



FI 000100723B



SUOMI-FINLAND

(FI)

Patentti- ja rekisterihallitus
Patent- och registerstyrelsen

(12) PATENTTIJULKAISU
PATENTSKRIFT

(10) FI 100723 B

(45) Patentti myönnetty - Patent beviljats 13.02.98

(51) Kv.lk.6 - Int.kl.6

C 12N 9/50, 15/57, C 12P 21/00
A 61K 39/02, 39/09, 39/095

(21) Patenttihakemus - Patentansökning 914356

(22) Hakemispäivä - Ansökningsdag 17.09.91

(24) Alkupaivä - Löpdag 16.03.90

(41) Tullut julkiseksi - Blivit offentlig 17.09.91

(86) Kv. hakemus - Int. ansökan PCT/DK90/00073

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

17.03.89 DK 1308/89 P

(73) Haltija - Innehavare

1. Kilian, Mogens, Lindevangsvej 15, 8240 Risskov, Danmark, (DK)
2. Poulsen, Knud, Kirketoften 5, 9260 Viby J., Danmark, (DK)

(72) Keksijä - Uppfinnare

1. Kilian, Mogens, Lindevangsvej 15, 8240 Risskov, Danmark, (DK)
2. Poulsen, Knud, Kirketoften 5, 9260 Viby J., Danmark, (DK)

(74) Asiamies - Ombud: Kolster Oy Ab, Iso Roobertinkatu 23, 00120 Helsinki

(54) Keksinnön nimitys - Uppfinningens benämning

Menetelmä immunoglobuliini-A1-proteaaasin (IgA1-proteaaasin) tai sen fragmenttien valmistamiseksi geenitekniikalla ja IgA1-proteaaasia koodaava geeni
Förfarande för framställning av immunoglobulin-A1-proteas (IgA1-proteas) eller fragment därav medelst genteknik och gen, som koder för IgA1-proteas

(56) Viitejulkaisut - Anförda publikationer

Chemical Abstracts, vol. 107, 192238f, Chemical Abstracts, vol. 104, 124204e,
Chemical Abstracts, vol. 99, 1983, 33850z, Chemical Abstracts, vol. 98, 1983, 66414t

(57) Tiivistelmä - Sammandrag

IgA₁-proteaaasin erityis on Haemophilus influenzaelle ja useille muille tartuntatauteja, meningiitti mukaan lukien, aiheuttaville patogeenisille bakteereille ominainen piirre. On olemassa epäsuoria todisteita, jotka viittaavat siihen, että proteaasit ovat tärkeitä virulenssitekijöitä. IgA₁-proteaaasia, erityisesti sel-laista, joka on peräisin Haemophilus influenzaen serotyypistä b, koodaava iga-geeni on kloonattu Escherichia coliin, jossa yhdistelmä-iga-geeni ilmentyi ja josta ilmennetty proteaasi erittyi. Iga-geenin sisältävän 7,5 ke:n suuruisen DNA-fragmentin sekvenssointi paljasti suuren avoimen lukukehyksen (ORF), jossa kodonien käyttö oli huomattavasti yleisestä suuntauksesta poikkeavaa ja jolla oli kyky koodata 1 541 aminohaposta koostuvaa ja molekyylimassaltaan 169 kd:n suuruista proteiinia. Otaksutut ORF:n viereiset promootori- ja terminaattorielementit on

tunnistettu. Muodostuneen H. influenzaen IgA₁-proteaaasin päätellyn aminohappojakson vertaaminen vastaavaan N. gonorrhoeaesta peräisin olevaan proteaasiin paljastaa useiden korkea-asteista homologiaa osoittavien rakennealueiden olemassaolon. Esitetään ehdotus N. gonorrhoeaan IgA₁-proteaaasin tunnettujen erittymismekanismien suhteen analogisesta H. influenzaen IgA₁-proteaaasin esimuodon translaation jälkeisestä modifikaatiotapahtumasarjasta, jonka johtaa valmiin IgA₁-proteaaasin M_r-arvoksi arvioitua 100 kd:n arvoa lähellä olevan 108 kd:n suuruisen proteaaasin erittymiseen.

Angivningen av IgA₁-proteaser som sekret är karakteristiskt för *Haemophilus influenzae* och ett flertal andra bakteriella patogener, vilka förorsakar infektioösa sjukdomar, inklusive meningitis. Indirekta bevis ger för handen att proteaserna är viktiga virulensfaktorer. Iga-genen, som kodar för IgA₁-proteas, speciellt från *H. influenzae* serotypen b, har klonats in i *Escherichia coli*, i vilken den rekombinerade iga-genen exprimerades och det resulterande proteaset avgavs som sekret. En sekvensbildning av en del av ett 7,5 kb DNA-fragment, vilket innehöll iga-genen, avslöjade en stor, öppen läsram (ORF) med starkt påverkad kodonanvändning och med potential att koda ett protein med 1 541 aminosyror och en molekylmassa av 169 kd. Förmodade promoter- och terminatorelement, vilka flankerar nämnda ORF, har identifierats. Genom jämförelse av den härledda aminosyrasekvensen hos det producerade *H. influenzae*-IgA₁-proteaset med sekvensen hos ett liknande proteas från *Neisseria gonorrhoeae* uppenbaras flera områden med hög homologigrad. Analogt med mekanismer kända från sekretionen av *N. gonorrhoeae*-proteas, föreslås ett schema för eftertranslationell modifiering av en föregångare till *H. influenzae*-IgA₁-proteas, vilket leder till ett som sekret avgivet proteas med en molekylmassa av 108 Kd, vilken ligger nära M_r-värdet 100 kd som uppskattats för det mogna IgA₁-proteaset.

Menetelmä immunoglobuliini-A1-proteaaasin (IgA₁-proteaaasin) tai sen fragmenttien valmistamiseksi geenitekniikalla ja IgA₁-proteaaasia koodaava geeni

5 Tämä keksintö koskee menetelmää immunoglobuliini-A1-proteaaasin (IgA₁-proteaaasin) ja sen fragmenttien valmistamiseksi geeniteknologisesti kloonamalla kyseistä IgA₁-proteaaasia vastaava geeni mikro-organismista, erityisesti b-serotyypin *Haemophilus influenzae*, *E. coli* -isäntä-
10 organismissa, josta muodostunut IgA₁-proteaaasi erittyy solun ulkopuolelle. Tämän keksinnön mukaisesti valmistettuja IgA₁-proteaaaseja voidaan käyttää rokotteissa erityisesti meningiitin ehkäisyyn mutta myöskin immunisointiin allergisia sairauksia, tippuria ja muita IgA-proteaaasia tuottavien bakteerien aiheuttamia tauteja vastaan. Keksintö
15 koskee myös IgA₁-proteaaasia koodaavaa geeniä.

Keksinnön tausta

 Immunoglobuliini-A1-proteaaasit (IgA₁-proteaaasit) ovat solunulkoisia bakteerientsyymejä, jotka pilkkovat spesifisesti ihmisen IgA₁-molekyylien raskaan ketjun sarana-alueen kohdalta. IgA₁-proteaaasin tuotto näyttää korreloivan näiden bakteerien kykyyn infektoida limakalvoja ja joissakin tapauksissa tunkeutua niihin. Erityisesti kolme merkittävintä bakteeriperäisen meningiitin aiheuttajaa (*Haemophilus influenzae*, *Neisseria meningitidis* ja *Streptococcus pneumoniae*) erittävät IgA₁-proteaaaseja, kun taas näille läheistä sukua olevilta ei-patogeenisilta lajeilta tällainen entsyymiaktiivisuus puuttuu. IgA₁-proteaaasin tuotto on myös niiden tiettyjen bakteerilajien ominaisuus, jotka aiheuttavat emätin- ja virtsatietulehduksia, kuten *Neisseria gonorrhoeae* ja *Ureaplasma urealyticum*. Lisäksi monet suubakteerit, jotka liittyvät hammasplakin muodostukseen (Streptococcus sanguis, Streptococcus oralis ja Streptococcus mitis biovar 1 (kaksi viimeainittua lajia ovat aiempia Streptococcus mitior -lajin osia) ja periodontaalisten
25
30
35

tautien patogeneesiin (*Bacteroides* ja *Capnocytophaga*-lajit) erittävät IgA₁-proteaaseja (Kilian, M. ja Reinholdt, J.: "Interference with IgA defence mechanisms by extracellular enzymes", s. 173 - 208 teoksessa C.S.F. Easmon ja J. Jeljaszewics (toim.) (1986) *Medical Microbiology*, osa 5, Academic Press, Inc. Lontoo; katso myös Mulks, M.H.: "Microbial IgA proteases" s. 81 -104 teoksessa I.A. Holder (toim.) (1985), *Bacterial enzymes and virulence*, CRC Press, ja Plaut, A.G., *Ann. Rev. Microbiol.* 37 (1983) 603 - 622. IgA₁-proteaasien ajatellaan olevan tärkeitä virulenssitekijöitä, koska niiden ansiosta bakteerit kykenevät ehkäisemään IgA₁:n, joka on kulloinkin kyseessä olevilla limakalvopinnoilla esiintyvä tärkein immunoglobuliinien luokka (Kett, K. et al., *J. Immunol.* 136, (1986) 3 631 - 3 635), normaalin suojausmekanismien toiminnan.

IgA₁-proteaasien humaani-IgA₁:een kohdistuvan spesifisyyden johdosta näiden entsyymien merkitystä patogeenisille ei voida vahvistaa eläinkokeilla. IgA:n on kuitenkin havaittu pilkkoutuneen suuressa mittakaavassa mainittujen bakteerien infektoimista potilaista peräisin olevissa eritteissä ja muissa ruumiinnesteissä. On lisäksi osoitettu, että IgA₁-proteaasien aiheuttama IgA:n, jonka tehtävänä oli ollut muodostaa suojaava este mahdollisia allergeeneja ja mikro-organismeja vastaan, pilkkoutuminen on määrällisesti suurempaa allergisia tauteja sairastaneiden lasten nenänielueritteissä kuin mitä havaittiin terveissä lapsissa (Christensen C.H. ja Kilian, M., *Acta Path. Microbiol. Immunol. Scand.* 92C, (1984) 85 - 87).

On julkaistu useita tutkimusraportteja entsymaattisesti erityyppisistä IgA₁-proteaaseista, joista kukin pilkkoo ihmisen IgA₁-molekyylin spesifisessä sarana-alueen kohdassa, kuten kuviosta 1 ilmenee (Kilian, M. ja Reinholdt, J., katso edellä). *Haemophilus influenzae*ssa tunetaan ainakin kaksi tällaista toisistaan poikkeavaa entsyymiä (tyyppi 1 pilkkoo asemassa 231 - 232 olevan PRO-SER-

peptididisidoksen ja tyyppi 2 pilkkoo asemassa 235 - 236 olevan PRO-SER-peptididisidoksen). Samoin Neisseria meningitidis ja Neisseria gonorrhoeae voivat erittää tyyppi 1:n tai tyyppi 2:n proteaaseja (katso kuvio 1).

5 **IgA₁-proteaasien polymorfia:**

Useista bakteereista peräisin olevia IgA₁-proteaaseja on verrattu geneettisin ja serologisin menetelmin. Hybridisointikokeista, joissa koettimena oli käytetty suurinta osaa gonokokkien IgA₁-proteaasigeenistä, on käynyt selville, että eri gonokokki- ja meningokokkikannoista saatujen IgA₁-proteaasigeenien välillä on huomattavaa homologiaa (> 78 %). Vielä hämmästyttävämpää on se, että gonokokkigeenin ja vastaavan Haemophilus influenzae (Rd) kannan geenin välillä arvioidaan olevan 67 - 75-%:inen homologia (Koomey, J.M. ja Falkow, S., Infect. Immun. 43 (1984) 101 - 107). IgA₁-proteaasigeenien restriktioentsyymikartoituksella havaittiin kuitenkin huomattava polymorfia-aste vieläpä samaan lajiin kuuluvien kantojen kesken (Halter, R. et al., EMBO J. 3, (1984) 1 595 - 1 601; Bricker, J. et al., Infect. Immun. 4 (1985) 370 - 374; Mulks, M.H. ja Knapp, J.S., Infect. Immun. 55 (1987) 987 - 931). Suuresta Haemophilus influenzaen b-serotyypin kantojen kokoelmasta havaittiin neljä erilaista IgA₁-proteaasigeenityyppiä restriktioentsyymianalyyseillä (Poulsen, K., Hjorth, J.P. ja Kilian, M., Infect. Immun. 56, (1988) 987 - 992).

