(19) SUOMI - FINLAND
(FI)

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

(10) **FI 923275 A7**

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

(21) Patenttihakemus - Patentansökan - Patent application 923275

(51) Kansainvälinen patenttiluokitus - Internationell patentklassifikation - International patent classification
C12N 9/42
C12N 15/56
C11D 3/386

(22) Tekemispäivä - Ingivningsdag - Filing date **18.01.1991**

(23) Saapumispäivä - Ankomstdag - Reception date **17.07.1992**

(41) Tullut julkiseksi - Blivit offentlig - Available to the public **17.07.1992**

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

(86) Kansainvälinen hakemus - Internationell ansökan - International application **18.01.1991 PCT/DK1991/000013**

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

19.01.1990 DK 0164/90

(71) Hakija - Sökande - Applicant

1 • Novo Nordisk A/S, Novo Alle, 2880 Bagsvaerd, TANSKA, (DK)

(72) Keksijä - Uppfinnare - Inventor

1 • Joergensen, Per Linå, Koebenhavn K, TANSKA, (DK)
2 • Schülein, Martin, Koebenhavn Oe, TANSKA, (DK)
3 • Hansen, Christian, France, RANSKA, (FR)

(74) Asiamies - Ombud - Agent

Berggren Oy Ab, Antinkatu 3 C, 00100 Helsinki

(54) Keksinnön nimitys - Uppfinningens benämning - Title of the invention
Entsyymi, jolla on sellulaasiaktiivisuutta
Enzym som har cellulasaktivitet

(83) Mikro-organismitalletus - Deposition av mikroorganism - Deposit of microorganism

40250 NCIMB

Entsyymi, jolla on sellulaasiaktiivisuutta - Enzym som har cellulasaktivitet

Esillä oleva keksintö kohdistuu entsyymiin, jolla on sellulaasiaktiivisuutta, DNA-rakenteeseen, joka koodaa entsyymiä, selluloosaa hydrolysoivaan aineeseen, joka käsittää entsyymin, ja entsyymiä sisältävään pesuaineeseen.

Biomassaan, joka koostuu pääasiallisesti selluloosasta, hemiselluloosasta ja ligniinistä, on kiinnitetty lisääntyvästi huomiota tärkeänä energian uusiutuvana lähteenä (mukaan lukien ravintoenergia). Fotosynteesin sitoman hiilen määrän on arvioitu olevan 100×10^9 tonnia vuodessa maailmanlaajuisesti ja puolet siitä sisältyy selluloosaan. Jos tämä materiaali tai vähintään merkittävä osa siitä voitaisiin muuttaa nestemäiseksi polttoaineeksi, kaasuksi tai ruokaproteiiniksi, se oli merkittävä apu ratkaistaessa voimavarojen kierrätys- ja säilyttämisongelmia. Kuitenkin on havaittu vaikeaksi kehittää taloudellisesti elinkykyinen menetelmä selluloosamateriaalin muuttamiseksi käymiskykyisiksi sokereiksi.

Yleisesti lupaavin ehdotettu menetelmä käsittää selluloosaa hajottamaan kykenevien entsyymien käytön. Lukuisat mikroorganismit mukaan lukien sienet (esim. Trichoderma reseei, Humicola insolens, Fusarium oxysporum jne) ja bakteerit (esim. Clostridium thermocellum, Cellulomonas spp., Thermonospora spp., Bacterioides spp., Microbispora bispora jne) valmistavat näitä entsyymeitä, jotka tunnetaan yleisesti sellulaaseina. Käymiskykyisten sokerien valmistaminen biomassasta tällaisten entsyymien avulla ei ole vielä kilpailukyistä esimerkiksi glukoosin valmistamisen maissista alfa-amylaasin avulla kanssa johtuen sellulaasientsyymien tehottomuudesta. Tärkeimmät sellulaasien käyttöön liittyvät ongelmat ovat niiden alhainen spesifinen aktiivisuus ja niiden tuotannon suuri hinta. Siksi esiintyy tarvetta kehittää sellulaaseja, jotka hajottavat tehokkaammin selluloosamateriaaleja käymiskykyisiksi sokereiksi.

Biomassan hajotuskäytön lisäksi sellulaaseja on ehdotettu käytettäviksi pesuainekoostumuksissa puuvillaa sisältävien tekstiilien käsittelemiseen, jotka koostuvat pääasiassa selluloosasta. Tiedetään hyvin, että toistuva puuvillaa sisältävien tekstiilien pesu aiheuttaa tavallisesti tekstiilien selvää epämiellyttävää karheutta johtuen amorfisten alueiden läsnäolosta selluloosakuiduissa, jotka alueet muodostavat esiinpullistuvia osia muutoin sileisiin kuituihin. Tämän ongelman ratkaisemiseksi on esitetty aiemmin useita menetelmiä. Esimerkiksi Unilever Ltd.:n US-patentissa 1 368 599 opetetaan sellulaasien käyttö puuvillaa sisältävien tekstiilien karheuden vähentämiseen. Myös US-patentissa 4 435 307 (Novo Industri A/S) opetetaan Humicola insolensista tai sen fraktiosta johdetun selluloosaa hydrolysoivan entsyymin käyttö karheutta vähentävänä pesuaineen lisäaineena. Muut alalla mainittujen selluloosien käytöt käsittävät lian poistamisen tekstiilistä ja tekstiilin värin kirkastamisen (ks. esimerkiksi EP 220 016).

Vaikkakin sellulaasientsyymien käyttö puuvillaa sisältävien tekstiilien karheuden vähentämiseen ehdotettiin ja esitettiin lähes 20 vuotta sitten, tämän prosessin mekanismia ei ole selvitetty ja sitä ei vieläkään tunneta yksityiskohtaisesti. Se johtuu muun muassa entsyymien ja entsyymillä katalysoitujen reaktioiden moninaisuudesta. Itse asiassa esimerkiksi mikrobikantojen luonnossa generoimat sellulaasit ovat sellulaasien kompleksiseoksia. Vastaavasti luonnossa esiintyvien materiaalien, kuten puuvillan, sellulaasien katalysoima konversio on äärimmäisen vaikea analysoida yksityiskohtaisesti.

Näistä olosuhteista johtuen sellulaasien hyväksikäyttö käytännössä karheuden vähentämiseen ja estämiseen samoin kuin värien kirkastamiseen ei ole levinnyt laajalle eikä sillä ole suurta käyttöä käytännössä siitä huolimatta, että se on toivottavaa: monien entsyymijärjestelmien tuotannon optimointi on vaikeaa, ja siksi vaikeudet, jotka johtuvat siitä, että sellulaaseja täytyy käyttää pikemminkin suuria määriä puuvillatekstiilien kovuuden toivottuun vähentämi-

seen ja estämiseen, estävät sellulaasientsyymien kalliin teollisen tuotannon toteuttamisen ja niiden varsinaisen käytön: esimerkiksi suurten entsyymimäärien lisääminen pesuainekoostumuksiin ei ole pesuaineformuloinnissa yhteensopivaa muiden aineosien optimaalisten funktioiden kanssa, eikä hyvin suurten entsyymimäärien lisääminen pesuainekoostumuksiin ole ole esimerkiksi kuluttajaturvallisuuden parhaaksi.

