(19) SUOMI - FINLAND
(FI)

PATENTTI- JA REKISTERIHALLITUS
PATENT- OCH REGISTERSTYRELSEN
FINNISH PATENT AND REGISTRATION OFFICE

(10) **FI 945221 A7**

(12) **JULKISEKSI TULLUT PATENTTIHAKEMUS**
PATENTANSÖKAN SOM BLIVIT OFFENTLIG
PATENT APPLICATION MADE AVAILABLE TO THE PUBLIC

(21) Patenttihakemus - Patentansökan - Patent application 945221

(51) Kansainvälinen patenttiluokitus - Internationell patentklassifikation - International patent classification
C12N 15/30
C12N 15/62
C12N 9/64
C07K 9/00
A61K 39/00
C07H 13/04

(22) Tekemispäivä - Ingivningsdag - Filing date **07.05.1993**

(23) Saapumispäivä - Ankomstdag - Reception date **07.11.1994**

(41) Tullut julkiseksi - Blivit offentlig - Available to the public **07.11.1994**

(43) Julkaisupäivä - Publiceringsdag - Publication date **13.06.2019**

(86) Kansainvälinen hakemus - Internationell ansökan - International application **07.05.1993 PCT/GB1993/000943**

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

08.05.1992 GB 9209993

(71) Hakija - Sökande - Applicant

1 • **The Biotechnology and Biological,** Babraham Hall Babraham, Cambridge CB2 4AT, United Kingdom, ISO-BRITANNIA, (GB)

(72) Keksijä - Uppfinnare - Inventor

1 • **Graham, Margaret,** United Kingdom, ISO-BRITANNIA, (GB)
2 • **Smith, Trevor Stanley,** United Kingdom, ISO-BRITANNIA, (GB)
3 • **Munn, Edward Albert,** United Kingdom, ISO-BRITANNIA, (GB)
4 • **Knox, David Patrick,** United Kingdom, ISO-BRITANNIA, (GB)
5 • **Oliver, Joanna Jane,** United Kingdom, ISO-BRITANNIA, (GB)
6 • **Newton, Susan Elizabeth,** Australia, AUSTRALIA, (AU)

(74) Asiamies - Ombud - Agent

Kolster Oy Ab, Salmisaarenaukio 1, 00180 Helsinki

(54) Keksinnön nimitys - Uppfinningens benämning - Title of the invention

Yhdistelmä-DNA-molekyylejä, jotka koodittavat aminopeptidaasientsyymej ä, ja niiden käyttö rokotteiden valmistamiseksi matotartuntoja vastaan

Rekombinant DNA-molekyler vilka kodar aminopeptidasenzymer och användn ing av dem vid framställning av vacciner mot maskinfektioner

Yhdistelmä-DNA-molekyylejä, jotka koodittavat aminopeptidaasientsyymejä, ja niiden käyttö rokotteiden valmistamiseksi matotartuntoja vastaan

Esillä oleva keksintö koskee suojaavien antigeenien valmistusta yhdistelmä-DNA-teknologialla käytettäviksi matovastaisina aineina ja suojaavina immunogeeneina matoloisten aiheuttamien sairauksien kontrolloinnissa.

Matoloiset ovat vastuussa laajasta valikoimasta kotieläinten sairauksia ja loistartuntoja, joilla johtaessaan niille ominaisesti tuotantohävikkiin ja jopa eläinten kuolleisuuteen on merkittävää taloudellistamerkitystä. Täten esimerkiksi verta nauttiva sukkulamato Haemonchus tartuttaa märehtijöiden ruoansulatuskanavan sisäseinän aiheuttaen anemiaa ja painon menetystä ja hoitamattomana usein johtaen kuolemaan. Eläimet, joilla on sukua oleva, ei-verta nauttiva sukkulamato Ostertagiatartunta, samoin eivät voi hyvin ja voivat hoitamattomina kuolla. Muita taloudellisesti merkittäviä matosukuja ovat Trichostrongylus ja Nematodirus, jotka aiheuttavat suolen tulehdusta erilaisissa eläimissä, ja imumadot.

Ongelmia aiheuttavat myös sukkulamadot, kuten koukkumadot (esim. Necator, Ancylostoma, Uncinaria ja Bunostomum spp) ja lampaan maksamadot (esim. Fasciola, Paramphistomum ja Dicrocoelium) ja niiden sukulaiset, jotka märehtijöiden ja lemikkikotiläinten lisäksi tartuttavat myös ihmisiä usein kohtalokkain seurauksin.

Matoloisten kontrollointi nojaa tällä hetkellä pääasiassa matovastaisten lääkkeiden käyttöön yhdistettynä laitumen hoitoon. Tällaisilla tekniikoilla on kuitenkin lukuisia huonoja puolia, lääkkeiden tiheä antaminen ja laitumen hoito eivät useinkaan ole käytännöllisiä ja lääkkeitä vastustavat matokannat ovat leviämässä yhä laajemmalle.

Täten tällä alalla on tarve saada tehokas matovastainen rokote, ja monia ponnistuksia on keskitetty tälle alalle viime vuosina. Kuitenkaan toistaiseksi ei ole kaupallisesti saatavia molekyyli- tai alayksikkörokotteita pääasiallisille matolajeille, erityisesti märehtijöiden ruoansulatuskanavan sukkulamadoille kuten Haemonchus ja Ostertagia.

Lupaavimmat tulokset tähän mennessä on saatu uusilla, Haemonchuksesta eristetyillä proteiineilla, joilla on mahdollisuuksia suojaavina antigeeneinä ei ainoastaan Haemonchusta vastaan vaan myös vastaan joukkoa muita matoja. Erityisesti proteiinidubletin H11OD, jota esiintyy H. contortuksen suolen luminaalipinnalla, on osoitettu tuottavan suojaavaa immuniteettia haemonchosis-sairautta vastaan lampaissa.

H11OD:n H. contortuksesta likimääräinen molekyylipaino on 110 kilodaltonia (kd) pelkistävissä ja ei-pelkistävissä olosuhteissa, määriteltynä SDS-PAGE:n mukaan, ja sitä kuvataan julkaisuissa WO 88/00 835 ja WO 90/11 086. Tässä käytettynä ilmaisu "H11OD" viittaa proteiinidublettiin H11OD sellaisena kuin tämä määritellään julkaisuissa WO 88/00 835 ja WO 90/11 086. Vastaavia proteiineja on myös vast'ikään osoitettu muissa matolajeissa, esim. Necator americanuksessa.

Julkaisussa WO 88/00 835 on kuvattu H11OD:n puhdistamiseksi lukuisia menetelmiä, jotka riittävät proteiinin karakterisoimiseksi, ja niiden mittakaavaa voidaan suurentaa aina siihen, että proteiinin tuottaminen kokeellisesti ja kaupallisesti käyttökelpoisina määrinä on mahdollista. Tarvitaan kuitenkin aiempaa parempi ja kätevä lähde, josta valmistaa ei ainoastaan H11OD:tä vaan myös antigeenisiä sukulaisproteiineja, etenkin menetelmään, joka perustuu yhdistelmä-DNA-teknologiaan ja proteiinien ilmentämiseen sopivasti transformoiduissa prokaryoottisissa tai eukaryoottisissa organismeissa.

Esillä oleva keksintö pyrkii antamaan käyttöön tällaisen aiempaa paremman menettelyn. Sekvenssimääritys cDNA-sekvensseille H110D:lle Haemonchus contortuksesta on suoritettu ja ennustettujen aminohapposekvenssien on todettu osoittavan homologiaa ryhmään integraalisia membraaniaminopeptidaaseja [systemaattinen nimi: α-aminoasyylipeptidihydrolaasi (mikrosomaalinen)] nähden.

Nisäkkään integraaliset membraaniaminopeptidaasit sijaitsevat useissa kudoksissa, esim. suolen seinämän ja munuaisen solukalvon ohuissa ulokkeissa. Niiden rooli munuaisessa on epäselvä, mutta suolessa niiden funktio on pilkkoa pienet peptidit, jotka ovat ruoansulatuksen lopputuotteita (katsauksien osalta katso Kenny & Maroux, 1982; Kenny & Turner, 1987; Noren et al., 1986; Semenza, 1986).

Yhtenä näkökohtana esillä oleva keksintö antaa täten käyttöön nukleiinihappomolekyylejä, jotka käsittävät yhtä tai useampaa nukleotidisekvenssiä, joka koodittaa matoaminopeptidaasientsyymejä tai niiden antigeenisiä osia, jotka oleellisesti vastaavat kokonaisuudessaan tai osittain kuvioissa 2, 3, 4 tai 5 esitettyjä nukleotidisekvenssejä (sekvenssit tunnusnrot 1 - 15), tai sekvenssejä, jotka koodittavat matoaminopeptidaasientsyymejä ja jotka ovat oleellisesti homologisia mihin tahansa mainittuihin sekvensseihin nähden tai hybridisoituvat niihin.

Keksinnön mukainen nukleiinihappo voi täten olla yksi- tai kaksisäikeinen DNA, cDNA tai RNA.

Variaatioita aminopeptidaasia koodittavissa nukleotidisekvensseissä voi esiintyä eri matokantojen välillä lajin puitteissa, madon elinsyklin eri vaiheiden välillä (esim. toukkavaiheen ja täysikasvuisen vaiheen välillä), samanlaisten kantojen välillä, joilla on eri maantieteellinen alkuperä, ja myös saman madon puitteissa. Tällaiset variaatiot sisältyvät tämän keksinnön suojapiiriin.

"Oleellisesti homologinen" tässä käytettynä sisältää ne sekvenssit, joiden sekvenssi-identtisyys on noin 50 % tai enemmän, esim. 60 % tai enemmän, ja myös funktionaalisesti ekvivalentit alleeliset variantit ja sukulaissekvenssit, joita on modifioitu yhdellä tai useammalla emässubstituutiolla, -lisäyksellä ja/tai -deleetiolla. "Funktionaalisesti ekvivalentilla" tarkoitetaan nukleiinihapposekvenssejä, jotka koodittavat polypeptidejä, joilla on aminopeptidaasiaktiivisuuksia, jotka ovat samalla tavalla immunoreaktiivisia, eli jotka synnyttävät isäntää suojaavia vasta-aineita matoja vastaan.

Nukleiinihappomolekyylit, jotka hybridisoituvat kuvioissa 2, 3, 4 tai 5 esitettyihin sekvensseihin (jotka käsittävät sekvenssit tunnusnrot 1 - 15), tai mitkä tahansa oleellisesti homologiset tai funktionaalisesti ekvivalentit sekvenssit edellä määritellyn mukaisesti sisältyvät samoin keksinnön suojapiiriin. "Hybridisaatio" tässä käytettynä määrittelee ne sekvenssit, jotka sitoutuvat eiankarissa olosuhteissa (6 x SSC/50 % formamidi huoneenlämpötilassa) ja pestyinä alhaisen ankaruuden olosuhteissa (2 x SSC, huoneenlämpötila, edullisemmin 2 x SCC, 42 °C), tai ankarammissa olosuhteissa, esim. 2 x SSC, 65 °C (jossa SSC = 0,15 M NaCl, 0,015 M natriumsitraatti, pH 7,2).

Menetelmät tällaisten sukua olevien johdannaissekvenssien tuottamiseksi, esimerkiksi paikkakohdennetusti mutagenisoimalla, umpimähkään mutagenisoimalla, tai entsymaattisesti pilkkomalla ja/tai ligatoimalla nukleiinihappoja, ovat alalla hyvin tunnettuja, kuten myös ovat menetelmät sen määrittämiseksi, onko täten modifioidulla nukleiinihapolla merkittävää homologiaa kyseiseen sekvenssiin nähden, esimerkiksi hybridisoimalla.

Keksinnön mukaisen nukleiinihappomolekyylin käyttöön antaminen mahdollistaa täten yhdistelmä-DNA-aminopeptidaasientsyymien tai niiden immunogeenisten fragmenttien saannin määrinä, joita ei tähän mennessä ole voitu

saada, mikä mahdollistaa matovastaisten rokotteiden kehittämisen.

Toisena näkökohtana esillä oleva keksintö antaa täten käyttöön nukleiinihappomolekyylejä, jotka käsittävät yhtä tai useampaa nukleotidisekvenssiä, joka koodittaa yhtä tai useampaa polypeptidiä, joka kykenee synnyttämään suojaavia vasta-aineita matoloisia vastaan, joihin sekvensseihin sisältyy yksi tai useampi antigeenimäärettä koodittava alue kuvioissa 2, 3, 4 tai 5 esitetyistä aminopeptidaasia koodittavista sekvensseistä (sekvenssit tunnusnrot 1 - 15).

Esillä oleva keksintö ulottuu samoin synteettisiin polypeptideihin, jotka käsittävät yhtä tai useampaa aminohapposekvenssiä, joka muodostaa aminopeptidaasientsyymin tai sen antigeenisiä osia, oleellisesti vastaten kokonaisuudessaan tai osittain kuvioissa 2, 3, 4 tai 5 esitettyjä nukleotidisekvenssejä (sekvenssit tunnusnrot 1 - 15), tai sen funktionaalisesti ekvivalentin variantin, joka on muu kuin synteettinen polypeptidi, joka vastaa proteiinidublettia H11OD, tai synteettinen polypeptidi, joka vastaa mitä tahansa julkaisussa WO 90/11 086 julkistettuja yksittäisiä polypeptidisekvenssejä.

Vaihtoehtoisesti tarkasteltuna keksintö antaa samoin käyttöön synteettisiä polypeptidejä, jotka käsittävät aminohapposekvenssin, joka muodostaa aminopeptidaasientsyymin tai sen antigeenisen osan, vastaten oleellisesti kokonaisuudessaan tai osittain kuvioissa 2, 3, 4 tai 5 esitettyjä nukleotidisekvenssejä (sekvenssit tunnusnrot 1 - 15), tai sen funktionaalisesti ekvivalentin variantin, joka on oleellisesti vapaa muista Haemonchus contortus -aineosista.

Keksintö ulottuu edelleen rokotekoostumuksiin immuunivasteiden matoloisia vastaan stimuloimiseksi ihmis- tai ei-ihmiseläimessä, jotka rokotekoostumukset käsittävät ainakin yhtä edellä määritellyn mukaista synteettistä

polypeptidiä yhdessä farmaseuttisesti hyväksyttävän kantajan kanssa.

Julkaisussa WO 90/11 086 julkistetaan lukuisia polypeptidi- tai osittaisia polypeptidisekvenssejä, jotka saadaan proteolyyttisesti pilkkomalla tai kemiallisesti pilkkomalla proteiinidublettia H110D seuraavasti:

(a) Met Gly Tyr Pro Val Val Lys Val Glu Glu Phe

(b) Met Gly Phe Pro Val Leu Thr Val Glu Ser

(c) Met Gly/Phe Asn Phe Lys Ile Glu/Val Thr/Glu Ala Gly

(d) Met Lys Pro/Glu Thr/Val Lys Asp/Ala Thr/Lys Leu - Ile Thr

(e) Met Leu Ala Leu Asp Tyr His Ser - Phe Val

(f) Met Leu Ala Glu/Tyr Asp Gln/Ala Glu Asp Val

(g) Met Gly Phe Pro Leu Val Thr Val Glu Ala Phe Tyr

(h) Met Lys Thr Pro Glu Phe Ala Val/Leu Gln Ala Phe/Thr Ala Thr Ser/Gly Phe Pro

(i) Lys His/Tyr Asn/Val Ser Pro Ala Ala Glu Asn/Leu Leu Asn/Gly

(j) Lys - Thr Ser Val Ala Glu Ala Phe Asn

(k) Lys Ala Ala Glu Val Ala Glu Ala Phe Asp - Ile - - - Lys Gly

(l) Lys Ala Val Glu Val/Pro Ala Glu Ala Phe Asp Asp Ile Thr? Tyr - - Gly Pro Ser

(m) Lys - Glu Glu Thr Glu Ile Phe Asn Met

(n) Lys - - - Pro Phe Asn/Asp Ile Glu Ala Leu

(o) Asp Gln Ala Phe Ser Thr Asp Ala Lys

(p) Met Gly Tyr Pro Val Val Lys Val Glu Glu Phe -Ala Thr Ala Leu

(q) Met Gly Phe Pro Val Leu Thr Val Glu Ser - Tyr? - Thr

(r) Met Glu/Phe Asn Phe Leu Ile Glu/Val Thr/Glu Ala Gly - Ile Thr

(s) Met Gly Phe Leu Val Thr Val Glu Ala Phe Tyr - Thr Ser

(t) Met Lys Thr Pro Glu Phe Ala Val/Leu Gln Ala Phe/Thr Ala Thr Ser/Gly Phe Pro

(u) Met Lys Pro/Glu Thr/Val Leu Asp/Ala Thr/Lys Leu - Ile Thr - Gly

(v) Met Leu Ala Leu Asp Tyr His Ser - Phe Val Gly?

(w) Met Leu Ala Glu/Tyr Asp Gln/Ala Glu Asp Val

(x) Lys His/Tyr Asn/Val Ser Pro Ala Ala Glu Asn/Leu Leu Asn/Gly

(y) Lys - Thr Ser Val Ala Glu Ala Phe Asn

(z) Lys Ala Ala Glu Val Ala Glu Ala Phe Asp - Ile - - - Lys Gly

(aa) Lys Ala Val Glu Val/Pro Ala Glu Ala Phe Asp Asp Ile Thr? Tyr - - Gly Pro Ser

(bb) Lys - Glu Gln Thr Glu Ile Phe Asn Met

(cc) Lys - - - Pro Phe Asn/Asp Ile Glu Ala Leu

(dd) Asp Gln Ala Phe Ser Thr Asp Ala Lys

Epävarmuudet esitetään joko muodolla Phe/Gly, jossa ensimmäinen kolmikirjainkoodi edustaa todennäköisintä oikeaa aminohappoa signaalin voimakkuuden perusteella, tai kysymysmerkillä; merkki "-" tarkoitata tuntematonta tähdettä.

Spesifisiä yksittäisiä polypeptidisekvenssejä, jotka julkistetaan julkaisussa WO 90/11 086, ei vaadita patentoitaviksi.

Ilmaisu "polypeptidi" sisältää tässä käytettynä sekä täysimittaisen proteiinin että lyhyemmät peptidisekvenssit.

Edellä suhteessa polypeptidiaminohapposekvensseihin käytettynä "funktionaalisesti ekvivalentti" määrittelee polypeptidit, jotka ovat sukua edellä mainituille polypeptidisekvensseille tai saadaan niistä, ja joissa aminohapposekvenssiä on modifioitu yhdellä tai useammalla aminohapposubstituutiolla, -lisäyksellä tai -deleetiolla, ja myös sekvenssit, joissa aminohappoja on kemiallisesti modifioitu, mukaan lukien glykosylaatio tai deglykosylaatio, mutta joilla kuitenkin on tallella suojaava antigeeninen (immunogeeninen) aktiivisuus. Tällaisia funktionaalisesti ekvivalentteja variantteja voi esiintyä luonnollisina biologisina variaatioina tai niitä voidaan valmistaa käyttäen tunnettuja tekniikoita, esimerkiksi funktionaalisesti ekvivalentteja yhdistelmä-DNA-polypeptidejä voidaan valmistaa käyttäen aminohappojan paikkakohdennetun mutagenisoinnin, umpimähkäisen mutagenisoinnin tai entsymaattisen pilkkomisen ja/tai ligatoinnin tunnettuja tekniikoita.

Yleensä keksinnön mukaiset synteettiset polypeptidit edustavat suojaavia antigeenisiä sekvenssejä. Tässä käytettynä ilmaisu "suojaava antigeeni" määrittelee ne antigeenit, jotka kykenevät synnyttämään isäntää suojaavan (immunogeenisen) immuunivasteen, eli isännässä vasteen, joka johtaa immuuniefektorimolekyylien, vasta-aineiden tai solujen syntymiseen, jotka steriloivat loisen hedelmällisyyden, vaurioittavat, inhiboivat sitä tai tappavat sen ja siten "suojaavat" isäntää kliiniseltä tai subkliiniseltä sairaudelta ja tuottavuuden laskulta. Tällainen suojaava immuunivaste voi tavanomaisesti ilmetä vasta-aineiden syntymisenä, jotka kykenevät inhiboimaan loisen metabolista funktiota, mikä johtaa surkastumiseen, munatuotannon puuttumiseen ja/tai kuolemaan.

Keksinnön tämän näkökohdan mukaiset synteettiset polypeptidit voidaan valmistaa ilmentämällä isäntäsolussa, joka sisältää yhdistelmä-DNA-molekyyliä, joka käsittää nukleotidisekvenssin edellä laajalti määriteltynä,

operatiivisesti kytkettynä ilmennystä kontrolloivaan sekvenssiin, tai yhdistelmä-DNA-kloonausvälinettä tai -vektoria, joka sisältää tällaista yhdistelmä-DNA-molekyyliä. Vaihtoehtoisesti polypeptidit voidaan ilmentää suoraan ruiskuttamalla isäntäsoluun paljasta keksinnön mukaista DNA-molekyyliä.

Näin ilmennetty synteettinen polypeptidi voi olla fuusiopolypeptidi, joka käsittää osan, joka osoittaa aminopeptidaasientsyymin kokonaisuudessaan tai osittain immunogeenisyyttä, ja lisäpolypeptidin, jota koodittaa ensin mainittuun fuusioidun yhdistelmä-DNA-molekyylin DNA. Esimerkiksi voi olla toivottavaa tuottaa fuusioproteiini, joka käsittää synteettisen aminopeptidaasin tai muun keksinnön mukaisen polypeptidin kytkettynä proteiiniin kuten ß-galaktosidaasi, fosfataasi, glutationi-S-transferaasi, ureaasi, hepatiitti-B-ydinantigeeni (Francis et al., 1989) ja muu vastaava. Useimmat fuusioproteiinit muodostetaan ilmentämällä yhdistelmä-DNA-geeni, jossa kaksi koodaussekvenssiä on liitetty yhteen lukualueet tahdittaen. Vaihtoehtoisesti polypeptidit voidaan kytkeä in vitro kemiallisin keinoin. Kaikki tällaiset aminopeptidaasia koodittavien nukleiinihappomolekyylien fuusio- tai hybridijohdannaiset ja niitä vastaavat aminohapposekvenssit kuuluvat esillä olevaan keksintöön. Tällaisia sopivia yhdistelmä-DNA- ja polypeptidi-ilmentämistekniikoita kuvataan esimerkiksi julkaisussa Sambrook et al., 1989. Vaihtoehtoisesti synteettiset polypeptidit voidaan tuottaa kemiallisin keinoin kuten hyvin tunnetulla Merrifieldin kiinteäfaasisynteesimenettelyllä.

Keksinnön muihin näkökohtiin sisältyy edellä määritellyn nukleiinihappomolekyylin tai synteettisen peptidin tai polypeptidin käyttäminen rokotekoostumuksen valmistamiseksi immuunivasteiden stimuloimiseksi ihmis- tai ei-ihmis-, edullisesti nisäkäs-, eläimessä matoloistartuntoja vastaan.

Vaihtoehtoisesti tarkasteltuna keksintö antaa myös käyttöön menetelmän immuunivasteen stimuloimiseksi ihmis- tai ei-ihmis-, edullisesti nisäkäs-, eläimessä matoloistartuntaa vastaan, jonka menetelmän mukaan mainitulle eläimelle annetaan rokotekoostumusta, joka käsittää yhtä tai useampaa edellä määritellyn nukleotidisekvenssin koodittamaa polypeptidiä.

Rokotekoostumus voidaan valmistaa keksinnön mukaan rokotevalmistuksen alalla hyvin tunnetuilla menetelmillä. Tavanomaiset rokotekoostumukset voivat käsittää yhtä tai useampaa keksinnön mukaista synteettistä polypeptidiä yhdessä, milloin tarkoituksenmukaista, yhden tai useamman sopivan apuaineen kanssa, [joita ovat] esim. aluminiumhydroksidi, saponiini, QuilA, tai sen puhdistetummat muodot, muramyylidipeptidi, mineraaliöljyt, tai Novasomes-valmisteet, yhden tai useamman farmaseuttisesti hyväksyttävän kantajan tai laimentimen läsnä ollessa. Sopivia kantajia ovat nestemäiset väliaineet, kuten suolaliuos, jotka sopivat käytettäviksi kantajina peptidien tai polypeptidien viemiseksi potilaaseen. Lisäaineosia kuten säilytysaineita voidaan sisällyttää mukaan.

Vaihtoehtoinen rokotekoostumus voi käsittää virusta tai isäntäsolua, esim. mikro-organismia (esim. lehmänrokkovirus, adenovirus, Salmonella), johon on insertoitu tämän keksinnön mukaista nukleiinihappomolekyyliä (esim. DNA-molekyyliä) immuunivasteen stimuloimiseksi, joka kohdistuu vastaan insertoidun nukleiinihappomolekyylin koodittamia polypeptidejä.

Rokotekoostumuksen antaminen voi tapahtua mitä tahansa tavanomaista tietä, esim. oraalisesti tai ruoansulatuskanavan ulkopuolisesti kuten lihaksen sisäisellä ruiskeella, mahdollisesti aikavälein, esim. kaksi ruisketta 7 - 28 päivän välein.

Kuten edellä mainittiin, kuvioissa 2, 3, 4 tai 5 esitettyjen nukleotidisekvenssien aminohappotranslaatio

osoittaa sekvenssihomologiaa ryhmään integraalisia membraaniaminopeptidaasientsyymejä nähden. Tämä määritettiin suorittamalla haku eri tietokannoissa, joita on saatavana Genetics Computer Group Sequence Analysis -ohjelmapakkauksessa, versio 7.01, marraskuu 1991 (Devereux et al., 1984) käyttäen kuvioissa 2, 3, 4 tai 5 esitettyjen sekvenssien translaatioita. Kaksi tällaista vertailua esitetään kuviossa 6.

Keksinnön mukaisten aminopeptidaasia koodittavien sekvenssien ilmentämiseen voidaan, kuten edellä mainittiin päästä käyttäen valikoimaa tunnettuja tekniikoita ja ilmentämissysteemejä, mukaan lukien ilmentäminen prokaryoottisissa soluissa, kuten E. coli ja eukaryoottisissa soluissa, kuten hiivat tai bakulovirus-hyönteissolusysteemissä tai transformoiduissa nisäkässoluissa ja siirtogeenisissä eläimissä ja kasveissa. Erityisen edullisesti nukleotidisekvenssit voidaan ilmentää käyttäen siirtogeeninen sukkulamato -systeemiä, kuten systeemiä sukkulamadolle Caenorhabditis, jota kuvataan esimerkiksi julkaisuissa Fire, 1986; Fire et al., 1989; Spieth et al., 1988; Han et al., 1990.

Keksinnön seuraava näkökohta antaa käyttöön menetelmän edellä määritellyn synteettisen polypeptidin valmistamiseksi, jonka menetelmän mukaan viljellään eukaryoottista tai prokaryoottista solua, joka sisältää edellä määriteltyä nukleiinihappomolekyyliä, olosuhteissa, joissa mainittu polypeptidi ilmentyy, ja otetaan täten tuotettu mainittu polypeptidi talteen.

Keksinnön lisänäkökohdat sisältävät siten kloonaus- ja ilmentämisvektoreita, jotka sisältävät keksinnön mukaisia nukleotidisekvenssejä. Tällaiset ilmentämisvektorit sisältävät sopivia kontrollisekvenssejä, kuten esimerkiksi translaation (esim. aloitus- ja lopetuskoodit) ja transkription kontrollielementit (esim. promoottori- operaattorialueet, ribosomisitoutumiskohdat, päätöksen stop-sek-

venssit) kytkettyinä yhteensopivalla lukualueella keksinnön mukaisiin nukleiinihappomolekyyleihin.

Keksinnön mukaiset vektorit voivat sisältää plasmideja ja viruksia (mukaan lukien sekä bakteriofagi- että eukaryoottiset virukset) alalla hyvin tunnettujen ja dokumentoitujen tekniikoiden mukaan, ja ne voidaan ilmentää lukuisissa erilaisissa ilmentämissysteemeissä, jotka myös ovat alalla hyvin tunnettuja ja dokumentoituja. Sopivia virusvektoreita ovat, kuten edellä mainittiin bakulovirus ja myös adenovirus ja lehmänrokkovirukset. Monia muita virusvektoreita kuvataan alalla.

Lukuisia tekniikoita tunnetaan ja voidaan käyttää tällaisten vektorien viemiseksi prokaryoottisiin tai eukaryoottisiin soluihin ilmentämistä varten, tai mikrobilinjaan tai somaattisiin soluihin siirtogeenisten eläinten muodostamiseksi. Sopivia transformointi- tai transfektiotekniikoita kuvataan perusteellisesti kirjallisuudessa.

Transformoidut tai transfektoidut eukaryoottiset tai prokaryoottiset isäntäsolut tai transgeeniset organismit, jotka sisältävät keksinnön mukaista, edellä määriteltyä nukleiinihappomolekyyliä, muodostavat keksinnön lisänäkökohdan.

Eukaryoottisilla ilmentämissysteemeillä yleensä ja sukkulamatoilmentämissysteemillä erityisesti on se etu, että translaation jälkeinen prosessointi ja erityisesti glykosylaatio voi tapahtua, siirtogeenisen sukkulamatosysteemin kyseessä ollen natiivissa proteiinissa tavattua vastaavaa glykosylaatiota voidaan odottaa. Tämä edustaa keksinnön tärkeätä näkökohtaa, koska monissa tapauksissa translaation jälkeistä prosessointia tarvitaan, jotta yhdistelmä-DNA-proteiini ilmentyisi optimaalisesti biologisesti aktiivisena.

Nisäkässoluilmentämissysteemeillä on myös lukuisia etuja. Nisäkäsisäntäsolut tarjoavat antigeenin natiivin muodon ja suojaavien epitooppien hyvän reproduktion, kos-

ka eukaryoottisessa ilmentämissysteemissä syntyy samankaltaisempia glykosylaatiokuvioita, disulfidin sillanmuodostusta ja muita translaation jälkeisia modifikaatioita kuin E. colissa, jossa voi muodostua liukenematon proteiini, joka edellyttää uudelleenlaskostumista ja jolla on huono natiivin muodon reproduktio. Lisäksi nisäkäsglykosylaatio epätodennäköisesti indusoi immuunivasteen, mikä poikkeaa suojaavasta antiproteiinivasteesta. Ihmisten ja kotieläinten suojaamiseksi on siten edullista käyttää ihmisen tai eläimen sidekudosmuodostussolu- tai myeloomasolulinjoja kuten HeLa - ihmissolulinja; BHK - vauvahamsterin munuaissolut; VERO, apinan munuaissolulinja; FR3T3, Fisher-rotan sidekudosmuodostussolut; NIH3T3, hiiren sidekudosmuodostussolulinja; C1227I, hiiren rintakasvainsolulinja; CV-1, afrikkalaisen apinan munuaisen sidekudosmuodostussolut; 3T6, hiiren alkion sidekudosmuodostussolut; L-solut, hiirisolulinja; CHO, kiinalaisen hamsterin munasarjasolulinja; NSO NSI, SP2 ja muut hiiren myeloomasolulinjat ja rotan myeloonasolulinjat kuten YB2/0 ja Y3.

Vektoreita, jotka sopivat eri nisäkässolulinjaluokille, tunnetaan alalla hyvin. Yleensä ottaen nämä käsittävät promoottorin ja/tai tehostimen, joka on operatiivisesti yhdistetty antigeenia tai sen fragmenttia koodittavaan nukleotidisekvenssiin. Sopivia promoottoreita ovat SV40-varhais- tai SV40-myöhäispromoottori, esim. PSVL-vektori, sytomegalovirus- (CMV-) promoottori, hiiren metallotioneiini-I-promoottori ja hiiren rintakasvainviruksen pitkä terminaalinen toisto. Vektori sisältää edullisesti sopivan merkin kuten geenin dihydrofolaattireduktaasille tai glutamiinisyntetaasille. Noiden tyyppisiä vektoreita kuvataan julkaisuissa WO 86/05 807, WO 87/04 462, WO 89/01 036 ja WO 89/10 404.

Isäntäsolujen transfektointi voidaan suorittaa käyttäen tavanomaisia tekniikoita, esimerkiksi käyttäen

kalsiumfosfaattia, DEAE-dekstraania, polybreeniä, protoplastifuusiota, liposomeja, suoraa mikroruiskutusta, geenitykkiä tai elektroporaatiota. Viimeksi mainittu tekniikka on edullinen ja menetelmiä nisäkässolulinjojen transfektoimiseksi elektroporaatiota käyttäen kuvaavat Andreason et al., 1980. Yleensä ottaen lineaarinen DNA viedään helpommin kuin ympyrän muotoinen DNA.

Proteiinilla H11OD on havaittu olevan yksilöllinen ja epätavallinen glykosylaatiokuvio, jonka arvellaan myötävaikuttavan immuuniaktiivisuuteen, koska tähän mennessä H11OD:lle Haemonchuksesta saadut monet monoklonaaliset vasta-aineet tunnistavat hiilihydraattiepitooppeja, jotka voivat olla merkityksellisiä käyttökelpoisten rokotteiden kehittämisessä.

Erityisesti seuraava glykosylaatiokuvio H11OD:lle Haemonchuksesta on osoitettu:

i. noin 65 % oligosakkarideista on N-kytketty, loput ovat O-kytkettyjä;

ii. valtaosa (esim. noin 48 %) N-kytketyistä oligosakkarideista edustaa kompleksista luokkaa;

iii. oleellisesti kaikki (esim. yli 95 %) oligosakkarideista on vailla varausta;

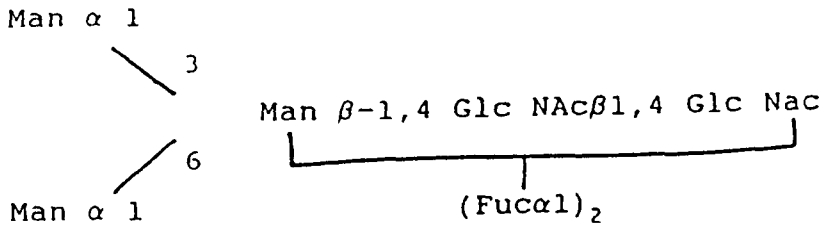
iv. aineosamonosakkaridien suhteellinen moolipitoisuus on N-asetyyligalaktosamiini 1,0, fukoosi 3,6, galaktoosi 4,1, glukoosi 4,4, mannoosi 6,2 ja N-asetyyliglukosamiini 5,2;

v. oligosakkaridit, muut kuin pääasiallinen oligosakkaridi (jota kutsutaan oligosakkaridiksi D), ovat huomattavan vastustuskykyisiä hajoamiselle laajan valikoiman eksoglykosidaaseja toimesta (esim. α-D-mannosidaasi, ß-D-mannosidaasi, ß-D-glukosidaasi, ß-D-galatosidaasi, α-D-galaktosidaasi, α-L-fukosidaasi, ß-D-ksylosidaasi, ß-D-N-asetyyliglukosaminidaasi).

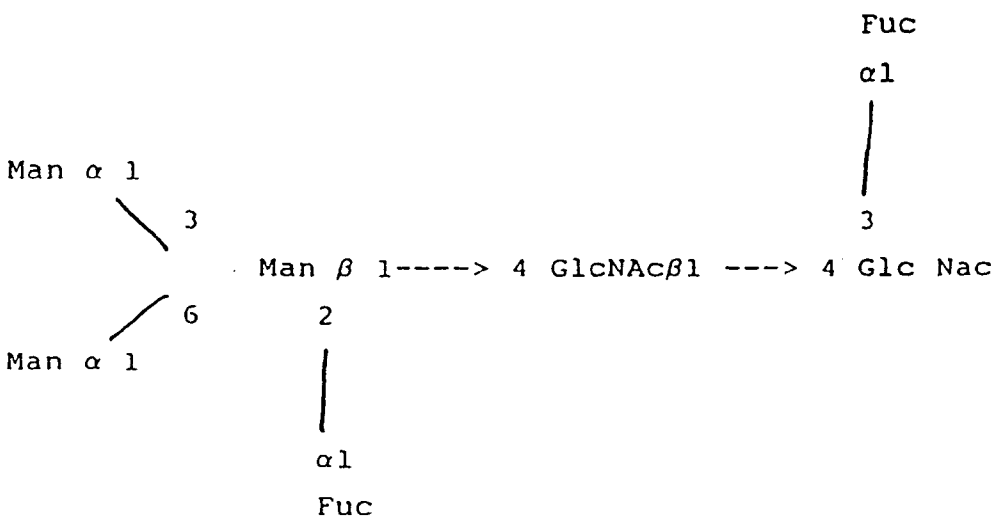
Tällaiset oligosakkardidit ja niitä sisältävät glykoproteiinit muodostavat tämän keksinnön lisänäkökohdan.

Oligosakkaridi D Haemonchus H11OD -glykoproteiinista on N-kytkettyä tyyppiä ja sillä on uusi rakenne, joka koostuu kahdesta fukoositähteestä kiinnittyneinä α-1,3-sidoksen ja α-1,2-sidoksen välityksellä mannoosi-(N-asetyyliglukosamiini)$_2$-ytimeen.

Keksinnön seuraava näkökohta antaa siten käyttöön oligosakkaridin, jonka rakenne on seuraava:



ja tarkemmin ottaen rakenne:



etenkin kytkettynä proteiiniin, esim. yhdistelmä-DNA-proteiiniin kuten matoaminopeptidaasiproteiini tai sen antigeeninen fragmentti, tai käytettynä anti-idiotyyppisten antigeenin muodostamiseksi etenkin hyvin nuorten eläinten immunointiin.

Eläinglykoproteiineissa on yleensä fukoosi-α-1,6-sidoksia ja esillä olevan keksinnön mukaisen oligosakkarin fukoosi-α-1,3-sidos on epätavallinen piirre.

Tätä keksintöä kuvataan hyt yksityiskohtaisemmin erityisesti viitaten proteiiniin H11OD Haemonchus contortuksesta. Kuitenkin erilasilla tekniikoilla, kuten histokemia ja DNA-hybridisaatio H11OD-ekvivalentteja on havaittu muissa loislajeissa. Arvellaan, että H11OD-proteiini on multigeenikompleksi ja että lisäksi sitä koodittavat nukleotidisekvenssit voivat osoittaa sekvenssivariaatioita eri kantojen ja madon eri elinsyklivaiheiden välillä. Lisäksi saattaa olla olemassa monientsyymimuotoja (isoentsyymejä), jotka voivat ilmentyä eri tavalla eri vaiheissa tai eri kannoissa. Tässä tutkimuksessa DNA-sekvenssit ja täten ennustetut aminohapposekvenssit on määritetty cDNA-klooneista ja PCR-tuotteista, jotka on saatu H11OD-geeniä vastaavasta mRNA:sta yhdistelmä-DNA-teknologialla eri lähteistä ja H. contortuksen elinsyklin eri loisvaiheissa.

cDNA- ja PCR-tuotteiden sekventointi on tehnyt meille mahdolliseksi identifioida kolme likeistä sukua olevaa H11OD-sekvenssiä, joita kutsutaan tässä H11-1:ksi (sekvenssi tunnusnro 19), H11-2:ksi (sekvenssi tunnusnro 20) ja H11-3:ksi (sekvenssi tunnusnro 21). H11-1 käsittää kolme jatkuvaa ja limittyvää sekvenssiä, cDNA-klooni AustB1 (sekvenssi tunnusnro 6), PCR-tuote A-648 (sekvenssi tunnusnro 9) ja 3-päässä PCR-tuote 014-178 (sekvenssi tunnusnro 12); H11-2 käsittää PCR-tuotteet A-650 ja 2,5ke (sekvenssit tunnusnrot 10 ja vastaavasti 7); H11-3 käsittää PCR-tuotteet 3,5ke ja A-649 (sekvenssit tunnusnrot 8 ja vastaavasti 11). Yksittäisten sekventoitujen cDNA- ja PCR-tuotekloonien ja H11-1:n, -2:n ja -3:n välisistä spesifisistä suhteista esitetään yhteenveto kuviossa 1 ja ne esitetään yksityiskohtaisesti kuvioissa 3, 4 ja 5.

Eroja ja variaatioita cDNA-klooneista ja PCR-tuotteista saaduissa sekvensseissä on havaittu kuten voidaan nähdä etenkin kuvioista 2, 3, 4 ja 5 (jotka koostuvat sek-

vensseistä tunnusnrot 1 - 15 ja 19 - 21) ja kuten yhteenvetona esitetään taulukossa 1.

Taulukko 1. Pääteltyjen aminohapposekvenssien homologioita, jotka on saatu kuviossa 2 esitettyjen nukleotidisekvenssien translaatiolla

	Samankaltaisuus-%	Identtisyys-%
H11-1:H11-2	77	63
H11-1:H11-3	79	65
H11-2:H11-3	82	69

Erot voidaan katsoa johtuviksi eri mRNA-sekvensseistä (multigeeniryhmästä). Lisäksi variaatiot voivat johtua ainakin osittain H11OD:tä koodittavan sekvenssin eri varianteista tai mRNA:sta, joka esiintyy elinsyklin eri vaiheissa tai kannoissa, jotka eroavat maantieteellisestä alkuperältään.

Taulukossa 2 esitetään lisäksi identtisyys- ja samankaltaisuustasot vastaavien ennustettujan aminohapposekvenssien ja kahden julkistetun nisäkäsaminopeptidaasisekvenssin välillä.

Taulukko 2. H11OD-aminohapposekvenssien homologiat rotan aminopeptidaasiin M (ApM) ja hiiren aminopeptidaasiin A (ApA) nähden

	Samankaltaisuus-%	Identtisyys-%
H11-1:ApM	55	32
H11-1:ApA	55	31
H11-2:ApM	52	31
H11-2:ApA	54	31
H11-3:ApM	53	32
H11-3:ApA	52	30

Kuviossa 1 esitetään kartta sekventoiduista H. contortus H11OD-cDNA- ja PCR-tuotekloneista ja niiden suhteet ja suhteelliset asemat H11OD-mRNA:n pituudella;

kuviossa 2 esitetään H11OD-nukleotidisekvenssit, jotka on nimetty H11-3 (sekvenssi tunnusnro 21, saatu kloonatuista PCR-tuotteista sekvenssit tunnusnrot 8 ja 11

ja cDNA-kloonista M1AUS, sekvenssi tunnusnro 5), H11-2 (sekvenssi tunnusnro 20, saatu kloonatuista PCR-tuotteista sekvenssit tunnusnrot 7 ja 10) ja H11-1 (sekvenssi tunnusnro 19, saatu kloonatuista PCR-tuotteista sekvenssit tunnusnrot 9 ja 12 ja cDNA-kloonista AustB1, sekvenssi tunnusnro 6);

kuviossa 3 esitetään sekvenssi H11-3 (sekvenssi tunnusnro 21) (esitetään kuviossa 2) rinnanasetteluin cDNA-klooneihin M1 ja M1AUS (sekvenssit tunnusnrot 1 ja 5);

kuviossa 4 esitetään sekvenssi H11-2 (sekvenssi tunnusnro 20, esitetään kuviossa 2) ja rinnanasettelu cDNA-klooniin B2 (sekvenssi tunnusnro 4);

kuviossa 5 esitetään sekvenssi nimeltä H11-1 (sekvenssi tunnusnro 19) ja cDNA-B1A:n ja -Aust-B1:n (sekvenssit tunnusnrot 2 ja vastaavasti 6) rinnanasettelu;

kuviossa 6 esitetään a) ennustetut aminohapposekvenssit (sekvenssit tunnusnrot 22, 23 ja 24), jotka on saatu kuviossa 2 esitetyistä DNA-sekvensseistä H11-1, H11-2 ja H11-3; bi) ja ii) esittävät H11-3:n ennustetun aminohapposekvenssin verrattuna rotan mikrosomaalisen aminopeptidaasin M (Watt et al., 1989) ja vastaavasti hiiren mikrosomaalisen aminopeptidaasin A (Wu et al., 1990) julkistettuihin aminohapposekvensseihin; identtisyydet on suljettu laatikoihin, viivat osoittavat tiloja, jotka on lisätty verrattujen sekvenssien välisen homologian tason maksimoimiseksi. Käytetään aminohappojen tavanomaista yksikirjainkoodia. Vaakasuora viiva sekvenssin päällä osoittaa transmembraanisen alueen aseman ja tähdet osoittavat sinkkiin sitoutuvan motiivin aseman. Samankaltaisuus tasot esitetään taulukoissa 1 ja 2;

kuviossa 7 esitetään rinnanasettelut aminohapposekvensseille (nimeltä Pep A, Pep B, Pep C, Pep D ja Pep E), jotka on saatu H11OD:n CNBr- ja Lys-C-fragmenteista, kuten on kuvattu aiemmin (kansainvälinen patenttihakemusjulkai-

su WO 90/11 086 ja kuten luetteloitu aiemmin, polypeptidisekvenssit (a), (b), (e), (k) ja vastaavasti (aa)) ja kolmelle uudelle sekvenssille (sekvenssit tunnusnrot 16, 17 ja 18), jotka on saatu H11OD:stä pilkkomalla elastaasilla tai termolysiinillä a) H11-1:n, b) H11-2:n ja c) H11-3:n translaatioilla.

Lisänäkökohtana keksintö antaa myös käyttöön nukleiinihappomolekyylejä, jotka käsittävät yhden tai useamman nukleotidisekvenssin, joka oleellisesti vastaa tai joka on oleellisesti komplementaarinen yhteen tai useampaan sekvenssiin nähden, joka valitaan kloonien M1, B1A, B1A-3', B2, M1AUS, AustB1, 014-015 (2,5PCR), 014-872 (3,5PCR-klooni 2), A-648 (5'-pää B1:stä), A-650 (5'-pää 2,5PCR:stä), A-649 (5'-pää 3,5PCR:stä), 014-178 (3'-pää AustB1-kloonista 2), 014-178 (3'-pää AustB1-klooneista 3 & 6), 014-872 (3,5PCR-klooni 10) ja 014-872 (3,5PCR-klooni 19), H11-1, H11-2 ja H11-3, sekvenssit tunnusnrot 1 - 15 ja vastaavasti 19 - 21 kuten esitetään kuvioissa 2, 3, 4 ja 5, sekvensseistä, tai sekvensseihin, jotka ovat oleellisesti homologisia mihin tahansa mainittuihin sekvensseihin nähden tai hybridisoituvat niihin.

Kuten esllä mainittiin, erilaisten edellä mainittujen kloonien sekvenssien vertaaminen tunnettujen sekvenssien tietokonetietokantoihin paljastaa oleellista homologiaa mikrosomaalisten aminopeptidaasientsyymien (EC. 3.4.11.-) ryhmään. H11OD-proteiinilla ja sen alafraktioilla suoritetut entsymologiset aktiivisuus- ja inhibiittoritutkimukset vahvistavat, että proteiini on itse asiassa mikrosomaalinen aminopeptidaasi (α-aminoasyylipeptidihydrolaasi (mikrosomaalinen)). Tällaiset tutkimukset ovat edelleen osoittaneet, että sekä aminopeptidaasi-A:n kaltaisia että aminopeptidaasi-M:m kaltaisia aktiivisuuksia esiintyy, ja että H11OD-dubletin kukin rakenneosa yksittäin osoittaa entsyymiaktiivisuutta.

Tutkimuksia H11OD:n proteolyyttisellä pilkkomisella on myös suoritettu. Käyttäen entsyymiä elastaasi todettiin, että H11OD voidaan osittain pilkkoa, jolloin muodostuu kaksi fraktiota, detergenttiin liukoinen fraktio (joka jäi membraaniin) ja vesiliukoinen fraktio (jota kutsutaan H11S:ksi). H11S esiintyy proteiinidimeerin muodossa, joka voidaan pelkistää kahdeksi rakenneosaksi. Kiinnostavasti todettiin, että vain aminopeptidaasi-M:n kaltaista aktiivisuutta liittyy vesiliukoiseen H11S-fraktioon, kun taas aminopeptidaasi-A:n kaltainen aktiivisuus liittyy vain detergenttiin liukoiseen fraktioon.

Seuraava esimerkki tarjoaa kuvauksen tutkimuksista, jotka johtavat kuvioissa 1 - 7 esitettyjen sekvenssien määritykseen, viitaten seuraaviin lisäkuvioihin, joissa:

kuviossa 8 esitetään Western-blot-määritykset integraalisille membraaniproteiineille, jotka esiintyvät täysikasvuisten Haemonchus contortus -organismien detergenttiuutteessa, jota on seulottu affiniteettipuhdistetuilla vasta-aineilla, jotka on eluoitu mahdollisista H11OD-klooneista; a) antigeenit Haemonchus-detergenttiuutteessa, jotka tunnisti uutteen vastaseerumi; b) vasta-aineet, jotka eluoitiin liuskasta kuten esitetään kohdassa a) uudelleentestattuina vastaan detergenttiuute-Blot-määritystä, vahvistavat eluutiovaiheen onnistumisen; c) vasta-aineet kuten b):ssä, jotka sitoutuvat kloonin M1 ilmentämään proteiiniin voimakkaasti, tunnistavat alueen kohdassa 110 kd (ja suhteellisen terävän nauhan kohdassa noin 205kd; d) ei tapahdu vasta-aineen sitoutumista käytettäessä ei-rekombinanttia seerumin adsorboimiseksi;

kuviossa 9 esitetään Northern-blot mRNA:lle, joka on puhdistettu 11, 15 ja 23 päivän ikäisistä Haemonchus contortus -organismeista seulottuna a) cDNA-kloonilla M1 (sekvenssi tunnusnro 1); b) cDNA-kloonilla M1 AUS (sekvenssi tunnusnro 5); c) cDNA-kloonilla B1A (sekvenssit tunnusnrot 2 ja 3); d) cDNA-kloonilla AustB1 (sekvenssi

tunnusnro 6); e) kloonatulla PCR-tuotteella 014-872 (3,5-2, sekvenssi tunnusnro 8); ja f) kloonatulla PCR-tuotteella 014-015 (sekvenssi tunnusnro 7). Numerot 11, 15 ja 23 osoittavat Haemonchuksen iän, josta mRNA saatiin;

kuviossa 10 esitetään Southern-blot-määritykset Haemonchus contortus -genomi-DNA:lle, jota on seulottu cDNA-klooneilla M1AUS (sekvenssi tunnusnro 5), B1A (sekvenssit tunnusnrot 2 ja 3) ja AustB1 (sekvenssi tunnusnro 6) ja PCR-tuotteilla 014-872 (3,5-2, sekvenssi tunnusnro 8) ja 014-015 (sekvenssi tunnusnro 7); a) blot-määritykset pestiin kohtalaisen ankaruuden olosuhteissa, b) blot-määritykset pestiin korkean ankaruuden olosuhteissa; kullekin koettimelle vyöhyke 1 sisälsi HindIII-pilkkomistuotteen lambda-DNA:sta merkkinä tai jätettiin tyhjäksi, vyöhykkeet 2 ja 3 sisälsivät vastaavasti EcoRI- ja HindIII-uutteita Haemonchus-genomi-DNA:sta;

kuviossa 11 esitetään Western-blot-määritykset yhdistelmä-GST-M1- ja -GST-B1A-fuusioproteiineille, jotka on seulottu affiniteettipuhdistetuilla vasta-aineilla elektroforeettisesti puhdistettua H11OD:tä (H11ODE) vastaan;

kuviossa 12 esitetään Western-blot-määritykset ConA-H11D-antigeenille, jota on seulottu ConA-H11OD:n ja yhdistelmä-GST-M1- ja -GST-B1A-fuusioproteiinien vastaseerumilla;

kuviossa 13 esitetään a) tulokset H11OD-proteiini- ja -aminopeptidaasientsyymiaktiivisuuksien analyysi fraktioissa, jotka on saatu ConA-H11OD:n ioninvaihtokromatografiasta MonoQ-kolonnissa;

b) SDS-PAGE kuviossa 13a) esitetyistä fraktioista;

kuviossa 14 esitetään a) pH-arvot, joilla fraktioita saatiin vapaavirtaus-isoelektrinen fokusointi-kokeessa;

b) SDS-PAGE pelkistävissä olosuhteissa fraktioille 14a):sta, jossa H11OD-dubletin alempi nauha tavataan fraktiosta 6 ja ylempi nauha fraktiosta 16, jolloin kumpaakin on vaihteleva määrä välissä olevissa fraktioissa;

c) Western-blot-määritykset 14b):ssä esitetyille fraktioille seulottuina i) monoklonaalisilla vasta-aineilla nimeltä TS 3/19.7 ja ii) affiniteettipuhdistetuilla polyklonaalisilla anti-M1-vasta-aineilla; kontrollivasta-aineet eivät antaa todettavaa reaktioita;

kuviossa 15 esitetään a) pH-arvot, joilla fraktioita saatiin toisessa vapaavirtaus-isoelektrinen fokusointikokeessa; b) SDS-PAGE pelkistävissä olosuhteissa fraktioille 15a):sta, joita käytettiin entsyymimäärityksissä, jossa H110D-dubletin alempi nauha tavataan fraktioissa 4 - 6 ja ylempi nauha fraktioissa 16 - 18, jolloin vaihtelevia määriä kumpaakin nauhaa on välissä olevissa fraktiosisa; c) 15b):ssä esitettyjen fraktioiden mikrosomaalisen aminopeptidaasin spesifiset aktiivisuudet;

kuviossa 16 esitetään lampaiden suojaaminen rokottamalla erillisillä ylemmillä (U), alemmilla (L), yhdistelmä- (U+L) ja välidubletti- (D) nauhoilla H110D:stä: a) loismunatulos ilmaistuna munina/g ulostetta, b) matokuorma postmortemissa, suhteessa verrokkeihin;

kuviossa 17 esitetään lampaiden suojaaminen rokottamalla vesiliukoisella fragmentilla (H11S), joka on saatu H110D:stä pilkkomalla elastaasilla ja H11A:lla, jäännöksenä saatu detergenttiliukoinen H110D. a) loismunatulos, ilmaistuna munina/g ulostetta; b) matokuorma postmortemissa, suhteessa verrokkeihin (C);

kuviossa 18 esitetään esimerkkejä suhteesta Ai), Bi) aminopeptidaasi-M-kaltaisten ja Aii), Bii) aminopeptidaasi-A-kaltaisten aktiivisuuksien välillä H110D:stä vastaseerumilla yksittäisistä lampaista, jotka on rokotettu H110D:llä, mitaten suojaustasot Ai:llä, ii) prosentuaalinen lasku matokuormassa postmortemissa ja Bi:llä, ii) prosentuaalinen lasku ulosteen munaluvussa; avoin neliö anti-H110D, umpinainen neliö anti-hevosferritiiniverrokki;

kuviossa 19 esitetään aminopeptidaasientsyymiaktiivisuuksien histokemiallinen paikannus täysikavsuisessa

Haemonchus contortus -organismissa - valomikrografit täysikasvuisen naaraspuolisen Haemonchus contortus -organismin kryo-sektioista osoittavat aminopeptidaasiaktiivisuutta (luettu reaktiotuote esiintyy tummana nauhana (varustettu nuolella) näissä mustavalkoisissa valokuvissa), joka liittyy vain suolen solukalvon ohuisiin ulokkeisiin (i). Mikään muista kudoksista (esim. ihon orvaskesi (c), ihonalainen kudos (h) sukupuolielinten alue (gt), seinälihas (wm)) ei osoita aktiivisuutta. a):ssa substraattina oli L-leusiini-4-metoksi-ß-naftyyliamidi, b):ssä substraatti oli L-glutamiinihappo-α-(4-metoksi-ß-naftyyliamidi);

kuviossa 20 esitetään kartta 3,5 PCR -tuotteesta (klooni 2) (sekvenssi tunnusnro 8) alakloonattuna bakulovirusilmentämisvektoriin pBlueBacII;

kuviossa 21 esitetään Western-blot uutteille bakuloviruksella tartutetuista hyönteisen Spodoptera frugiperda (Sf) soluista seulottuina anti-H11ODN-vasta-aineilla. Kaksi kloonattua plakkia, P3A ja P4A, ilmensivät täysimittaisen immuunipositiivisen H11OD:n (merkitty nuolella), verrokit eivät.