Tämä huomattava polymorfia on vahvistettu vertaamalla IgA₁-proteaaseja kaniineissa muodostettuja neutraloivia vasta-aineita käyttäen. IgA₁-proteaasipreparaateilla immunisoituina kaniinit reagoivat tuottamalla vasta-aineita, jotka kykenivät neutraloimaan immunisointiin käytetyn entsyymin. Näillä vasta-aineilla on tunnistettu ainakin 15 antigeeniensa suhteen erilaista IgA₁-proteaasityyppiä H. influenzae -isolaatioiden joukosta (Kilian, M., et al., Molec. Immunol. 20 (1983) 1 051 - 1 058) ja kaksi tyyppiä H. influenzaen b-serotyypin joukosta (Poulsen, K., et al.,

katso yllä). Antigeenien osittaista sukulaisuutta on havaittu kahden *N. meningitidis* -kannan toisen edustajan ja yhden *N. gonorrhoeae* -kannan välillä (Kilian, M., et al., Ann. N.Y. Acad. Sci. 409 (1983) 612 - 624). Neisseriae-bakteerisuvun tyyppin 1 ja tyyppin 2 IgA₁-proteaasit voidaan lisäksi erottaa toisistaan näiden välillä olevien inaktivoitumiserojen perusteella käyttäen seerumeja, jotka on saatu joko proteaasityyppejä 1 tai 2 tuottavien bakteerien aiheuttamista meningokokki-infektioista toipuvista potilaista (Stafford ja Plaus, Abstr. Annu. Met. Am. Soc. Microbiol. (1982) B125, s. 38).

IgA₁-proteaasigeenien kloonauk

On kloonattu sellaisia IgA₁-proteaaseja koodaavia geenejä (iga-geenejä), joita tuottavat *Neisseria gonorrhoeae* tyyppi 1 (Fishman, Y., Bricker, J., Gilbert, J.V., Plaut, A.G. ja Wright, A., "Cloning of the type 1 immunoglobulin A1 protease from *Neisseria gonorrhoeae* and secretion of the enzyme from *Escherichia coli*" (1985), sivut 164 - 168 teoksessa G.K. Schoolnik (toim.), The Pathogenic *Neisseria*, Am. Soc. Microbiology, Washington DC) ja tyyppi 2 (Koomey, J.M., Gill, R.E. ja Falkow, S., Proc. Natl. Acad. Sci. USA 79 (1982) 7 881 - 7 885; Halter, R. Pohlner, J. ja Meyer, T.F., (EMBO J. 3 (1984) 1 595 - 1 601), *Haemophilus influenzae* tyyppi 1, joka on peräisin d-serotyypistä (Brickner, J. et al., Proc. Natl. Acad. Sci. USA 80 (1983) 2 681 - 2 685; ja Koomey, J.M. ja Falkow, S. (ks. edellä) ja tyyppi 2, joka on peräisin C-serotyypistä (Grundy, J.F. et al., J. Bacteriol. 169 (1987) 4 442 - 4 450), *Neisseria meningitidis* (Koomey ja Falkow, ks. edellä) ja *Streptococcus sanguis* (Gilbert J.V. et. al., Infect. Immun. 56 (1988) 1 961 - 1 966).

Kullakin näillä iga-geeneillä transformoitu *E. coli* ilmentää IgA₁-proteaasiaktiivisuutta ja *E. coli*-isäntien, jotka sisältävät *H. influenzae* ja *N. gonorrhoeae* iga-geenejä, on havaittu erittävän proteaasia. Geeni, joka on

esimerkki *N. gonorrhoeae*sta peräisin olevasta tyypin 2 IgA₁-proteaasigeenistä, on nykyisin ainoa iga-geeni, jonka nukleotidisekvenssi tunnetaan (Pohlner, J., et al., Nature 325 (1987) 458 - 462; katso myös DE OS 36 22 221). Proteiinin päätelty primäärirakenne ja välituotteena olevien IgA₁-proteaasimuotojen analysointi on paljastanut sen, että proteaasi ilmentyy esimuotoa olevana proteiinina, joka sisältää signaalipeptidin, proteaasin rakenteellisen osan ja avustajaproteiinin, joka prosessoituu autoproteolyytisesti erityystapahtuman aikana. Neisseria sp:n ja *H. influenzae*n entsyymien samanlaisuudesta huolimatta on tehty se johtopäätös, että *Haemophilus*in entsyymin kuljetus ei tapahdu mekanismilla, jota käytetään Neisserian IgA₁-proteaasin erittämiseen (Fishman, et al., katso edellä; katso myös DE OS 36 22 221).

Nyt on kuitenkin tehty se yllättävä havainto, että *Haemophilus influenzae*sta peräisin oleva IgA₁-proteaasigeeni voidaan kloonata *E. coli*ssa siten, että entsyymin erittyminen solun ulkopuoliseen tilaan voidaan saada aikaan, mikä tekee mahdolliseksi uuttaa entsyymi kasvatusliuoksesta. Keksinnön mukaiselle menetelmälle on tunnusomaista se, mitä patenttivaatimuksessa 1 esitetään. Keksinnön mukaiselle geenille on tunnusomaista se, mitä patenttivaatimuksessa 3 esitetään.

*H. influenzae*n d-serotyypistä peräisin olevan kloonatun tyypin 1 IgA₁-proteaasigeenin fragmentteja on aiemmin käytetty koettimina Southern-blottauskokeissa tutkittaessa iga-geenin eri *H. influenzae*n b-serotyypin kannoissa esiintyvää restriktiokohtapolymorfiaa (Poulsen, K., Hjort, J.P. ja Kilian, M., Infect. Immun. 56 (1988) 987 -992). Nämä kannat voitiin jakaa iga-geenin restriktiotulosten mukaan neljään ryhmään, jotka korreloivat aiemmin havaittuihin monilokuksisten genotyyppien (elektroforeettiset tyypit) muodostamiin ryhmiin. Kolmeen näistä iga-geenin restriktiotyypeistä, jotka näyttävät edustavan 98 % *H. in-*

fluenzaen b-serotyypipopulaatiosta, sisältyi yksinomaisesti samaa ainutlaatuista IgA₁-proteaasi-"inhibitiotyyppiä" olevia kantoja ja siksi nämä voisivat muodostaa perustan rokotteen kehittämiseksi H. influenzaen aiheuttamaa meningiittiä vastaan.

Inhiboivilla vasta-aineilla suoritetuissa kokeissa on havaittu, että nämä kolme geenityyppiä ohjaavat sellaisten proteaasien muodostusta, joilla on kullekin näille kolmelle proteaasille yhteiseksi havaittu epitooppi.

10 **Keksinnön ensimmäisen kohteen yksityiskohtainen kuvaus**

Tässä keksinnössä valittiin H. influenzaen kanta HK368 edustamaan yleisintä b-serotyypin iga-geenityyppiä (Poulsen, K., et al., katso yllä). Seuraavassa kuvataan iga-geenin kloonaus ja sekvenssointi tästä kannasta. Tämän proteaasin päätellyn aminohappojakson vertaaminen N. gonorrhoeaeen vastaavaan proteaasiin paljastaa mielenkiintoisten homologisten alueiden esiintymisen.

Seuraavassa tätä keksintöä kuvataan yksityiskohtaisemmin käyttäen apuna piirustuksia, joissa

kuvio 1 esittää humaani-IgA₁:n sarana-alueen primäärirakenteen;

kuvio 2 esittää H. influenzaen b-serotyypin kannan HK368 IgA₁-proteaasigeenin rakenteen;

25 kuvio 3 esittää H. influenzaen b-serotyypin kannasta HK368 peräisin olevan kloonatun iga-geenin nukleotidisekvenssin;

30 kuvio 4 esittää päätellyn H. influenzaen iga-geenituotteen ja N. gonorrhoeae-iga-geenituotteen väliset aminohappohomologiat, ja

 kuvio 5 esittää fysikaalisen kartan lambda HF13iga-1:sta, joka ilmentää meningokokin IgA₁-proteaasigeeniä kannasta HF13.

35 Tätä keksintöä kuvataan yksityiskohtaisemmin seuraavassa. Kokeita suoritettaessa on käytetty seuraavia materiaaleja ja menetelmiä:

Bakteerikannat

Villityypin H. influenzaen b-serotyypin kanta HK368, joka on eristetty meningiittipotilaan selkäydinnesteestä, oli kokonaisten solujen DNA-lähteenä molekulaarista kloonausta varten. Se kasvatettiin Brain Heart Infusion -lihalieimessä (Difco, Detroit, Mich.), jota oli täydennetty hemiinillä ja nikotiiniamidiadeniinidinukleotidillä (NAD) (Kilian, M.J., Gen. Microbiol. 93 (1976) 9 - 62).

E. coli -kantaa K802 käytettiin lambdafaagien kasvatukseen. M13mp19-faagia ja tämän yhdistelmäjohdannaisia kasvatettiin E. coli -kannassa JM109, julkaisussa Yanisch-Perron, C., et al., Gene 33 (1985) 103 - 119 kuvatulla tavalla.

Entsyymit ja kemikaalit

Restriktioendonukleaasit hankittiin yhtiöstä Amersham International (Amersham, Englanti) ja Boehringer GmbH (Mannheim, Saksa). T4-ligaasi, proteinaasi K, RNAasi A ja DNAasi I hankittiin Boehringeriyhtiöstä; DNA-polymeraasi I, DNA-polymeraasi-I:n Klenowfragmentti ja radioaktiivisella leimalla varustetut deoksinukleotidit saatiin Amersham-yhtiöstä. Eksonukleaasi III saatiin Pharmacialta (Uppsala, Ruotsi), lysotsyymi Sigmalta (St. Louis, MS), M13-pentadekameerialuke New England Biolabs -yhtiöstä (Beverly, CA), ja Sequenase-DNA-sekvenssointitestipakkaus saatiin yhtiöstä United States Biochemical Corporation (Cleveland, OH).

H. influenzae HK368 DNA:n kloonauksen faagissa λ L47.1.

H. influenzaesta saatu DNA, joka oli eristetty aiemmin kuvatulla tavalla (Poulsen, K., et al., et al., katso edellä), pilkottiin osittain Sau3A:lla ja fraktioitiin agarosigeelielektroforeesilla. Kooltaan 12 - 20 kilonemästä (ke) olevat fragmentit uutettiin geelistä eluoimalla sähkövirralla ja kloonattiin käyttäen λ L47.1:tä BamHI-korvausvektorina.

In vitro -olosuhteissa teoksessa Maniatis et al., "Molecular Cloning, A laboratory manual", Cold Spring

Harbor Laboratory, N.Y., kuvatulla tavalla pakatut faagit maljattiin *E. coli* -kannalla K802 ja positiiviset plakit tunnistettiin in situ -hybridisoinnilla. Positiiviset plakit puhdistettiin maljaamalla uudelleen *E. coli* -kannalle K802 ja DNA eristettiin 20 ml:n faagiviljelmistä (Mikkelsen, B.M., et al., Biochem. Genet. 29 (1985) 511 - 524).

Nukleiinihappohybridisoinnit ja DNA-käsittely

Hybridisoinneissa käytettiin leimana (α -³²P)-dATP:tä, joka liitettiin DNA-koettiin katkostranslaatiolla (Rigby, P.W.J. et al., J. Mol. Biol. 11 (1977) 237 - 251). Nitroselluloosasuodattimille kiinnitettyjen faagien in situ -koetinkäsittely suoritettiin julkaisussa Benton ja Davies, Science 196, (1977) 180 - 182 kuvatulla tavalla ja Southern-blottaus suoritettiin aiemmin kuvatulla tavalla (Poulsen, K. et al., katso edellä). H. influenzaen d-tyypin iga-geenin sisältävän plasmidin pVD116 fragmentteja käytettiin koettimina. Elektroeluoinnilla eristettyjen restriktiofragmenttien jatkokloonaus M13mp19:aan ja DNA-puhdistustoimenpiteet ja käsittely suoritettiin teoksessa Maniatis, et al. (katso yllä) kuvatulla tavalla.

