAGTA TCTACACGAC GGGGAGTCAG GCACTATGG ATGAACGAAA	1020
TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT	1080
TTACTCATAT AACTTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAA GGATCTAGGT	1140
GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTCCACTG	1200
AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT	1260
AATCTGCTGC TTGCAARCAA AAAAACCACC GCTACCAGCG GTGGTTTGTG TGCCCGATCA	1320
AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAATAAC	1380
TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC	1440
ATACCTCGCT CTGCTAATCC TGTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT	1500
TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCCG GCTGAACGGG	1560
GGGTTGCTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAAGTGA GATACCTACA	1620
GCGTGAGCAT TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT	1680
AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGAA ACGCCTGTA	1740
TCTTTATAGT CCGTCTGGGT TCGCCACCT CTGACTTGAG CGTCGATTTT TGTGATGCTC	1800
GTCAGGGGGG CGGAGCCCTAT GGAAAAACGC CAGCAACCGG GCCTTTTAC GGTTCCTGGC	1860
CFTTTGCTGG CCTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCTGATT CTGTGGATAA	1920
CCGTATTACC GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG	1980
CGAGTCAGTG AGCGAGGAAG CGGAAGAGCG CCTGATGCGG TATTTCTCC TTACGCATCT	2040
GTGCGGTATT TCACACCGCA TATATGGTGC ACTCTCAGTA CAATCTGCTC TGATGCCGCA	2100
TAGTTAAGCC AGTATAACT CCGCTATCGC TACGTGACTG GGTGATGGCT GCGCCCCGAC	2160
ACCCGCCAAC ACCCGCTGAC GCGCCCTGAC GGGCTTGTCT GCTCCCGCA TCCGCTTACA	2220
GACAAGCTGT GACCGTCTCC GGGAGCTGCA TGTGTCAGAG GTTTTCACCG TCATCACCGA	2280
AACGCGCGAG GCAGCTGCGG TAAAGCTCAT CAGCGTGGTC GTGAAGCGAT TCACAGATGT	2340
CTGCCTGTTT ATCCGCGTCC AGCTCGTTGA GTTTCTCCAG AAGCGTTAAT GTCTGGCTTC	2400

TGATAAAGCG GGCCATGTTA AGGGCGGTTT TTCTCTGTTT GGTCACTTGA TGCCTCCGTC	2460
TAAGGGGGAA TTTCTGTTC TGGGGGTAAT GATACCGATG AAACGAGAGA GGATGCTCAC	2520
GATACGGGTT ACTGATGATG AACATGCCCG GTTACTGGAA CGTTGTGAGG GTAAACAAC	2580
GGCGGTATGG ATCGGGCGGG ACCAGAGAAA AATCACTCAG GGTCAATGCC AGCGCTTCGT	2640
TAATACAGAT GTAGGTGTTT CACAGGGTAG CCAGCAGCAT CCTGCCGATG AGATCCGGAA	2700
CATAATGGTG CAGGGCGCTG ACTTCCGCGT TTCCAGACTT TACGAAACAC GGAAACCGAA	2760
GACCATTTCAT GTTGTGCTC AGGTCCGAGA CGTTTTGCAG CAGCAGTCGC TTCACGTTCC	2820
CTCGCGTATC GGTGATTTCAT TCTGCTAACC AGTAAGGCAA CCCC GCCAGC CTAGCCGGGT	2880
CCTCAACGAC AGGAGCACGA TCATGCCAC CCGTGGCCAG GACCCAACGC TGCCCCGAGT	2940
GCGCCCGGTC CGGCTGCTGG AGATGGCGGA CCGATGGAT ATGTTCTGCC AAGGGTTGGT	3000
TTGGCGATTC ACAGTTCTCC GCAAGATTG ATTGGCTCCA ATTCTGGAG TGGTGATCC	3060
GTTAGCGAGG TGCCGCCGGC TTCCATTTCAG GTCGAGGTGG CCCGGCTCCA TGCACCCGA	3120
CGCAACCGCG GGAGGCAGAC AAGGTATAGG GCGGCGCCTA CAATCCATGC CAACCCGTT	3180
CATGTGCTCG CCGAGGCGGC ATAAATCGCC GTGACGATCA GCGGTCCAGT GATCGAAGTT	3240
AGGCTGGTAA GAGCCCGGAG CGATCCTTGA AGCTGTCCCT GATGGTCTG ATCTACCTGC	3300
CTGGACAGCA TGGCCTGCAA CCGGGGCATC CCGATGCCGC CGGAAGCGAG AAGAATCATA	3360
ATGGGGAAGG CCATCCAGCC TCGCGTCGG AACCAGCA AGACGTAGCC CAGCGGCTCG	3420
GCCGCCATGC CGGCGATAAT GGCCTGCTTC TCGCCGAAAC GTTTGGTGGC GGGACCACTG	3480
ACGAAGGCTT GAGCGAGGGC GTCCAAGATT CCGAATACCG CAAGCGACAG GCCGATCTC	3540
GTCGCGCTCC AGCGAAAGCG GTCCCTCGCCG AAAATGACCC AGAGCGCTGC CGGCACCTGT	3600
CCTACGAGTT GCATGATAAA GAAGACAGTC ATAAGTCCGG CGACGATAGT CATGCCCGC	3660
GCCCACCGGA AGGAGCTGAC TGGGTTGAAG GCTCTCAAGG GCATCGGTGC ACGCTCTCCC	3720
TTATGCGACT CCTGCATTAG GAAGCAGCCC AGTAGTAGGT TGAGGCCGTT GAGCACCGCC	3780
GCCGCAAGGA ATGGTGATG CAAGGAGATG GCGCCCAACA GTCCCCGGC CACGGGGCCT	3840
GCCACCATAC CCACGCCGAA ACAAGCGCTC ATGAGCCCGA AGTGGCGAGC CCGATCTTCC	3900
CCATCGGTGA TGTCGGCGAT ATAGGCGCCA GCAACCGCAC CTGTGGCGCC GGTGATGCCC	3960
GCCACGATGC GTCCGGCGTA GAGGATCGAG ATCTCGATCC CCGAAATTA ATACCACTCA	4020
CTATAGGGAG ACCACAACGG TTTCCCTCTA GAAATAATT TGTTTA ACTT TAAGAAGGAG	4080

ATATACATAT GGAACCGGTC GACCCCGGTC TGGAAACCATG GAAACACCCC GGGTCCCAGC	4140
CGAAAACCGC GTGCACCAAC TGCTACTGCA AAAAATGCTG CTTCCTACTGC CAGGTTTGCT	4200
TCATCACCAA AGCCCTAGGT ATCTCTTACG GCGGTAAAAA ACGTCGTCAG CGACGTCGTC	4260
CGCCGCAGGG ATCCCAGACC CACCAGGTTT CTCTGTCCGG CCGGCCGGAC AGCGGCGACG	4320
CCCTGCTGGA GCGCAACTAT CCCACTGGCG CGGAGTTCCT CGGCGACGGC GCGACGTC	4380
GCTTCAGCAC CCGCGGCACG CAGAACTGGA CGGTGGAGCG GCTGCTCCAG GCGCACCGCC	4440
AACTGGAGGA GCGCGGCTAT GTGTTGCTCG GCTACCACGG CACCTTCCTC GAAGCGGCGC	4500
AAAGCATCGT CTTCGGCGGG GTGCGCGCGC GCAGCCAGGA CCTCGACGGC ATCTGGCGCG	4560
GTTTCTATAT CCGCGGCGAT CCGCGGCTGG CCTACGGCTA CGCCAGGAC CAGGAACCCG	4620
ACGCACGGCG CCGGATCCGC AACGGTGCCC TGCTGCGGGT CTATGTGCCG CGCTCGAGCC	4680
TGCCGGGCTT CTACCGCACC AGCCTGACCC TGGCCGCGCC GGAGGCGGGC GCGGAGGTCG	4740
AACGGCTGAT CGGCCATCCG CTGCCGCTGC GCCTGGACGC CATEACCGGC CCCGAGGAGG	4800
AAGGCGGGCG CCTGGAGACC ATTCTCGGCT GCGCGCTGGC CGAGCGCACC GTGGTGATTC	4860
CCTCGGCGAT CCCCACCGAC CCGCGCAACG TCGCGGGCGA CCTCGACCCG TCCAGCATCC	4920
CCGACAAGGA ACAGGCGATC AGCGCCCTGC CGGACTACGC CAGCCAGCCC GGCAAACCGC	4980
CGCGCGAGGA CCTGAAGTAA CTGCCGCGAC CCGCGGGCTC CCTCGCAGG AGCCGGCCTT	5040
CTCGGGCCTT GGCCATACAT CAGGTTTTCC TGATGCCAGC CCAATCGAAT ATGAATTC	5098

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4910 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTGAAGACGA AAGGGCCTCG TGATACGCTT ATTTTATAG GTTAATGTCA TGATAATAAT	60
GGTTTCTTAG ACGTCAGGTG GCACTTTTCG GGGAAATGTG CGCGGAACCC CTATTTGTTT	120
ATTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCG GATAAATGCT	180
TCAATAATAT TGAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCC CCCTTATTC	240

CTTTTTGCG GCATTTGCC TTCCTGTTT TGCTACCCA GAAACGCTGG TGAAAGTAAA	300
AGATGCTGAA GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG	360
TAAGATCCTT GAGAGTTTT CCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTAAAGT	420
TCTGCTATGT GGGCGGTAT TATCCCGTGT TGACGCCGGG CAAGAGCAAC TCGGTCCCGG	480
CATACACTAT TCTCAGAATG ACTTGTTGA GACTCACCA GTCACAGAA AGCATCTTAC	540
GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC	600
GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGCACAA	660
CATGGGGGAT CATGTAAGTC GCCTTGATCG TTGGGAACCG GAGCTGATG AAGCCATACC	720
AAACGACGAG CGTGACACCA CGATGCCTGC AGCAATGGCA ACAACGTTGC GCAAACATTT	780
AACTGGCGAA CTACTTACTC TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA	840
TAAAGTTGCA GGACCACTTC TGCGCTCGGC CCTTCCGGCT GGCTGGTTTA TTGCTGATAA	900
ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATTGCA GCACTGGGGC CAGATGGTAA	960
GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA	1020
TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT	1080
TTACTCATAT ATACTTTAGA TTGATTTAAA ACTTCATTT TAATTTAAA GGATCTAGGT	1140
GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTCCACTG	1200
AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT	1260
AATCTGCTGC TTGCAAACAA AAAAACCACC GCTACCAGCG GTGGTTTGT TTGCCGGATCA	1320
AGAGCTACCA ACTCTTTTT CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAATAC	1380
TGTCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC	1440
ATACCTCGCT CTGCTAATCC TGTTACCAGT GCCTGCTGCC AGTGGCGATA AGTCGTGTCT	1500
TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCCG GCTGAACGGG	1560
GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA	1620
GCGTGAGCAT TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGCGGACA GGTATCCGGT	1680
AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGAA ACGCCTGGTA	1740
TCTTTAAGT CCGTCCGGGT TTCGCCACCT CTGACTTGAG CGTGGATTTT TGTGATGCTC	1800
GTCAGGGGGG CGGAGCCTAT GGAAAACCG CAGCAACCGG GCCTTTTTAC GGTCCCTGGC	1860
CTTTGCTGG CCTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCCTGATT CTGTGGATAA	1920

CCGTATTACC GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCCGAG	1980
CGAGTCAGTG AGCGAGGAAG CGGAAGAGCG CCGATGCGG TATTTCTCC TTACGCATCT	2040
GTCCGGTATT TCACACCCGA TATATGGTGC ACTCTCAGTA CAATCTGCTC TGATGCCGCA	2100
TAGTTAAGCC AGTATACT CCGCTATCGC TACGTGACTG GGTTCATGGCT GCGCCCCGAC	2160
ACCCGCCAAC ACCCGCTGAC GCGCCCTGAC GGGCTTGTCT GCTCCCGGCA TCCGCTTACA	2220
GACAAGCTGT GACCGTCTCC GGGAGCTGCA TGTGTCAGAG GTTTTCACCG TCATCACCGA	2280
AACCGCCGAG GCAGCTGCGG TAAAGCTCAT CAGCGTGGTC GTGAAGCGAT TCACAGATGT	2340
CTGCCCTGTC ATCCCGCTCC AGCTCGTTGA GTTTCTCCAG AAGCGTTAAT GTCTGGCTTC	2400
TGATAAAGCG GGCCATGTTA AGGGCGGTTT TTTCCCTGTTT GGTCACTGA TGCCTCCGTG	2460
TAAGGGGGAA TTTCTGTTCA TGGGGGTAAT GATACCGATG AAACGAGAGA GGATGCTCAC	2520
GATACGGGTT ACTGATGATG AACATGCCCG GTTACTGGAA CGTTGTGAGG GTAACAACCT	2580
GGCGGTATGG ATGCGGGGGG ACCAGAGAAA AATCACTCAG GGTCAATGCC AGCGCTTCGT	2640
TATACAGAT GTAGGTGTTT CACAGGTTAG CCAGCAGCAT CCTGCGATGC AGATCCGGAA	2700
CATAATGGTG CAGGGCGCTG ACTTCCGCTT TTCCAGACTT TACGARACAC GGAACCGGAA	2760
GACCATTCAT GTTGTGCTC AGGTCCGAGA CGTTTTGCAG CAGCAGTCCG TTCAGTTCCG	2820
CTCGCGTATC GGTGATTCAT TCTGCTAACC AGTAAGGCAA CCCC GCCAGC CTAGCCGGGT	2880
CCTCAACGAC AGGAGCACGA TCATGCCGAC CCGTGGCCAG GACCCAACGC TGCCCCAGAT	2940
GCGCCCGCTG CGGCTGCTGG AGATGGCGGA CGCGATGAT ATGTTCTGCC AAGGGTTGGT	3000
TTGCGCATTC ACAGTTCTCC GCAAGAATTG ATTGGCTCCA ATTCTTGGAG TGGTGAATCC	3060
GTTAGCGAGG TGCCGCCGGC TTCCATTGAG GTCGAGGTGG CCGGCTCCA TGCACCGCGA	3120
CGCAACGCGG GGAGGCAGAC AAGGTATAGG GCGGCGCTA CAATCCATGC CAACCCGTTT	3180
CATGTGCTCG CCGAGGCGGC ATAAATCGCC GTGACGATCA GCGGTCCAGT GATCGAAGTT	3240
AGGCTGGTAA GAGCCGCGAG CGATCCTTGA AGCTGTCCCT GATGGTCGTC ATCTACCTGC	3300
CTGGACAGCA TGGCCTGCAA CGCGGGCATC CCGATGCCGC CGGAAGCGAG AAGAATCATA	3360
ATGGGGAAGG CCATCCAGCC TCGCGTCCG AACGCCAGCA AGACGTAGCC CAGCCCGTCC	3420
GCCGCCATGC CGGCCATAAT GGCCTGCTT TCGCCGAAAC GTTTGGTGGC GGGACCACTG	3480
ACGAAGGCTT GAGCGAGGGC GTGCAAGATT CCGAATACCG CAAGCGACAG GCCGATCATC	3540
GTCCGCTCC AGCGAAAGCG GTCCTCGCCG AAAATGACCC AGAGCGCTGC CGGCACCTGT	3600