Siksi esillä olevan keksinnön kohteena on saada aikaan sellulaasientsyymeitä, joilla on erittäin spesifinen aktiivisuus.

Esillä oleva keksintö kohdistuu entsyymiin, jolla on sellulaasiaktiivisuutta ja jonka entsyymin voivat tuottaa Bacillus spp. -kannat, NCIMB 40250, tai niiden kaltainen Bacillus spp. -kanta, tai sanotun sellulaasin johdannainen. Kanta NCIMB 40250 talletettiin 18. tammikuuta 1990 National Collection of Industrial and Marine Bacteriaan, 23 St Machar Drive, Aberdeen, Skotlanti, UK, Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure mukaisesti.

Tässä yhteydessä ilmauksella "entsyymi, jolla on sellulaasiaktiivisuutta" on tarkoitettu ymmärrettävän entsyymiä, joka osallistuu selluloosan hajotusprosessiin. On olemassa kolme erilaista sellulaasityyppiä, jotka toimivat synergisesti tuottaen liukenevia sokereita: endoglukanaaseja, joissa on selluloosa-affiniteetti ja jotka tunkeutuvat selluloosakuidussa oleviin alhaisen kiteyden amorfisiin alueisiin saaden aikaan vapaiden päiden muodostumisen; eksoglukanaasit, jotka panevat alulle hajoamisen ei-pelkistävistä ketjun päistä poistamalla sellobioosiyksiköt; ja ß-galaktosidaasit, jotka hydrolysoivat sellobioosin glukoosiksi.

Ilmauksen "kaltainen Bacillus spp. -kanta" on tarkoitettu merkitsevän kantaa, joka kuuluu samaan Bacillus-lajiin kuin kanta NCIMB 40250, tai kantaa, joka on erittäinen samanlainen laji. Laji, johon kanta NCIMB 40250 kuuluu, on identi-

fioitu Bacillus lautukseksi. Esillä olevan keksinnön piiri on myös tarkoitettu käsittämään ainakin sellulaasientsyymit, joita muut Bacillus lautus -kannat kuin NCIMB 40250 voivat valmistaa.

Termin "johdannainen" on tarkoitettu merkitsevän proteiinia, joka on johdettu luonnollisesta proteiinista lisäämällä yhtä tai useampaa aminohappoa joko jompaankumpaan tai kumpaankin C- ja N-terminaalisista päistä, substituoimalla yksi tai useampi aminohappo yhdessä tai useassa luonnollisen aminohapposekvenssin erilaisessa kohdassa, deletoimalla yksi tai useampi aminohappo luontaisen proteiinin jommastakummasta tai kummastakin päästä tai aminohapposekvenssissä olevasta yhdestä tai useammasta kohdasta tai insertoimalla yksi tai useampi aminohappo luonnollisessa aminohapposekvenssissä olevaan yhteen tai useampaan kohtaan.

Vaikkakin keksinnön mukaista entsyymiä voidaan valmistaa viljemällä Bacillus spp. -kantaa NCIMB 40250 tai sen kaltaista kantaa ja eristämällä entsyymi viljelmästä, on tavallisesti edullisempaa valmistaa entsyymi yhdistelmä-DNA-tekniikoilla, jotka mahdollistavat tuotetun entsyymin saannon optimoimisen. Lisäksi entsyymeitä koodaavat kloonatut geenit voidaan modifioida entsyymien saamiseksi, joilla on parantuneet ominaisuudet.

Täten esillä oleva keksintö kohdistuu toisessa kohteessaan DNA-rakenteeseen, joka käsittää DNA-sekvenssin, joka koodaa entsyymiä, jolla on sellulaasiaktiivisuutta ja joka entsyymi on johdettavissa Bacillus spp. -kannasta NCIMB 40250 tai sen kaltaisesta Bacillus spp. -kannasta tai sanotun sellulaasin johdannaisesta. Keksintö kohdistuu lisäksi ekspressiovektoriin, jossa on edellä esitetty insertoitu DNA-rakenne, sekä soluun, joka on transformoitu DNA-rakenteella tai vektorilla.

Lisäkohteena keksintö kohdistuu selluloosaa hydrolysoivaan aineeseen, joka kykenee hajottamaan selluloosakuitujen

amorfiset alueet aineen käsittäessä entsyymin, jolla on edellä määriteltyä sellulaasiaktiivisuutta.

Keksintö kohdistuu myös pesuainekoostumukseen, joka käsittää selluloosaa hydrolysoivaa ainetta. Keksinnön mukaisen sellulaasientsyymin on yllättäen havaittu olevan vakaampaa pesemisen aikana (60 minuuttia 40 °C:ssa) konventionaalisten pesuaineiden läsnäollessa kuin kaupallisen sellulaasivalmisteen (Celluzyme™, sellulaasivaste, joka on eristetty Huminocola insolensistä, saatavissa Novo Nordisk A/S:stä). Parantuneen stabiiliuden syy voi olla entsyymin alkalofiilisessä luonteessa (katso esimerkki 5). Lisäksi on spekuloitu, että se voi olla olla seurausta oksidointistabiiliudesta tai pesuaineisiin yleisesti sisällytetyistä proteaaseista. Jos näin on, keksinnön mukaisella sellulaasientsyymillä on myös parantunut stabiilius nestemäisissä pesuaineissa, jotka sisältävät proteaaseja.

Esillä olevan keksinnön mukainen sellulaasientsyymi on edullisesti sellainen, jossa on endoglukanaasiaktiivisuutta (nimitetääm seuraavassa endoglukanaasiksi), erityisesti sellainen, jossa on vähintään noin 10, edullisemmin vähintään noin 20 ja edullisimmin vähintään noin 25, kuten noin 30, endoglukanaasiaktiivisuus CMC-endoaasiyksikköä milligrammaa kohti kokonaisproteiinia alkalisissa olosuhteissa. Endoglukanaasiaktiivisuus määritetään karboksimetyyliselluloosan liuoksen viskositeetin laskuna esillä olevan keksinnön mukaisella entsyymillä inkuboinnin jälkeen seuraavissa olosuhteissa:

Valmistetaan substraattiliuos, joka sisältää 35 mg/l CMC (Hercules 7 LFD) 0,1 M Tris-puskurissa pH:ssa 9,0. Analysoitava entsyyminäyte liuotetaan samaan puskuriin.

10 ml substraattiliuosta ja 0,5 ml entsyymiliuosta sekoitetaan ja siirretään viskosimetriin (esim. Haake VT 181, NV-sensori, 1818 rpm), lämpösäädetään 40 °C.

Viskositeettilukemat otetaan niin pian kuin mahdollista sekoittamisen jälkeen ja uudelleen 30 min kuluttua. Entsyymin määrä, joka vähentää näissä olosuhteissa viskositeettia puolella, määritetään 1 CMC-endoaasiyksikkönä.