Sekvenssointimenetelmässä käytettävät m13mp19-yhdistelmäfaagien progressiiviset deleetit tuotettiin vaihtelemalla BamHI/SacI-käsittelyllä avatun replikoituvan muodon DNA:n eksonukleaasi III:lla suoritettavaan pilkkomiseen käytettävää aikaa (Yanisch-Perron et al., katso yllä). Muodostuneiden yksinauhaisten päiden poistamiseen käytettiin eksonukleaasi VIII:n tilalla S1-nukleaasia.

DNA:n sekvensointi

Yksittäisten kloonien DNA-jaksot määritettiin dideoksinukleotidiketjuterminaatiomenetelmällä (Sanger F. et al., Proc. Natl. Acad. Sci. USA 74 (1977) 5 463 - 5 467 ja Biggin, M.D., et al., ibid. 80 (1983) 3 963 - 3 965) käyttäen 15-meeristä yleisaluketta ja (α -³⁵S)-dATP:tä. Joidenkin sekvenssien osalta käytettiin Sequenase-DNA-testipakkausta valmistajan suositusten mukaisesti. Jaksoista saa-

tujen koetulosta kokoamiseen käytettiin Larsonin ja Messingin (Nucl. Acids Res. 10, (1982) 39 - 49 ohjelmaa Apple IIe -tietokonetta käyttäen.

Sekvenssien järjestäminen keskenään

5 Aminohapposekvenssejä koskevat tulokset analysoitiin tietokoneella julkaisussa J. Mol. Biol. 195, (1987) 43 - 61 kuvattua ohjelmaa käyttäen.

IgA₁-proteaasiaktiivisuuden määrittäminen

10 IgA₁-proteaasiaktiivisuus faagilysaateissa havaittiin sekoittamalla keskenään 1 tilavuusosa lymfaattista saattua supernatanttia ja yksi tilavuusosa ihmisen myelooma-IgA₁:n liuosta, jonka pitoisuus oli 2 mg/ml. Yön yli suoritettun inkuboinnin jälkeen osoitettiin IgA₁-substraatin pilkkoutuminen immuno-elektroforeesilla julkaisussa Kilian M., et al., Mol. Immunol. 220 (1983) 1 051 - 1 058 kuvatulla tavalla.

Immunisointi ja inhibitiomääritys

20 Kaniini immunisoitiin iga-geenin sisältävän yhdistelmäfaagin E. coli -lysaatin supernatantilla. Immunisoinnin suoritus ja proteaasin inhibitiomääritys tehtiin aiemmin kuvatulla tavalla (Kilian, M. et al., katso yllä). Saatiin seuraavat tulokset:

H. influenzae:n b-serotyypin iga-geenin kloonaukseen E. coli:ssa

25 E. coli -plasmidi pVD116 sisältää H. influenzae-kannasta Rd-b+ peräisin olevan iga-geenin (Koomey, J.M., Proc. Natl. Acad. Sci. USA 79, (1982) 7 881 - 7 885). Tämän plasmidin 2,0 kb:n suuruinen ClaI/PstI-fragmentti, joka sisältää d-serotyypistä peräisin olevan iga-geenin 5'-pään, 30 käytettiin radioaktiivisella leimalla varustettuna koettimena iga-geenin eristämiseen H. influenzae b-serotyypin kannan HK368 DAN -faagikirjastosta käyttäen vektoria λL47.1. Koettimella seulottiin amplifioimattomasta kirjastosta peräisin olevat noin 2×10^{-3} yhdistelmä-λfaagia 35 sisältäviä alkuperäisten maljojen suodattimelle otettuja

kaksoiskappaleita. Havaittiin 16 positiivista kloonia, jotka merkittiin tunnuksilla λ 368iga-1 - λ 368iga-16. Olettaen, että iga on yksi yhtenäinen geeni, tämä frekvenssi sopii hyväksyttävällä tavalla yhteen niiden arvioiden kanssa, joiden mukaan H. influenzaen genomien koko on $1,8 \times 10^{-3}$ ke ja yhdistelmä- λ -faageissa olevan keskimääräisen liitosjakson suuruus 15 ke. 11 näistä positiivisista geeneistä puhdistettiin ja yhdistelmäfaagin DNA eristettiin. Liitetyssä DNA:ssa olevien restriktioendonukleaasien EcoRI ja HindIII tunnistuskohtien paikat määritettiin yksittäisistä λ 368-iga-faageista saadun DNA:n täydellisellä ja osittaisella pilkkomisella aiemmin kuvatulla tavalla (Mikkelsen, B.M. et. al., katso yllä). Tämän lisäksi λ 368iga-8 ja λ 368iga-16:sta saatu DNA kartoitettiin restriktioentsyymien BamHI, ClaI ja PstI suhteen. Näiden liitosjaksojen restriktiokartat menevät limittäin, kuten kuvioista 2 käy ilmi. Tämä viittaa vahvasti siihen, että H. influenzae -kannassa HK368 oleva iga-geeni on yksi yhtenäinen geeni.

Iga-geenin paikantaminen.

Aiemmin kuvattu 5'-iga-spesifinen koetin yhdessä 3'-iga-spesifisen pVD116:n 2,8 ke:n PstI/EcoRI-fragmentin kanssa hybridisoitiin Southern-blottien kaksoiskappaleisiin, jotka oli saatu kustakin analysoidusta 11 λ 368iga-kloonista peräisin olevasta EcoRI, HindIII ja Sau3A -restriktioentsyymeillä käsitellystä DNA:sta. Iga-geenin pääteltiin näin sijoittuvan noin 7,5 ke:n suuruiseen EcoRI/HindIII-fragmenttiin ja olevan orientoituneena kuviossa 2 esitetyllä tavalla.

Restriktioentsyymi Sau3A-kohdat, jotka sijaitsevat liitosjakson ja vektorin yhtymäkohdassa, säilyvät kloonauksessa. Analysoidusta yhdestätoista yhdistelmäklonista Sau3A:lla muodostuneiden DNA-fragmenttien koko ja lukumäärä ja hybridisoituminen kahteen iga-koettiin olivat seitsemän kloonin osalta identtisiä. H. influenzaen HK368-

kannasta saadun Sau3A:lla pilkotun solun kokonais-DNA:n, joka oli hybridisoitu näillä samoilla kahdella koettimella, Southern-bloteissa havaittiin vyöhykkeiden syntyvän samalla tavalla (tuloksia ei ole esitetty). Tämä viittasi siihen, että näissä -faageissa ei ollut tapahtunut hybridisoituneen alueen uudelleenjärjestymistä kloonaustoimenpiteiden aikana. Lisäksi tämä tulos vahvistaa sen, että sama iga-spesifinen sekvenssi on kloonattu näissä faageissa.

Näistä seitsemästä yhdistelmäfaagikloonista saaduissa nestelysaateissa havaittiin olevan IgA₁-proteaasiaktiivisuutta. Tämä vahvistaa iga-geenin kloonautuneen. Kolmessa näistä yhdestätoista kloonista saadussa lysaatissa, toisin sanoen λ368iga-2:ssa, λ368iga-4:ssä ja λ368iga-5:ssä, ei voitu havaita IgA₁:tä pilkkovaa aktiivisuutta. Näistä kolmesta kloonista saadusta DNA:sta puuttuivat jotkin Sau3A-fragmentit, jotka havaittiin HK368-kannan genomien DNA:ssa, joka hybridisoitui 3'-iga-koettimeen ja tämän vuoksi ne sisältävät ainoastaan iga-geenin 5'-osan (katso kuvio 2). Kloonista λ368iga-8, joka koodaa IgA-1-proteaasiaktiivisuutta, puuttuivat jotkin KH368-genomissa olevat Sau3-fragmentit, jotka hybridisoituvat 5'-iga-koettimeen. Tämä viittaa siihen, että koettimena käytetty plasmidi pVD116 sisältää sekvenssejä, jotka ovat peräisin H. influenzaen d-serotyypin genomista, joka sijaitsee iga-geenin ulkopuolella.

λ368iga-16:n 7,5 ke:n suuruinen iga-geenin sisältävä EcoRI/HindIII-fragmentti jatkokloonattiin M13mp19:ään. Tällä M13-yhdistelmäfaagilla transformoitujen E. coli JM109-nesteviljelmien supernatanteissa havaitsimme IgA₁-proteaasiaktiivisuutta tasoissa, jotka olivat verrattavissa H. influenzaen kannan HK368:n supernatanteista havaittuun IgA₁:tä pilkkovaan aktiivisuuteen. Koska iga-geeni kloonattiin M13mp19:ssa olevan lac Z -geenin suhteen päinvastaiseen suuntaan, tämä tulos vahvistaa sen, että H. influenzaen iga-geeni ilmentyy E. colissa. Se viittaa myös siihen,

että IgA₁-proteaasin erityismekanismit toimivat E. colissa, koska M13-faagit eivät hajoita bakteerisoluja.

5 λ368iga-16-lysaatista saatua raakauutetta käytettiin kaniinien immunisointiin. Tuloksena olevien antiseerumien immunoglobuliinifraktiolla oli täydellinen inhiboiva vaikutus H. influenzae -kannan HK368 viljelmäsupernatanteista saatuun IgA₁-proteaasiaktiivisuuteen. Tämä osoittaa sen, että H. influenzae -kanta HK368 ei eritä yhtään sellaista IgA₁-proteaasia, joka ei muistuttaisi λ368iga-16:ssa
10 kloonattua iga-geenin koodaamaa proteaasia.

Iga-geenin nukleotidisekvenssi

λ368iga-16:n 7,5 ke:n suuruinen iga-geenin sisältävä EcoRI/HindIII-fragmentti oli ajoittain pysymätön M13mp19:ään jatkokloonattuna ja E. coli JM109:ssa kasvatettuna. Sekvenssointistrategiassa käytettiin tämän vuoksi
15 sisäistä PstI-kohtaa tämän fragmentin jakamiseksi (katso kuvio 2). λ368iga-16:n 5,9 ke:n HindIII/PstI- ja 1,6 ke:n PstI/EcoRI-fragmenttien tarttuvat päät (kuvio 21) tasoitettiin Klenow-entsyymillä ja jatkokloonattiin molemmissa
20 orientaatioissa M13mp19:n HincII-kohtaan. Näiden jatkokloonien deleetiojohdannaiset muodostettiin pilkkomalla SacI/BamHI-käsittelyllä avatun replikoituvan muodon DNA:ta Exo III -nukleaasilla, jonka jälkeen käsiteltiin S1-nukleaasilla ja ligaation annettiin tapahtua itsestään. Kuviossa
25 3 esitetyn 7,5 ke:n EcoRI/HindIII-fragmentin 5 091 nukleotidin (nt) jakso saatiin sekvensoimalla yhteensä 79 tällaista deleetiokloonia, joista saatiin molempien nauhojen fragmentin tässä keskiosassa olevat limittäiset jaksot. Jaksot, jotka sijaitsivat sisäisen PstI-kohdan molemmilla
30 puolella, toisin sanoen kuviossa 3 olevalla alueella nt 4856 - 4861, varmistettiin suorittamalla jatkokloonaus M13mp19:ään ja sekvenssoimalla 300 nukleotidia 1 640 emäsparin suuruisesta HhaI-fragmentin 3'-päästä välillä nt 3 349 - 4 988 olevalta alueelta.

Jakso paljasti suuren avoimen lukukehyksen (ORF), jossa esiintyi homologiaa aiemmin julkaistun *N. gonorrhoeae* iga-geenisekvenssin suhteen (Pohlner, J., et al., Nature 325 (1987) 458 - 462). *H. influenzae* -kannan HK368:n ehdotuksen mukainen iga-geeni, joka kuviossa 3 esitetyllä tavalla alkaa ensimmäisestä ORF:ssä olevasta ATG:stä, koostuu 4646 nukleotidista, mukaanlukien TAA-lopetuskodonin, ja koodaa proteiinia, jossa päätellään olevan 1 541 aminohappoa. Tämän primäärisen translaatiotuotteen molekyylimassan päätellään olevan 169 kd, jota vastoin valmiin IgA₁-proteaasin arvioitu koko on noin 100 kd. Tämän eron selittämiseksi seuraavassa tarkastelussa tehdään ehdotus proteiinin esimuodon translaation jälkeisiä modifikaatioita koskevasta kaaviosta.