Esimerkki

Menetelmät

UK-lambda-GT11-kirjaston rakentaminen

mRNA:n eristäminen

Nestetypessä pikajäädytettyjä, UK:sta peräisin olevia täysikasvuisia Haemonchus contortus -organismeja (0,5 g) jauhettiin nestetypessä käyttäen esijäähdytettyä huhmaretta ja survinta. RNA uutettiin survoksesta 10 tilavuudella 4 M guanidiinihydrokloridiliuosta 25 mM natriumsitraatissa, joka sisälsi 0,5 % w/v sarkosyyliä ja 0,7 % w/v 2-merkaptoetanolia, minkä jälkeen uutettiin fenolilla ja kloroformilla käyttäen Chomczynskin ja Sacchin menetelmää (1987). Lähdetti-RNA (mRNA) valmistettiin tästä affiniteettikromatografisesti oligo-dT-selluloosassa (kahdes-

ti) kuten kuvataan Maniatiksen et al. julkaisussa (1982) ja laatu määritettiin in vitro -translaatiolla käyttäen kaniinin retikulosyyttilysaattipakkausta ja ^{35}S-metioniinia valmistajalta Amersham International plc valmistajan ohjeita noudattaen. Kooltaan aina 120 kd:n polypeptidejä todettiin.

Komplementaarisen DNA:n valmistus

Ensimmäinen säie komplementaarista DNA:ta (cDNA) syntetisoitiin 1 µg:sta mRNA:ta käyttäen satunnaispohjustusta ja linnun käänteiskopioijaentsyymiä ja toinen säie syntetisoitiin käyttäen korvausreaktiota RN-aasi-H:lla ja E. colin DNA-polymeraasi-I:llä, mitä seurasi 3'-ylijäämäpäiden korjaaminen käyttäen T4-DNA-polymeraasia Gublerin ja Hoffmanin menetelmän mukaan (1983). Kaksisäikeisen (ds) cDNA:n saanto oli noin 400 ng 1 µg:sta mRNA:ta. ds-cDNA:ta tutkittiin elektroforeesilla 1-%:isessa agaroosigeelissä, mitä seurasi autoradiokuvan otto. ds-cDNA oli kooltaan 0,2 - 9,4 kiloemästä (ke) valtaosan ollessa kooltaan 0,5 - 2,3 ke.

cDNA:n kloonuas lambda-gt11:een

Ei-kokovalittua cDNA:ta käytettiin kirjaston rakentamiseksi lambda-gt11:een käyttäen Amershamin cDNA-kloonaussysteemiä (pakkaus nro RPN 1280, Amersham International plc) ja in vitro -pakkausuutteita (pakkaus nro N334, Amersham International plc), kuten kuvataan valmistajien ohjeissa, ja EcoRI-liitto-oligonukleotidejä (5'GGAATTCC). Saatu kirjasto maljattiin E. coli -kannalle Y1090 isopropyylitio-ß-D-galaktosidin (IPTG) ja 5-bromi,4-kloori,3-indolyyli-ß-D-galaktosidin (X-gal) läsnä ollessa, joissa olosuhteissa yhdistelmä-DNA-lambda-gt11 ilmenee kirkkaina ("valkosina") plakkeina ja villityypin ei-yhdistelmä-DNA-lambda-gt11 sinisinä plakkeina. Kirjasto sisälsi 90 % valkoisia plakkeja ja kloonaustehon laskettiin olevan 4 x 10^7 plakin muodostavaa yksikköä (pfu)/µg cDNA:ta ja kirjaston tiitterin 2 x 10^6 plakin muodostavaa yksikköä millilitraa

kohden. DNA:n analyysi 20:sta satunnaisesti noukitusta rekombinantista paljasti insertin keskimääräiseksi kooksi 0,51 ke. Kuitenkin tätä keskiarvoa vääristi yksi klooni, jossa insertti oli 3,5 ke. Valtaosa inserteistä oli > 300 emäsparia (ep). Tämä ei-vahvistettu UK-mato-mRNA:sta saatu lambda-gt11-kirjasto immuuniseulottiin sitten.

Vasta-ainekoettimien valmistaminen

Vastaseerumi integraalisille membraaniproteiineille

Suolia leikattiin täysikasvuisista Haemonchus contortus -organismeista (UK-alkuperä) ja homogenisoitiin jääkylmässä fosfaattipuskuroidussa suolaliuoksessa (PBS), pH 7,4, joka sisälsi pitoisuudet 1 mM etyleenidiamiinitetraetikkahappo (EDTA) ja 1 mM fenyylimetyylisulfonyylifluoridi (PMSF). Homogenaattia sentrifugoitiin 10 minuutin ajan käyttäen mikrofugia ja pelletti uudelleensuspendoitiin samaan puskuriin, joka sisälsi 0,1 % v/v Tween 20 -valmistetta (Tween on tavaramerkki). Uudelleensentrifugoinnin jälkeen pelletti uudelleensuspendoitiin samaan puskuriin, joka sisälsi 2 % v/v Triton X-100 -valmistetta, ja uutettiin 2 tunnin ajan 4 °C:ssa. Tämä uute sentrifugoitiin kuten edellä supernatantin saamiseksi, joka sisälsi integraalisia membraaniproteiineja (IMP).

Lammas hyperimmunoitiin IMP:llä Freundin täydellisessä tehosteessa (FCA) ruiskuttamalla lihaksensisäisesti 50, 50, 120 ja 130 µg IMP:tä annettuina viikkoina 0, 7, 11 ja 15. 6 viikon kuluttua viimeisestä ruiskeesta seerumi kerättiin talteen ja nimettiin seerumiksi EE-068.

Integraalisten membraaniproteiinien valmistus detergenttiuutolla Haemonchus contortuksesta

Uute valmistettiin homogenisoimalla matoja 5 - 10 tilavuudessa PBS:ää, joka sisälsi pitoisuudet 1 mM EDTA ja 1 mM PMSF. Suspensiota sentrifugoitiin 10 000 x g:ssä 20 minuutin ajan 4 °C:ssa ja pelletti pestiin samassa puskurissa, joka sisälsi 0,1 % v/v Tween 20 -valmistetta, minkä

jälkeen uutettiin 5 tilavuudella 2-%:ista v/v Triton X-100 -valmisteliuosta kuten edellä kuvattiin. Supernatanttia uudelleensentrifugoitiin 100 000 x g:ssä 1 tunnin ajan ja saatua supernatanttia, joka oli rikastunut H11OD:n suhteen mutta sisälsi muuta IMP:tä, käytettiin Western-blot-kokeissa sekä ei-denaturoidun H11OD:n valmistukseen (katso alla).

H11OD:n ja affiniteettipuhdistetun anti-H11ODN:n valmistus

H11OD:n suhteen rikastetulla uutteella suoritettiin affiniteettikromatografia ConA-agaroosissa, mitä seurasi ioninvaihtokromatografia MonoQ-kolonnissa (kuten kuvataan julkaisuissa WO88/00835 ja WO90/11086). Puhdistettua H11OD:tä ruiskutettiin lihaksensisäisesti karitsoihin FCA:ssa. Kolme 100 µg:n annosta annettiin 3 viikon välein. Seerumia, joka kerättiin karitsoista 4 viikon kuluttua viimeisestä ruiskeesta, affiniteettipuhdistettiin absorboimalla kolonniin, joka sisälsi puhdistettua H11OD:tä, joka oli kyketty syaanibromidilla aktivoituun Sepharose-valmisteeseen (Pharmacia). H11OD:n kytkeminen Sepharose-valmisteeseen, vastaseerumin sitominen ja anti-H11OD-vasta-aineiden eluointi suoritettiin Pharmacian toimittamien ohjeiden mukaan. Nämä affinitettipuhdistetut vasta-aineet nimetään anti-H11ODN:ksi. "N" erottaa nämä vasta-aineet niistä, jotka on tuotettu vastaan denaturoitua, elektroforeettisesti puhdistettua H11OD:tä ja joita kutsutaan anti-H11ODE:ksi.

Western-blot-määritys

Western-blot-määritykset suoritettiin käyttäen tavanomaisia menettelyjä (Johnstone et al., 1982).

Kloonien eristäminen ja karakterisointi

UK-lambda-gt11-kirjaston immuuniseulonta

Menetelmä, jota käytettiin kirjaston immuuniseulontaan, oli oleellisesti Bowtellin et al. (1986) kuvaaman mukainen. Ennen käyttöä seerumista (EE-068) poistettiin

anti-E. coli-vasta-aineet absorboimalla E. colin Y1090 lysaateilla ja kokonaisilla soluilla. Kirjasto maljattiin E. coli Y1090 -soluille tiheydeksi 10^3 pfu kohden maljaa, jonka halkaisija oli 90 mm. Maljapinnoille asetettiin nitroselluloosasuodattimet, jotka oli imeytetty IPTG:llä, ja inkuboitiin yön yli. Suodattimet pestiin TBST:llä (50 mM Tris, pH 7,4, 150 mM NaCl, 0,05 % v/v Tween 20 -valmistetta), minkä jälkeen niitä salvattiin 5-%:isella v/v hevosseerumilla TBST:ssä 2 tunnin ajan. Seerumia EE-068 laimennettuna 1:200 TBST:hen, joka sisälsi 5 % hevosseerumia, lisättiin ja suodattimia inkuboitiin 4 tunnin ajan kevyesti keinuttaen. Suodattimet pestiin jälleen TBST:llä, minkä jälkeen niitä inkuboitiin piparjuuriperoksidaasiin (HPR) konjugoidulla hevosen anti-lammas-IgG:llä laimennettuna 1:500 TBST:hen, joka sisälsi 5 % v/v hevosseerumia, 2 tunnin ajan. (Vastaseerumi lammas-IgG:lle tuotettiin hevosessa, anti-lammas-IgG:tä puhdistettiin affiniteettikromatografisesti lammas-IgG-Sepharose-kolonissa ja vasta-aineet konjugoitiin HPR:ään Nakanen ja Kawaoin menetelmällä, 1974.) Suodattimet pestiin edelleen TBST:llä ja positiiviset plakit todettiin käyttäen 0,6 mg/ml 3,3'-diaminobentsidiini- (DAB-) liuosta ja 0,1-%:ista v/v vetyperoksidia. 25 oletettua positiivista plakkia noukittiin ja ne uudelleenseulottiin affiniteettipuhdistetulla anti-H110DN:llä kuten edellä kuvattiin. Tämän toisen seulonnan jälkeen 5 rekombinanttia oli edelleen positiivisia, jolloin M1:ksi nimetty klooni antoi voimakkaimman signaalin.

Vasta-aineen affiniteettipuhdistus yhdistelmä-DNA-fagilla

Yhteenkasvaneita maljoja valmistettiin E. coli Y1090 -kerroksille maljaamalla 10^3 pfu kutakin vasta-ainepositiivista lambda-kloonia tai negatiivista ei-yhdistelmä-DNA-lambda-gt11-verrokkifagia. Kerroksia inkuboitiin 4 tunnin ajan 42 °C:ssa, minkä jälkeen niille asetettiin IPTG:llä imeytetyt suodattimet ja inkuboitiin edelleen yön

yli 37 °C:ssa. Suodattimet poistettiin maljoista ja pestiin TBST:ssä, minkä jälkeen niitä salvattiin 5-%:isella v/v hevosseerumilla 1 tunnin ajan. Sitten suodattimia inkuboitiin 1:100 laimennoksella vastaseerumia EE-068 6 tunnin ajan, minkä jälkeen ne huuhdottiin perusteellisesti TBST:llä. Sitoutuneet vasta-aineet eluoitiin suodattimilta käyttäen 2 kertaa 2 ml eluutiopuskuria (5 mM glysiini, 500 mM NaCl, 0,2 % Tween 20 -valmistetta, pH 2,3) kulloinkin 2 - 3 minuutin ajan, minkä jälkeen ne neutraloitiin lisäämällä 200 µl 1 M tris-HCl-liuosta, pH 7,4, niitä laimennettiin 1:200 ja käytettiin H11OD-rikastetun uutteen Western-blot-määrityksen immuuniseulomiseksi.

M1-kloonin DNA-sekventointi

Lambda-DNA eristettiin M1-kloonista menetelmien mukaan, joita kuvataan julkaisussa Maniatis et al. (1982). 2,38 ke:n KpnI-SstI-fragmentti, joka sisältää 300 ep:n M1-fragmentin, eristettiin geelielektroforeettisesti, sitä puhdistettiin käyttäen Geneclean-pakkausta (Stratagene) (Geneclean on BI0101:n rekisteröity tavaramerkki) ja se alakloonattiin pBluescriptII SK -vektoriin (Stratagene). EcoRI-fragmenttia puhdistettiin käyttäen samoja menetelmiä ja se uudelleenalakloonattiin samaan vektoriin.

M1-insertin nukleotidisekvenssi määritettiin käyttäen T7-sekventointipakkausta (Pharmacia, UK) käyttäen sekä M13-eteenpäin- että käänteisalukkeita.

Austraalialaisten lambda-GT11- ja lambda-ZAP-cDNA-kirjastojen valmistus

mRNA:n eristäminen

5 g nestetypessä pikajäädytettyjä Haemonchus contortus -organismeja (austraalialainen altis McMaster-kanta) jauhettiin nestetypessä ja RNA uutettiin käyttäen kuumaa fenolia Cordingleyn et al. (1983) menetelmän mukaan. Kokonais-RNA-saanto oli 10,35 g. 1,3 mg:sta tätä RNA:ta valmistettiin mRNA:ta affiniteettikromatografisesti oligo-dT-selluloosassa (2 peräkkäistä puhdistusta)

käyttäen Maniatiksen et al. (1982) kuvaamaa menetelmää. mRNA-saanto oli 21,6 µg. mRNA laatu määritettiin in vitro -translaatiolla kaniinin retikulosyyttilysaatissa ^{35}S-metioniinin (Amersham) läsnä ollessa valmistajan ohjeiden mukaan. Saaduilla translaatiotuotteilla oli selvästi erilliset nauhat mukaan lukien kooltaan > 200 kd:n nauhat, kuten osoitettiin elektroforeettisesti SDS-polyakryyliamidigeeleissä ja tätä seuraavalla fluorografialla.

cDNA:n synteesi ja kirjaston valmistus

1 µg:sta mRNA:ta valmistettiin cDNA:ta pohjustaen oligo-dT- tai satunnaisalukkeilla käyttäen cDNA-synteesipakkausta valmistajalta Amersham International plc valmistajan ohjeita noudattaen. Saanto oli 115 ng kaksisäikeista (ds) cDNA:ta. cDNA:n laatua tutkittiin ^{32}P-leimatun DNA:n elektroforeesilla alkalisessa agaroosigeelissä kuten kuvataan Amersham cDNA-pakkauksen ohjeissa. cDNA:n koko (verrattuna lambda-HindIII-merkkeihin, New England Biolabs) oli 150 ep - >10 ke, jolloin valtaosa tuotteista oli kokoalueella 0,6 - 5 ke. Oligo-dT-pohjustetut ja satunnaispohjustetut ds-cDNA:t yhdistettiin pooliksi ja ligatoitiin ylimäärään EcoRI-8-meeri-liittäjiä (5'GGAATTCC3', New England Biolabs, kataloginro 1018), jotka oli leimattu gamma-^{32}P-ATP:llä ja T4-polynukleotidikinaasilla. Liittäjillä varustettua cDNA:ta pilkottiin EcoRI:llä ja liittäjäylimäärä poistettiin kromatografia-ajolla Sepharose 4B -valmisteessa (Pharmacia) Maniatiksen et al. (1982) kuvaamien menetelmien mukaan. Fraktiot kolonnista yhdistettiin pooliksi kahtena eränä, joista yksi sisälsi kooltaan yli 2 ke:n cDNA:ta ja toinen kooltaan alle 2 ke:n cDNA:ta. Kutakin poolia ligatoitiin sitten erikseen 1 µg:aan EcoRI:llä pilkottuja, fosfataasilla käsiteltyjä lambda-ZapII-varsia (Stratagene) ja ne pakattiin erikseen käyttäen Gigapack Gold -tuotetta (Stratagene, rekisteröity tavaramerkki). Kooltaan suuremmasta cDNA:sta saatiin 1,3 x 10^5 rekombinanttia ja pienemmästä cDNA:sta 1,4 x 10^5 rekom-

binanttia; nämä yhdistettiin pooliksi, jolloin saatiin kirjasto, jonka koko oli 2,7 x 10^5. Lambda-Zap-kirjastoa vahvistettiin maljaamalla XL1-Blue-soluille (Stratagene) määränä 2 x 10^4 pfu kohden maljaa, jonka halkaisija oli 135 mm. Vahvistetun kirjaston tiitteri oli 7 x 10^7 pfu/ml.

Edelleen 2 µg mRNA:ta käytettiin cDNA:n valmistamiseksi kuten kuvattiin edellä, mutta käyttäen vain oligo-dT:tä alukkeena. ds-cDNA-saanto oli 740 ng. Tätä cDNA:ta käsiteltiin EcoRI-metylaasilla kuten kuvataan julkaisussa Maniatis et al. (1982) ennen EcoRI-liittäjien lisäämistä, jolloin tässä tapauksessa käytettiin 12-meeriliittäjiä (5'CCGGAATTCCGG3', New England Biolabs, kataloginro 1019). Liittäjät käsittävän cDNA:n pilkkomisen EcoRI:llä jälkeen kaikki fraktiot Sepharose 4B -kolonnista, jotka sisälsivät cDNA:ta, yhdistettiin pooliksi, ja ligatoitiin 2 µg:aan EcoRI:llä pilkottuja, fosfataasilla käsiteltyjä lambda-gt11-varsia (Stratagene). Ligatointiseos jaettiin kahtia ja pakattiin käyttäen kahta Gigapack Gold -erää (Stratagene); nämä pooliksi yhdistämällä saatiin lambda-gt11-kirjasto, jonka koko oli 7 x 10^6 pfu. Kirjastoa vahvistettiin maljaamalla ST9-soluille määränä 5 x 10^5 pfu kohden maljaa, jonka halkaisija oli 135 mm. Vahvistetun lambda-gt11-kirjaston tiitteri oli 4,5 x 10^{11} pfu/ml.

Austraalialaisen lambda-gt11-kirjaston seulominen H11OD:n vastaseerumilla

Vastaseerumia muodostettiin ruiskuttamalla lampaisiin H11OD-proteiinia (UK-alkuperä), joka oli elektroeluoitu polyakryyliamidista elektroforeesin SDS:ssä jälkeen seuraavan menetelmän mukaan: ConA-H11OD:llä, joka oli valmistettu kuten kuvataan julkaisuissa WO 88/00 835 ja WO 90/11 086, suoritettiin elektroforeesiajo SDS-polyakryyliamidigeeleissä (Laemmli 1970) elektroeluoidun H11OD:n saamiseksi. Elektroforeesin jälkeen polyakryyliamidigeelin H11OD:tä sisältävä alue leikattiin irti, se sijoitettiin elektroeluoijaan (Atto) ja eluoitiin 3 tunnin ajan 10 wa-

tin teholla. Elektroeluoitu H110D (joka nimettiin H110DE:ksi) konsentroitiin Centriprep 10 -laitteessa (Amicon) ja puskurivaihdettiin PD10-kolonnissa (Pharmacia) 50 mM ammoniumbikarbonaattiliuokseen 0,07-%:isessa SDS:ssä, sekoitettiin tehosteisiin ja ruiskutettiin sitten lampaisiin. Immunoglobuliinit seerumista saostettiin ammoniumsulfaatilla (Johnstone ja Thorpe, 1982). Saostetut vasta-aineet uudelleensuspendoitiin pitoisuudeksi 60 mg/ml fosfaattipuskuroituun suolaliuokseen, dialysoitiin vastaan fosfaattipuskuroitua suolaliuosta ja laimennettiin 1:10 Tris-puskuroituun suolaliuokseen (TBS), joka sisälsi 5 % w/v rasvatonta maitojauhetta. 10 mg ConA-H110D:tä laimennettiin pitoisuudeksi 0,5 % SDS:ään, liuosta kuumennettiin 100 °C:ssa 3 minuutin ajan ja kuivattiin nitroselluloosasuodattimelle. Suodatin pestiin TBS:llä, joka sisälsi 0,2 % v/v Tween 20 -valmistetta ja 0,5 % Triton X-100 -valmistetta (TBSTT), minkä jälkeen sitä inkuboitiin 1 - 2 tunnin ajan huoneenlämpötilassa H110DE-vastaisilla vasta-aineilla. Suodatinta pestiin 2 tunnin ajan TBSTT:llä, minkä jälkeen sitoutuneita vasta-aineita eluoitiin 3 ml:lla liuosta, joka käsitti pitoisuudet 0,1 M glysiini, 0,15 M NaCl, pH 2,6, 2 minuutin ajan, minkä jälkeen pH säädettiin välittömästi neutraaliksi lisäämällä 75 µl 1,5 M Tris-liuosta, pH 8,0. Näillä affiniteettipuhdistetuilla vasta-aineilla, jotka nimettiin anti-H110DE:ksi, seulottiin 5 x 10^5 pfu austraalialaista lambda-gt11-cDNA-kirjastoa kuten edellä kuvattiin.

5 x 10^5 rekombinanttia austraalialaisesta Haemonchus contortuksesta saadusta lambda-gt11-kirjastosta immuuniseulottiin ja talteen saatiin kolme positiivista. Lisäseulonnan jälkeen 2 näistä rekombinanteista oli edelleen positiivisia ja ne nimettiin B1A:ksi ja B2:ksi.

B1A- ja B2-kloonien sekventointi

Nämä kaksi kloonia pilkottiin EcoRI:llä, jolloin saatiin yksittäinen noin 500 ep:n insertti B1A:lle ja

kolme fragmenttia, B2A (noin 400 ep), B2B (noin 100 ep) ja B2C (noin 100 ep), B2:lle. Nämä alakloonattiin pBluescript SK -vektoriin (Stratagene) ja sekventoitiin käyttäen Sequenase 2.0 -pakkausta (United States Biochemicals).

Kloonien M1 ja B1A ilmentäminen

M1- (sekvenssi tunnusnro 1) ja B1A- (sekvenssit tunnusnrot 2 ja 3) insertit ilmennettiin E. colissa käyttäen pGEX-vektoria (Smith ja Johnson, 1988). Tämä vektori ilmentää proteiineja Schistosoma japonicumin glutationi-S-transferaasin (GST) C-päässä. M1- ja B1A-EcoRI-insertit ligatoitiin EcoRI:llä pilkottuun, fosfataasikäsiteltyyn pGEX1:een ja transformoitiin E. coli -kantaan JM101 menetelmän mukaan, jota kuvataan julkaisussa Maniatis et al., 1982. Kahdeksan jälkeläistä noukittiin kustakin transformaatiosta ja 2 ml:n viljelmiä kasvatettiin 6 tunnin ajan 37 °C:ssa. IPTG:tä lisättiin fuusioproteiinisynteesin indusoimiseksi ja inkubointi jatkui yön yli. Solut kerättiin talteen sentrifugoimalla, ne rikottiin keittämällä näytepuskurissa (Laemmli, 1974) ja uutteet analysoitiin SDS-PAGE:lla ja Western-blot-määrityksellä käyttäen affiniteettipuhdistettuja lammasvasta-aineita, jotka ovat spesifisiä SDS-denaturoidulle H11OD-dubletille (anti-H11ODE, katso edellä). Sitoutuneet vasta-aineet todettiin käyttäen alkaliseen fosfataasiin konjugoitua kaniinin anti-lammas-IgG-alkalinen fosfataasi-konjugaattia (Jackson Immunoresearch), mitä seurasi värin kehittäminen käyttäen 5-bromi,4-kloori,3-indolyyli-fosfaattia (BCIP) ja nitrosinitetratsoliumia (NBT). Immuunipositiivisten kloonien viljelmiä kasvatettiin ja indusoitiin kuten edellä ja solut rikottiin sonikoimalla. Sonikaatit erotettiin liukoisiin ja ei-liukoisiin fraktioihin sentrifugoimalla (Sorvall RC-2B -sentrifugi, HS4-roottori, 7 000 kierr./min, 30 minuuttia, 4 °C). Ei-liukoiset pelletit uudelleensuspendoitiin 8 M urealiuokseen sonikoimalla ja fraktionäytteitä tutkittiin SDS-PAGE:lla. Fuusioproteiinien todettiin olevan ei-liu-

koisessa inkluusiokappalefraktiossa. Kullakin näillä valmisteilla rokotettiin 2 lammasta 3 kertaan määrällä 150 µg fuusioproteiinia annosta kohden Freundin tehosteissa. Positiiviset verrokkilampaat immunoitiin natiivilla ConA-H11OD-proteiinilla, ja negatiiviset verrokkilampaat immunoitiin liuotetulla proteiinilla E. colista, joka sisälsi pGEX-vektorin ilman Haemonchus-inserttiä. Seerumia rokotetuista lampaista analysoitiin Western-blot-määrityksellä H11OD:tä vastaan.

Austraalialaisen lambda-ZAP-kirjaston seulominen DNA-hybridisaatiolla M1- ja B1A-inserteillä

M1- ja B1A-plasmidi-DNA:t (kloonattu pBluescript-vektoriin) pilkottiin EcoRI:llä ja insertit eristettiin elektroforeettisesti TBE-puskurissa (tris-boraatti-EDTA; 89 mM tris-boraatti, 89 mM boorihappo, 2 mM EDTA, pH noin 8,3) 1-%:isessa agaroosigeelissä, mitä sseurasi puhdistaminen käyttäen Geneclean-pakkausta. Eristäminen ja puhdistaminen toistettiin, jottei koetin kontaminoituisi plasmidi-DNA-sekvensseillä, jotka hybridisoituisivat lambda-ZAP-sekvensseihin aiheuttaen ei-hyväksyttäviä taustatasoja. Puhdistetut insertti-DNA:t leimattiin α-^{32}P-dCTP:llä käyttäen Nick Translation -pakkausta valmistajalta Promega Biotech valmistajan ohjeita noudattaen. Leimattu DNA erotettiin ei sisällytetyksi tulleesta leimasta spin-kolonnikromatografisesti (Maniatis et al., 1982). Kahdeksan 135 mm:m maljaa lambda-ZAP-kirjastoa maljattiin määränä 10^5 pfu/malja, ja suoritettiin plakkinostot nitroselluloosasuodattimille (Maniatis et al., 1982). Suodattimia paistettiin tyhjöuunissa 2 tunnin ajan 80 °C:ssa, minkä jälkeen suodattimia esihybridisoitiin 2 tunnin ajan, minkä jälkeen niitä hybridisoitiin 42 °C:ssa yön yli (kuten kuvataan alla Southern blot -analyysiä käsittelevässä osassa). Neljä suodatinta seulottiin M1-koettimella ja neljä B1A-koettimella. Suodattimet pestiin kahdesti 2 x SSC:ssä, joka sisälsi 0,5 % SDS:ää, kerran 1 x SSC:ssä, joka sisälsi 0,5 %

SDS:ää, ja kerran 0,5 x SSC:ssä, joka sisälsi 0,5 % SDS:ää, kaikki 50 °C:ssa, ja otettiin autoradiokuva. Mahdolliset positiiviset plakit noukittiin ja uudelleenseulottiin koettimilla. Korkeatiitterisiä fagivarastoja valmistettiin vahvistetuista positiivisista (jotka nimettiin M1AUS M1:een hybridisoituvalle kloonille ja AustB1 B1A:han hybridisoituvalle kloonille) ja kloonit siirrettiin pBluescript-vektoriin lambda-ZAP-valmistajan (Stratagene) ohjekirjan mukaan käyttäen BB4:ää E. coli -isäntäkantana. Plasmidi-DNA-miniprep-valmisteita saaduista jälkeläisistä valmistettiin alkalisella lyysillä (Maniatis et al., 1982) ja pilkottiin EcoRI:llä. Pilkkomistuotteet analysoitiin elektroforeettisesti agaroosigeelissä.

M1AUS-insertin sekventointi

Puhdistettu pBluescript-plasmidi-DNA DNA-sekventoitiin käyttäen valmistajan United States Biochemicals versio 2.0 Sequenase -pakkausta valmistajan ohjeiden mukaan. Ensimmäisille sekventointireaktioille alukkeita vektorisekvenssin päistä käytettiin reaktioiden pohjustamiseksi. Näistä reaktioista saatua sekventointitietoa käytettiin toisen alukeparin suunnittelemiseksi ja näillä toisilla alukkeilla saadulla tiedolla suunniteltiin kolmas pari. Tällä tavalla DNA sekventoitiin "kävelemällä pitkin" sekä 5'- että 3'-päästä.

AustB1-insertin sekventointi

Tämä suoritettiin käyttäen Sequenase 2.0 T7-polymeraasia (USB Biochemicals) kuten kuvattiin M1AUS-insertin sekventoinnille.

Polymeraasiketjureaktiot

cDNA:n valmistus

mRNA:ta (1 µg) 11 päivän ikäisistä tartutuksen jälkeisistä UK H. contortus -organismeista valmistettuna kuten kuvattiin täysikasvuisille UK-madoille sekoitettiin T17-adaptorialukkeeseen (5'GACTCGAGTCGACATCGATTTTTTTTTTT-TTTTTT3') dietyylipyrokarbonaatilla (DEPC) käsitellyssä

vedessä, minkä jälkeen seos kuumennettiin 65 °C:seen 5 minuutiksi ja sijoitettiin välittömästi jäihin. Metyylielohopeahydroksidia lisättiin loppupitoisuuteen 28,6 mM ja seosta inkuboitiin huoneenlämpötilassa 3 minuutin ajan. 2-merkaptoetanolia lisättiin loppupitoisuuteen 14,2 mM ja seos sijoitettiin jäihin. cDNA:n syntetisoimiseksi RNaasi-Guard-valmistetta (Pharmacia) lisättiin 1 yksikkö/µl, käänteiskopioijaentsyymipuskuria (Life Sciences) 1 kertaa pitoisuus, dATP:tä, dGTP:tä, dCTP:tä ja DTTP:tä kutakin pitoisuuteen 1 mM ja AMV-käähteiskopioijaentsyymiä (Life Sciences) pitoisuuteen 2 yksikköä/µl (kaikki esitettyinä loppupitoisuuksina). Reaktiota inkuboitiin 41 °C:ssa 75 minuutin ajan, minkä jälkeen seosta uutettiin fenoliliuoksella ja kloroformilla ja puhdistettiin spin-kolonnikromatografisesti (Maniatis et al., 1982). Puhdistettua reaktioseosta laimennettiin 2,5-kertaisesti ja se varastoitiin 4 °C:seen.

cDNA:n PCR-vahvistaminen käyttäen M1AUS-spesifisiä alukkeita

PCR-reaktiot suoritettiin käyttäen laitetta Programmable Thermal Cycler (M.J. Research Inc.). Reaktioseos sisälsi 1 µl:n edellä kuvatun mukaisesti valmistetusta 250 µl:sta laimennettua cDNA:ta, 25 pmol ensimmäisen säikeen T17-adaptorialuketta, 25 pmol toisen säikeen vahvistusaluketta [joko sitä, joka perustuu asemiin 865 - 884 (5'ACGGGTGTTCGGTTTCCGTAT3') tai sitä, joka perustuu asemiin 30 - 49 (5'GCTGAATCTAACTCCAATCC3') M1AUS-sekvenssistä (sekvenssi tunnusnro 5)], 1 x Taq-puskuria (Northumbria Biologicals Ltd) ja pitoisuuden 0,5 mM kutakin dATP:tä, dTTP:tä, dGTP:tä ja dCTP:tä 100 µl:n reaktiotilavuudessa ja peitettynä 40 µl:lla mineraaliöljyä haihtumisen estämiseksi. Tämä seos kuumennettiin sitten lämpökierrättimessä 95 °C:seen 2 minuutiksi, minkä jälkeen sitä pidettiin 72 °C:ssa. Seoksen ollessa 72 °C:ssa lisättiin 2 yksikköä Taq-polymeraasia (Northumbria Biologicals Ltd) ja sekoi-

tettiin varovaisesti reaktion muihin aineosiin. Sitten lämpökierrättimessä suoritettiin seuraava ohjelma:

vaihe 1: juottaminen 50 °C:ssa 5 minuutin ajan
vaihe 2: laajentaminen 72 °C:ssa 40 minuutin ajan
vaihe 3: denaturointi 94 °C:ssa 40 sekunnin ajan
vaihe 4: juottaminen 50 °C:ssa 2 minuutin ajan
vaihe 5: laajentaminen 72 °C:ssa 3 minuutin ajan
vaihe 6: 39 kierrosta vaiheita 3 - 5
vaihe 7: viimeinen laajentaminen 72 °C:ssa 15 minuutin ajan
vaihe 8: pidetään 4 °C:ssa.

Nämä olosuhteet saatiin julkaisusta Frohman et al. (1988).

PCR-tuotteiden kloonaus

PCR-tuotteet edeltävistä reaktioista erotettiin elektroforeettisesti agaroosigeelissä. Noin 2,5 ja 3,5 ke:n DNA-nauhoja elektroeluoitiin lasikuidulle (Whatman), uutettiin fenoliliuoksella ja puhdistettiin G50-kromatografisesti (Pharmacia) (Sambrook et al., 1989). Puhdistettu DNA ligatoitiin pT7Blue T -vektoriin (Novagene) noudattaen valmistajan ohjeita.

2,5 ke:n ja 3,5 ke:n PCR-tuotteiden sekventointi

DNA-sekventointi suoritettiin Sequenase 2.0 -pakkauksella (US Biochemicals) käyttäen "oligonukleotidikävely"-tekniikkaa, jota kuvataan osassa, joka käsittelee M1AUS:in sekventointia.

Polymeraasiketjureaktiot 5'-päille

Ensimmäisen säikeen cDNA:n valmistus

1 µg mRNA:ta 11 päivän ikäisistä tartutuksen jälkeisistä UK Haemonchus contortus -organismeista valmistettuna kuten kuvattiin täysikasvuisille madoille sekoitettiin pysyvään alukkeeseen (5'AAIGAAAGCGGATGGCTTGAIGC3'), joka oli suunniteltu säilyvästä alueesta AustB1:ssä, ja 2,5 ke:n ja 3,5 ke:n PCR-tuotteisiin (sekvenssit tunnusnrot 6, 7 ja vastaavasti 8). Seosta kuumennettiin 65 °C:ssa

5 minuutin ajan, se sijoitettiin jäihin ja metyylielohopeahydroksidia lisättiin loppupitoisuuteen 28,6 mM. Seosta inkuboitiin huoneenlämpötilassa 5 minuutin ajan, minkä jälkeen 2-merkaptoetanolia lisättiin loppupitoisuuteen 14,2 mM ja seos sijoitettiin jäihin. Ensimmäisen säikeen DNA valmistettiin käyttäen reagensseja 5' RACE -systeemistä (Gibco/BRL) loppupitoisuudessa 20 mM Tris/HCl, pH 8,4, 50 mM KCl, 2,5 mM $MgCl_2$, 100 µg/ml BSA:ta, pitoisuus 0,5 mM kutakin dATP:tä, dCTP:tä, dGTP:tä, dTTP:tä. 200 yksikköä Superscript-käänteiskopioijaentsyymiä lisättiin ja reaktiota inkuboitiin 42 °C:ssa 30 minuutin ajan, minkä jälkeen sitä kuumennettiin 55 °C:ssa 5 minuutin ajan. RN-aasi-H:ta lisättiin loppupitoisuuteen 100 yksikköä/ml ja reaktiota inkuboitiin 55 °C:ssa 10 minuutin ajan, minkä jälkeen se sijoitettiin jäihin. cDNA puhdistettiin käyttämällä se Glassmax-spin-kolonnissa (Gibco/BRL) ja se varastoitiin -20 °C:seen.

cDNA:n C-päädyttäminen

1/5:aa ensimmäisen säikeen cDNA:ta kuumennettiin 70 °C:ssa 5 minuutin ajan, minkä jälkeen sitä jäähdytettiin jäissä 1 minuutin ajan. Reagensseja 5' RACE -systeemistä (Gibco/BRL) lisättiin loppupitoisuuteen 10 mM Tris/HCl, pH 8,4, 25 mM KCl, 1,25 mM MgCl, 50 µg/ml BSA:ta, 0,2 mM dCTP:tä. 500 yksikköä/ml päätytransferaasia lisättiin ja reaktiota inkuboitiin 37 °C:ssa 10 minuutin ajan, minkä jälkeen sitä kuumennettiin 70 °C:ssa 15 minuutin ajan ja varastoitiin jäihin.

PCR-vahvistaminen käyttäen AustB1-, 2,5kePCR- ja 3,5kePCR-spesifisiä alukkeita

PCR-reaktiot suoritettiin ohjelmoitavassa Thermal Cycler -laitteessa (M.J. Research Inc.). 3'-päätä varten yhtä kolmesta alukkeesta käytettiin.

1. Aluke, joka on spesifinen 2,5 ke:n PCR-tuotteelle ja joka perustuu asemiin 374 - 394 (5'TGTTGTGGCTAATTTCGTCCA3').

2. Aluke, joka on spesifinen 3,5 ke:n tuotteelle ja joka perustuu asemiin 1210 - 1229 (5'CATCTTIAGTTATCTGACC-AG3').

3. Aluke, joka on spesifinen cDNA-kloonille AustB1 ja perustuu asemiin 357 - 377 (5'GACCATCGCTGATGAAGTCGG-3').

Reaktioiden 5'-päätä varten käytettiin yleistä "ankkurialuketta" (5'CUACUACUACUAGGCCACGCGTCGACTAGTACGGG-IIGGGIIGGGIIG3'). Kukin reaktioseos sisälsi 4 µl 50 µl:sta C-päädytettyä cDNA:ta, 25 pmol tarkoituksenmukaisia kahta aluketta, 1 x Taq-polymeraasipuskuria (Boehringer/Mannheim) ja 0,5 mM kutakin dATP:tä, dCTP:tä, dGTP:tä ja dTTP:tä 100 µl:n lopputilavuudessa. Tämä seos peitettiin 50 µl:lla mineraaliöljyä ja kuumennettiin 95 °C:ssa kierrättimessä 2 minuutin ajan. Reaktioseosta pidettiin 80 °C:ssa, kun 0,6 yksikköä Taq-polymeraasia lisättiin, minkä jälkeen suoritettiin seuraava ohjelma:

1. Juottaminen 50 °C:ssa 5 minuutin ajan.
2. Laajentaminen 72 °C:ssa 10 minuutin ajan.
3. Denaturointi 94 °C:ssa 45 sekunnin ajan.
4. Juottaminen 50 °C:ssa 1 minuutin ajan.
5. Laajentaminen 72 °C:ssa 2,5 minuutin ajan.
6. 39 kierrosta 3 - 5.
7. Laajentaminen 72 °C:ssa 15 minuutin ajan.
8. Pidetään 4 °C:ssa.

5'-PCR-tuotteiden kloonaaminen

PCR-tuotteet erotettiin elektroforeettisesti agaroosigeelissä ja odotettua kokoa olevat nauhat, noin 1,3 ke, leikattiin irti, DNA:ta puhdistettiin käyttäen Geneclean-pakkausta ja ligatoitiin PT7Blue T -vektoriin (Novagene) valmistajan ohjeiden mukaan.

Polymeraasiketjureaktio AustB1:n 3'-pään tuottamiseksi

Käytetty ensimmäisen säikeen cDNA oli se, jota kuvattiin cDNA:n tuottamiseksi käytettäväksi M1AUS-alukkei-

den kanssa. Spesifistä aluketta 1414 - 1435 (5'TCTTGAAGA-AATGAAAAAGCTT3') AustB1:ssä (sekvenssi tunnusnro 6) käytettiin T17-adaptorialukkeen kanssa, jota käytettiin M1AUS-PCR:ään, ja reaktiot suoritettiin lämpökierrättimessä (M.J. Research Inc). Reaktioseos käsitti 25 pmol kutakin aluketta, 2 µl cDNA:ta, 1 x Taq-polymeraasipuskuria (Boehringer Mannheim), 0,5 mM dATP:tä, dCTP:tä, dGTP:tä ja dTTP:tä 100 µl:ssa. Nämä peitettiin 50 µl:lla mineraaliöljyä ja kuumennettiin 95 °C:ssa 2 minuutin ajan. 0,6 µl Taq-polymeraasia lisättiin ja suoritettiin sama ohjelmasykli kuin edellä kuvatussa 5'PCR:ssä.

3'-tuotteiden kloonaus ja sekventointi

PCR-tuotteet erotettiin elektroforeettisesti agaroosigeelissä ja odotetun kokoinen, 1,3 ke:n, nauha leikattiin irti ja DNA puhdistettiin käyttäen Geneclean-pakkausta ja ligatoitiin PCR-Script-tuotteeseen (Stratagene) valmistajan ohjeiden mukaan.

Kloonattujen PCR-tuotteiden sekventointi

DNA PCR-klooneista sekventoitiin käyttäen Sequenase 2.0 -pakkausta (United States Biochemical) valmistajan ohjeiden mukaan. Oligonukleotidialukkeita käytettiin "kävelemiseksi pitkin" kloonien DNA:ta sekä 5'- että 3'-päistä.

Kaikkien DNA-sekvenssien analyysi

Sekvenssit analysoitiin käyttäen GCG (Genetics Computer Group) Sequence Analysis Software Package -pakkausta, Devereux et al., 1984.

Northern- ja Southern-blot-analyysit

Northern-blot-tuotteiden valmistus

Northern-blot-määritykset suoritettiin formaldehydigeeleissä oleellisesti kuten kuvaavat Maniatis et al., 1982. mRNA-näytteitä (11, 15 ja 23 päivän ikäisistä täysikasvuisista H. contortus -organismeista) käsiteltiin 17,5-%:isella v/v formaldehydiliuoksella ja 50-%:isella v/v formamidiliuoksella MOPS-puskurissa (20 mM 3-(N-morfo-

lino)propaanisulfonihappo, pH 7,0, 8 mM natriumasetaatti, 1 mM EDTA) 65 °C:ssa 15 minuutin ajan ja jäähdytettiin jäissä. Geeleille suoritettiin elektroforeesi MOPS-puskurissa ja suoritettiin blot-siirto Duralon-membraaneille kapillaarikuljetuksella kuten kuvataan julkaisussa Sambrook et al., 1989.

Southern-blot-tuotteiden valmistus

2 g täysikasvuisia Haemonchus contortus -organismeja, jotka oli pikajäädytetty nestetypessä, jauhettiin hienoksi jauheeksi nestetypessä. Jauhe lisättiin hitaasti 25 ml:aan lyysipuskuria (0,05 M Tris-HCl, pH 8, 0,1 % EDTA:ta, 1 % w/v sarkosyyliä, 0,05 mg/ml proteinaasi-K:ta (Boehringer Mannheim)) ja inkuboitiin 2 tunnin ajan 65 °C:ssa. Suspensiota uutettiin sitten kahdesti yhdellä tilavuudella fenolia + kloroformia, kahdesti 2 tilavuudella kloroformia, minkä jälkeen saostettiin etanolilla. Saostettu genomi-DNA uudelleensuspendoitiin 20 ml:aan Tris, EDTA -puskuria (TE, pH 8) yön aikana 4 °C:ssa keinupöydällä, minkä jälkeen dialysoitiin vastaan kahta vaihtoa 1 litra TE:tä. RNA poistettiin inkuboimalla DN-aasi-vapaalla RN-aasi-A-tyypillä 1 (Sigma) loppupitoisuutena 20 µg/ml 37 °C:ssa 1 tunnin ajan, minkä jälkeen uutettiin kerran fenoli-kloroformilla, kerran kloroformilla ja saostettiin etanolilla kuten edellä. Saostettu genomi-DNA-pelletti pestiin kahdesti 70-%:isella v/v etanoliliuoksella ja uudelleensuspendoitiin 1 ml:aan TE:tä kuten edellä.

Genomi-DNA:ta pilkottiin EcoRI:llä tai HindIII:lla (25 µg DNA:ta kummassakin pilkkomisessa) yön yli 37 °C:ssa, minkä jälkeen suoritettiin elektroforeesi määrällä 5 µg/-vyöhyke 1-%:isessa w/v agaroosigeelissä Tris-asetaattipuskurissa. Geelille suoritettiin Southern blot -siirto kapillaarikuljetuksella kuten kuvaavat Maniatis et al., 1982, Hybond-N-membraanille (Amersham International). DNA kiinnitettiin membraaniin käyttäen ultraviolettivaloa valmistajan suosituksia noudattaen.

Koettimien valmistus

pBluescript-plasmideja, jotka sisälsivät M1AUS-, B1A- tai AustB1-inserttejä, pilkottiin EcoRI:llä. pT7Blue-plasmideja, jotka sisälsivät 3,5kePCR-tuoteinserttejä, pilkottiin BamHI:llä ja niitä, jotka sisälsivät 2,5kePCR-tuoteinserttejä, pilkottiin BamHI:llä ja XbaI:llä. Pilkkomistuotteilla suoritettiin elektroforeesi, insertit otettiin talteen ja leimattiin radioaktiivisesti α-^{32}P-dCTP:llä nick-translaatiolla kuten kuvattiin edellä lambda-ZAP-kirjaston seulomista koskevassa kohdassa.

Hybridisaatio-olosuhteet

Southern-blot-määrityksille

Membraanit leikattiin liuskoiksi ja esihybridisoitiin hybridisaatiopuskurissa kuten kuvattiin aiemmin 3 tunnin ajan 28 °C:ssa. Genomi-DNA-Southern-blot-liuskoja hybridisoitiin kuhunkin edeltävään alukkeeseen yön yli 28 °C:ssa, pestiin kahdesti huoneenlämpötilassa (24 °C) ja sitten kahdesti 42 °C:ssa 2 x SSC:ssä, joka sisälsi 0,1 % w/v SDS:ää (kohtuullinen ankaruus) ja otettiin autoradiokuva. Autoradiografien kehityksen jälkeen liuskat pestiin jälleen korkean ankaruuden olosuhteissa (0,1 x SSC, 0,1 % w/v SDS:ää 65 °C:ssa) ja otettiin uusi autoradiokuva.

Northern-blot-määrityksille

Koettimille M1, M1AUS ja B1A (sekvenssit tunnusnrot 1, 5 ja vastaavasti 2)

Northern-blot mRNA:lle 11, 15 ja 23 päivän ikäisistä Haemonchus contortus -organismeista seulottiin ensin M1-insertillä. Suodatinta esihybridisoitiin 2 tunnin ajan 42 °C:ssa 2 x SSC:ssä (jossa 20 x SSC = 3 M NaCl, 0,3 M natriumsitraatti, pH 7,2), joka sisälsi 5 x Denhardtin liuosta (0,1 % w/v Ficoll 400 -valmistetta (Pharmacia), 0,1 % polyvinyylipyrrolidonia, 0,1 % nautaseerumialbumiinia fraktio V (Sigma Chemical Corp)), 0,5 % SDS:ää (natriumdodesyylisulfaatti), 10 % dekstraanisulfaattia, 0,1 mg/ml lohenkives-DNA:ta ja 50 % deionisoitua formami-

dia. Hybridisaatio koettimeen suoritettiin samassa puskurissa yön yli 42 °C:ssa. Suodattimia pestiin kahdesti 30 minuutin ajan 2 x SSC:ssä, joka sisälsi 0,5 % SDS:ää ja 50 % formamidia, kahdesti 30 minuutin ajan 2 x SSC:ssä, joka sisälsi 0,5 % SDS:ää, ja kahdesti 30 minuutin ajan 2 x SSC:ssä. Ensimmäinen pesu tapahtui 42 °C:ssa ja kaikki loput pesut suoritettiin 37 °C:ssa. Autoradiokuvanoton jälkeen blot riisuttiin pesemällä kiehuvassa 0,1-%:isessa SDS:ssä ja otettiin uusi autoradiokuva koettimen poistumisen varmistamiseksi. Samaa blot-kopiota seulottiin sitten M1AUS-insertillä, pestiin ja otettiin autoradiokuva. Blot riisuttiin jälleen ja tarkastettiin ja sitten, kun se oli puhdas, seulottiin B1A-insertillä. Jälleen riisumisen jälkeen blot-kopio seulottiin edelleen kuten alla kuvataan.

Koettimille AustB1, 2,5 ke - ja 3,5 ke -PCR-tuotteet (sekvenssit tunnusnrot 6, 7 ja 8)

Northern-blot hybridisoitiin AustB1-inserttiin käyttäen kohtalaisen ankaruuden olosuhteita Southern-blot-hybridisaatiolle kuvatun mukaisesti. Autoradiokuvan oton jälkeen blot riisuttiin keittämällä 0,1-%:isessa SDS:ssä membraanivalmistajan (Amersham) ohjeita noudattaen, minkä jälkeen se seulottiin 2,5 ke:n PCR-insertillä (klooni 2), riisuttiin toistamiseen ja seulottiin 3,5 ke:n PCR-insertillä (klooni 2).

H11OD:n pilkkominen ja entsyymiaktiivisuuden määritykset

H11OD:n valmistus

Natiivia H11OD:tä (H11ODN) valmistettiin menetelmien mukaan, joita kuvataan julkaisuissa WO 88/00 835 ja WO 90/11 086.

H11OD:n elastaasifragmentin (H11S) valmistaminen

Täysikasvuisia Haemonchus-organismeja homogenisoitiin 10 tilavuudessa jääkylmää PBS:ää, joka sisälsi 0,02 % natriumatsidia, minkä jälkeen sentrifugoitiin 20 minuutin ajan kierrosnopeudella 13 000 kierr./min. Pelletti uudel-

leensuspendoitiin 10 tilavuuteen PBS/atsidia, uudelleenhomogenisoitiin ja toistettiin sentrifugointi. Pellettimateriaali uudelleensuspendoitiin 50 mM MOPS-puskuriin, pH 7,4 (suspendointitilavuus on 1 ml kohden kutakin 0,17 g matoja) ja suoritettiin 30 minuutin esilämmitys 37 °C:ssa, minkä jälkeen sitä pilkottiin elastaasilla (800 µl/20 ml suspensiota; 1 mg/ml tuoretta varastoliuosta, johon on suoritettu lisäykset pitoisuuksiin 1 mM HCl/2 mM Ca^{2+}) 1 tunnin ajan. Pilkkominen keskeytettiin lisäämällä 3,4-dikloori-isokumariinia (300 µl 10 mM varastoliuosta DMSO:ssa/20 ml pilkkomistuotetta). Seosta sentrifugoitiin kierrosnopeudella 13 000 kierr./min 20 minuutin ajan ja pelletoitu materiaali otettiin talteen. Supernatanttia ultrasentrifugoitiin 100 000 x g:ssä 1 tunnin 20 minuutin ajan. Saatu supernatanttineste panostettiin ConA-kolonniin ja saatiin sitoutumisfraktiot. Analyysiä varten tämä fraktio ajettiin SDS-polyakryyliamidigeelissä ja siirrettiin elektroforeettisesti polyvinylideenidifluoridimembraanille (Immobilon-P, Millipore), värjättiin kevyesti Coomassie-sinisellä, leikattiin irti 105 kd:n nauha ja analysoitiin se kaasufaasiaminohapposekventoijassa. Rokotustutkimuksia varten ConA:han sitoutuvia fraktioita puhdistettiin edelleen konsentroimalla ja käyttämällä Superose 12 - (Pharmacia) geelisuodatuskolonnissa keräten aminopeptidaasiaktiivisuutta sisältävät fraktiot.