CCTACGAGTT GCATGATAAA GAAGACAGTC ATAAGTGCGG CGACGATAGT CATGCCCCGC	3660
GCCCACCGGA AGGAGCTGAC TGGGTTGAAG GCTCTCAAGG GCATCGGTCG ACGCTCTCCC	3720
TTATGCGACT CCTGCATTAG GAAGCAGCCC AGTAGTAGGT TGAGGCCGTT GAGCACCGCC	3780
GCCGCAAGGA ATGGTGCATG CAAGGAGATG GCGCCCAACA GTCCCCCGGC CACGGGGCCT	3840
GCCACCATAC CCACGCCGAA ACAAGCGCTC ATGAGCCCGA AGTGGCGAGC CCGATCTTCC	3900
CCATCGGTGA TGTCGGCGAT ATAGGGGCCA GCAACCGCAC CTGTGGCGCC GGTGATGCCG	3960
GCCACGATGC GTCCGGCGTA GAGGATCGAG ATCTCGATCC CGCGAAATTA ATACGACTCA	4020
CTATAGGGAG ACCACAACCG TTTCCCTCTA GAAATAATTT TGTTAACTT TAAGAAGGAG	4080
ATATATATGG AACCGGTCGT TTCTCTGTG GGGCCGGCGG ACAGCGGCGA CGCCCTGCTG	4140
GAGCGCAACT ATCCCACTGG CGCGGAGTTC CTCGGCGACG GCGGCGACGT CAGCTTCAGC	4200
ACCCGCGGCA CGCAGAACTG GACGGTGGAG CGGCTGCTCC AGGCGCACCG CCAACTGGAG	4260
GAGCGCGGCT ATGTGTTCGT CGGCTACCAC GGCACCTTCC TCGAAGCGGC GCAAAGCATC	4320
GTCTTCGGCG GGTGCGCGC GCGCAGCCAG GACCTCGACG CGATCTGGCG CGGTTTCTAT	4380
ATCGCCGGCG ATCCGGCGCT GGCCTACGCC TACGCCCAGG ACCAGGAACC CGACGCACGC	4440
GGCCGGATCC GCAACGGTGC CCTGCTGCGG GTCTATGTG CGCGCTCGAG CCTGCCGGGC	4500
TTCTACCGCA CCAGCCTGAC CCTGGCCGCG CCGGAGGCCG CGGGCGAGGT CGAACGGCTG	4560
ATCGGCCATC CGCTGCCGCT GCGCCTGGAC GCCATCACCG GCCCCGAGGA GGAAGGCGGG	4620
CGCCTGGAGA CCATTCTCGG CTGGCCGCTG GCCGAGCGCA CCGTGGTGAT TCCCTCGGCC	4680
ATCCCCACCG ACCCGCGCAA CGTCGGCGGC GACCTCGACC CGTCCAGCAT CCCCACAAAG	4740
GAACAGGCGA TCAGCGCCCT GCCGACTAC GCCAGCCAGC CCGGCAAACC GCCGCGCGAG	4800
GACCTGAAGT AACTGCCGCG ACCGGCCGGC TCCCTCGCA GGAGCCGGCC TTCTCGGGGC	4860
CTGGCCATAC ATCAGTTTTT CCTGATGCCA GCCCAATCGA ATATGAATTC	4910

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TATGGAACCG GTCGTTTCTC TGTCGGGCC

29

## (2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGACAGAGAA ACGACCGGTT CCA

23

## (2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4977 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTCTTGAAGA CGAAAGGGCC TCCTGATACG CCTATTTTTA TAGGTTAATG TCATGATAAT	60
AATGGTTTCT TAGACGTCAG GTGGCACTTT TCGGGGAAAT GTGCGCGGAA CCCCTATTTG	120
TTTATTTTTC TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT	180
GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATPCAA CATTTCCTG TCGCCCTTAT	240
TCCCTTTTTT GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT	300
AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTTAC ATCGAACTGG ATCTCRACAG	360
CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT CCAATGATGA GCACTTTTAA	420
AGTTCIGCTA TGTGGCGCGG TATTATCCCG TTTGACGCC GGGCAAGAGC AACTCGGTCC	480
CEGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCACTCACAG AAAAGCATCT	540
TACGGATGGC ATGACAGTAA GAGAATTATG CACTGCTGCC ATAACCATGA GTGATAACAC	600
TGCGGCCAAC TTA CTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTCGA	660

CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT	720
ACCAAACGAC GAGCGTGACA CCACGATGCC TGCAGCAATG GCAACAACGT TCGCCAAAC	780
ATTAAGTGGC GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC	840
GGATAAAGTT GCAGGACCAC TTCTGGCGTC GGCCCTTCCG GCTGGCTGCT TTATTGCTGA	900
TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGSTATCATT GCAGCACTGG GGCCAGATGG	960
TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG	1020
AAATAGACAG ATCGCTGAGA TAGGTGCCCTC ACTGATTAAG CATTGGTAAC TGTCAGACCA	1080
AGTTTACTCA TATATACTTT AGATTGATTT AAAACTTCAT TTTAATTTA AAAGGATCTA	1140
GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTCCA	1200
CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG	1260
CGTAATCTGC TGCTTGCAAA CAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTCGGGA	1320
TCAAGAGCTA CCAACTCTTT TTCGGAAGT AACTGGCTTC AGCAGAGCGC AGATACCAA	1380
TACTGTCTT CTAGTGTAGC CGTAGTTAGG CCACCCTTC AAGAACTCTG TAGCACCGCC	1440
TACATACCTC GCTCTGCTAA TCCTGTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCTG	1500
TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC	1560
GGGGGTTTCG TGCACACAGC CCAGCTTGGG GCGAACCACC TACCCGAAC TGAGATACCT	1620
ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC	1680
GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCCTG	1740
GTAICTTTAT AGTCTGTGCG GGTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG	1800
CTCGTCAGGG GGGCGGAGCC TATGGAAAA CGCCAGCAAC GCGGCCTTTT TACGGTTCCT	1860
GGCCTTTTGC TGGCCTTTG CTCACATGTT CTTTCTGCG TTATCCCCTG ATTCTGTGGA	1920
TAACCGTATT ACCGCCTTTG AGTGAGCTGA TACCGCTGCG CGCAGCCGAA CGACCGAGCG	1980
CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCTGATG CCGTATTTTC TCCTTACGCA	2040
TCTGTGCGGT ATTTACACACC GCATATATGG TGCACCTCTA GTACAATCTG CTCTGATGCC	2100
GCATAGTTAA GCCAGTATAC ACTCCGCTAT CGCTACGTA CTGGGTCTG GCTGGCCCC	2160
GACACCCGCC AACACCCGCT GACGCGCCCT GACGGGCTTG TCTGCTCCCG GCATCCGCTT	2220
ACAGACAAGC TGTGACCGTC TCCGGGAGCT GCATGTCTCA GAGGTTTTCA CCGTCATCAC	2280
CGAAACCGCG GAGGCAGCTG CGGTAAAGCT CATCAGCGTG GTCGTGAAGC GATTCACAGA	2340



TGTCFGCCTG TTCATCCGGC TCCAGCTCGT TGAGTTTCTC CAGAAGCGTT AATGTCTGGC	2400
TTCTGATAAA GCGGGCCATG TTRAGGGCCG TTTTTTCCTG TTTGGTCACT TGATGCCTCC	2460
GTGTAAGGGG GAATTTCTGT TCATGGGGGT AATGATACCG ATGAAACGAG AGAGGATGCT	2520
CACGATACGG GTTACTGATG ATGAACATGC CCGGTTACTG GAACGTTGTG AGGGTAAACA	2580
ACTGGCGGTA TGGATGCGGC GGGACCAGAG AAAATCACT CAGGGTCAAT GCCAGCGCTT	2640
CGTTAATACA GATGTAGGTG TTCCACAGGG TAGCCAGCAG CATCCTGCCA TGCAGATCCG	2700
GAACATAATG GTGCAGGGCG CTGACTTCCG CGTTTCCAGA CTTTACGAAA CACGGAAACC	2760
GAAGACCATT CATGTTGTTG CTCAGGTCCG AGACGTTTGG CAGCAGCAGT CGCTTCACGT	2820
TCGCTCGCGT ATCGGTGATT CATTCTGCTA ACCAGTAAGG CAACCCCGCC AGCCTAGCCG	2880
GSTCCTCAAC GACAGGAGCA CGATCATGCG CACCCGTGGC CAGGACCCAA CGCTGCCCGA	2940
GATGCGCCGC GTGCGGCTGC TGGAGATGGC GGACCGGATG GATATGTTCT GCCAAGGGTT	3000
GTTTTCGCA TTCACAGTTC TCCGCAAGAA TTGATTGGCT CCAATTCTTG GAGTGGTGAA	3060
TCCGTTAGCG AGGTGCCGCC GGCTTCCATT CAGGTCCAGG TGGCCCGGCT CCATGCACCG	3120
CGACCAACG CGGGGAGGCA GACAAGGTAT AGGGCGGGCG CTACAATCCA TGCCAACCCG	3180
TTCCATGTGC TCGCCGAGGC GGCATAATC GCCGTGACGA TCAGCGGTCC AGTGATCGAA	3240
GTTAGGCTGG TAAGAGCCGC GAGCGATCCT TGAAGCTGTC CCTGATGGTC GTCATCTACC	3300
TGCCTGGACA GCATGGCCTG CAACGCGGGC ATCCCGATGC CGCCGGAAGC GAGAAGAATC	3360
ATAATGGGGA AGGCCATCCA GCCTCGCGTC GCGAAGCCCA GCAAGACGTA GCCCAGCGCG	3420
TCGGCCGCCA TGCCCGCGAT AATGGCCTGC TTCTCGCCGA AACGTTTGGT GCGGGGACCA	3480
GTGACGAAG CTTGAGCGAG GCGGTGCAAG ATTCCGAATA CCGCAAGCGA CAGGCCGATC	3540
ATCGTCCGCG TCCAGCGAAA GCGGTCCCTG CCGAAAATGA CCCAGAGCCG TGCCGGCACC	3600
TGTCCTACGA GTTGCATGAT AAAGAAGACA GTCATAAGTG CGGCGACGAT AGTCATGCCC	3660
CGCGCCACC GGAAGGAGCT GACTGGGTTG AAGGCTCTCA AGGGCATCGG TCGACGCTCT	3720
CCCTTATGG ACTCCTGCAT TAGGAAGCAG CCCAGTAGTA GGTGAGGCC GTTGAGCACC	3780
GCCGCCCAA GGAATGGTGC ATGCAAGGAG ATGGCGCCCA ACAGTCCCC GGCACGGGG	3840
CCTGCCACCA TACCCACGCC GAAACAAGCG CTCATGAGCC CGAAGTGGCG AGCCCGATCT	3900
TCCCATCCG TGATGTCCGC GATATAGGCG CCAGCAACCG CACCTGTGGC GCCGGTGATG	3960
CCGGCCACGA TCGGTCCGGC GTAGAGGATC GAGATCTCGA TCCCGCGAAA TTAATACGAC	4020

TCACTATAGG GAGACCACAA CGGTTTCCT CTAGAAATAA TTTTGTAA CTTTAAGAAG	4080
GAGATATAAC ATGGTACCAG ACACCGGAAA CCCCTGCCAC ACCACTAAGT TGTTCACAG	4140
AGACTCAGTG GACAGTGCTC CAATCCTCAC TGCATTTAAC AGCTCACACA AAGGACGGAT	4200
TAACTGTAAT AGTAACACTA CACCCATAGT ACATTTAAA GGTGATGCTA ATACTTTAA	4260
ATGTTTAAGA TATAGATTA AAAAGCATTG TACATTGTAT ACTGCAGTGT CGTCTACATG	4320
GCATTGGACA GGACATAATG TAAAACATAA AAGTGCAATT GTTACACTTA CATATGATAG	4380
TGAATGGCAA CGTGACCAAT TTTTGTCTCA AGTTAAAATA CCAAAAATA TTACAGTGTC	4440
TACTGGATTT ATGTCTATAT GAGGATCCGG CTGCTAACAA AGCCCCGAAAG GAAGCTGAGT	4500
TGGCTGCTGC CACCGCTGAG CAATAACTAG CATAACCCCT TGGGGCCTCT AAACGGGTCT	4560
TGAGGGGTTT TTTGCTGAAA GGAGGAATA TATCCGGATA TCCACAGGAC GGTGTGGTC	4620
GCCATGATCG CGTAGTCGAT AGTGGCTCCA AGTAGCGAAG CGAGCAGGAC TGGCCGGCGG	4680
CCAAAGCGGT CGGACAGTGC TCCGAGAACC GGTGCGCATA GAAATTGCAT CAACGCATAT	4740
AGCGCTAGCA GCACGCCATA GTGACTGGCG ATGCTGTCCG AATGGACGAT ATCCCGCAAG	4800
AGCCCCGCA GTACCGGCAT AACCAAGCCT ATGCCTACAG CATCCAGGAT GACGGTGCCG	4860
AGGATGACGA TGAGCGCATT GTTAGATTC ATACACGGTG CCTGACTGCC TTAGCAATTT	4920
AACTGTGATA AACTACCGCA TTAAGCTTA TCGATGATAA GCTGTCAAAC ATGAGAA	4977

## (2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTCCCATGGT ACCAGACACC GGAAACC

27

## (2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGGGGATCCT CATATAGACA TAAATCC

27

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4977 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTCTTGAAGA CGAAAGGGCC TCGTGATACG CCTATTTTAA TAGGTTAATG TCATGATAAT	60
AATGGTTTCT TAGACGTCAG GTGGCACTTT TCGGGGAAAT GTGCGCGGAA CCCCTATTTG	120
TTTATTTTTC TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT	180
GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTCCCGTG TCGCCCTTAT	240
TCCCTTTTTT GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT	300
AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTTAC ATCGAACTGG ATCTCAACAG	360
CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT CCAATGATGA GCACTTTTAA	420
AGTTCGCTA TGTGGCGGGG TATTATCCCG TGTGACGCC GGGCAAGAGC AACTCGGTCC	480
CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT	540
TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC	600
TGCGGCCAAC TTA CTCTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTGCA	660
CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT	720
ACCAAACGAC GAGCGTGACA CCACGATGCC TGCAGCAATG GCAACAACGT TCGGCAAACT	780
ATTA ACTGGC GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC	840
GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTCCG GCTGGCTGGT TTATTGCTGA	900
TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GCCCAGATGG	960
TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG	1020

AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTGGTAAC TGTCAGACCA	1080
AGTTFIACCTCA TATATACTTT AGATTGATT AAAACTTCAT TTTAATTTA AAAGGATCTA	1140
GGTGAAGATC CTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTTCCA	1200
CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG	1260
CGTAATCTGC TGCTTGCAA CAAAAAACC ACCGCTACCA GCGGTGGTT GTTTGCCGGA	1320
TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTC AGCAGAGCGC AGATACCAA	1380
TACTGTCTCT CTAGTGTAGC CGTAGTTAGG CCACCACTC AARGAAGCTG TAGCACCGCC	1440
TACATACCTC GCTCTGCTAA TCCTGTACC AGTGGCTGCT GCCAGTGGC ATAAGTCGTG	1500
TCTTACCGGG TTGGACTCAA GACGATAGT ACCGGATAAG GCGCAGCGGT CCGGCTGAAC	1560
GGGGGGTTCG TGCACACAGC CCAGCTTGA GCGAAGCACC TACACCGAAC TGAGATACCT	1620
ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAGGCGG ACAGGTATCC	1680
GGTAGCGGC AGGGTGGAA CAGGAGAGCG CACGAGGGAG CTCCAGGGG GAAACGCCTG	1740
GTATCTTTAT AGTCCTGTG GGTTCGCCA CCTCTGACTT GAGCGTCGAT TTTGTGATG	1800
CTCGTCAGGG GGGCGGAGCC TATGAAAAA CGCCAGCAAC GCGGCCTTTT TACGGTCTCT	1860
GGCCTTTTGC TGCCCTTTTG CTCACATGTT CTTTCCTGGG TTATCCCTG ATTCTGTGGA	1920
TAACCGTATT ACCGCCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG	1980
CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCTGATG CCGTATTTT TCCTTACGCA	2040
TCTGTCCGGT ATTTACACACC GCATATATGG TGCACCTCA GTACAATCTG CTCTGATGCC	2100
GCATAGTTAA GCCAGTATAC ACTCCGCTAT CGCTACGTGA CTGGGTCATG GCTGCGCCCC	2160
GACACCCGCC AACACCCGCT GACCGCCCT GACGGGCTTG TCTGCTCCCG GCATCCGCTT	2220
ACAGACAAGC TGTGACCGTC TCCGGGAGCT GCATGTGTCA GAGGTTTTCA CCGTCATCAC	2280
CGAAACGCGC GAGGCAGCTG CGGTAAAGCT CATCAGCGTG GTCGTGAAGC GATTCACAGA	2340
TGTCGTCCCG TTCATCCCG TCCAGCTCGT TGAGTTTCTC CAGAAGCGTT AATGTCTGGC	2400
TTCTGATAAA GCGGGCCATG TTAAGGGCGG TTTTTCCTG TTTGGTCACT TGATGCCCTCC	2460
GTGTAAGGGG GAATTTCTGT TCATGGGGGT AATGATACCG ATGAAACGAG AGAGGATGCT	2520
CACGATACGG GTTACTGATG ATGAACATGC CCGGTTACTG GAACGTTGTG AGGTTAAACA	2580
ACTGGCGGTA TGGATCCGCC GGGACCAGAG AAAAATCACT CAGGTCAT GGCAGCGCTT	2640
CGTTAATACA GATGTAGGTG TTCCACAGGG TAGCCAGCAG CATCCTGCGA TGCAGATCCG	2700