Tämän keksinnön yhteydessä on *H. influenzae* b-serotyypin HK368-kannasta peräisin oleva IgA₁-proteaasia koodaava geeni kloonattu ja sekvensoitu, jotta tätä proteiinia voitaisiin luonnehtia tarkemmin ja ymmärtää paremmin sen osuus bakteeri-infektion aikana. Kuviossa 3 esitetty iga:n nukleotidisekvenssi ja siitä päätelty IgA₁-proteaasin primaarirakenne paljastaa useita mielenkiintoisia piirteitä.

Transkription ja translaation potentiaaliset aloitus- ja lopetussignaalit

Havaittu ilmentyminen *E. coli*ssa tarkoittaa sitä, että *H. influenzae* iga-geenillä täytyy olla transkriptio ja translaatiosignaalit, jotka *E. coli* -solu tunnistaa. Potentiaalinen ribosomeja sitovan Shine-Dalgarno-jakson 5'-TAAAGA-3' (kuviossa 3 oleva jakso nt 248 - 253) voidaan havaita olevan kahdeksan nukleotidia ennen ehdotettua translaation alkukohtaa avoimessa lukukehyksessä olevan ensimmäisen ATG:n kohdalla (kuviossa 3 oleva A nt 262:n kohdalla). Tämä sekvenssi on viidessä identtisessä asemassa homologinen kuuden *E. coli* n 16 S ribosomaalisen RNA:n komplementaarisen 3'-jakson suhteen (Shine, J. ja Dalgarno,

I., Proc. Natl. Sci. USA 71, (1974) 1342 - 1346). Tämän lisäksi se sijaitsee Stormo et al.'in, Nucl. Acids Res. 10, (1982) 2 971 - 2 996, ehdottamien geenien alkukohtien sääntöjen mukaisessa hyväksyttävässä asemassa. Jakso 5'-TAAACT-3' (nt 235 - 240) ja 5'-TTGTG-3' (nt 221 -225) voisivat muodostaa transkriptiosignaali kohdissa -10 ja -35, tässä järjestyksessä mainittuna. Ehdotettu -10-jakso on identtinen E. colin konsensusjakson suhteen neljässä parhaiten konservoituneessa asemassa (Hawley, D.K. ja McClure, W.R., Nucl. Acids Res. 11, (1983) 2 237 - 2 255).

Tästä 19 emäsparia vastasuuntaan sijaitsee -35-jakso, joka on sallitun etäisyyden päässä ja joka on homologinen kolmessa kohdassa kaikkiaan viisi kohtaa sisältävän E. colin konsensus-jakson suhteen (Hawley, D.K. et al., katso yllä). Näiden kolmen promoottorielementtiehdokkaan tunnistaminen antaa vahvan tuen ehdotukselle siitä, että translaatio alkaisi nukleotidin 262 kohdalla.

ORF päättyy kahteen peräkkäiseen TAA-lopetuskodoniin nukleotidien 4 885 - 4 890 kohdalla. Tyypillisen Rho:sta riippumattoman transkription lopetusjakson suhteen samanlainen jakso havaitaan kohdalla nt 4 919 - 4 943 ja se sisältää täydellisen käänteisen toistuman, jolla on kyky muodostaa hiusneularakenne ja jolla on 5 nt:n silmukka ja 10 nt:n varsiosa, mukaanlukien 5 CG-paria (kuvio 3). Tyypillisten E. colin terminaattorien tavoin tätä sekvenssirakennetta seuraa runsaasti T-ryhmiä sisältävä jakso (Rosenberg, M. ja Court, D., Ann. Rev. Genet. 13, (1979) 314 - 353).

Kodonien käyttö

H. influenzaen iga-geenin kodonien käyttö on esitetty taulukossa 1. Tripleteillä on silmiinpistävänä taipumuksena päättyä ensisijaisesti A:han tai T:hen. Tämä voi olla heijastusta yksinkertaisesti korkeasta A + T -pitoisuudesta (A+T -mooliprosentti = 62). Iga-geenin kodonien frekvenssit ovat samanlaisia kuin neljässä muussa H. in-

fluenzae -geenissä, joiden nukleotidit tunnetaan (Chandra-segaran, S. et al., Gene 70, (1988) 387 - 392). Sekvenssoidun 4 626 nt:n suuruisen H. influenzaen iga-geenin A + T -mooliprosentin havaittiin olevan 63 %, joka on lähellä
5 koko genomille julkaistua arviota 62 % A + T:tä (Kilian, M., J. Gen. Microbiol. 93, (1976) 9 - 62). Julkaisussa Bibb et al., Gene 30, (1984) 157 - 166 olevien havaintojen mukaan tämän genomien geeneillä tulisi kodonien ensimmäisen, toisen ja kolmannen aseman osalta olla A + T joka 51 %,
10 61 % ja 75 %. Näissä asemissa havaittiin A + T-arvot 53 %, 61 % ja 75 %.

Perustuen siihen, että kahdessa rinnakkaiskodonin kolmannessa emäksessä käytetään ensisijaisesti T:tä C:n kustannuksella ja että neljän rinnakkaisen ja kahden rinnakkaisen kodonin esiintymistiheyksien välinen suhde on
15 suhteellisen alhainen kuudessa rinnakkaisessa kodonissa, jotka koodaavat Arg:ia, Leu:ta ja Ser:ää, tulisi iga-geenin ilmentyä heikosti niiden sääntöjen mukaan, jotka on esitetty julkaisussa Grantham et al., Nucl. Acids Res., 9,
20 (1981) 43 - 74. Tämä on yhdenmukaista sen suhteen, että H. influenzae tuottaa IgA₁-proteaaasia suhteellisen pieniä määriä.

Tämä H. influenzaen iga-geenissä esiintyvä vaihtoehtoisten kodonien kolmannessa asemassa oleva normaalista poikkeava suuntaus eroaa jyrkästi huomattavassa määrin normaalista kodonien käytöstä poikkeamattomasta käytöstä, joka
25 esiintyy genomien 51-%:isen A + T-pitoisuuden omaavasta N. gonorrhoeasta peräisin olevan iga-geenin tapauksessa.

Näillä kahdella geenillä on todennäköisimmin yhteinen esimuoto, joka viittaa siihen, että näiden kahden
30 organismin kehityksen aikana on ilmeisesti esiintynyt kodonien käyttöön vaikuttavia erilaisia rajoituksia.

H. influenzaen ja N. gonorrhoeaeen proteaasisekvenssien vertaus

Koomey ja Falkow, Infect. Immun. 43, (1984) 101 - 107 ovat aiemmin osoittaneet, että kloonattu 2-tyypin N. gonorrhoeaeen iga-geeni hybridisoituu H. influenzaen 1-tyypin iga-geeniin. Tämä näiden kahden iga-geenin nukleotidijaksojen vertaaminen keskenään paljastaa alueiden esiintymisen, jotka osoittavat suurta homologia-astetta sekä alueita, joilla ei ole yhtään homologiaa. Näiden kahden toisiaan vastaavien IgA₁-proteaasiproteiinien pääteltyjen aminohappojaksojen suurin homologia saatiin asettelemalla ne toisiinsa näiden kuviossa 4 esitetyllä tavalla.

Polner et al., Nature 325, (1987) 458 - 462 havaitsivat että N. gonorrhoeaeen IgA₁-proteaasin esimuoto sisältää kolme toiminnallista rakennealuetta; aminoterminaalisen signalipeptidin, keskusosan IgA₁-proteaasin ja karboksiterminaalisen avustavan alueen. Johtojakso vapautuu proteiinin esimuodon siirtymisen aikana periplasmiseen tilaan, kun taas avustavan rakennealueen otaksutaan muodostavan reiän ulkokalvoon proteaasialueen erittymistä varten ja tämä jää kalvoon liittyneeksi autoproteolyttisen pilkkoutumisen tapahduttua.

Kuviossa 4 esitetystä H. influenzaen iga-sekvenssistä päätellyn primäärisen translaatiotuotteen molekyyli-massa on 169 kd. H. influenzaesta peräisin olevan valmiin 1-tyyppisen IgA₁-proteaasin kooksi on arvioitu noin 100 kd. Grundy et al., (J. Bacteriol. 169 (1987) 4 442 - 4 450) havaitsivat, että H. influenzaen d-serotyypin kannan 1-tyypin IgA₁-proteaasigeenin 3'-pään alue, jonka koko on 2,2 - 3,1 ke, on tarpeen proteaasin erittymiselle, mutta ei sen aktiivisuudelle. He tekevät sen ehdotuksen, että tämä alue pilkkoutuu pois proteaasin kypsymisen aikana. Näihin havaintoihin ja kuviossa 4 esitettyihin aminohappohomologioihin perustuen teemme sen ehdotuksen, että H. influenzaen IgA₁-proteaasi erittyy samanlaisella mekanismilla

kuin mikä on Polner et al.:n (katso edellä) N. gonorrhoeaen IgA₁-proteaasille ehdottama mekanismi.

5 H. influenzaen IgA₁-proteaasin aminotermiinalinen osa sisältää positiivisesti varautuneet glysiinit aminohappoasemissa (aa) 4, 5 ja 7, joita seuraa hydrofobinen alue, johon kuuluu kierteen katkaiseva proliini asemassa 21
10 neljä ryhmää ennen alaniinia (kuvio 4). Nämä piirteet ovat luonteenomaisia signaalijaksoille, jotka pilkotaan pois ja jotka vapautuvat sisältäen C-terminaalisen aminohapponaan alaniinin (Watson, M.E.E., Nucl. Acids Res., 12, (1984) 5 145 - 5 164). Nämä kaksi proteaasisekvenssiä, jotka on aseteltu rinnakkain kuviossa 4 esitetyllä tavalla, ovat identtisiä N. gonorrhoeaen proteaasin johtopeptidin pilkkoutumiskohdan ympäriltä, jonka on määrittänyt Polner et
15 al., katso edellä. Tämän vuoksi ehdotetaan, että H. influenzaen IgA₁-proteaasin esiproteiinin 25 aminotermiinalista aminohappoa muodostavat signaaliptidin.

Jaksot, joissa esiintyy runsaasti proliineja ja jotka muistuttavat suuresti humaani-IgA₁-molekyylin sarana-
20 alueessa olevaa 1-tyyppin IgA₁-proteaasin kohdealuetta (kuvio 1), esiintyy yksinomaan kolmessa H. influenzaen proteaasijakson alueessa (A, B ja D kuviossa 4). Yksi sekvenssi, joka on homologinen IgA₁-proteaasin 2-tyyppin kohdetta kohtaan (kuvio 1), esiintyy samalla alueella proteaasissa
25 (C kuviossa 4). Teemme sen ehdotuksen, että yksi tai useampi näistä neljästä jakson elementistä vastaa toiminnallisesti N. gonorrhoeae -proteaasissa olevia autoproteolyyt-tisiä kohtia a, b ja c, kuten kuviosta 4 ilmenee.

Kun otetaan huomioon N-terminaalinen 25:stä amino-
30 haposta muodostuva signaalijakso, saataisiin asemassa A, B, C tai D tapahtuvasta autoproteolyytisistä pilkkoutumisesta se tulos, että solu erittäisi proteaaseja, joiden päätellyt molekyylimassat ovat 107,8, 108,1, 109,9 tai 110,3 kd, tässä järjestyksessä mainittuna. Kukin näistä arvoista
35 sopii hyväksyttävällä tavalla yhteen arvioidun Mr:n kanssa, joka on 100 kd.

On silmiinpistävää, että näissä kahdessa proteaasissa oleva 32 aminohapon mittainen alue (aminohapot 16 - 47) on identtinen yhtä säilyttävällä tavalla tapahtunutta substituutiota lukuunottamatta. Tämä osoittaa, että kypsän

5 IgA₁-proteaasin N-terminaalinen osa on säilynyt evoluutiossa, mikä viittaa siihen, että se on oleellinen proteaasimolekyylin entsymaattiselle toiminnalle tai sen spesifisyydelle.

Toinen huomattavan hyvin säilynyt sekvenssi havaitaan alueella, joka ulottuu aminohaposta 785 aminohappoon 797. Se sisältää kaksi säilyvää kysteiiniä, joiden Polner et al., katso edellä, ehdottivat kuuluvan osana N. gonorrhoeaen proteaasin aktiiviseen keskukseen. Kolmas kysteiini havaitaan H. influenzaen proteaasin esiproteiini muodon

10

15 avustavalla alueella aminohappoasemassa 1 250.