H110D:n pilkkominen termolysiinillä

H110D-dublettia puhdistettiin elektroeluutiolla preparatiivisen mittakaavan 8-%:isesta SDS-polyakryyliamidigeelistä, jolloin saatiin H110DE:tä kuten kuvataan julkaisussa WO90/11086, ja dimeerisen muodon, joka on juuri yli 200 kd, elektroeluutiolla (josta saatiin uudelleenajettuna SDS-PAGE:ssa karakteristinen dubletti kohtaan 110 kd) H110DE:n saamiseksi. H110DE-liuokset konsentroitiin 200 µl:ksi, kalsiumkloridia lisättiin pitoisuuteen 5 mM ja seos lämmitettiin 37 °C:seen. Vastavalmistettua ter-

molysiiniliuosta, jonka pitoisuus oli 1 mg/ml (1 mM HCl:ssä, 2 mM $CaCl_2$) lisättiin suhteessa 0,1 µg termolysiiniä kohden µg H11ODE:tä. Seosta inkuboitiin 37 °C:ssa 120 minuutin ajan, minkä jälkeen reaktio keskeytettiin lisäämällä 5 µl 0,5 M EDTA-liuosta.

Proteiinifragmentit erotettiin elektroforeettisesti 15-%:isessa SDS-polyakryyliamidigeelissä ja siirrettiin elektroforeettisestipolyvinylideenidifluoridimembraanille (Immobilon-P, Millipore). Värjäyksen Coomassie-sinisellä jälkeen voimakkaimmat erilliset nauhat leikattiin irti ja analysoitiin kaasufaasiaminohapposekventoijassa.

Aminopeptidaasi-A-aktiivisuuden suhteen rikastuneen H11OD-fraktion (H11A) valmistaminen

Pelletoitu materiaali, joka saatiin elastaasikäsittelyn jälkeen sentrifugoimalla 17 000 x g:ssä 20 minuutin ajan (katso edellä), uudelleensuspendoitiin PBS:ään 4 °C:ssa ja uudelleenpelletoitiin sentrifugoimalla, minkä jälkeen se uudelleensuspendoitiin 1-%:iseen Tween-liuokseen PBS/atsidissa, ja jätettiin seisomaan (samalla sekoittaen) 1 tunniksi. Suspensiota sentrifugoitiin 17 000 g:ssä 20 minuutin ajan ja supernatantti poistettiin. Pellettiä uutettiin toistuvasti 1-%:isella Thesit-liuoksella PBS/atsidissa. Supernatantit sentrifugoinneista 17 000 x g:ssä 20 minuutin ajan yhdistettiin ja ultrasentrifugoitiin 1 tunnin 20 minuutin ajan 100 000 x g:ssä. Supernatantti panostettiin ConA-affiniteettikolonniin (Affigel, Biorad) ja sitoutunut materiaali eluoitiin ja fraktioitiin edelleen ioninvaihtokromatografisesti MonoQ-kolonnissa.

Entsyymiaktiivisuuksien määritykset

α-aminoasyylipeptidihydrolaasi- (mikrosomaalisia) aminopeptidaasiaktiivisuuksia H11OD-valmisteissa karakterisoitiin määrityksillä liuoksessa käyttäen L-leusiini-, metioniini-, fenyylialaniini-, α-glutamiinihappo- ja lysiini-p-nitroanilideja (pNA). Kaikki aminohappo-p-nitro-

anilidisubstraatit (α-glutamiinihappoa lukuun ottamatta) saatiin valmistajalta Sigma, Poole, Dorset, UK, α-glutamiinihappo oli Flukalta, Dorset, UK. Yksittäiset hydrofobiset (leusiini- tai fenyylialaniini-) ja varatut (α-glutamiinihappo-) aminohappo-pNA:t, joiden tiedetään olevan substraatteja nisäkkään aminopeptidaasi-M:lle (ApM) ja vastaavasti -A:lle (ApA), valittiin entsyymi-inhibiittorien ja seerumi-inhibition vaikutuksen H110D-aminopeptidaasiaktiivisuuksiin mittaamiseksi.

(a) Mikromaljamääritys

Mikro-ELISA-levy- (Dynatech Immulon 1, Virginia, USA) kuopat täytettiin kukin 250 µl:lla joko 50 mM Hepes- tai MOPS-liuosta, pH 7, plus 1 - 10 µl määritettävää fraktiota. Maljoja esi-inkuboitiin sitten 37 °C:ssa 10 minuutin ajan, minkä jälkeen lisättiin 10 µl 25 mM aminohappo-p-nitroanilidisubstraattia kuoppaa kohden. Ajankohdan 0 optinen tiheys (OD) 405 nm:ssä mitattiin sitten käyttäen ELISA-levylukijaa, minkä jälkeen maljoja inkuboitiin 37 °C:ssa 15 - 30 minuutin ajan. Viimeinen OD-lukema otettiin sitten kuten aiemmin ja laskettiin OD:n muutos minuutissa kohden milligrammaa proteiinia.

(b) Inhibiittoriherkkyys

Enstyymimääritykselle käytetty menetelmä oli kuten kuvattiin kohdassa (a) edellä lukuun ottamatta, että inhibiittorit lisättiin 250 µl:aan puskuria plus 10 µl 1 mg/ml ConA-H110D-liuosta ja esi-inkuboitiin 10 minuutin ajan 37 °C:ssa ennen substraattien lisäämistä (leusiini-pNA tai α-glutamiinihappo-pNA). Prosentuaalinen inhibitio laskettiin seuraavasti:

$$\frac{x - y}{x} \times 100 = \text{prosentuaalinen inhibitio}$$

jossa x = entsyymin delta-OD/min ilman lisättyä inhibiittoria ja y = entsyymin delta-OD/min inhibiittorin

läsnä ollessa. Yhdeksän yhdistettä, joilla oli erilaiset entsyymiluokkainhibitiot, testattiin erikseen: Amastatin, Bestatin (metalloproteaasi, aminopeptidaasi), 1,10-fenantroliini, EDTA (metalloproteaasi), fosforamidon (metalloproteaasi, termolysiini, kollagenaasi), Aprotinin (seriiniproteaasi), Pepstatin (asparagiinihappoproteaasi), PMSF (seriini- ja kysteiiniproteaasi) ja E64 (kysteiiniproteaasi. Amastatin, bestatin, 1,10-fenantroliini, PMSF ja pepstatin olivat Sigmalta, Dorset, UK. Kaikki muut inhibiittorit saatiin Boehringer-Mannheimilta. Kutakin inhibiittoria käytettiin kahtena pitoisuutena, jotka olivat yhtä kuin tai suurempia kuin Boehringer-Mannheimin suosittelemat maksimipitoisuudet.

(c) Entsyymin vastaseerumi-inhibition määritys

Määritys oli kuten kuvattiin kohdassa (a) edellä lukuun ottamatta, että valmistettiin kaksi mikro-ELISA-maljaa. Yhdessä maljoista 10 µl ConA-H11OD:tä (1 mg/ml) plus 10 µl vastaseerumia kustakin lampaasta/kuoppa esi-inkuboitiin 37 °C:ssa 15 minuutin ajan. Toisessa maljassa 250 µl Hepes/bikarbonaattipuskuria plus 10 µl joko fenyylialaniini-pNA- tai α-glutamiinihappo-pNA-substraattia/-kuoppa samoin esi-inkuboitiin 37 °C:ssa 15 minuutin ajan. ConA-H11OD-seerumiseokset lisättiin sitten puskurisubstraattiseoksiin toisessa kuopassa multipipetillä. Määritettiin delta-OD/min. Korrelaatiotarkoituksissa prosentuaalinen inhibitio laskettiin käyttäen kaavaa edellä kohdassa (b), jossa x merkitsee keskimääräistä delta-OD/-min-arvoa kuopille, jotka sisälsivät entsyymiä plus seerumia negatiivisesta verrokkilampaasta (rokotettu hevosenpernaferritiinillä), ja y = delta-OD/min kuopille, jotka sisälsivät entsyymiä plus seerumia yksittäisistä H11OD:llä rokotetuista lampaista. Korrelaatiotarkoituksissa (kuvio 18) prosentuaalinen suojaus laskettiin käyttäen kaavaa edellä kohdassa (b), jossa x = joko keskimääräinen ulosteen munatulos/g tai verrokkien keskimääräinen matokuorma,

ja y = joko ulostetulos/g kokeen aikana tai yksittäisten, H11OD:llä rokotettujen lampaiden matokuorma.

Enstyymiaktiivisuuden paikantaminen histokemiallisesti

Aminopeptidaasiaktiivisuus osoitettiin täysikasvuisten Haemonchus contortus -organismien 10 µm:n kryostaattiviipaleilla käyttäen L-leusiini-4-metoksi-ß-naftyyliamidia ja L-glutamiini-α-4-metoksi-ß-naftyyliamidia substraatteina Nachlaksen et al. (1957), Nakanen et al. (1974) ja Lojdan et al. (1980) menetelmien mukaan.

Yhdistelmä-DNA-H11OD:n ilmentäminen eukaryoottisessa bakulovirus-hyönteissolusysteemissä

Ilmentämisplasmidien rakentaminen

Edellä kuvatut 3,5k PCR-fragmentit muodostettiin vahvistamalla oligo-dT-adaptorin (joka sisälsi SalI-kohdan) ja oligon 872, joka edustaa emäsnumeroita 30 - 49 M1AUS-sekvenssissä, välillä, ja tuotteet kloonattiin pT7-Blue-vektoriin. Klooni pT73.5-2 on suuntautunut vektoripolyliittäjä-BamHI-kohtaan sen 5'-päässä ja HindIII-, XbaI- ja SalI-kohtiin 3'-päässä. Sekvenssi 5'-päässä on:

```
BamHI      *!-----oligo 872-----------!--3.5K-----
5' GG ATC CGA TTG CTG AAT CTA ACT CCA ATC C .......3'
                Leu Asn Leu Thr Pro Ile .........
```

Tähti osoittaa 3'-dT/dA-ylijäämäpään, jota käytettiin kloonaukseen pT7 Blue -vektoriin.

3,5 PCR -klooni 2 pilkottiin HindIII:lla (yksittäinen kohta vektorin polyliittäjäsekvenssissä 3,5K -geenin 3'-päässä) ja päät täytettiin deoksinukleotideillä käyttäen DNA-polymeraasi-I:tä (Klenow-fragmentti Maniatiksen et al. (1982) mukaan. BamHI-liittäjä (5'CGGATCCG3', New England Biolabs kataloginro 1021) ligatoitiin tylppiin päihin, ja kloonit, joissa oli ylimääräinen BamHI-kohta, joka mahdollisti täysimittaisen H11OD-geenisekvenssin leikkaamisen BamHI-fragmenttina, valittiin. BamHI-frag-

mentti eristettiin elektroforeettisesti 0,6-%:isessa w/v agaroosigeelissä tris-asetaattipuskurissa, mitä seurasi puhdistaminen käyttäen Geneclean-tuotetta (BIO101); tämä menettely suoritettiin kahdesti. Sitten puhdistettu fragmentti ligatoitiin BamHI:llä leikattuun, fostaasikäsiteltyyn pBlueBac-II:een (Invitrogen Corp.), ja fragmentin oikein suuntautuneena (eli siten, että 3,5 PCR -kloonin 2 5'-pää on sijoitettu bakuloviruspolyhedriinipromoottorin kontrolliin) käsittävät kloonit määritettiin pilkkomalla NheI:llä ka XbaI:llä. Saatua plasmidia pilkottiin osittain BamHI:llä ja päät täytettiin kuten edellä kuvattiin. NcoI-liittäjä, joka sisälsi ATG:n (5'CCCATGGG3'; New England Biolabs kataloginro 1040), lisättiin ja seos ligatoitiin. Kloonit, joissa oli liittäjä ligatoituneena 3,5 PCR -kloonin 2 BamHI-kohdan 5'-päähän paremminkin kuin 3'-kohtaan, määritettiin pilkkomalla NcoI:llä. Saatu plasmidi, joka nimettiin pBB3.5-2(N), esitetään diagrammina kuviossa 19.

Tämä rakenne johtaa lukualueella olevan ATG:n insertointiin 3,5 PCR -klooni 2-insertin 5'-päähän translaation käynnistämiseksi. Tätä lähtö-ATG:tä ympäröivä sekvenssi on:

```
  BamHI--NcoI link-BamHI----*!-----oligo 872----------
  !--
5'GGATCCCC ATG GGG ATC CGA TTG CTG AAT CTA ACT CCA ATC
C..
           Met Gly Ile Arg Leu Leu Asn Leu Thr Pro Ile
...        ...
```

Ilmentyneestä proteiinista puuttuvat aminohapot 2 - 9 vastaavasta H11OD-sekvenssistä, ja siinä on 3 aminohappoa liittäjäsekvenssistä, joka välittömästi seuraa ATG:tä.

H11OD-sekvenssejä sisältävien yhdistelmä-DNA-bakulovirusten muodostaminen

Plasmidi pBB3.5-2(N) transfektoitiin Spodoptera frugiperda - (Sf9-) soluihin (joita voidaan saada yrityk-

seltä Invitrogen Corp) käyttäen lineaarista Autographica californica -tumapolyhedroosiviruksen (ACNPV) DNA:ta ja kationisia liposomeja (Invitrogen Corp. -transfektiomoduuli) noudattaen valmistajan ohjeita. Soluja viljeltiin TC-100-kasvualustassa (Sigma), jota oli täydennetty vasikansikiöseerumilla (CSL Ltd.; lämpöinaktivoitu 56 °C:ssa 45 minuutin ajan) ja antibiooteilla (penisilliini/streptomysiini, gentamysiini; CSL Ltd). Kontrollitransfektointi käyttäen pBB3.5-2(N)-plasmidia, jossa ATG on insertoitu 3,5 PCR -klooni 2-sekvenssin 3'-päähän, suoritettiin myös. Yhdistelmä-DNA-plakit valittiin sillä perusteella, että pBlueBac-II-vektori myös koodittaa E. coli ß-galaktosidaasia (ß-gal), sisällyttämällä X-gal:ia agaroosipintakerrokseen suositeltuna tasona. Valikoima sinisiä plakkeja noukittiin ja niille suoritettiin kaksi lisäkierrosta plakkipuhdistusta, minkä ajan jälkeen tartutetut yksikerrokset eivät osoittaneet mitään todisteita kontaminoivasta villityypin viruksesta (minkä todistaisi tumapolyhedran läsnäolo). Puhdistetut virukset nimettiin 3.5-2-P2A, -P3A ja -P4A, ja niitä vahvistettiin Sf9-solujen kahdella peräkkäisellä tartutuksella ennen käyttöä. Verrokkitransfektoinnista puhdistettu plakki nimettiin 3.5-2-rev.

H11OD-ilmentymisen määritys hyönteissoluissa, jotka on tartutettu yhdistelmä-DNA-bakuloviruksella

Sf9-solujen yksikerroksia (1 x 10^6 solua 25 ml:n pulloissa) tartutettiin 3.5-2-viruksilla, villityypin (wt) viruksella, ß-gal:in ilmentävällä verrokkiviruksella tai niitä ei tartutettu. Neljän päivän kasvatuksen 26 °C:ssa jälkeen yksikerrokset irrotettiin kevyesti ravistamalla, solut otettiin talteen sentrifugoimalla (2 000 kierr./min, 10 minuuttia) ja solupelletit rikottiin kolmella jäädytys-sulatuskierroksella. Lysaatit uudelleensuspendoitiin 500 µl:aan PBS:ää, ja 25 µl:n eristä määritettiin ApM-aktiivisuus mikrokuoppamäärityksellä.

15 µl:n erillä (3 x 10^4 soluekvivalenttia) edeltävistä lysaateista suoritettiin elektroforeesiajo denaturoivissa 7,5-%:isissa SDS-polyakryyliamidigeeleissä. Yksi geeli värjättiin sitten Coomassie-sinisellä ilmentymistasojen määrittämiseksi. Toisesta geelistä suoritettiin Western-blot-määritys, ja blot-kopiota seulottiin anti-H11ODN:llä (kuten on aiemmin kuvattu).

Tulokset

Immuunipositiivisten kloonien analyysi

Affiniteettipuhdistettujen vasta-aineiden analyysi kloonilla M1

Affiniteettipuhdistettuja vasta-aineita, jotka ovat spesifisiä kullekin viidestä vasta-ainepositiivisesta kloonista, valmistettiin ja niillä seulottiin H11OD-rikastetun uutteen Western-blot. Kuten esitetään kuviossa 8, kaikki 5 kloonia vaikuttivat tunnistavan H11OD-dubletin. Kuitenkin reaktio kloonin M1 kanssa antoi voimakkaimman signaalin (kuvio 8d) verrattuna lambda-gt11-negatiiviseen verrokki-blot-määritykseen (kuvio 8e). Tätä kloonia tutkittiin tämän vuoksi edelleen.

Northern blot -analyysi kloonilla M1

M1-insertillä seulotun Haemonchus contortus -mRNA:n Northern blot -analyysi esitetään kuviossa 9. Yksi mRNA-nauha tunnistettiin kohdassa noin 3,5 ke. Tämä on kooltaan riittävä koodittamaan noin 110 kd:n proteiinia.

Kloonin M1 sekvenssianalyysi

DNA:n restriktiopilkkomistuotteiden EcoRI:llä analyysi osoitti M1-insertin kooksi noin 300 ep. M1-fragmentin DNA-sekvenssi määritettiin (sekvenssi tunnusnro 10 ja esitetään kuviossa 3). Fragmentti käsittää 295 ep avoimen lukualueen alkaessa emäksestä nro 3.

Northern blot -analyysi kloonilla B1A

B1A-insertillä seulotun Haemonchus contortus -mRNA:n Northern blot -analyysi esitetään kuviossa 9c.

Kuten M1:lle, yksi mRNA-nauha todettiin kohdassa noin 3,5 ke.

Kloonien B1A ja B2 sekventointi

B1A-kloonit sekventoitiin (sekvenssit tunnusnrot 2 ja 3) ja täysimittainen sekvenssi (sekvenssi tunnusnro 2) esitetään rinnan H11-1:n (sekvenssi tunnusnro 19) ja AustB1:n (sekvenssi tunnusnro 6) kanssa kuviossa 5. Insertin koko on 484 ep ja sillä on täysi ORF ensimmäisestä emäksestä. Kolme B2-fragmenttia, jotka saadaan pilkkomisesta EcoRI:llä, sekventoitiin ja täydellinen sekvenssi B2:lle (sekvenssi tunnusnro 4) on 581 ep. Se esitetään rinnan H11-2:n kanssa kuviossa 4. Sekvenssissä on ORF asemissa 3 - 213 ep, jolloin stop-kodoni ja ei-luettu alue sopivat 2,5 ke -PCR-tuotesekvenssin (sekvenssi tunnusnro 7) vastaaviin.

M1:n ja B1A:n ilmentäminen *E. colissa*

GST-ilmentämisvektoriin alakloonaamalla saatiin klooneja, jotka ilmentsivät 38 - 40 kd:n M1:lle ja 45 kd:n B1A:lle fuusioproteiineja. Nämä sopivat näille inserteille ennustettuihin kokoihin ottaen huomioon glutationi-S-transferaasin molekyylipaino. Molemmat fuusioproteiinit reagoivat hyvin voimakkaasti Western-blot-määrityksissä H110DE-vastaisilla affiniteettipuhdistetuilla vasta-aineilla (kuvio 11). Fuusioproteiinit ilmentyivät ei-liukoisina inkluusiokappaleina.

Vasta-ainevasteet lampaissa, jotka on rokotettu M1-GST- ja B1A-GST-fuusioproteiineilla

Vasta-seerumia fuusioproteiineilla ruiskutetuista lampaista testattiin Western-blot-määrityksellä H110D-valmisteita vastaan. Sekä GST-M1 että GST-B1A tuottivat vasta-aineita, jotka spesifisesti tunnistivat H110D-dubletin (kuvio 12). Seerumit negatiivisista verrokkilampaista eivät tunnistaneet H110D-dublettia.

Kloonien eristäminen ja karakterisointi, jotka on valittu hybridisaatiolla M1- tai B1A-insertti-DNA:han

Vahvistettu positiivinen klooni, joka hybridisoituu M1-koettimeen, nimettiin M1AUS:ksi (sekvenssi tunnusnro 5) ja B1A:han hybridisoituva klooni nimettiin AustB1:ksi (sekvenssi tunnusnro 6). Puhdistettujen plasmidi-DNA-sekvenssien restriktiopilkkominen EcoRI:llä osoitti insertin kooksi noin 900 ep M1AUS:lle ja noin 1,6 ke AustB1:lle. Kuten esitetään kuvioissa 9b) ja 9d), Northern-blot-määrityksissä M1AUS ja AustB1 hybridisoituivat samankokoiseen mRNA:han (noin 3,5 ke) kuin M1 ja B1A.

M1AUS:in sekvenssianalyysi

M1AUS-fragmentin täydellinen sekventointi suoritettiin käyttäen synteettisiä oligonukleotidejä "kävelemiseksi" pitkin DNA:ta kummasta tahansa päästä. Saadun sekvenssin analyysi paljasti, että M1AUS-insertti oli 948 ep kuten esitetään kuviossa 3. Sekvenssi (sekvenssi tunnusnro 5) alkaa ATG:llä (joka koodittaa metioniinia) ja siinä on avoin lukualue (ORF) sen koko pituudella. Sekvenssi on 19 emäsparia pitempi kuin M1-sekvenssi 5'-päässä, ja 634 ep pitempi 3'-päässä. Näille kahdelle kloonille yhteinen sekvenssi (emäkset 20 - 314) olivat identtiset lukuun ottamatta kahta nukleotidieroa kolmannen kodonin asemassa. Kaikkien mahdollisten lukualueiden vertaaminen eri tietokantoihin osoitti, että lukualueella, joka alkaa ATG:llä emäksestä nro 1, on homologiaa mikrosomaalisten aminopeptidaasien ryhmän jäsenten kanssa.

AustB1:n sekvenssi

AustB1-fragmentin täydellinen sekventointi suoritettiin käyttäen synteettisiä oligonukleotidejä kävelemiseksi pitkin DNA:ta kummasta tahansa päästä. DNA-sekvenssi (sekvenssi tunnusnro 6) esitetään kuviossa 5. Klooni on 1 689 ep:n mittainen ja siinä on ORF tähteestä 2. Tämä sekvenssi muodostaa osan H11-1:tä kuten esitetään kuvios-

sa 1. Tämän sekvenssi aminohappotranslaatio osoitti aminopeptidaaseille tunnusomaisen sinkkiin sitoutuvan motiivin.

H11OD-mRNA-sekvenssien cDNA:n PCR-vahvistaminen käyttäen M1AUS-alukkeita

cDNA syntetisoitiin Haemonchus contortus -mRNA:sta käyttäen alukkeena adaptorisekvenssin sisältävää oligo-dT:tä DNA:n myöhemmän kloonauksen ja manipuloinnin helpottamiseksi. Tätä cDNA:ta käytettiin sitten M1AUS-sekvenssin vahvistamiseksi PCR:llä käyttäen 5'-pääalukkeena asemiin 865 - 885 perustuvaa synteettistä oligonukleotidia. Vahvistettiin noin 2,5 ke:n PCR-fragmentti. Tämä on suunnilleen odotuksia vastaavan kokoinen fragmentti perustuen mRNA:n tunnettuun kokoon ja nisäkäsaminopeptidaasi-cDNA-sekvensseihin.

Toinen sarja PCR-reaktioita suoritettiin käyttäen aluketta lähellä M1AUS:in 5'-päätä (emäkset 30 - 49). Agaroosigeelissä todettiin neljä nauhaa. Suurin näistä, kohdassa 3,5 ke, vastaa PCR-tuotteelle ennustettua kokoa.

2,5 ke:n ja 3,5 ke:n PCR-tuotteiden kloonaaminen ja sekventointi M1AUS-alukkeista

2,5 ke:n ja 3,5 ke:n PCR-tuotteet kloonattiin ja nimettiin 2.5PCR:ksi (sekvenssi tunnusnro 7) ja 3.5PCR:ksi (sekvenssit tunnusnrot 8, 14 ja 15 klooninumeroille 2, 10 ja vastaavasti 19). Northern blot -määrityksissä 2.5PCR ja 3.5PCR (klooni 2, 3.5PCR-2) hybridisoitui noin 3,5 ke:n mRNA:han (kuvio 9e, f) samalla kuviolla (suhteessa mRNA:n saamiseksi käytetyn Haemonchuksen ikään) kuin M1, B1A, M1AUS ja AustB1.

Kloonien täydellinen sekventointi suoritettiin "oligonukleotidikävelyllä". Kuten esitetään kuviossa 1, sekvenssi 2,5 ke:n tuotteelle (sekvenssi tunnusnro 7) on osa H11-2:sta (sekvenssi tunnusnro 20) ja sekvenssi 3,5 ke:n tuotteelle (sekvenssi tunnusnro 8) on valtaosa H11-3:sta (sekvenssi tunnusnro 21). Kummankin näiden sekvenssin aminohappotranslaatiot (esitetään kuviossa 6) si-

sältävät sinkkiin sitoutuvan motiivin His Glu Xaa Xaa His Xaa Trp (HEXXHXW), joka on karakteristinen mikrosomaalisille aminopeptidaaseille.

5'-pää-PCR-kloonien sekventointi

cDNA syntetisoitiin käyttäen aluketta, joka sopi säilyvään sekvenssiin cDNA-kloonissa AustB1, 2.5PCR ja 3.5PCR (sekvenssit tunnusnrot 6, 7 ja 8), joka hybridisoituu mRNA:han näille sekvensseille noin 1,3 ke 5'-päästä. cDNA-sekvenssit C-päädytettiin 5'-päässä, minkä jälkeen suoritettiin PCR-reaktiot käyttäen yleistä ankkuri- (A) aluketta 5'-päälle ja kolmea kullekin sekvensseistä AUSTB1, 2.5PCR ja 3.5PCR-klooni 2 spesifistä aluketta (sekvenssit tunnusnrot 6, 7 ja 8) 3'-päälle. Kustakin reaktiosta saatiin ennustetun kokoinen tuote, juuri alle 1,3 ke: 1 301 ep (sekvenssi tunnusnro 9), 1 280 ep (sekvenssi tunnusnro 10) ja vastaavasti 1 292 (sekvenssi tunnusnro 11). Kaikille kolmelle sekvenssille saatiin ei-luettu alue 5'-päähän (kuvio 2). Kaikki alkavat samalla 22 ep:n sekvenssillä (5'GGTTTAATTACCCAAGTTTGAG3'), joka tunnetaan silmukoituneena johtosekvenssinä 1 (SL1) ja esiintyy ei-luetulla 5'-alueella laajassa valikoimassa sukkulamatoja, Huang et al., 1990. Sekvensseissä tunnusnrot 9 ja 10 SL1-sekvenssi on välittömästi ennen alku-ATG:tä. Sekvenssissä nro 11 esiintyy 13 ep SL1:n ja alku-ATG:n välissä. Kaikille kolmelle sekvenssille saatiin täydelliset ORF:t.

AustB1-3'-pää-PCR-kloonin sekventointi

Käyttäen spesifistä aluketta, joka sopi asemiin 1 414 - 1 438 AustB1:ssä (sekvenssi tunnusnro 6), PCRtuotteelle saatiin ennustusten mukaan noin 1,3 ke:n nauha. Sekventoimalla kloonattu nauha saatiin sekvenssit tunnusnrot 12 ja 13. Niille saatiin 1 - 615 ep:n ORF ja huomattava ei-luettu alue.

Kloonattujen PCR-tuotteiden sekvenssianalyysi

Edellä kuvattujen sekvenssien komposiitit nimeltä H11-1, H11-2 ja H11-3 esitetään kuviossa 2. Näistä ennustetut aminohapposekvenssit esitetään kuviossa 6a. Esitettyjen DNA-sekvenssien ennustettujen translaatioiden oikeellisuuden vahvistaa oleellisesti sopivuudet aminohapposekvensseihin määritettyinä Edman-pilkkomisella CNBr- ja proteolyyttisen pilkkomisen fragmenteista (kuvio 7). Täten 27 tähdettä H11S:n 29 tähteen N-pääsekvenssistä (sekvenssi tunnusnro 16) sopii H11-2:een tähteissä 61 - 90 (kuvio 7b). Valiinin (V) asemassa 78 ja glysiinin (G) asemassa 90 sopivuudet ovat karakteristisia H11-2:lle, koska H11-3:ssa on asparagiini (N) asemassa 90 ja H11-1:ssä on leusiini (L) asemassa 86 (joka vastaa asemaa 78 H11-2:ssa). Kaksi tähdettä H11S:n N-pääaminohapposekvenssistä (sekvenssi tunnusnro 16) ei sovi mihinkään kolmesta tässä esitetystä H11OD-sekvenssistä. Erityisesti huomion arvoisia ovat hyvin samankaltaisten, mutta erillisten sekvenssien Pep A ja B (aiemmin kuvattu julkaisussa WO 90/11 086) tarkat sopivuudet H11-2:een ja vastaavasti H11-1:een tähteiden 540 - 555 alueella. Vastaavasti 450 - 470 tähdealueella Pep D on tarkka sovitus H11-2:lle, kun taas samankaltainen, mutta erillinen Pep E sopii paremmin H11-3:een.

Esimerkkinä yhden täysimittaisista sekvensseistä (H11-3) luettua aminohapposekvenssiä verrataan kuviossa 6b kahteen sekvenssiin nisäkkään mikrosomaalisille aminopeptidaaseille. H11OD-translaation homologia näihin aminopeptidaaseihin nähden on esitetty sijoittamalla identtiset aminohapot laatikoihin. Mikrosomaalisten aminopeptidaasien karakteristinen motiivi on aminohapposekvenssi HEXXHXW, joka toimii sinkkiä sitovana kohtana (Jongeneel et al., 1989; Vallee et al., 1990); tämä esitetään tähdillä kuviossa 6. Tämä sekvenssi, jonka osoitetaan esiintyvän H11-1:n, H11-2:n ja H11-3:n translaatioissa, on säilyvä kaikissa mikrosomaalisissa aminopeptidaaseissa. Muita ni-

säkkään ja Haemonchuksen mikrosomaalisille aminopeptidaaseille yhteisiä piirteitä ovat verraten lyhyen solunsisäisen alueen esiintyminen, yksittäinen transmembraaninen sekvenssi N-pään vieressä ja useat mahdolliset glykosylaatiokohdat. H11-1:n, -2:n ja -3:n homologiatasot (52 - 55 %:n samankaltaisuudet ja 30 - 32 %:n identtisyydet) nisäkkään mikrosomaalisiin aminopeptidaaseihin nähden esitetään taulukossa 2.

Southern blot -analyysit

Tulokset H. contortuksen genomi-DNA:n Southern-blot-määrityksistä seulottuina eri H11OD-cDNA-klooneilla ja PCR-tuotteilla esitetään kuviossa 10. Kaikki koettimet osoittavat monia hybridisaationauhoja; tämä on tyypillistä multigeeniryhmälle. Odotuksia vastaten B1A ja AustB1 osoittivat toisiinsa nähden samankaltaisia hybridisaatiokuvioita kuten myös M1AUS ja 3.5kePCR. Kuitenkin nämä kuviot poikkesivat selvästi toisistaan ja 2.5kePCR-koettimella havaitusta jopa kohtalaisen ankaruuden olosuhteissa (kuvio 10A), mikä heijastaa vaihtelevia homologiatasoja näiden kolmen cDNA:n välillä.

H11OD:hen liittyvien aminopeptidaasiaktiivisuuksien osoittaminen

Mikrosomaalisen aminopeptidaasiaktiivisuuden todettiin liittyvän H11OD:tä sisältäviin fraktioihin, eli supernatantteihin ultrasentrifugoinnista Thesit-uutteilla, ConA:han sitoutuvalla fraktiolla (ConA H11OD) ja MonoQ-kolonnissa ioninvaihtokromatografia suorittamalla saaduilla H11OD:tä sisältävillä fraktioilla (taulukko 3). Spesifiset aktiivisuudet kaikilla testatuilla substraateilla kasvoivat H11OD:n puhtauden lisääntyessä.

Taulukko 3

Fraktioiden tyypillisestä H11OD-valmistuksesta entsyymiaktiivisuudet

Fraktio	Spesifiset entsyymiaktiivisuudet (OD/minuutti/mg proteiinia)				
	Fenyyli-alaniini-pNA	Leusiini-pNA	Lysiini-pNA	α-Glutamiini-pNA	Metioniini-pNA
fosfaatti-puskuroitu suolaliuos (PBS)	0,20	0,08	0,12	0,01	0,16
1-%:inen Tween 20/PBS	0,15	0,15	0,09	0,01	0,09
1-%:inen Thesit/PBS	1,94	1,25	0,57	0,54	1,91
ConA H11OD	4,08	3,01	1,43	2,09	3,84
CamQ H11OD	6,55	5,01	3,01	3,90	6,41

Nisäkäsaminopeptidaasien inhibiittorien vaikutukset H11OD-aminopeptidaasiaktiivisuuksiin

Inhibiittorin bestatin (joka inhiboi nisäkkään mikrosomaalista aminopeptidaasia) lisääminen ConA H11OD:hen valmistajan (Boehringer Mannheim) suosittelemana pitoisuutena laski aktiivisuutta leusiini-p-nitroanilidia vastaan noin 70 %. Suoritettiin sarja kokeita aktiivisuuden inhibition testaamiseksi valikoimalla proteaasi-inhibiittoreita. Niillä inhibiittoreilla, jotka eivät olleet spesifisiä metalloproteaaseille tai aminopeptidaaseille, ei ollut inhibitiovaikutuksia reaktionopeuteen. Inhibiittoreilla, joiden tiedetään vaikuttavan metalloproteaaseihin tai aminopeptidaaseihin, oli vaikutus reaktionopeuksiin kuten esitetään taulukossa 4. Tehokkain inhibiittori oli 5 mM fenantroliini.

Taulukko 4

H110D-aminopeptidaasiaktiivisuuksien inhibointi käyttäen erilaisia proteaasi-inhibiittoreita

Inhibiittori	Pitoisuus	Inhibitio-% α-glutamiinihappo-p-nitroanilidi-substraatti[1]	Inhibitio-% Leusiini-p-nitroanilidi-substraatti[2]
Amastatin	10 µM	0	61
	50 µM	0	66
Bestatin	50 µM	23	63
	100 µM	35	68
EDTA	1 mM	17	8
	10 mM	21	16
1,10-fenantroliini	1 mM	78	78
	10 mM	85	87
fosforamidon	10 µM	1	0,8
	50 µM	1	7

[1] Nisäkäsaminopeptidaasi-A-substraatti

[2] Nisäkäsaminopeptidaasi-M-substraatti

H11OD:n subfraktiointi

Aktiivisuuksien jakauma, jotka liittyvät fraktioihin ConAH11OD:n ioninvaihtokromatografiasta MonoQ-kolonnissa, esitetään kuviossa 13a, ja fraktioiden SDS-PAGE kuviossa 13b.

Edelleen entsymaattista aktiivisuutta liittyi H11OD:n alafraktioihin, jotka erotettiin uudelleenkierrätetyllä vapaan virtauksen isoelektrisellä fokusoinnilla (kuviot 14 ja 15). Alhaisemmilla pI-arvoilla (pH 4,5) nämä alafraktiot sisältävät vain suuremman nauhoista, jotka muodostavat SDS-PAGE:ssa nähdyn H11OD-dubletin, ja korkeammilla arvoilla (pH 6,5) ne sisältävät vain pienemmän nauhoista, jotka muodostavat H11OD-dubletin. Välifraktiot sisältävät molemmat nämä nauhat. Pienempi nauha voidaan myös saada erillisenä fraktiona ioninvaihtokromatografisesti MonoQ-kolonnissa käyttäen Pharmacian SMART^{R}-laitteistoa. Kaikki nämä alafraktiot sitovat lammasvasta-aineita H11OD:lle, jota on affiniteettipuhdistettu lambda-gt11-kloonin M1 ilmentämällä proteiinilla, kun taas vasta-aineet, jotka on eluoitu lambda-gt11:stä, joka ei sisällä inserttiä, eivät sitoudu (kuvio 14c). Kaikki alafraktiot sitovat hiiren monoklonaalisia vasta-aineita nimeltä TS 3/19.7 (kuvio 14c), jotka myös sitoutuvat kloonin M1 ilmentämään yhdistelmä-DNA-proteiiniin. Kaikki alafraktiot osoittavat mikrosomaalista aminopeptidaasiaktiivisuutta (kuvio 15c), vaikka tämä aktiivisuus on suhteellisen alhaista fraktioissa, jotka on saatu korkeimmilla ja matalimmilla pI-arvoilla. Tämän alhaisemman aktiivisuuden voidaan katsoa johtuvan laskeneista proteiinipitoisuuksista, ääri-pH-arvojen vaikutuksista subfraktioinnin aikana tai edellytyksestä, että maksimaalista aktiivisuutta varten läsnä on sekä suurempia että pienempiä nauhoja.

Rokottaminen H11OD:n erotetuilla komponenteilla

Erotetut ylemmät ja alemmat nauhat, jotka saadaan vapaan virtauksen isoelektrisellä fokusoinnilla tai ionin-

vaihtokromatografisesti, indusoivat suojaavien vasta-aineiden muodostuksen ruiskutettuina lampaisiin kuten esimerkkinä esitetään seuraavassa kokeessa. 30 noin 6 kuukauden ikäistä lammasta jaettiin viiteen kuuden ryhmään siten, että kukin ryhmä on sovitettu samalle lampaiden painoalueelle. Kuhunkin eläimeen ruiskutettiin kaikkiaan 150 µg proteiinia annettuna kolmena yhtä suurena annoksena, kuten kuvataan julkaisussa Munn et al. (1992) ja Tavernor et al. (1992a, b) 54 päivän aikajakson aikana. Eläimiin ryhmässä L ruiskutettiin H110-dubletin alempaa (pienempää) nauhaa, niihin, jotka olivat ryhmässä U, ylempää nauhaa, U+L:ään yhdistettyjä ylempiä ja alempia nauhoja, D:hen kahta (ei erotettua) nauhaa, jotka saatiin vapaan virtauksen isoelektrisellä fokusoinnilla väli-pH-arvoilla, ja verrokkeihin (ryhmä C)) hevosenpernaferritiiniä (antigeenisesti ei sukua oleva proteiini). Lampaille annettiin haasteena 10 000 tartutuskykyistä toukkaa kolmen viikon kuluttua 3. ruiskeesta ja koe päätettiin 29 - 31 päivän kuluttua tartutuksesta. Kokeen tuloksesta esitetään yhteenveto kuviossa 16. Minkä tahansa alafraktioista ruiskuttaminen laski loismunatulosta kautta kokeen noin 90 %:lla ja laski kokonaismatolukuja 63 - 84 %:lla, jotka kaikki osoittivat merkitsevää eroa ($p<0,05$) verrokkeihin käytettäessä ei-parametrisiä tilastollisia analyysejä. Laskut (70 - 88 %) naaraspuolisten matojen lukumäärissä olivat suurempia kuin laskut urospuolisten matojen lukumäärissä ja (lukuun ottamatta laskua urospuolisten matojen lukumäärissä lampaissa, joihin oli ruiskutettu yhdistettyjä ylempiä ja alempia nauhoja, jolloin $p<0,07$) kummallekin sukupuolelle laskut olivat merkitseviä ($p<0,05$).

Rokottaminen H11S:llä ja H11A:lla

Typistetty, vesiliukoinen muoto H11OD:tä (H11S; jolla on tallella entsyymiaktiivisuutensa) voidaan saada natiivista molekyylistä käsittelemällä elastaasilla. Tämän

muodon todettiin sisältävän pääasiassa ApM:n kaltaista entsyymiaktiivisuutta ja elastaasilla pilkotun pelletin Thesit-uute (H11A) oli rikastunut ApA:n kaltaisen aktiivisuuden suhteen (katso taulukko 5).

Taulukko 5

	Suhde
Aminopeptidaasi-M (leusiini-pNA)	**: aminopeptidaasi-A (α-glutamiinihappo-pNA)**
H11OD	1,44:1
H11S	26,0:1
H11A	0,48:1

Seuraava koe osoittaa, että lampaiden rokottaminen joko H11S:llä tai H11A:lla antaa suojaa vastaan Haemonchus-haastetta tai -tartuntaa. 18 lammasta, jotka olivat noin 8 kuukauden ikäisiä, jaettiin kolmeksi kuuden ryhmäksi siten, että kussakin ryhmässä lampaiden painoalue oli sama. Kuhunkin eläimeen ruiskutettiin kaikkiaan 100 µg proteiinia annettuna 2 samansuuruisena annoksena, kuten kuvataan julkaisussa Munn et al. (1992) ja Tavernor et al. (1992a, b) 23 päivän aikajakson aikana. Eläimiin ryhmässä A ruiskutettiin H11A:ta, niihin, jotka olivat ryhmässä S, H11S:ää ja niihin, jotka olivat ryhmässä C, hevosenperna-ferritiiniä (antigeenisesti ei sukua oleva proteiini) negatiivisena verrokkina. Sitten lampaille annettiin haasteena 10 000 tartutuskykyistä toukkaa 25 päivän kuluttua toisesta ruiskeesta ja koe päätettiin 34 - 36 päivän kuluttua tartutuksesta. Kokeen tuloksesta esitetään yhteenveto kuviossa 17. H11S:n ruiskuttaminen laski loismunatulosta kautta kokeen 89 %:lla ja laski kokonaismatolukuja 76 %:lla. H11A:n ruiskuttaminen laski loismunatulosta kautta kokeen 98 %:lla ja laski kokonaismatolukuja 84 %:lla. Nämä osoittivat merkitsevää eroa ($p<0,05$) verrokeihin käytettäessä ei-parametrisiä tilastollisia analyysejä.

H110D-aminopeptidaasiaktiivisuuksien inhibointi vasta-aineilla

H110D:tä sisältäviä liuoksia inkuboitiin seerumilla yksittäisistä lampaista, joihin oli ruiskutettu H110D:tä sisältäviä fraktioita, tai verokkilampaista. Liuoksista määritettiin sitten aminopeptidaasiaktiivisuudet käyttäen fenyylialaniini- ja α-glutamiinihappo-pNA-yhdisteitä substraatteina. Aktiivisuuden inhibition aste (maksimissaan 80 %) korreloi suojaustasoon, jota osoittivat yksittäiset lampaat, joista seerumit oli saatu (katso kuvio 18).

Entsyymiaktiivisuuden paikantaminen

Jäädytettyjä viipaleita täysikasvuisista Haemonchus contortus -organismeista tarkasteltiin aminopeptidaasiaktiivisuuden suhteen. Kuten esitetään kuviossa 19, aminopeptidaasientsyymiaktiivisuudet sijaitsevat suolen luminaalipinnalla. H110D-proteiinia todetaan myös spesifisesti tässä paikassa.

H110D:n (3.5PCR -klooni 2) ilmentäminen käyttäen eukaryoottista bakulovirus-hyönteissolusysteemiä

Aminopeptidaasiaktiivisuuden ilmentäminen hyönteissoluissa

Tartutettuja soluja otettiin talteen ja niistä määritettiin aminopeptidaasiaktiivisuudet käyttäen phe-, leu-, met- ja lys-pNA-yhdisteitä substraatteina. Määritystä mutkisti se havainto, että hyönteissoluilla on aminopeptidaasiaktiivisuutta, joka selvästi suosii lys-kytkettyjä amidisidoksia. Kuitenkin solu-uutteet, joka sisälsivät ilmennettyä H110D:tä, pilkkoivat lisäksi leu-, met- ja phe-pNA:ta tässä suosituimmuusjärjestyksessä.

Ilmennetyn H110D- (3.5PCR-klooni 2) proteiinin molekyylipaino ja immuunireaktiivisuus

Näytteillä tartutetuista tai verrokkisolu-uutteista suoritettiin elektroforeesiajo 7,5-%:isessa SDS-polyakryyliamidigeelissä, joka värjättiin sitten Coomassie-sini-

sellä. 3.5-2-P3A- ja 3.5-2-P4A-tartutetuilla solulysaateilla oli molemmilla samankokoinen nauha kuin H11OD:llä, 110 kd, joka kulki suoraan koilmennetyn ß-galin alla, jonka molekyylipaino on 120 kd. Tätä 110 kd:n nauhaa ei esiintynyt missään negatiivisista verrokkilysaateista. (Se puuttui myös P2A-lysaatista, joka ei ilmentänyt entsyymiaktiivisuutta.)

Rinnakkaisgeelille suoritettiin Western-blot-siirto, ja seulottiin anti-H11ODN:llä (kuvio 21). Saatiin hyvin voimakas, spesifinen positiivinen immuunireaktio 110 kd:n nauhalle, jonka ilmensivät 3.5-2-P3A ja 3.5-2-P4A, ja natiiville H11OD-dubletille verrokkivyöhykkeessä, joka sisälsi ConA H11OD:tä, kun taas missään negatiivisista verrokkivyöhykkeistä ei nähty reaktiota.

References

Andreason, G.L. & Evans, G.A. (1980) Biotechniques 6, 650.

Blin, N., & Stafford, D.W. (1976). Nucleic Acids Research 3, 2303-2308.

Bowtell, D.D.L., Saint, R.B., Rickard, M.D. & Mitchell, G.F. (1986). Parasitology 93, 599-610.

Chomcyznski, P. & Sacchi, N. (1987). Analytical Biochemistry 162, 156-159.

Cordingley, J.S., Taylor, D.W., Dunne, D.W. & Butterworth, A.E. (1983) Gene 26, 25-39.

Devereux, J., Haerberli, S. & Smithies, O.(1984), Nucleic Acids Research 12, 387-395

Feinberg, A.P. & Vogelstein, B. (1983). Analytical Biochemistry 132, 6-13.

Fire, A. (1986) EMBO Journal 5, 2673-2680.

Fire, A. & Waterston, R.H. (1989) EMBO Journal 8, 3419-3428.

Francis, M.J. & Clarke, B.E. (1989) Methods in Enzymology 178, 659

Frohman, M.A., Dush, M.K. and Martin, G.R. (1988). Proceedings of the National Academy of Sciences, Washington 85, 8998-9002.

Gubler, U. & Hoffman, B.J. (1983). Gene 25, 263-265.

Han, M. & Sternberg, P.W. (1990) Cell 63, 921-931.

Huang, X.-Y. & Hirsch, D. (1990) Proceedings of the National Academy of Sciences Washington 86, 8640-8644.

Johnstone, A. and Thorpe, R. (1982) Immunochemistry in Practice. Blackwell Scientific. London

Jongeneel, C.V., Bouvier, J. & Bairoch, A. (1989). FEBS Letters 242, 211-214.

Kenny, A.J. & Maroux, S. (1982). Physiological Reviews, 62, 91-128.

Kenny, A.J. & Turner, A.J. (1987). Mammalian Ectoenzymes. Elsevier Press, Amsterdam. Volume 14 of Research Monographs in Cell and Tissue Biology. (editors J.T.Dingle & J.L.Gordon).

Laemmli, U.K. (1970) Nature 227, 680-685.

Lojda, Z. & Gossrau, R. (1980) Histochemistry 67, 267-290.

Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982). Molecular Cloning, a Laboratory Manual. Cold Spring Harbor Laboratory Publication.

Merrifield, R.B. (1964) Biochemistry 3, 1385-1390.

Munn, E.A., Smith T.S., Graham, M., Greenwood, C.A., Tavernor, A.S. & Coetzee, G. (1992) Parasitology 106, 63-66.

Nachlas, M.M., Crawford, D.T. and Seligman, A.M. (1957). J Histochem & Cytochem. 5, 264-278.

Nakane, P.K. & Kawoi, A.K. (1974). Journal of Histochemistry and Cytochemistry, 22, 1084-1091.

Noren, O., Sjostrom, H., Danielsen, E.M., Cowell, G.M. & Skovbjerg,H. (1986). The Enzymes of the Enterocyte Plasma Membrane. In Molecular and Cellular Basis of Digestion. Elsevier Press, Amsterdam, (editors P.Desnuelle, H.Sjostrom & O.Noren).

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). Molecular Cloning, a Laboratory Manual, 2nd Edition. Cold Spring Harbor Laboratory Press.

Semenza, G. (1986). Annual Review of Cell Biology, 2, 255-313.

Smith, D.B. & Johnson, K.S. (1988) Gene 67, 31-40.

Spieth, J. MacMorris, M., Broverman, S., Greenspoon, S. & Blumenthal, T. (1988) Dev. Biol. 130, 285-293.

Tavernor, A.S., Smith, T.S. Langford, C.F., Graham, M. & Munn, E.A. (1992a) Parasite Immunol. 14, 671-675.

Tavernor, A.S., Smith, T.S., Langford, C.F., Munn, E.A. & Graham, M. (1992b) Parasite Immunol 14, 645-655.

Vallee, B.L. & Auld, D.S. (1990) Biochemistry 29, 5647-5659.

Watt, V.M. & Yip, C.C. (1989) Journal of Biological Chemistry 264, 5480-5487.

Woodward, MP, Young, WW Jr and Bloodgood, RA (1985). Detection of Monoclonal Antibodies specific for carbohydrate epitopes using periodate oxidation. J. Immunol. Meth. 78, 143-153.

Wu, Q., Lahti, J.M., Air, G.M., Burrows, P.D. & Cooper, M.D. (1990) Procedings of the National Academy of Sciences, Washington 87, 993-997.