GAACATAATG GTGCAGGGCG CTGACTTCCG CGTTTCCAGA CTTTACGAAA CACGGAAACC	2760
GAAGACCATT CATGTTGTG CTCAGGTCCG AGACGTTTTG CAGCAGCAGT CGCTTCACGT	2820
TCGCTCCGCT ATCGGTGATT CATTCTGCTA ACCAGTAAGG CAACCCCGCC AGCCTAGCCG	2880
GGTCCTCAAC GACAGGAGCA CGATCATGCG CACCCGTGGC CAGGACCCAA CGCTGCCCGA	2940
GATGCCCCGC GTGCCGCTGC TGGAGATGGC GGACGCGATG GATATGTTCT GCCAAGGGTT	3000
GGTTTGCGCA TTCACAGTTC TCCGCAAGAA TTGATTGGCT CCAATTCTTG GAGTGGTGAA	3060
TCCGTTIAGCG AGGTGCCGCC GGCTTCCATT CAGGTGAGG TGGCCCGGCT CCATGCACCG	3120
CGACGCAACG CGGGGAGGCA GACAAGGTAT AGGGCGGCGC CTACAATCCA TGCCAACCCG	3180
TTCCATGTGC TCGCCGAGGC GGCATAAATC GCCGTGACGA TCAGCGGTCC AGTGATCGAA	3240
GTTAGGCTGG TAAGAGCCGC GAGCGATCCT TGAAGCTGTC CCTGATGGTC GTCATCTACC	3300
TGCCTGGACA GCATGGCCTG CAACGCGGGC ATCCCGATGC CGCCGGAAGC GAGAAGAATC	3360
ATAATGGGGA AGGCCATCCA GCCTCGCGTC GCGAACGCCA GCAGACGTA GCCCAGCCCG	3420
TCGGCCGCCA TGCCGGCGAT AATGGCCTGC TTCTCGCCGA AACGTTGGT GGCGGGACCA	3480
GTGACGAAGG CTTGAGCGAG GGCCTGCAAG ATTCCGAATA CCGCAAGCGA CAGGCCGATC	3540
ATCGTCCGCG TCCAGCGAAA GCGGTCCCTCG CCGAAATGA CCCAGAGCCG TGCCGGCACC	3600
TGTCTACGA GTTGCATGAT AAAGAAGACA GTCATAAGTG CGGCGACGAT AGTCATGCCC	3660
CGCGCCACC GGAAGGAGCT GACTGGGTTG AAGGCTCTCA AGGGCATCGG TCGACGCTCT	3720
CCCTTATCGG ACTCCTGCAT TAGGAAGCAG CCCAGTAGTA GGTGAGGCC GTTGAGCACC	3780
GCCGCCGCAA GGAATGGTGC ATGCAAGGAG ATGGCGCCCA ACAGTCCCCC GGCCACGGGG	3840
CCTGCCACCA TACCCACGCC GAAACAAGCG CTCATGAGCC CGAAGTGGCG AGCCCCATCT	3900
TCCCCATCGG TGATGTCGGC GATATAGGCG CCAGCAACCG CACCTGTGGC GCCGGTGATG	3960
CCGGCCACGA TCGTCCGGC GTAGAGGATC GAGATCTCGA TCCCGCGAAA TTAATAGAC	4020
TCACTATAGG GAGACCACAA CGGTTTCCCT CTAGAAATAA TTTTGTAA CTTTAAGAAG	4080
GAGATATACC ATGGTACCAG ACACCGGAAA CCCCTGCCAC ACCACTAAGT TGTTGCACAG	4140
AGACTCAGTG GACAGTGCTC CAATCCTCAC TGCATTTAAC AGCTCACACA AAGGACGGAT	4200
TAACTGTAAT AGTAACCTA CACCCATAGT ACATTTAAAA GGTGATGCTA ATACTTTAAG	4260
ATCTTTAAGA TATAGATTA AAAAGCATT CACATTTGAT ACTGCAGTGT CGTCTACATG	4320
GCATTGGACA GGACATAATG TAAAACATAA AAGTGCATT GTTACACTTA CATATGATAG	4380

TGAATGGCAA CGTGACCAAT TTTTGTCTCA AGTTAAATA CCAAAAATA TTACAGTGTG	4440
TACTGGATTT ATGTCTATAT GAGGATCCGG CTGCTAACA AGCCCGAAG GAAGCTGACT	4500
TGGCTGCTGC CACCCGTGAG CAATAACTAG CATAACCCCT TGGGGCCTCT AACCGGCTCT	4560
TGAGGGGTTT TTTGCTGAAA GGAGGAATA TATCCGGATA TCCACAGGAC GGGTGTGGTC	4620
GCCATGATCG CGTAGTCGAT AGTGGCTCCA AGTAGCGAAG CGAGCAGGAC TGGGCGGGG	4680
CCAAGCGGT CGGACAGTGC TCCGAGAACG GGTGCGCATA GAAATTGCAT CAACGCATAT	4740
AGCGCTAGCA GCACGCCATA GTGACTGGCG ATGCTGTCCG AATGGACGAT ATCCCGCAAG	4800
AGGCCCCGCA GTACCGGCAT AACCAAGCCT ATGCCTACAG CATCCAGGGT GACGGTGCCG	4860
AGGATGACGA TGAGCGCATT GTTAGATTTT ATACACGGTG CCTGACTGCG TTAGCAATTT	4920
AACTGTGATA AACTACCGCA TTAAGCTTA TCGATGATAA GCTGTCAAAC ATGAGAA	4977

## (2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 59 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGACACTGCA GTATACAATG TAGAATGCTT TTTAAATCTA TATCTTAAAG ATCTTAAAG	59
--	----

## (2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCGTCGGCCG CCATGCCGGC GATAAT	26
------------------------------	----

## (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4819 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TTCTGAAGA CGAAAGGGCC TCGTGATACG CCTATTTTTA TAGGTTAATG TCATGATAAT	60
AATGGTTTCT TAGACGTCAG GTGGCACTTT TCGGGGAAAT GTGCCGGGAA CCCCTATTTG	120
TTTATTTTTT TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT	180
GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTTCCGTG TCGCCCTTAT	240
TCCCTTTTTT GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT	300
AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTTAC ATCGAACTGG ATCTCAACAG	360
CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT CCAATGATGA GCACCTTTAA	420
AGTTCTGCTA TGTGGCGCGG TATTATCCCG TGTTGACGCC GGGCAAGAGC AACTCGGTCC	480
CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT	540
TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC	600
TGCGGCCAAC TFACTTCTGA CAACGATCGG AGGACCGAAG GAGCTRACCG CTTTTTTGCA	660
CAACATGGGG GATCATGTAA CTCGCCTTGA TCCTTGGGAA CCGGAGCTGA ATGAAGCCAT	720
ACCAAACGAC GAGCGTGACA CCACGATGCC TGCAGCAATG GCAACAACGT TCGGCAAACT	780
ATTAAGTGGC GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC	840
GGATAAAGTT GCAGGACCAC TTCTGGGCTC GGGCCCTCCG GCTGGCTGGT TTATTGCTGA	900
TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG	960
TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG	1020
AARTAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTGGTAAC TGTCAGACCA	1080
AGTTACTCA TATATACTTT AGATTGATTT AAAACTTCAT TTTAATTTA AAAGGATCTA	1140
GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCTGTTCCA	1200
CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG	1260
CGTAATCTGC TGCTTGCAA CAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA	1320
TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGTACCAAA	1380

TACTGTCCCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC	1440
TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG	1500
TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC	1560
GGGGGGTTCC TGCACACAGC CCAGCTTGA GCGAACGACC TACACCGAAC TGAGATACCT	1620
ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC	1680
GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG	1740
GSTATCTTAT AGTCCTGTCC GGTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG	1800
CTCGTCAGGG GGGCGGAGCC TATGGAAAA CGCCAGCAAC GCGGCCTTTT TACGGTTCCT	1860
GGCCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCCTGCG TTATCCCTG ATTCTGTGGA	1920
TAACCGTATT ACCGCCTTTG AGTGAGCTGA TACCGCTCGC GCGAGCCGAA CGACCGAGCG	1980
CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCTGATG CGSTATTTTC TCCTTACGCA	2040
TCTGTGCGGT ATTTACACCC GCATATATGG TGCACCTCTCA GTACAATCTG CTCTGATGCC	2100
GCATAGTTAA GCCAGTATAC ACTCCGCTAT CGCTACGTGA CTGGGTCATG GCTGCCCCCC	2160
GACACCCGCC AACRCCCGCT GACGCGCCCT GACGGGCTTG TCTGCTCCCG GCATCCGCTT	2220
ACAGACAAGC TGTGACCGTC TCCGGGAGCT GCATGTGTCA GAGGTTTTCA CCGTCATCAC	2280
CGAAACGCGC GAGGCAGCTG CGGTAAAGCT CATCAGCGTG GTCGTGAAGC GATTCACAGA	2340
TGCTGCCTG TTCATCCGCG TCCAGCTCGT TGASTTTCTC CAGAAGCGT AATGTCTGGC	2400
TTCTGATAAA GCGGGCCATG TTAAGGGCGG TTTTTTCCTG TTTGGTCACT TGATGCCTCC	2460
GTGTAAGGGG GAATTTCTGT TCATGGGGT AATGATACCG ATGAAACGAG AGAGGATGCT	2520
CACGATACGG GTTACTGATG ATGAACATGC CCGTTACTG GAACGTTGTG AGGGTAAACA	2580
ACTGGCGGTA TGGATGCGGC GGGACCAGAG AAAAATCACT CAGGGTCAAT GCCAGCGCTT	2640
CGTTAATACA GATGTAGGTG TTCCACAGGG TAGCCAGCAG CATCCTCGGA TGCAGATCCG	2700
GAACATAATG GTGCAGGGCG CTGACTTCG CGTTCCAGA CTTTACGAAA CACGGAAACC	2760
GAAGACCATT CATGTTGTTG CTCAGGTCGC AGACGTTTTG CAGCAGCAGT CGCTTCACGT	2820
TCGCTCGCGT ATCGGTGATT CATTCTGCTA ACCAGTAAGG CAACCCCGCC AGCCTAGCCG	2880
GGTCTCAAC GACAGGAGCA CGATCATGCG CACCCGTGGC CAGGACCCAA CGCTGCCCGA	2940
GATGCCCGGC GTGCGGCTGC TGGAGATGGC GGACGCGATG GATATGTTCT GCCAAGGGTT	3000
GGTTTGGCA TTCACAGTTC TCCGCAAGAA TTGATTGGCT CCAATCTTG GAGTGGTGAA	3060



TCCGTTAGCG AGGTGCCGCC GGCTTCCATT CAGGTCGAGG TGGCCCGGCT CCATGCACCG	3120
CGACGCAACG CGGGGAGGCA GACAAGGTAT AGGGCGGGCG CTACAATCCA TGCCAACCCG	3180
TTCCATGTGC TCGCCGAGGC GGCATAAATC GCCGTGACGA TCAGCGGTCC AGTGATCGAA	3240
GTTAGGCTGG TAAGAGCCGC GAGCGATCCT TGAAGCTGTC CCTGATGGTC GTCATCTACC	3300
TGCCTGGACA GCATGGCCTG CAACGCGGGC ATCCCGATGC CGCCGGAAGC GAGAAGAATC	3360
ATAATGGGGA AGGCCATCCA GCCTCGCGTC GCGAACGCCA GCARGACGTA GCCCAGCGCG	3420
TCGGCCGCCA TGCCGGCGAT AATGGCCTGC TTCTCGCCGA AACGTTTGGT GCGGGGACCA	3480
GTGACGAAGG CTTGAGCGAG GCGGTGCAAG ATTCCGAATA CCGCAAGCGA CAGGCCGATC	3540
ATCGTGGGCG TCCAGCGAAA GCGGTCCTCG CCGAAAATGA CCCAGAGCGC TGCCGGCACC	3600
TGTCCTACGA GTTGCAFGAT AAAGAAGACA GTCATAAGTG CGGCGACGAT AGTCATGCCC	3660
CGCGCCACC GGAAGGAGCT GACTGGGTG AAGGCTCTCA AGGGCATCGG TCGACGCTCT	3720
CCCTTATGCG ACTCCTGCAT TAGGAAGCAG CCCAGTAGTA GGTTGAGGCC GTTGAGCACC	3780
GCCGCCGCAA GGAATGGTGC ATGCAAGGAG ATGGCGCCCA ACAGTCCCCC GGCCACGGGG	3840
CCTGCCACCA TACCCACGCC GAAACAAGCG CTCATGAGCC CGAAGTGGCG AGCCCGATCT	3900
TCCCCATCGG TGATGTGGC GATATAGGCG CCAGCAACCG CACCTGTGGC GCCGGTGATG	3960
CCGGCCACGA TCGTCCGGC GTAGAGGATC GAGATCTCGA TCCCGCGAAA TTAATACGAC	4020
TCACTATAGG GAGACCACAA CGGTTTCCCT CTAGAAATAA TTTGTTTAA CTTTAAGAAG	4080
GAGATATACA TATGGAACCG GTCGACCCGC GTCGGAACC ATGGAACAC CCCGGGTCCC	4140
AGCCGAAAAC CGCGTTCATC ACCAAAGCCC TAGGTATCTC TTACGGCCGT AAAAAACGTC	4200
GTCAGCGACG TCGTCCGCCG CAGGGATCCC AGACCCACCA GGTTCCTCTG TCTAAACAGT	4260
GATCAGCATT GGCTAGCATG ACTGGTGGAC AGCAAATGGG TCGCGGATCC GGCTGCTAAC	4320
AAAGCCCGAA AGGAAGCTGA GTTGGCTGCT GCCACCCGCTG AGCAATAACT AGCATAACCC	4380
CTTGGGGCCT CTAAACGGGT CTTGAGGGGT TTTTGTCTGA AAGGAGGAAC TATATCCGGA	4440
TATCCACAGG ACGGGTGTGG TCGCCATGAT CCGGTAGTCC ATAGTGGCTC CAAGTAGCGA	4500
AGCGAGCAGG ACTGGGCGGC GCCCAAAGCG GTCGGACAGT GCTCCGAGAA CGGGTGCGCA	4560
TAGAAATGTC ATCAACGCAT ATAGCGCTAG CAGCACGCCA TAGTACTGG CGATGCTGTC	4620
GGAATGGACG ATATCCCGCA AGAGGCCCGG CAGTACCGGC ATAACCAAGC CTATGCCTAC	4680
AGCATCCAGG GTGACSGTGC CGAGGATGAC GATGAGCGCA TTGTTAGATT TCATACACGG	4740

TGCCTGACTG CGTTAGCAAT TTAAGTGTGA TAACTACCG CATTAAAGCT TATCGATGAT 4800  
 AAGCTGTCAA ACATGAGAA 4819

## (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 45 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TTTACGGCCG TAAGAGATAC CTAGGGCTTT GGTGATGAAC GCGGT 45

## (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 5574 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TTCTGAAGA CGAAGGGCC TCGTGATACG CCTATTTTAA TAGGTTAATG TCATGATAAT 60  
 AATGGTTTCT TAGACGTCAG GTGGCACTTT TCGGGGAAAT GTGGCGGGAA CCCCTATTG 120  
 TTTATTTTTC TAATACATT CAAATATGTA TCCGCTCATG AGACRATAAC CCTGATAAAT 180  
 GCTCAATAA TATTGAAAAA GGAAGAGTAT GACTATTCAA CATTCCGTCG TCGCCCTTAT 240  
 TCCCTTTTTT GCGGCATTTT GCCTTCCTGT TTTGCTCAC CCAGAAACGC TGGTGAAGT 300  
 AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTAC ATCGAACTGG ATCTCAACAG 360  
 CGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT CCAATGATGA GCACTTTTAA 420  
 AGTCTGCTA TGTGGCGCGG TATTATCCCC GTTTGACGCC GGGCAAGAGC AACTCGGTG 480  
 CCGCATAAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT 540  
 TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC 600  
 TGCGGCCAAC TTAATTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTGCA 660

CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT	720
ACCAAACGAC GAGCGTGACA CCACGATGCC TGCAGCAATG GCAACAACGT TCGGCAAACT	780
ATTAAGTGGC GAACACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC	840
GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCTTCCG GCTGGCTGGT TTATTGCTGA	900
TAAATCTGGA GCGCGTGAGC GTGGGTCTCG CCGTATCATT GCAGCACTGG GGCAGATGG	960
TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGACT CAGGCAACTA TGGATGAACG	1020
AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTGGTAAC TGTCAGACCA	1080
AGTTTACTCA TATATACTTT AGATTGATTT AAAACTTCAT TTTTAATTA AAAGGATCTA	1140
GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAATCCCT TAACGTGAGT TTTCGTTCCA	1200
CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTTCTGCG	1260
CGTAATCTGC TGCTTGCAA CAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA	1320
TCAAGAGCTA CCAACTCTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAA	1380
TACTGTCCCT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC	1440
TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG	1500
TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC	1560
GGGGGGTTCC TGCACACAGC CCAGCTTGA GCGAACCACC TACACCGAAC TGAGATACCT	1620
ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC	1680
GSTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCTG	1740
GSTATTTTAT AGTCCTGTCC GGTTCGCCA CCTCTGACTT GAGCGTCGAT TTTGTGATG	1800
CTCGTCAGGG GGGCGGAGCC TATGGAAAA CGCCAGCAAC GCGGCCTTTT TACGGTTCCT	1860
GGCCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCTGCG TTATCCCCTG ATTCTGTGGA	1920
TRACCGTATF ACCGCCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG	1980
CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCTGATG CGGTATTTTC TCCTTAGCCA	2040
TCTGTGCGGT ATTTACACCC GCATATATGG TGCACCTCTA GTACAATCTG CTCTGATGCC	2100
GCATAGTTAA GCCAGTATAC ACTCCGCTAT CGCTACGTGA CTGGGTCATG GCTGCGCCCC	2160
GACACCCGCC AACACCCGCT GACCGCCCT GACGGGCTTG TCTGCTCCCG GCATCCGCTT	2220
ACAGACAAGC TGTGACCGTC TCCGGGAGCT GCATGTGTCA GAGGTTTTCA CGTCATCAC	2280
CGAAACGCGC GAGGCAGCTG CCGTAAAGCT CATCAGCGTG GTCGTGAAGC GATTCACAGA	2340

TGTCTGCCTG TTCATCCCGG TCCAGCTCCT TGAGTTTCTC CAGAAGCGTT AATGTCTGGC	2400
TTCTGATAAA GCGGGCCATG TTAAGGGCGG TTTTTTCCTG TTTGGTCACT TGATGCCTCC	2460
GTGTAAGGGG GAATTTCTGT TCATGGGGGT AATGATACCG ATGAAACGAG AGAGGATGCT	2520
CACGATACGG GTTACTGATG ATGAACATGC CCGGTTACTG GAACGTTGTG AGGGTAAACA	2580
ACTGGCGGTA TGGATGCGGC GGGACCAGAG AAAAATCACT CAGGGTCAAT GCCAGCGCTT	2640
CGTTAATACA GATGTAGGTG TTCCACAGGG TAGCCAGCAG CATCCTGCGA TGCAGATCCG	2700
GAACATAATG GTGCAGGGCG CTGACTTCCG CGTTCCAGA CTTTACGAAA CACGGAAACC	2760
GAAGACCATT CATGTTGTTG CTCAGGTCCG AGACGTTTTG CAGCAGCAGT CGCTTCACGT	2820
TCGCTCGCGT ATCGGTGATT CATTCTGCTA ACCAGTAAGG CAACCCCGCC AGCCTAGCCG	2880
GGTCCTCAAC GACAGGAGCA CGATCATGCG CACCCGTGGC CAGGACCCAA CGCTGCCCGA	2940
GATGCGCCCGC GTGCGGCTGC TGGAGATGGC GGACGCGATG GATATGTTCT GCCAAGGGTT	3000
GGTTTCCGCA TTCACAGTTC TCCGCAAGAA TTGATTGGCT CCAATTCTTG GAGTGGTGAA	3060
TCCGTTAGCG AGGTGCCGCC GGCTTCCATT CAGGTCCAGG TGGCCCGGCT CCATGCACCG	3120
CGACGCAACG CGGGGAGGCA GACAAGGTAT AGGGCGGCGC CTACAATCCA TGCCAACCCG	3180
TTCCATGTGC TCGCCGAGGC GGCATAAATC GCCGTGACGA TCAGCCGTCC AGTGATCGAA	3240
GTTAGGCTGG TAAGAGCCCG GAGCGATCCT TGAAGCTGTC CCTGATGGTC GTCATCTACC	3300
TGCCCTGGACA GCATGGCCTG CAACGCGGGC ATCCCGATGC CGCCGGAAGC GAGAAGAATC	3360
ATAATGGGGA AGGCCATCCA GCCTCGCGTC GCGAACGCCA GCAAGACGTA GCCCAGCGCG	3420
TCGGCCGCCA TGCCCGCGAT AATGGCCTGC TTCTCGCCGA AACGTTTGGT GCGGGGACCA	3480
GTGACGAAGG CTTGAGCGAG GGCSTGCAAG ATTCCGAATA CCGCAAGCGA CAGGCCGATC	3540
ATCGTCCGCG TCCAGCGAAA GCGGTCTCTG CCGAAAATGA CCCAGAGCGC TGCCGGCACC	3600
TGTCCTIACGA GTTGCATGAT AAGAAGACA GTCATAACTG CCGCGACGAT AGTCATGCCC	3660
CGCGCCCAAC GGAAGGAGCT GACTGGGTTG AAGGCTCTCA AGGGCATCGG TCGACGCTCT	3720
CCCTTATGCG ACTCCTGCAT TAGGAAGCAG CCCAGTAGTA GGTTGAGGCC GTTGAGCACC	3780
GCCGCGCAA GGAATGGTGC ATGCAAGGAG ATGGCGCCA ACAGTCCCCC GGCCACGGGG	3840
CCTGCCACCA TACCCACGCC GAAACAAGCG CTCATGAGCC CGAAGTGGCG AGCCCGATCT	3900
TCCCATCGG TGATGTCCGC GATATAGGCG CCAGCAACCG CACCTGTGGC GCCGGTGATG	3960
CCGGCCACGA TCGTCCGGC GTAGAGGATC GAGATCTCGA TCCCGCGAAA TTAATACGAC	4020

TCACTATAGG GAGACCACAA CGGTTTCCT CTAGAAATAA TTTTGTAA CTTTAAGAAG	4080
GAGATATACA TATGGAACCG GTCGACCCGC GTCTGGAACC ATGGAAACAC CCCGGGTCCC	4140
AGCCGAAAAC CCGTTCATC ACCAAAGCCC TAGGTATCTC TTACGGCCGT AAAAAACGTC	4200
GTCAGCGACG TCGTCCCGG CAGGGATCTT CCATGGCCCG TGCTGGACGC ATTTACTATT	4260
CTCGCTTTGG TGACGAGGCA GCCAGATTTA GTACAACAGG GCATTACTCT GTAAGAGATC	4320
AGGACAGAGT GTATGCTGGT GTCTCATCCA CCTCTTCTGA TTTTAGAGAT CGCCCAGACG	4380
GAGTCTGGGT CGCATCCGAA GGACCTGAAG GAGACCCTGC AGGAAAAGAA GCCGAGCCAG	4440
CCCAGCCTGT CTCTTCTTTG CTCGGCTCCC CCGCCTCGG TCCCATCAGA GCAGGCCTCG	4500
GTGGGTACG GGACGGTCTT CGTCCGACC CCTACAATT TCGTGCAGGC TCGGGGGGCT	4560
CTATTCTCCG CTCTTCTCC ACCCCGGTGC AGGGCAGGT ACCGGTGGAC TTGGCATCAA	4620
GGCAGGAAGA AGAGGAGCAG TCGCCCGACT CCACAGAGGA AGAACCAGTG ACTCTCCCAA	4680
GGCGCACAC CAATGATGGA TTCCACCTGT TAAAGGCAGG AGGGTCATGC TTTGCTCTAA	4740
TTTCAGGAAC TGCTAACCAG GTAAAGTGCT ATCGCTTTCC GGTGAAAAG AACCATAGAC	4800
ATCGCTACGA GAACCTGCACC ACCACCTGGT TCACAGTTGC TGACAACGGT GCTGAAAGAC	4860
AAGGACAAGC ACAAATACTG ATCACCTTTG GATCGCCAAG TCAAAGGCAA GACTTTCTGA	4920
AACATGTACC ACTACCTCCT GGAATGAACA TTTCCGGCTT TACAGCCAGC TTGGACTTCT	4980
GATCACTGCC ATTGCCTTTT CTCATCTGA CTGGTGTACT ATGCCAAATC TATGGTTTCT	5040
ATTGTTCTTG GGACTAGGAA GATCCGGCTG CTAACAAAGC CCGAAAGGAA GCTGAGTTGG	5100
CTGCTGCCAC CGCTGAGCAA TAACTAGCAT AACCCCTGG GGCCTCTAA CGGGTCTTGA	5160
GGGGTTTTTT GCTGAAAGGA GGAACTATAT CCGGATATCC ACAGGACGGG TGTGCTCGCC	5220
ATGATCGCGT AGTCGATAGT GGCTCCAAGT AGCGAAGCGA GCAGGACTGG GCGCGGCCA	5280
AAGCGGTCCG ACAGTGCTCC GAGAACGGGT GCGCATAGAA ATTGCATCAA CGCATATAGC	5340
GCTAGCAGCA CGCCATAGTG ACTGGCGATG CTGTCCGAAT GGACGATATC CCGCAAGAGG	5400
CCCGGCAGTA CCGGCATAAC CAAGCCTATG CCTACAGCAT CCAGGGTGAC GGTGCCGAGG	5460
ATGACGATGA GCGCATTGTT AGATTTTATA CACGGTGCCT GACTGCGTTA GCAATTTAAC	5520
TGTGATAAAC TACCGCATTA AAGCTTATCG ATGATAAGCT GTCAAACATG AGAA	5574

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GATCCCAGAC CCACCAGTT

20

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GAACCTGCTG GGTCTGG

17

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 50 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGTCCGCCGC AGGGATCGCA GACCCACCAG GTTTCTCTGT CTAAACAGGC

50

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 58 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CATGGCCTGT TTAGACAGAG AAACCTGGTG GGTCTGGGAT CCCTGCGGCG GACGACGT 58

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 48 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CATGTACGGC CGTAAAAAAC GTCGTCAGCG ACGTCGTCGG CCGGACAC 48

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 46 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CGGTGTCCGG CGGACGACGT CGCTGACGAC GTTTTTTACG GCCGTA 46

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ATCATCGATA AGCTTTAATG CGGTAG 26

(2) INFORMATION FOR SEQ ID NO:30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ACTTTAAGAA GGAGATATAC ATATGTTTAT CACCAAAGCC CTAGGTATCT CT

52

## (2) INFORMATION FOR SEQ ID NO:31:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ACTTTAAGAA GGAGATATAC ATATGTACCG CCGTAAAAAA CGTCGTCAGC G

51

## (2) INFORMATION FOR SEQ ID NO:32:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

AACGTCGTCA GCGACGTCGT CCGCCGGACA CCGGAAACCC CTGCCACACC AC

52

## (2) INFORMATION FOR SEQ ID NO:33:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CGAAAAGTGC CACCTGACGT CTAAGAAACC

30

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CTCCCATGGC TAGCAACACT ACACCC

26

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GAAGATCTTC

10

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CAGAGGAAGC CATGGTGACT CTCCCAA

27

(2) INFORMATION FOR SEQ ID NO:37:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

AAGGCAATGG ATCCGATCAG AAGTCCA

27

## (2) INFORMATION FOR SEQ ID NO:38:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 134 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Met Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Asp Thr  
 1                    5                    10                    15

Gly Asn Pro Cys His Thr Thr Lys Leu Leu His Arg Asp Ser Val Asp  
                   20                    25                    30

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

Asn Cys Asn Ser Asn Thr Thr Pro Ile Val His Leu Lys Gly Asp Ala  
   50                    55                    60

Asn Thr Leu Lys Cys Leu Arg Tyr Arg Phe Lys Lys His Cys Thr Leu  
 65                    70                    75                    80

Tyr Thr Ala Val Ser Ser Thr Trp His Trp Thr Gly His Asn Val Lys  
                   85                    90                    95

His Lys Ser Ala Ile Val Thr Leu Thr Tyr Asp Ser Glu Trp Gln Arg  
                   100                    105                    110

Asp Gln Phe Leu Ser Gln Val Lys Ile Pro Lys Thr Ile Thr Val Ser  
 115 120 125

Thr Gly Phe Met Ser Ile  
 130

## (2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 55 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CATGTACGGC CGTAAAAAC GTCGTCAGCG ACGTCGTCGG CTGAGTCAGG CCCAG 55

## (2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 51 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CTGGGCCTGA CTCAGCGGAC GACGTCGCTG ACGACGTTTT TTACGGCCGT A 51

## (2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 46 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

TCCTTCCTGT CCGCTGGTCA GCGCCCGCGC CGCCTGTCCA CCTAAG 46

## (2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 54 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:-

AATTCTTAGG TGGACAGGCG GCGCGGCGC TGACCAGCGG ACAGGAAGGA CATG

54

## (2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GGGGACTTTC CGCTGGGGAC TTTCCACGGG GGACTTTCC

39

## (2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GGAAAGTCCC CCGTGGAAG TCCCAGCGG AAAGTCCC

39

## (2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GTCTACTTTC CGCTGTCTAC TTTCCACGGT CTACTTTCC

39

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GGAAAGTAGA CCGTGAAAG TAGACAGCGG AAAGTAGAC

39

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro  
1                   5                           10

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln Gly Ser  
 1                                5                                10                                15

Gln Thr His Gln Val Ser Leu Ser Lys Gln  
                               20                                25

## (2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 35 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg Arg  
 1                                5                                10                                15

Gln Arg Arg Arg Pro Pro Gln Gly Ser Gln Thr His Gln Val Ser Leu  
                               20                                25                                30

Ser Lys Gln  
                               35

## (2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg Arg  
 1                                5                                10                                15

Gln Arg Arg Arg Pro  
                               20

## (2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 121 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Pro Asp Thr Gly Asn Pro Cys His Thr Thr Lys Leu Leu His Arg Asp  
 1                   5                   10                   15

Ser Val Asp Ser Ala Pro Ile Leu Thr Ala Phe Asn Ser Ser His Lys  
                  20                   25                   30

Gly Arg Ile Asn Cys Asn Ser Asn Thr Thr Pro Ile Val His Leu Lys  
          35                   40                   45

Gly Asp Ala Asn Thr Leu Lys Cys Leu Arg Tyr Arg Phe Lys Lys His  
   50                   55                   60

Cys Thr Leu Tyr Thr Ala Val Ser Ser Thr Trp His Trp Thr Gly His  
 65                   70                   75                   80

Asn Val Lys His Lys Ser Ala Ile Val Thr Leu Thr Tyr Asp Ser Glu  
                  85                   90                   95

Trp Gln Arg Asp Gln Phe Leu Ser Gln Val Lys Ile Pro Lys Thr Ile  
          100                   105                   110

Thr Val Ser Thr Gly Phe Met Ser Ile  
       115                   120

(2) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 15 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln Gly Ser  
 1 5 10 15

## (2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg Arg  
 1 5 10 15

Gln Arg Arg Arg Pro Pro Gln Gly Ser  
 20 25

## (2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 85 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Cys Asn Ser Asn Thr Thr Pro Ile Val His Leu Lys Gly Asp Ala Asn  
 1 5 10 15