Jyrkkänä vastakohtana H. influenzaen IgA₁-proteaasin muuhun osaan nähden ei alueella, joka ulottuu aminohaposta 980 aminohappoon 1 240, ole huomattavaa sekvenssihomologiaa N. gonorrhoeaen proteaasin kanssa (katso kuvio 4). Tämän

20 alueen ehdotetaan muodostavan avustavan alueen N-terminaalisen osan. Se on kummassakin proteiinissa hyvin hydrofiilinen, joka viittaa siihen, että sillä on yhteinen toiminto näiden kahden proteaasin erittymisessä. Näiden kahden proteaasijakson tähän toisistaan poikkeavaan alueeseen ei

25 ollut tarpeen muodostaa yhtään kehityksen kuluessa syntyneitä deleetioita tai insertioita kuviossa 4 esitettyjen reunustavien alueiden homologisten kohtien aikaansaamiseksi. Tämä havainto poikkeaa Grundy et al.:n saamista tuloksista, jotka tutkijat havaitsivat tällä alueella olevan deleetio-substituutiosilmukan analysoidessaan kloonatun H.

30 influenzaen tyyppi 1:n ja H. influenzaen tyyppi 2:n IgA₁-proteaasigeenien välille muodostuneita heteroduplekseja.

H. influenzaen b-serotyypistä saadun IgA₁-proteaasigeenin kloonaus ja sekvenssointi antaa käyttöön työvälineen

35 tämän proteaasin valmistamiseksi suurina määrinä tarkempia rakenteellisia ja toiminnallisia analyyseja varten.

Kuten jo mainittiin, kyseinen IgA₁-proteaasi on ensisijaisesti käyttökelpoinen aktiivisena komponenttina rokotteissa, joita käytetään bakteeriperäistä meningiittiä vastaan suoritettavaan immunisointiin. Rokotus kyseisen

5 tyypin infektioita vastaan suoritetaan periaatteessa parhaiten antamalla rokote suuontelon kautta siten, että rokote pääsee yhteyteen suolen seinän lymfoidikudosten kanssa. Esimerkkejä rokotteen antamisesta ovat a) esittäminen transformoidun E. colin tai muun sopivan isännän pinnalla

10 (vide infra), b) liitettyä mikrosfääreihin yksinään tai yhdessä muiden sopivien rokotekomponenttien tai immunologisesti stimuloivan aineen kanssa, ja c) osana fuusioproteiinia koleratoksiinin B-alayksikön kanssa. Rokotteiden antaminen voi myös tapahtua parenteraalista tietä käyttäen.

15 Nämä esimerkit esitetään ainoastaan mahdollisten sovellustusten kuvaamiseksi eikä näiden rajoittamiseksi.

Se seikka, että IgA₁-geenin ilmentymissignaalit ja H. influenzaen IgA₁-proteaasien eritysmekanismit toimivat E. colissa, tarkoittaa sitä, että IgA₁-proteaasigeeniellä on

20 mahdollinen käyttötapana vektorina vieraiden proteiinien tuottamiseksi E. coli-isäntäviljelmissä tai muiden sopivien isäntien viljelmissä.

Seuraavat esimerkit esitetään keksinnön tämän puolen kuvaamiseksi tarkemmin. Nämä esimerkit on tarkoitettu

25 kuvaamaan tätä keksintöä eikä rajoittamaan sitä:

Iga-geenin rakenteellisen osan muutokset eivät vaikuta proteiinin erittymiseen transformoidun E. coli -kloonnin pinnasta. Vieraita polypeptidejä koodaavien prokaryoottisten tai eukaryoottisten geenijaksojen liittäminen

30 IgA₁-proteaasigeeniin johtaa hybridiproteiinin kuljettamiseen E. coli -solun pintaan. Hybridiproteiini voi tulla esille solun pintaan ja pysyä siinä tai se voi vapautua autoproteolyysin vaikutuksesta kasvatusliuokseen vieraan geenin liitosjakson sijainnista ja koosta riippuen. Tällaisia transformoituja E. coli -klooneja tai muita sopivia

35

isäntiä, joissa nämä fuusioproteiinit esiintyvät solun pinnalla tai vapautuvat tästä, voidaan käyttää suun kautta otettavina rokotteina yhden tai useamman rokotepolypeptidin esittämiseksi kulloinkin asiaankuuluvalla lymfoidikudokselle ihmisten tai eläinten suolessa. Näitä klooneja voidaan vaihtoehtoisesti käyttää vieraiden polypeptidien tuottamiseen *E. coli*ssa tai muissa isännissä. Vapautuminen kasvatusliuokseen voi tarvittaessa tapahtua aktiivisen IgA₁-proteaasin lisäyksellä.

10 **Keksinnön toisen kohteen yksityiskohtainen kuvaus**
 Bakteerikannat

Tutkittiin 19 maata edustava 97 potilaasta ja 36 terveestä kantajasta saatu 133 *Neisseria meningitidis* -isolaation suuruinen kokoelma. Nämä 133 isolaatiota edustivat 15 88 monilokuksista entsyymigenotyyppiä (ET:t), jotka määriteltiin 15 entsyymejä koodaavan geenin elektroforeettisesti havaittavien alleelisten variaatioiden perusteella (Caugant. D. et al., J. Bacteriol. 169 (1987) 2 781 - 2 792; Infect. Immun. 56 (1988) 2 060 - 2 068). Lisäksi 20 tutkittiin 8 gonokokkikantaa, joista kumpaakin IgA₁-proteaasityyppiä 1 ja 2 erikseen edusti 4 kantaa. Aiemmin kuvattua menetelmää (Kilian, M. et al., Molec. Immunol. 20 (1983) 1 051 - 1 058) käyttäen immunisoitiin kaniineja neljästä meningokokkikannasta ja yhdestä gonokokkikannasta 25 peräisin olevilla IgA₁-proteaasipreparaateilla, jotka oli valmistettu julkaisussa Higard, T.B. et al. (J. Immunol. Methods 18 (1977) 245 - 249) kuvatulla menetelmällä. Saa-
duista antiseerumeista tutkittiin niiden kyky estää kaikista 133 meningokokkikannasta ja 8 gonokokkikannasta saatuja 30 IgA₁-proteaaseja käyttäen kuvattuja menetelmiä (Kilian, M. et al., katso edellä). Kun kolme yhteensä neljästä meningokokkiproteaasia vastaan valmistetusta antiseerumista ja gonokokkien IgA₁-proteaasia vastaan muodostettu antiseerumi inhiboivat vain joitakin näistä 141 kannasta saatuja IgA₁-
35 proteaaseja, niin jäljelle jäänyt antiseerumi aiheutti häm-

mästyttävästi IgA₁-proteaasien inhibition riippumatta siitä, olivatko ne tyyppi 1:tä vai tyyppi 2:ta tai siitä, oliko ne saatu meningokokeista tai gonokokeista. Inhibi-
5 tiotiitteri, joka määritettiin edustavia IgA₁-proteaasipre-
paraatteja vastaan, jotka oli säädetty vakioaktiivisuuteen
aiemmin kuvatulla tavalla (Kilian, M. et al., katso yllä),
vaihteli arvosta 1 024 arvoon 4 096. Meningokokkikanta
(HF13), jota käytettiin tämän antiseerumin tuottamiseksi,
10 oli Y-seroryhmän alatyypin 2c ja se tuotti 2-tyypin IgA₁-
proteaasia. Tämän suhteen identtisiä IgA₁-proteaaseja
tunnistettiin muista 2c-alatyyppeistä olevista meningokokki-
kannoista.

Neisseria meningitidis -kannasta HF13:sta saatu
IgA₁-proteaasigeeni kloonattiin E. colissa käyttäen edellä
15 kuvattuja menetelmiä H. influenzaen iga-geenin kloonami-
seksi. Kuviossa 5 esitetään faagi lambda L47.1:ssä oleva
tämän geenin ja sen vieressä olevien genomien alueiden
kartta. Transformoitu E. coli -klooni eritti meningokokkien
IgA₁-proteaasia kasvatusliuokseen ja kasvatusliuoksesta
20 eristetty tuloksena oleva IgA₁-proteaasi kykeni
aiheuttamaan kaniineissa neutraloivien vasta-aineiden muo-
dostuksen, joilla oli samanlainen inhibitiospektri kuin
lähtömuotoa oleva meningokokkikannan HF13 IgA₁-proteaasia
vastaan muodostetuilla vasta-aineilla. Nämä tulokset osoit-
25 tavat, että IgA₁-proteaasi, joka on eristetty meningokokki-
kannasta HF13, joka tuottaa identtistä tai samanlaista
IgA₁-proteaasia, tai E. colista tai muista sopivista iga-
geenillä transformoiduista isännistä eristetty proteaasi on
ehdokas rokotteeksi kaikkia meningokokki- ja gonokok-
30 kityyppejä vastaan.

**Tämän keksinnön kolmannen kohteen yksityiskohtainen
kuvaus:**

Mainittu IgA:n laajamittainen pilkkoutuminen al-
lergisia sairauksia sairastavien lasten nenä-nieluerit-
35 teissä viittaa siihen, että IgA₁-proteaasin indusoima IgA:n

pilkkoutuminen on osallisena allergisten sairauksien kehittymiseen ja jatkumiseen. Kaksi mahdollista selitystä saattaisi selittää allergisissa potilaissa havaitun IgA:n laajemmassa mittakaavassa tapahtuvan pilkkoutumisen: 1) 5 nenä-nielualeelle on asettunut aiempaa suurempi määrä IgA₁-proteaasia tuottavia bakteereja ja 2) allergiaan taipuvaisten potilaiden kyvyttömyys reagoida IgA₁-proteaaseja vastaan entsyymejä neutraloivilla vasta-aineilla.

Tämän selvittämiseksi tutkittiin nenä-nielualueen 10 floora 24 lapsesta yhden vuoden ja noin 30 kuukauden iässä. Edustavat bakteeri-isolaatiot tunnistettiin ja niistä tutkittiin IgA₁-proteaasiaktiivisuus tavanomaisia menetelmiä käyttäen. Kaikki lapset tutkittiin kliinisesti ja immunologisesti allergisten sairauksien mahdollisen diagnoosin 15 vahvistamiseksi.

Yhteensä 11:een 24:stä lapsesta oli kehittynyt 1 vuoden iässä allergisia tauteja. Taulukko 1 esittää kliinisesti ja immunologisesti vahvistetun allergiadiagnoosin ja bakteriologisten löytöjen välisen korrelaation. Luku- 20 arvot osoittavat, että merkittävästi suurempi osuus IgA₁-proteaasia tuottavia bakteereita esiintyy allergisista lapsista peräisin olevissa nenä-nielunäytteissä. On hämmästyttävää, että numeerisesti merkittävin IgA₁-proteaasia tuottava bakteeri oli *Streptococcus mitis* biovar 1. Tämä 25 organismi on aiemmin ollut "*Streptococcus mitior*"-lajin jäsen (Kilian, M., Mikkelsen, L. & Henrichsen, J., Int. J. Syst. Bacteriol. 39 (1989) 471 - 484. Tämä organismi esiintyy IgA₁-proteaasia tuottavassa muodossa ja tätä entsyymiä tuottamattomassa muodossa. Kontrolliryhmässä ha- 30 vaittujen havaintojen mukaan aiemmissa tutkimuksissa on osoitettu, että *S. mitis* biovar 1 esiintyy nenä-nielualueella tavallisesti muodossa, josta puuttuu IgA₁-proteaasiaktiivisuus (Kilian, M. & Holmgren, K., Infect. Immun. 31 (1981) 868 - 873). 18 kuukautta myöhemmin tehdyssä tut- 35 kimuksessa allergisista lapsista peräisin olevat näytteet

sisälsivät Streptococcus pneumoniaeta ja Streptococcus oralista taulukossa 1 lueteltujen bakteerien lisäksi. Nämä koetulokset osoittavat että lapsiin, joissa kehittyy allergisia sairauksia, on nenä-nielualueelle asettunut suuremmassa määrin IgA₁-proteaasia tuottavia bakteereja kuin terveissä lapsissa. Rokote, joka sisältää mainituista asianmukaisista bakteereista peräisin olevia IgA₁-proteaaseja, varmistaa todennäköisesti sen, että immunologinen suoja pysyy ehyenä ja estää näin allergisten tautien kehittymisen ja jatkumisen.