GENERAL INFORMATION:

APPLICANT:
NAME: Agricultural and Food Research Council
STREET: Babraham Hall
CITY: Babraham
STATE OR PROVINCE: Cambridgeshire
COUNTRY: United Kingdom
POSTAL CODE: CB2 4AT
TELEPHONE: 223-832312
TELEFAX: 223-837952

TITLE OF INVENTION: Recombinant DNA molecules encoding aminopeptidase enzymes and their use in the preparation of vaccines against helminth infections

NUMBER OF SEQUENCES: 24

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 295 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
CGCGGACATT GCTGAATCTA ACTCCAATCC GTCTTATTGT CGCATTATTT CTAGTAGCTG   60
CTGCAGTCGG CCTCTCTATT GGTCTCACCT ATTACTTTAC TCGCAAAGCG TTCGATACCT  120
CAGAAAAGCC AGGGAAGGAT GATACTGGTG GCAAGGACAA AGACAATTCT CCCTCTGCGG  180
CGGAACTACT TCTTCCAAGT AATATAAAAC CATTGTCTTA CGACTTGACG ATCAAAACAT  240
ATCTACCTGG TTATGTGGAC TTCCCACCGG AGAAAAACCT CACATTCGAC GGGCG       295
```

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 484 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
GCAATATCAG GATTCAATAC AATTTTGGAC TTTTTCGGCA GCGAACCCGA ATCTCAATGG   60
GCTTCGGAAT ACATGCGAAA ACTGATGAAG CCAATTTATG ACAAGAGTAG CATCAAGTTT  120
ATAGCGGAGA ACTACAAAAA AGATTCGCTT TTCTTCAAAA ATAATCTCCA AATAGCTGTT  180
ATTGACACAT ACTGTGGTCT TGGAGGCAAA GAATGTCTTG AAGAAATGAA AAAGCTTTTT  240
GACAAGGAGG TCATGAAATG TCAACCTGGT CAGCAAGCGA CCGACTGCGT AAAGGTAACT  300
GCTCCTCTCC GAAAAACTGT TTACTGCTAT GGGGTCCAGG AAGGCGGTGA TGAGGCATTC  360
GACAAGGTGA TGGAACTATA TAATGCGGAA CAAGTGCAGT TGGAGAAAGA CAGTCTACGT  420
GAAGCATTGG GATGCCATAA AGACGTTACA GCTCTAAAGG GACTTCTTAT GCTGGCTTTG  480
GATC                                                               484
```

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 216 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
ACCTGGTCAG CAAGCGACCG ACTGCGTAAA GGTAACTGCT CCTCTCAAAA CTGTTTACTG   60
CTATGGGGTC CAGGAAGGCG GTGATGAGGC ATTCGACAAG GTGATGGAAC TATATAATGC  120
GGAACAAGTG CAGTTGGAGA AAGACAGTCT ACGTGAAGCA TTGGGATGCC ATAAAGACGT  180
TACAGCTCTA AAGGGACTTC TTATGCTGGC TTTGGA                             216
```

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 581 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
GGGAGGAAAT CACTGCGAGC TTGGAAACAG AACACAGAGC AGTTGATAAA GTGGTCGGCG   60
CTTGTTGCAC AGGAATTCGC TCCCAACAAC AAATAGATCA GCTGAAGAAG AATCTACAGA  120
AGAACAATGC GCAGGCTAAG AAGTTCCATA AAATTGCCTG GATCAAGAAA CATTTTCACA  180
GATTATCGGA ATTCTTCAAG AGAGCAAGAT CATAGCTTTT CACACTGAGC TCCAATTTTA  240
ACGTCTTCAA ACTAGGAGAC AGTTTTGCTG AAAAGTCAGT TTCACATTTT CCGTTTGAAT  300
GCCATCCATT CGAATACAAC CAACCCCATT TTAAGTACCT TTCATTCACA GTGATTACTA  360
AATTTCGAAT ATATTATGAA GCTTGTATCT TGAACGTTAT GATCGGTGAC TTTCAATTTA  420
TAGAGCTCAC TCTCCATTTT GTAGCTGTGA TGACTTGCAT TTAAGACCCA CCATTTACCA  480
GCCTATAATC TTTCCCCAAT ACATTCCAAA CTCCGATCAC CTCCACCGCT GACAATGCCC  540
AGATTTGTTT CTTTGTCTGC TATCCATCTA ACTGTTTCGA T                      581
```

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 948 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
ATGACGTCGC AGGGGAGAAC GCGGACATTG CTGAATCTAA CTCCAATCCG TCTTATTGTC   60
GCATTATTTC TAGTAGCTGC TGCAGTCGGC CTCTCTATTG GTCTCACCTA TTACTTTACT  120
CGCAAAGCGT TCGATACCTC AGAAAAGCCA GGGAAGGATG ATACTGGTGG CAAGGACAAA  180
GACAATTCTC CCTCTGCGGC GGAACTACTC CTTCCAAGTA ATATAAAACC ATTGTCTTAC  240
GACTTGACGA TCAAAACATA TCTACCTGGT TATGTGGACT TCCCACCGGA GAAAAACCTC  300
ACATTCGATG GGCGTGTGGA AATATCAATG GTTGTAATTG AGCCAACAAA GAGTATCGTA  360
CTCAATTCAA AGAAGATCTC TGTAATACCC CAAGAATGTG AACTGGTATC GGGCGATAAA  420
AAATTCGAAA TTGAAAGTGT AAAGGAGCAC CCAAGACTGG AAAAGGTTGA GTTTCTTATC  480
AAAAGCCAAC TGGAAAAAGA TTCACAAATC TTGCTCAAGT CGGCTTACAT CGGTCTCATC  540
AGCAACAGCC TTGGTGGAAT CTACCAGACC ACTTATACCA CCCCGGATGG CACCCCTAAG  600
ATCGCTGCAG TTTCACAAAA TGAGCCCATA GATGCTCGTC GAATGGTACC ATGCATGGAT  660
GAACCGAAAT ACAAAGCAAA CTGGACCGTT ACTGTCATTC ATCCAAAAGG CACCAAAGCC  720
GTCTCGAATG GAATCGAAGT GAACGGAGAT GGAGAGATCA GTGGTGATTG GATCACATCG  780
AAGTTCTTGA CTACTCCACG GATGTCATCC TACTTGTTGG CAGTTATGGT TTCAGAATTT  840
GAATACATCG AAGGTGAAAC AAAGACGGGT GTTCGGTTCC GTATATGGTC ACGCCCAGAG  900
GCAAAGAAGA TGACACAATA TGCTCTGCAA TCTGGTATCA AGTGCATA               948
```

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 1689 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
AATTCCGGCT CGTCCGGAAG CTATGAAGAT GACAGAATAT GCCATGATAG CTGGAATCAA    60
ATGTTTGGAT TACTATGAGG ACTTCTTCGG GATCAAATTC CCACTTCCAA AACAAGATAT   120
GGTTGCTCTT CCTGACTTCT CATCTGGTGC TATGGAGAAC TGGGGTCTCA TCACATACAG   180
GGAGGGTTCC GTGCTCTACG ATGAAAACCT CTACGGACCA ATGAATAAGG AGCGGGTTGC   240
AGAAGTGATC GCGCACGAAC TTGCACATCA GTGGTTCGGT AATTTGGTCA CGATGAAGTG   300
GTGGGATAAC CTATGGCTGA ACGAAGGATT CGCGTCATTC GTGGAATACA TCGGAGCCGA   360
CTTCATCAGC GATGGTCTAT GGGAAATGAA AGATTTCTTC CTGCTGGCAC CGTACACAAG   420
TGGTATTACG GCTGATGCAG TAGCTTCAAG CCATCCGCTT TCCTTCAGAA TAGATAAGGC   480
TGCAGATGTA TCAGAAGCGT TCGATGATAT CACATACCGT AAAGGAGCAT CCGTTCTTCA   540
AATGCTATTG AATTTAGTTG GGGACGAAAA TTTCAAGCAG TCTGTTTCGC GTTACCTCAA   600
GAAGTTTTCA TATGATAATG CGGCTGCTGA AGATTTATGG GCAGCATTCG ACGAAACCGT   660
CCAAGGTATA ACCGGACCTA ATGGTGGACC ATTGAAAATG TCCGAGTTTG CGCCACAATG   720
GACAACTCAG ATGGGGTTCC CTGTTCTTAC TGTCGAGTCG GTTAACGCAA CGACTTTGAA   780
AGTCACCCAA AAACGATACA GGCAGAACAA GGATGCAAAG GAACCAGAGA AGTACCGTCA   840
TCCAACTTAT GGGTTCAAAT GGGATGTTCC TCTGTGGTAT CAGGAAGATG AACAGCAAGT   900
GAAAAGAACT TGGTTAAAAA GAGAGGAACC GCTCTATTTC CATGTAAGCA ATTCTGATTC   960
GTCAGTTGTG GTGAATGCCG AACGTCGTGC TTTTTGCCGA TCAAACTATG ACGCTAACGG  1020
TTGGAGGAAC ATTATGAGAA GACTCAAGCA GAATCATAAG GTCTATGGTC CACGAACAAG  1080
AAACGCTCTC ATAAGTGATG CGTTTGCAGC AGCTGCAGTT GAGGAAATGA ATTACGAGAC  1140
CGTATTTGAA ATGCTCAAAT ACACCGTGAA AGAAGAGGAT TACTTACCAT GGAAGGAGGC  1200
AATATCAGGA TTCAATACAA TTTTGGACTT TTTCGGCAGC GAACCCGAAT CTCAATGGGC  1260
TTCGGAATAC ATGCGAAAAC TGATGAAGCC AATTTATGAC AAGAGTAGCA TCAAGTTTAT  1320
AGCGGAGAAC TACAAAAAAG ATTCGCTTTT CTTCAAAAAT AATCTCCAAA TAGCTGTTAT  1380
TGACACATAC TGTGGTCTTG GAGGCAAAGA ATGTCTTGAA GAAATGAAAA AGCTTTTTGA  1440
CAAGGAGGTC ATGAAATGTC AACCTGGTCA GCAAGCGACC GACTGCGTAA AGGTAACTGC  1500
TCCTCTCCGA AAAACTGTTT ACTGCTATGG GGTCCAGGAA GGCGGTGATG AGGCATTCGA  1560
CAAGGTGATG GAACTATATA ATGCGGAACA AGTGCAGTTG GAGAAAGACA GTCTACGTGA  1620
AGCATTGGGA TGCCATAAAG ACGTTACAGC TCTAAAGGGA CTTCTTATGC TGGCTTTGGA  1680
TCGGAATTC                                                          1689
```

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 2472 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
ACGGGTGTTC GGTTCCGTAT ATGGTCTCGA CCAGAGGCGA AACGAATGAC GGCATACGCT    60
TTGGATGCTG GCATCAGATG CCTGGAGTTC TATGAGAAGT TCTTTGACAT AAAATTCCCT   120
CTGGAAAAAC AAGATATGAT TGCTCTTCCT GATTTCACCG CTGGTGCCAT GGAAAACTGG   180
GGCCTTATCA CTTATAGAGA GGATTCTCTC CTATACGATG AAAAAATTTA TGCACCGATG   240
AATAAACAGC GGGTTGCTCT CGTAGTTGCT CACGAGCTTG CTCATCAGTG GTTCGGCAAT   300
CTGGTCACAC TGAAGTGGTG GGATGATACG TGGTTGAACG AAGGTTTTGC AACATTTGTT   360
GAGTATCTTG GAATGGACGA AATTAGCCAC AACAATTTCA GAACGCAAGA TTTCTTCTTG   420
CTCGATGGAA TGGATCGCGG AATGAGAGCT GACTCGGCAG CATCGAGCCA TCCGCTTTCG   480
TTTAGGATTG ACAAAGCGGC AGAAGTTGCC GAAGCCTTTG ACGATATTTC ATACGCCAAG   540
GGAGCGTCAG TTCTCACTAT GCTACGGGCT TTGATTGGAG AGGACAATTA CAGGAATGCT   600
GTTGTGCAAT ACCTCAAGAA GTTCTCCTAC AGCAATGCAC AAGCAGCCGA TCTGTGGAAC   660
GTCTTCAATG AAGTTGTCAA AGGTGTTAAG GGTCCTGACG GCAACGTCAT GAAAATCGAC   720
CAATTTACCG ATCAGTGGAC GTATCAGATG GGTTATCCTG TGGTTAAAGT AGAAGAATTT   780
AATGCGACCG CCCTAAAGGT TACGCAGAGC CGGTACAAGA CAAATAAAGA CGCCTTGGAA   840
```

```
CCAGAGAAAT ATCGTAATCC AAAATACGGG TTCAAGTGGG ATGTTCCCCT ATGGTATCAG   900
GAAGGCAATA GCAAAGAGGT GAAGCGAACA TGGCTAAAAA GAGATGAACC GCTGTACTTG   960
AACGTCAACA ATCGGGATAC ATCCCTTGTG GTGAACGCTG ATCGACATGG ATTTTATCGA  1020
CAAAACTATG ATGCCAACGG TTGGAAAAAG ATAATCAAGC AGCTCAAGAA AGATCACAAG  1080
GTCTTCGGTC CAAGGACAAG GAACGCTATC ATAAGCGATG CATTTGCTGC AGCTACGATT  1140
GACGCAATCG ACTATGAAAC TGTATTCGAA CTACTTGAAT ATGCCAAAAA TGAAGAGGAA  1200
TTCTTGCCTT GGAAGGAAGC TCTGTCCGGC ATGTTCGCAG TTTTAAAGTT CTTCGGTAAT  1260
GAGCCGGAGA CAAAACCAGC TAGAGCTTAC ATGATGAGCA TATTAGAACC GATGTATAAT  1320
AAGAGCAGCA TTGATTACAT CGTCAAGAAT TATTTGGATG ATACGTTATT CACAAAAATT  1380
AATACTCAAA AGGATATCAT TGATGCATAT TGTTCCCTTG GATCAAAGGA CTGTATAAAG  1440
CAATATAAGG ATATCTTCTA CGATGAGGTT ATGCCCAAGT GTAAGGCCGG GGAAGCAGCA  1500
ACCAAATGCG TTAAGGTTTC CGCTCCTCTT CGAGCCAATG TTTACTGTTA TGGTGTACAG  1560
GAAGGTGGTG AAGAAGCTTT TGAAAAGGTG ATGGGGCTGT ATCTAGCAGA AGATGTTCAA  1620
CTGGAGAAGG GTATCCTGTT CAAAGCCTTG GCATGCCACA AAGATGTTAC AGCTCTAAAA  1680
GAACTTCTTT TGCGAGCCCT GGACCGTAAA TCGTCGTTTG TGCGTCTTCA GGATGTCCCT  1740
ACCGCTTTCC GTGCTGTATC TGAAAACCCT GTGGGCGAAG AATTCATGTT CAATTTCCTA  1800
ATGGAGAGAT GGGAGGAAAT CACTGCGAGC TTGGAAACAG AACACAGAGC AGTTGATAAA  1860
GTGGTCGGCG CTTGTTGCAC AGGAATTCGC TCCCAACAAC AAATAGATCA GCTGAAGAAT  1920
CTACAGAAGA ACAATGCGCA GGCTAAGAAG TTCGGCTCAT TCACCCAGGA AATCGAAAAA  1980
GGAGAACATA AAATTGCCTG GATCAAGAAA CATTTTCACA GATTATCGGA ATTCTTCAAG  2040
AGAGCAAGAT CATAGCTTTT CACACTGAGC TCCAATTTTA ACGTCTTCAA ACTAGGAGAC  2100
AGTTTTGCTG AAAAGTCAGT TTCACATTTT CCGTTTGAAT GCCATCCATT CGAATACAAC  2160
CAATAATACC ATTTTAAGTA CCTTTCATTC ACAGTGATTA CTGAATTTCG AATATATCAT  2220
GAAGCTTGTA TCTTGAACGT TATGATCGGT GACTTTCAAT TTATAGAGCT CACTCTCCAT  2280
TTTGTAGCTG TGATGACTTG CATTTAAGAC CCACCATTTA CCAGCCTAGA ATCTTTCCCC  2340
AATACATTCC AAACTCCGAT CACCTCCACC GCTGACAATG CCCAGATTTG TTTTTTTGTC  2400
TGCTATCCAT CTAACTGTTT CGATCGCCGG TTGTTTGTCA ATTGCTTATC TGATAAATAT  2460
TGACGTTGGT GT                                                      2472
```

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 3305 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```
GCTGAATCTA ACTCCAATCC GTCTTATTGT CGCATTATTT CTAGTAGCTG CTGCAGTCGG    60
CCTCTCTATT GGTCTCACCT ATTACTTTAC TCGCAAAGCG TTCGATACCT CAGAAAAGCC   120
AGGGAAGGAT GATACTGGTG GCAAGGACAA AGACAATTCT CCCTCTGCGG CGGAACTACT   180
CCTTCCAAGT AATATAAAAC CATTGTCTTA CGACTTGACG ATCAAAACAT ATCTACCTGG   240
TTATGTGGAC TTCCCACCGG AGAAAAACCT CACATTCGAT GGGCGTGTGG AAATATCAAT   300
GGTTGTAATT GAGCCAACAA AGAGTATCGT ACTCAATTCA AAGAAGATCT CTGTAATACC   360
CCAAGAATGT GAACTGGTAT CGGGCGATAA AAAACTCGAA ATTGAAAGTG TAAAGGAGCA   420
CCCAAGACTG GAAAAGGTTG AGTTTCTTAT CAAAAGCCAA CTGGAAAAAG ATCAACAAAT   480
CTTGCTCAAG GTCGGCTACA TCGGTCTCAT CAGCAACAGC TTTGGTGGAA TCTACCAGAC   540
CACTTATACC ACCCCGGATG GCACCCCTAA GATCGCTGCA GTTTCACAAA ATGAGCCCAT   600
AGATGCTCGT CGAATGGTAC CATGCATGGA TGAACCGAAA TACAAAGCAA ACTGGACCGT   660
TACTGTCATT CATCCAAAAG GCACCAAAGC CGTCTCGAAT GGAATCGAAG TGAACGGAGA   720
TGGAGAGATC AGTGGTGATT GGATCACATC GAAGTTCTTG ACTACTCCAC GGATGTCATC   780
CTACTTGTTG GCAGTTATGG TTTCAGAATT TGAATACATC GAAGGTGAAA CAAAGACGGG   840
TGTTCGGTTC CGTATATGGT CACGCCCAGA GGCAAAGAAG ATGACACAAT ATGCTCTGCA   900
ATCTGGTATC AAGTGCATAG AATTCTACGA AGATTTCTTT GATATCAGAT TCCCTCTGAA   960
GAAACAAGAT ATGATTGCCC TTCCTGATTT CTCTGCCGGT GCCATGGAGA ATTGGGGCCT  1020
CATCACTTAC AGGGAAAACT CTTTGTTGTA CGATGACAGA TTCTATGCAC CGATGAATAA  1080
ACAGCGAATT GCTCGCATTG TTGCTCATGA GCTTGCTCAT CAGTGGTTCG GCGACTTGGT  1140
TACGATGAAG TGGTGGGATA ATTTGTGGTT GAATGAAGGT TTTGCAAGAT TCACAGAATT  1200
TATTGGAGCT GGTCAGATAA CTCAAGATGA CGCCAGAATG AGGAACTACT TCCTGATTGA  1260
TGTACTTGAA CGCGCTTTGA AAGCTGATTC GGTAGCGTCA AGCCATCCAC TTTCCTTCAG  1320
AATCGACAAA GCTGCAGAAG TTGAAGAAGC CTTTGATGAT ATCACATACG CCAAAGGAGC  1380
TTCTGTTCTT ACTATGCTGA GAGCCTTGAT TGGAGAAGAA AAACATAAGC ATGCAGTATC  1440
GCAGTACCTC AAGAAGTTCT CGTATAGCAA TGCAGAAGCG ACTGATCTAT GGGCAGTTTT  1500
TGATGAAGTT GTCACTGACG TCGAAGGTCC AGACGGCAAA CCTATGAAAA CCACAGAGTT  1560
```

```
TGCAAGTCAG TGGACGACTC AGATGGGCTT CCCAGTTATT TCCGTAGCAG AGTTTAACTC   1620
GACTACTTTG AAATTAACGC AAAGTCGATA TGAGGCGAAT AAAGACGCTG TGGAGAAAGA   1680
GAAGTACCGT CACCCGAAAT ACGGATTTAA ATGGGATATT CCACTGTGGT ATCAGGAAGG   1740
CGATAAGAAG GAGATAAAGC GAACATGGTT GAGAAGAGAT GAACCGCTTT ACTTGCATGT   1800
TAGTGATGCT GGCGCTCCCT TTGTGGTGAA CGCAGACCGC TATGGATTTT ATCGACAAAA   1860
TCATGACGCT AATGGTTGGA AAAAGATAAT CAAGCAGCTC AAGGATAATC ATGAGGTTTA   1920
CAGTCCCCGG ACAAGGAATG TCATCATTAG CGATGCGTTT GCTGCGGCTG CAACTGACGC   1980
AATTGAGTAT GAGACTGTAT TTGAACTTCT GAATTATGCC GAAAAAGAAA CGGAATATCT   2040
ACCATTAGAA ATCGCAATGT CCGGGATCTC TTCGATTTTG AAATACTTCC CTACCGAGCC   2100
AGAGGCAAAG CCAGCTCAAA CATACATGAT GAACATATTG AAACCGATGT ATGAAAAAAG   2160
CAGTATCGAC TTCATTGCCA ATAACTACAG AAATGACAAG CTGTTTTTCC AAATCAACCT   2220
CCAAAAAGAT GTCATTGATA TGTTCTGCGC CCTCGGATCG CAAGACTGCA GGAAGAAATA   2280
TAAAAAACTT TTCGATGACG AAGTCATGAA CAAATGCAGG GATGGTCAAG CAGCAACCGA   2340
ATGCGTAAGA ATCGCCGCTC CTCTCCGATC AAGTGTTTAT TGTTATGGTG TGAAGGAAGG   2400
CGGTGATTAT GCTTCCGACA AGGTGATGGA GCTTTATACG GCCGAAACAC TCGCCCTAGA   2460
AAAAGACTTC CTACGCCTAG CATTGGGATG TCATAAAGAT GTTACTGCTT TGAAAGGACT   2520
TCTCTTGCGG GCTCTGGACA GGAATTCGTC GTTCGTACGT ATGCAGGATA TCCCAAGTGC   2580
TTTCAACGAT GTAGCAGCAA ATCCTATTGG CGAAGAATTC ATTTTCAATT TCCTTATTGA   2640
GAGATGGCCA GATATCATTG AAAGTATAGG AACGAAGCAC ACATATGTTG AGAAAGTGAT   2700
ACCAGCCTGC ACTTCAGGAA TCCGCTCACA ACAGCAGATT GACCAGCTGA AGAATCTGCA   2760
GAAAAATGGC ATGAACGCTC GTCAATTCGG TGCATTCGAT AAAGCAATCG AACGAGCACA   2820
AAATAGGGTG GATTGGATTA AAAAACATTT CCAAAAATTA GCGGCTTTCT TCAAGAAAGC   2880
CACCTTGTAA TTCGAATTAC ATTGCCAGTA ATCCAGATCT TAAAGTTCAT GAAGGAATAT   2940
GACAGGGAAC TGACTGTCTG TTGGTCACTG TTCCACTGAA TGGAAGTTTT TACCTACAAA   3000
AATTTTTATC GTTATATTTG CCTTCCGTGA GGGGTCATTG TTGTCACTTG AATAGTAAAC   3060
AAAGCTCAGT ATTGGCAACC GTAGAACAAT ATTACTTTCG CTTCATCAAA TTGTTATCTT   3120
CCCTATACCC TCTTCCTAAC TGAATTCGGA AATTTGTTCA TATTCGTTTG TAGTCTGTTG   3180
CTCAGAACAC TTTCTCCTCA ATAGCTTCTT GTTTGTTTTT TTTTTGATTG TATTGATCGT   3240
TTTACAATTG TATAGATTAG TTATCTTATA AATATTGATG GTTAAAAAAA AAAAAAAAAA   3300
AAAAA                                                               3305
```

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 1301 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```
GGTTTAATTA CCCAAGTTTG AGATGACAGC AGAGGAGAGT CAGGAGCAGG AGACGCAGCA    60
ACCACGAAAA AATACAGTGC TACGGCTCAC CCCAATCAAG TCTCTCTTTG CTTTGTTAGT   120
GGTAGCTGCT GCCGTCGGCC TCTCAATCGG TCTCACCTAT TACTTTACAA GGAAAGCTTT   180
TGATACTACT GGCGGAAATG GAAAAGGGGA TCAACCTATT GTCGATGATA ATTCCCCATC   240
AGCTGAAGAA TTACGTCTCC CAACAACCAT AAAACCTTTG ACATACGACT TAGTAATCAA   300
AACGTATCTG CCAAACTATG TAAACTATCC ACCTGAGAAA GATTTCGCTA TTGATGGGAC   360
TGTGGTGATT GCTATGGAAG TTGTGGAGCC AACAAAGTCT ATTGTGCTCA ACTCGAAAAA   420
TATTCCTGTA ATTGCAGACC AGTGCGAACT GTTTTCTAAC AACCAAAAAC TCGACATCGA   480
AAAGGTTGTG GATCAGCCAA GGCTGGAGAA AGTCGAATTC GTTTTGAAGA AAAAGCTGGA   540
GAAGAATCAG AAAATCACGC TCAAGATTGT ATACATTGGC CTTATCAACG ACATGCTTGG   600
AGGACTTTAT CGAACAACCT ACACGGATAA AGATGGTACA ACCAAGATTG CTGCATGCAC   660
TCATATGGAA CCGACGGACG CCCGTCTTAT GGTCCCCTGT TTCGACGAGC CGACGTTTAA   720
GGCAAACTGG ACTGTGACTG TGATTCATCC GAAGGGCACC AGTGCCGTGT CGAATGGAAT   780
AGAAAAGGGA GAAGGAGAAG TCTCTGGCGA TTGGGTCACA ACCAGATTCG ATCCAACCCC   840
GCGAATGCCT TCGTATTTGA TTGCTCTTGT GATTTCCGAA TTTAAGTACA TTGAAAATTA   900
TACGAAAAGC GGTGTTCGAT TCCGAATATG GGCTCGTCCG GAAGCTATGA AGATGACAGA   960
ATATGCCATG GTAGCTGGAA TCAAATGCTT GGATTACTAT GAGGACTTCT TCGGGATCAA  1020
ATTCCCTCTT CCAAAACAAG ATATGGTTGC TCTTCCTGAC TTCTCATCTG GTGCTATGGA  1080
GAACTGGGGT CTCATCACAT ACAGGGAGGG TTCCGTACTA TACGATGAAA ACCTCTATGG  1140
ACCAATGAAT AAGGAGCGGG TTGCAGAAGT GATTGCACAC GAGCTTGCAC ATCAGTGGTT  1200
CGGTAATTTG GTCACGATGA AGTGGTGGGA TAACCTATGG CTAAACGAAG GATTCGCGTC  1260
ATTCGTAGAA TACATTGGAG CCGACTTCAT CAGCGATGGT C                       1301
```

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1280 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GGTTTAATTA CCCAAGTTTG AGATGACGGC GGAGTGGCAG AAGCGTCGAA TCTTGGGCTT 60
CTCACCTATC AGCCTACTTT GTACATTATT TGTATTAGCT GCTGCCGTTG GACTCTCCAT 120
TGGTCTTACC TATTACTTCA CTCGTAAAGC ATTCGATACC ACACAAAAAG AACAGAAGGA 180
TGACAGTGGT GGTAAAGAAA AGGATAATTC TCCTTCTGCA GAAGAACTAC TTCTTCCAAC 240
GAACATAAAA CCAGTCTCGT ACGACTTGAA CATCAAAACA TATCTACCGG GTTACGTGAA 300
CTTTCCACCA GAAAAGAATC TCACATTTGA TGCCCATGTG GAGATTGCTA TGGTTGTGGT 360
TGAGCCTACA AATAGTATTG TGCTGAACTC GAAGAAAATC ACTTTGGCAC AAGGAGGATG 420
CGAACTGTTC TCAGGTAATC AGAAACTTGA CATCGAAAGT GTAAAGATGC AGGAAAGACT 480
TGACAAGCTT GAGATTACCC TCAAAAATCA GCTGCAAAAA GATCTGAAAA TCCTGCTCAA 540
GATCACTTAC ACCGGCCTTA TTAGCGACAC TCTCGGTGGG CTCTACCAGT CCATCTACAC 600
TGATAAGGAC GGAAAAACTA AGATCGTTGC TGTTTCACAA AATGAACCAT CAGACGCTCG 660
TCGTATAGCG CCATGCTTTG ACGAACCGAA GTACAAGGCA ACATGGACTG TCACCGTCGT 720
TCATCCCAAA GGTACAAAGG CTGCATCGAA CGGCATTGAA GCAAATGGAA AAGGGGAGCT 780
CAAGGGTGAT TGGATAACGT CTAAATTTAA AACTACCCCA CCGATGTCGT CCTATTTATT 840
GGCTATTATT GTTTGTGAAT TTGAATACAT TGAAGGATTT ACAAAAACAG GTGTACGGTT 900
CCGTATATGG TCTCGACCAG AGGCGATGGC AATGACGGGA TATGCCCTGG ATGCTGGCAT 960
CAGATGTCTG GAGTTCTATG AGAGATTCTT TGACATCAAA TTCCCTCTGG AAAAACAAGA 1020
TATGATTGCT CTACCTGATT TCACCGCTGG TGCTATGGAA AACTGGGGTC TTATCACTTA 1080
CAGAGAGGAT TCTCTTCTAT ACGATGAGAA AATTTATGCG CCGATGAATA AGCAGCGGGT 1140
TGCTCTCGTA GTTGCACACG AGCTTGCTCA TCAGTGGTTC GGCAATCTGG TCACATTGAA 1200
GTGGTGGGAT GATACGTGGT TGAACGAAGG TTTTGCGACA TTTGTTGAAT ATCTTGGAAT 1260
GGACGAAATT AGCCACAACA 1280

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 1293 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGTTTAATTA CCCAAGTTTG AGGGTCTCCA TCTAGATGAC GTCGCAGGGG AGAACGCGGA 60
CATTGCTGAA TCTAACTCCA ATCCGTCTTA TTGTCGCATT ATTTCTAGTA GCTGCTGCAG 120
TCGGCCTCTC TATTGGTCTC ACCTATTACT TTACTCGCAA AGCGTTCGAT ACCTCAGAAA 180
AGCCAGGGAA GGATGATACT GGTGGCAAGG ACAAAGACAA TTCTCCCTCT GCGGCGGAAC 240
TACTCCTTCC AAGTAATATA AAACCATTGT CTTACGACTT GACGATCAAA ACATATCTAC 300
CTGGTTATGT GGACTTCCCA CCGGAGAAAA ACCTCACATT CGATGGGCGT GTGGAAATAT 360
CAATGGTTGT AATTGAGCCA ACAAAGAGTA TCGTACTCAA TTCAAAGAAG ATCTCTGTAA 420
TACCCCAAGA ATGTGAACTG GTATCGGGCG ATAAAAAACT CGAAATTGAA AGTGTAAAGG 480
AGCACCCAAG ACTGGAAAAG GTTGAGTTTC TTATCAAAAG CCAACTGGAA AAAGATCAAC 540
AAATCTTGCT CAAGGTCGGC TACATCGGTC TCATCAGCAA CAGCCTTGGT GGAATCTACC 600
AGACCACTTA TACCACCCCG GATGGCACCC CTAAGATCGC TGCAGTTTCA CAAAATGAGC 660
CCATAGATGC TCGTCGAATG GTACCATGCA TGGATGAACC GAAATACAAA GCAAACTGGA 720
CCGTTACTGT CATTCATCCA AAAGGCACCA AAGCCGTCTC GAATGGAATC GAAGTGAACG 780
GAGATGGAGA GATCAGTGGT GATTGGATCA CATCGAAGTT CTTGACTACT CCACGGATGT 840
CATCCTACTT GTTGGCAGTT ATGGTTTCAG AATTTGAATA CATCGAAGGT GAAACAAAGA 900
CGGGTGTTCG GTTCCGTATA TGGTCACGCC CAGAGGCAAA GAAGATGACA CAATATGCTC 960
TGCAATCTGG TATCAAGTGC ATAGAATTCT ACGAAGATTT CTTTGATATC AGATTCCCTC 1020
TGAAGAAACA AGATATGATT GCCCTTCCTG ATTTCTCTGC CGGTGCCATG GAGAATTGGG 1080
GCCTCATCAC TTACAGGGAA AACTCTTTGT TGTACGATGA CAGATTCTAT GCACCGATGA 1140
ATAAACAGCG AATTGCTCGC ATTGTTGCTC ATGAGCTTGC TCATCAGTGG TTCGGCGACT 1200
TGGTTACGAT GAAGTGGTGG GATAATTTGT GGTTGAATGA AGGTTTTGCA AGATTCACAG 1260
AATTCACTGG AGCTGGTCAG ATAACTCAAG ATG 1293

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 746 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

```
CTTGAAGAAA TGAAAAAGCT TTTTGACAAG GAGGTCATGA AATGTCAACC TGGTCAGCAA    60
GCGACCGACT GCGCGAAGGT AACTGCTCCT CTCCGAAAAA CTGTTTACTG CTATGGGGTC   120
CAGGAAGGCG GTGATGAGGC ATTCGACAAG GTGATGGAAC TATATAATGC GGAACAAGTC   180
CAGTTGGAGA AAGACAGTCT ACGTGAAGCA TTGGGATGTC ATAAAGACGT TACTGCTCTA   240
AAGGGACTTC TTATGCTGGC GTTGGATCGG AATTCGTCAT TTGTGCGTCT TCAAGATGCT   300
CATGATGTGT TTAACATTGT ATCCAGAAAT CCTGTTGGAA ACGAACTGCT GTTCAATTTC   360
CTCACAGAGC GATGGGAAGA GATACTTGAA AGTTTGTCAA TACGACACAG ATCAGTTGAT   420
CGAGTGATCA AAGCCTGTAC TCGAGGACTA CGATCCAGGG AACAAGTACA ACAGTTGAAG   480
AATCTATACA AAAATGACAA GCGTGCTCGC GAATACGGTG CATTTGGTGG GGCAATAGAA   540
AGATCGGAAC ACAGAGTCAA ATGGATTGAG AAACATTTCC GAAAACTAGC AGCTTTCTTC   600
AAAAAATCTA ATTCATAATT CTGAAATGGC TATAACTAGC ACACTGGATA GTTGTCTCGA   660
ATCATCCAAA AAGATTAATG ATGTTTTTTT ACTAGATAAT ATGGAGATAT TCTGTAAATT   720
TGTCATCGAT TCAAGTGTCT GTATTG                                        746
```

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 1274 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