Thr Leu Lys Cys Leu Arg Tyr Arg Phe Lys Lys His Cys Thr Leu Tyr  
 20 25 30

Thr Ala Val Ser Ser Thr Trp His Trp Thr Gly His Asn Val Lys His  
 35 40 45

Lys Ser Ala Ile Val Thr Leu Thr Tyr Asp Ser Glu Trp Gln Arg Asp  
 50 55 60



Gln Phe Leu Ser Gln Val Lys Ile Pro Lys Thr Ile Thr Val Ser Thr  
 65 70 75 80

Gly Phe Met Ser Ile  
 85

## (2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 121 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Pro Asp Thr Gly Asn Pro Cys His Thr Thr Lys Leu Leu His Arg Asp  
 1 5 10 15

Ser Val Asp Ser Ala Pro Ile Leu Thr Ala Phe Asn Ser Ser His Lys  
 20 25 30

Gly Arg Ile Asn Cys Asn Ser Asn Thr Thr Pro Ile Val His Leu Lys  
 35 40 45

Gly Asp Ala Asn Thr Leu Lys Ser Leu Arg Tyr Arg Phe Lys Lys His  
 50 55 60

Ser Thr Leu Tyr Thr Ala Val Ser Ser Thr Trp His Trp Thr Gly His  
 65 70 75 80

Asn Val Lys His Lys Ser Ala Ile Val Thr Leu Thr Tyr Asp Ser Glu  
 85 90 95

Trp Gln Arg Asp Gln Phe Leu Ser Gln Val Lys Ile Pro Lys Thr Ile  
 100 105 110

Thr Val Ser Thr Gly Phe Met Ser Ile  
 115 120

## (2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 161 amino acids

(B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

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

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 249 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

```

Met Ala Gly Ala Gly Arg Ile Tyr Tyr Ser Arg Phe Gly Asp Glu Ala
 1          5          10          15

Ala Arg Phe Ser Thr Thr Gly His Tyr Ser Val Arg Asp Gln Asp Arg
          20          25          30

Val Tyr Ala Gly Val Ser Ser Thr Ser Ser Asp Phe Arg Asp Arg Pro
          35          40          45

Asp Gly Val Trp Val Ala Ser Glu Gly Pro Glu Gly Asp Pro Ala Gly
          50          55          60

Lys Glu Ala Glu Pro Ala Gln Pro Val Ser Ser Leu Leu Gly Ser Pro
65          70          75          80

Ala Cys Gly Pro Ile Arg Ala Gly Leu Gly Trp Val Arg Asp Gly Pro
          85          90          95

Arg Ser His Pro Tyr Asn Phe Pro Ala Gly Ser Gly Gly Ser Ile Leu
          100          105          110

Arg Ser Ser Ser Thr Pro Val Gln Gly Thr Val Pro Val Asp Leu Ala
          115          120          125

Ser Arg Gln Glu Glu Glu Glu Gln Ser Pro Asp Ser Thr Glu Glu Glu
          130          135          140

Pro Val Thr Leu Pro Arg Arg Thr Thr Asn Asp Gly Phe His Leu Leu
          145          150          155          160

Lys Ala Gly Gly Ser Cys Phe Ala Leu Ile Ser Gly Thr Ala Asn Gln
          165          170          175

```

Val Lys Cys Tyr Arg Phe Arg Val Lys Lys Asn His Arg His Arg Tyr  
 180 185 190

Glu Asn Cys Thr Thr Thr Trp Phe Thr Val Ala Asp Asn Gly Ala Glu  
 195 200 205

Arg Gln Gly Gln Ala Gln Ile Leu Ile Thr Phe Gly Ser Pro Ser Gln  
 210 215 220

Arg Gln Asp Phe Leu Lys His Val Pro Leu Pro Pro Gly Met Asn Ile  
 225 230 235 240

Ser Gly Phe Thr Ala Ser Leu Asp Phe  
 245

(2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 385 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Met Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Leu Ser Gln  
 1 5 10 15

Ala Gln Leu Met Pro Ser Pro Pro Met Pro Val Pro Pro Ala Ala Leu  
 20 25 30

Phe Asn Arg Leu Leu Asp Asp Leu Gly Phe Ser Ala Gly Pro Ala Leu  
 35 40 45

Cys Thr Met Leu Asp Thr Trp Asn Glu Asp Leu Phe Ser Gly Phe Pro  
 50 55 60

Thr Asn Ala Asp Met Tyr Arg Glu Cys Lys Phe Leu Ser Thr Leu Pro  
 65 70 75 80

Ser Asp Val Ile Asp Trp Gly Asp Ala His Val Pro Glu Arg Ser Pro  
 85 90 95

Ile Asp Ile Arg Ala His Gly Asp Val Ala Phe Pro Thr Leu Pro Ala  
 100 105 110

Thr Arg Asp Glu Leu Pro Ser Tyr Tyr Glu Ala Met Ala Gln Phe Phe  
 115 120 125

Arg Gly Glu Leu Arg Ala Arg Glu Glu Ser Tyr Arg Thr Val Leu Ala  
 130 135 140

Asn Phe Cys Ser Ala Leu Tyr Arg Tyr Leu Arg Ala Ser Val Arg Gln  
 145 150 155 160

Leu His Arg Gln Ala His Met Arg Gly Arg Asn Arg Asp Leu Arg Glu  
 165 170 175

Met Leu Arg Thr Thr Ile Ala Asp Arg Tyr Tyr Arg Glu Thr Ala Arg  
 180 185 190

Leu Ala Arg Val Leu Phe Leu His Leu Tyr Leu Phe Leu Ser Arg Glu  
 195 200 205

Ile Leu Trp Ala Ala Tyr Ala Glu Gln Met Met Arg Pro Asp Leu Phe  
 210 215 220

Asp Gly Leu Cys Cys Asp Leu Glu Ser Trp Arg Gln Leu Ala Cys Leu  
 225 230 235 240

Phe Gln Pro Leu Met Phe Ile Asn Gly Ser Leu Thr Val Arg Gly Val  
 245 250 255

Pro Val Glu Ala Arg Arg Leu Arg Glu Leu Asn His Ile Arg Glu His  
 260 265 270

Leu Asn Leu Pro Leu Val Arg Ser Ala Ala Ala Glu Glu Pro Gly Ala  
 275 280 285

Pro Leu Thr Thr Pro Pro Val Leu Gln Gly Asn Gln Ala Arg Ser Ser  
 290 295 300

Gly Tyr Phe Met Leu Leu Ile Arg Ala Lys Leu Asp Ser Tyr Ser Ser  
 305 310 315 320



Leu Lys Cys Leu Arg Tyr Arg Phe Lys Lys His Cys Thr Leu Tyr Thr  
 85 90 95

Ala Val Ser Ser Thr Trp His Trp Thr Gly His Asn Val Lys His Lys  
 100 105 110

Ser Ala Ile Val Thr Leu Thr Tyr Asp Ser Glu Trp Gln Arg Asp Gln  
 115 120 125

Phe Leu Ser Gln Val Lys Ile Pro Lys Thr Ile Thr Val Ser Thr Gly  
 130 135 140

Phe Met Ser Ile  
 145

(2) INFORMATION FOR SEQ ID NO:60:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 157 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Met Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg  
 1 5 10 15

Arg Gln Arg Arg Arg Pro Pro Gln Gly Ser Gln Thr His Gln Val Ser  
 20 25 30

Leu Ser Lys Gln Pro Asp Thr Gly Asn Pro Cys His Thr Thr Lys Leu  
 35 40 45

Leu His Arg Asp Ser Val Asp Ser Ala Pro Ile Leu Thr Ala Phe Asn  
 50 55 60

Ser Ser His Lys Gly Arg Ile Asn Cys Asn Ser Asn Thr Thr Pro Ile  
 65 70 75 80

Val His Leu Lys Gly Asp Ala Asn Thr Leu Lys Cys Leu Arg Tyr Arg  
 85 90 95

Phe Lys Lys His Cys Thr Leu Tyr Thr Ala Val Ser Ser Thr Trp His  
 100 105 110

Trp Thr Gly His Asn Val Lys His Lys Ser Ala Ile Val Thr Leu Thr  
 115 120 125

Tyr Asp Ser Glu Trp Gln Arg Asp Gln Phe Leu Ser Gln Val Lys Ile  
 130 135 140

Pro Lys Thr Ile Thr Val Ser Thr Gly Phe Met Ser Ile  
 145 150 155

(2) INFORMATION FOR SEQ ID NO:61:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 177 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

Met Gly Arg Lys Lys Arg Arg Gln Arg Arg Pro Pro Gln Gly Ser  
 1 5 10 15

Leu Gly Trp Val Arg Asp Gly Pro Arg Ser His Pro Tyr Asn Phe Pro  
 20 25 30

Ala Gly Ser Gly Gly Ser Ile Leu Arg Ser Ser Ser Thr Pro Val Gln  
 35 40 45

Gly Thr Val Pro Val Asp Leu Ala Ser Arg Gln Glu Glu Glu Gln  
 50 55 60

Ser Pro Asp Ser Thr Glu Glu Glu Pro Val Thr Leu Pro Arg Arg Thr  
 65 70 75 80

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



Leu Ile Ser Gly Thr Ala Asn Gln Val Lys Cys Tyr Arg Phe Arg Val  
 100 105 110

Lys Lys Asn His Arg His Arg Tyr Glu Asn Cys Thr Thr Thr Trp Phe  
 115 120 125

Thr Val Ala Asp Asn Gly Ala Glu Arg Gln Gly Gln Ala Gln Ile Leu  
 130 135 140

Ile Thr Phe Gly Ser Pro Ser Gln Arg Gln Asp Phe Leu Lys His Val  
 145 150 155 160

Pro Leu Pro Pro Gly Met Asn Ile Ser Gly Phe Thr Ala Ser Leu Asp  
 165 170 175

Phe

(2) INFORMATION FOR SEQ ID NO:62:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 187 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(\*1) SEQUENCE DESCRIPTION: SEQ ID NO:62:

Met Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg  
 1 5 10 15

Arg Gln Arg Arg Arg Pro Pro Gln Gly Ser Leu Gly Trp Val Arg Asp  
 20 25 30

Gly Pro Arg Ser His Pro Tyr Asn Phe Pro Ala Gly Ser Gly Gly Ser  
 35 40 45

Ile Leu Arg Ser Ser Ser Thr Pro Val Gln Gly Thr Val Pro Val Asp  
 50 55 60

Leu Ala Ser Arg Gln Glu Glu Glu Glu Gln Ser Pro Asp Ser Thr Glu  
 65 70 75 80

Glu Glu Pro Val Thr Leu Pro Arg Arg Thr Thr Asn Asp Gly Phe His  
85 90 95

Leu Leu Lys Ala Gly Gly Ser Cys Phe Ala Leu Ile Ser Gly Thr Ala  
100 105 110

Asn Gln Val Lys Cys Tyr Arg Phe Arg Val Lys Lys Asn His Arg His  
115 120 125

Arg Tyr Glu Asn Cys Thr Thr Thr Trp Phe Thr Val Ala Asp Asn Gly  
130 135 140

Ala Glu Arg Gln Gly Gln Ala Gln Ile Leu Ile Thr Phe Gly Ser Pro  
145 150 155 160

Ser Gln Arg Gln Asp Phe Leu Lys His Val Pro Leu Pro Pro Gly Met  
165 170 175

Asn Ile Ser Gly Phe Thr Ala Ser Leu Asp Phe  
180 185

(2) INFORMATION FOR SEQ ID NO:63:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 143 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

Met Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg  
1 5 10 15

Arg Gln Arg Arg Arg Pro Pro Asp Thr Gly Asn Pro Cys His Thr Thr  
20 25 30

Lys Leu Leu His Arg Asp Ser Val Asp Ser Ala Pro Ile Leu Thr Ala  
35 40 45

Phe Asn Ser Ser His Lys Gly Arg Ile Asn Cys Asn Ser Asn Thr Thr  
 50 55 60

Pro Ile Val His Leu Lys Gly Asp Ala Asn Thr Leu Lys Cys Leu Arg  
 65 70 75 80

Tyr Arg Phe Lys Lys His Cys Thr Leu Tyr Thr Ala Val Ser Ser Thr  
 85 90 95

Trp His Trp Thr Gly His Asn Val Lys His Lys Ser Ala Ile Val Thr  
 100 105 110

Leu Thr Tyr Asp Ser Glu Trp Gln Arg Asp Gln Phe Leu Ser Gln Val  
 115 120 125

Lys Ile Pro Lys Thr Ile Thr Val Ser Thr Gly Phe Met Ser Ile  
 130 135 140

**Patenttivaatimukset**

1. Fuusioproteiini, joka koostuu karboksiterminaalista lastiosasta ja aminotermiinaalisesta kuljetusosasta,  
5 tunnettu siitä, että

a) kuljetusosassa on:

i) läsnä HIV:n tat-proteiinin aminohapot 49 - 57,

ii) poissa HIV:n tat-proteiinin aminohapot 22 - 36 ja

iii) poissa HIV:n tat-proteiinin aminohapot 73 - 86,

10 ja

b) lastiosassa säilyy biologinen aktiivisuus sen jälkeen, kun se on viety kuljetusosasta riippuvaisella tavalla solun sisään.

2. Patenttivaatimuksen 1 mukainen fuusioproteiini,  
15 tunnettu siitä, että lastiosa on joko terapeuttinen molekyyli, profylaktinen molekyyli tai diagnostinen molekyyli.

3. Patenttivaatimuksen 1 tai 2 mukainen fuusioproteiini, tunnettu siitä, että lastiosa koostuu ihmisen papilloomaviruksen E2-repressorista, joka säilyttää biologisen aktiivisuutensa sen jälkeen, kun se on viety kohdesoluun, ja kuljetusosa on joko:

a) HIV:n tat-proteiinin aminohapot 47 - 58 (sekvenssi tunnusnumero 47),

25 b) HIV:n tat-proteiinin aminohapot 47 - 72 (sekvenssi tunnusnumero 48),

c) HIV:n tat-proteiinin aminohapot 38 - 72 (sekvenssi tunnusnumero 49) tai

30 d) HIV:n tat-proteiinin aminohapot 38 - 58 (sekvenssi tunnusnumero 50).

4. Patenttivaatimuksen 3 mukainen fuusioproteiini, tunnettu siitä, että kuljetusosaa edeltää aminotermiinaalinen metioniini.

5. Minkä tahansa patenttivaatimuksen 1 - 4 mukainen fuusioproteiini, tunnettu siitä, että lastiosa koostuu ihmisen papilloomaviruksen E2-proteiinin aminohapoista 245 - 365 (sekvenssi tunnusnumero 51).

5           6. Patenttivaatimuksen 1 mukainen fuusioproteiini, tunnettu siitä, että se on fuusioproteiini JB106 (sekvenssi tunnusnumero 38).

10           7. Patenttivaatimuksen 1 mukainen fuusioproteiini, tunnettu siitä, että se on fuusioproteiini JB117 (sekvenssi tunnusnumero 59).

8. Patenttivaatimuksen 1 mukainen fuusioproteiini, tunnettu siitä, että se on fuusioproteiini JB118 (sekvenssi tunnusnumero 60).

15           9. Patenttivaatimuksen 1 mukainen fuusioproteiini, tunnettu siitä, että se on fuusioproteiini JB122 (sekvenssi tunnusnumero 63).

20           10. Patenttivaatimuksen 1 mukainen fuusioproteiini, tunnettu siitä, että lastiosa koostuu naudan papilloomaviruksen E2-repressorista, joka säilyttää biologisen aktiivisuuteensa kohdesoluun viemisen jälkeen, ja kuljetusosa on joko:

a) HIV:n tat-proteiinin aminohapot 47 - 62 (sekvenssi tunnusnumero 52) tai

25           b) HIV:n tat-proteiinin aminohapot 38 - 62 (sekvenssi tunnusnumero 53).

11. Patenttivaatimuksen 10 mukainen fuusioproteiini, tunnettu siitä, että kuljetusosaa edeltää aminoterminaalinen metioniini.