On huomattava, että keksinnön mukainen endonukleaasi on sellainen, joka on aktiivinen (CMC-endonukleaasiaktiivisuus) alkalisissa olosuhteissa. Endoglukanaasi on erityisemmin sellainen, jonka pH-optimi on pH:ssa noin 7,5-10,5. Poiketen useista tunnetuista sellulaaseista, jotka ovat aktiivisia happamassa pH:ssa ja suhteellisen inaktiivisia alkalisissa pH-arvoissa, tämä piirre tekee keksinnön mukaisen endonukleaasin erityisen käyttökelpoiseksi pesemistarkoituksiin, erityisesti pesuainekoostumuksen aineosaan, koska vaatteiden peseminen suoritetaan tavallisesti alkalisissa olosuhteissa johtuen useimpien pyykinpesuaineiden alkalisuudesta. Alkalofiilisia sellulaaseja tunnetaan esimerkiksi EP 271 004:stä mutta niillä ei ole osoitettu olevan suurta selluloosa-affiniteettia, joka on esillä olevan keksinnön mukaisen sellulaasientsyymin tapauksessa ja jolla on myös suurempi spesifinen aktiivisuus.

Esillä olevan keksinnön mukainen entsyymi on edullisesti sellainen, joka on aktiivista lämpötilassa, jossa vaatteita tavallisesti pestään ja joka on tavallisesti lämpötila noin 60 °C:een saakka. Täten kannasta NCIMB 40250 eristetty luonnollinen entsyymi on aktiivista lämpötiloissa noin 45-65 °C. Tämä ei kuitenkaan estä sitä, että entsyymi voi olla tietyissä olosuhteissa aktiivista lämpötiloissa yli 65 °C.

Eräs keksinnön mukainen entsyymi on endoglukanaasi, jonka näennäismolekyylipaino on 75 kD, tai sen lohkontatuote, jolla on endoglukanaasiaktiivisuutta. Termin "lohkontatuote" on tarkoitettu merkitsevän entsyymin lyhyempää muotoa, joka on tulosta esimerkiksi kemiallisesta tai entsymaattisesta lohkaisusta (esim. sopivalla proteaasilla) entsyymin talteenoton jälkeen tai entsyymiä tuottavan organismin posttranslationaalisesta käsittelystä, esim. N- ja/tai C-terminaalin käsittelystä, joka voi synnyttää entsyymin kyp-

sän muodon. ~75 kD entsyymin lohkontatuotteen erityinen esimerkki, joka on esillä oleviin tarkoituksiin mielenkiintoinen, on suunnilleen 58 kD:n tuote, joka on valmistettu viljelemällä isäntäorganismia, joka on transformoitu ~75 kD:n entsyymiä koodaavalla DNA:lla. ~75 kD:n entsyymiä (ja sen ~58 kD:n lohkontatuotetta) nimitetään seuraavissa esimerkeissä Endo1:ksi.

Keksinnön mukainen entsyymi voi olla endoglukanaasi, jota koodaa oheistetussa sekvenssilistassa ID#1 esitetty DNA-sekvenssi (esittää ~75 kD:n entsyymin sekvenssin) tai sen modifikaatio, joka koodaa sanotun endonukleaasin johdannaista.

Endoglukanaasijohdannaiset voidaan saada tavanomaisesti modifioimalla sopivasti luonnollista endoglukanaasia koodaava DNA-sekvenssi. Esimerkkejä sopivista DNA-sekvenssin modifikaatioista ovat nukleotidisubstituutiot, jotka eivät synnytä endoglukanaasin muuta aminohapposekvenssiä mutta jotka vastaavat isäntäorganismin kodonin käyttöä, johon isäntäorganismiin DNA-rakenne viedään, tai nukleotidisubstituutiot, jotka synnyttävät erilaisen aminohapposekvenssin ja siten mahdollisesti erilaisen polypeptidirakenteen ilman, että ne kuitenkaan heikentävät endoglukanaasin ominaisuuksia. Muut esimerkit mahdollisista modifikaatioista ovat yhden tai useamman nukleotidin liittäminen sekvenssiin, yhden tai useamman nukleotidin lisääminen sekvenssin jompaankumpaan päähän ja yhden tai useamman nukleotidin deletoiminen joko sekvenssin päästä tai sen sisältä. Tällaiset luonnollisia proteiineja koodaavan DNA:n modifikaatiot ovat erittäin tunnettuja ja niitä käytetään alalla laajasti.

Toinen keksinnön mukainen entsyymi on endoglukanaasi, jonka näennäismoolipaino on 56 kD, tai sen lohkontatuote (kuten edellä määritettiin), jolla on endoglukanaasiaktiivisuutta. Tätä entsyymiä nimitetään jäljempänä esimerkissä 2 Endo2:ksi.

Keksinnön mukainen entsyymi voi olla endoglukanaasi, jota koodaa oheistetussa sekvenssilistassa ID#3 esitetty DNA-sekvenssi (esittää ~56 kD:n entsyymin sekvenssin) tai sen modifikaatio, joka koodaa sanotun endonukleaasin johdannaista (kuten edellä määritettiin).

Toinen keksinnön mukainen entsyymi on endoglukanaasi, jonka näennäismoolipaino on 45 kD, tai sen lohkontatuote (kuten edellä määritettiin), jolla on endoglukanaasiaktiivisuutta. Esimerkki tällaisesta spesifisestä lohkontatuotteesta on suunnilleen 30 kD:n lohkontatuote, joka valmistettu viljelemällä isämtäorganismia, joka on transformoitu ~45 kD:n entsyymiä koodaavalla DNA:lla. Tätä entsyymiä nimitetään jäljempänä esimerkissä 3 Endo3A:ksi.

Keksinnön mukainen entsyymi voi olla endoglukanaasi, jota koodaa oheistetussa sekvenssilistassa ID#5 esitetty DNA-sekvenssi (esittää DNA-sekvenssin, joka koodaa ~45 kD:n tuotetta) tai sen modifikaatio (kuten edellä määritettiin), joka koodaa sanotun endonukleaasin johdannaista (kuten edellä määritettiin).

Muut endoglukanaasikloonista 3 (ks. esimerkki 3) valmistetut entsyymit, joilla on endoglukanaasiaktiivisuutta ovat suunnilleen 60 ja 56 kD:n proteiineja, joita nimitetään vastaavasti Endo3B:ksi ja Endo3C:ksi.

Lisäksi vielä eräs keksinnön mukainen entsyymi on endoglukanaasi, jonka näennäismoolipaino on 92 kD, tai sen lohkontatuote (kuten edellä määritettiin), joilla on endoglukanaasiaktiivisuutta. Tätä entsyymiä nimitetään jäljempänä esimerkissä 4 Endo4:ksi. Muut endoglukanaasikloonista 4 (ks. esimerkki 4) valmistetut entsyymit, joilla on endoglukanaasiaktiivisuutta, ovat suunnilleen 74 ja 71 kD:n proteiineja, Kumpikin näistä voi olla yksittäinen entsyymi tai 92 kD:n entsyymin lohkontatuote.