```
TCTTGAAGAA ATGAAAAAGC TTTTCGACGA AGAGGTCATG AAAAAATGTA GACCTGGTCA    60
GCAAGCGACC GACTGCGTAA AGGTAACTGC TCCTCTCCGA AAAACTGTTT ACTGCTATGG   120
GGTCCAGGAA GGCGGTGATG AGGCATTCGA CAAGGTGATG GAACTATATA ATGCGGAACA   180
AGTCCAATTG GAGAAAGACA GTCTACGTGA AGCATTGGGA TGTCATAAAG ACGTTACTGC   240
TCTAAAGGGA CTTCTTATGC TGGCTTTGGA TCGGAATTCG TCATTTGTGC GTCTTCAAGA   300
TGCTCATGAT GTGTTTAACA TTGTATCCAG AAATCCTGTT GGAAACGAAC TGCTGTTCAA   360
TTTCCTCACA GAGCGATGGG AAGAGATACT TGAAAGTTTG TCAATACGAC ACAGATCAGT   420
TGATCGAGTG ATCAAAGCCT GTACTCGAGG ACTACGATCC AGGGAACAAG TACAACAGTT   480
GAAGAATCTA TACAAAAATG ACAAGCGTGC TCGCGAATAC GGTGCATTTG GTGGGGCAAT   540
AGAAAGATCG GAACACAGAG TCAAATGGAT TGAGAAACAT TTCCGAAAAC TAGCAGCTTT   600
CTTCAAAAAA TCTAATTCAT AATTCTGAAA TGGCTATAAC TAGCACACTG GATAGTTGTC   660
TCGAATCATC CAAAAAGATT AATGATGTTT TTTTACTAGA TAATATGGAG ATATTCTGTA   720
AATTTGTCAT CGATTCAAGT GTCTGTATTG CAGCCACATT ACATATCTCG ATGGTTCTGT   780
GAATTTTTGA TGGAATTATT TTCTCCTCAA AATAGACACT ATGCGCTAAC TCCCATTATT   840
ACCAATCTTT GAGAGAAATC TTTTGCAATA TACCCTAAAT AGCCCTTGGG AACTAGCTTT   900
TTTCATTATT GTAATTTTTG TACTCTTCAA ATGACGTATT TCCAACATGA CACATTCTCA   960
GTGATTTACT CGAGTAATTT TATTCTTCTC AATTGCAGTG CCTTATTGTT ATTCGCTTTG  1020
AGACTTTGTA GCTCAACTGT TTTCTGCCCT GCTGTCTTCT TCTCTTTATC TACTACTTCA  1080
GTGATGAACT TACCTGAAGT TGTAGGTTTT AAGAAAGAAA TAACTATTTT CCATAACTCA  1140
TCTTCATGCC ATTGTTCTTT GGACTTTCTC ATGCTTCACA TTGTAGAGAT ATTTACTAAA  1200
AATGGAATTC TAATTTTCGT TTACTTACAT AAAAATCACT TATTGCCTAA AAAAAAAAAA  1260
AAAAAAAAAA AAAA                                                    1274
```

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 3296 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GCTGAATCTA ACTCCAATCC GTCTAATTTT TGCATTATTT CTAGTAGCTG CTGCAGTCGG 60
CCTCTCTATT GGTCTCACCT ATTACTTTAC TCGCAAAGCG TTCGATACCT CAGAAAAGCC 120
AGGGAAGGAT GATACTGGTG GCAAGGGCAA AGACAATTCT CCCTCTGCGG CGGAACTACT 180
TCTTCCAACC AATATAAAAC CATTGTCTTA CGATTTGACG ATCAAAACAT ATCTACCTGG 240
TTATGTGAAC TTCCCACCGG AGAAGAATCT CACATTCGAT GGGCGTGTGG AAATATCAAT 300
GGTTGTAATT GAGCCAACAA AGAGTATCGT GCTCAATTCA AAGAAGATCT CTGTGATACC 360
CCAAGAATGT GAACTGGTAT CGGGCGATAA AAAACACGAA ATTGAAAGTG TAAAGGAGCA 420
CCCAAGACTG GAAAAGGTCG AGTTTCTTCT TAAGAACCAA CTGGAAAAAG ATCAACAAAT 480
CTTGCTCAAG GTCGGCTATA TCGGCCTCAT CAGCAACAGT CTTGGAGGAA TCTACCAGAC 540
CACTTACACC ACCCCGAATG GCACCCCTAA GATCGCTGCA GTTTCACAAA ATGAGCCCAT 600
AGATGCTCGT CGAATGGTAC CATGCATGGA CGAACCGAAA TACAAAGCAA ACTGGACCGT 660
TACTGTCATT CATCCAAAAG GCACCAAAGC CGTCTCGAAT GGAATCGAAG TGAACGGAGA 720
TGGAGAGATC AGTGGTGATT GGATCACATC GAAGTTCTTG ACTACTCCAC GGATGTCATC 780
CTACTTGTTG GCAGTTATGG TTTCAGAATT TGAATACATT GAAGGTGAAA CAAGGACGGG 840
TGTCCGGTTC CGCATATGGT CACGCCCAGA GGCCAAGAAG ATGACAAAAC TTGCTTTGGA 900
TTATGGTATC AAATGCATAG AGTTCTACGA AGATTTCTTT GATATCAAAT TCCCTCTGAA 960
AAAACAAGAT ATGATCGCCC TTCCTGATTT CTCAGCAGGA GCCATGGAGA ACTGGGGTCT 1020
TATCACTTAC AGGGAAAACT CTTTGTTGTA CGATGACAGA TTCTATGCAC CGATGAATAA 1080
ACAGCGAATT GCTCGCATTG TTGCTCATGA GCTTGCCCAT CAGTGGTTTG GGGACTTGGT 1140
TACAATGAAG TGGTGGGATA ATCTGTGGTT GAATGAAGGT TTTGCAAGAT TCACGGAATT 1200
CATTGGAGCT GGTCAGATAA CTAAAGATGA CGCCAGAATG AGGAACTACT TTCTGATTGA 1260
TGTACTTGAA CGCGCTTTGA AAGCTGATTC GGTAGCGTCA AGCCATCCAC TTTCCTTCAG 1320
AATCGACAAA GCTGCAGAAG TTGAAGAAGC CTTTGATGAT ATCACATACG CCAAAGGAGC 1380
TTCTGTTCTT ACTATGTTGA GAGCCTTGAT TGGAGAAGAA AAACATAAGC ATGCAGTATC 1440
GCAGTACCTC AAGAAGTTCT CGTATAGCAA TGCAGAAGCG ACTGATCTAT GGGCAGTTTT 1500
CGATGAAGTT GTCACTGACG TCGAAGGTCC AGACGGCAAA CCTATGAAAA CCACGGAATT 1560
TGCAAGTCAG TGGACGACTC AGATGGGCTT CCCAGTTATT TCCGTAGCAG AGTTTAACTC 1620
GACTACTTTG AAATTAACGC AAAGTCGATA TAAGGCGAAT AAAGACGCTG TGGAGAAAGA 1680
GAAGTACCGT CATCCGAAAT ACGGATTTAA ATGGGATATT CCATTGTGGT ATCAGGAAGG 1740
CGATAAGAAG GAGATAAAGC GAACATGGCT GAGAAGAGAT GAACCGCTTT ACTTGCATGT 1800
TAGTAATCCT GGTGCTCCAT TTGTGGTGAA CGCAGACCGC TATGGATTTT ATCGACAAAA 1860
TCATGACGCT AATGGTTGGA AAAAGATAAT CAAGCAGCTC AAGGACAATC ATGAGGTTTA 1920
TAGTCCTCGG ACAAGGAATG TCATCATTAG CGATGCGTTT GCTGCAGCCG CAACTGACGC 1980
AATTGAGTAT GAGACTGTTT TTGAACTTCT GAAATATGCC GAAAAAGAAA CGGAATACCT 2040
ACCGTTGGAA ATAGCAATGT CCGGGATCTC TTCGATTTTG AAGTACTTCG GTACCGAGCC 2100
GGAAGCAAAG CCAGCTCAAG TGTACATGAT GAACATATTG AAGCCGATGT ATGAAAAAAG 2160
CAGTATCGAG TTCATTACCA ATAACTACAG AAACGACACG CTGTTTTTCC AAATCAACCT 2220
CCAAAAGGAT GTCGTTGATA TGTTCTGCGC CCTTCGATCG CAAGACTGCA GGCAGAAATA 2280
TAAAAAACTT TTCGATGACG AAGTCATGGC GAAATGCAGG GATGGTCAAG CAGCAACCGA 2340
ATGCGTGAGA ATCGCCGCTC CTCTCCGATC AAGTGTTTAT TGTTATGGTG TGAAGGAAGG 2400
CGGTGATTAT GCTTTCGACA AGGTGATGGA GCTTTATACG GCCGAAACAC TTGCCCTAGA 2460
AAAAGACTTC CTACGCCTAG CATTAGGATG TCACAAAGAT GTTACTGCTT TGAAAGGACT 2520
TCTCTTGCGG GCTCTGGACA GGAATTCGTC ATTCGTACGT ATGCAGGATA TCCCAAGTGC 2580
TTTCAACGAT GTAGCAGCAA ATCCTATTGG CGAAGAATTC ATTTTCAATT TCCTCATTGA 2640
GAGATGGCCA GATATCATTG AAAGTATAGG AACGAAGCAC ACATATGTTG AGAAAGTGAT 2700
ACCAGCCTGC ACTTCAGGAA TCCGCTCACA ACAGCAGATT GACCAGCTGA AGAATCTGCA 2760
GAAAAATGGC ATAAATGCTC GTCAATTTGG TGCATTCGAT AAAGCAATCG AACGAGCACA 2820
AAATAGGGTG GATTGGATTA AAAAACATTT CCAAAAATTA GCGGCTTTCT TCAAGAAAGC 2880
CACCTTGTAG TTTGAATTAC GTCGCCATTA ATCCAGATCT TAAAGCTCGC TAAGGAATAT 2940
GTGGGAACTG ACTGTGTGTT GGTTACTGTT CCACTGAATG GAAGTTTTTA CCCACAAAAA 3000
TTTTTACCAT TTGCCTTCCA TGAGGGGTCA TTGTTGTCAC TTGAATAGTA AACAAAGCTC 3060
AGTATTAGGA CCCAGTGATC AATATTACTT TTGCTTCATC AAATTGTTAC CTTCTCTATA 3120
CCCTCTTCCT ACCTGAATTC AGAAATTTGT TCATATTCGT TTGTAGTCTG TTGCTCAGAA 3180
CACTTTCTCC TCGATAGCTT TTTGTTTGTT TTTCTTTTGA TTGTATTGAT CGTTTTACAA 3240
TTGTATAGAT TAGTTATCTG ATAAATATTG ATGGCTAAAA AAAAAAAAAA AAAAAA 3296

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 3319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GCTGAATCTA ACTCCAATCC GTCTAATTTT TGCATTATTT CTAGTAGCCG CTGCAGTCGG 60
CCTCTCTATT GGTCTCACCT ATTACTTTAC TCGCAAAGCG TTCGATACCT CAGAAAAGCC 120
AGGGAAAGAT GATACTGGTG GTAAAGACAA AGATAATTCT CCCTCTGCGG CGGAACTACT 180
TCTTCCAACC AACATAAAAC CATTGTCTTA CGATTTGACA ATCAAAACAT ATCTACCTGG 240
TTATGTGAAC TTCCCACCGG AGAAGAATCT CACATTTGAT GGGCGTGTCG AAATTTCAAC 300
GGTTGTCATT GAGCCAACAA AGAGTATCGT GCTCAATTCA AAGAAGATCT CAGTAATACC 360
CCCTGAATGT GAACTGGTAT CGGGCGGTAA AAAACTCGAA ATCGAAAATG TAAAGGATCA 420
CCCAAGACTG GAAAAGGTTG AGTTTCTTCT TAAGAACCAA CTGGAAAAAG ATCAACAAAT 480
CTTGCTCAAG GTTGCCTACA TCGGCCTCAT CAGCAACAGC CTTGGCGGAA TCTACCAGAC 540
CACTTACACA ACCCCGGATG GCACCCCCAA GATCGCTACA GTGTCACAAA ATGAGCCCAT 600
AGATGCTCGT CGGATGGTGC CATGCATGGA TGAACCGAAA TACAAAGCGA ATTGGACCGT 660
TACTGTCATT CATCCAAAAG GTACAAAAGC CGTCTCGAAT GGCATCGAAA CGAACGGAGA 720
TGGAGAGATC AGTGGTGATT GGATTACGTC GAAGTTCTTG ACTACTCCGA GGATGTCATC 780
CTACTTGTTG GCAGTTATGG TATCAGAATT TGAATTTATC GAGGGTAAAA CAAAGACAGA 840
TGTTCGGTTC CGTATATGGT CACGCCCAGA GGCCAAGAAG ATGACAAAAC TTGCTTTGGA 900
TTATGGTATC AAATGCATAG AGTTCTACGA AGATTTCTTT GATATCAGAT TCCCCTTAAA 960
GAAACAAGAT ATGATCGCCC TTCCTGATTT CTCAGCAGGA GCCATGGAGA ACTGGGGTCT 1020
TATCACTTAC AGGGAAAACC CTTTGTTGTA CGATGACAGA TTCTATGCAC CGATGAATAA 1080
ACAGCGAATT GCTCGCATTG TTGCTCATGA GCTTGCCCAT CAGTGGTTTG GCGACTTGGT 1140
TACGATGAAG TGGTGGGATA ATCTGTGGTT GAATGAAGGT TTTGCAAGAT TCACAGAATT 1200
CATTGGAGCT GGTAAGATAA CTGAAGATGA CGCCAGAATG AGGAACTACT TCCTGATTGA 1260
TGTACTTGAA CGCGCGTTGA AAGCTGATTC CGTAGCGTCA AGCCATCCAC TTTCCTTCAG 1320
AATCGACAAA GCTGCAGAAG TTGAAGAAGC GTTTGATGAT ATCACATACG CCAAAGGAGC 1380
TTCTGTTCTT ACGATGCTGA GAGCGTTGAT CGGAGAAGAA AAACATAAGC ATGCGGTATC 1440
GCAGTATCTC AAGAAGTTCT CGTATAGCAA TGCAGAAGCG ACTGATCTAT GGGCAGTTTT 1500
CGATGAAGTT GTCACTGATG TCGAGGGTCC AGACGGCAAA CCTATGAAAA CCACGGAATT 1560
CGCAAGTCAG TGGACAACTC AGATGGGCTT TCCAGTAATT TCTGTGGCAG AGTTTAACTC 1620
GACTACTCTG AAACTAACGC AAAGTCGATA TAAGGCGAAT AAGGACGCTG TTGAGAAAGA 1680
GAAATACCGT CATCCGAAAT ACGGATTCAA GTGGGATATT CCATTGTGGT ATCAGGAAGG 1740
CGATAAGAAG GAGGTAAAGC GAGCATGGTT AAGAAGAGGT GAACCGCTTT ACTTGCATGT 1800
GAGTGATCCT GGCGCTCCAT TTGTGGTGAA TGCGGACCGC TATGGATTTT ACCGACAAAA 1860
CCACGACACT AATGGTTGGA AAAAGATAAT CAAGCAGCTC AAGGATAATC ATGAGGTTTA 1920
CAGTCCCCGG ACAAGGAATG CCATCATTAG CGATGCGTTT GCTGCGGCTG CAACTGACGC 1980
GATTGAGTAC GAGACTGTAT TCGAACTTCT GAAATATGCC GAAAAAGAAA CGGAATACCT 2040
ACCGTTGGAA ATAGCAATGT CTGGAATCTC TTCGATTTTG AAGTACTTCG GTACCGAGCC 2100
CGAGGCAAAG CCAGCTCAAA CATACATGAT GAACATATTG AAGCCGATGT ATGAGAAAAG 2160
CGATATCGAC TTCATTGCCA AAAACTACAA GGACGACAAG CTGTTTTTCC AAATCAACCT 2220
CCAAAAAGAT GTCATTGATA TGTTCTGCGC CCTTGGATCG CAAGACTGCA GGAAGAAATA 2280
TAAAAAACTT TTCGATGACA AAGTCATGGC GAAATGCAGG GATGGCCAAG CAGCAACCGA 2340
ATGCGTGAAA ATCGCCGCTC CTCTCCGATC AAGTGTTTAT TGTTATGGTG TGAAGGAAGG 2400
CGGTGATTAT GCTTTCGACA AGGTGATGGA GCTTTATACG GCCGAAACAC TCGCCCTAGA 2460
AAAAGACTTC CTACGCCTAG CATTAGGATG TCACAAAGAT GTTACCGCTT TGAAAGGACT 2520
TCTCCTGCGG GCTCTGGACA GGAATTCGTC GTTCGTACGA ATGCAGGATA TCCCAAGTGC 2580
TTTCAACGAT GTAGCAGCAA ATCCTATTGG CGAAGAATTC ATTTTCAATT TCCTTATTGA 2640
GAGATGGCCA GATATCGTTG AAAGTATAGG AACGAAACAC ACATATGTTG AAAAAGTGAT 2700
ACCAGCTTGC ACTTCAGGAA TCCGCTCACA ACAACAGATT GACCAGCTGA AGAATCTGCA 2760
GAAAAATGGC ATAAACGCTC GTCAATTTGG TGCATTCGAT AAAGCGATCG AACGAGCACA 2820
AAATAGGGTG GATTGGATTA AAAAACATTT CCAAAAATTA GCGGCTTTCT TCAAGAAAGC 2880
CACCTTGTAA TTCGTATTAC ATCACCATGA ATCCAGATCT TAAAACTCAC TAAGGAATGT 2940
GTGGGAACTG ACTGTCTGTT GCTTACTGTT CCACTGAATG GAAGTTTTTA CCCATAAAAA 3000
TTTTTTACCAT TTGCCTTCCG TGAGGGGTCA TTGTTGTCAC TTGAATAGTA AACAAAGCTC 3060
AGTATTGGAC CCAGTGATCA ATATTACTTT CGCTTCATCG AATTGTTACC TTCTCTATAC 3120
CCTCGTCCTA CCTGAATTCA CACATTTGTT CATATTTGTT TGTAGTCTGT TGCTCAGAAC 3180
ACTTTCTCTT CGATAGCTTT TTGTTTGTTT TTCTTTTGAT TGTATTGATC GTTTTACAAT 3240
TGTATAGATT AGTTATCTGA TAAATATTGA TGGCTAAGGA AAAAAAAAAA AAAAAAAAAA 3300
AAAAAAAAAA AAAAAAAAA 3319

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 amino acids
TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Asp Asn Ser Pro Ser Ala Glu Glu Leu Leu Leu Pro Thr Asn Ile
1 5 10 15

Lys Pro Val Ser Tyr Asp Leu Lys Ile Ala Thr Tyr Leu Pro Gly
20 25 30

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids
TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Leu Tyr Leu Ala Glu Asp Val Gln Leu Xaa Lys Gly Ile Leu Phe
1 5 10 15

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids
TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Leu Ala Tyr Asp Glu Lys Ser Tyr Ala Pro Asp Asn Lys Gln Tyr
1 5 10 15

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 3084 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 19:

GGTTTAATTA CCCAAGTTTG AGATGACAGC AGAGGAGAGT CAGGAGCAGG AGACGCAGCA 60
ACCACGAAAA AATACAGTGC TACGGCTCAC CCCAATCAAG TCTCTCTTTG CTTTGTTAGT 120
GGTAGCTGCT GCCGTCGGCC TCTCAATCGG TCTCACCTAT TACTTTACAA GGAAAGCTTT 180
TGATACTACT GGCGGAAATG GAAAAGGGGA TCAACCTATT GTCGATGATA ATTCCCCATC 240
AGCTGAAGAA TTACGTCTCC CAACAACCAT AAAACCTTTG ACATACGACT TAGTAATCAA 300
AACGTATCTG CCAAACTATG TAAACTATCC ACCTGAGAAA GATTTCGCTA TTGATGGGAC 360
TGTGGTGATT GCTATGGAAG TTGTGGAGCC AACAAAGTCT ATTGTGCTCA ACTCGAAAAA 420
TATTCCTGTA ATTGCAGACC AGTGCGAACT GTTTTCTAAC AACCAAAAAC TCGACATCGA 480
AAAGGTTGTG GATCAGCCAA GGCTGGAGAA AGTCGAATTC GTTTTGAAGA AAAAGCTGGA 540
GAAGAATCAG AAAATCACGC TCAAGATTGT ATACATTGGC CTTATCAACG ACATGCTTGG 600
AGGACTTTAT CGAACAACCT ACACGGATAA AGATGGTACA ACCAAGATTG CTGCATGCAC 660
TCATATGGAA CCGACGGACG CCCGTCTTAT GGTCCCCTGT TTCGACGAGC CGACGTTTAA 720
GGCAAACTGG ACTGTGACTG TGATTCATCC GAAGGGCACC AGTGCCGTGT CGAATGGAAT 780
AGAAAAGGGA GAAGGAGAAG TCTCTGGCGA TTGGGTCACA ACCAGATTCG ATCCAACCCC 840
GCGAATGCCT TCGTATTTGA TTGCTCTTGT GATTTCCGAA TTTAAGTACA TTGAAAATTA 900
TACGAAAAGC GGTGTTCGAT TCCGAATTCC GGCTCGTCCG GAAGCTATGA AGATGACAGA 960
ATATGCCATG ATAGCTGGAA TCAAATGTTT GGATTACTAT GAGGACTTCT TCGGGATCAA 1020
ATTCCCACTT CCAAAACAAG ATATGGTTGC TCTTCCTGAC TTCTCATCTG GTGCTATGGA 1080
GAACTGGGGT CTCATCACAT ACAGGGAGGG TTCCGTGCTC TACGATGAAA ACCTCTACGG 1140
ACCAATGAAT AAGGAGCGGG TTGCAGAAGT GATCGCGCAC GAACTTGCAC ATCAGTGGTT 1200
CGGTAATTTG GTCACGATGA AGTGGTGGGA TAACCTATGG CTGAACGAAG GATTCGCGTC 1260
ATTCGTGGAA TACATCGGAG CCGACTTCAT CAGCGATGGT CTATGGGAAA TGAAAGATTT 1320
CTTCCTGCTG GCACCGTACA CAAGTGGTAT TACGGCTGAT GCAGTAGCTT CAAGCCATCC 1380
GCTTTCCTTC AGAATAGATA AGGCTGCAGA TGTATCAGAA GCGTTCGATG ATATCACATA 1440
CCGTAAAGGA GCATCCGTTC TTCAAATGCT ATTGAATTTA GTTGGGGACG AAAATTTCAA 1500
GCAGTCTGTT TCGCGTTACC TCAAGAAGTT TTCATATGAT AATGCGGCTG CTGAAGATTT 1560
ATGGGCAGCA TTCGACGAAA CCGTCCAAGG TATAACCGGA CCTAATGGTG GACCATTGAA 1620
AATGTCCGAG TTTGCGCCAC AATGGACAAC TCAGATGGGG TTCCCTGTTC TTACTGTCGA 1680
GTCGGTTAAC GCAACGACTT TGAAAGTCAC CCAAAAACGA TACAGGCAGA ACAAGGATGC 1740
AAAGGAACCA GAGAAGTACC GTCATCCAAC TTATGGGTTC AAATGGGATG TTCCTCTGTG 1800
GTATCAGGAA GATGAACAGC AAGTGAAAAG AACTTGGTTA AAAAGAGAGG AACCGCTCTA 1860
TTTCCATGTA AGCAATTCTG ATTCGTCAGT TGTGGTGAAT GCCGAACGTC GTGCTTTTTG 1920
CCGATCAAAC TATGACGCTA ACGGTTGGAG GAACATTATG AGAAGACTCA AGCAGAATCA 1980
TAAGGTCTAT GGTCCACGAA CAAGAAACGC TCTCATAAGT GATGCGTTTG CAGCAGCTGC 2040
AGTTGAGGAA ATGAATTACG AGACCGTATT TGAAATGCTC AAATACACCG TGAAAGAAGA 2100
GGATTACTTA CCATGGAAGG AGGCAATATC AGGATTCAAT ACAATTTTGG ACTTTTTCGG 2160
CAGCGAACCC GAATCTCAAT GGGCTTCGGA ATACATGCGA AAACTGATGA AGCCAATTTA 2220
TGACAAGAGT AGCATCAAGT TTATAGCGGA GAACTACAAA AAAGATTCGC TTTTCTTCAA 2280
AAATAATCTC CAAATAGCTG TTATTGACAC ATACTGTGGT CTTGGAGGCA AAGAATGTCT 2340
TGAAGAAATG AAAAAGCTTT TTGACAAGGA GGTCATGAAA TGTCAACCTG GTCAGCAAGC 2400
GACCGACTGC GTAAAGGTAA CTGCTCCTCT CCGAAAAACT GTTTACTGCT ATGGGGTCCA 2460
GGAAGGCGGT GATGAGGCAT TCGACAAGGT GATGGAACTA TATAATGCGG AACAAGTGCA 2520
GTTGGAGAAA GACAGTCTAC GTGAAGCATT GGGATGCCAT AAAGACGTTA CAGCTCTAAA 2580
GGGACTTCTT ATGCTGGCTT TGGATCGGAA TTCGTCATTT GTGCGTCTTC AAGATGCTCA 2640
TGATGTGTTT AACATTGTAT CCAGAAATCC TGTTGGAAAC GAACTGCTGT TCAATTTCCT 2700
CACAGAGCGA TGGGAAGAGA TACTTGAAAG TTTGTCAATA CGACACAGAT CAGTTGATCG 2760
AGTGATCAAA GCCTGTACTC GAGGACTACG ATCCAGGGAA CAAGTACAAC AGTTGAAGAA 2820
TCTATACAAA AATGACAAGC GTGCTCGCGA ATACGGTGCA TTTGGTGGGG CAATAGAAAG 2880
ATCGGAACAC AGAGTCAAAT GGATTGAGAA ACATTTCCGA AAACTAGCAG CTTTCTTCAA 2940
AAAATCTAAT TCATAATTCT GAAATGGCTA TAACTAGCAC ACTGGATAGT TGTCTCGAAT 3000
CATCCAAAAA GATTAATGAT GTTTTTTTAC TAGATAATAT GGAGATATTC TGTAAATTTG 3060
TCATCGATTC AAGTGTCTGT ATTG 3084

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 3358 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 20:

GGTTTAATTA CCCAAGTTTG AGATGACGGC GGAGTGGCAG AAGCGTCGAA TCTTGGGCTT 60
CTCACCTATC AGCCTACTTT GTACATTATT TGTATTAGCT GCTGCCGTTG GACTCTCCAT 120
TGGTCTTACC TATTACTTCA CTCGTAAAGC ATTCGATACC ACACAAAAAG AACAGAAGGA 180
TGACAGTGGT GGTAAAGAAA AGGATAATTC TCCTTCTGCA GAAGAACTAC TTCTTCCAAC 240
GAACATAAAA CCAGTCTCGT ACGACTTGAA CATCAAAACA TATCTACCGG GTTACGTGAA 300
CTTTCCACCA GAAAAGAATC TCACATTTGA TGCCCATGTG GAGATTGCTA TGGTTGTGGT 360
TGAGCCTACA AATAGTATTG TGCTGAACTC GAAGAAAATC ACTTTGGCAC AAGGAGGATG 420
CGAACTGTTC TCAGGTAATC AGAAACTTGA CATCGAAAGT GTAAAGATGC AGGAAAGACT 480
TGACAAGCTT GAGATTACCC TCAAAAATCA GCTGCAAAAA GATCTGAAAA TCCTGCTCAA 540
GATCACTTAC ACCGGCCTTA TTAGCGACAC TCTCGGTGGG CTCTACCAGT CCATCTACAC 600
TGATAAGGAC GGAAAAACTA AGATCGTTGC TGTTTCACAA AATGAACCAT CAGACGCTCG 660
TCGTATAGCG CCATGCTTTG ACGAACCGAA GTACAAGGCA ACATGGACTG TCACCGTCGT 720
TCATCCCAAA GGTACAAAGG CTGCATCGAA CGGCATTGAA GCAAATGGAA AAGGGGAGCT 780
CAAGGGTGAT TGGATAACGT CTAAATTTAA AACTACCCCA CCGATGTCGT CCTATTTATT 840
GGCTATTATT GTTTGTGAAT TTGAATACAT TGAAGGATTT ACAAAAACGG GTGTTCGGTT 900
CCGTATATGG TCTCGACCAG AGGCGAAACG AATGACGGCA TACGCTTTGG ATGCTGGCAT 960
CAGATGCCTG GAGTTCTATG AGAAGTTCTT TGACATAAAA TTCCCTCTGG AAAAACAAGA 1020
TATGATTGCT CTTCCTGATT TCACCGCTGG TGCCATGGAA AACTGGGGCC TTATCACTTA 1080
TAGAGAGGAT TCTCTCCTAT ACGATGAAAA AATTTATGCA CCGATGAATA AACAGCGGGT 1140
TGCTCTCGTA GTTGCTCACG AGCTTGCTCA TCAGTGGTTC GGCAATCTGG TCACACTGAA 1200
GTGGTGGGAT GATACGTGGT TGAACGAAGG TTTTGCAACA TTTGTTGAGT ATCTTGGAAT 1260
GGACGAAATT AGCCACAACA ATTTCAGAAC GCAAGATTTC TTCTTGCTCG ATGGAATGGA 1320
TCGCGGAATG AGAGCTGACT CGGCAGCATC GAGCCATCCG CTTTCGTTTA GGATTGACAA 1380
AGCGGCAGAA GTTGCCGAAG CCTTTGACGA TATTTCATAC GCCAAGGGAG CGTCAGTTCT 1440
CACTATGCTA CGGGCTTTGA TTGGAGAGGA CAATTACAGG AATGCTGTTG TGCAATACCT 1500
CAAGAAGTTC TCCTACAGCA ATGCACAAGC AGCCGATCTG TGGAACGTCT TCAATGAAGT 1560
TGTCAAAGGT GTTAAGGGTC CTGACGGCAA CGTCATGAAA ATCGACCAAT TTACCGATCA 1620
GTGGACGTAT CAGATGGGTT ATCCTGTGGT TAAAGTAGAA GAATTTAATG CGACCGCCCT 1680
AAAGGTTACG CAGAGCCGGT ACAAGACAAA TAAAGACGCC TTGGAACCAG AGAAATATCG 1740
TAATCCAAAA TACGGGTTCA AGTGGGATGT TCCCCTATGG TATCAGGAAG GCAATAGCAA 1800
AGAGGTGAAG CGAACATGGC TAAAAAGAGA TGAACCGCTG TACTTGAACG TCAACAATCG 1860
GGATACATCC CTTGTGGTGA ACGCTGATCG ACATGGATTT TATCGACAAA ACTATGATGC 1920
CAACGGTTGG AAAAAGATAA TCAAGCAGCT CAAGAAAGAT CACAAGGTCT TCGGTCCAAG 1980
GACAAGGAAC GCTATCATAA GCGATGCATT TGCTGCAGCT ACGATTGACG CAATCGACTA 2040
TGAAACTGTA TTCGAACTAC TTGAATATGC CAAAAATGAA GAGGAATTCT TGCCTTGGAA 2100
GGAAGCTCTG TCCGGCATGT TCGCAGTTTT AAAGTTCTTC GGTAATGAGC CGGAGACAAA 2160
ACCAGCTAGA GCTTACATGA TGAGCATATT AGAACCGATG TATAATAAGA GCAGCATTGA 2220
TTACATCGTC AAGAATTATT TGGATGATAC GTTATTCACA AAAATTAATA CTCAAAAGGA 2280
TATCATTGAT GCATATTGTT CCCTTGGATC AAAGGACTGT ATAAAGCAAT ATAAGGATAT 2340
CTTCTACGAT GAGGTTATGC CCAAGTGTAA GGCCGGGGAA GCAGCAACCA AATGCGTTAA 2400
GGTTTCCGCT CCTCTTCGAG CCAATGTTTA CTGTTATGGT GTACAGGAAG GTGGTGAAGA 2460
AGCTTTTGAA AAGGTGATGG GGCTGTATCT AGCAGAAGAT GTTCAACTGG AGAAGGGTAT 2520
CCTGTTCAAA GCCTTGGCAT GCCACAAAGA TGTTACAGCT CTAAAAGAAC TTCTTTTGCG 2580
AGCCCTGGAC CGTAAATCGT CGTTTGTGCG TCTTCAGGAT GTCCCTACCG CTTTCCGTGC 2640
TGTATCTGAA AACCCTGTGG GCGAAGAATT CATGTTCAAT TTCCTAATGG AGAGATGGGA 2700
GGAAATCACT GCGAGCTTGG AAACAGAACA CAGAGCAGTT GATAAAGTGG TCGGCGCTTG 2760
TTGCACAGGA ATTCGCTCCC AACAACAAAT AGATCAGCTG AAGAATCTAC AGAAGAACAA 2820
TGCGCAGGCT AAGAAGTTCG GCTCATTCAC CCAGGAAATC GAAAAAGGAG AACATAAAAT 2880
TGCCTGGATC AAGAAACATT TTCACAGATT ATCGGAATTC TTCAAGAGAG CAAGATCATA 2940
GCTTTTCACA CTGAGCTCCA ATTTTAACGT CTTCAAACTA GGAGACAGTT TTGCTGAAAA 3000
GTCAGTTTCA CATTTTCCGT TTGAATGCCA TCCATTCGAA TACAACCAAT AATACCATTT 3060
TAAGTACCTT TCATTCACAG TGATTACTGA ATTTCGAATA TATCATGAAG CTTGTATCTT 3120
GAACGTTATG ATCGGTGACT TTCAATTTAT AGAGCTCACT CTCCATTTTG TAGCTGTGAT 3180
GACTTGCATT TAAGACCCAC CATTTACCAG CCTAGAATCT TTCCCCAATA CATTCCAAAC 3240
TCCGATCACC TCCACCGCTG ACAATGCCCA GATTTGTTTT TTTGTCTGCT ATCCATCTAA 3300
CTGTTTCGAT CGCCGGTTGT TTGTCAATTG CTTATCTGAT AAATATTGAC GTTGGTGT 3358

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:
LENGTH: 3369 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GGTTTAATTA CCCAAGTTTG AGGGTCTCCA TCTAGATGAC GTCGCAGGGG AGAACGCGGA 60
CATTGCTGAA TCTAACTCCA ATCCGTCTTA TTGTCGCATT ATTTCTAGTA GCTGCTGCAG 120
TCGGCCTCTC TATTGGTCTC ACCTATTACT TTACTCGCAA AGCGTTCGAT ACCTCAGAAA 180
AGCCAGGGAA GGATGATACT GGTGGCAAGG ACAAAGACAA TTCTCCCTCT GCGGCGGAAC 240
TACTCCTTCC AAGTAATATA AAACCATTGT CTTACGACTT GACGATCAAA ACATATCTAC 300
CTGGTTATGT GGACTTCCCA CCGGAGAAAA ACCTCACATT CGATGGGCGT GTGGAAATAT 360
CAATGGTTGT AATTGAGCCA ACAAAGAGTA TCGTACTCAA TTCAAAGAAG ATCTCTGTAA 420
TACCCCAAGA ATGTGAACTG GTATCGGGCG ATAAAAAAACT CGAAATTGAA AGTGTAAAGG 480
AGCACCCAAG ACTGGAAAAG GTTGAGTTTC TTATCAAAAG CCAACTGGAA AAAGATCAAC 540
AAATCTTGCT CAAGGTCGGC TACATCGGTC TCATCAGCAA CAGCTTTGGT GGAATCTACC 600
AGACCACTTA TACCACCCCG GATGGCACCC CTAAGATCGC TGCAGTTTCA CAAAATGAGC 660
CCATAGATGC TCGTCGAATG GTACCATGCA TGGATGAACC GAAATACAAA GCAAACTGGA 720
CCGTTACTGT CATTCATCCA AAAGGCACCA AAGCCGTCTC GAATGGAATC GAAGTGAACG 780
GAGATGGAGA GATCAGTGGT GATTGGATCA CATCGAAGTT CTTGACTACT CCACGGATGT 840
CATCCTACTT GTTGGCAGTT ATGGTTTCAG AATTTGAATA CATCGAAGGT GAAACAAAGA 900
CGGGTGTTCG GTTCCGTATA TGGTCACGCC CAGAGGCAAA GAAGATGACA CAATATGCTC 960
TGCAATCTGG TATCAAGTGC ATAGAATTCT ACGAAGATTT CTTTGATATC AGATTCCCTC 1020
TGAAGAAACA AGATATGATT GCCCTTCCTG ATTTCTCTGC CGGTGCCATG GAGAATTGGG 1080
GCCTCATCAC TTACAGGGAA AACTCTTTGT TGTACGATGA CAGATTCTAT GCACCGATGA 1140
ATAAACAGCG AATTGCTCGC ATTGTTGCTC ATGAGCTTGC TCATCAGTGG TTCGGCGACT 1200
TGGTTACGAT GAAGTGGTGG GATAATTTGT GGTTGAATGA AGGTTTTGCA AGATTCACAG 1260
AATTTATTGG AGCTGGTCAG ATAACTCAAG ATGACGCCAG AATGAGGAAC TACTTCCTGA 1320
TTGATGTACT TGAACGCGCT TTGAAAGCTG ATTCGGTAGC GTCAAGCCAT CCACTTTCCT 1380
TCAGAATCGA CAAAGCTGCA GAAGTTGAAG AAGCCTTTGA TGATATCACA TACGCCAAAG 1440
GAGCTTCTGT TCTTACTATG CTGAGAGCCT TGATTGGAGA AGAAAAACAT AAGCATGCAG 1500
TATCGCAGTA CCTCAAGAAG TTCTCGTATA GCAATGCAGA AGCGACTGAT CTATGGGCAG 1560
TTTTTGATGA AGTTGTCACT GACGTCGAAG GTCCAGACGG CAAACCTATG AAAACCACAG 1620
AGTTTGCAAG TCAGTGGACG ACTCAGATGG GCTTCCCAGT TATTTCCGTA GCAGAGTTTA 1680
ACTCGACTAC TTTGAAATTA ACGCAAAGTC GATATGAGGC GAATAAAGAC GCTGTGGAGA 1740
AAGAGAAGTA CCGTCACCCG AAATACGGAT TTAAATGGGA TATTCCACTG TGGTATCAGG 1800
AAGGCGATAA GAAGGAGATA AAGCGAACAT GGTTGAGAAG AGATGAACCG CTTTACTTGC 1860
ATGTTAGTGA TGCTGGCGCT CCCTTTGTGG TGAACGCAGA CCGCTATGGA TTTTATCGAC 1920
AAAATCATGA CGCTAATGGT TGGAAAAAGA TAATCAAGCA GCTCAAGGAT AATCATGAGG 1980
TTTACAGTCC CCGGACAAGG AATGTCATCA TTAGCGATGC GTTTGCTGCG GCTGCAACTG 2040
ACGCAATTGA GTATGAGACT GTATTTGAAC TTCTGAATTA TGCCGAAAAA GAAACGGAAT 2100
ATCTACCATT AGAAATCGCA ATGTCCGGGA TCTCTTCGAT TTTGAAATAC TTCCCTACCG 2160
AGCCAGAGGC AAAGCCAGCT CAAACATACA TGATGAACAT ATTGAAACCG ATGTATGAAA 2220
AAAGCAGTAT CGACTTCATT GCCAATAACT ACAGAAATGA CAAGCTGTTT TTCCAAATCA 2280
ACCTCCAAAA AGATGTCATT GATATGTTCT GCGCCCTCGG ATCGCAAGAC TGCAGGAAGA 2340
AATATAAAAA ACTTTTCGAT GACGAAGTCA TGAACAAATG CAGGGATGGT CAAGCAGCAA 2400
CCGAATGCGT AAGAATCGCC GCTCCTCTCC GATCAAGTGT TTATTGTTAT GGTGTGAAGG 2460
AAGGCGGTGA TTATGCTTCC GACAAGGTGA TGGAGCTTTA TACGGCCGAA ACACTCGCCC 2520
TAGAAAAAGA CTTCCTACGC CTAGCATTGG GATGTCATAA AGATGTTACT GCTTTGAAAG 2580
GACTTCTCTT GCGGGCTCTG GACAGGAATT CGTCGTTCGT ACGTATGCAG GATATCCCAA 2640
GTGCTTTCAA CGATGTAGCA GCAAATCCTA TTGGCGAAGA ATTCATTTTC AATTTCCTTA 2700
TTGAGAGATG GCCAGATATC ATTGAAAGTA TAGGAACGAA GCACACATAT GTTGAGAAAG 2760
TGATACCAGC CTGCACTTCA GGAATCCGCT CACAACAGCA GATTGACCAG CTGAAGAATC 2820
TGCAGAAAAA TGGCATGAAC GCTCGTCAAT TCGGTGCATT CGATAAAGCA ATCGAACGAG 2880
CACAAAATAG GGTGGATTGG ATTAAAAAAC ATTTCCAAAA ATTAGCGGCT TTCTTCAAGA 2940
AAGCCACCTT GTAATTCGAA TTACATTGCC AGTAATCCAG ATCTTAAAGT TCATGAAGGA 3000
ATATGACAGG GAACTGACTG TCTGTTGGTC ACTGTTCCAC TGAATGGAAG TTTTTACCTA 3060
CAAAAATTTT TATCGTTATA TTTGCCTTCC GTGAGGGGTC ATTGTTGTCA CTTGAATAGT 3120
AAACAAAGCT CAGTATTGGC AACCGTAGAA CAATATTACT TTCGCTTCAT CAAATTGTTA 3180
TCTTCCCTAT ACCCTCTTCC TAACTGAATT CGGAAATTTG TTCATATTCG TTTGTAGTCT 3240
GTTGCTCAGA ACACTTTCTC CTCAATAGCT TCTTGTTTGT TTTTTTTTTG ATTGTATTGA 3300
TCGTTTTACA ATTGTATAGA TTAGTTATCT TATAAATATT GATGGTTAAA AAAAAAAAAA 3360
AAAAAAAAA 3369

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 977 amino acids
TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 22:

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

Leu Ala Pro Tyr Thr Ser Gly Ile Thr Ala Asp Ala Val Ala Ser
440 445 450
Ser His Pro Leu Ser Phe Arg Ile Asp Lys Ala Ala Asp Val Ser
455 460 465
Glu Ala Phe Asp Asp Ile Thr Tyr Arg Lys Gly Ala Ser Val Leu
470 475 480
Gln Met Leu Leu Asn Leu Val Gly Asp Glu Asn Phe Lys Gln Ser
485 490 495
Val Ser Arg Tyr Leu Lys Lys Phe Ser Tyr Asp Asn Ala Ala Ala
500 505 510
Glu Asp Leu Trp Ala Ala Phe Asp Glu Thr Val Gln Gly Ile Thr
515 520 525
Gly Pro Asn Gly Gly Pro Leu Lys Met Ser Glu Phe Ala Pro Gln
530 535 540
Trp Thr Thr Gln Met Gly Phe Pro Val Leu Thr Val Glu Ser Val
545 550 555
Asn Ala Thr Thr Leu Lys Val Thr Gln Lys Arg Tyr Arg Gln Asn
560 565 570
Lys Asp Ala Lys Glu Pro Glu Lys Tyr Arg His Pro Thr Tyr Gly
575 580 585
Phe Lys Trp Asp Val Pro Leu Trp Tyr Gln Glu Asp Glu Gln Gln
590 595 600
Val Lys Arg Thr Trp Leu Lys Arg Glu Glu Pro Leu Tyr Phe His
605 610 615
Val Ser Asn Ser Asp Ser Ser Val Val Val Asn Ala Glu Arg Arg
620 625 630
Ala Phe Cys Arg Ser Asn Tyr Asp Ala Asn Gly Trp Arg Asn Ile
635 640 645
Met Arg Arg Leu Lys Gln Asn His Lys Val Tyr Gly Pro Arg Thr
650 655 660
Arg Asn Ala Leu Ile Ser Asp Ala Phe Ala Ala Ala Ala Val Glu
665 670 675
Glu Met Asn Tyr Glu Thr Val Phe Glu Met Leu Lys Tyr Thr Val
680 685 690
Lys Glu Glu Asp Tyr Leu Pro Trp Lys Glu Ala Ile Ser Gly Phe
695 700 705
Asn Thr Ile Leu Asp Phe Phe Gly Ser Glu Pro Glu Ser Gln Trp
710 715 720
Ala Ser Glu Tyr Met Arg Lys Leu Met Lys Pro Ile Tyr Asp Lys
725 730 735
Ser Ser Ile Lys Phe Ile Ala Glu Asn Tyr Lys Lys Asp Ser Leu
740 745 750
Phe Phe Lys Asn Asn Leu Gln Ile Ala Val Ile Asp Thr Tyr Cys
755 760 765
Gly Leu Gly Gly Lys Glu Cys Leu Glu Glu Met Lys Lys Leu Phe
770 775 780
Asp Lys Glu Val Met Lys Cys Gln Pro Gly Gln Gln Ala Thr Asp
785 790 795
Cys Val Lys Val Thr Ala Pro Leu Arg Lys Thr Val Tyr Cys Tyr
800 805 810
Gly Val Gln Glu Gly Gly Asp Glu Ala Phe Asp Lys Val Met Glu
815 820 825
Leu Tyr Asn Ala Glu Gln Val Gln Leu Glu Lys Asp Ser Leu Arg
830 835 840
Glu Ala Leu Gly Cys His Lys Asp Val Thr Ala Leu Lys Gly Leu
845 850 855
Leu Met Leu Ala Leu Asp Arg Asn Ser Ser Phe Val Arg Leu Gln
860 865 870
Asp Ala His Asp Val Phe Asn Ile Val Ser Arg Asn Pro Val Gly
875 880 885
Asn Glu Leu Leu Phe Asn Phe Leu Thr Glu Arg Trp Glu Glu Ile
890 895 900
Leu Glu Ser Leu Ser Ile Arg His Arg Ser Val Asp Arg Val Ile
905 910 915
Lys Ala Cys Thr Arg Gly Leu Arg Ser Arg Glu Gln Val Gln Gln
920 925 930
Leu Lys Asn Leu Tyr Lys Asn Asp Lys Arg Ala Arg Glu Tyr Gly

```
                935                 940                 945
Ala Phe Gly Gly Ala Ile Glu Arg Ser Glu His Arg Val Lys Trp
                950                 955                 960
Ile Glu Lys His Phe Arg Lys Leu Ala Ala Phe Phe Lys Lys Ser
                965                 970                 975
Asn Ser
```

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 972 amino acids
TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 23:

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

```
Val Val Ala His Glu Leu Ala His Gln Trp Phe Gly Asn Leu Val
                380                 385                 390
Thr Leu Lys Trp Trp Asp Asp Thr Trp Leu Asn Glu Gly Phe Ala
                395                 400                 405
Thr Phe Val Glu Tyr Leu Gly Met Asp Glu Ile Ser His Asn Asn
                410                 415                 420
Phe Arg Thr Gln Asp Phe Phe Leu Leu Asp Gly Met Asp Arg Gly
                425                 430                 435
Met Arg Ala Asp Ser Ala Ala Ser Ser His Pro Leu Ser Phe Arg
                440                 445                 450
Ile Asp Lys Ala Ala Glu Val Ala Glu Ala Phe Asp Asp Ile Ser
                455                 460                 465
Tyr Ala Lys Gly Ala Ser Val Leu Thr Met Leu Arg Ala Leu Ile
                470                 475                 480
Gly Glu Asp Asn Tyr Arg Asn Ala Val Val Gln Tyr Leu Lys Lys
                485                 490                 495
Phe Ser Tyr Ser Asn Ala Gln Ala Ala Asp Leu Trp Asn Val Phe
                500                 505                 510
Asn Glu Val Val Lys Gly Val Lys Gly Pro Asp Gly Asn Val Met
                515                 520                 525
Lys Ile Asp Gln Phe Thr Asp Gln Trp Thr Tyr Gln Met Gly Tyr
                530                 535                 540
Pro Val Val Lys Val Glu Glu Phe Asn Ala Thr Ala Leu Lys Val
                545                 550                 555
Thr Gln Ser Arg Tyr Lys Thr Asn Lys Asp Ala Leu Glu Pro Glu
                560                 565                 570
Lys Tyr Arg Asn Pro Lys Tyr Gly Phe Lys Trp Asp Val Pro Leu
                575                 580                 585
Trp Tyr Gln Glu Gly Asn Ser Lys Glu Val Lys Arg Thr Trp Leu
                590                 595                 600
Lys Arg Asp Glu Pro Leu Tyr Leu Asn Val Asn Asn Arg Asp Thr
                605                 610                 615
Ser Leu Val Val Asn Ala Asp Arg His Gly Phe Tyr Arg Gln Asn
                620                 625                 630
Tyr Asp Ala Asn Gly Trp Lys Lys Ile Ile Lys Gln Leu Lys Lys
                635                 640                 645
Asp His Lys Val Phe Gly Pro Arg Thr Arg Asn Ala Ile Ile Ser
                650                 655                 660
Asp Ala Phe Ala Ala Ala Thr Ile Asp Ala Ile Asp Tyr Glu Thr
                665                 670                 675
Val Phe Glu Leu Leu Glu Tyr Ala Lys Asn Glu Glu Glu Phe Leu
                680                 685                 690
Pro Trp Lys Glu Ala Leu Ser Gly Met Phe Ala Val Leu Lys Phe
                695                 700                 705
Phe Gly Asn Glu Pro Glu Thr Lys Pro Ala Arg Ala Tyr Met Met
                710                 715                 720
Ser Ile Leu Glu Pro Met Tyr Asn Lys Ser Ser Ile Asp Tyr Ile
                725                 730                 735
Val Lys Asn Tyr Leu Asp Asp Thr Leu Phe Thr Lys Ile Asn Thr
                740                 745                 750
Gln Lys Asp Ile Ile Asp Ala Tyr Cys Ser Leu Gly Ser Lys Asp
                755                 760                 765
Cys Ile Lys Gln Tyr Lys Asp Ile Phe Tyr Asp Glu Val Met Pro
                770                 775                 780
Lys Cys Lys Ala Gly Glu Ala Ala Thr Lys Cys Val Lys Val Ser
                785                 790                 795
Ala Pro Leu Arg Ala Asn Val Tyr Cys Tyr Gly Val Gln Glu Gly
                800                 805                 810
Gly Glu Glu Ala Phe Glu Lys Val Met Gly Leu Tyr Leu Ala Glu
                815                 820                 825
Asp Val Gln Leu Glu Lys Gly Ile Leu Phe Lys Ala Leu Ala Cys
                830                 835                 840
His Lys Asp Val Thr Ala Leu Lys Glu Leu Leu Leu Arg Ala Leu
                845                 850                 855
Asp Arg Lys Ser Ser Phe Val Arg Leu Gln Asp Val Pro Thr Ala
                860                 865                 870
Phe Arg Ala Val Ser Glu Asn Pro Val Gly Glu Glu Phe Met Phe
```

```
                    875                 880                 885
Asn Phe Leu Met Glu Arg Trp Glu Glu Ile Thr Ala Ser Leu Glu
                    890                 895                 900
Thr Glu His Arg Ala Val Asp Lys Val Val Gly Ala Cys Cys Thr
                    905                 910                 915
Gly Ile Arg Ser Gln Gln Gln Ile Asp Gln Leu Lys Asn Leu Gln
                    920                 925                 930
Lys Asn Asn Ala Gln Ala Lys Lys Phe Gly Ser Phe Thr Gln Glu
                    935                 940                 945
Ile Glu Lys Gly Glu His Lys Ile Ala Trp Ile Lys Lys His Phe
                    950                 955                 960
His Arg Leu Ser Glu Phe Phe Lys Arg Ala Arg Ser
                965                 970
```

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 972 amino acids
TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 24:

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

```
                320                 325                 330
Lys Gln Asp Met Ile Ala Leu Pro Asp Phe Ser Ala Gly Ala Met
                335                 340                 345
Glu Asn Trp Gly Leu Ile Thr Tyr Arg Glu Asn Ser Leu Leu Tyr
                350                 355                 360
Asp Asp Arg Phe Tyr Ala Pro Met Asn Lys Gln Arg Ile Ala Arg
                365                 370                 375
Ile Val Ala His Glu Leu Ala His Gln Trp Phe Gly Asp Leu Val
                380                 385                 390
Thr Met Lys Trp Trp Asp Asn Leu Trp Leu Asn Glu Gly Phe Ala
                395                 400                 405
Arg Phe Thr Glu Phe Ile Gly Ala Gly Gln Ile Thr Gln Asp Asp
                410                 415                 420
Ala Arg Met Arg Asn Tyr Phe Leu Ile Asp Val Leu Glu Arg Ala
                425                 430                 435
Leu Lys Ala Asp Ser Val Ala Ser Ser His Pro Leu Ser Phe Arg
                440                 445                 450
Ile Asp Lys Ala Ala Glu Val Glu Glu Ala Phe Asp Asp Ile Thr
                455                 460                 465
Tyr Ala Lys Gly Ala Ser Val Leu Thr Met Leu Arg Ala Leu Ile
                470                 475                 480
Gly Glu Glu Lys His Lys His Ala Val Ser Gln Tyr Leu Lys Lys
                485                 490                 495
Phe Ser Tyr Ser Asn Ala Glu Ala Thr Asp Leu Trp Ala Val Phe
                500                 505                 510
Asp Glu Val Val Thr Asp Val Glu Gly Pro Asp Gly Lys Pro Met
                515                 520                 525
Lys Thr Thr Glu Phe Ala Ser Gln Trp Thr Thr Gln Met Gly Phe
                530                 535                 540
Pro Val Ile Ser Val Ala Glu Phe Asn Ser Thr Thr Leu Lys Leu
                545                 550                 555
Thr Gln Ser Arg Tyr Glu Ala Asn Lys Asp Ala Val Glu Lys Glu
                560                 565                 570
Lys Tyr Arg His Pro Lys Tyr Gly Phe Lys Trp Asp Ile Pro Leu
                575                 580                 585
Trp Tyr Gln Glu Gly Asp Lys Lys Glu Ile Lys Arg Thr Trp Leu
                590                 595                 600
Arg Arg Asp Glu Pro Leu Tyr Leu His Val Ser Asp Ala Gly Ala
                605                 610                 615
Pro Phe Val Val Asn Ala Asp Arg Tyr Gly Phe Tyr Arg Gln Asn
                620                 625                 630
His Asp Ala Asn Gly Trp Lys Lys Ile Ile Lys Gln Leu Lys Asp
                635                 640                 645
Asn His Glu Val Tyr Ser Pro Arg Thr Arg Asn Val Ile Ile Ser
                650                 655                 660
Asp Ala Phe Ala Ala Ala Ala Thr Asp Ala Ile Glu Tyr Glu Thr
                665                 670                 675
Val Phe Glu Leu Leu Asn Tyr Ala Glu Lys Glu Thr Glu Tyr Leu
                680                 685                 690
Pro Leu Glu Ile Ala Met Ser Gly Ile Ser Ser Ile Leu Lys Tyr
                695                 700                 705
Phe Pro Thr Glu Pro Glu Ala Lys Pro Ala Gln Thr Tyr Met Met
                710                 715                 720
Asn Ile Leu Lys Pro Met Tyr Glu Lys Ser Ser Ile Asp Phe Ile
                725                 730                 735
Ala Asn Asn Tyr Arg Asn Asp Lys Leu Phe Phe Gln Ile Asn Leu
                740                 745                 750
Gln Lys Asp Val Ile Asp Met Phe Cys Ala Leu Gly Ser Gln Asp
                755                 760                 765
Cys Arg Lys Lys Tyr Lys Lys Leu Phe Asp Asp Glu Val Met Asn
                770                 775                 780
Lys Cys Arg Asp Gly Gln Ala Ala Thr Glu Cys Val Arg Ile Ala
                785                 790                 795
Ala Pro Leu Arg Ser Ser Val Tyr Cys Tyr Gly Val Lys Glu Gly
                800                 805                 810
Gly Asp Tyr Ala Ser Asp Lys Val Met Glu Leu Tyr Thr Ala Glu
                815                 820                 825
```

```
Thr Leu Ala Leu Glu Lys Asp Phe Leu Arg Leu Ala Leu Gly Cys
                830                 835                 840
His Lys Asp Val Thr Ala Leu Lys Gly Leu Leu Leu Arg Ala Leu
                845                 850                 855
Asp Arg Asn Ser Ser Phe Val Arg Met Gln Asp Ile Pro Ser Ala
                860                 865                 870
Phe Asn Asp Val Ala Ala Asn Pro Ile Gly Glu Glu Phe Ile Phe
                875                 880                 885
Asn Phe Leu Ile Glu Arg Trp Pro Asp Ile Ile Glu Ser Ile Gly
                890                 895                 900
Thr Lys His Thr Tyr Val Glu Lys Val Ile Pro Ala Cys Thr Ser
                905                 910                 915
Gly Ile Arg Ser Gln Gln Gln Ile Asp Gln Leu Lys Asn Leu Gln
                920                 925                 930
Lys Asn Gly Met Asn Ala Arg Gln Phe Gly Ala Phe Asp Lys Ala
                935                 940                 945
Ile Glu Arg Ala Gln Asn Arg Val Asp Trp Ile Lys Lys His Phe
                950                 955                 960
Gln Lys Leu Ala Ala Phe Phe Lys Lys Ala Thr Leu
                965                 970
```

Patenttivaatimukset:

1. Nukleiinihappomolekyylit, jotka käsittävät yhden tai useamman nukleotidisekvenssin, joka koodittaa/jotka koodittavat matoaminopeptidaasientsyymejä tai niiden antigeenisiä osia ja joka vastaa/jotka vastaavat kokonaisuudessaan tai osittain kuvioissa 2, 3, 4 tai 5 esitetyn mukaisia nukleotidisekvenssejä (sekvenssit tunnusnrot 1 - vastaavasti 15) tai sekvenssejä, jotka koodittavat matoaminopeptidaasientsyymejä, joilla on mainittuihin nähden 50 %:n tai yli sekvenssi-identtisyys, tai jotka hybridisoituvat mihin tahansa mainittuihin sekvensseihin olosuhteissa 2 x SSC, huoneenlämpötila (jossa SCC = 0,15 M NaCl, 0,015 M natriumsitraatti, pH 7,2).

2. Nukleiinihappomolekyylit, jotka käsittävät yhden tai useamman nukleotidisekvenssin, joka koodittaa/jotka koodittavat polypeptidiä, joka kykenee synnyttämään suojaavia vasta-aineita vastaan matoloisia, joihin sekvensseihin sisältyy yksi tai useampi antigeenimäärettä koodittava alue aminopeptidaasia koodittavista sekvensseistä, jotka esitetään kuvioissa 2, 3, 4 tai 5 (koostuvat sekvensseistä tunnusnrot 1 - 15).

3. Nukleiinihappomolekyylit, jotka käsittävät yhden tai useamman nukleotidisekvenssin, joka vastaa/jotka vastaavat yhtä tai useampaa sekvenssiä tai joka on/jotka ovat komplementaarisia yhteen tai useampaan sekvenssiin nähden, joka/jotka valitaan kloonien sekvensseistä, joilla on mainittuihin nähden 50 %:n tai yli sekvenssi-identtisyys kuvioissa 2, 3, 4 ja 5 esitetyn mukaisesti, tai sekvensseistä, joilla on mainittuihin nähden 50 %:n tai yli sekvenssi-identtisyys tai jotka hybridisoituvat mihin tahansa mainittuihin sekvensseihin olosuhteissa 2 x SSC, huoneenlämpötila (jossa SCC = 0,15 M NaCl, 0,015 M natriumsitraatti, pH 7,2).

4. Ilmentämis- tai kloonausvektori, joka käsittää jossain patenttivaatimuksessa 1 - 3 määritellyn mukaisen nukleiinihappomolekyylin.

5. Prokaryoottinen tai eukaryoottinen solu, tai siirtogeeninen organismi, joka sisältää jossain patenttivaatimuksessa 1 - 3 määritellyn mukaisen nukleiinihappomolekyylin.

6. Synteettiset polypeptidit, jotka käsittävät aminohapposekvenssin, joka muodostaa aminopeptidaasientsyymin tai sen antigeenisen osan, ja joka vastaa kokonaisuudessaan tai osittain kuviossa 2, 3, 4 tai 5 esitetyn mukaisia nukleotidisekvenssejä (sekvenssit tunnusnrot 1 - vastaavasti 15), tai mainitun funktionaalisesti ekvivalentin variantin, joka on muu kuin synteettinen polypeptidi, joka vastaa proteiinidublettia H110D, tai synteettinen polypeptidi, joka valitaan seuraavista:

(a) Met Gly Tyr Pro Val Val Lys Val Glu Glu Phe

(b) Met Gly Phe Pro Val Leu Thr Val Glu Ser

(c) Met Gly/Phe Asn Phe Lys Ile Glu/Val Thr/Glu Ala Gly

(d) Met Lys Pro/Glu Thr/Val Lys Asp/Ala Thr/Lys Leu - Ile Thr

(e) Met Leu Ala Leu Asp Tyr His Ser - Phe Val

(f) Met Leu Ala Glu/Tyr Asp Gln/Ala Glu Asp Val

(g) Met Gly Phe Pro Leu Val Thr Val Glu Ala Phe Tyr

(h) Met Lys Thr Pro Glu Phe Ala Val/Leu Gln Ala Phe/Thr Ala Thr Ser/Gly Phe Pro

(i) Lys His/Tyr Asn/Val Ser Pro Ala Ala Glu Asn/Leu Leu Asn/Gly

(j) Lys - Thr Ser Val Ala Glu Ala Phe Asn

(k) Lys Ala Ala Glu Val Ala Glu Ala Phe Asp - Ile - - - Lys Gly

(l) Lys Ala Val Glu Val/Pro Ala Glu Ala Phe Asp Asp Ile Thr? Tyr - - Gly Pro Ser

(m) Lys - Glu Glu Thr Glu Ile Phe Asn Met

(n) Lys - - - Pro Phe Asn/Asp Ile Glu Ala Leu

(o) Asp Gln Ala Phe Ser Thr Asp Ala Lys

(p) Met Gly Tyr Pro Val Val Lys Val Glu Glu Phe - Ala Thr Ala Leu

(q) Met Gly Phe Pro Val Leu Thr Val Glu Ser - Tyr? - Thr

(r) Met Glu/Phe Asn Phe Leu Ile Glu/Val Thr/Glu Ala Gly - Ile Thr

(s) Met Gly Phe Leu Val Thr Val Glu Ala Phe Tyr - Thr Ser

(t) Met Lys Thr Pro Glu Phe Ala Val/Leu Gln Ala Phe/Thr Ala Thr Ser/Gly Phe Pro

(u) Met Lys Pro/Glu Thr/Val Leu Asp/Ala Thr/Lys Leu - Ile Thr - Gly

(v) Met Leu Ala Leu Asp Tyr His Ser - Phe Val Gly?

(w) Met Leu Ala Glu/Tyr Asp Gln/Ala Glu Asp Val

(x) Lys His/Tyr Asn/Val Ser Pro Ala Ala Glu Asn/Leu Leu Asn/Gly

(y) Lys - Thr Ser Val Ala Glu Ala Phe Asn

(z) Lys Ala Ala Glu Val Ala Glu Ala Phe Asp - Ile - - - Lys Gly

(aa) Lys Ala Val Glu Val/Pro Ala Glu Ala Phe Asp Asp Ile Thr? Tyr - - Gly Pro Ser

(bb) Lys - Glu Gln Thr Glu Ile Phe Asn Met

(cc) Lys - - - Pro Phe Asn/Asp Ile Glu Ala Leu

(dd) Asp Gln Ala Phe Ser Thr Asp Ala Lys

7. Patenttivaatimuksen 6 mukaiset synteettiset polypeptidit, jotka käsittävät aminohapposekvenssin, joka muodostaa aminopeptidaasientsyymin tai sen antigeenisen osan, ja joka vastaa kokonaisuudessaan tai osittain ku-

viossa 2, 3, 4 tai 5 esitetyn mukaisia nukleotidisekvenssejä (sekvenssit tunnusnrot 1 - 15), tai mainitun funktionaalisesti ekvivalentin variantin, joka oleellisesti ei sisällä muita Haemonchus contortus -komponentteja.

8. Patenttivaatimuksen 6 tai patenttivaatimuksen 7 mukaiset synteettiset polypeptidit fuusiopolypeptidin muodossa, joka käsittää ylimääräisen polypeptidin fuusioituna mainittuun patenttivaatimuksessa 6 tai patenttivaatimuksessa 7 määritellyn mukaiseen aminohapposekvenssiin.

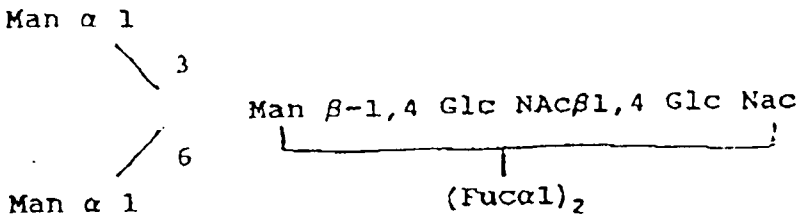
9. Menetelmä jossain patenttivaatimuksessa 6 - 8 määritellyn mukaisen synteettisen polypeptidin valmistamiseksi, jossa viljellään prokaryoottista tai eukaryoottista solua, joka sisältää jossain patenttivaatimuksessa 1 - 3 määritellyn mukaista nukleiinihappomolekyyliä, olosuhteissa, joissa mainittu polypeptidi ilmentyy, ja otetaan talteen näin tuotettu mainittu polypeptidi.

10. Rokotekoostumus immuunivasteiden vastaan matoloisia stimuloimiseksi ihmis- tai ei-ihmiseläimessä, joka käsittää ainakin yhtä jossain patenttivaatimuksessa 6 - 8 määritellyn mukaista synteettistä polypeptidiä, tai virusta tai isäntäsolua, johon on insertoitu jossain patenttivaatimuksessa 1 - 3 määritellyn mukaista nukleiinihappomolekyyliä, insertoidun nukleiinihappomolekyylin koodittamien polypeptidien immuunivasteen stimuloimiseksi, yhdessä farmaseuttisesti hyväksyttävän kantajan kanssa.

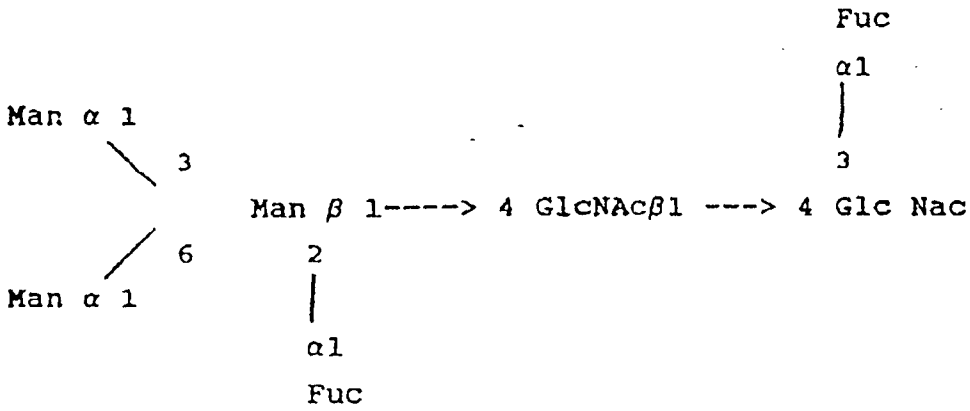
11. Jossain patenttivaatimuksessa 1 - 3 määritellyn mukaisen nukleiinihappomolekyylin, tai jossain patenttivaatimuksessa 6 - 8 määritellyn mukaisen synteettisen polypeptidin, käyttö rokotekoostumuksen valmistamiseksi immuunivasteen stimuloimiseksi vastaan matoloisia ihmis- tai ei-ihmiseläimessä.

12. Menetelmä immuunivasteen stimuloimiseksi vastaan matoloisia ihmis- tai ei-ihmiseläimessä, jossa mainitulle eläimelle annetaan patenttivaatimuksessa 10 määritellyn mukaista rokotekoostumusta.

13. Oligosakkaridi, jonka rakenne on:



14. Patenttivaatimuksen 13 mukainen oligosakkaridi, jonka rakenne on:



15. Patenttivaatimuksessa 13 tai patenttivaatimuksessa 14 määritellyn mukainen oligosakkaridi, joka on kytketty jossain patenttivaatimuksessa 6 - 8 määritellyn mukaiseen synteettiseen polypeptidiin.

Patentkrav:

1. Nukleinsyramolekyler, vilka omfattar en eller flera nukleotidsekvenser som kodar för mask-aminopeptidasenzymer eller antigeniska partier därav, vilka motsvarar hela eller en del av nukleotidsekvenserna som visas i fig. 2, 3, 4 eller 5 (SEQ ID nr 1 - 15) eller sekvenser som kodar för mask-aminopeptidasenzymer, vilka har en sekvens som är identisk med dessa i en grad av 50 % eller mera eller vilka hybridiserar med någon av dessa sekvenser under förhållanden av 2 x SSC, rumstemperatur (varvid SCC = 0,15 M NaCl, 0,015 M natriumcitrat, pH 7,2).

2. Nukleinsyramolekyler, vilka omfattar en eller flera nukleotidsekvenser som kodar för en polypeptid som förmår bilda skyddande antikroppar mot maskparasiter, varvid sekvenserna innehåller en eller flera antigeniska determinantkodande regioner från de aminopeptidaskodande sekvenserna som visats i fig. 2, 3, 4 eller 5 (sammansatta från SEQ ID nr 1 - 15).

3. Nukleinsyramolekyler, vilka omfattar en eller flera nukleotidsekvenser som motsvarar eller är komplementära till en eller flera sekvenser som valts bland klonsekvenser med en sekvensidentitet med dessa i en grad av 50 % eller mera såsom visats i fig. 2, 3, 4 och 5 eller sekvenser som har en sekvensidentitet med dessa i en grad av 50 % eller mera eller vilka hybridiserar med någon av nämnda sekvenser under förhållanden av 2 x SSC, rumstemperatur (vari SCC = 0,15 M NaCl, 0,015 M natriumcitrat, pH 7,2).

4. Expressions- eller kloningsvektor som omfattar en nukleinsyramolekyl enligt något av patentkraven 1 - 3.

5. Prokaryotisk eller eukaryotisk cell eller en transgenisk organism som innehåller en nukleinsyramolekyl enligt något av patentkraven 1 - 3.

6. Syntetiska polypeptider, vilka omfattar en aminosyrasekvens som bildar ett aminopeptidasenzym eller ett antigeniskt parti därav, och som motsvarar hela eller ett parti av nukleotidsekvenserna som visats i fig. 2, 3, 4 eller 5 (SEQ ID nr 1 - 15), eller en funktionellt ekvivalent variant därav annan än en syntetisk polypeptid som motsvarar proteindubletten H110D, eller en syntetisk polypeptid som valts bland

(a) Met Gly Tyr Pro Val Val Lys Val Glu Glu Phe

(b) Met Gly Phe Pro Val Leu Thr Val Glu Ser

(c) Met Gly/Phe Asn Phe Lys Ile Glu/Val Thr/Glu Ala Gly

(d) Met Lys Pro/Glu Thr/Val Lys Asp/Ala Thr/Lys Leu - Ile Thr

(e) Met Leu Ala Leu Asp Tyr His Ser - Phe Val

(f) Met Leu Ala Glu/Tyr Asp Gln/Ala Glu Asp Val

(g) Met Gly Phe Pro Leu Val Thr Val Glu Ala Phe Tyr

(h) Met Lys Thr Pro Glu Phe Ala Val/Leu Gln Ala Phe/Thr Ala Thr Ser/Gly Phe Pro

(i) Lys His/Tyr Asn/Val Ser Pro Ala Ala Glu Asn/Leu Leu Asn/Gly

(j) Lys - Thr Ser Val Ala Glu Ala Phe Asn

(k) Lys Ala Ala Glu Val Ala Glu Ala Phe Asp - Ile - - - Lys Gly

(l) Lys Ala Val Glu Val/Pro Ala Glu Ala Phe Asp Asp Ile Thr? Tyr - - Gly Pro Ser

(m) Lys - Glu Glu Thr Glu Ile Phe Asn Met

(n) Lys - - - Pro Phe Asn/Asp Ile Glu Ala Leu

(o) Asp Gln Ala Phe Ser Thr Asp Ala Lys

(p) Met Gly Tyr Pro Val Val Lys Val Glu Glu Phe - Ala Thr Ala Leu

(q) Met Gly Phe Pro Val Leu Thr Val Glu
Ser - Tyr? - Thr
(r) Met Glu/Phe Asn Phe Leu Ile Glu/Val Thr/Glu
Ala Gly - Ile Thr
(s) Met Gly Phe Leu Val Thr Val Glu Ala Phe
Tyr - Thr Ser
(t) Met Lys Thr Pro Glu Phe Ala Val/Leu Gln
Ala Phe/Thr Ala Thr Ser/Gly Phe Pro
(u) Met Lys Pro/Glu Thr/Val Leu Asp/Ala Thr/Lys
Leu - Ile Thr - Gly
(v) Met Leu Ala Leu Asp Tyr His Ser - Phe Val
Gly?
(w) Met Leu Ala Glu/Tyr Asp Gln/Ala Glu Asp Val
(x) Lys His/Tyr Asn/Val Ser Pro Ala Ala Glu
Asn/Leu Leu Asn/Gly
(y) Lys - Thr Ser Val Ala Glu Ala Phe Asn
(z) Lys Ala Ala Glu Val Ala Glu Ala Phe Asp -
Ile - - - Lys Gly
(aa) Lys Ala Val Glu Val/Pro Ala Glu Ala Phe
Asp Asp Ile Thr? Tyr - - Gly Pro Ser
(bb) Lys - Glu Gln Thr Glu Ile Phe Asn Met
(cc) Lys - - - Pro Phe Asn/Asp Ile Glu
Ala Leu
(dd) Asp Gln Ala Phe Ser Thr Asp Ala Lys

7. Syntetiska polypeptider enligt patentkrav 6 och omfattande en aminosyrasekvens som bildar ett aminopeptidasenzym eller ett antigeniskt parti därav och som motsvarar hela eller ett parti av nukleotidsekvenserna som visats i fig. 2, 3, 4 eller 5 (SEQ ID nr 1 - 15) eller en funktionellt ekvivalent variant därav, vilken är väsentligen fri från andra Haemonchus Contortus -komponenter.

8. Syntetiska polypeptider enligt patentkrav 6 eller 7 i form av en fusionspolypeptid och omfattande en ytterligare polypeptid som fusionerats med aminosyrasekvensen enligt patentkrav 6 eller 7.

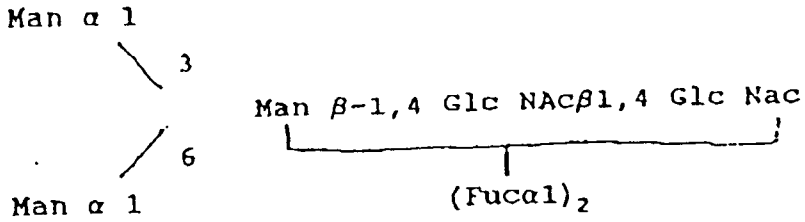
9. Förfarande för framställning av en syntetisk polypeptid enligt något av patentkraven 6 - 8, k ä n n e t e c k n a t av, att man odlar en prokaryotisk eller eukaryotisk cell, vilken innehåller en nukleinsyramolekyl enligt något av patentkraven 1 - 3 under förhållanden, vid vilka polypeptiden exprimeras, och tillvaratager den sålunda producerade polypeptiden.

10. Vaccinkomposition för stimulering av immungensvar mot maskparasiter hos en människa eller ett icke-mänskligt djur, varvid den omfattar åtminstone en syntetisk polypeptid enligt något av patentkraven 6 - 8, eller ett virus eller en värdcell, i vilka införts en nukleinsyramolekyl enligt något av patentkraven 1 - 3, för stimulering av ett immungensvar hos popypeptider genom den insatta nukleinsyramolekylen, samt en farmaceutiskt godtagbar bärare.

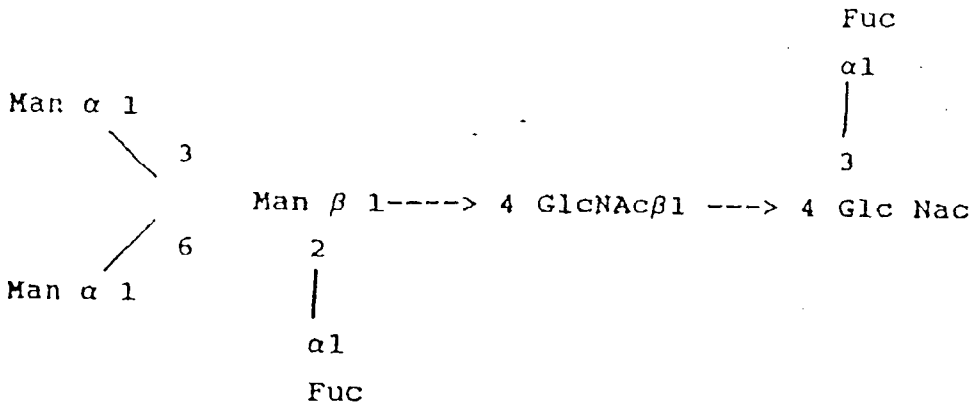
11. Användningen av en nukleinsyramolekyl enligt något av patentkraven 1 - 3 eller av en syntetisk polypeptid enligt något av patentkraven 6 - 8 för framställning av en vaccinkomposition för stimulering av ett immungensvar mot maskparasiter hos en människa eller hos ett icke-mänskligt djur.

12. Förfarande för stimulering av ett immungensvar mot maskparasiter hos en människa eller ett icke-mänskligt djur, k ä n n e t e c k n a t av, att man åt djuret administrerar en vaccinkomposition enligt patentkrav 10.

13. Oligosackarid med strukturen

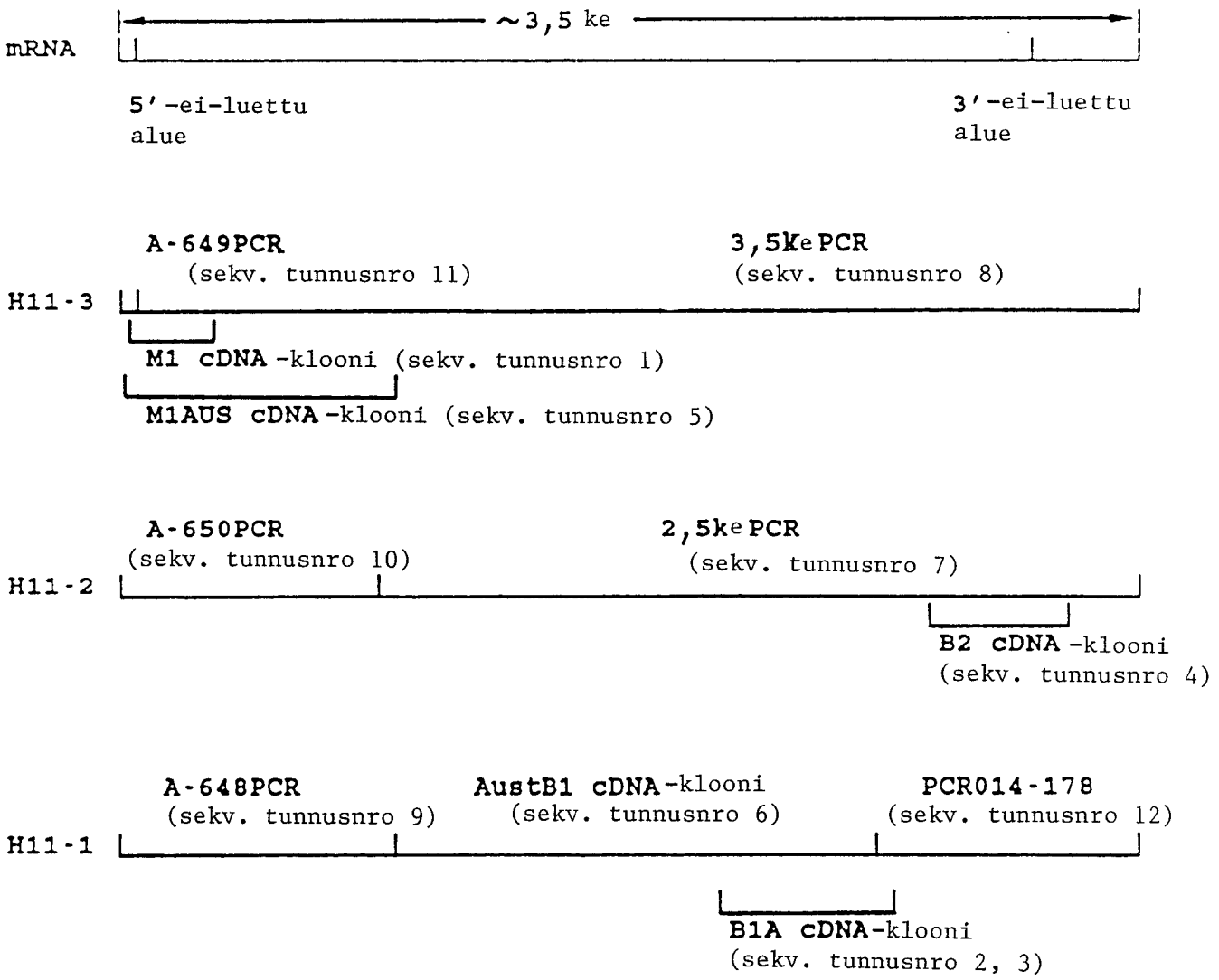


14. Oligosackarid enligt patentkrav 13 med strukturen



15. Oligosackarid enligt patentkrav 13 eller 14 och bunden vid en syntetisk polypeptid enligt något av patentkraven 6 - 8.

FIG. 1


mRNA
~3,5 ke
5'-ei-luettu alue
3'-ei-luettu alue
H11-3
A-649PCR
(sekv. tunnusnro 11)
3,5ke PCR
(sekv. tunnusnro 8)
M1 cDNA-klooni (sekv. tunnusnro 1)
M1AUS cDNA-klooni (sekv. tunnusnro 5)
H11-2
A-650PCR
(sekv. tunnusnro 10)
2,5ke PCR
(sekv. tunnusnro 7)
B2 cDNA-klooni
(sekv. tunnusnro 4)
H11-1
A-648PCR
(sekv. tunnusnro 9)
AustB1 cDNA-klooni
(sekv. tunnusnro 6)
PCR014-178
(sekv. tunnusnro 12)
B1A cDNA-klooni
(sekv. tunnusnro 2, 3)