30           12. Minkä tahansa patenttivaatimuksen 1, 2, 10 tai 11 mukainen fuusioproteiini, tunnettu siitä, että lastiosa on E2-repressorin, joka koostuu naudan papilloomaviruksen E2-proteiinin aminohapoista 250 - 410 (sekvenssi tunnusnumero 56).

35           13. Patenttivaatimuksen 12 mukainen fuusioproteiini, tunnettu siitä, että se on JB119 (sekvenssi tunnusnumero 61).

14. Patenttivaatimuksen 12 mukainen fuusioproteiini, tunnettu siitä, että se on JB120 (sekvenssi tunnusnumero 62).

5 15. Patenttivaatimuksen 1 tai 2 mukainen fuusioproteiini, tunnettu siitä, että lastiosa koostuu HSV:n VP16-proteiinin aminohapoista 43 - 412 ja kuljetusosa koostuu HIV:n tat-proteiinin aminohapoista 47 - 58.

10 16. Patenttivaatimuksen 15 mukainen fuusioproteiini, tunnettu siitä, että kuljetusosaa edeltää aminoterminaalinen metioniini.

17. Minkä tahansa patenttivaatimuksen 1 - 16 mukaisen fuusioproteiinin käyttö lastin viemiseen solun sisään ei-terapeuttisessa tarkoituksessa.

15 18. Kovalenttisesti liitetty kemiallinen konjugaatti, joka koostuu kuljetuspolyeptidiosasta ja lastiosasta, tunnettu siitä, että

a) konjugaatin kuljetuspolyeptidiosassa on:

i) läsnä HIV:n tat-proteiinin aminohapot 49 - 57,

ii) poissa HIV:n tat-proteiinin aminohapot 22 - 36 ja

20 iii) poissa HIV:n tat-proteiinin aminohapot 73 - 86,  
ja

b) konjugaatin lastiosassa säilyy biologinen aktiivisuus sen jälkeen, kun se on viety solun sisään, kuljetusosasta riippuvaisella tavalla.

25 19. Patenttivaatimuksen 18 mukainen kovalenttisesti liitetty kemiallinen konjugaatti, tunnettu siitä, että kuljetuspolyeptidiosa koostuu HIV:n tat-proteiinin aminohapoista 37 - 72 (sekvenssi tunnusnumero 2).

30 20. Patenttivaatimuksen 19 mukainen kovalenttisesti liitetty kemiallinen konjugaatti, tunnettu siitä, että lastiosa on:

a) ihmisen papilloomaviruksen E2-proteiinin aminohapot 245 - 365 (sekvenssi tunnusnumero 51) tai

35 b) ihmisen papilloomaviruksen E2-proteiinin aminohapot 245 - 365, jossa aminohapot 300 ja 309 on muutettu kysteiniiksi (sekvenssi tunnusnumero 55).

21. Patenttivaatimuksen 20 mukainen kovalenttisesti liitetty kemiallinen konjugaatti, tunnettu siitä, että lastiosa on kaksijuosteinen DNA, joka on:

5 a) oligonukleotidi NF1 (sekvenssi tunnusnumero 43),  
joka on hybridisoitu oligonukleotidiin NF2 (SEQ ID nro 44)  
tai

b) oligonukleotidi NF3 (sekvenssi tunnusnumero 45),  
joka on hybridisoitu oligonukleotidiin NF4 (sekvenssi tun-  
nusnumero 46).

10 22. Minkä tahansa patenttivaatimuksen 18 - 21 mu-  
kaisen kovalenttisesti liitetyn kemiallisen konjugaatin  
käyttö lastin viemiseen solun sisään ei-terapeuttisessa tar-  
koituksessa.

15 23. DNA-molekyyli, joka käsittää nukleotidisekvens-  
sin, joka koodittaa patenttivaatimuksen 1 mukaista fuusio-  
proteiinia, joka on:

a) JB106 (sekvenssi tunnusnumero 38),  
b) JB117 (sekvenssi tunnusnumero 59),  
c) JB118 (sekvenssi tunnusnumero 60),  
20 d) JB119 (sekvenssi tunnusnumero 61),  
e) JB120 (sekvenssi tunnusnumero 62),  
f) JB122 (sekvenssi tunnusnumero 63) tai  
g) tat-VP16R.GF (sekvenssi tunnusnumero 58).

25 24. Patenttivaatimuksen 23 mukainen DNA-molekyyli,  
tunnettu siitä, että fuusioproteiinia koodittava nukleo-  
tidisekvenssi on toimivalla tavalla liitetty ekspression  
kontrollisekvensseihin.

30 25. Yksisoluinen isäntä, tunnettu siitä, että  
se on transformoitu patenttivaatimuksen 23 tai 24 mukaisel-  
la DNA-molekyyllillä.

26. Menetelmä patenttivaatimuksen 1 mukaisen fuusio-  
proteiinin tuottamiseksi, joka on:

a) JB106 (SEQ ID nro 38),  
b) JB117 (SEQ ID nro 59),  
35 c) JB118 (SEQ ID nro 60),  
d) JB119 (SEQ ID nro 61),

- e) JB120 (SEQ ID nro 62),
  - f) JB122 (SEQ ID nro 63) tai
  - g) HIV:n tat-proteiinin aminohapot 47 - 58, joita seuraa HSV:n VP16-proteiinin aminohapot 43 - 412,
- 5 tunnettu siitä, että
- a) viljellään patenttivaatimuksen 23 tai 24 mukaisella DNA-molekyylillä transformoitua yksisoluiista isäntää ja
  - b) kerätään fuusioproteiini mainitusta viljelmästä.



**Patentkrav**

1. Fusionsprotein, som består av en karboxiterminal lastdel och en aminoterminal transportdel, kännetecknat av att
- 5 a) transportdelen karakteriseras av:
- i) förekomsten av HIV tat-proteinets aminosyror 49 - 57,
- ii) avsaknaden av HIV tat-proteinets aminosyror 22 - 10 36 och
- iii) avsaknaden av HIV tat-proteinets aminosyror 73 - 86, och
- b) lastdelen bibehåller en väsentlig biologisk aktivitet efter att den har förts in i cellen på ett sätt, 15 som är beroende av transportdelen.
2. Fusionsprotein enligt patentkrav 1, kännetecknat av att lastdelen är en terapeutisk molekyl, profylaktisk molekyl eller diagnostisk molekyl.
3. Fusionsprotein enligt patentkrav 1 eller 2, 20 kännetecknat av att lastdelen består av en human papillomavirus E2-repressor som bibehåller sin biologiska aktivitet efter transporten till målcellen och transportdelen är:
- a) HIV tat-proteinets aminosyror 47 - 58 (SEQ ID nr 25 47),
- b) HIV tat-proteinets aminosyror 47 - 72 (SEQ ID nr 48),
- c) HIV tat-proteinets aminosyror 38 - 72 (SEQ ID nr 49) eller
- 30 d) HIV tat-proteinets aminosyror 38 - 58 (SEQ ID nr 50).
4. Fusionsprotein enligt patentkrav 3, kännetecknat av att transportdelen föregås av en aminoterminal metionin.

5. Fusionsprotein enligt vilket som helst av patentkraven 1 - 4, kännetecknat av att lastdelen består av human papillomavirus E2-proteinets aminosyror 245 - 365 (SEQ ID nr 51).

5           6. Fusionsprotein enligt patentkrav 1, kännetecknat av att det är fusionsprotein JB106 (SEQ ID nr 38).

10           7. Fusionsprotein enligt patentkrav 1, kännetecknat av att det är fusionsprotein JB117 (SEQ ID nr 59).

8. Fusionsprotein enligt patentkrav 1, kännetecknat av att det är fusionsprotein JB118 (SEQ ID nr 60).

15           9. Fusionsprotein enligt patentkrav 1, kännetecknat av att det är fusionsprotein JB122 (SEQ ID nr 63).

20           10. Fusionsprotein enligt patentkrav 1, kännetecknat av att lastdelen består av en nöt papillomavirus E2-repressor som bibehåller sin biologiska aktivitet efter transporten till en målcell och transportdelen är:

a) HIV tat-proteinets aminosyror 47 - 62 (SEQ ID nr 52) eller

b) HIV tat-proteinets aminosyror 38 - 62 (SEQ ID nr 53).

25           11. Fusionsprotein enligt patentkrav 10, kännetecknat av att transportdelen föregås av en aminoterminal metionin.

30           12. Fusionsprotein enligt vilket som helst av patentkraven 1, 2, 10 eller 11, kännetecknat av att lastdelen är en E2-repressor bestående av nöt papillomavirus E2-proteinets aminosyror 250 - 410 (SEQ ID nr 56).

13. Fusionsprotein enligt patentkrav 12, kännetecknat av att det är JB119 (SEQ ID nr 61).

35           14. Fusionsprotein enligt patentkrav 12, kännetecknat av att det är JB120 (SEQ ID nr 62).

15. Fusionsprotein enligt patentkrav 1 eller 2, kännetecknat av att lastdelen består av HSV VP16-proteinets aminosyror 43 - 412 och transportdelen består av HIV tat-proteinets aminosyror 47 - 58.

5 16. Fusionsprotein enligt patentkrav 15, kännetecknat av att transportdelen föregås av en aminoterminal metionin.

10 17. Användning av ett fusionsprotein enligt vilket som helst av kraven 1 - 14, för intracellulär tillförsel av last i icke-terapeutiskt ändamål.

18. Kovalent bundet kemiskt konjugat bestående av en transport polypeptiddel och en lastdel, kännetecknat av att

15 a) konjugatets transportpolypeptiddel karakteriseras av:

i) förekomsten av HIV tat-proteinets aminosyror 49 - 57,

ii) avsaknaden av HIV tat-proteinets aminosyror 22 - 36 och

20 iii) avsaknaden av HIV tat-proteinets aminosyror 73 - 86, och

b) lastdelen bibehåller en väsentlig biologisk aktivitet efter att den har förts in i cellen på ett sätt, som är beroende av transportdelen.

25 19. Kovalent bundet kemisk konjugat enligt patentkrav 18, kännetecknat av att transportpolypeptiddelen består av HIV tat-proteinets aminosyror 37 - 72 (SEQ ID nr 2).

30 20. Kovalent bundet kemiskt konjugat enligt patentkrav 19, kännetecknat av att lastdelen är:

a) human papillomavirus E2-proteinets aminosyror 245 - 365 (SEQ ID nr 51) eller

35 b) human papillomavirus E2-proteinets aminosyror 245 - 365, vari aminosyrorna 300 och 309 har ändrats till cystein (SEQ ID nr 55).

21. Kovalent bundet kemiskt konjugat enligt patentkrav 20, kännetecknat av att lastdelen är en dubbelsträngad DNA-molekyl, som är:

5 a) oligonukleotiden NF1 (SEQ ID nr 43) kopplad till oligonukleotiden NF2 (SEQ ID nr 44) eller

b) en oligonukleotiden NF3 (SEQ ID nr 45) kopplad till oligonukleotiden NF4 (SEQ ID nr 46).

10 22. Användning av ett kovalent bundet kemisk konjugat enligt vilket som helst av patentkraven är 18 - 21 för intracellulär tillförsel av last i icke-terapeutiskt ändamål.

23. DNA-molekyl, kännetecknad av att den består av en nukleotidsekvens som kodar ett fusionsprotein enligt patentkrav 1 som är:

15 a) JB106 (SEQ ID nr 38),

b) JB117 (SEQ ID nr 59),

c) JB118 (SEQ ID nr 60),

d) JB119 (SEQ ID nr 61),

e) JB120 (SEQ ID nr 62),

20 f) JB122 (SEQ ID nr 63) eller

g) tat-VP16R.GF (SEQ ID nr 58).

24. DNA-molekyl enligt patentkrav 23, kännetecknad av att nukleotidsekvensen, som kodar fusionsproteinet, är funktionellt bunden till expressionskontrollsekvenser.

25 25. Encellig värd, kännetecknad av att den är transformerad med en DNA-molekyl enligt patentkrav 23 eller 24.

30 26. Förfarande för framställning av ett fusionsprotein enligt patentkrav 1, som är:

a) JB106 (SEQ ID nr 38),

b) JB117 (SEQ ID nr 59),

c) JB118 (SEQ ID nr 60),

d) JB119 (SEQ ID nr 61),

35 e) JB120 (SEQ ID nr 62),

f) JB122 (SEQ ID nr 63) eller

g) aminosyror 47 - 58 av HIV tat-protein följt av aminosyror 43 - 412 av HSV VP16-protein, kännetecknat av att

5 a) man odlar en encellig värd transformerad med en DNA-molekyl enligt patentkrav 23 eller 24 och

b) tillvaratar fusionsproteinet från nämnda odling.

# FIG. 1

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

FIG. 2

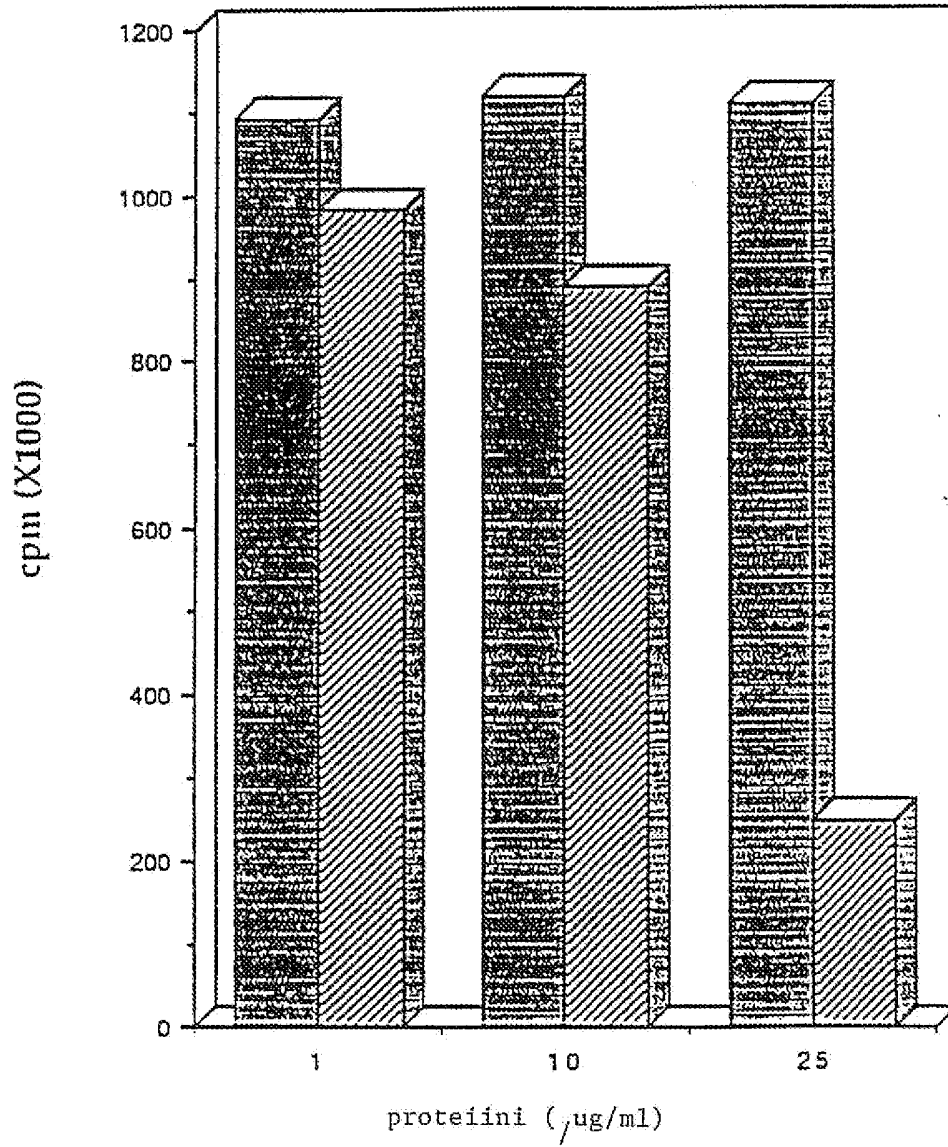
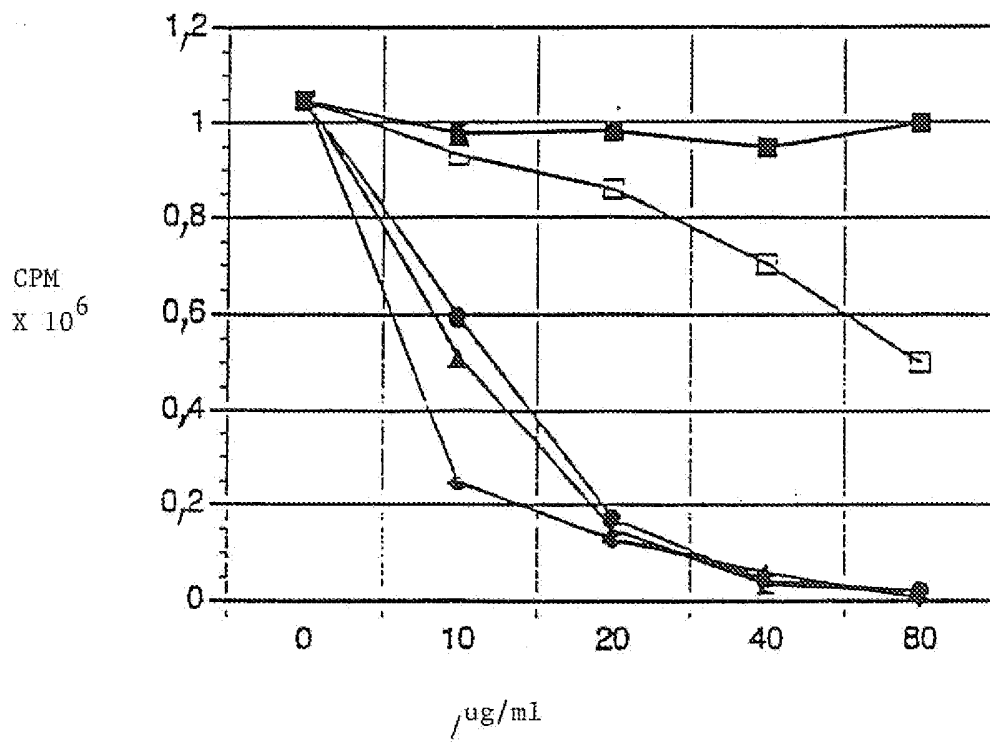