On havaittu, että keksinnön mukaiset entsyymit, esim. Endo1 ja Endo3A koostuvat ydinalueesta, joka käsittää katalyyt-

tisesti aktiivisen domeenin ja alueen, joka käsittää domeenin, jonka funktio on välittää sitoutumista selluloosasubstraatteihin (so. selluloosaa sitova domeeni; tämä vastaa Bacillus subtilisista peräisin olevassa endosellulaasissa olevaa samanlaista domeenia (Nakamura et al., 1987). Esimerkiksi Endo1:n täysipituinen ~75 kD-muoto käsittää ydinalueen ja C-terminaalisen selluloosaa sitovan domeenin, joka joissakin tapauksissa voidaan lohkaista pois proteolyyttisesti jättämällä ~58 kD:n ydinalue. Selluloosaa sitovan domeenin läsnäolon on havaittu olevan tärkeää värien kirkastumisen aikaansaamiseksi esipestyissä tekstiileissä (ks. esimerkki 6).

Tähän havaintoon perustuen voi olla mahdollista generoida uusia sellulaasientsyymien johdannaisia, jolloin esillä olevan keksinnön mukaisesta endoglukanaasista johdettuun ydinalueeseen yhdistyy toisesta sellulaasientsyymistä johdettu selluloosaa sitova domeeni (esim. Cellulomonas fimi-sellulaasista johdettu). Vaihtoehtoisesti voi olla mahdollista yhdistää toisesta sellulaasientsyymistä johdettuun ydinalueeseen esillä olevan keksinnön mukaisesta endoglukanaasista johdettu selluloosaa sitova domeeni. Erityisessä suoritusmuodossa ydinalue voi olla johdettu sellulaasientsyymistä, joka ei luonnossa käsitä selluloosaa sitovaa domeenia ja joka ulottuu C-terminaalisesti esillä olevan keksinnön mukaisesta endoglukanaasista johdettuun selluloosaa sitovaan alueeseen. Tällä tavoin on mahdollista konstruoida sellulaasientsyymeitä, joilla on paremmat sitoutumisominaisuudet.

Keksinnön DNA-rakenne voi olla sellainen, joka käsittää DNA-sekvenssin, jokaa koodaa mitä tahansa edelllä kuvattua entsyymiä tai näiden edellä kuvattujen entsyymien johdannaisia. DNA-rakenne voi olla genomista tai cDNA-alkuperää, esimerkiksi sellainen, joka on saattu valmistamalla genominen tai cDNA-kirjasto sopivasta Bacillus spp. -kannasta (esim. kanta NCIMB 40250 tai sen kaltainen kanta) ja seulomalla DNA-sekvenssit, jotka koodaavat sopivaa sellulaasia kokonaan tai osittain, hybridoimalla käyttäen synteettisiä

oligonukleotidikoettimia standardien tekniikoiden mukaisesti (ks. Maniatis et al., 1982).

Keksinnön mukainen ekspressiovektori, johon on insertoitu keksinnön mukaista entsyymiä koodaava DNA-rakenne, voi olla mikä tahansa vektori, joka kykenee replikoitumaan itsenäisesti tietyssä isäntäeläimessä, tavallisesti plasmidi tai bakteriofagi. Keksinnön mukaista entsyymiä koodaava DNA-sekvenssi tulee olla vektorissa kytkettynä toimivasti sopivaan promoottorisekvenssiin. Promoottori voi olla mikä tahansa DNA-sekvenssi, jossa on transkriptionaalista aktiivisuutta isäntäsolussa ja se voi olla johdettu geeneistä, jotka koodaavat proteiineja, jotka ovat joko homologisia tai heterologisia isäntäorganismin kanssa.

Promoottori johdetaan edullisesti geenistä, joka koodaa proteiinia, joka on homologinen isäntäorganismin kanssa. Esimerkkejä sopivista promoottoreista ovat E. colin lacm Streptomyces coelicorin dagA ja Bacillus licheniformisin amyL.

Keksinnön mukainen ekspressiovektori käsittää lisäksi DNA-sekvenssin, joka mahdollistaa vektorin replikoitumisen isäntäsolussa. Esimerkkejä tällaisista sekvensseistä ovat plasmidien pUC19, pACYC177, pUB110 ja pIJ702 replikaation alkuperät.

Keksinnön mukainen ekspressiovektori voi käsittää lisäksi DNA-sekvenssin, joka koodaa signaalipeptidiä entsyymin solunulkoisen ilmentymisen aikaansaamiseksi. DNA-sekvenssi voi olla esimerkiksi B. licheniformiksen alfa-amylaasigeenistä peräisin oleva signaalisekvenssi.

Vektori voi käsittää lisäksi valikoitavan merkitsimen, esimerkiksi geenin, jonka tuote tuottaa antibiottiresistenssin, kuten ampisilliini-, kloramfenikoli- tai tetrasykliiniresistenssin, tai B. subtiliksesta tai B. licheniformisista peräisin olevat dal-geenit.

Menetelmät, joita käytetään keksinnön mukaista entsyymiä koodaavien DNA-sekvenssien ja vastaavasti promoottorin ligatoimiseen ja niiden insertoimiseen sopiviin vektoreihin, jotka sisältävät informaation, joka on tarpeellista isäntäsolussa replikoitumiseen, ovat alan ammattilaisten hyvin tuntemia (ks. esim. Maniatis et al., op.cit.)

Esillä olevan keksinnön mukainen isäntäsolu voidaan transformoida keksinnön mukaisella DNA-rakenteella, joka koodaa edellä kuvattua sellulaasientsyymiä. Tässä tapauksessa voidaan DNA-rakenne integroida sopivasti isäntäkromosomiin, mikä voi olla etu, koska sellulaasia koodaava DNA-sekvenssi on helpommin pidettävissä vakaana solussa. DNA-rakenteiden integrointi isäntäkromosomiin voidaan suorittaa konventionaalisten menetelmien mukaisesti esim. homogeenisella rekombinaatiolla.

Vaihtoehtoisesti voidaan isäntäsolu transformoida edellä kuvatulla ekspressiovektorilla.

Keksinnön mukaisessa menetelmässä käytetty isäntäsolu voi olla mikä tahansa sopiva bakteeri, joka tuottaa viljeltäessä suuria määriä keksinnön mukaista entsyymiä. Esimerkit sopivista bakteereista voivat olla gram-positiivisia bakteereita, kuten bakteereita, jotka kuuluvat sukuun Bacillus, esim. Bacillus subtilis, Bacillus licheniformis, Bacillus lentus, Bacillus brevis, Bacillus stearothermophilus, Bacillus alkalophilus, Bacillus amyloquefaciens, Bacillus coagulans, Bacillus circulans tai Bacillus lautus, tai gram-negatiivisia bakteereita, kuten Escherichia coli. Bakteerien transformointi voidaan saada aikaan protoplastitransformoinnilla tai käyttämällä kompetentteja soluja tavalla, joka on tunnettu per se.

Vielä eräässä kohteessaan keksintö kohdistuu menetelmään keksinnön mukaisen sellulaasientsyymin valmistamiseksi, joka menetelmä käsittää edellä kuvatun isäntäsolun viljelemisen olosuhteissa, jotka johtavat sellulaasin tai sen joh-

dannaisen tuotantoon, ja sellulaasin tai sen johdannaisen talteenottamisen soluista ja/tai viljelyväliaineesta.