Taulukko 1

IgA₁-proteaaasia tuottavat bakteerit prosentteina kokonaisbakteerifloorasta 24 sellaisen 1-vuotiaan lapsen kurkussa, joilla on ja ei ole allergisia sairauksia

IgA₁-proteaaasia tuottavat bakteerit (% kokonaisfloorasta)

Potilas	Kaikki lajit	H. influenzae	H. parahaemolyticus	S. mitis biovar 1	N. meningitidis
A	30%	0	0	30	0
I	38%	2,4	0	33,4	2,2
N	64%	1,0	0	63	0
A	36%	0	2	34	0
+ allergia	52%	5,1	0	46,9	0
a	58%	0,4	0	0	57,6
D	14%	0	0	14	0
O	38%	0	0	38	0
R	18%	0	0	18	0
E	31%	23	0	8	0
Q	22%	11,7	0	8,3	0
Keskiarvo (alue)	36% (14-64%)	0,4% (0-23%)	0 (0-2%)	30% (0-63%)	0 (0-57,5%)

100723

Taulukko 1 (jatkoa)

IgA₁-proteaaasia tuottavat bakteerit prosentteina kokonaisbakteerifloorasta 24 sellaisen 1-vuotiaan lapsen kurkussa, joilla on ja ei ole allergisia sairauksia

IgA₁-proteaaasia tuottavat bakteerit (% kokonaisfloorasta)

Potilas	Kaikki lajit	H.influenzae	H.paragaenolyticus	S. mitis biovar 1	N. meningitidis
K	8%	0,05	0	8	0
M	6%	0	0	6,4	0
P	4%	0	0	4	0
T	2%	1,2	0	0,8	0
Z	2%	0,1	0	1,9	0
b	14%	0	3,6	10,4	0
- allergia C	2%	0	0	0	2
E	0,4%	0,4	0	0	0
F	14%	2,1	0	11,9	0
H	0,8%	0,8	0	0	0
L	2%	0	0	2	0
U	10%	0,4	0	9,6	0
V	10%	1,1	2,2	6,7	0

Keskiarvo (aiue)	4%	0,1%	0	4%	0
	(0,4-14%)	(0-2,1%)	(0-3,6%)	(0-10,4%)	(0-2%)

100723

Kuvioiden selitykset

Kuvio 1 ihmisen IgA₁:n sarana-alueen primaarirakenne. Nuolet osoittavat yksittäisten IgA₁-proteasien pilkkomia peptidisidoksia.

5 Kuvio 2 H. influenzaen b-serotyypin IgA₁-proteasigeenin rakenne. A. H. influenzaen serotyypin d-iga-geenikoettimen suhteen homologisia jaksoja sisältävien
10 analysoitujen lambda-yhdistelmäfaagien restriktiokartat. Symbolit: (E) EcoRI; (H) HindIII; lopussa oleva palkki tarkoittaa yksittäisten faagien vasenta käsivartta; (+) tai
15 (-) tarkoittavat humaani-IgA₁:tä pilkkovan aktiivisuuden esiintymistä tai tämän puuttumista yksittäisiä klooneja sisältävien E. colin viljelmäsupernatanteissa. B. EcoRI/HindIII:lla suoritetulla yhtäaikaisella käsittelyllä
saatujen fragmenttien hybridisoituminen kahteen d-serotyypin iga-geenikoettimeen. Iga-geenin pääteltyä orientaat-
20 tiota merkitään 5':lla ja 3':lla. C. λ368iga-8:n ja λ369iga-16:n yhdistetty restriktiokartta, joka on A-kohdan lisätutkimus, johon on otettu mukaan restriktioentsyymit
BamHI, (B), ClaI (C) ja PstI (P). D. λ368iga-16:n restriktiofragmentit, jotka on jatkokloonattu M13mp19:aan sek-
venssin analysointia varten.

Kuvio 3 H. influenzaen b-serotyypin kannan HK368 kloonatun iga-geenin nukleotidisekvenssi. Esitetty sek-
25venssi on peräisin λ368iga-16-liitosjakson 7,5 ke:n suuruisen EcoRI-HindIII-fragmentin osasta. Suuren ORF:n päätelty aminohappojakso esitetään nukleotidijakson päällä. Kaksi peräkkäistä lopetuskokodia on merkitty asteriskeilla. Otaksutut -35 ja -10 -promoottorielementit sekä mahdollinen
30 Shine-Dalgarno (SD) -jakso ovat varustettuja yläpuolisella viivalla. Mahdollinen Rho:sta riippumaton kaksinkertaisena esiintyvä terminaattori on merkitty eri suuntaan osoittavilla nuolilla.

Kuvio 4 pääteltyjen H. influenzaen iga-geenituotteen (H11GAP) ja N. gonorrhoeaen iga-geenituotteen (NG1GAP)

väliset aminohappohomologiat. N. gonorrhoeaen sekvenssointitulokset saatiin julkaisusta Pohlner et al. (ks. edellä). Asteriskit tarkoittavat identtisiä tai toiminnallisesti samanarvoisia ryhmiä. Näihin kahteen jaksoon muodostettiin viivoilla merkityt aukot maksimaalisen homologian aikaansaamiseksi. N. gonorrhoeaen proteaasin aminoterminaalisen signaaliptidin pilkkoutumiskohta on merkitty s-kirjaimella; S tarkoittaa H. influenzaen proteaasin esimuodolle ehdotettua vastaavaa pilkkoutumiskohtaa. Asemat a, b ja c edustavat N. gonorrhoeaen proteaasin autoproteolyyttisiä kohtia; asemat A, B, C ja D tarkoittavat samanlaisia H. influenzaen proteaasille ehdotettuja vastaavia autoproteolyyttisiä kohtia.

Kuvio 5 kannasta HF13 peräisin olevaa meningokokkien IgA₁-proteaasigeeniä ilmentävän lambda HF13iga-1:n fysikaalinen kartta. Symbolit: E, Eco RI, H, Hind III, Xb, Xba I, Xh, Xho I. Viivoitettu palkki, fragmentti, joka hybridisoituu 4,3 ke:n suuruiseen pVD 105:n Hind III -fragmenttiin, joka sisältää suurimman osan N. gonorrhoeaen iga-geenistä, mukaanlukien tämän 3'-pään (Koomey, J.M. ja Falkow, S., 1984, katso edellä). BamHI- ja SalI-entsyymien ei havaittu pilkkovan lambda HF13 iga-1:n DNA:ta.

Patenttivaatimukset:

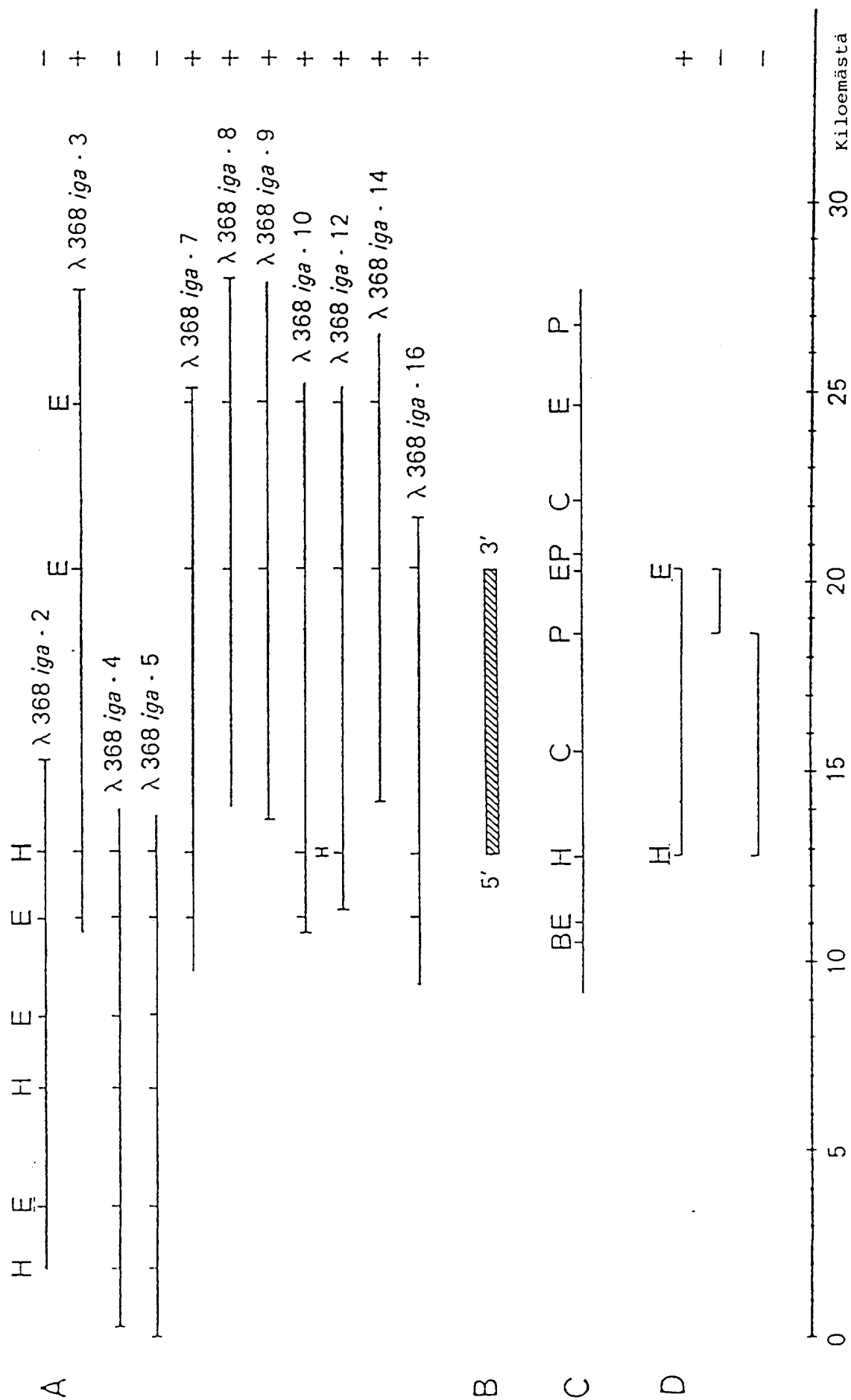
1. Menetelmä immunoglobuliini A1-proteaasin (IgA1-proteaasin) tai sen immunogeenisten fragmenttien valmistamiseksi, t u n n e t t u siitä, että ilmentämisvektori, joka sisältää geenin, joka koodaa IgA1-proteaasia ja joka on peräisin bakteerista Haemophilus influenzae serotyyppi b, kanta HK368 tai Neisseria meningitis-kannasta, joka kuuluu seroryhmään Y, alatyyppiin 2c (nimetty HF13:ksi) ja jonka geenin restriktiokartta ja siihen rajoittuvat geenialueet faagissa lambda L 47 on esitetty kuviossa 5, vietään isäntäorganismiin, joka edullisesti on Escherichia coli, Salmonella tai Saccharomyces, minkä jälkeen isäntäorganismia viljellään IgA1-proteaasin tuottamiseksi ja proteaasi eristetään viljelyalustasta.

2. Patenttivaatimuksen 1 mukainen menetelmä, t u n n e t t u siitä, että isäntäorganismi on Escherichia coli.

3. Eristetty ja puhdistettu geeni, joka koodaa patenttivaatimuksen 1 mukaisesti valmistettavaa IgA1-proteaasia, t u n n e t t u siitä, että sillä on kuviossa 3 esitetty nukleotidisekvenssi.

Patentkrav

1. Förfarande för framställning av immunoglobulin
A1-proteas (IgA1-proteas) eller immunogena fragment därav,
5 k ä n n e t e c k n a t därav, att en expressionsvektor,
som innehåller en gen, som kodar för IgA1-proteas och som
härstammar från bakterien Haemophilus influenzae serotyp b,
stam HK368, eller en Neisseria meningitis-stam ur serogrupp
Y, subtyp 2c (med betckningen HF13), och genens restriktionskarta med vidhängande genområden i fagen lambda L 47
10 visas i figur 5, införs i en värdorganism, som företrädesvis är Escherichia coli, Salmonella eller Saccharomyces, varefter värdorganismen odlas för att producera IgA1-proteas och att proteaset isoleras ur odlingsmediet.
- 15 2. Förfarande enligt patentkrav 1, k ä n n e t e c k n a t därav, att värdorganismen är Escherichia coli.
3. Isolerad och renad gen, som kodar IgA1-proteas som framställs enligt patentkrav 1, k ä n n e t e c k n a t
20 därav, att den har en nukleotidsekvens som anges i figur 3.