```
          1                                                        50
H11-3     GGTTTAATTA CCCAAGTTTG AGGGTCTCCA TCTAG..... .........A
H11-2     GGTTTAATTA CCCAAGTTTG AG..... ..... .........A
H11-1     GGTTTAATTA CCCAAGTTTG AGATGACAGC AGAGGAGAGT CAGGAGCAGG

          51                                                       100
H11-3     TGACGTCGCA GGGGAGAACG CGGACATTGC TGAATCTAAC TCCAATCCGT
H11-2     TGACGGCGGA GTGGCAGAAG CGTCGAATCT TGGGCTTCTC ACCTATCAGC
H11-1     AGACGCAGCA ACCACGAAAA AATACAGTGC TACGGCTCAC CCCAATCAAG

          101                                                      150
H11-3     CTTATTGTCG CATTATTTCT AGTAGCTGCT GCAGTCGGCC TCTCTATTGG
H11-2     CTACTTTGTA CATTATTTGT ATTAGCTGCT GCCGTTGGAC TCTCCATTGG
H11-1     TCTCTCTTTG CTTTGTTAGT GGTAGCTGCT GCCGTCGGCC TCTCAATCGG

          151                                                      200
H11-3     TCTCACCTAT TACTTTACTC GCAAAGCGTT CGATACCTCA GAAAAGCCAG
H11-2     TCTTACCTAT TACTTCACTC GTAAAGCATT CGATACCACA CAAAAAGAAC
H11-1     TCTCACCTAT TACTTTACAA GGAAAGCTTT TGATACTACT GGCGGAAATG

          201                                                      250
H11-3     GGAAGGATGA TACTGGTGGC AAGGACAAAG ACAATTCTCC CTCTGCGGCG
H11-2     AGAAGGATGA CAGTGGTGGT AAAGAAAAGG ATAATTCTCC TTCTGCAGAA
H11-1     GAAAAGGGGA TCAACCTATT GTCGAT...G ATAATTCCCC ATCAGCTGAA

          251                                                      300
H11-3     GAACTACTCC TTCCAAGTAA TATAAAACCA TTGTCTTACG ACTTGACGAT
H11-2     GAACTACTTC TTCCAACGAA CATAAAACCA GTCTCGTACG ACTTGAACAT
H11-1     GAATTACGTC TCCCAACAAC CATAAAACCT TTGACATACG ACTTAGTAAT

          301                                                      350
H11-3     CAAAACATAT CTACCTGGTT ATGTGGACTT CCCACCGGAG AAAAACCTCA
H11-2     CAAAACATAT CTACCGGGTT ACGTGAACTT TCCACCAGAA AAGAATCTCA
H11-1     CAAAACGTAT CTGCCAAACT ATGTAAACTA TCCACCTGAG AAAGATTTCG

          351                                                      400
H11-3     CATTCGATGG GCGTGTGGAA ATATCAATGG TTGTAATTGA GCCAACAAAG
H11-2     CATTTGATGC CCATGTGGAG ATTGCTATGG TTGTGGTTGA GCCTACAAAT
H11-1     CTATTGATGG GACTGTGGTG ATTGCTATGG AAGTTGTGGA GCCAACAAAG

          401                                                      450
H11-3     AGTATCGTAC TCAATTCAAA GAAGATCTCT GTAATACCCC AAGAATGTGA
H11-2     AGTATTGTGC TGAACTCGAA GAAAATCACT TTGGCACAAG GAGGATGCGA
H11-1     TCTATTGTGC TCAACTCGAA AAATATTCCT GTAATTGCAG ACCAGTGCGA

          451                                                      500
H11-3     ACTGGTATCG GGCGATAAAA AACTCGAAAT TGAAAGTGTA AAGGAGCACC
H11-2     ACTGTTCTCA GGTAATCAGA AACTTGACAT CGAAAGTGTA AAGATGCAGG
H11-1     ACTGTTTTCT AACAACCAAA AACTCGACAT CGAAAAGGTT GTGGATCAGC

          501                                                      550
H11-3     CAAGACTGGA AAAGGTTGAG TTTCTTATCA AAAGCCAACT GGAAAAAGAT
H11-2     AAAGACTTGA CAAGCTTGAG ATTACCCTCA AAAATCAGCT GCAAAAAGAT
H11-1     CAAGGCTGGA GAAAGTCGAA TTCGTTTTGA AGAAAAAGCT GGAGAAGAAT

          551                                                      600
H11-3     CAACAAATCT TGCTCAAGGT CGGCTACATC GGTCTCATCA GCAACAGCTT
H11-2     CTGAAAATCC TGCTCAAGAT CACTTACACC GGCCTTATTA GCGACACTCT
```

FIG. 2

```
H11-1    CAGAAAATCA CGCTCAAGAT TGTATACATT GGCCTTATCA ACGACATGCT

          601                                                  650
H11-3    TGGTGGAATC TACCAGACCA CTTATACCAC CCCGGATGGC ACCCCTAAGA
H11-2    CGGTGGGCTC TACCAGTCCA TCTACACTGA TAAGGACGGA AAAACTAAGA
H11-1    TGGAGGACTT TATCGAACAA CCTACACGGA TAAAGATGGT ACAACCAAGA

          651                                                  700
H11-3    TCGCTGCAGT TTCACAAAAT GAGCCCATAG ATGCTCGTCG AATGGTACCA
H11-2    TCGTTGCTGT TTCACAAAAT GAACCATCAG ACGCTCGTCG TATAGCGCCA
H11-1    TTGCTGCATG CACTCATATG GAACCGACGG ACGCCCGTCT TATGGTCCCC

          701                                                  750
H11-3    TGCATGGATG AACCGAAATA CAAAGCAAAC TGGACCGTTA CTGTCATTCA
H11-2    TGCTTTGACG AACCGAAGTA CAAGGCAACA TGGACTGTCA CCGTCGTTCA
H11-1    TGTTTCGACG AGCCGACGTT TAAGGCAAAC TGGACTGTGA CTGTGATTCA

          751                                                  800
H11-3    TCCAAAAGGC ACCAAAGCCG TCTCGAATGG AATCGAAGTG AACGGAGATG
H11-2    TCCCAAAGGT ACAAAGGCTG CATCGAACGG CATTGAAGCA AATGGAAAAG
H11-1    TCCGAAGGGC ACCAGTGCCG TGTCGAATGG AATAGAA... AAGGGAGAAG

          801                                                  850
H11-3    GAGAGATCAG TGGTGATTGG ATCACATCGA AGTTCTTGAC TACTCCACGG
H11-2    GGGAGCTCAA GGGTGATTGG ATAACGTCTA AATTTAAAAC TACCCCACCG
H11-1    GAGAAGTCTC TGGCGATTGG GTCACAACCA GATTCGATCC AACCCCGCGA

          851                                                  900
H11-3    ATGTCATCCT ACTTGTTGGC AGTTATGGTT TCAGAATTTG AATACATCGA
H11-2    ATGTCGTCCT ATTTATTGGC TATTATTGTT TGTGAATTTG AATACATTGA
H11-1    ATGCCTTCGT ATTTGATTGC TCTTGTGATT TCCGAATTTA AGTACATTGA

          901                                                  950
H11-3    AGGTGAAACA AAGACGGGTG TTCGGTTCCG TATATGGTCA CGCCCAGAGG
H11-2    AGGATTTACA AAAACGGGTG TTCGGTTCCG TATATGGTCT CGACCAGAGG
H11-1    AAATTATACG AAAAGCGGTG TTCGATTCCG AATTCCGGCT CGTCCGGAAG

          951                                                 1000
H11-3    CAAAGAAGAT GACACAATAT GCTCTGCAAT CTGGTATCAA GTGCATAGAA
H11-2    CGAAACGAAT GACGGCATAC GCTTTGGATG CTGGCATCAG ATGCCTGGAG
H11-1    CTATGAAGAT GACAGAATAT GCCATGATAG CTGGAATCAA ATGTTTGGAT

          1001                                                1050
H11-3    TTCTACGAAG ATTTCTTTGA TATCAGATTC CCTCTGAAGA AACAAGATAT
H11-2    TTCTATGAGA AGTTCTTTGA CATAAAATTC CCTCTGGAAA AACAAGATAT
H11-1    TACTATGAGG ACTTCTTCGG GATCAAATTC CCACTTCCAA AACAAGATAT

          1051                                                1100
H11-3    GATTGCCCTT CCTGATTTCT CTGCCGGTGC CATGGAGAAT TGGGGCCTCA
H11-2    GATTGCTCTT CCTGATTTCA CCGCTGGTGC CATGGAAAAC TGGGGCCTTA
H11-1    GGTTGCTCTT CCTGACTTCT CATCTGGTGC TATGGAGAAC TGGGGTCTCA

          1101                                                1150
H11-3    TCACTTACAG GGAAAACTCT TTGTTGTACG ATGACAGATT CTATGCACCG
H11-2    TCACTTATAG AGAGGATTCT CTCCTATACG ATGAAAAAAT TTATGCACCG
H11-1    TCACATACAG GGAGGGTTCC GTGCTCTACG ATGAAAACCT CTACGGACCA

          1151                                                1200
```

FIG. 2 (jatkuu)

```
H11-3    ATGAATAAAC AGCGAATTGC TCGCATTGTT GCTCATGAGC TTGCTCATCA
H11-2    ATGAATAAAC AGCGGGTTGC TCTCGTAGTT GCTCACGAGC TTGCTCATCA
H11-1    ATGAATAAGG AGCGGGTTGC AGAAGTGATC GCGCACGAAC TTGCACATCA

          1201                                                1250
H11-3    GTGGTTCGGC GACTTGGTTA CGATGAAGTG GTGGGATAAT TTGTGGTTGA
H11-2    GTGGTTCGGC AATCTGGTCA CACTGAAGTG GTGGGATGAT ACGTGGTTGA
H11-1    GTGGTTCGGT AATTTGGTCA CGATGAAGTG GTGGGATAAC CTATGGCTGA

          1251                                                1300
H11-3    ATGAAGGTTT TGCAAGATTC ACAGAATTTA TTGGAGCTGG TCAGATAACT
H11-2    ACGAAGGTTT TGCAACATTT GTTGAGTATC TTGGAATGGA CGAAATTAGC
H11-1    ACGAAGGATT CGCGTCATTC GTGGAATACA TCGGAGCCGA CTTCATCAGC

          1301                                                1350
H11-3    CAAGATGACG CCAGAATGAG GAACTACTTC CTGATTGATG TACTTGAACG
H11-2    CACAACAATT TCAGAACGCA AGATTTCTTC TTGCTCGATG GAATGGATCG
H11-1    GATGGTCTAT GGGAAATGAA AGATTTCTTC CTGCTGGCAC CGTACACAAG

          1351                                                1400
H11-3    CGCTTTGAAA GCTGATTCGG TAGCGTCAAG CCATCCACTT TCCTTCAGAA
H11-2    CGGAATGAGA GCTGACTCGG CAGCATCGAG CCATCCGCTT TCGTTTAGGA
H11-1    TGGTATTACG GCTGATGCAG TAGCTTCAAG CCATCCGCTT TCCTTCAGAA

          1401                                                1450
H11-3    TCGACAAAGC TGCAGAAGTT GAAGAAGCCT TTGATGATAT CACATACGCC
H11-2    TTGACAAAGC GGCAGAAGTT GCCGAAGCCT TTGACGATAT TTCATACGCC
H11-1    TAGATAAGGC TGCAGATGTA TCAGAAGCGT TCGATGATAT CACATACCGT

          1451                                                1500
H11-3    AAAGGAGCTT CTGTTCTTAC TATGCTGAGA GCCTTGATTG GAGAAGAAAA
H11-2    AAGGGAGCGT CAGTTCTCAC TATGCTACGG GCTTTGATTG GAGAGGACAA
H11-1    AAAGGAGCAT CCGTTCTTCA AATGCTATTG AATTTAGTTG GGGACGAAAA

          1501                                                1550
H11-3    ACATAAGCAT GCAGTATCGC AGTACCTCAA GAAGTTCTCG TATAGCAATG
H11-2    TTACAGGAAT GCTGTTGTGC AATACCTCAA GAAGTTCTCC TACAGCAATG
H11-1    TTTCAAGCAG TCTGTTTCGC GTTACCTCAA GAAGTTTTCA TATGATAATG

          1551                                                1600
H11-3    CAGAAGCGAC TGATCTATGG GCAGTTTTTG ATGAAGTTGT CACTGACGTC
H11-2    CACAAGCAGC CGATCTGTGG AACGTCTTCA ATGAAGTTGT CAAAGGTGTT
H11-1    CGGCTGCTGA AGATTTATGG GCAGCATTCG ACGAAACCGT CCAAGGTATA

          1601                                                1650
H11-3    GAAGGTCCAG ACGGCAAACC TATGAAAACC ACAGAGTTTG CAAGTCAGTG
H11-2    AAGGGTCCTG ACGGCAACGT CATGAAAATC GACCAATTTA CCGATCAGTG
H11-1    ACCGGACCTA ATGGTGGACC ATTGAAAATG TCCGAGTTTG CGCCACAATG

          1651                                                1700
H11-3    GACGACTCAG ATGGGCTTCC CAGTTATTTC CGTAGCAGAG TTTAACTCGA
H11-2    GACGTATCAG ATGGGTTATC CTGTGGTTAA AGTAGAAGAA TTTAATGCGA
H11-1    GACAACTCAG ATGGGGTTCC CTGTTCTTAC TGTCGAGTCG GTTAACGCAA

          1701                                                1750
H11-3    CTACTTTGAA ATTAACGCAA AGTCGATATG AGGCGAATAA AGACGCTGTG
H11-2    CCGCCCTAAA GGTTACGCAG AGCCGGTACA AGACAAATAA AGACGCCTTG
H11-1    CGACTTTGAA AGTCACCCAA AAACGATACA GGCAGAACAA GGATGCAAAG
```

FIG. 2 (jatkuu)

```
          1751                                                    1800
H11-3     GAGAAAGAGA AGTACCGTCA CCCGAAATAC GGATTTAAAT GGGATATTCC
H11-2     GAACCAGAGA AATATCGTAA TCCAAAATAC GGGTTCAAGT GGGATGTTCC
H11-1     GAACCAGAGA AGTACCGTCA TCCAACTTAT GGGTTCAAAT GGGATGTTCC

          1801                                                    1850
H11-3     ACTGTGGTAT CAGGAAGGCG ATAAGAAGGA GATAAAGCGA ACATGGTTGA
H11-2     CCTATGGTAT CAGGAAGGCA ATAGCAAAGA GGTGAAGCGA ACATGGCTAA
H11-1     TCTGTGGTAT CAGGAAGATG AA...CAGCA AGTGAAAAGA ACTTGGTTAA

          1851                                                    1900
H11-3     GAAGAGATGA ACCGCTTTAC TTGCATGTTA GTGATGCTGG CGCTCCCTTT
H11-2     AAAGAGATGA ACCGCTGTAC TTGAACGTCA ACAATCGGGA TACATCCCTT
H11-1     AAAGAGAGGA ACCGCTCTAT TTCCATGTAA GCAATTCTGA TTCGTCAGTT

          1901                                                    1950
H11-3     GTGGTGAACG CAGACCGCTA TGGATTTTAT CGACAAAATC ATGACGCTAA
H11-2     GTGGTGAACG CTGATCGACA TGGATTTTAT CGACAAAACT ATGATGCCAA
H11-1     GTGGTGAATG CCGAACGTCG TGCTTTTTGC CGATCAAACT ATGACGCTAA

          1951                                                    2000
H11-3     TGGTTGGAAA AAGATAATCA AGCAGCTCAA GGATAATCAT GAGGTTTACA
H11-2     CGGTTGGAAA AAGATAATCA AGCAGCTCAA GAAAGATCAC AAGGTCTTCG
H11-1     CGGTTGGAGG AACATTATGA GAAGACTCAA GCAGAATCAT AAGGTCTATG

          2001                                                    2050
H11-3     GTCCCCGGAC AAGGAATGTC ATCATTAGCG ATGCGTTTGC TGCGGCTGCA
H11-2     GTCCAAGGAC AAGGAACGCT ATCATAAGCG ATGCATTTGC TGCAGCTACG
H11-1     GTCCACGAAC AAGAAACGCT CTCATAAGTG ATGCGTTTGC AGCAGCTGCA

          2051                                                    2100
H11-3     ACTGACGCAA TTGAGTATGA GACTGTATTT GAACTTCTGA ATTATGCCGA
H11-2     ATTGACGCAA TCGACTATGA AACTGTATTC GAACTACTTG AATATGCCAA
H11-1     GTTGAGGAAA TGAATTACGA GACCGTATTT GAAATGCTCA AATACACCGT

          2101                                                    2150
H11-3     AAAAGAAACG GAATATCTAC CATTAGAAAT CGCAATGTCC GGGATCTCTT
H11-2     AAATGAAGAG GAATTCTTGC CTTGGAAGGA AGCTCTGTCC GGCATGTTCG
H11-1     GAAAGAAGAG GATTACTTAC CATGGAAGGA GGCAATATCA GGATTCAATA

          2151                                                    2200
H11-3     CGATTTTGAA ATACTTCCCT ACCGAGCCAG AGGCAAAGCC AGCTCAAACA
H11-2     CAGTTTTAAA GTTCTTCGGT AATGAGCCGG AGACAAAACC AGCTAGAGCT
H11-1     CAATTTTGGA CTTTTTCGGC AGCGAACCCG AATCTCAATG GGCTTCGGAA

          2201                                                    2250
H11-3     TACATGATGA ACATATTGAA ACCGATGTAT GAAAAAAGCA GTATCGACTT
H11-2     TACATGATGA GCATATTAGA ACCGATGTAT AATAAGAGCA GCATTGATTA
H11-1     TACATGCGAA AACTGATGAA GCCAATTTAT GACAAGAGTA GCATCAAGTT

          2251                                                    2300
H11-3     CATTGCCAAT AACTACAGAA ATGACAAGCT GTTTTTCCAA ATCAACCTCC
H11-2     CATCGTCAAG AATTATTTGG ATGATACGTT ATTCACAAAA ATTAATACTC
H11-1     TATAGCGGAG AACTACAAAA AAGATTCGCT TTTCTTCAAA AATAATCTCC

          2301                                                    2350
H11-3     AAAAAGATGT CATTGATATG TTCTGCGCCC TCGGATCGCA AGACTGCAGG
H11-2     AAAAGGATAT CATTGATGCA TATTGTTCCC TTGGATCAAA GGACTGTATA
```

FIG. 2 (jatkuu)

```
H11-1    AAATAGCTGT TATTGACACA TACTGTGGTC TTGGAGGCAA AGAATGTCTT

          2351                                                2400
H11-3    AAGAAATATA AAAAACTTTT CGATGACGAA GTCATGAACA AATGCAGGGA
H11-2    AAGCAATATA AGGATATCTT CTACGATGAG GTTATGCCCA AGTGTAAGGC
H11-1    GAAGAAATGA AAAAGCTTTT TGACAAGGAG GTCATG...A AATGTCAACC

          2401                                                2450
H11-3    TGGTCAAGCA GCAACCGAAT GCGTAAGAAT CGCCGCTCCT CTCCGATCAA
H11-2    CGGGGAAGCA GCAACCAAAT GCGTTAAGGT TTCCGCTCCT CTTCGAGCCA
H11-1    TGGTCAGCAA GCGACCGACT GCGTAAAGGT AACTGCTCCT CTCCGAAAAA

          2451                                                2500
H11-3    GTGTTTATTG TTATGGTGTG AAGGAAGGCG GTGATTATGC TTCCGACAAG
H11-2    ATGTTTACTG TTATGGTGTA CAGGAAGGTG GTGAAGAAGC TTTTGAAAAG
H11-1    CTGTTTACTG CTATGGGGTC CAGGAAGGCG GTGATGAGGC ATTCGACAAG

          2501                                                2550
H11-3    GTGATGGAGC TTTATACGGC CGAAACACTC GCCCTAGAAA AAGACTTCCT
H11-2    GTGATGGGGC TGTATCTAGC AGAAGATGTT CAACTGGAGA AGGGTATCCT
H11-1    GTGATGGAAC TATATAATGC GGAACAAGTG CAGTTGGAGA AAGACAGTCT

          2551                                                2600
H11-3    ACGCCTAGCA TTGGGATGTC ATAAAGATGT TACTGCTTTG AAAGGACTTC
H11-2    GTTCAAAGCC TTGGCATGCC ACAAAGATGT TACAGCTCTA AAAGAACTTC
H11-1    ACGTGAAGCA TTGGGATGCC ATAAAGACGT TACAGCTCTA AAGGGACTTC

          2601                                                2650
H11-3    TCTTGCGGGC TCTGGACAGG AATTCGTCGT TCGTACGTAT GCAGGATATC
H11-2    TTTTGCGAGC CCTGGACCGT AAATCGTCGT TTGTGCGTCT TCAGGATGTC
H11-1    TTATGCTGGC TTTGGATCGG AATTCGTCAT TTGTGCGTCT TCAAGATGCT

          2651                                                2700
H11-3    CCAAGTGCTT TCAACGATGT AGCAGCAAAT CCTATTGGCG AAGAATTCAT
H11-2    CCTACCGCTT TCCGTGCTGT ATCTGAAAAC CCTGTGGGCG AAGAATTCAT
H11-1    CATGATGTGT TTAACATTGT ATCCAGAAAT CCTGTTGGAA ACGAACTGCT

          2701                                                2750
H11-3    TTTCAATTTC CTTATTGAGA GATGGCCAGA TATCATTGAA AGTATAGGAA
H11-2    GTTCAATTTC CTAATGGAGA GATGGGAGGA AATCACTGCG AGCTTGGAAA
H11-1    GTTCAATTTC CTCACAGAGC GATGGGAAGA GATACTTGAA AGTTTGTCAA

          2751                                                2800
H11-3    CGAAGCACAC ATATGTTGAG AAAGTGATAC CAGCCTGCAC TTCAGGAATC
H11-2    CAGAACACAG AGCAGTTGAT AAAGTGGTCG GCGCTTGTTG CACAGGAATT
H11-1    TACGACACAG ATCAGTTGAT CGAGTGATCA AAGCCTGTAC TCGAGGACTA

          2801                                                2850
H11-3    CGCTCACAAC AGCAGATTGA CCAGCTGAAG AATCTGCAGA AAAATGGCAT
H11-2    CGCTCCCAAC AACAAATAGA TCAGCTGAAG AATCTACAGA AGAACAATGC
H11-1    CGATCCAGGG AACAAGTACA ACAGTTGAAG AATCTATACA AAAATGACAA

          2851                                                2900
H11-3    GAACGCTCGT CAATTCGGTG CATTCGATAA AGCAATCGAA CGAGCACAAA
H11-2    GCAGGCTAAG AAGTTCGGCT CATTCACCCA GGAAATCGAA AAAGGAGAAC
H11-1    GCGTGCTCGC GAATACGGTG CATTTGGTGG GGCAATAGAA AGATCGGAAC

          2901                                                2950
```

FIG. 2 (jatkuu)

```
H11-3    ATAGGGTGGA TTGGATTAAA AAACATTTCC AAAAATTAGC GGCTTTCTTC
H11-2    ATAAAATTGC CTGGATCAAG AAACATTTTC ACAGATTATC GGAATTCTTC
H11-1    ACAGAGTCAA ATGGATTGAG AAACATTTCC GAAAACTAGC AGCTTTCTTC

          2951            2958
H11-3    AAGAAAGCCA CCTTGTAA
H11-2    AAGAGAGCAA GATCATAG
H11-1    AAAAAATCTA ATTCATAA
```

FIG. 2 (jatkuu)

```
         1                                                  50
M1                          C GCGGACATTG CTGAATCTAA CTCCAATCCG
AustM1   ATGACGTCGC AGGGGAGAAC GCGGACATTG CTGAATCTAA CTCCAATCCG
H11-3    ATGACGTCGC AGGGGAGAAC GCGGACATTG CTGAATCTAA CTCCAATCCG

         51                                                100
M1       TCTTATTGTC GCATTATTTC TAGTAGCTGC TGCAGTCGGC CTCTCTATTG
AustM1   TCTTATTGTC GCATTATTTC TAGTAGCTGC TGCAGTCGGC CTCTCTATTG
H11-3    TCTTATTGTC GCATTATTTC TAGTAGCTGC TGCAGTCGGC CTCTCTATTG

         101                                               150
M1       GTCTCACCTA TTACTTTACT CGCAAAGCGT TCGATACCTC AGAAAAGCCA
AustM1   GTCTCACCTA TTACTTTACT CGCAAAGCGT TCGATACCTC AGAAAAGCCA
H11-3    GTCTCACCTA TTACTTTACT CGCAAAGCGT TCGATACCTC AGAAAAGCCA

         151                                               200
M1       GGGAAGGATG ATACTGGTGG CAAGGACAAA GACAATTCTC CCTCTGCGGC
AustM1   GGGAAGGATG ATACTGGTGG CAAGGACAAA GACAATTCTC CCTCTGCGGC
H11-3    GGGAAGGATG ATACTGGTGG CAAGGACAAA GACAATTCTC CCTCTGCGGC

         201                                               250
M1       GGAACTACTT CTTCCAAGTA ATATAAAACC ATTGTCTTAC GACTTGACGA
AustM1   GGAACTACTC CTTCCAAGTA ATATAAAACC ATTGTCTTAC GACTTGACGA
H11-3    GGAACTACTC CTTCCAAGTA ATATAAAACC ATTGTCTTAC GACTTGACGA

         251                                               300
M1       TCAAAACATA TCTACCTGGT TATGTGGACT TCCCACCGGA GAAAAACCTC
AustM1   TCAAAACATA TCTACCTGGT TATGTGGACT TCCCACCGGA GAAAAACCTC
H11-3    TCAAAACATA TCTACCTGGT TATGTGGACT TCCCACCGGA GAAAAACCTC

         301                                               350
M1       ACATTCGACG GGCG
AustM1   ACATTCGATG GGCGTGTGGA AATATCAATG GTTGTAATTG AGCCAACAAA
H11-3    ACATTCGATG GGCGTGTGGA AATATCAATG GTTGTAATTG AGCCAACAAA

         351                                               400
AustM1   GAGTATCGTA CTCAATTCAA AGAAGATCTC TGTAATACCC CAAGAATGTG
H11-3    GAGTATCGTA CTCAATTCAA AGAAGATCTC TGTAATACCC CAAGAATGTG

         401                                               450
AustM1   AACTGGTATC GGGCGATAAA AAATTCGAAA TTGAAAGTGT AAAGGAGCAC
H11-3    AACTGGTATC GGGCGATAAA AAACTCGAAA TTGAAAGTGT AAAGGAGCAC

         451                                               500
AustM1   CCAAGACTGG AAAAGGTTGA GTTTCTTATC AAAAGCCAAC TGGAAAAAGA
H11-3    CCAAGACTGG AAAAGGTTGA GTTTCTTATC AAAAGCCAAC TGGAAAAAGA

         501                                               550
AustM1   TTCACAAATC TTGCTCAAGT CGGCTTACAT CGGTCTCATC AGCAACAGCC
H11-3    TCAACAAATC TTGCTCAAGG TCGGCTACAT CGGTCTCATC AGCAACAGCT

         551                                               600
AustM1   TTGGTGGAAT CTACCAGACC ACTTATACCA CCCCGGATGG CACCCCTAAG
H11-3    TTGGTGGAAT CTACCAGACC ACTTATACCA CCCCGGATGG CACCCCTAAG

         601                                               650
AustM1   ATCGCTGCAG TTTCACAAAA TGAGCCCATA GATGCTCGTC GAATGGTACC
H11-3    ATCGCTGCAG TTTCACAAAA TGAGCCCATA GATGCTCGTC GAATGGTACC
```

FIG.3

```
          651                                                       700
AustM1    ATGCATGGAT GAACCGAAAT ACAAAGCAAA CTGGACCGTT ACTGTCATTC
H11-3     ATGCATGGAT GAACCGAAAT ACAAAGCAAA CTGGACCGTT ACTGTCATTC

          701                                                       750
AustM1    ATCCAAAAGG CACCAAAGCC GTCTCGAATG GAATCGAAGT GAACGGAGAT
H11-3     ATCCAAAAGG CACCAAAGCC GTCTCGAATG GAATCGAAGT GAACGGAGAT

          751                                                       800
AustM1    GGAGAGATCA GTGGTGATTG GATCACATCG AAGTTCTTGA CTACTCCACG
H11-3     GGAGAGATCA GTGGTGATTG GATCACATCG AAGTTCTTGA CTACTCCACG

          801                                                       850
AustM1    GATGTCATCC TACTTGTTGG CAGTTATGGT TTCAGAATTT GAATACATCG
H11-3     GATGTCATCC TACTTGTTGG CAGTTATGGT TTCAGAATTT GAATACATCG

          851                                                       900
AustM1    AAGGTGAAAC AAAGACGGGT GTTCGGTTCC GTATATGGTC ACGCCCAGAG
H11-3     AAGGTGAAAC AAAGACGGGT GTTCGGTTCC GTATATGGTC ACGCCCAGAG

          901                                                       950
AustM1    GCAAAGAAGA TGACACAATA TGCTCTGCAA TCTGGTATCA AGTGCATA
H11-3     GCAAAGAAGA TGACACAATA TGCTCTGCAA TCTGGTATCA AGTGCATAGA

          951                                                      1000
H11-3     ATTCTACGAA GATTTCTTTG ATATCAGATT CCCTCTGAAG AAACAAGATA

          1001                                                     1050
H11-3     TGATTGCCCT TCCTGATTTC TCTGCCGGTG CCATGGAGAA TTGGGGCCTC

          1051                                                     1100
H11-3     ATCACTTACA GGGAAAACTC TTTGTTGTAC GATGACAGAT TCTATGCACC

          1101                                                     1150
H11-3     GATGAATAAA CAGCGAATTG CTCGCATTGT TGCTCATGAG CTTGCTCATC

          1151                                                     1200
H11-3     AGTGGTTCGG CGACTTGGTT ACGATGAAGT GGTGGGATAA TTTGTGGTTG

          1201                                                     1250
H11-3     AATGAAGGTT TTGCAAGATT CACAGAATTT ATTGGAGCTG GTCAGATAAC

          1251                                                     1300
H11-3     TCAAGATGAC GCCAGAATGA GGAACTACTT CCTGATTGAT GTACTTGAAC

          1301                                                     1350
H11-3     GCGCTTTGAA AGCTGATTCG GTAGCGTCAA GCCATCCACT TTCCTTCAGA

          1351                                                     1400
H11-3     ATCGACAAAG CTGCAGAAGT TGAAGAAGCC TTTGATGATA TCACATACGC

          1401                                                     1450
H11-3     CAAAGGAGCT TCTGTTCTTA CTATGCTGAG AGCCTTGATT GGAGAAGAAA

          1451                                                     1500
H11-3     AACATAAGCA TGCAGTATCG CAGTACCTCA AGAAGTTCTC GTATAGCAAT

          1501                                                     1550
```

FIG. 3 (jatkuu)

```
H11-3   GCAGAAGCGA CTGATCTATG GGCAGTTTTT GATGAAGTTG TCACTGACGT

        1551                                                1600
H11-3   CGAAGGTCCA GACGGCAAAC CTATGAAAAC CACAGAGTTT GCAAGTCAGT

        1601                                                1650
H11-3   GGACGACTCA GATGGGCTTC CCAGTTATTT CCGTAGCAGA GTTTAACTCG

        1651                                                1700
H11-3   ACTACTTTGA AATTAACGCA AAGTCGATAT GAGGCGAATA AAGACGCTGT

        1701                                                1750
H11-3   GGAGAAAGAG AAGTACCGTC ACCCGAAATA CGGATTTAAA TGGGATATTC

        1751                                                1800
H11-3   CACTGTGGTA TCAGGAAGGC GATAAGAAGG AGATAAAGCG AACATGGTTG

        1801                                                1850
H11-3   AGAAGAGATG AACCGCTTTA CTTGCATGTT AGTGATGCTG GCGCTCCCTT

        1851                                                1900
H11-3   TGTGGTGAAC GCAGACCGCT ATGGATTTTA TCGACAAAAT CATGACGCTA

        1901                                                1950
H11-3   ATGGTTGGAA AAAGATAATC AAGCAGCTCA AGGATAATCA TGAGGTTTAC

        1951                                                2000
H11-3   AGTCCCCGGA CAAGGAATGT CATCATTAGC GATGCGTTTG CTGCGGCTGC

        2001                                                2050
H11-3   AACTGACGCA ATTGAGTATG AGACTGTATT TGAACTTCTG AATTATGCCG

        2051                                                2100
H11-3   AAAAAGAAAC GGAATATCTA CCATTAGAAA TCGCAATGTC CGGGATCTCT

        2101                                                2150
H11-3   TCGATTTTGA AATACTTCCC TACCGAGCCA GAGGCAAAGC CAGCTCAAAC

        2151                                                2200
H11-3   ATACATGATG AACATATTGA AACCGATGTA TGAAAAAAGC AGTATCGACT

        2201                                                2250
H11-3   TCATTGCCAA TAACTACAGA AATGACAAGC TGTTTTTCCA AATCAACCTC

        2251                                                2300
H11-3   CAAAAAGATG TCATTGATAT GTTCTGCGCC CTCGGATCGC AAGACTGCAG

        2301                                                2350
H11-3   GAAGAAATAT AAAAAACTTT TCGATGACGA AGTCATGAAC AAATGCAGGG

        2351                                                2400
H11-3   ATGGTCAAGC AGCAACCGAA TGCGTAAGAA TCGCCGCTCC TCTCCGATCA

        2401                                                2450
H11-3   AGTGTTTATT GTTATGGTGT GAAGGAAGGC GGTGATTATG CTTCCGACAA

        2451                                                2500
H11-3   GGTGATGGAG CTTTATACGC CCGAAACACT CGCCCTAGAA AAAGACTTCC
```

FIG. 3 (jatkuu)

```
          2501                                                    2550
H11-3     TACGCCTAGC ATTGGGATGT CATAAAGATG TTACTGCTTT GAAAGGACTT

          2551                                                    2600
H11-3     CTCTTGCGGG CTCTGGACAG GAATTCGTCG TTCGTACGTA TGCAGGATAT

          2601                                                    2650
H11-3     CCCAAGTGCT TTCAACGATG TAGCAGCAAA TCCTATTGGC GAAGAATTCA

          2651                                                    2700
H11-3     TTTTCAATTT CCTTATTGAG AGATGGCCAG ATATCATTGA AAGTATAGGA

          2701                                                    2750
H11-3     ACGAAGCACA CATATGTTGA GAAAGTGATA CCAGCCTGCA CTTCAGGAAT

          2751                                                    2800
H11-3     CCGCTCACAA CAGCAGATTG ACCAGCTGAA GAATCTGCAG AAAAATGGCA

          2801                                                    2850
H11-3     TGAACGCTCG TCAATTCGGT GCATTCGATA AAGCAATCGA ACGAGCACAA

          2851                                                    2900
H11-3     AATAGGGTGG ATTGGATTAA AAAACATTTC CAAAAATTAG CGGCTTTCTT

          2901                2919
H11-3     CAAGAAAGCC ACCTTGTAA
```

FIG. 3 (jatkuu)

```
H11-2  GGTTTAATTA CCCAAGTTTG AGATGACGGC GGAGTGGCAG AAGCGTCGAA   50
H11-2  TCTTGGGCTT CTCACCTATC AGCCTACTTT GTACATTATT TGTATTAGCT  100
H11-2  GCTGCCGTTG GACTCTCCAT TGGTCTTACC TATTACTTCA CTCGTAAAGC  150
H11-2  ATTCGATACC ACACAAAAAG AACAGAAGGA TGACAGTGGT GGTAAAGAAA  200
H11-2  AGGATAATTC TCCTTCTGCA GAAGAACTAC TTCTTCCAAC GAACATAAAA  250
H11-2  CCAGTCTCGT ACGACTTGAA CATCAAAACA TATCTACCGG GTTACGTGAA  300
H11-2  CTTTCCACCA GAAAAGAATC TCACATTTGA TGCCCATGTG GAGATTGCTA  350
H11-2  TGGTTGTGGT TGAGCCTACA AATAGTATTG TGCTGAACTC GAAGAAAATC  400
H11-2  ACTTTGGCAC AAGGAGGATG CGAACTGTTC TCAGGTAATC AGAAACTTGA  450
H11-2  CATCGAAAGT GTAAAGATGC AGGAAAGACT TGACAAGCTT GAGATTACCC  500
H11-2  TCAAAAATCA GCTGCAAAAA GATCTGAAAA TCCTGCTCAA GATCACTTAC  550
H11-2  ACCGGCCTTA TTAGCGACAC TCTCGGTGGG CTCTACCAGT CCATCTACAC  600
H11-2  TGATAAGGAC GGAAAAACTA AGATCGTTGC TGTTTCACAA AATGAACCAT  650
H11-2  CAGACGCTCG TCGTATAGCG CCATGCTTTG ACGAACCGAA GTACAAGGCA  700
H11-2  ACATGGACTG TCACCGTCGT TCATCCCAAA GGTACAAAGG CTGCATCGAA  750
H11-2  CGGCATTGAA GCAAATGGAA AAGGGGAGCT CAAGGGTGAT TGGATAACGT  800
H11-2  CTAAATTTAA AACTACCCCA CCGATGTCGT CCTATTTATT GGCTATTATT  850
H11-2  GTTTGTGAAT TTGAATACAT TGAAGGATTT ACAAAAACGG GTGTTCGGTT  900
H11-2  CCGTATATGG TCTCGACCAG AGGCGAAACG AATGACGGCA TACGCTTTGG  950
H11-2  ATGCTGGCAT CAGATGCCTG GAGTTCTATG AGAAGTTCTT TGACATAAAA 1000
H11-2  TTCCCTCTGG AAAAACAAGA TATGATTGCT CTTCCTGATT TCACCGCTGG 1050
H11-2  TGCCATGGAA AACTGGGGCC TTATCACTTA TAGAGAGGAT TCTCTCCTAT 1100
H11-2  ACGATGAAAA AATTTATGCA CCGATGAATA AACAGCGGGT TGCTCTCGTA 1150
H11-2  GTTGCTCACG AGCTTGCTCA TCAGTGGTTC GGCAATCTGG TCACACTGAA 1200
H11-2  GTGGTGGGAT GATACGTGGT TGAACGAAGG TTTTGCAACA TTTGTTGAGT 1250
H11-2  ATCTTGGAAT GGACGAAATT AGCCACAACA ATTTCAGAAC GCAAGATTTC 1300
H11-2  TTCTTGCTCG ATGGAATGGA TCGCGGAATG AGAGCTGACT CGGCAGCATC 1350
H11-2  GAGCCATCCG CTTTCGTTTA GGATTGACAA AGCGGCAGAA GTTGCCGAAG 1400
H11-2  CCTTTGACGA TATTTCATAC GCCAAGGGAG CGTCAGTTCT CACTATGCTA 1450
```

FIG. 4

```
H11-2  CGGGCTTTGA TTGGAGAGGA CAATTACAGG AATGCTGTTG TGCAATACCT 1500
H11-2  CAAGAAGTTC TCCTACAGCA ATGCACAAGC AGCCGATCTG TGGAACGTCT 1550
H11-2  TCAATGAAGT TGTCAAAGGT GTTAAGGGTC CTGACGGCAA CGTCATGAAA 1600
H11-2  ATCGACCAAT TTACCGATCA GTGGACGTAT CAGATGGGTT ATCCTGTGGT 1650
H11-2  TAAAGTAGAA GAATTTAATG CGACCGCCCT AAAGGTTACG CAGAGCCGGT 1700
H11-2  ACAAGACAAA TAAAGACGCC TTGGAACCAG AGAAATATCG TAATCCAAAA 1750
H11-2  TACGGGTTCA AGTGGGATGT TCCCCTATGG TATCAGGAAG GCAATAGCAA 1800
H11-2  AGAGGTGAAG CGAACATGGC TAAAAAGAGA TGAACCGCTG TACTTGAACG 1850
H11-2  TCAACAATCG GGATACATCC CTTGTGGTGA ACGCTGATCG ACATGGATTT 1900
H11-2  TATCGACAAA ACTATGATGC CAACGGTTGG AAAAAGATAA TCAAGCAGCT 1950
H11-2  CAAGAAAGAT CACAAGGTCT TCGGTCCAAG GACAAGGAAC GCTATCATAA 2000
H11-2  GCGATGCATT TGCTGCAGCT ACGATTGACG CAATCGACTA TGAAACTGTA 2050
H11-2  TTCGAACTAC TTGAATATGC CAAAAATGAA GAGGAATTCT TGCCTTGGAA 2100
H11-2  GGAAGCTCTG TCCGGCATGT TCGCAGTTTT AAAGTTCTTC GGTAATGAGC 2150
H11-2  CGGAGACAAA ACCAGCTAGA GCTTACATGA TGAGCATATT AGAACCGATG 2200
H11-2  TATAATAAGA GCAGCATTGA TTACATCGTC AAGAATTATT TGGATGATAC 2250
H11-2  GTTATTCACA AAAATTAATA CTCAAAAGGA TATCATTGAT GCATATTGTT 2300
H11-2  CCCTTGGATC AAAGGACTGT ATAAAGCAAT ATAAGGATAT CTTCTACGAT 2350
H11-2  GAGGTTATGC CCAAGTGTAA GGCCGGGGAA GCAGCAACCA AATGCGTTAA 2400
H11-2  GGTTTCCGCT CCTCTTCGAG CCAATGTTTA CTGTTATGGT GTACAGGAAG 2450
H11-2  GTGGTGAAGA AGCTTTTGAA AAGGTGATGG GGCTGTATCT AGCAGAAGAT 2500
H11-2  GTTCAACTGG AGAAGGGTAT CCTGTTCAAA GCCTTGGCAT GCCACAAAGA 2550
H11-2  TGTTACAGCT CTAAAAGAAC TTCTTTTGCG AGCCCTGGAC CGTAAATCGT 2600
H11-2  CGTTTGTGCG TCTTCAGGAT GTCCCTACCG CTTTCCGTGC TGTATCTGAA 2650

B2                                                    GGGA    4
H11-2  AACCCTGTGG GCGAAGAATT CATGTTCAAT TTCCTAATGG AGAGATGGGA 2700

B2     GGAAATCACT GCGAGCTTGG AAACAGAACA CAGAGCAGTT GATAAAGTGG   54
H11-2  GGAAATCACT GCGAGCTTGG AAACAGAACA CAGAGCAGTT GATAAAGTGG 2750

B2     TCGGCGCTTG TTGCACAGGA ATTCGCTCCC AACAACAAAT AGATCAGCTG  104
H11-2  TCGGCGCTTG TTGCACAGGA ATTCGCTCCC AACAACAAAT AGATCAGCTG 2800

B2     AAGAAGAATC TACAGAAGAA CAATGCGCAG GCTAAGAAGT TC.......  154
```

FIG. 4 (jatkuu)

```
H11-2  AAGAA...TC TACAGAAGAA CAATGCGCAG GCTAAGAAGT TCGGCTCATT 2850

B2     .......... .......... .....CATAA AATTGCCTGG ATCAAGAAAC  204
H11-2  CACCCAGGAA ATCGAAAAAG GAGAACATAA AATTGCCTGG ATCAAGAAAC 2900

B2     ATTTTCACAG ATTATCGGAA TTCTTCAAGA GAGCAAGATC ATAGCTTTTC
H11-2  ATTTTCACAG ATTATCGGAA TTCTTCAAGA GAGCAAGATC ATAGCTTTTC 2950

B2     ACACTGAGCT CCAATTTTAA CGTCTTCAAA CTAGGAGACA GTTTTGCTGA
H11-2  ACACTGAGCT CCAATTTTAA CGTCTTCAAA CTAGGAGACA GTTTTGCTGA 3000

B2     AAAGTCAGTT TCACATTTTC CGTTTGAATG CCATCCATTC GAATACAACC
H11-2  AAAGTCAGTT TCACATTTTC CGTTTGAATG CCATCCATTC GAATACAACC 3050

B2     AA...CCCCA TTTTAAGTAC CTTTCATTCA CAGTGATTAC TAAATTTCGA
H11-2  AATAATACCA TTTTAAGTAC CTTTCATTCA CAGTGATTAC TGAATTTCGA 3100

B2     ATATATTATG AAGCTTGTAT CTTGAACGTT ATGATCGGTG ACTTTCAATT
H11-2  ATATATCATG AAGCTTGTAT CTTGAACGTT ATGATCGGTG ACTTTCAATT 3150

B2     TATAGAGCTC ACTCTCCATT TTGTAGCTGT GATGACTTGC ATTTAAGACC
H11-2  TATAGAGCTC ACTCTCCATT TTGTAGCTGT GATGACTTGC ATTTAAGACC 3200

B2     CACCATTTAC CAGCCTATAA TCTTTCCCCA ATACATTCCA AACTCCGATC
H11-2  CACCATTTAC CAGCCTAGAA TCTTTCCCCA ATACATTCCA AACTCCGATC 3250

B2     ACCTCCACCG CTGACAATGC CCAGATTTGT TTCTTTGTCT GCTATCCATC
H11-2  ACCTCCACCG CTGACAATGC CCAGATTTGT TTTTTTGTCT GCTATCCATC 3300

B2     TAACTGTTTC GAT
H11-2  TAACTGTTTC GATCGCCGGT TGTTTGTCAA TTGCTTATCT GATAAATATT 3350

H11-2  GACGTTGGTG T                                          3361
```

FIG. 4 (jatkuu)

```
          1                                                        50
H11-1     GGTTTAATTA CCCAAGTTTG AGATGACAGC AGAGGAGAGT CAGGAGCAGG

          51                                                       100
H11-1     AGACGCAGCA ACCACGAAAA AATACAGTGC TACGGCTCAC CCCAATCAAG

          101                                                      150
H11-1     TCTCTCTTTG CTTTGTTAGT GGTAGCTGCT GCCGTCGGCC TCTCAATCGG

          151                                                      200
H11-1     TCTCACCTAT TACTTTACAA GGAAAGCTTT TGATACTACT GGCGGAAATG

          201                                                      250
H11-1     GAAAAGGGGA TCAACCTATT GTCGATGATA ATTCCCCATC AGCTGAAGAA

          251                                                      300
H11-1     TTACGTCTCC CAACAACCAT AAAACCTTTG ACATACGACT TAGTAATCAA

          301                                                      350
H11-1     AACGTATCTG CCAAACTATG TAAACTATCC ACCTGAGAAA GATTTCGCTA

          351                                                      400
H11-1     TTGATGGGAC TGTGGTGATT GCTATGGAAG TTGTGGAGCC AACAAAGTCT

          401                                                      450
H11-1     ATTGTGCTCA ACTCGAAAAA TATTCCTGTA ATTGCAGACC AGTGCGAACT

          451                                                      500
H11-1     GTTTTCTAAC AACCAAAAAC TCGACATCGA AAAGGTTGTG GATCAGCCAA

          501                                                      550
H11-1     GGCTGGAGAA AGTCGAATTC GTTTTGAAGA AAAAGCTGGA GAAGAATCAG

          551                                                      600
H11-1     AAAATCACGC TCAAGATTGT ATACATTGGC CTTATCAACG ACATGCTTGG