Solujen viljelemiseen käytetty väliaine voi olla mikä tahansa konventionaalinen väliaine, joka sopii bakteerien kasvattamiseen. Sellulaasi voidaan ottaa talteen väliaineesta konventionaalisilla menetelmillä mukaan lukien solun erottaminen väliaineesta sentrifugoinnilla tai suodattamalla, ja jos tarpeellista solujen hajottamisen jälkeen presipitoimalla supernatantin tai filtraatin proteiinikomponentit suolan, esim. ammoniumsulfaatin, avulla, mitä seuraa puhdistaminen erilaisilla kromatografiamenetelmillä, esim. ioninvaihtokromatografialla, affiniteettikromatografialla tai niiden kaltaisella.

Keksinnön mukaisen menetelmän erityisessä suoritusmuodossa sellulaasi otetaan talteen kypsässä muodossa, joko tuloksena edellä selitetystä proentsyymin posttranslationaalisesta käsittelystä tai tuloksena entsyymiä koodaavan DNA-sekvenssin sopivista modifikaatioista entsyymin N- ja/tai C-terminaalisissa sekvensseissä olevia typistyksiä vastaavien DNA:n deleetioiden muodossa.

On syytä uskoa, että erilaisilla sellulaaseilla on synergistä vaikutusta selluloosan hajoamisen suhteen. Keksinnön mukainen selluloosaa hydrolysoiva aine voi siten käsittää edullisesti yhden tai useamman keksinnön mukaisen sellulaasientsyymin kombinaation tai yhden tai useamman keksinnön mukaisen sellulaasientsyymin kombinaation yhdessä yhden tai useamman muun sellulaasiaktiivisuutta omaavan entsyymin kanssa. Tällaiset sellulaasit voivat olla endosellulaaseja tai eksosellulaaseja riippuen selluloosaa hydrolysoivan aineen aiotusta käytöstä (esim. aiotun selluloosahajoamisen tasosta). Muut sellulaasit voidaan valita niistä, jotka voidaan erottaa Humicola-lajista, kuten Humicola insolens (esim. kanta DSM 1800), Fusarium-lajista, kuten Fusarium oxysporum (esim. kanta DSM 2672), Myceliopthora-laji, kuten Myceliopthora thermophile, Erwinia-laji, kuten Erwinia chrysanthermis (ks. M.H. Boyer et al., Eur. J. Biochem.

162, 1987, s. 311-316), Trichoderma-lajista, kuten Trichoderma reseei, Microbispora-lajista, kuten Micorbispora bispora, Neocallimastix-lajista, kuten Neocallimastix frontalis, Piromonas-lajista, kuten Piromonas communis, Robillarda spp., Cellulomonas-lajista, kuten Cellulomonas fimi, Clostridium-lajista, kuten Clostridium thermoclellum, Pseudomonas spp., Thermonospora spp., Bacterioides spp. tai Ruminococcus spp.

Keksinnön mukainen selluloosaa hydrolysoiva aine voi olla sopivasti pölyämättömän granulaatin, stabiloidun nestemäisen tai suojatun entsyymin muodossa. Pölyämättömät granulaatit voidaan valmistaa esim. US-4 106 991:n ja 4 661 452:n (Novo Industri A/S) mukaisesti ja ne voidaan päällystää optionaalisesti alalla tunnetuilla menetelmillä. Nestemäiset entsyymivalmisteet voidaan stabiloida esimerkiksi lisäämällä polyolia, kuten propyleeniglykolia, sokeria, tai sokerialkoholia, maitohappoa tai boorihappoa vakiintuneiden menetelmien mukaisesti. Suojatut entsyymit voidaan valmistaa EP-238 216:ssa selostetun menetelmän mukaisesti.

Selluloosaa hydrolysoivalla aineella voi olla sopivasti endoglukanaasiaktiivisuutta 500-10000 CMC-endoaasiyksikköä (kuten edellä määritettiin) ainegrammaa kohti. Selluloosaa hydrolysoiva aine on sopivasti pesuaineen lisäaine, joka voi käsittää yhtä tai useampaa entsyymiä, kuten proteaasia. lipaasia ja/tai amylaasia, joita sisällytetään konventionaalisesti pesuaineen lisäaineisiin.

Keksinnön mukainen pesuainekoostumus, joka käsittää edellä kuvattua selluloosaa hydrolysoivaa ainetta, käsittää lisäksi yhtä tai useampaa pinta-aktiivista ainetta, joka voi olla tyypiltään anioninen, ei-ioninen, kationinen, amfoteerinen tai zwitterioninen sekä näiden pesuaineluokkien seoksia. Tyypilliset esimerkit anionisista pinta-aktiivisista aineista ovat lineaariset alkyylibentseenisulfonaatit (LAS), alfa-olefiinisulfonaatit (AOS), alkoholietoksisulfaatit (AES) ja luontaisten rasvahappojen alkalimetallisuolat.

Keksinnön mukainen pesuainekoostumus voi sisältää muita alalla tunnettuja pesuaineen aineosia esim. rakenneaineita, valkaisuaktivaattoreita, korroosion estoaineita, sekvestrausaineita, lian uudelleenkerrostumisen estoaineita, hajusteita, entsyymin stabilointiaineita jne.

Keksinnön mukainen pesuainekoostumus voidaan formuloida mihin tahansa sopivaan muotoon, esimerkiksi jauheeksi tai nesteeksi. Entsyymi voivaan tarvittaessa stabiloida nestemäiseen pesuaineeseen sisällyttämällä edellä mainittuja entsyymin stabilointiaineita. Tavallisesti pesuainekoostumusliuoksen pH on 7-12 ja joissakin tapauksissa 7,0-10,5. Muita pesuaine-entsyymeitä, kuten proteaaseja, lipaaseja tai amylaaseja voidaan sisällyttää keksinnön mukaiseen pesuainekoostumukseen joko erikseen tai edellä kuvattuna yhdistettynä lisäaineena.

Keksinnön mukaisella sellulaasientsyymillä saatavat pehmentävät, likaa poistavat ja väreja kirkastavat vaikutukset edellyttävät tavallisesti pesuliuoksen sellulaasikonsentraatiota, joka vastaa 5-200 CMC-endoaasiyksikköä/litra endoglukanaasiaktiivisuutta. Keksinnön mukaista pesuainekoostumusta käytetään pesuliuoksessa tavallisesti konsentraatioina 0,5-20 g/l. Tavallisesti keksinnön mukaisen pesuaineen sellulaasikonsentraatio on noin 0,3-400 CMC-endoaasiyksikköä grammaa kohti. Yleensä on sopivampaa lisätä pesuaineen lisäainetta määrinä 0,1-5 % paino/paino tai edullisesti määrinä 0,2-2 % pesuainekoostumuksesta. Erityiskäytöissä, esimerkiksi silloin kun koostumusta käytetään värin kirkastamiseen tai toistuvan pesemisen vahingoittaman tekstiilin karheuden poistoon, on sopivaa sisällyttää suurempia määriä, kuten noin 20 % paino/paino, selluloosaa hydrolysoivaa ainetta.