FIG. 3





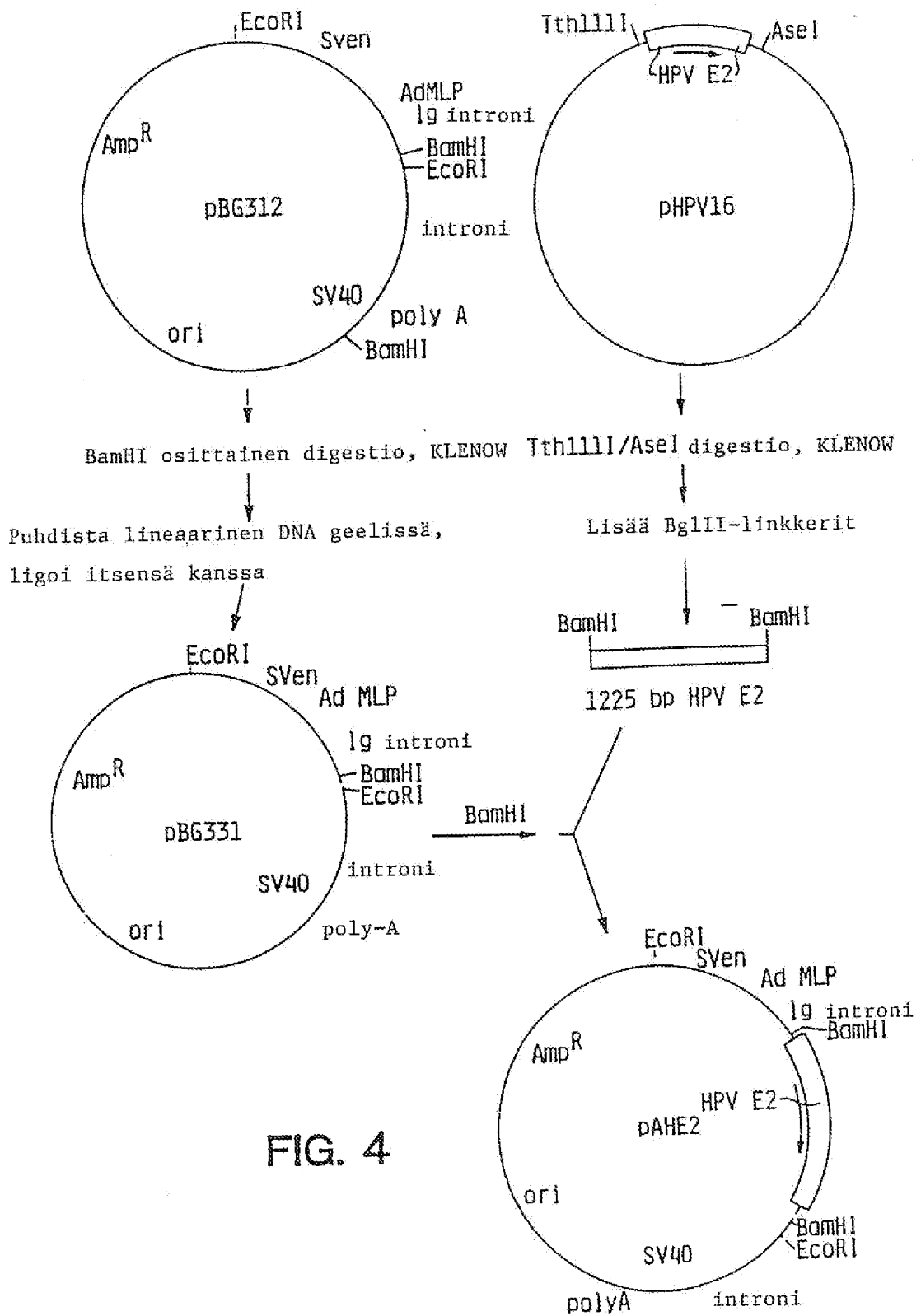


FIG. 4

FIG. 5

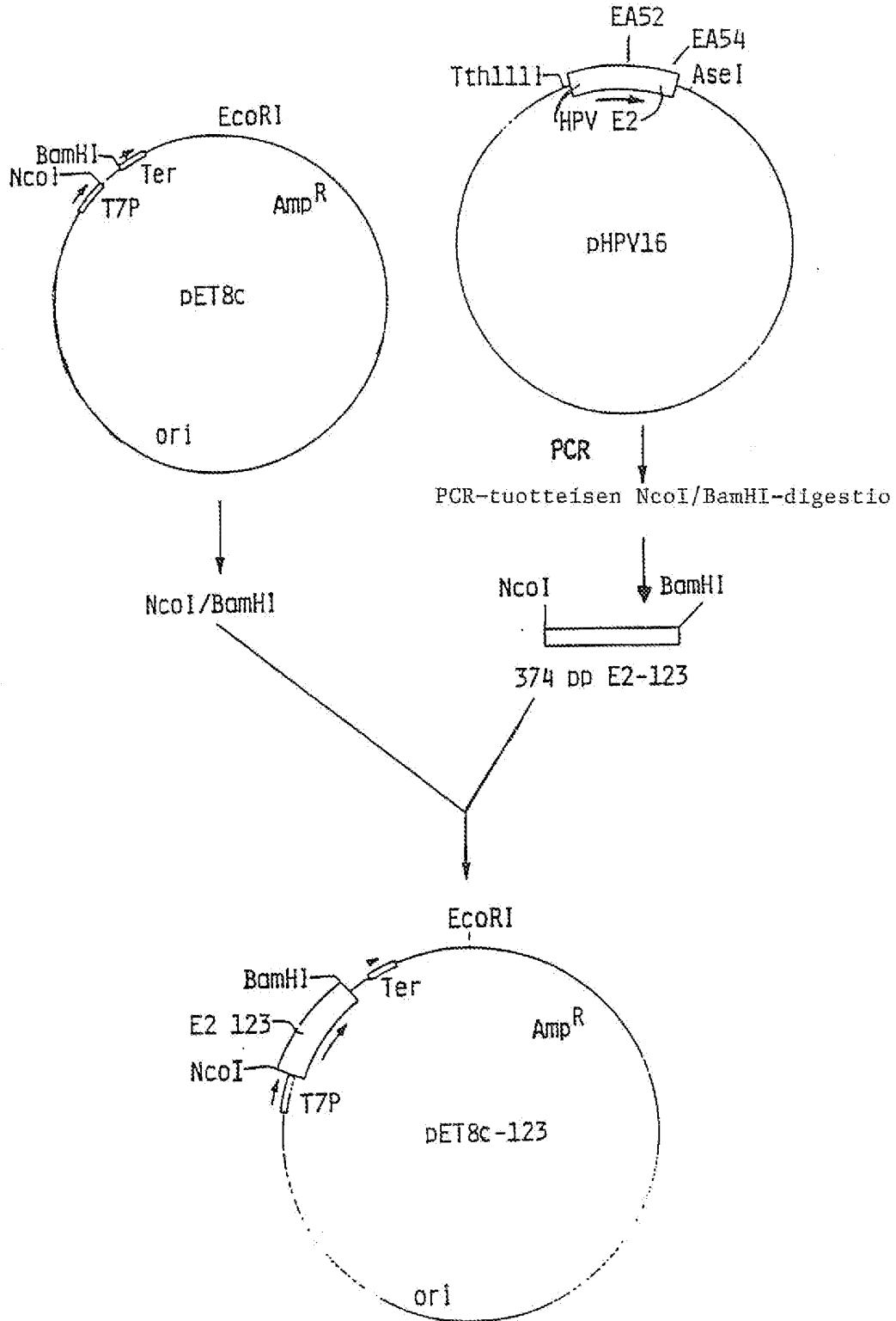


FIG. 6

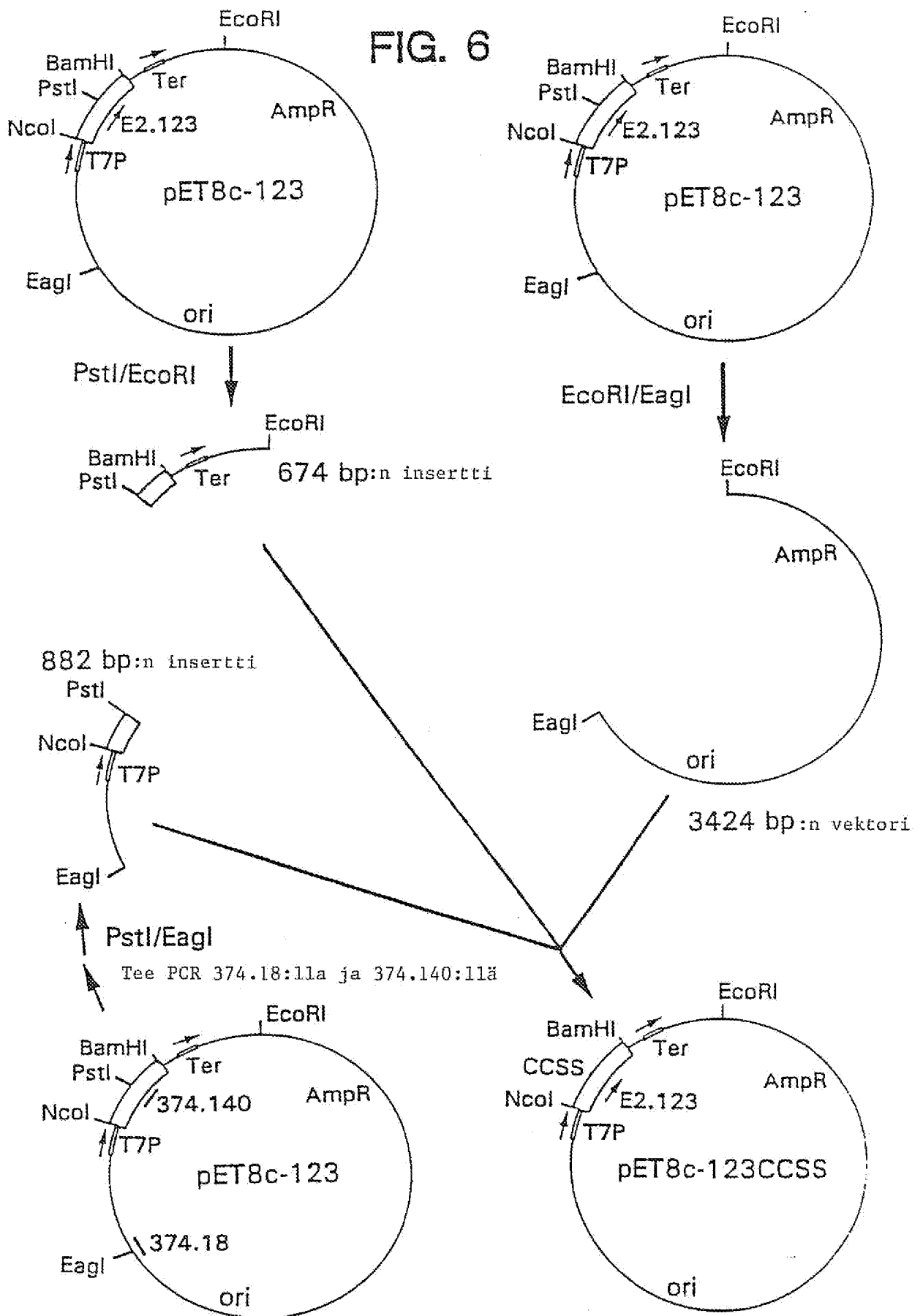
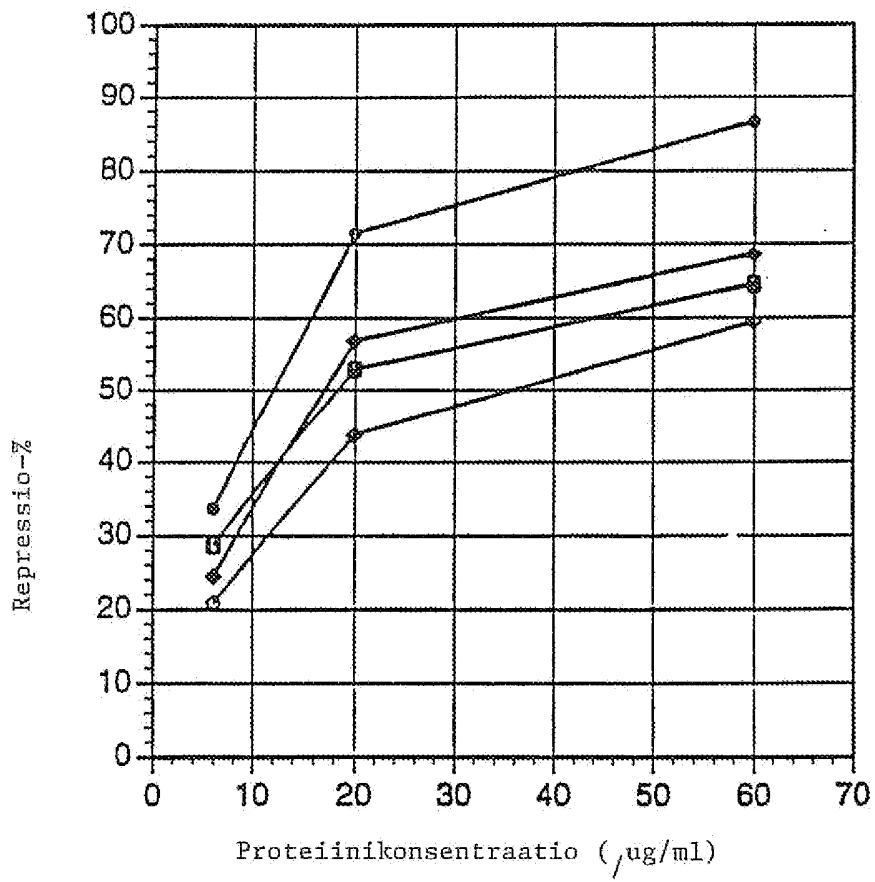
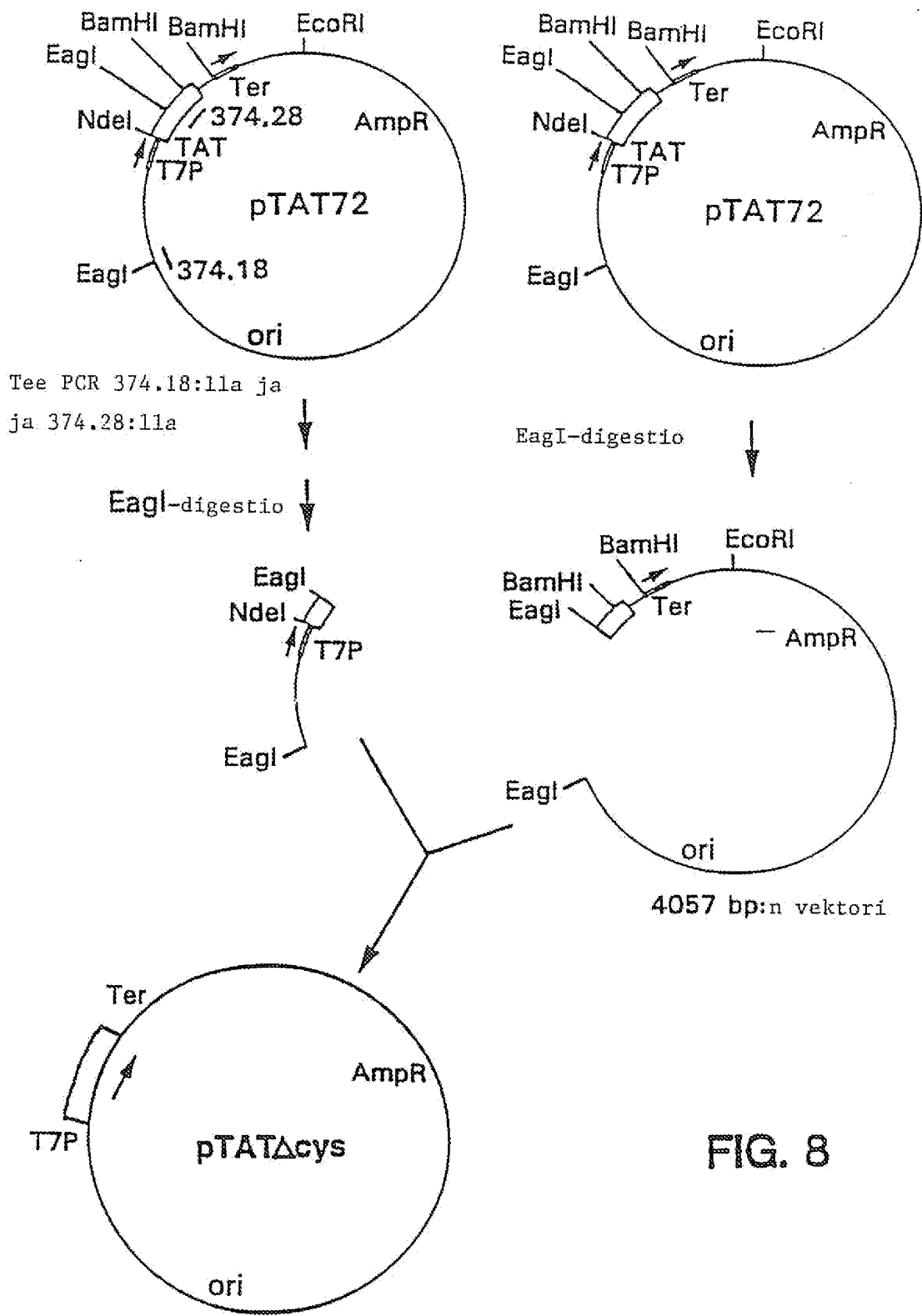