100723

Kuvio 3

3/15

CATCATTACGCCTTAATTAAACCAAAGTGCGGTAAAAATAGCCRATTTTTTAAAAAT
 10 20 30 40 50 60

ATCAAAATCATTATCATTACCTCGTAATGTAAATTCAGTTTTGAGCGATTTATGCTAT
 70 80 90 100 110 120

AAARCCCACTCATTATAAAATGARGACAACCTTGCAACTATCAATGCARACAGCCAAG
 130 140 150 160 170 180

ATTGAACTCACTTGTAAGCCGTATATCAAAATTTGTGCTTATCAATCTACTTTTTAACT
 190 200 210 220 230 240

SD M L N K K F K L N F I A L
 TAATTAATAAGACAGCTTCTATGCTAAATAAAAAATTCAAACTCAATTTTATTGCACTT
 250 260 270 280 290 300

T U A Y A L T P Y T E A A L U A D D U D
 ACTGTGCGCTACGCATTAAACCCCTTATACAGAGCCGCGTTAGTGAGAGCGATGTGGAT
 310 320 330 340 350 360

Y Q I F R D F A E N K G K F S U G A T H
 TATCAATATTCCGTGATTTTGCAGAGAAATAAGGGAGTTTTCTGTTGGTGCAACAAT
 370 380 390 400 410 420

U L U K D K N N K D L G T A L P H G I P
 GTGCTGGTAAGGATAAAATAATAAGATTTGGGCACTGCCTTACCTAACGGTATTCCG
 430 440 450 460 470 480

M I D F S U U D U D K R I A T L I H P Q
 ATGATTGATTTTAGCGTGGTAGATGTAGATAACGCATTGCCACATTGATAAATCCACAA
 490 500 510 520 530 540

Y U U G U K H U S N G U S E L H F G N L
 TATGTAGTAGGTGTAAACACGTTAGTAACGGCGTGAGTGAACACTACATTTTGGTAACTTA
 550 560 570 580 590 600

N G N M N N G N A K A H A D U S S E E N
 AACGGCAATATGAATAATGGCAATGCCAAGGCACACCGAGATGTATCTTCAGAGAAAT
 610 620 630 640 650 660

R Y F S U E K H E Y P T K L H G K T U T
AGATATTTTCCGTTGAGAAAATGAGTATCCAACTAATTGAATGGAAAACAGTACT
670 680 690 700 710 720

T E D Q T Q K R R E D Y Y M P R L D K F
ACGGAGATCAAACTCAAAACGCCGTGAGACTACTATATGCCACGCTCTTGATAATTT
730 740 750 760 770 780

U T E U A P I E A S T A S S D A G T Y H
GTTACGGAGTTGCACCAATAGAGGCTTCAACCGCAAGTAGTGATGCTGGCACATATAAT
790 800 810 820 830 840

D Q H K Y P A F U A L G S G S Q F I Y K
GATCAGATAAATATCCTGCTTTTGTAGACTAGGAAGTGGTAGTCAATTTATTTATAAA
850 860 870 880 890 900

K G D H Y S L I L N H H E U G G H N L K
AAAGGAGATAATTACAGCTTAATTTTAATAATCATGAGGTTGGAGGCAATAATCTTAA
910 920 930 940 950 960

L U G D A Y T Y G I A G T P Y K U N H E
TTGGTGGGCGATGCCTATACCTATGGTATTGCAGGTACACCTTATAAAGTAACCACGAG
970 980 990 1000 1010 1020

H N G L I G F G H S K E E H S D P K G I
AATAATGGACTAATTGGTTTTGGCAATTCAAAAGAGGACACAGCGATCCAAAGGAATA
1030 1040 1050 1060 1070 1080

L S Q D P L T H Y A U L G D S G S P L F
TTATCTCAGATCCTCTTACCAATTATGCTGTTTTAGGCGACAGTGGCTCCCCATTATTT
1090 1100 1110 1120 1130 1140

U Y D R E K G K W L F L G S Y D F W A G
GTATATGATAGAGAAAAGGAAATGGCTTTTTCTTGGGTCTTATGATTTTTGGCGGGT
1150 1160 1170 1180 1190 1200

Y H K K S W Q E W H I Y K S Q F T K D U
TATAACRAAAATCTTGGCAAGATGGAATATTTATAAATCTCAATTTACTAAGATGTT
1210 1220 1230 1240 1250 1260

L H K D S A G S L I G S K T D Y S W S S
CTCAATAAGATAGTGCAGGTTCTTTAATTGGTTCCAAGACAGATTATAGTTGGTCTTCT
1270 1280 1290 1300 1310 1320

H G K T S T I T G G E K S L N U D L A D
 RATGGCAGACAAGTACGATTACGGGAGGGGAGAAATCTTTAATGTTGATTTAGCTGAC
 1330 1340 1350 1360 1370 1380

G K D K P H H G K S U T F E G S G T L T
 GGAAAAGATAAACCTAATCACGGGAAAAGTGTTACATTTGAGGGAGTGGARCGCTTACC
 1390 1400 1400 1420 1430 1440

L H H H I D Q G A G G L F F E G D Y E U
 TTAATAATAATATCGATCAGGTGCAGGCGGATTATTCTTTGAGGGCGATTATGAGTT
 1450 1460 1470 1480 1490 1500

K G T S D H T T W K G A G U S U A E G K
 RAAGGTACTTCTGATAATACTACTTGGAAAGGAGCAGGTGTCTCTGTTGCCGAGGAAAA
 1500 1520 1530 1540 1550 1560

T U T W K U H H P Q Y D A L A K I G K G
 ACTGTARCGTGGAAAGTGCATAATCCTCAATATGATCGTTTAGCAAAATTTGGCAAGGG
 1570 1580 1590 1600 1610 1620

T L I U E G T G D H K G S L K U G D G T
 ACATTAATTGTTGAGGAACAGGAGATAATAAGGTTTCGCTAAAGTGGGCGATGGCACC
 1630 1640 1650 1660 1670 1680

U I L K Q Q T H G S G Q H A F A S U G I
 GTTATTTTAAACACACAAACAAATGGTTCCGGACACACGCTTTTGCTTCTGTAGGGATT
 1690 1700 1710 1720 1730 1740

U S G A S T L U L H D D K Q U D P H S I
 GTAGTGGTTCGCTCACTCTTGCTTAATGATGATAACAGTAGATCCAAATTCATT
 1750 1760 1770 1780 1790 1800

Y F G F A G G R L D L H G H S L T F D H
 TACTTTGGCTTTAGAGGCGGTGATTAGACTTAACGGTATTCACTAACCTTTGATCAC
 1810 1820 1830 1840 1850 1860

I R H I O D G A R L U H H H M T H A S H
 ATCAGAAATATTGATGATGGTGCAGACTAGTTAATCATATATGACTAATGCCTCAAT
 1870 1880 1890 1900 1910 1920

I T I T G E S L I T D P H T I T -P Y H I
 ATACGATTACTGGGGAAGTCTAATTACAGATCCAATACATTACTCCATATAATATA
 1930 1940 1950 1960 1970 1980

D A P D E D H P Y A F A R A I K D G G Q L
 GACGCACCAGATGAAGATAATCCTTATGCCTTTCGACGGATTAAAGATGGAGGACAGCTC
 1990 2000 2010 2020 2030 2040

Y L N L E N Y T Y Y A L R K G A S T R S
TATTTAAATTTGGAAATTACACTTATTATGCGTTAAGAAAAGGTGCGAGCACTCGTTCA
2050 2060 2070 2080 2090 2100

E L P K N S G E S N E N W L Y M G K T S
GAATTACCTAAAAATAGTGGCGAAGCAATGAAAATTGGCTATATATGGGTAAACTTCC
2110 2120 2130 2140 2150 2160

D E A K R N U M N H I N N E R M N G F N
GATGAAGCCAAAAGAATGTARTGAACCATATCAACACGAGCGTATGAATGGCTTTAAT
2170 2180 2190 2200 2210 2220

G Y F G E E E G K N N G N L N U T F K G
GGTTATTTTGGCGAGGAGAGGGTAAAAATAACGGTARTCTAATGTGACTTTTAAAGGC
2230 2240 2250 2260 2270 2280

K S E Q N R F L L T G G T N L N G D L T
AAAAGTGAGCAAATCGCTTTTTATTACAGGCGGTACAAACCTTARTGGCGATTTAACG
2290 2300 2310 2320 2330 2340

U E K G T L F L S G R P T P H A R R D I A
GTTGAAAAGGCGACCTTATTCCTTTCAGGCAGACCAACACCGCACGCAAGAGATATTGCA
2350 2360 2370 2380 2390 2400

G I S S T K K D P H F A E N N E U U U E
GGTATTTCTTCGACAAAAAAGATCCTCACTTTGCTGAAAATAATGAAGTGGTAGTAGAA
2410 2420 2430 2440 2450 2460

D D W I N R N F K A T T M N U T G N A S
GATGACTGGATTACCGCAATTTTAAAGCAACCACAATGAACGTGACTGGCAATGCCTCA
2470 2480 2490 2500 2510 2520

L Y S G R N U A N I T S N I T A S N K A
CTTTATTCAGGTGCGAATGTTGCAACATTACGTCAATATCACAGCTTCTAATAAGCA
2530 2540 2550 2560 2570 2580

Q U H I G Y K T G D T U C U A S D Y T G
CAAGTTCATATCGGCTATAAAACAGGCGATACCGTTTGTGTACGTTCTGACTATACGGGC
2590 2600 2610 2620 2630 2640

Y U T C T T D K L S D K A L N S F N P T
TATGTGACTTGTACTACTGACAGTTATCCGATAAAGCCCTTARTAGCTTTAATCCACCC
2650 2660 2670 2680 2690 2700

N L R G N U N L T E S A N F U L G K A N
ATCTACGCGGCAATGTAAATTTAACCAGAAAGTGCAAACTTTGTCTTAGGCCAAGCCAC
2710 2720 2730 2740 2750 2760

L F G T I Q S R G H S Q U A L T E H S H
 TTATTTCGGCACRATTCAAGCAGAGGAATAGCCAGTACGTTTARCCGAAATAGCCAT
 2770 2780 2790 2800 2810 2820

W H L T G H S D U H Q L D L A H G H I H
 TGGCATTTRAACAGGAATAGTGATGTTTCATCAATTAGATCTAGCAATGGGCATATTCAT
 2830 2840 2850 2860 2870 2880

L H S R D H S H H U T K Y H T L T U H S
 TTAARTTCAGCAGACRATTCAACCAATGTGACAAATATAACACGCTGACTGTGAATAGC
 2890 2900 2910 2920 2930 2940

L S G H G S F Y Y L T D L S H K Q G D K
 TTATCAGGTRACGGTTCTTTCTATTATTTTAACTGATCTTTCCAATAARCAAGGCGACAA
 2950 2960 2970 2980 2990 3000

U U U T K S A T G H F T L Q U A D K T G
 GTTGTGTAACTAARTCCGCCACAGGTAACCTTACATTACAAGTGGCAGATAAAACAGGC
 3010 3020 3030 3040 3050 3060

E P H H H E L T L F D A S K A Q A D H L
 GAGCCAAATCATAATGAACCTCACACTTTTGTATGCTTCAAAAGCTCAAGAGATCATTTG
 3070 3080 3090 3100 3110 3120

H U S L U G H T U D L G A W K Y K L R H
 RATGTGTCATTAGTTGGGAATACCGTTGATTTAGGTGCTTGGAATATAAATTACGTAT
 3130 3140 3150 3160 3170 3180

U H G A Y D L Y H P E U E K A H Q T U D
 GTTARTGGACGTTACGATTTGTATAACCCAGAGGTGGAAAAAGGAATCAAACTGTTCAT
 3190 3200 3210 3220 3230 3240

T T H I T T P H H I Q A D U P S U P S H
 ACGACAAATATCACACACCTAATAATATTCAAGCTGATGTGCCTAGCGTACCAAGTAAC
 3250 3260 3270 3280 3290 3300

H E E I A R U D E A P U P P P A P A T P
 RATGAAGAAATAGCCCGTGTGATGAAGCACCAGTTCCACCACCTGCGCCTGCTACACCA
 3310 3320 3330 3340 3350 3360