Vielä eräässä kohteessa esillä oleva keksintö kohdistuu menetelmään, jossa vähennetään määrää, jona selluloosaa sisältävät tekstiilit tulevat karheaksi, tai selluloosaa sisältävien tekstiilien karheuden vähentämiseksi, jolloin menetelmä käsittää selluloosaa sisältävän tekstiilin käsit-

telemisen edellä kuvatulla selluloosaa hydrolysoivalla aineella. Keksinnön mukainen menetelmä voidaan suorittaa käsittelemällä selluloosaa sisältävät tekstiilit pesemisen aikana. Selluloosaa hydrolysoivaa ainetta voidaan lisätä sellaisenaan määränä, joka tarvitaan saatavaksi haluttuun vaikutukseen, tai sitä voidaan lisätä pesuainekoostumuksen aineosana, kuten edellä selostettiin. Kuitenkin haluttaessa voidaan tekstiilien käsittely suorittaa liottamisen tai huuhtelemisen aikana tai yksinkertaisesti lisäämällä selluloosaa hydrolysoivaa ainetta veteen, johon tekstiilit on upotettu tai tullaan upottamaan.

Keksinnön mukaista selluloosaa hydrolysoivaa ainetta voidaan käyttää selluloosaa sisältävien tekstiilien värin kirkastamiseen. Toistuvan pesemisen jälkeen kehittyy tällaisiin tesktiileihin usein harmahtava ulkonäkö. Tämä vaikutus on erityisen selvää värillisten tekstiilien, erityisesti tummien tekstiilien, yhteydessä ja tämä johtuu luultavasti selluloosakuitujen värjäämättömien osien tulemisesta näkyviin, kun mekaaniset voimat vahingoittavat selluloosakuituja, joista tekstiili koostuu. Kuitujen vahingoittuneiden osien oletetaan olevan amorfisempia kuin vahingoittumattomien selluloosakuitujen ja siten alttiimpia esillä olevan keksinnön mukaisten sellulaasien vaikutukselle. Värin kirkastusvaikutus on selvempää, kun selluloosa hydrolysoiva aine sisältää endoglukanaasia, joka käsittää selluloosaa sitovan alueen (ks. esimerkiksi esimerkki 6).

Vastaavasti esillä oleva keksintö kohdistuu lisäksi menetelmään värillisten selluloosaa sisältävien tekstiilien käsittelemiseksi värin kirkastumisen aikaan saamiseksi, jolloin menetelmä käsittää selluloosaa sisältävän tekstiilin käsittelemisen keksinnön mukaisella selluloosaa hydrolysoivalla aineella. Keksinnön mukainen menetelmä voidaan suorittaa käsittelemällä selluloosaa sisältävät tekstiilit vesiväliaineessa pesemisen aikana. Selluloosaa hydrolysoiva aine voidaan lisätä sellaisenaan määränä, joka on riittävä halutun vaikutuksen saamiseen, tai sitä voidaan lisätä edellä kuvatun pesuainekoostumuksen aineosana. Haluttaessa

voidaan tekstiilien käsittely suorittaa myös liottamisen tai huuhtelemisen aikana tai yksinkertaisesti lisäämällä selluloosaa hydrolysoivaa ainetta veteen, jossa vaatteet ovat tai johon ne upotetaan. Värinkirkastuksen yhteydessä voi vesiväliaineen endoglukanaasiaktiivisuus olla sopivasti yli 250 CMC-endoaasiyksikköä vesiväliainelitraa kohti.

Lisäksi on mahdollista käyttää keksinnön mukaista selluloosaa hydrolysoivaa ainetta tekstiilin värin paikallisten muutosten saamiseksi, jotta saadaan aikaan tekstiilin "kivipesty" ulkonäkö (mitä tulee sellulaasientsyymien käyttöön yleisesti tässä tarkoituksessa, katso esimerkiksi EP 307 564).

Keksinnön mukainen selluloosaa hydrolysoiva aine on myös käyttökelpoinen paperimassan valmistuksessa, esim. massan vedenpoistossa (sellulaasientsyymien käytön osalta tavanomaisesti tähän tarkoitukseen katso esimerkiksi EP 262 040) samoin kuin musteen poistoon uudelleenkierrätykseen tarkoitetusta paperista (tähän tarkoitukseen yleisesti käytettyjen sellulaasientsyymien osalta katso esimerkiksi JP 59-9299 tai JP- 63-59494).

Keksintöä valaistaan lisäksi seuraavilla esimerkeillä, joiden ei ole tarkoitettu rajoittavan millään tavalla keksinnön piiriä, viitaten oheistettuihin piirroksiin, joissa:

Kuvio 1 esittää neljässä endoglukanaasipositiivisessa E. coli -kloonissa oleviin plasmideihin sisällytetyn Bacillus spp. -DNA-insertin restriktiokarttoja. Endoglukanaasin 1 (Endo1, pPL517) ja endoglukanaasin 2 (Endo2, pP1382) transkription positio ja suunta on osoitettu. Restiktioentsyymikohdat on osoitettu seuraavasti: Pst1 (P), HindIII (H), SmaI (S), SalI (Sa), BamHI (Ba), BglII (B), SphI (sp), EcoRI (E). ▬ osoittaa pBR322 DNA.

Kuvio 2 on erilaisten Endo1-geenin (□) sisältävien plasmidien restriktiokartta. pBR322:n ß-laktamaasipromoottori osoitetaan nuolella (Pb) ja Endo1-geenin trasnkription

suunta osoitetaan nuolella (Pb). (▬▬): pBR322, (▓▓): pBR322:lla koodattujen aminohappojen "häntä". Restriktioentsyymikohdat on osoitettu seuraavasti: PstI (P), BglII (B), EcoRI (E), ClaI (C). Osoitettuja plasmideja sisältävien E. coli MC1000 -uutteiden endoglukanaasiaktiivisuus on esitetty oikealle (A) sellulaasiyksikköä /ml viljelyväliainetta.

Kuvio 3 esittää lämpötilan vaikutusta E. coli MC1000 (pPL517) -sellulaasiuutteisiin ja moninkertaisiin sellulaasiaktiivisuuksiin Bacillus spp. PL236 -supernatantissa. Aktiivisuus mitattiin lämpötilassa, joka esiintyi 30 min. inkubointiaikana. Katso Materiaalit ja menetelmät. Erilaisten lämpötilojen aktiivisuus on esitetty sellulaasiyksikönä/ml alkuperäistä viljelytilavuutta.

Kuvio 4 esittää endoglukanaasin Endo1, joka on esitetty kuviossa endosellulaasina 1, relevantteja restriktiokohtia ja sekvennointistrategiaa. (⟼): sekvenssireaktioiden laajuus ja suunta. Lyhennykset: EcoRI (E), PstI (P), BglII (Bg), BstNI (Bs), AvaI (A), ClaI (C), BamHI (Ba).

Kuvio 5 on plasmidin pDN 2801 restriktiokartta. Restriktioentsyymikohdat on esitetty seuraavasti: EcoRI (E), BglII (Bg), HindIII (H), SmaI (Sm), SalI (Sa), SphI (Sp), PstI (P), EagI (Ea), ClaI (C), BamHI (Ba). CAT osoittaa geenin, joka välittää kloramfenikoliresistenssin. P_m osoittaa Bacillus maltogeenisen alfa-amylaasipromoottorin.