          601                                                      650
H11-1     AGGACTTTAT CGAACAACCT ACACGGATAA AGATGGTACA ACCAAGATTG

          651                                                      700
H11-1     CTGCATGCAC TCATATGGAA CCGACGGACG CCCGTCTTAT GGTCCCCTGT

          701                                                      750
H11-1     TTCGACGAGC CGACGTTTAA GGCAAACTGG ACTGTGACTG TGATTCATCC

          751                                                      800
H11-1     GAAGGGCACC AGTGCCGTGT CGAATGGAAT AGAAAAGGGA GAAGGAGAAG

          801                                                      850
H11-1     TCTCTGGCGA TTGGGTCACA ACCAGATTCG ATCCAACCCC GCGAATGCCT

          851                                                      900
H11-1     TCGTATTTGA TTGCTCTTGT GATTTCCGAA TTTAAGTACA TTGAAAATTA

          901                                                      950
AustB1                                AATTCC GGCTCGTCCG GAAGCTATGA
H11-1     TACGAAAAGC GGTGTTCGAT TCCGAATTCC GGCTCGTCCG GAAGCTATGA
```

FIG. 5

```
          951                                                      1000
AustB1    AGATGACAGA ATATGCCATG ATAGCTGGAA TCAAATGTTT GGATTACTAT
 H11-1    AGATGACAGA ATATGCCATG ATAGCTGGAA TCAAATGTTT GGATTACTAT

         1001                                                     1050
AustB1    GAGGACTTCT TCGGGATCAA ATTCCCACTT CCAAAACAAG ATATGGTTGC
 H11-1    GAGGACTTCT TCGGGATCAA ATTCCCACTT CCAAAACAAG ATATGGTTGC

         1051                                                     1100
AustB1    TCTTCCTGAC TTCTCATCTG GTGCTATGGA GAACTGGGGT CTCATCACAT
 H11-1    TCTTCCTGAC TTCTCATCTG GTGCTATGGA GAACTGGGGT CTCATCACAT

         1101                                                     1150
AustB1    ACAGGGAGGG TTCCGTGCTC TACGATGAAA ACCTCTACGG ACCAATGAAT
 H11-1    ACAGGGAGGG TTCCGTGCTC TACGATGAAA ACCTCTACGG ACCAATGAAT

         1151                                                     1200
AustB1    AAGGAGCGGG TTGCAGAAGT GATCGCGCAC GAACTTGCAC ATCAGTGGTT
 H11-1    AAGGAGCGGG TTGCAGAAGT GATCGCGCAC GAACTTGCAC ATCAGTGGTT

         1201                                                     1250
AustB1    CGGTAATTTG GTCACGATGA AGTGGTGGGA TAACCTATGG CTGAACGAAG
 H11-1    CGGTAATTTG GTCACGATGA AGTGGTGGGA TAACCTATGG CTGAACGAAG

         1251                                                     1300
AustB1    GATTCGCGTC ATTCGTGGAA TACATCGGAG CCGACTTCAT CAGCGATGGT
 H11-1    GATTCGCGTC ATTCGTGGAA TACATCGGAG CCGACTTCAT CAGCGATGGT

         1301                                                     1350
AustB1    CTATGGGAAA TGAAAGATTT CTTCCTGCTG GCACCGTACA CAAGTGGTAT
 H11-1    CTATGGGAAA TGAAAGATTT CTTCCTGCTG GCACCGTACA CAAGTGGTAT

         1351                                                     1400
AustB1    TACGGCTGAT GCAGTAGCTT CAAGCCATCC GCTTTCCTTC AGAATAGATA
 H11-1    TACGGCTGAT GCAGTAGCTT CAAGCCATCC GCTTTCCTTC AGAATAGATA

         1401                                                     1450
AustB1    AGGCTGCAGA TGTATCAGAA GCGTTCGATG ATATCACATA CCGTAAAGGA
 H11-1    AGGCTGCAGA TGTATCAGAA GCGTTCGATG ATATCACATA CCGTAAAGGA

         1451                                                     1500
AustB1    GCATCCGTTC TTCAAATGCT ATTGAATTTA GTTGGGGACG AAAATTTCAA
 H11-1    GCATCCGTTC TTCAAATGCT ATTGAATTTA GTTGGGGACG AAAATTTCAA

         1501                                                     1550
AustB1    GCAGTCTGTT TCGCGTTACC TCAAGAAGTT TTCATATGAT AATGCGGCTG
 H11-1    GCAGTCTGTT TCGCGTTACC TCAAGAAGTT TTCATATGAT AATGCGGCTG

         1551                                                     1600
AustB1    CTGAAGATTT ATGGGCAGCA TTCGACGAAA CCGTCCAAGG TATAACCGGA
 H11-1    CTGAAGATTT ATGGGCAGCA TTCGACGAAA CCGTCCAAGG TATAACCGGA

         1601                                                     1650
AustB1    CCTAATGGTG GACCATTGAA AATGTCCGAG TTTGCGCCAC AATGGACAAC
 H11-1    CCTAATGGTG GACCATTGAA AATGTCCGAG TTTGCGCCAC AATGGACAAC

         1651                                                     1700
AustB1    TCAGATGGGG TTCCCTGTTC TTACTGTCGA GTCGGTTAAC GCAACGACTT
```

FIG. 5 (jatkuu)

```
H11-1   TCAGATGGGG TTCCCTGTTC TTACTGTCGA GTCGGTTAAC GCAACGACTT

        1701                                                1750
AustB1  TGAAAGTCAC CCAAAAACGA TACAGGCAGA ACAAGGATGC AAAGGAACCA
H11-1   TGAAAGTCAC CCAAAAACGA TACAGGCAGA ACAAGGATGC AAAGGAACCA

        1751                                                1800
AustB1  GAGAAGTACC GTCATCCAAC TTATGGGTTC AAATGGGATG TTCCTCTGTG
H11-1   GAGAAGTACC GTCATCCAAC TTATGGGTTC AAATGGGATG TTCCTCTGTG

        1801                                                1850
AustB1  GTATCAGGAA GATGAACAGC AAGTGAAAAG AACTTGGTTA AAAAGAGAGG
H11-1   GTATCAGGAA GATGAACAGC AAGTGAAAAG AACTTGGTTA AAAAGAGAGG

        1851                                                1900
AustB1  AACCGCTCTA TTTCCATGTA AGCAATTCTG ATTCGTCAGT TGTGGTGAAT
H11-1   AACCGCTCTA TTTCCATGTA AGCAATTCTG ATTCGTCAGT TGTGGTGAAT

        1901                                                1950
AustB1  GCCGAACGTC GTGCTTTTTG CCGATCAAAC TATGACGCTA ACGGTTGGAG
H11-1   GCCGAACGTC GTGCTTTTTG CCGATCAAAC TATGACGCTA ACGGTTGGAG

        1951                                                2000
AustB1  GAACATTATG AGAAGACTCA AGCAGAATCA TAAGGTCTAT GGTCCACGAA
H11-1   GAACATTATG AGAAGACTCA AGCAGAATCA TAAGGTCTAT GGTCCACGAA

        2001                                                2050
AustB1  CAAGAAACGC TCTCATAAGT GATGCGTTTG CAGCAGCTGC AGTTGAGGAA
H11-1   CAAGAAACGC TCTCATAAGT GATGCGTTTG CAGCAGCTGC AGTTGAGGAA

        2051                                                2100
AustB1  ATGAATTACG AGACCGTATT TGAAATGCTC AAATACACCG TGAAAGAAGA
H11-1   ATGAATTACG AGACCGTATT TGAAATGCTC AAATACACCG TGAAAGAAGA

        2101                                                2150
AustB1  GGATTACTTA CCATGGAAGG AGGCAATATC AGGATTCAAT ACAATTTTGG
  B1a                             GCAATATC AGGATTCAAT ACAATTTTGG
H11-1   GGATTACTTA CCATGGAAGG AGGCAATATC AGGATTCAAT ACAATTTTGG

        2151                                                2200
AustB1  ACTTTTTCGG CAGCGAACCC GAATCTCAAT GGGCTTCGGA ATACATGCGA
  B1a   ACTTTTTCGG CAGCGAACCC GAATCTCAAT GGGCTTCGGA ATACATGCGA
H11-1   ACTTTTTCGG CAGCGAACCC GAATCTCAAT GGGCTTCGGA ATACATGCGA

        2201                                                2250
AustB1  AAACTGATGA AGCCAATTTA TGACAAGAGT AGCATCAAGT TTATAGCGGA
  B1a   AAACTGATGA AGCCAATTTA TGACAAGAGT AGCATCAAGT TTATAGCGGA
H11-1   AAACTGATGA AGCCAATTTA TGACAAGAGT AGCATCAAGT TTATAGCGGA

        2251                                                2300
AustB1  GAACTACAAA AAAGATTCGC TTTTCTTCAA AAATAATCTC CAAATAGCTG
  B1a   GAACTACAAA AAAGATTCGC TTTTCTTCAA AAATAATCTC CAAATAGCTG
H11-1   GAACTACAAA AAAGATTCGC TTTTCTTCAA AAATAATCTC CAAATAGCTG

        2301                                                2350
AustB1  TTATTGACAC ATACTGTGGT CTTGGAGGCA AAGAATGTCT TGAAGAAATG
  B1a   TTATTGACAC ATACTGTGGT CTTGGAGGCA AAGAATGTCT TGAAGAAATG
H11-1   TTATTGACAC ATACTGTGGT CTTGGAGGCA AAGAATGTCT TGAAGAAATG
```

FIG. 5 (jatkuu)

```
          2351                                                      2400
AustB1    AAAAAGCTTT TTGACAAGGA GGTCATGAAA TGTCAACCTG GTCAGCAAGC
   B1a    AAAAAGCTTT TTGACAAGGA GGTCATGAAA TGTCAACCTG GTCAGCAAGC
 H11-1    AAAAAGCTTT TTGACAAGGA GGTCATGAAA TGTCAACCTG GTCAGCAAGC

          2401                                                      2450
AustB1    GACCGACTGC GTAAAGGTAA CTGCTCCTCT CCGAAAAACT GTTTACTGCT
   B1a    GACCGACTGC GTAAAGGTAA CTGCTCCTCT CCGAAAAACT GTTTACTGCT
 H11-1    GACCGACTGC GTAAAGGTAA CTGCTCCTCT CCGAAAAACT GTTTACTGCT

          2451                                                      2500
AustB1    ATGGGGTCCA GGAAGGCGGT GATGAGGCAT TCGACAAGGT GATGGAACTA
   B1a    ATGGGGTCCA GGAAGGCGGT GATGAGGCAT TCGACAAGGT GATGGAACTA
 H11-1    ATGGGGTCCA GGAAGGCGGT GATGAGGCAT TCGACAAGGT GATGGAACTA

          2501                                                      2550
AustB1    TATAATGCGG AACAAGTGCA GTTGGAGAAA GACAGTCTAC GTGAAGCATT
   B1a    TATAATGCGG AACAAGTGCA GTTGGAGAAA GACAGTCTAC GTGAAGCATT
 H11-1    TATAATGCGG AACAAGTGCA GTTGGAGAAA GACAGTCTAC GTGAAGCATT

          2551                                                      2600
AustB1    GGGATGCCAT AAAGACGTTA CAGCTCTAAA GGGACTTCTT ATGCTGGCTT
   B1a    GGGATGCCAT AAAGACGTTA CAGCTCTAAA GGGACTTCTT ATGCTGGCTT
 H11-1    GGGATGCCAT AAAGACGTTA CAGCTCTAAA GGGACTTCTT ATGCTGGCTT

          2601                                                      2650
AustB1    TGGATCGGAA TTC
   B1a    TGGATC
 H11-1    TGGATCGGAA TTCGTCATTT GTGCGTCTTC AAGATGCTCA TGATGTGTTT

          2651                                                      2700
 H11-1    AACATTGTAT CCAGAAATCC TGTTGGAAAC GAACTGCTGT TCAATTTCCT

          2701                                                      2750
 H11-1    CACAGAGCGA TGGGAAGAGA TACTTGAAAG TTTGTCAATA CGACACAGAT

          2751                                                      2800
 H11-1    CAGTTGATCG AGTGATCAAA GCCTGTACTC GAGGACTACG ATCCAGGGAA

          2801                                                      2850
 H11-1    CAAGTACAAC AGTTGAAGAA TCTATACAAA AATGACAAGC GTGCTCGCGA

          2851                                                      2900
 H11-1    ATACGGTGCA TTTGGTGGGG CAATAGAAAG ATCGGAACAC AGAGTCAAAT

          2901                                                      2950
 H11-1    GGATTGAGAA ACATTTCCGA AAACTAGCAG CTTTCTTCAA AAAATCTAAT

          2951                                                      3000
 H11-1    TCATAATTCT GAAATGGCTA TAACTAGCAC ACTGGATAGT TGTCTCGAAT

          3001                                                      3050
 H11-1    CATCCAAAAA GATTAATGAT GTTTTTTTAC TAGATAATAT GGAGATATTC

          3051                                        3084
 H11-1    TGTAAATTTG TCATCGATTC AAGTGTCTGT ATTG
```

FIG. 5 (jatkuu)

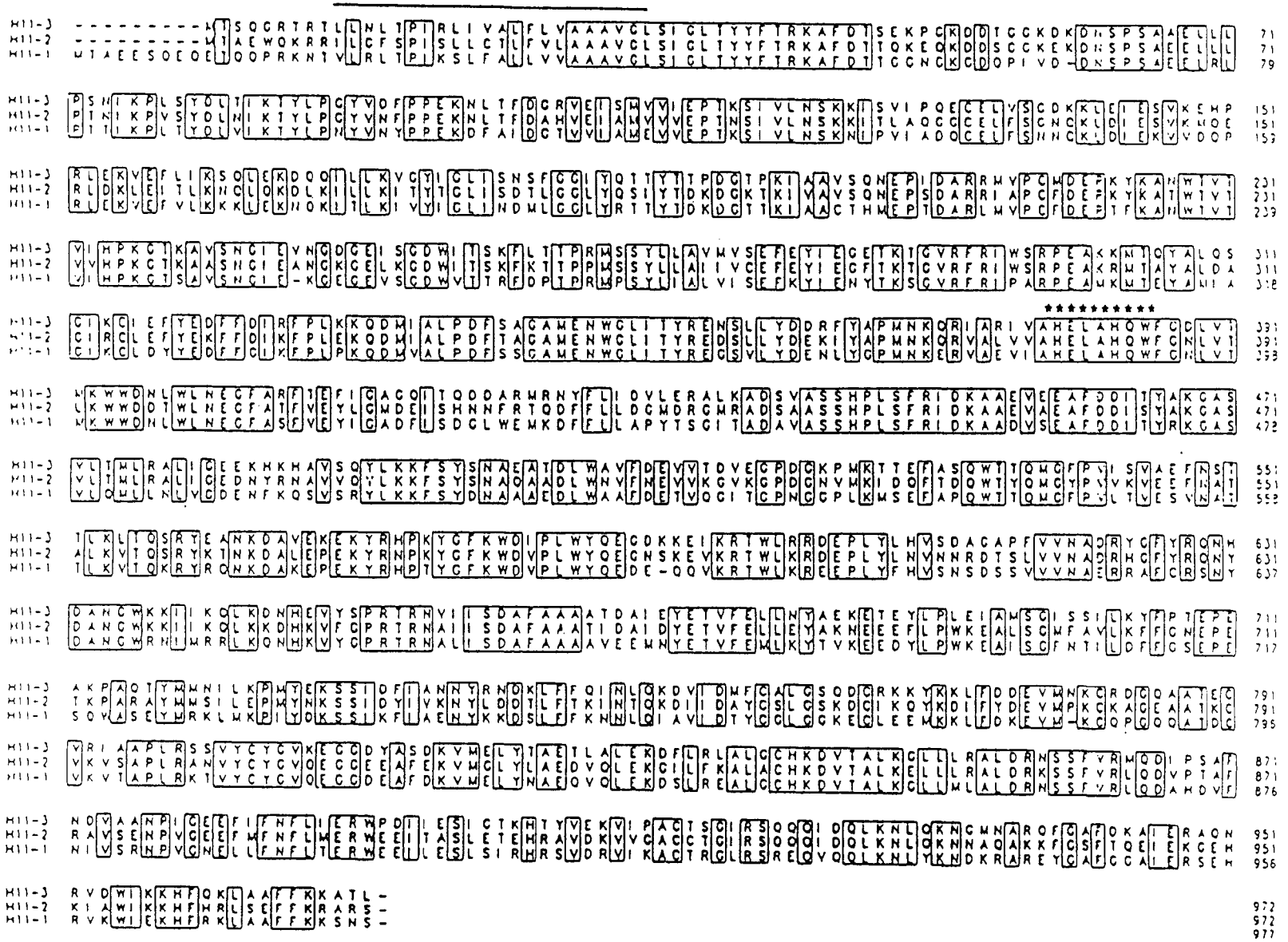


FIG. 6a



```
H11-3   MTSQGRTRTLLNLTPIRLIVALFLVAAAVGLSIGLTYYFTRKAFDTSEK-----------PGKDDTGGKDKDNSPSAAEL   69
Rat-Apm ------MAKGFYISKTLGILGILLGVAAVCTIIALSVVYAQEKNRNAENSAIAPTLPGSTSATTSTTNPAIDESKPANQY   74

H11-3   LLPSNIKPLSTDLIIKTYLPGYVDFPPEKNL-TFDGRVEISMVVIEPTKSIVLNSKKISVIPQEGELV-----SGDKKLE   143
Rat-Apm RLPKTLIPDSYQVTLRPYLT-----PNEQGLYIFKGSSTVRFTCNETTNVIIIHSKKLNYTNKGNHRVALRALGDTPAPN   143

H11-3   IESVKEHPRLEKVEFLIKSQLEKDQQILLKVGYIGLISNSFGGIYQTTYTTPDGTPKIAAVSQNEPIDARRMVPCMDEPK   223
Rat-Apm IDTTELVERTEYLVVHLQGSLVKGHQYEMDSEFQGELADDLAGFYRSEYME-GGNKKVVATTQMQAADARKSFPCFDEPA   225

H11-3   YKANWTVTVIHPKGTKAVSNGIEVNGDG-EISGDWITSKFLTTPRMSSYLLAVMVSEFETIEGETKTGVRFRIWSRPEAK   302
Rat-Apm MKASFNITLIHPNNLTALSNMLPKDSRTLQEDPSWNVTEFHPTPKMSTYLLAYIVSEFKYVEAVSPNRVQIRIWARPSAI   305

H11-3   KM--TQYALQSGIKCIEFYEDFFDIRFPLKKQDMIALPDFSAGAMENWGLITYRENSLLYDDRFYAPMNKQRIARIVAHE   380
Rat-Apm DEGHGDYALQVTGPILNFFAQHYNTAYPLEKSDQIALPDFNAGAMENWGLVTYRESALVFDPQSSSISNKERVVTVIAHE   388

H11-3   LAHQWFGDLVTMKWWDNLWLNEGFARFTEFIGAGQITQDDARMRNYFLIDVLERALKADSVASSHPLSFR---IDKAAEV   457
Rat-Apm LAHQWFGNLVTVDWWNDLWLNEGFASYVEFLGA-DYAEPTWNLKDLIVLNDVYRVMAVDALASSHPLSSPANEVNTPAQI   457

H11-3   EEAFDDITYAKGASVLTMLRALIGEEKHKHAVSQYLKKFSYSNAEATDLWAVFDEVVTDVEGPDGKPMKTTEFASQWTTQ   537
Rat-Apm SELFDSITYSKGASVLRMLSSFLTEDLFKKGLSSYLHTFQYSNTIYLDLWEHLQQAV-DSQTAIKLPASVSTIMDPWILQ   540

H11-3   MGFPVISVAEFNSTILKLTQSRYEANKDAVEKEKYRHPKYGFKWDIPLWYQEGDKKEIKRTWLRRDEPLYLHVSDAGAPF   517
Rat-Apm MGFPVITV---NTSTGEIYQEHFLLDPTS---KPTRPSDFNYLWIVPIPYLKNGKEDHYWLETEKNQSAEFQTS-SNEWL   519

H11-3   VVNADRYGFYRQNHDANGWKKIIKQLKDNHEVYSPRTRNVIISDAFAAAATDAIEYETVFELLNYAEKETEYLPLEIAMS   597
Rat-Apm LLNINVTGYYQVNYDENNWRKIQNQLQTDLSVIPVINRAQIIHDSFNLASAGKLSITLPLSNTLFLASETEYMPWEAALS   599

H11-3   GISSILKYFPTEPEAKPAQTYMMNILKPMYEKSSIDFIANNYRNDKLFFQINLQKDVIDMFGALGSQDCRKKYKKLFDDE   777
Rat-Apm SLNYFKLMFDRSEVYGPMKRYLKKQVTPLFAYFKIKTNNWLDRPPTLMEQYN-EINAISTACSSGLEECR--------DL   770

H11-3   VMNKCRDGQAATECVRIAAPLRSSVYCYGV---KEGQDYASDKVMELYTAETLALEKDFLRLALGCHKDVTALKGLLLRA   854
Rat-Apm VVGLYSQWMNNSDNNPIHPNLRSTVYCNAIASCEEAWNFAWATVPE----RTLVNEADKLRSAVGRSNEVWILNRYLSYT   845

H11-3   LDRNSSFYRMQDIPSAFNDYAANPIGEEFIFNFLIERWPDIIESIGTKHTYVEKVIPACTSGIRSQQQIDQLKNLQKNGM   934
Rat-Apm L--NPDYIRKQDATSTIVSIANNVVGQTLVWDFVRSNWKKLFEDYGGGSFSFANLIQGVTRRFSSEFELQQLEQF-KEDN   923

H11-3   NARQFG----AFDKAIERAQNRVDWIKKHFQKLAAFFKKATL-                                          972
Rat-Apm SATGFGSGTRALEQALEKTKANIKWVKENKDVVLKWFTENS--                                          964
```

FIG. 6b1



```
H11-3   MTSQGRTRTLLNLTPIRLIVALFLVAAAVGLSIGLTYYFTRKAFD-------------TSEKPGKDDTGGKDKDNSPSA   65
Mouse-Apo  MNFAEEEPSKKYCIKGKHVAIICGVVVAVGLIVGLSVGLTRSCEQDTTPAPSQPPPEASTALPPQDQNVCPDSEDESGEW   80

H11-3   AELLLPSNIKPLSYDLTIKTYLPGYVDFPPEKNLTFDGRVEISMVVIEPTKSIVLNSKKISVIPQECELVSGDKKLEIES  145
Mouse-Apo  KNFRLPDFINPVHYDLEVKALM---------EEDRYTGIVTISVNLSKPTRDLWLHIRETKITKLPELRRPSGEQVPIRR  151

H11-3   VKEHPRLEKVEFLIKSQLEK---DQQILLKVGYIGLISNSFGGIYQTTYTTPDGTPKIAAVSGNEPIDARRMVPCMDEPK  223
Mouse-Apo  CFTYKKGEYVVIQAAEDLAATSGDSVYRLTMEFKGWLNGSLVGFYKTTYMEDGQIRSIAA-TDHEPTDARKSFPCFDEPN  223

H11-3   YKANWTVTVIHPKGTKAVSNGIEVNGDGEISGDWITSKFLTTPRMSSYLLAVMVSEFEYIEGETKTGVRFRIWSRPEAKK  303
Mouse-Apo  KKSTYSISIIHPKEYSALSNMPEEKSE-MVDDNWKKTTFVKSVPMSTYLVCFAVHRFTAIERKSRSGKPLKVYVQPNQKE  309

H11-3   MTQYALQSGIKCIEFYEDFFDIRFPLKKQDMIALPDFSAGAMENWGLITYRENSLLYDDRFYAPMNKQRIARIVAHELAH  383
Mouse-Apo  TAEYAANITQAVFDYFEDYFAMEYALPKLDKIAIPDFGTGAMENWGLVTYRETNLLYDPLLSASSNQQRVASVVAHELVH  389

H11-3   QWFGDLVTMKWWDNLWLNEGFARFTEFIGAGQITQDDARMRNYFLIDVLERALKADSVASSHPLSFRIDKAAEVEEAFDD  453
Mouse-Apo  QWFGNTVTMDWWDDLWLNEGFASFFEFLGVNHAEKDWQMLSQVLLEDVF-PVQEDDSLMSSHPVVVTVSTPAEITSVFDG  453

H11-3   ITYAKGASVLTMLRALIGEEKHKHAVSQYLKKFSYSNAEATDLWAVFDEVVTDVEGPDGKPMKTTEFASQWTTQMGFPVI  543
Mouse-Apo  ISYSKGASILRMLQDWITPEKFQKGCQIYLKKFQFANAKTSDFWDSLQE---------ASNLPVKEVMDTWTSQMGYPVV  539

H11-3   SVAEFNSTTLKLTQSRYEANKDAVEKEKYRHPKYGFKWDIPLWYQEGDKKEIKR-TWLRRDEPLYLHVSDAGAPFVVNAD  622
Mouse-Apo  TV----SGRQNITQKRFLLDSKADPSQ--PPSELGYTWNIPVRWADNDNSRITVYNRLDKGGITLNANLSGDAFLKINPD  613

H11-3   RYGFYRQNHDANGWKKIIKQLKDNHEVYSPRTRNVIISDAFAAAATDAIEYETVFELLNYAEKETEYLPLEIAMSGISSI  702
Mouse-Apo  HIGFYRVNYEGGTWDWIAEALSSNHTRFSAADRSSFIDDAFALARAQLLNYKIALNLTMYLKSEEDFLPWERVISSVSYI  693

H11-3   LKYFPTEPEAKP-AQTYMMNILKPMYEKSSIDFIANNYRNDKLFFQINLQKDVIDMFCALGSQDCRKKYKKLFDDEVMNK  781
Mouse-Apo  ISMFEDDRELYPMIETYFQGQVKPVADL--LGWQDTGSHITKL-----LRASILGFACKMGDREALGNASQLFDSWL---  763

H11-3   CRDGQAATECVRIAAPLRSSVYCYGVKEGG-DYASDKVMELYTAETLALEKDFLRLALGGHKDVTALKGLLLRALDRNSS  850
Mouse-Apo  -------KGSASIPVNLRLLVYRYGMQNSGNEAAWNYTLEQYQKTSLAQEKEKLLYGLASVKDVKLLARYL--EMLKDPN  831

H11-3   FVRMQDIPSAFNDVAANPIGEEFIFNFLIERWPDIIESIGTKHTYVEKVIPACTSGIRSQQQIDQLKN-LQKNGMNAROF  930
Mouse-Apo  IIKTQDVFTVIRYISYNSYGKTMAWNWIQLNWDYLVSRFTINDRYLGRIV-TIAEPFNTELQLWQMQSFFAKYPNAGAGA  913