FIG. 7





**FIG. 8**

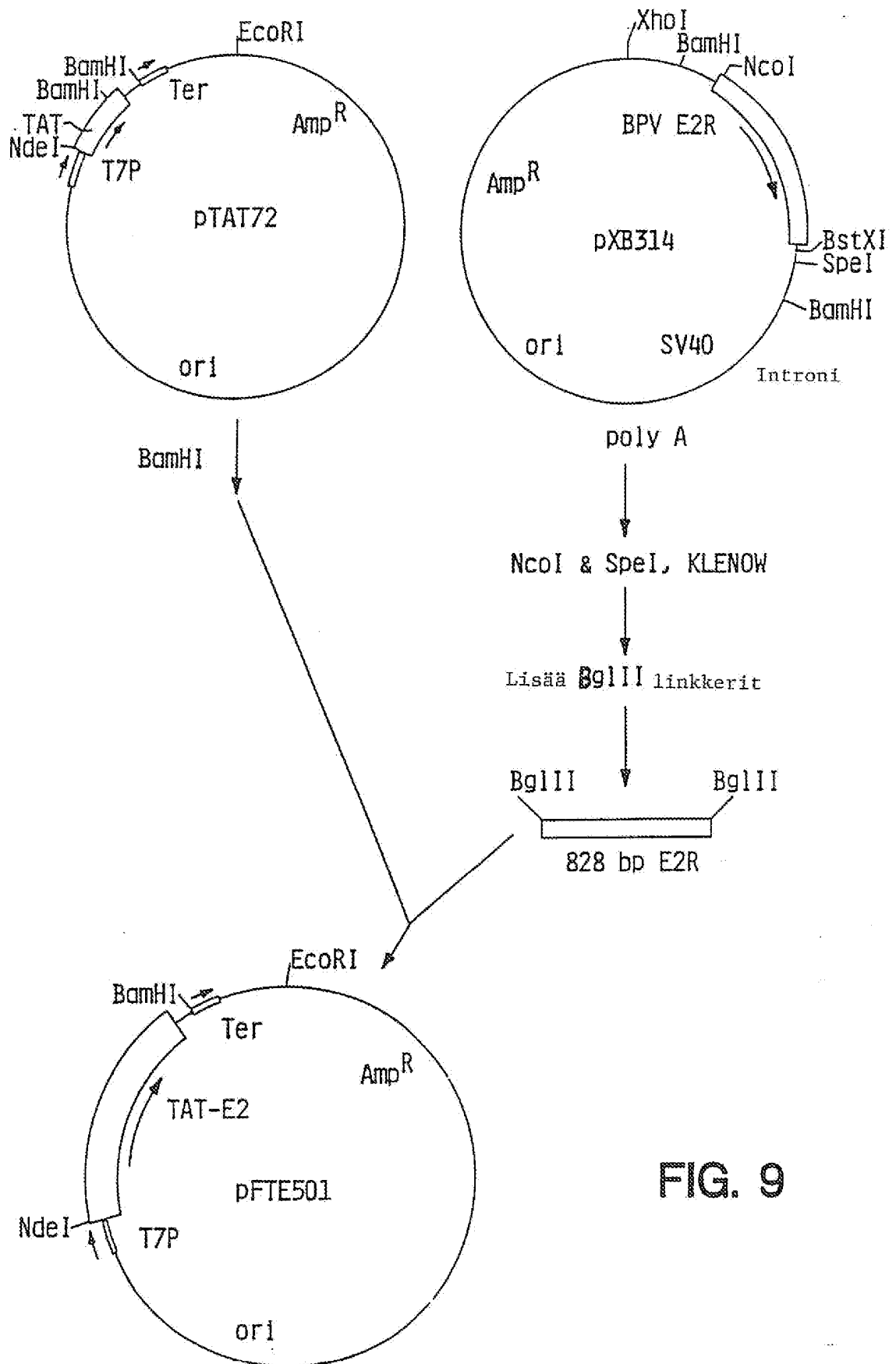
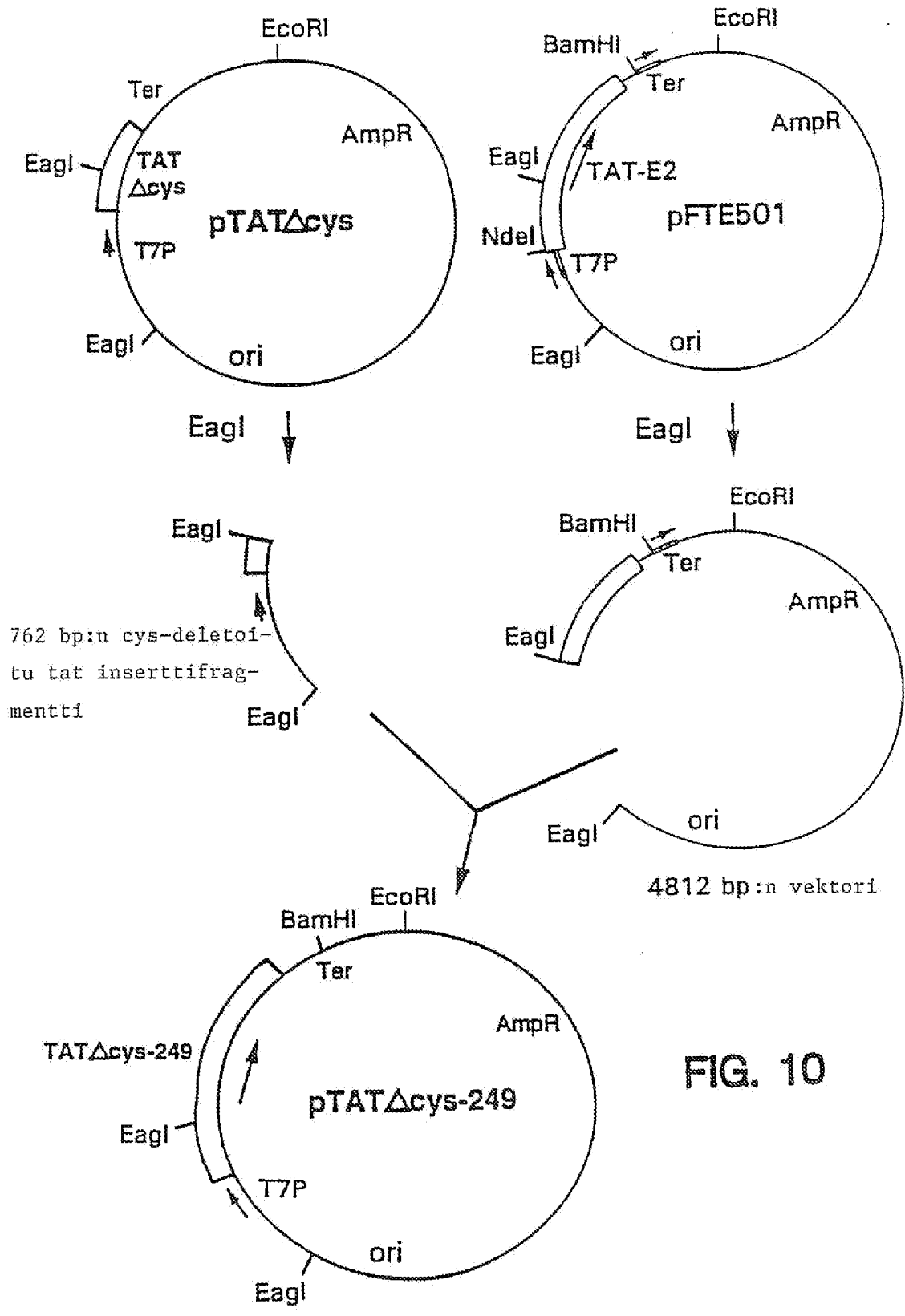
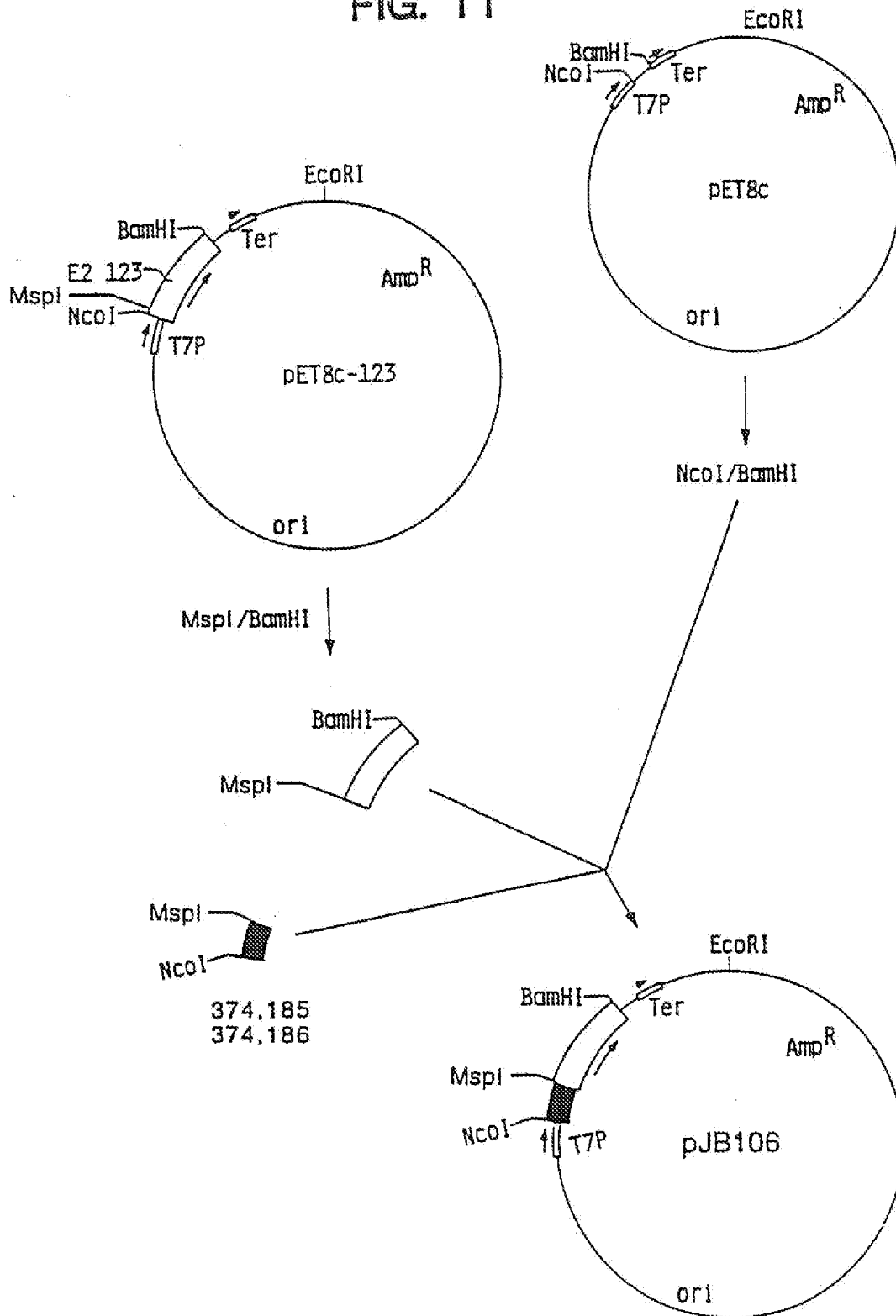


FIG. 9



**FIG. 10**

FIG. 11





## FIG. 12

MET TYR GLY ARG LYS LYS ARG ARG GLN ARG ARG  
47

ARG PRO PRO ASP THR GLY ASN PRO CYS HIS THR THR  
58 245

LYS LEU LEU HIS ARG ASP SER VAL ASP SER ALA PRO  
255

ILE LEU THR ALA PHE ASN SER SER HIS LYS GLY ARG  
267

ILE ASN CYS ASN SER ASN THR THR PRO ILE VAL HIS  
279

LEU LYS GLY ASP ALA ASN THR LEU LYS CYS LEU ARG  
291

TYR ARG PHE LYS LYS HIS CYS THR LEU TYR THR ALA  
303

VAL SER SER THR TRP HIS TRP THR GLY HIS ASN VAL  
315

LYS HIS LYS SER ALA ILE VAL THR LEU THR TYR ASP  
327

SER GLU TRP GLN ARG ASP GLN PHE LEU SER GLN VAL  
339

LYS ILE PRO LYS THR ILE THR VAL SER THR GLY PHE  
351

365  
MET SER ILE

FIG. 13