Kuvio 6 esittää erilaisten Endo1-geenin sisältävien plasmidien (B. subtilis-replikointialkuperää) restriktiokarttoja. Maltogeeninen alfa-amylaasipromoottori on osoitettu nuolella (P_m), joka osoittaa myös Endo1-geenin transkription suunnan, (▬▬): pDN2801, (▓▓): pDN2801:llä koodattujen sekvenssien häntä. Restriktioentsyymipaikat on osoitettu seuraavasti: Eag I (Ea), BglII (bg), PstI (P), BamHI (B). Osoitettuja plasmideja sisältävän B. substilis DN 1815:n uutteissa oleva endosellulaasiaktiivisuus on osoi-

tettu oikealle (A) endoglukanaasiyksikköinä/ml viljelyväliainetta.

Kuvio 7a ja b esittävät pTL05:n ja pLA03:n konstruktiota. (▩) esittää vektorisekvensseillä koodattuja C-terminaalisia "häntiä". (⬟) osoittaa odotetun C-terminaalisen lohkaisukohdan. Restriktiokohdat on lyhennetty seuraavasti: PstI (P), Pvu (Pv), BamHI (B), BglII (Bg), SalI (S).

Kuvio 8a ja b esittävät pCH57:n konstruktiota. (▨): Alfa-amylaasin signaalisekvenssi. (▥): Endo1-glukanaasin signaalisekvenssi (▬): "suora toisto" esim. kypsän Endo1-geenin alku. (●): Alfa-amylaasin ribosomin sitomiskohta (○): Endo1-glukanaasin ribosomin sitomiskohta (➡): alfa-amylaasipromoottori. Restriktiokohdat on lyhennetty seuraavasti: PstI (P), SalI (S), KpnI (K), EagI (Ea), BglII (Bg).

Kuvio 9 esittää endoglukanaasin Endo2 (esitetty kuviossa endosellulaasina 2). (⟼): sekvenssireaktioiden laajuus ja suunta. Lyhennykset: ScaI (Sc), SacI (Sa), XmnI (X), HindIII (H), PstI (P).

Kuvio 10 on tsymogrammi, joka esittää erilaisista plasmideista, joissa on alkuperäinen endoglukanaasi 3 -klooni (pPL591), saatujen aktiivisten proteiinien molekyylipainon sekä alkuperäisessä insertissä olevat deleetiot.

Endoglukaansi 3 -kloonin (esitetty kuviossa endosellulaasina 3) (pPL591 ja alkuperäisessä insertissä olevia deleetioita sisältävien plasmidien resktriktiokartat. Endoglukanaasigeenien Endo3A, Endo3B ja Endo3C asema plasmidissa pPL591 olevassa insertissä on esitetty restriktiokartassa. Restriktioentsyymikohdat on osoitettu seuraavasti: HindIII (H), SmaI (S), EcoRI (E).

Esimerkit

Materiaalit ja menetelmät

a) Bakteerikannat ja plasmidit

Luovuttajakanta, Bacillus spp., kanta PL236, eristettiin kompostinäytteestä Lyngbystä, Tanskasta, sen suuren selluloosaa hydrolysoivan aktiivisuuden perusteella. Tämän lajin näyte talletettiin 18. tammikuutaa 1990 National Collection of Industrial and Marine Bacteriaan, 23 St. Machar Drive, Aberdeen, Skotlanti, talletusnumerolla NCIMB 40250.

Seuraavia E. coli -lajeja ja plasmideita käytettiin MC 000 (araD139), (ara, leu)7967, lacX74, galU, galK, rpsL) (Casabadan et al., 1980), CSR603 (F-, thr-1. leuB6, proA2, phr-1, RecA1, argE3, thi-1, uvrA6, ara-14, lacY1, galK2, xyl-5, mt1-1, rpsL31, tsx-33, -, supE44) (Sancar et al., 1979), PL248 on MC1000, joka sisältää plasmidin pNF2690, joka sisältää replikaatioalkuperän ja kanamysiiniresistenssigeenin, joka on peräisin pACYC177:stä (Chang and Cohen, 1978) ja cI857-repressorigeenin, joka on peräisin kolifagi lambdasta, pBR322 (Bolivare, 1977), pUN121 (Nilson et al., 1983), pUC18 (Yanisch-Perron et al., 1985), pPLc28 (Remaut et al., 1981), pPL170 (Jørgensen, P.L., 1983).

B. subtilis -kokeiden osalta käytettiin seuraavia B. subtilis -lajeja ja plasmideita: DN1885 (amyE, amyR, spo+, pro+) (Diderichsen, Novo Industri A/S) on B. subtilis 168:n johdannainen, PL18001 on DN1885:n johdannainen, josta puuttuu kaksi pääeksoproteaasia (apr-, npr-), DN2801:ssä on pUB110:stä peräisin oleva alkuperä (Keggins et al., 1987), pC194:n Cat-geeni (Horinouchi and Weisblum, 1982) ja B. stearothermophiluksesta peräisin oleva maltogeeninen alfa-amylaasipromoottori (Pm) (Diderichsen and Christiansen 1988), jota seuraa polyliittäjä, Kreft et al. ovat kuvanneet B. subtilis/E.coli sukkulavektorin pJKK3-1, pPL1759:ssä on pUB110:n alkuperä ja kanamysiiniresistenssi ja promoottori, ribosomin sitomiskohta ja signaalisekvenssi, joka on peräisin B. licheniformiksesta (Stephens et al., 1984).

b. Väliaineet

Fosforihapolla paisutettu selluloosa (ASC) valmistettiin kromatografiaselluloosasta (MN 300, Machery, Nagel), kuten

Walseth (1952) kuvasi, poikkeuksella että selluloosajauhe suspendoitiin asetoniin ennen fosforihapolla käsittelyä. Sellulaasiaktiivisuuden havaitsemiseen tarkoitettu väliaine valmistettiin standardi m)-vääliaineena (Maniatis et al.), joka sisältää 0,2 % ASC tai mikrokiteistä selluloosaa (Avicel, Merck).

Bacillus spp.-, B. subtilis- ja E. coli -soluja kasvatettiin NY-väliaineessa (von Meyenburg et al., 1982), LB-väliaineessa (Maniatis et al., 1982) tai BPX-väliaineessa (100 g/l perunatärkkelystä, 50 g/l ohrajauhoa, 0,1 g/l BAN 5000 SKB, 10g/l natriumkaseinaattia, 20 g/l soijapapujauhoa, 9 g/l Na_2HPO_4, $12H_2O$, 0,1 g/l Pluronic).

Väliaineet tehtiin kiinteiksi lisäämällä agaria (20 g/l).