H11-3   GAFDKAIERAQNRVDWIKKHFQKLAAFFKKATL-                                                  972
Mouse-Apo  KPREQVLETVKNNIEWLNVNRQSIREWFASLP--                                                  945
```

FIG. 6bII

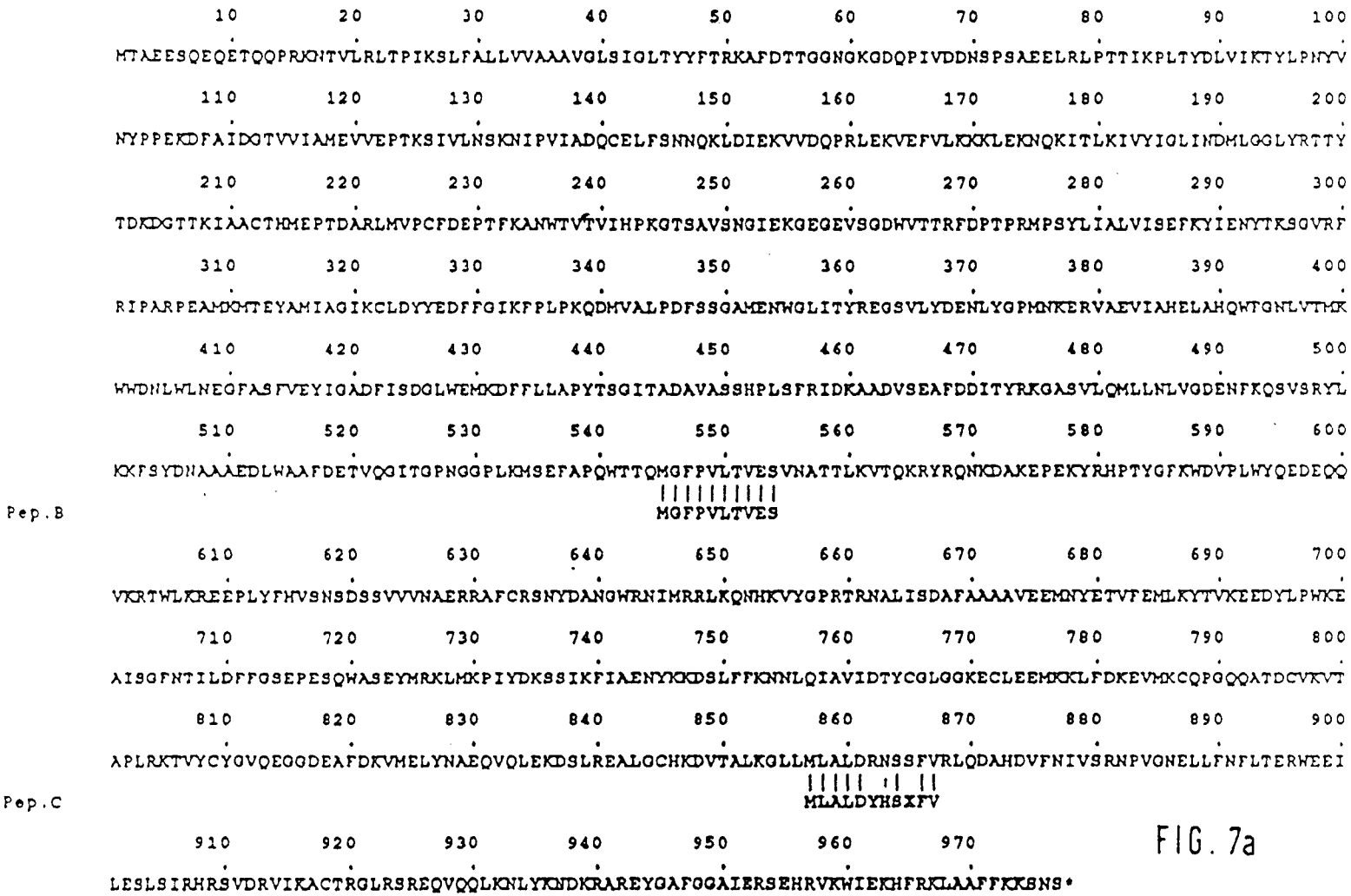

10 20 30 40 50 60 70 80 90 100
MTAEESQEQETQQPRKNTVLRLTPIKSLFALLVVAAAVGLSIGLTYYFTRKAFDTTGGNGKGDQPIVDDNSPSAEELRLPTTIKPLTYDLVIKTYLPNYV
110 120 130 140 150 160 170 180 190 200
NYPPEKDFAIDGTVVIAMEVVEPTKSIVLNSKNIPVIADQCELFSNNQKLDIEKVVDQPRLEKVEFVLKKKLEKNQKITLKIVYIGLINDMLGGLYRTTY
210 220 230 240 250 260 270 280 290 300
TDKDGTTKIAACTHMEPTDARLMVPCFDEPTFKANWTVTVIHPKGTSAVSNGIEKGEGEVSGDWVTTRFDPTPRMPSYLIALVISEFKYIENYTKSGVRF
310 320 330 340 350 360 370 380 390 400
RIPARPEAMDQHTEYAMIAGIKCLDYYEDFFGIKFPLPKQDMVALPDFSSGAMENWGLITYREGSVLYDENLYGPMNKERVAEVIAHELAHQWTGNLVTMK
410 420 430 440 450 460 470 480 490 500
WWDNLWLNEGFASFVEYIGADFISDGLWEMKDFFLLAPYTSGITADAVASSHPLSFRIDKAADVSEAFDDITYRKGASVLQMLLNLVGDENFKQSVSRYL
510 520 530 540 550 560 570 580 590 600
KKFSYDNAAAEDLWAAFDETVQGITGPNGGPLKMSEFAPQWTTQMGFPVLTVESVNATTLKVTQKRYRQNKDAKEPEKYRHPTYGFKWDVPLWYQEDEQQ
Pep.B
MGFPVLTVES
610 620 630 640 650 660 670 680 690 700
VKRTWLKREEPLYFHVSNSDSSVVVNAERRAFCRSNYDANGWRNIMRRLKQNHKVYGPRTRNALISDAFAAAAVEEMNYETVFEMLKYTVKEEDYLPWKE
710 720 730 740 750 760 770 780 790 800
AISGFNTILDFFGSEPESQWASEYMRKLMKPIYDKSSIKFIAENYKKDSLFFKNNLQIAVIDTYCGLGGKECLEEMKKLFDKEVMKCQPGQQATDCVKVT
810 820 830 840 850 860 870 880 890 900
APLRKTVYCYGVQEGGDEAFDKVMELYNAEQVQLEKDSLREALGCHKDVTALKGLLMLALDRNSSFVRLQDAHDVFNIVSRNPVGNELLFNFLTERWEEI
Pep.C
MLALDYHSXFV
910 920 930 940 950 960 970
LESLSIRHRSVDRVIKACTRGLRSREQVQQLKNLYKNDKRARFYGAFGGAIERSEHRVKWIEKHFRKLAAFFKKSNS*

FIG. 7a

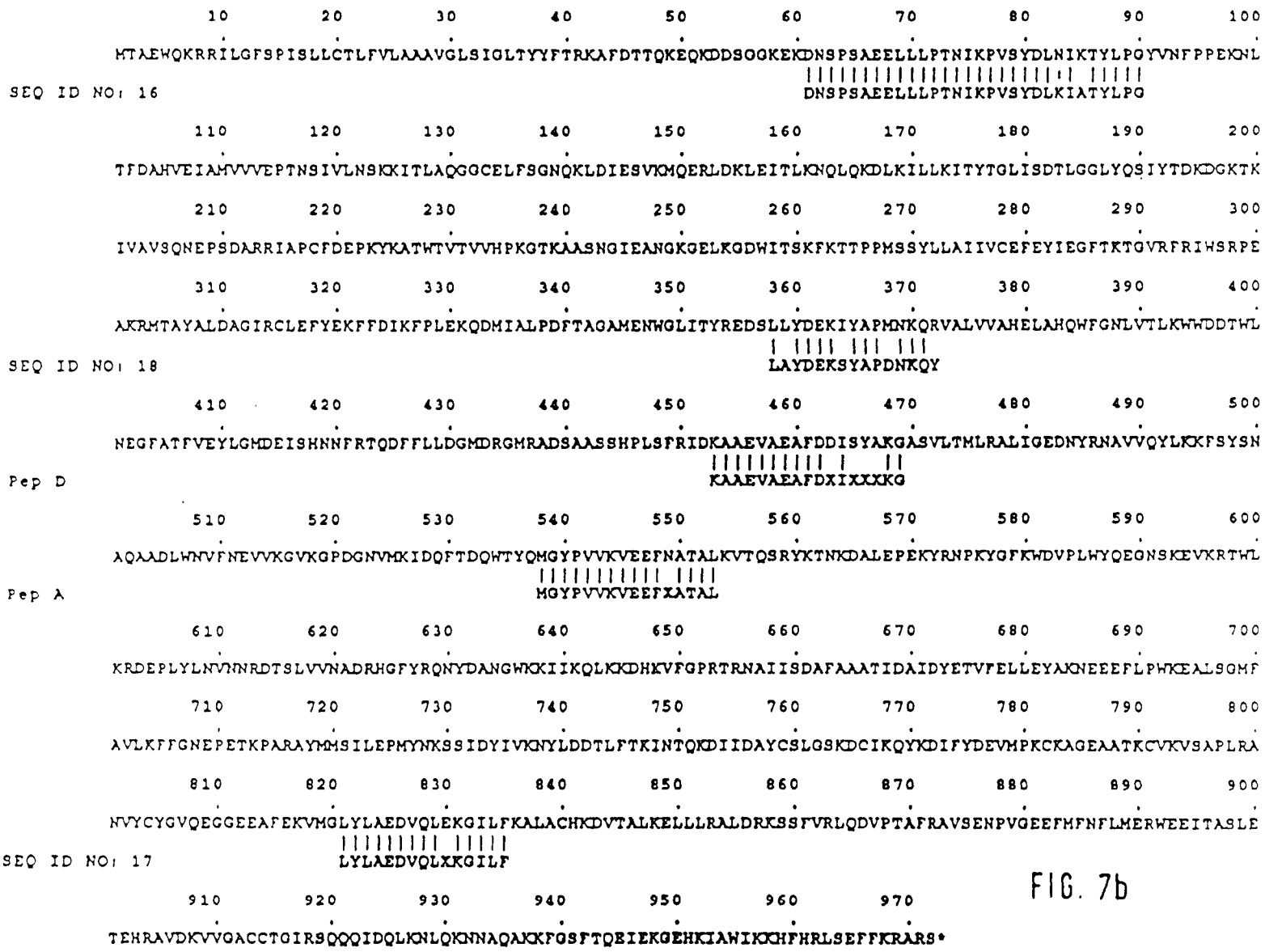

10 20 30 40 50 60 70 80 90 100
MTAEWQKRRILGFSPISLLCTLFVLAAAVGLSIGLTYYFTRKAFDTTQKEQKDDSGGKEKDNSPSAEELLLPTNIKPVSYDLNIKTYLPGYVNFPPEKNL
SEQ ID NO: 16
DNSPSAEELLLPTNIKPVSYDLKIATYLPG
110 120 130 140 150 160 170 180 190 200
TFDAHVEIAMVVVEPTNSIVLNSKKITLAQGGCELFSGNQKLDIESVKMQERLDKLEITLKNQLQKDLKILLKITYTGLISDTLGGLYQSIYTDKDGKTK
210 220 230 240 250 260 270 280 290 300
IVAVSQNEPSDARRIAPCFDEPKYKATWTVTVVHPKGTKAASNGIEANGKGELKGDWITSKFKTTPPMSSYLLAIIVCEFEYIEGFTKTGVRFRIWSRPE
310 320 330 340 350 360 370 380 390 400
AKRMTAYALDAGIRCLEFYEKFFDIKFPLEKQDMIALPDFTAGAMENWGLITYREDSLLYDEKIYAPMNKQRVALVVAHELAHQWFGNLVTLKWWDDTWL
SEQ ID NO: 18
LAYDEKSYAPDNKQY
410 420 430 440 450 460 470 480 490 500
NEGFATFVEYLGMDEISHNNFRTQDFFLLDGMDRGMRADSAASSHPLSFRIDKAAEVAEAFDDISYAKGASVLTMLRALIGEDNYRNAVVQYLKKFSYSN
Pep D
KAAEVAEAFDXIXXXKG
510 520 530 540 550 560 570 580 590 600
AQAADLWNVFNEVVKGVKGPDGNVMKIDQFTDQWTYQMGYPVVKVEEFNATALKVTQSRYKTNKDALEPEKYRNPKYGFKWDVPLWYQEGNSKEVKRTWL
Pep A
MGYPVVKVEEFXATAL
610 620 630 640 650 660 670 680 690 700
KRDEPLYLNVVNNRDTSLVVNADRHGFYRQNYDANGWKKIIKQLKKDHKVFGPRTRNAIISDAFAAATIDAIDYETVFELLEYAKNEEEFLPWKEALSGMF
710 720 730 740 750 760 770 780 790 800
AVLKFFGNEPETKPARAYMMSILEPMYNKSSIDYIVKNYLDDTLFTKINTQKDIIDAYCSLGSKDCIKQYKDIFYDEVMPKCKAGEAATKCVKVSAPLRA
810 820 830 840 850 860 870 880 890 900
NVYCYGVQEGGEEAFEKVMGLYLAEDVQLEKGILFKALACHKDVTALKELLLRALDRKSSFVRLQDVPTAFRAVSENPVGEEFMFNFLMERWEEITASLE
SEQ ID NO: 17
LYLAEDVQLXKGILF
910 920 930 940 950 960 970
TEHRAVDKVVGACCTGIRSQQQIDQLKNLQKNNAQAKKFGSFTQEIEKGEHKIAWIKKHFHRLSEFFKRARS*

FIG. 7b

FIG. 8

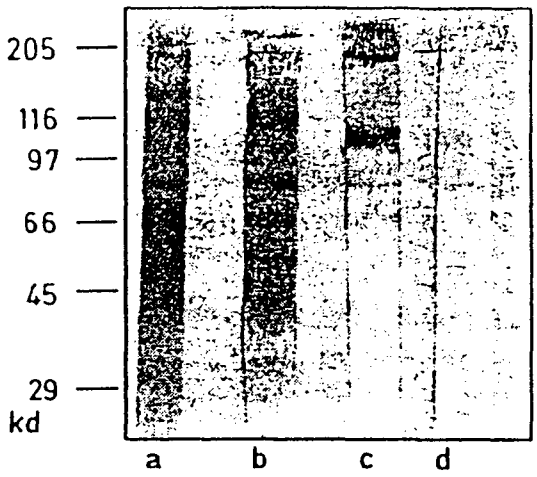
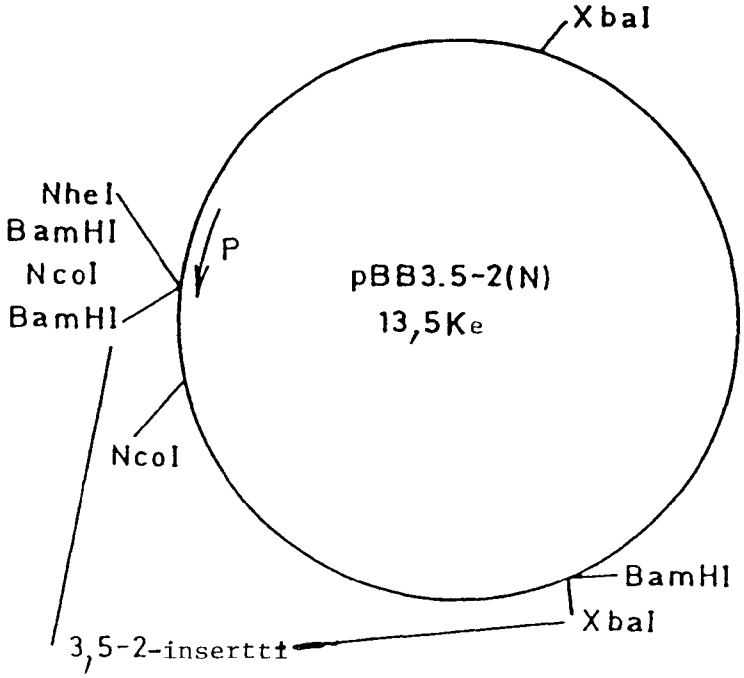


FIG. 20



P = polyhedriinipromoottori

FIG. 9

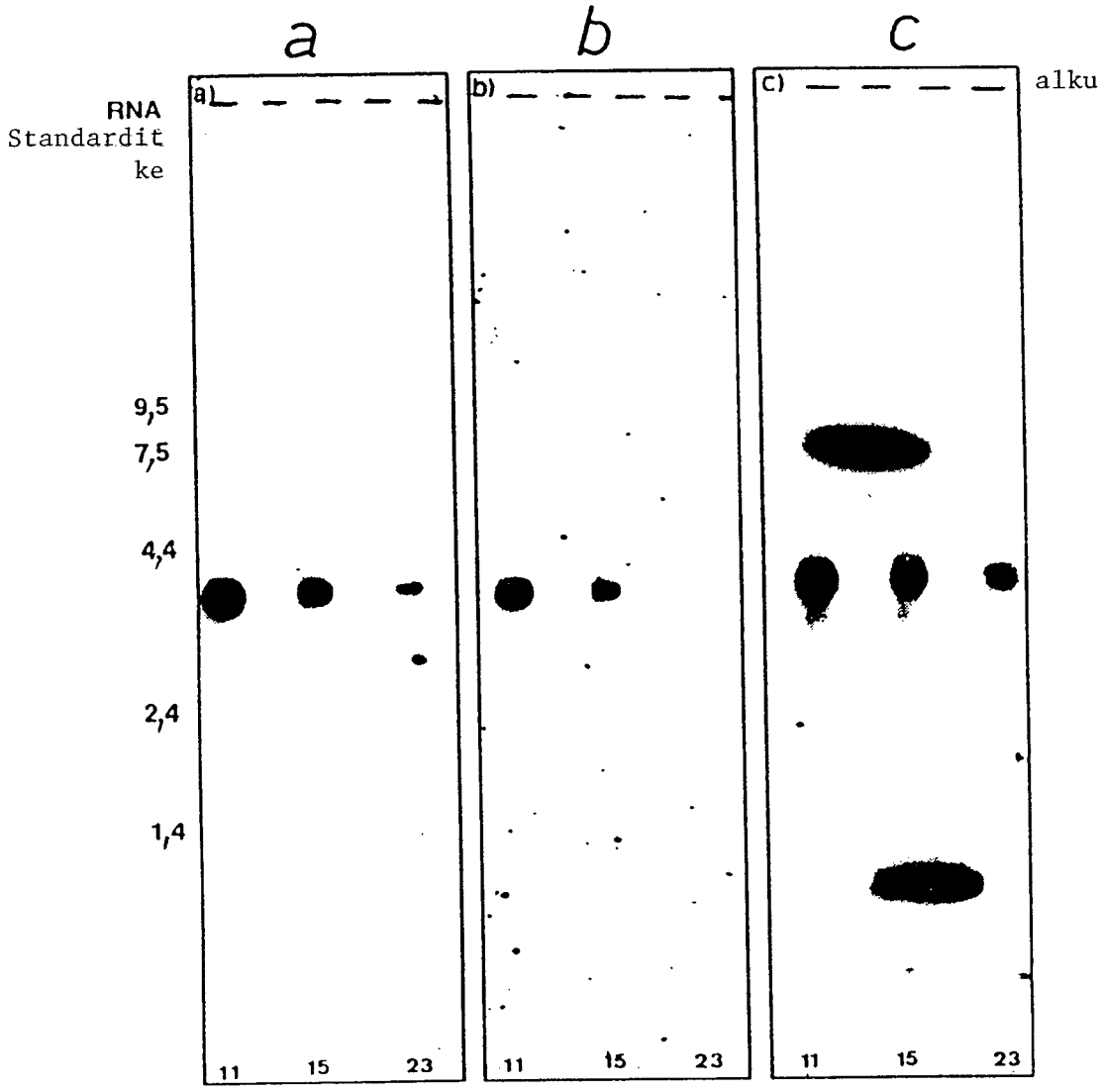

a
b
c
RNA
Standardit
ke
a)
b)
c)
alku
9,5
7,5
4,4
2,4
1,4
11
15
23
11
15
23
11
15
23

FIG. 9

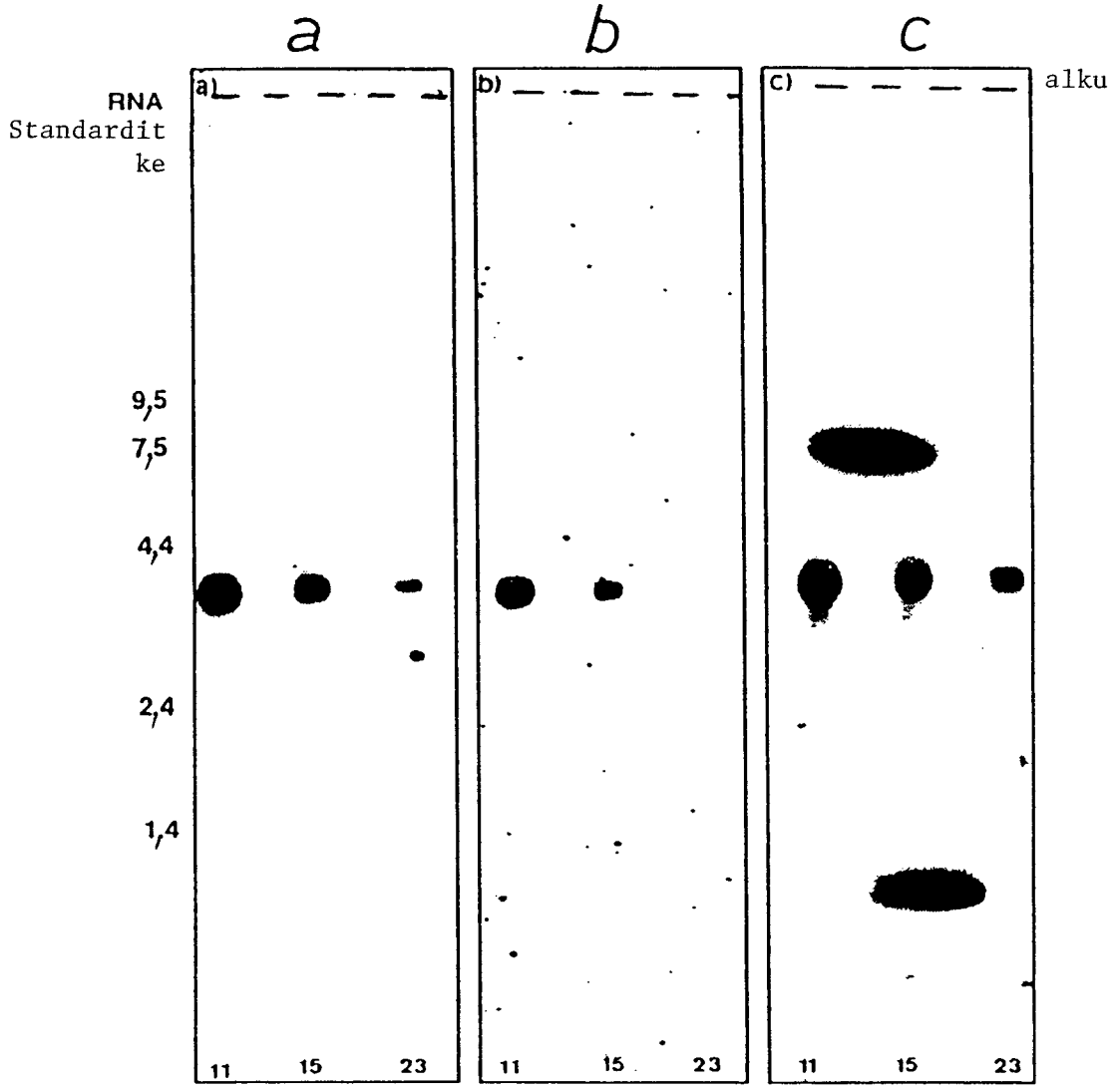

a
b
c
RNA
Standardit
ke
a)
b)
c)
alku
9,5
7,5
4,4
2,4
1,4
11
15
23
11
15
23
11
15
23

FIG. 9

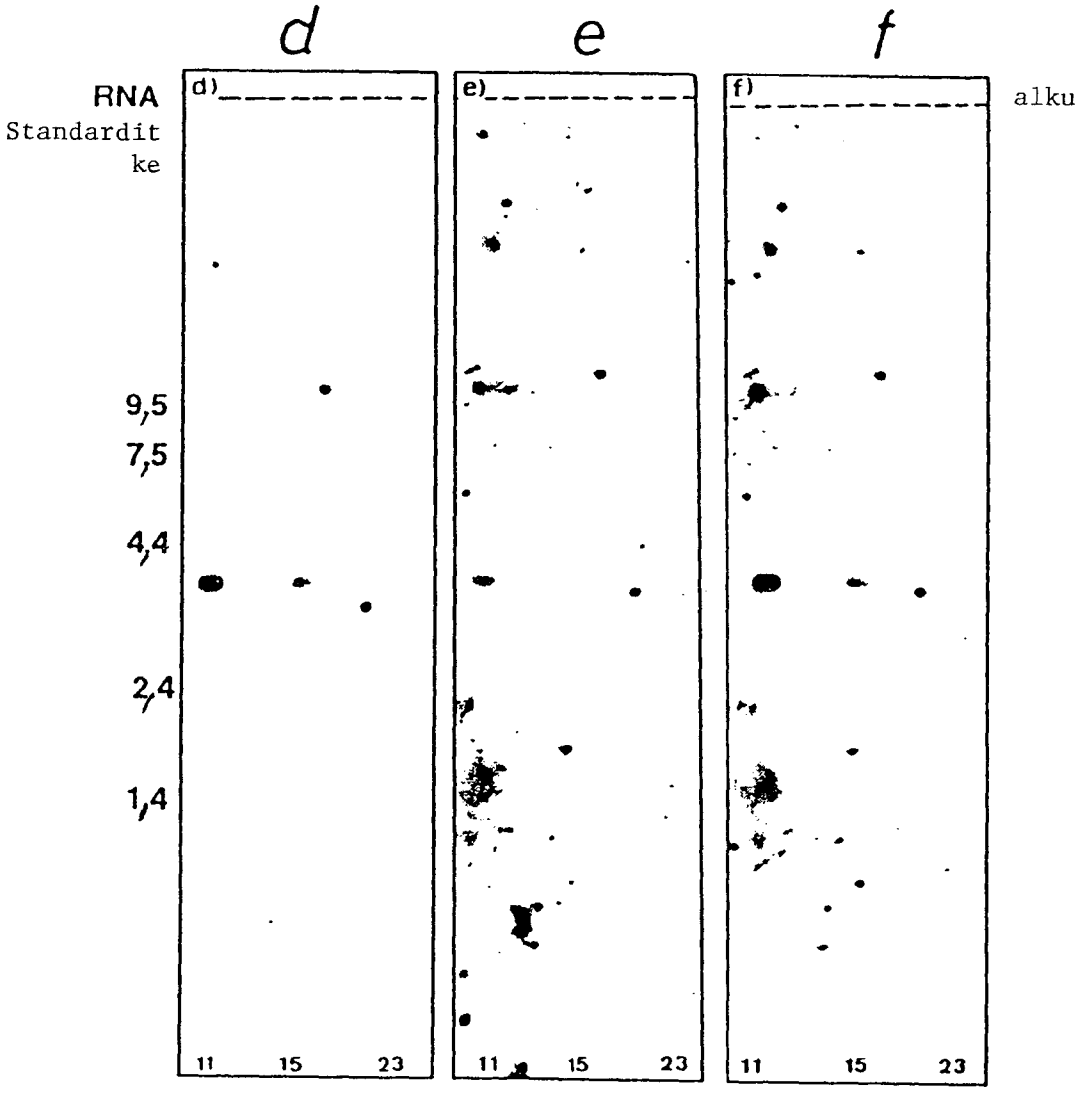

d
e
f
RNA
Standardit
ke
d)
e)
f)
alku
9,5
7,5
4,4
2,4
1,4
11
15
23
11
15
23
11
15
23

FIG .10a

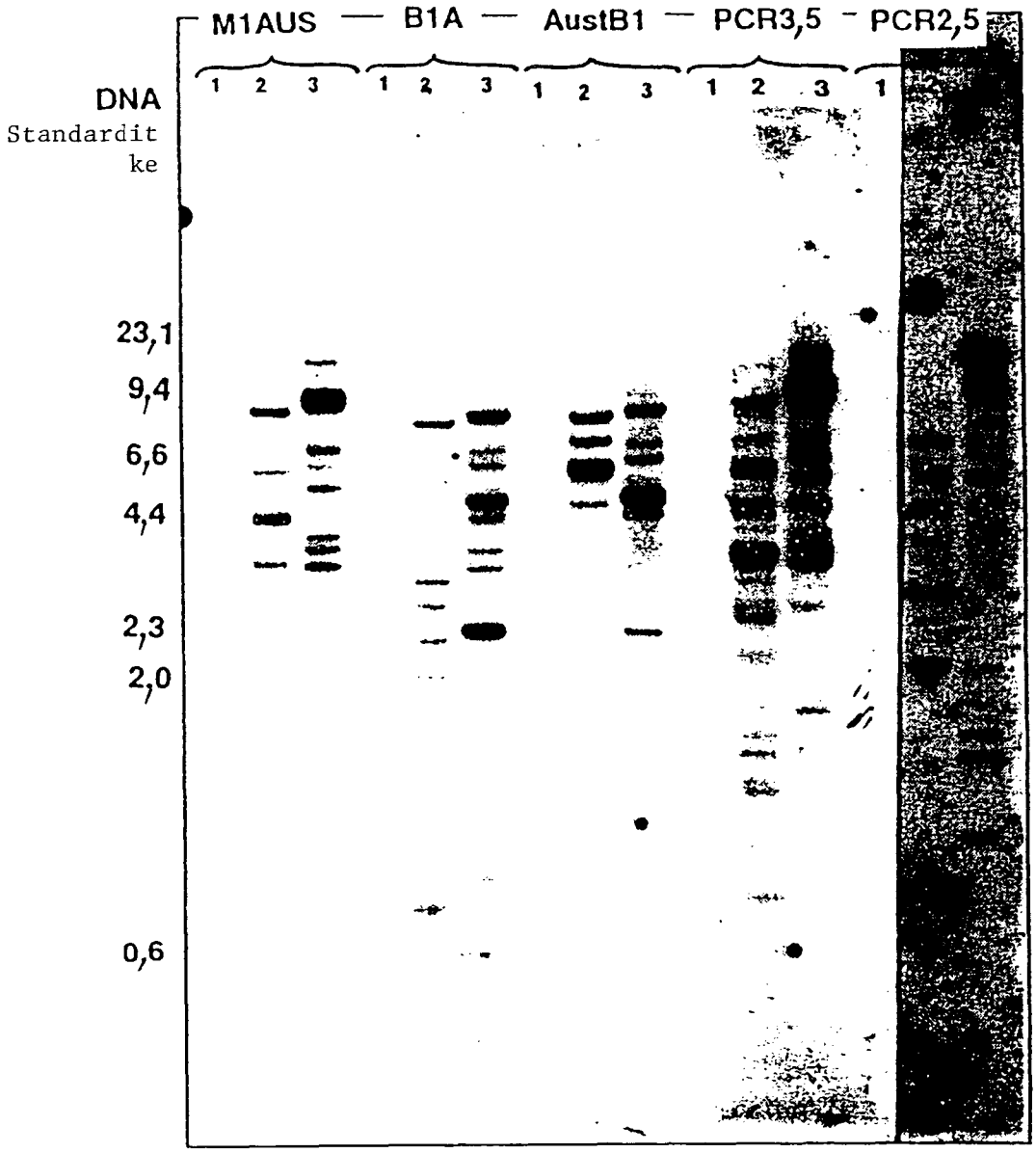

M1AUS
B1A
AustB1
PCR3,5
PCR2,5
1 2 3 1 2 3 1 2 3 1 2 3 1
DNA
Standardit
ke
23,1
9,4
6,6
4,4
2,3
2,0
0,6

FIG. 10b

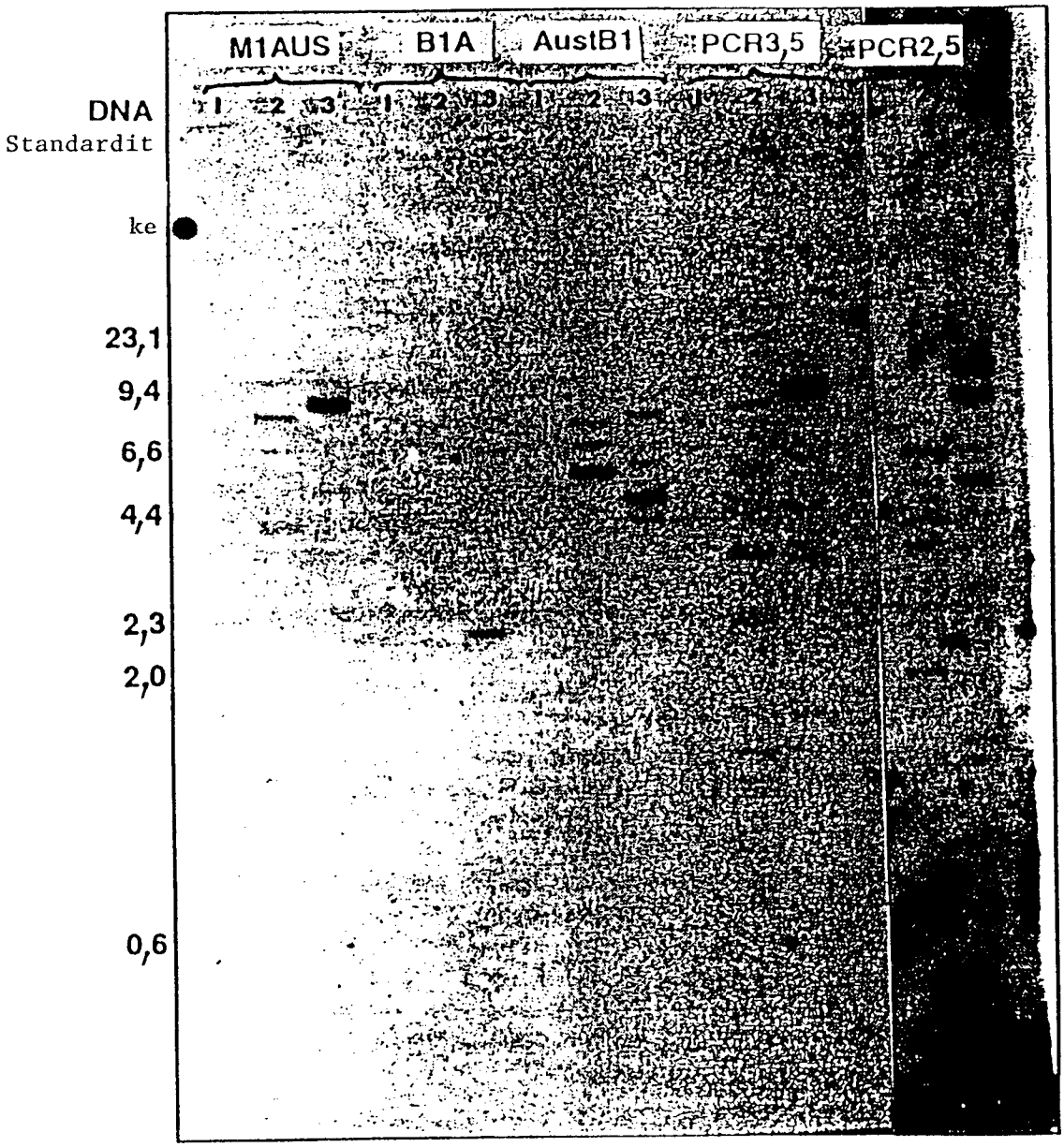

M1AUS
B1A
AustB1
PCR3,5
PCR2,5
1 2 3 1 2 3 1 2 3 1 2 3
DNA
Standardit
ke
23,1
9,4
6,6
4,4
2,3
2,0
0,6

FIG. 11

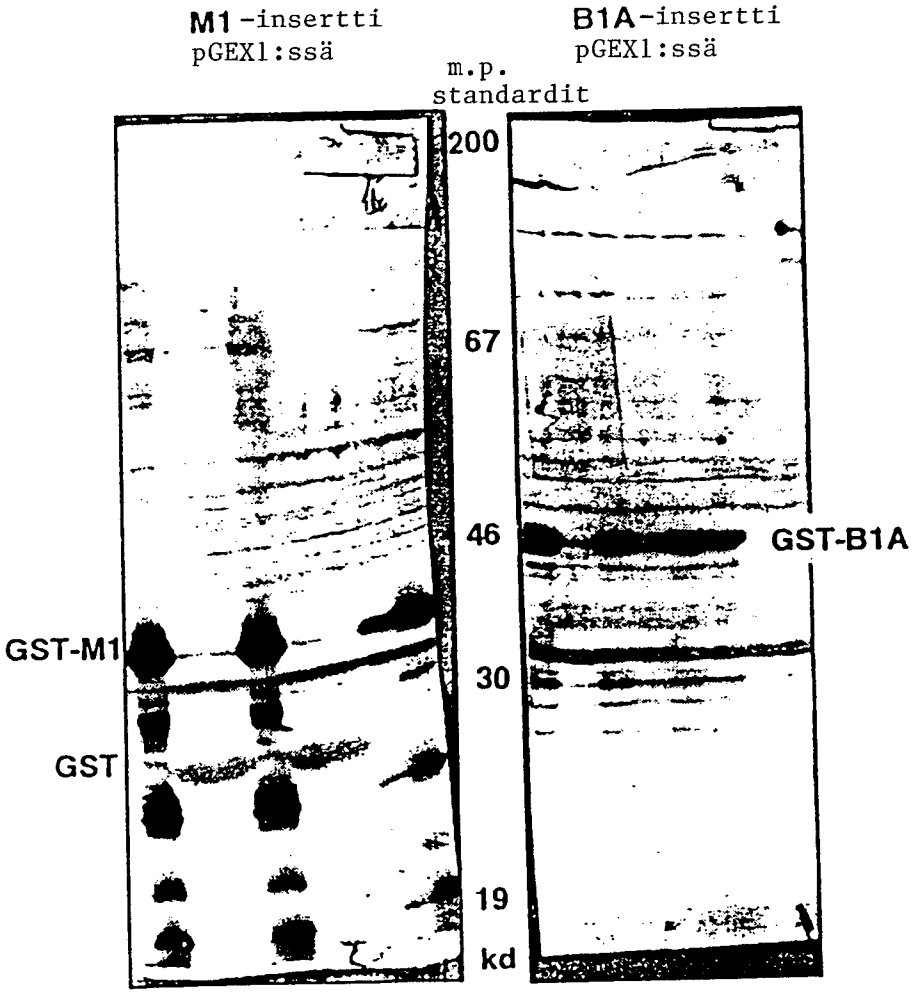

M1-insertti
pGEX1:ssä
B1A-insertti
pGEX1:ssä
m.p.
standardit
200
67
46
30
19
kd
GST-B1A
GST-M1
GST

FIG. 12

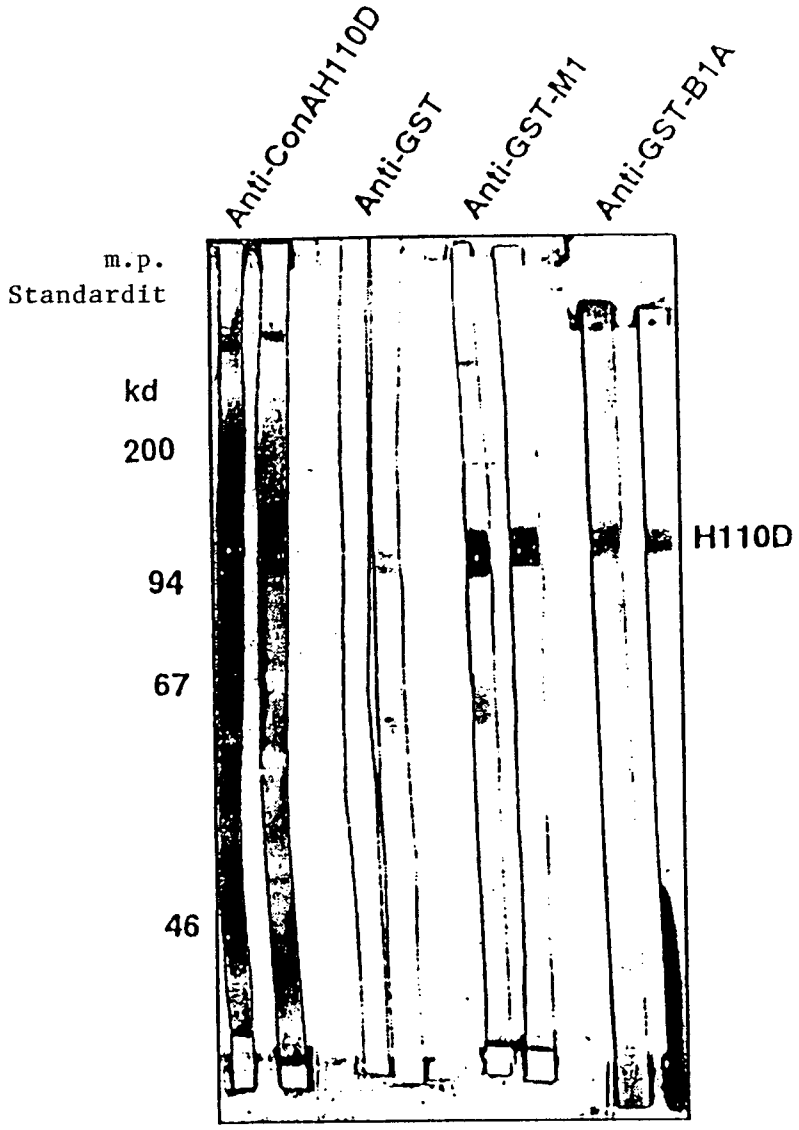

Anti-ConAH110D
Anti-GST
Anti-GST-M1
Anti-GST-B1A
m.p.
Standardit
kd
200
94
67
46
H110D

FIG. 13a

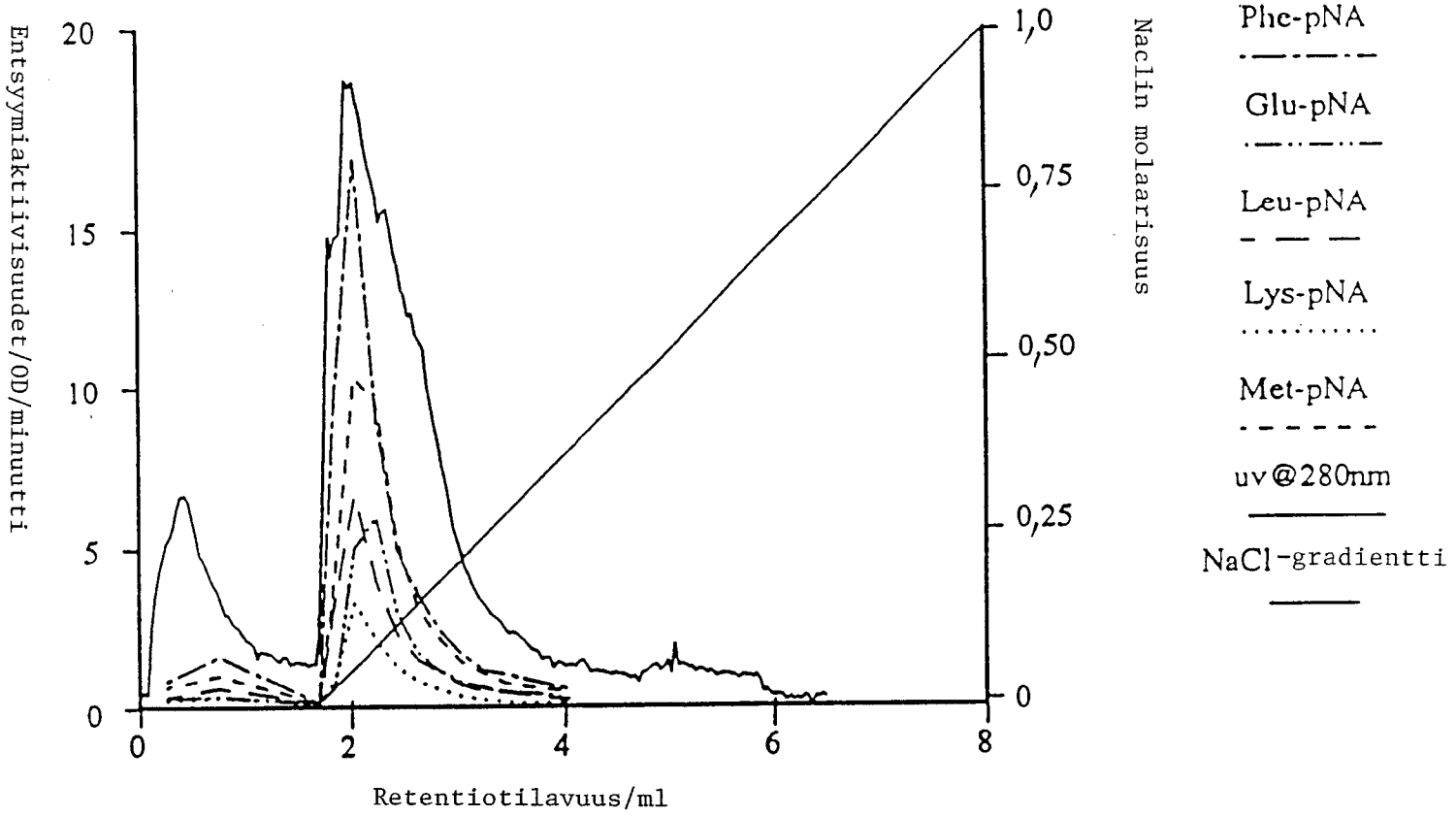
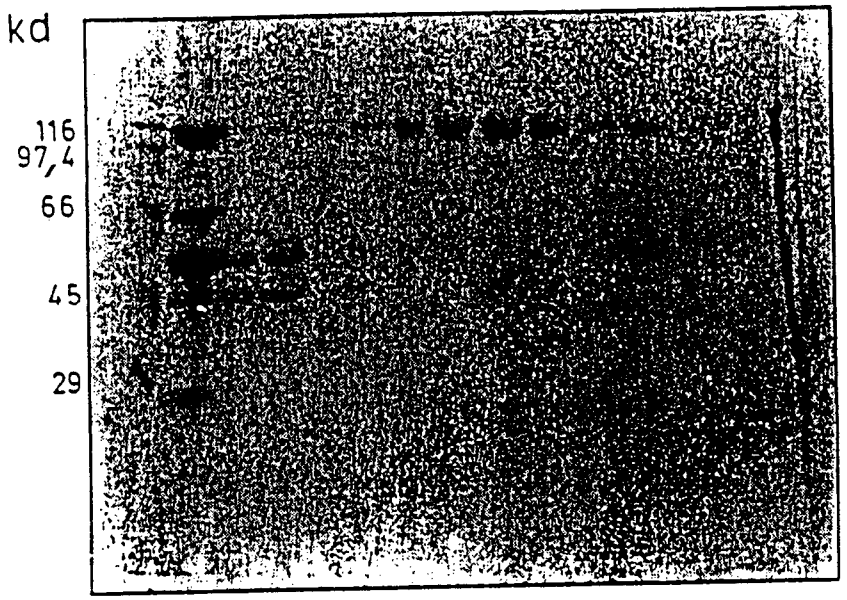

Entsyymiaktiivisuudet/OD/minuutti
20
15
10
5
0
0
2
4
6
8
Retentiotilavuus/ml
1,0
0,75
0,50
0,25
0
Naclin molaarisuus
Phe-pNA
Glu-pNA
Leu-pNA
Lys-pNA
Met-pNA
uv@280nm
NaCl-gradientti

FIG. 13b



A= molekyyli-
paino-
merkit

B= ConA H110D-
lähtömateriaali

FIG. 14a

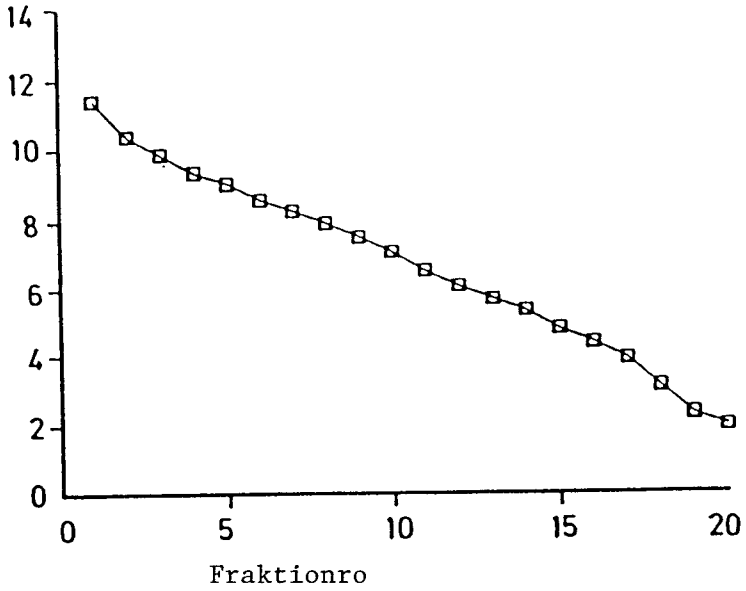
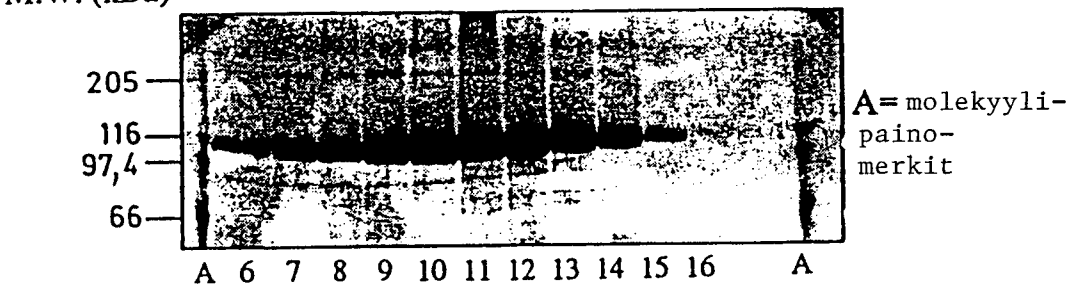


FIG. 14b



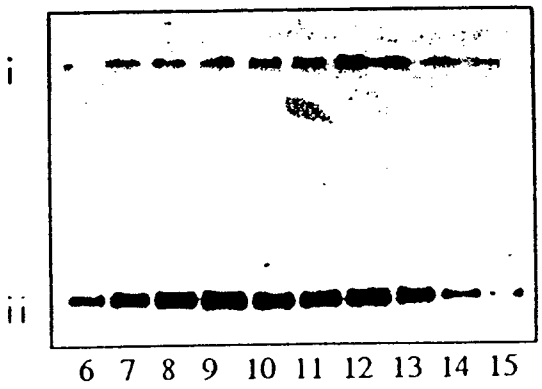


FIG. 14c

FIG. 15a

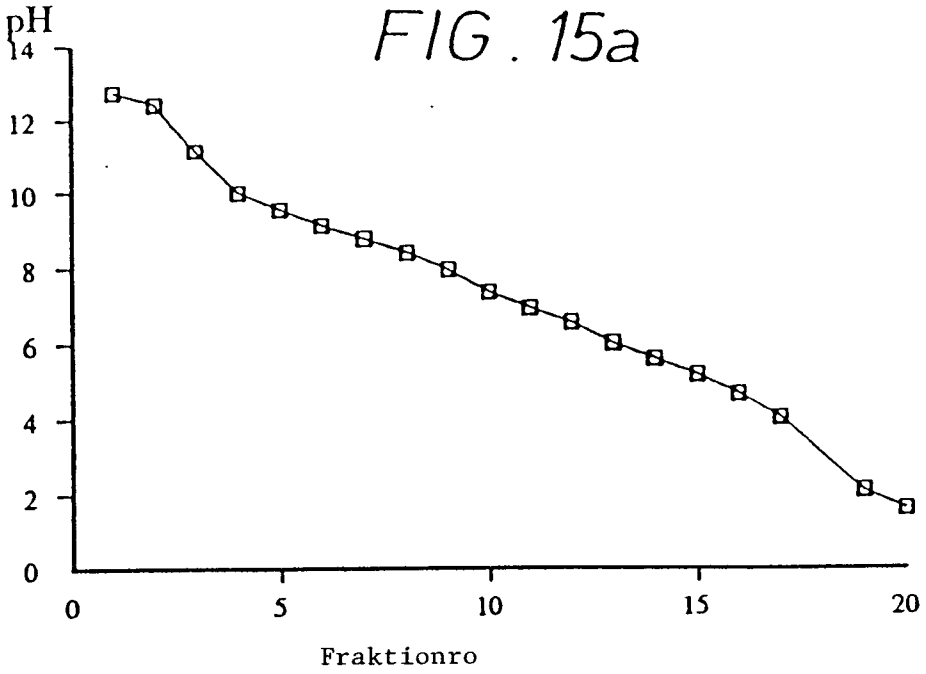
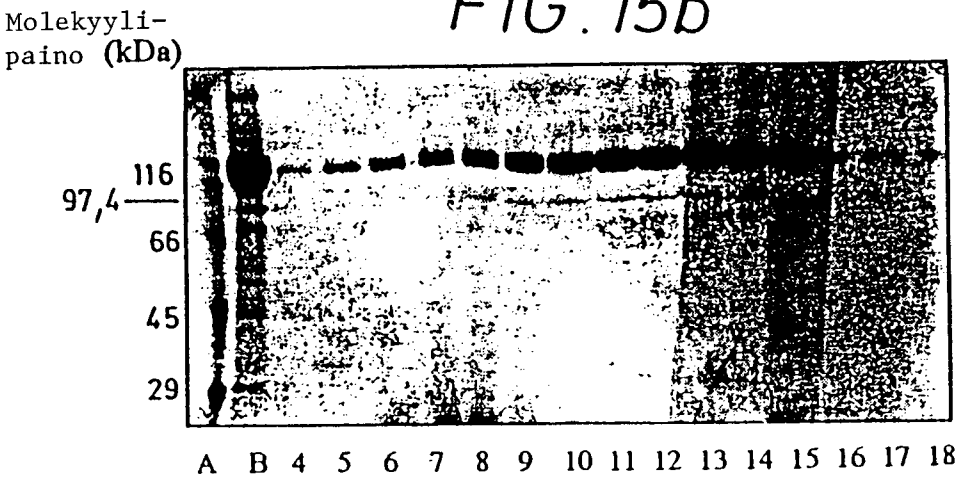


FIG. 15b



A = Molekyyli-painomerkit

B = ConA H110D-lähtömateriaali

FIG.16a

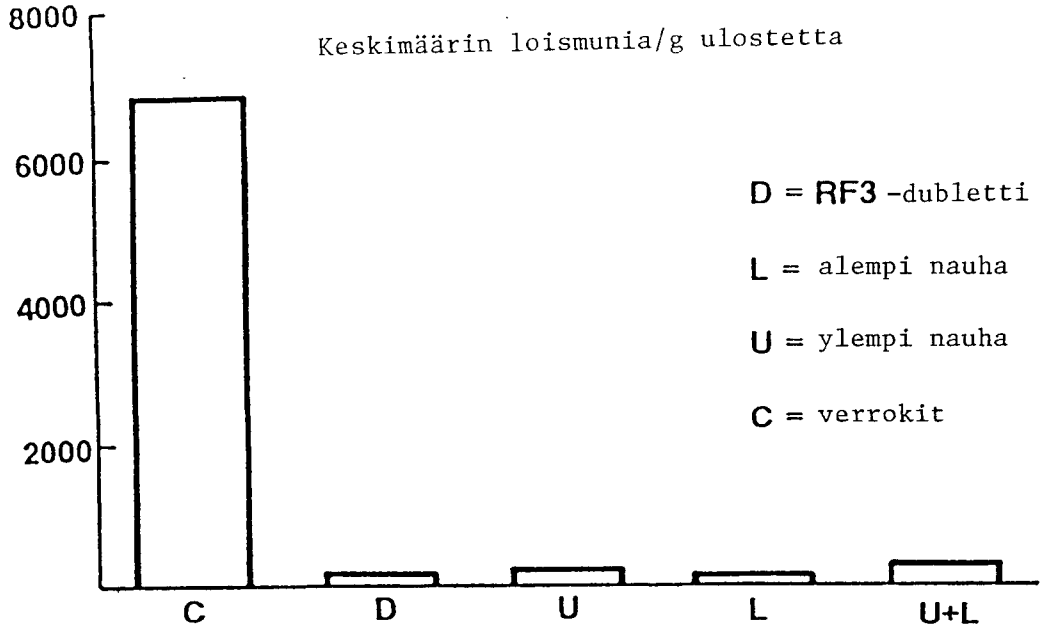
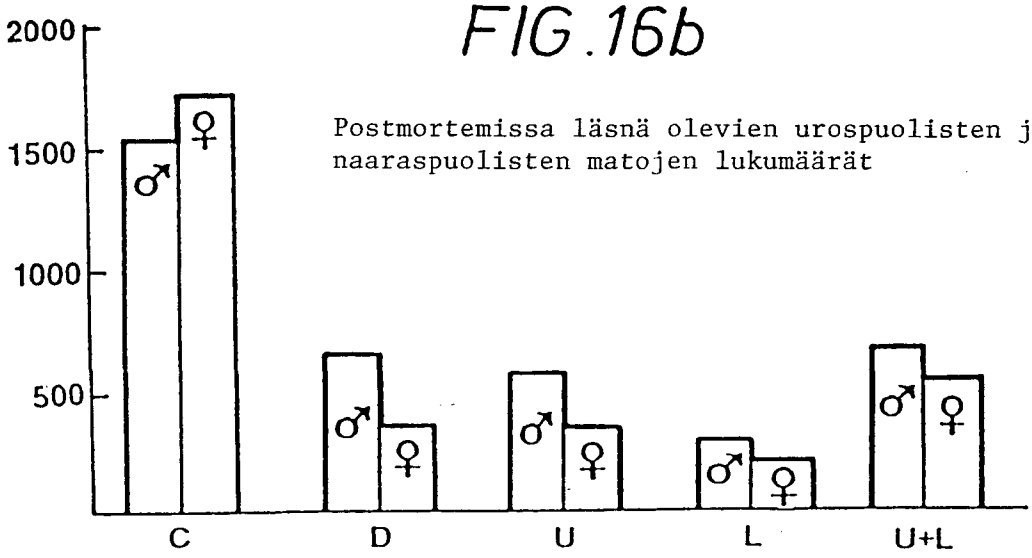


FIG.16b



Keskimäärin loismunia/g ulostetta

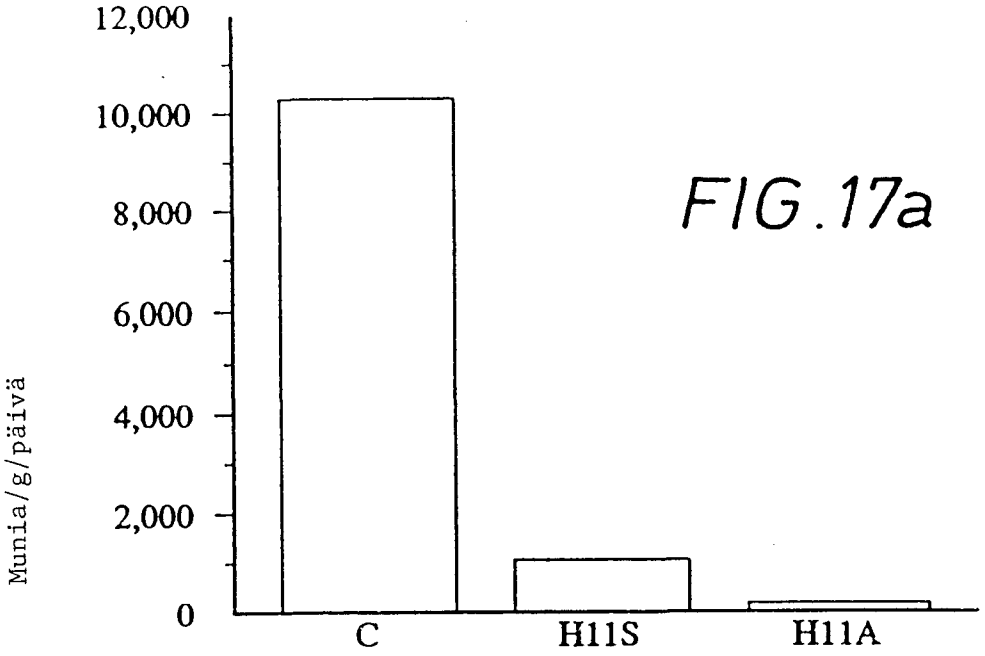


FIG.17a

Postmortemissa läsnä olevien urospuolisten ja naaras-puolisten matojen keskimääräinen lukumäärä

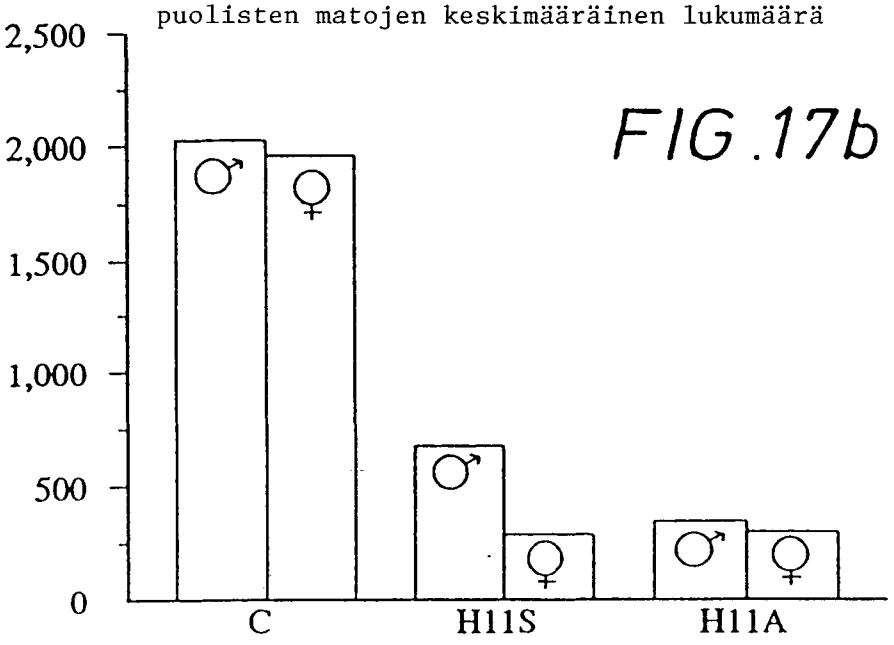


FIG.17b

FIG. 18a)i

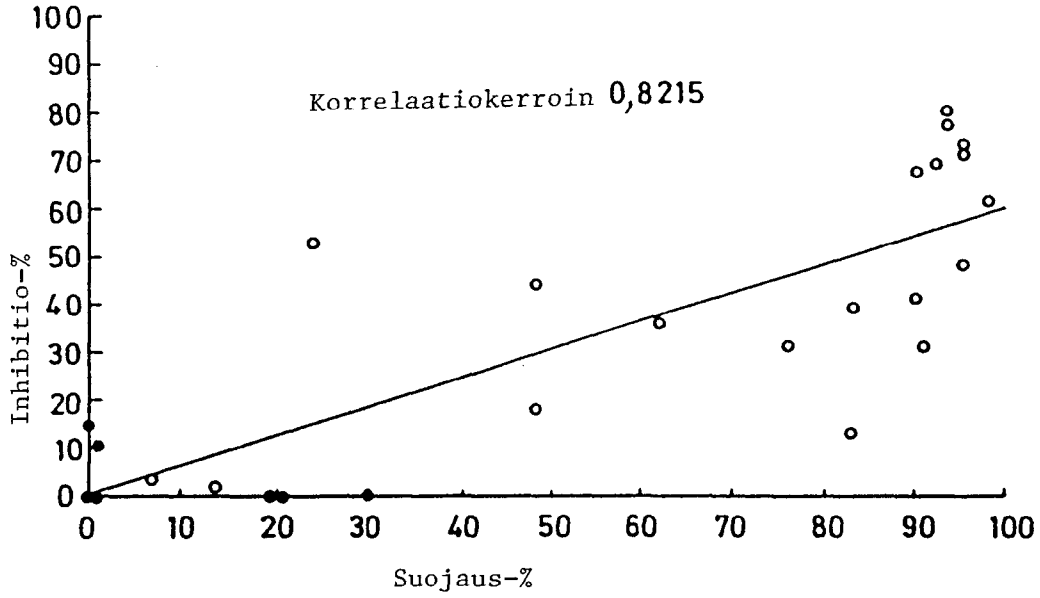


FIG. 18a)ii

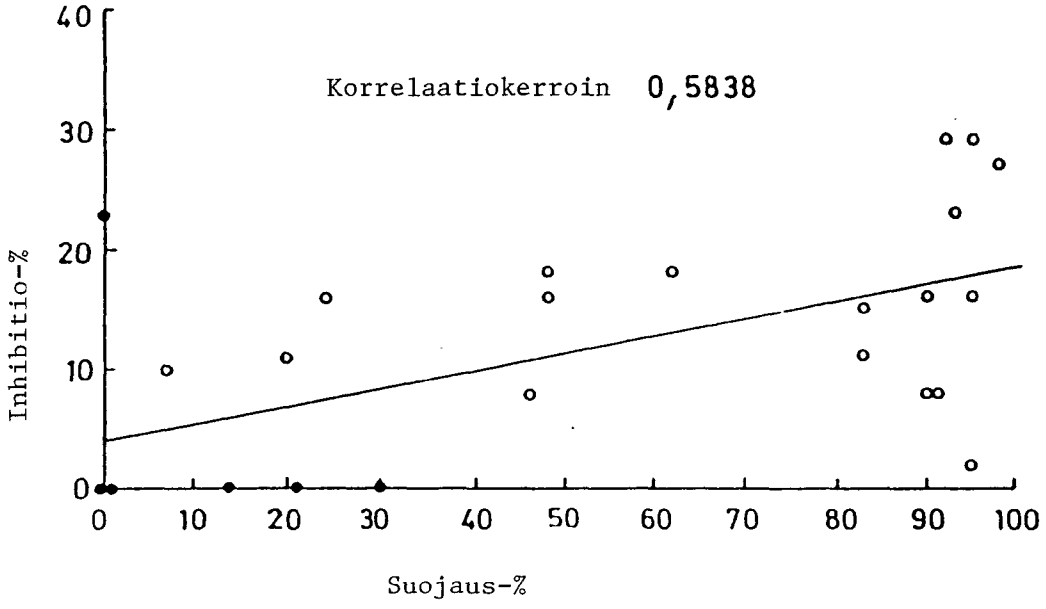


FIG. 18b)i

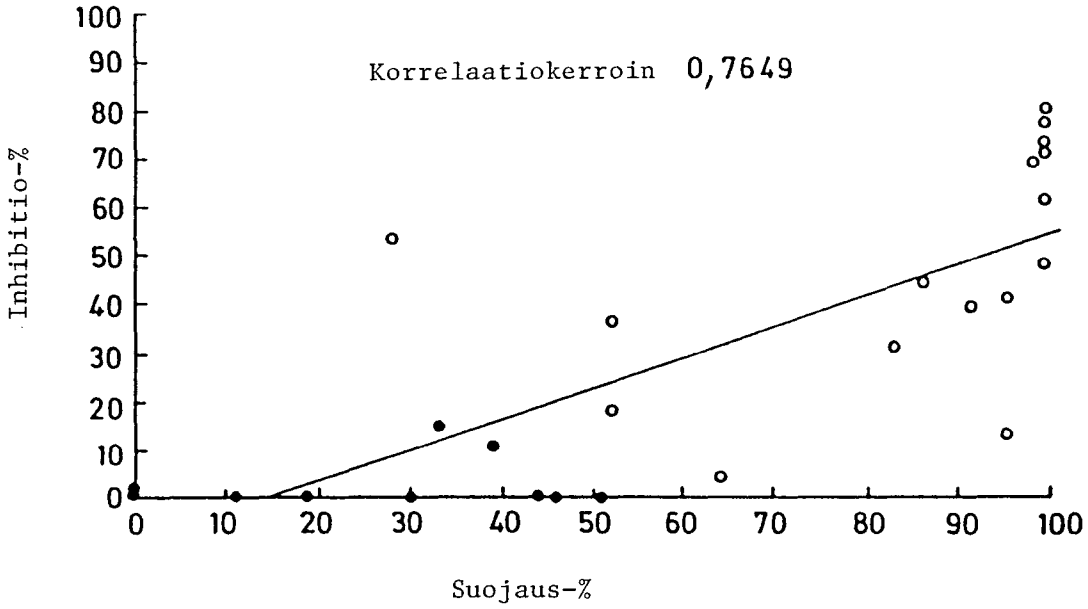


FIG. 18b)ii

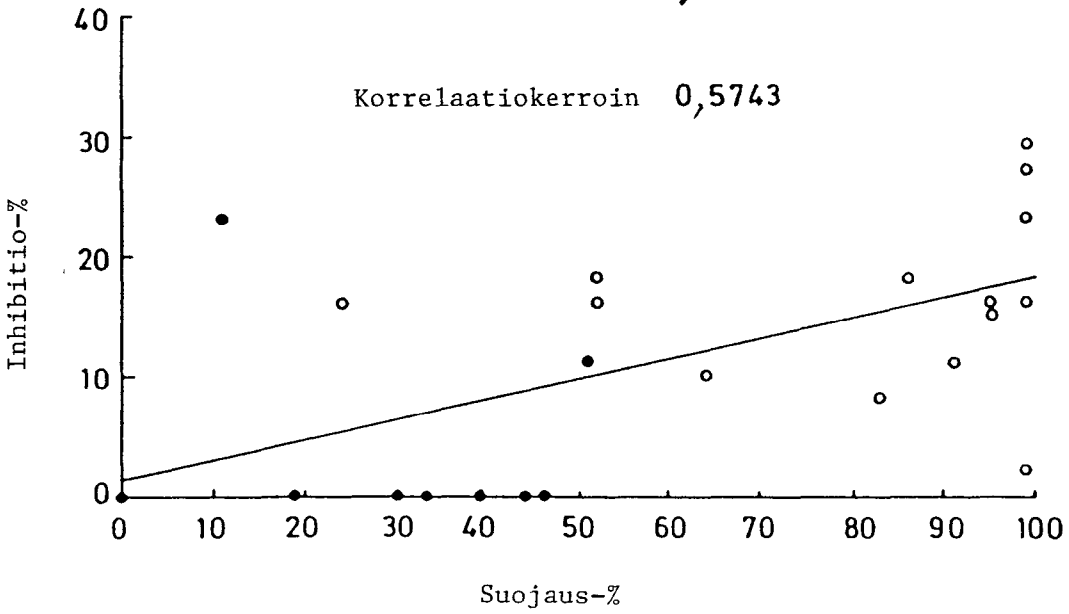


FIG. 19a

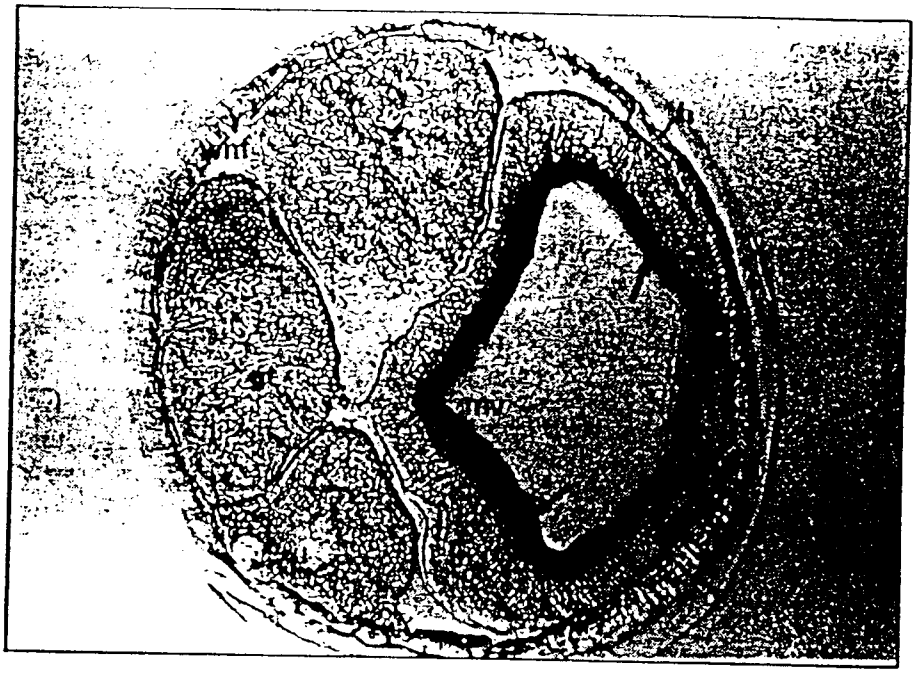


FIG. 19b

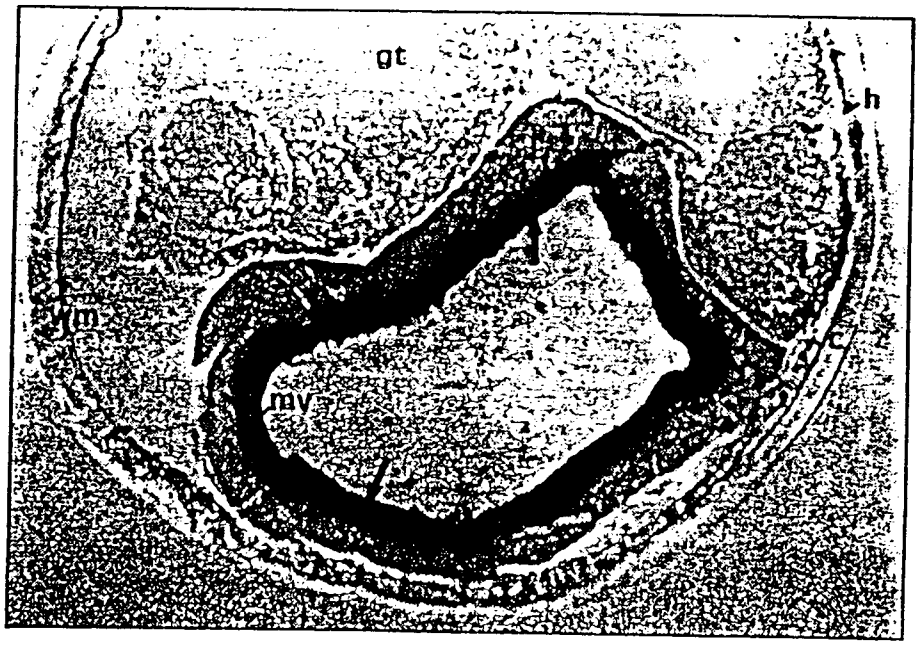
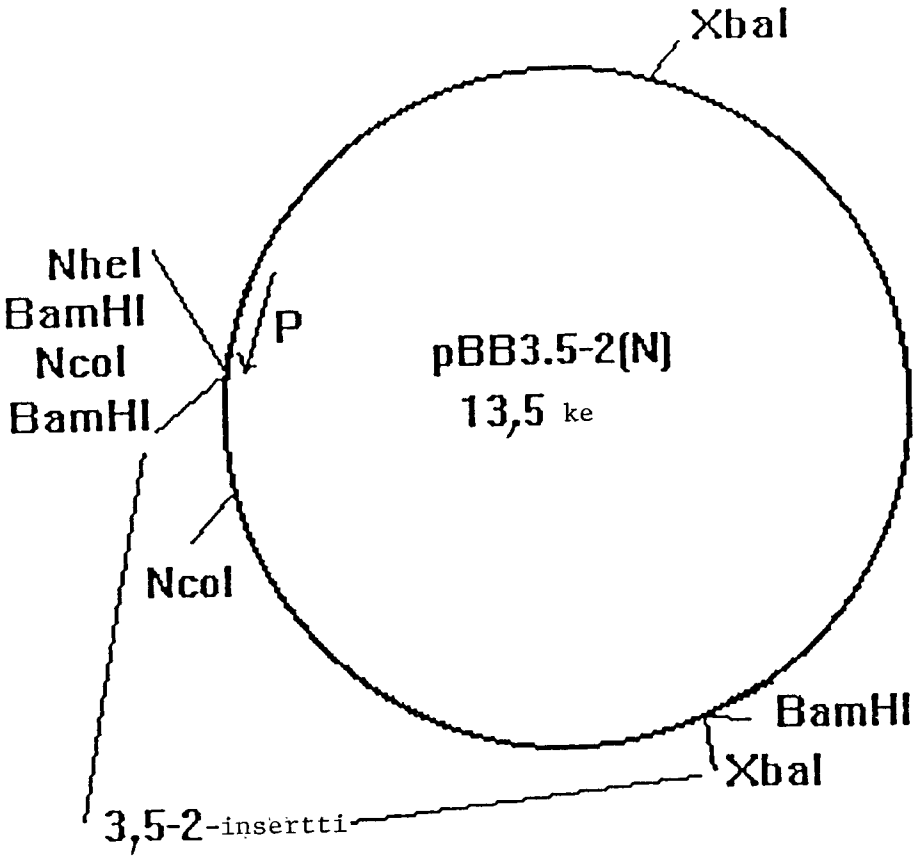
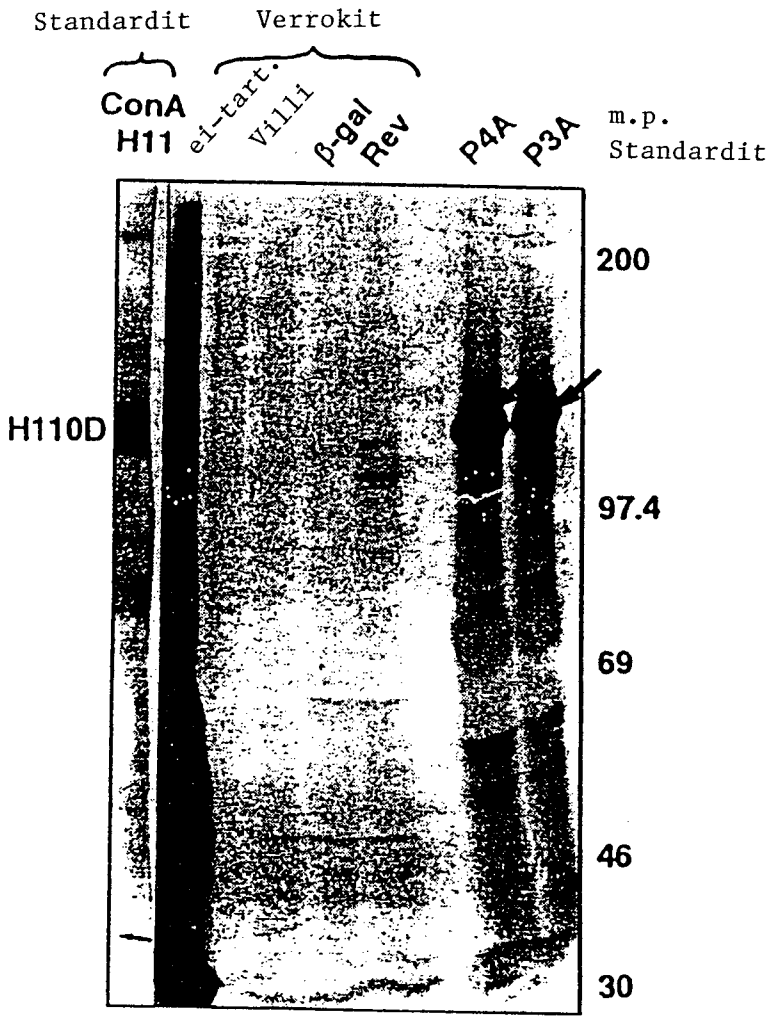


Fig. 20



P = polyhedriinipromoottori

FIG. 21


Standardit
Verrokit
ConA
H11
ei-tart.
Villi
β-gal
Rev
P4A
P3A
m.p.
Standardit
H110D
200
97.4
69
46
30