S E T T E T U A E H S K Q E S K T U E K
 TCAGAGACAACTGAACACGTGGCTGAAATAGTAAGCAAGAAAGTAAACAGTAGAGAA
 3370 3380 3390 3400 3410 3420

H E Q D A T E T T A Q H A E U A K E A K
 AACGAGCAAGACGCAACCGAGACAACAGCTCAAAATAGAGAGTTGCAAAAGAGCTAA
 3430 3440 3450 3460 3470 3480

100723

8/15

Kuvio 3 (jatkoa)

S H U K A H T Q T H E U A Q S G S E T K
 TCRAATGTAAAGCTAATACTCAAACRAATGAAGTAGCTCAAAGTGGAGTGAAACCAAG
 3490 3500 3510 3520 3530 3540

E T Q T T E T K E T A T U E K E E K A K
 GAAACTCAACGACTGAACAAAAGAACAGCTACGGTAGAAAAGAAGAAAAGGCTAA
 3550 3560 3570 3580 3590 3600

U E T E K T Q E U P K U T S Q U S P K Q
 GTAGAACAGAGAAACTCAGAAGTCCCTAAGTGACTTCTCAAGTGTCTCCGAACAG
 3610 3620 3630 3640 3650 3660

E Q S E T U Q P Q A E P A A E H D P T U
 GAACAGTCTGAACTGTTCAACCGCAGCAGAGCCTGCTCGTGAATGATCCGACTGTT
 3670 3680 3690 3700 3710 3720

H I K E P Q S Q T H T T A D T E Q P A K
 AATATAAAGAGCCTCAATCTCAACRAATACACAGCAGACACTGAACRACTGCGAA
 3730 3740 3750 3760 3770 3780

E T S S H U E Q P U T E S T T U H T G H
 GAGACTAGCTCAATGTTGAACARCCAGTGACAGAAAGCACACAGTAARCACTGGAAC
 3790 3800 3810 3820 3830 3840

S U U E H P E H T T P A T T Q P T U H S
 TCTGTAGTGGAAATCCAGAGAATACACACCTGCTACAACTCAACCTACGGTTAATTCA
 3850 3860 3870 3880 3890 3900

E S S H K P K H R H R R S U R S U P H H
 GAAGCAGTAATAAGCCAAAGATAGCATAGAGAGTGTTCCGCTCAGTTCCGCATAAT
 3910 3920 3930 3940 3950 3960

U E P A T T S S H D R S T U A L C D L T
 GTTGARCCAGCTACACRAAGTAGCAACGATCGTTCTACAGTAGCATTGTGCGATCTCACA
 3970 3980 3990 4000 4010 4020

S T H T H A U L S D A R A K A Q F U A L
 AGTACAAACACAATGCGGTACTTTCTGATGCARGGGCAAAAGCACAAATTTGTTGCATTA
 4030 4040 4050 4060 4070 4080

H U G K A U S Q H I S Q L E H H H - E G Q
 AATGTGGGGAAGCAGTTTCTCAACATATTAGCCAGTTAGAAATGAATAACGAGGGGCA
 4090 4100 4110 4120 4130 4140

Y H U W U S H T S M H K H Y S S S Q Y R
 TATAACGTTTGGGTATCTAATACTTCAATGAACAAAATTATTCCTCAAGTCAATATCGT
 4150 4160 4170 4180 4190 4200

R F S S K S T Q T Q L G W D Q T I S H H
CGTTTTAGTTCTAAAAGTACGCRAACTCAACTGGGTTGGGATCAACAATCTCAACAAT
4210 4220 4230 4240 4250 4260

U Q L G G U F T Y U R H S H H F D K A T
GTTCAAGTTAGGTGGCGTGTCTTACTTATGTTTCGCAATAGTACCAACTTTTGATRAGGCAACA
4270 4280 4290 4300 4310 4320

S K H T L A Q U H F Y S K Y Y A D H H W
AGTAAAAATACTCTAGCACAGTTAATTTCTATTCTAATATTATGCGGATATCATTTGG
4330 4340 4350 4360 4370 4380

Y L G I D L G Y G K F Q S K L Q T H H H
TATTTGGGCATTGATTTAGGCTACGGCAAGTTCCAAAGCAATTACCAACTAATCATAT
4390 4400 4410 4420 4430 4440

A K F A A H T A Q F G L T A G K A F H L
GCGAATTTTGCTCGCCATACTGCACAATTTGGTTTAACCGCAGGCAAGCATTTTAATCTT
4450 4460 4470 4480 4490 4500

G H F G I T P I U G U R Y S Y L S H A D
GGCAATTTTGGTATTACGCCAATAGTAGGCGTGGCTTATAGCTATTTATCAACGCTGAT
4510 4520 4530 4540 4550 4560

F A L D Q A R I K U H P I S U K T A F A
TTTGCATTAGATCAAGCTCGCATTAAAGTAAATCCAATATCTGTCAARACAGCCTTTGCT
4570 4580 4590 4600 4610 4620

Q U D L S Y T Y H L G E F S U T P I L S
CAAGTTGATTTAAGTTATACTTATCACTTAGGCGAGTTTTCCGTTACGCCAATTTTGCT
4630 4640 4650 4660 4670 4680

A A Y D A H Q G S G K I H U H G Y D F A
GCTCGATATGATGCARACCAAGGCAGCGGAAAAATTAATGTAATGGATATGATTTTGCT
4690 4700 4710 4720 4730 4740

Y H U E H Q Q Q Y H A G L K L K Y H H U
TACAACGTGGAAACCAACAGCAATATAACGCAGGGCTTAATTTGAATATCATATGTG
4750 4760 4770 4780 4790 4800

K L S L I G G L T K A K Q A E K Q K T A
AAATTAAGTCTAATAGGCGGATTACCAAAAGCGAAACAGCGGAAAAACAAAAAACTGCA
4810 4820 4830 4840 4850 4860

E L K L S F S F * *
GAATTAAACTRAGTTTTAGTTTTTAATRAGCCTGTTTGAATTACGTTATAARCAACAA
1870 1880 1890 1900 1910 1920

AGCCCTGTGTATTACAGGGCTTTATTTTTGAATGAATTCAGTGATTAGTGCGGTGAAA
1930 1940 1950 1960 1970 1980

AATCAGCGCATTTTTTATTTTTAACGTAAAAACGCTGGAAATTTTTTCTCATATGCTGAG
1990 5000 5010 5020 5030 5040

ATTTTGTCTTCGTGCTGAAGGGTTAAACCGATATTATCTAAACCGTTTAGC
5050 5060 5070 5080 5090

100723

Kuvio 4

11/15

S

50

MLHKKFKLNFIALTU--AYALTPYTEARLUADDDUDYQIFADFAENKKGKFSUGATH

* * * * *

MKAKRFKIHAIISLSIFLAYALTPYSEARLUADDDUDYQIFADFAENKKGKFFUGATO

S

100

ULUKDKHKKDLGTALPHGIPMIDFSUDDUDKRIATLIHPQYUUGUKHUSHGUSEL

* * * * *

LSUKHKAGQHIGNALSH-UPMIDFSUADUNKRIATUUDPQYRUSUKHAKREUHTF

150

HFGHLHGHHHGHAKAHADUSSEENRYFSUEKHEY-PTKLHGKTUTTEDQTQKAR

* * * * *

YYGQYHG-----HNDUADKEHEYRUVEQHNHYEPHKAWGASHLG-----AL

200

EDYYMPALOKFUTEUAPIEASTASSDAGTYHQKYPAFUPLGSGSQFIYKKGDHY

* * * * *

EDYHMAHFHKFUTEUAPIAPTDAGGLOTYKDKHAFSSFURIGAGRQLUYEKGUYH

250

SLILHHKEUGGHHKLUGDAYTYGIAGTPYKUHHEHH-----GLIGFGHSKEEHS

* * * * *

Q----EGNEKGYOLRQLSQAYAYAIAGTPYKDIHQDTMHTTEGLIGFGHHNKQYS

300

--DPKGILSQOPLTHYAULGDSGSPLFUYDREKKGKULFLGSYDFWAGYHKKSWQE

* * * * *

REELKQALSQDALTHYGULGDSGSPLFAFOKQKHQWUFLGTYDYWAGYGKKSWE

350

WHIYKSQFTKOUHKKDSAGSLIGSKTOYSWSSHGKTSTITGGEKSLHVDLADGKO

* * * * *

WHIYKKEFAOKIKQHDHAGTUKG-NGEHHWKTGTTHSHI----GSTAURLANNEG

400

KPHHGKSUTFEGSGTLTHHHIQGAGGLFFEGDYUUKGTSDHTTWKGAGUSUAE

* * * * *

DAHHGQHUTFEDHGTULHQNINQGAGGLFFKGOYTUKGANNHDTWLGAGIDUAD

150

GKTUTWKUHPQYDALAKIGKGTLEIEGTGDHKGSLKUGDGTUILKQQTHGSGQH
 * * * * *
 GKUUWQKHPHGDALAKIGKGTLEIHGTGVNQGLKUGDGTUILHQAADKKU

500

-AFASUGIUSGASTLULHODKQUOPHSIYFGFAGGALDLHGNSLTFOHIAHIDG
 * * * * *
 QAFSQUGIUSGAGTLULHSSHQINPDHLYFGFAGGALDANGHCLTFEHAHUEG

550

ARLUHHMTHASHITITGESLITDPHTITPYHI-DAPDEOHYPYF--RAIKDGGQ
 * * * * *
 ARIUHHNTDASTITLTGKSLITNPHLSUHSIQHOYDEDDYSYYYRPARPIPQG

600

LYLHLEHYTYVALAKGASTRSELPKNSGESHEWLYMGKTSDEAKAHUMHHIHNE
 * * * * *
 KDLYYKHRYRYALKSGGALHAPMPENGUAENHDIIFMGYTQEERAKHAMHHKHHR

650

700

AMHGFHGYFGEEEGK-HHGH LHUTFKGKSEQHAFLLTGGLHLHGDLTUEKGTFL
 * * * * *
 RIGDFGGFFDEENGKGHNGALHLHFNGKSAQHAFLLTGGAHLHGKISUTQGHULL

750

SGRPTPHARDIAGISSTKKOPHFAENHEUUUEODWIHAFKATTHUTGHASLYS
 * * * * *
 SGRPTPHARDFUNKSSARKDAHFSEKHHEUUFEDOWIHRTFKAREIAHQSAFSS

800

GAHUAHITSHITASHKAQUHIGYKTGOTUCURSDYTGUTCTTOKLSOKALNSFH
 * * * * *
 GAHUSDITAHITATONAKUHLGYKNGOEUCURSDYTGUTCHTGHLSOKALNSFO

850

PTHLAGHULTESANFULGKANLFGTIQSAGHSQUALTEHSHWHLTGNSDUHQLO
 * * * * *
 ATAHNGHULHQAALULGKAALWGKIQQQGNSRUSLHGHSHWHLTGDSQUHHLS

100723

1350

HYSSSQYARFSSKSTQTQLGWDQTI SHVUQLGGUFTYURHSHNFOKATSKHTLAQ
* * ***** * *** * ***** ** * **** *
DYASQYARFSSKRTQTQIGIORSLSEHMQIGGULTYSOSQHTFDQAGGKHTFUQ

1400

UHFYSKYYADHHWYLGIDLGYGKFQSKLQTHHNAKFARHTAQFGLTAGKAFNLGH
 *** ** * * * * * * * * * *
 ANLYGKYYLHDWWYVAGDIGAGSLASALQTQQKANFHATSIGTGLTLGHTLKIHO

1450

FGITPIUGURYSYLSHADFALDQARIKUNPI SUKTAF AQUOLSYTYHLGEFSUTP
* * * * *
FEIUPSAGIRYSRLSSADYKLGODSUKUSSMAUKTLTAGLOFAYRFXUGHLTKRP

1500

ILSARYDANQGSCKIHUNGYDFAYHVENQQQYNAGLKLKYHNKLSLIGGLTKAK

 LLSARYFANYGKGGUNUGGKSFAYKADNQQQYSAGUALLYAKUTLHUNGSITKGK

1540

QREKQKTRELKLSFSF HIIGAP
* * * * * * * *
QLEKQKSGOIKIQIRF HGIGAP

Kuvio 5