Tetrasykliiniä (20 µg/ml) tai kanamysiinia (10 µg/l) lisättiin tarvittavasti.

c) DNA:n eristäminen

Bacillus spp. PL236 -soluista peräisin olevan kromosomaalisen DNA:n eristämiseksi 250 ml yön yli viljelmästä suspendoitiin uudelleen 10 ml:aan (50 mM Tris-HCl, pH 8,0, 100 mM EDTA) ja inkuboitiin 25 mg:lla lysotsyymiä 20 min. 37 °C:ssa. Seokseen lisättiin 2 ml 10 % (paino/til.) SDS, se sekoitettiin ja pantiin jäälle 10 minuutiksi. Sitten liuokseen lisättiin 15 ml fenolia, joka oli kyllästetty TE-puskurilla (10 mM Tris-HCl, pH 8,0, 1 mM EDTA), se kuumennettiin 65 °C:een, sekoitettiin varovasti ja jäähdytettiin jään päällä. 30 minuutin 40000x g:ssä sentrifugoinnin jälkeen vesifaasi uutettiin eetterillä, presipitoitiin etanolilla ja suspendoitiin uudelleen TE-puskuriin. DNA puhdistettiin lisäksi liittämällä CsCl-tiheysgradienttiin (Maniatis et al., 1982).

E. coli -plasmidi-DNA valmistettiin SDS-lyysimenetelmällä (Maniatis et al., 1982), plasmidi-DNA:n minivalmisteet restriktioanalyysiä ja transformaatioita varten valmistettiin Holmesin ja Quigleyn (1981) mukaisesti. B. subtilis -plas-

midi-DNA valmistettiin alkalilyysimenetelmällä (Maniatis et al., 1982).

d) Bacillus spp. PL236:sta peräisin olevan kromosomaalisen d) DNA:n kloonaminen E. coliin.

Restriktioentsyymit ja T4 DNA-ligaasi saatiin New England Biolabsista ja niitä käytettiin valmistajien kuvaamasti. Pst1-, EcoRI- tai HindIII-restriktioentsyymeillä hajottamisen jälkeen DNA kuumennettiin 65 °C:een 10 minuutiksi ja uutettiin etanolilla. 10 µ:oon linearisoitua pBR322, pJKK3-1 tai pUN121 ja 20 g:aan fragmentoitua B. spp. kromosomaalista DNA:ta ligatoitiin 5 yksikköä DNA-ligaasia lopullisessa tilavuudessa 100 µl (16 h, 15 °C). Ligatoitua DNA:ta käytettiin kompetentin E. coli MC1000:n transformoimiseen tetrasykliiniresistentiksi oleellisesti siten kuin Mandel ja Higa (1970) ovat kuvanneet. (pBR322: 20 µg/ml, pUN 121, 7 µg/ml ja pJKK3-1: 10 µg/ml).

e) Sellulaasipositiivisten E. coli -kloonien havaitseminen

Levyillä olevan sellulaasiaktiivisuuden havaitseminen suoritettiin käyttämällä Teatherin ja Woodin (1982) tekniikan modifiointia. E. coli -klooneja viljeltiin yön yli kiinteässä NY-väliaineessa 37 °C:ssa. Solut lysoitiin päällystämällä pesäkkeet peiteagarilla, joka sisältää fosfaattipuskuria (100 mM, pH 7,0), agaria (0,7 %), CMC (0,2 %), SDS (0,25 mg/ml) ja kloramfenikolia (200 µ/ml) ja inkuboimalla yön yli 37 °C:ssa. Levyt peitettiin sitten Congo-punaisen vesiliuoksen alle (1 mg/ml) 15 minuutiksi ja pestiin seuraavaksi 1 M NaCl. Sellulaasipositiivisia pesäkkeitä ympäröi keltainen halo punaisella taustalla.

f) Maksisolut

Plasmidilla koodatut proteiinit analysoitiin käyttämällä Sancar et al.:n (1979) maksisolumenetelmää muuntaen seuraavasti. UV-säteilyttämisen jälkeen eloon jääneet solut tapettiin inkuboimalla soluja D-sykloseriinillä (150 µg/ml) 48 h 37 °C:ssa.

g) Geelielektroforeesi

(^{35}S)-L-metioniinilla merkityt maksisoluproteiinit ja muut proteiinit analysoitiin elektroforesoimalla 15 % (0,075 % bisakrylamidia) SDS-polyakryyliamidigeeleillä (Laemmli, 1970). Proteiinit tehtiin näkyviksi joko Coomassie Blue G 250:llä värjäämällä tai autoradiografialla.

DNA:n analysointi tehtiin elektroforesoimalla agaroosigeeleillä, joissa on Loeningin (1967) kuvaamaa puskuria.

h) Sellulaasiaktiivisuuden havaitseminen polyakryyliamidigeeleistä

SDS-polyakryyliamidigeelielektroforeesilla erotetuissa proteiininauhoissa olevan sellulaasiaktiivisuuden havaitseminen suoritettiin modifioidulla zymogrammitekniikalla, jonka on kuvannut Beguin (1983). Proteiinivalmisteet elektroforesoitiin SDS-polyakryyliamidigeelillä edellä kuvatusti ja geeliä pestiin 3 kertaa 30 min. fosfaattipuskurissa (100 mM, pH 7,0) kerrostettuna ohuelle (0,8 mm) agaroosigeelille (agaroosi, (1,8 %), CMC (0,2 %), fosfaattipuskuri (100mM, pH 7,0) ja sitä inkuboitiin 4 h 42 °C:ssa. Sellulaasiaktiivisuus tehtiin näkyväksi värjäämällä agaroosigeeliä 30 min. Congo-punaisen vesiliuoksessa (1 mg/ml), mitä seurasi geelin peseminen 1 M NaCl:ssa.

i) Kolorimetrinen sellulaasiaktiivisuus

Solu-uutteissa oleva sellulaasiaktiivisuus analysoitiin mittaamalla CMC:n hydrolyysin vapauttamien pelkistävien ryhmien lisäys (Miller, 1959). Sopiva määrä entsyymiä inkuboitiin 1,5 ml:lla 1 % CMC fosfaattipuskurissa (100 mM, pH 7,0). 30 min 55 °C:ssa inkuboinnin jälkeen lisättiin 1,5 ml dinitrosalisyylihapporeagenssia ja näytteitä keitettiin 5 min. Absorbanssi luettiin 550 nm:ssä vasten kohtia, jotka sisälsivät ekvivalenttisia määriä E. coli -resipienttikannasta peräisin olevaa uutetta. Yksi sellulaasiyksikkö vapautti 1 nM glukoosiekvivalenttia sekunnissa viitaten standardikäyrään.

j) DNA-sekvennointi

Yksipäiset merkityt DNA-fragmentit eristettiin ja sekvennoitiin kemiallisella modifiointimenetelmällä (Maxam ja Gilbert, 1980). Lohkontatuotteet erotettiin 8-%:isella tai 20-%:isella polyakryyliamidigeelillä ja autoradiografoitiin sen jälkeen -70 °C:ssa käyttämällä vahvistusseuloja.

Sanger et al.:n (1977) dideoksiribonukleotidimenetelmää käytettiin Endo3A:n sekvennointiin käyttäen pUC18:n johdannaisia (Yanisch-Perron, 1985).

k) Southern-analyysi

Bacillus spp.:stä (PL236) peräisin oleva kromosomaalinen DNA hajotettiin restriktioentsyymeillä tarvittavasti ja fraktonoitiin 1 % (paino/til.) agaroosigeeleillä. Sitten DNA blotattiin nitrosellulosasuodattimille. ^{32}p-merkityt DNA-koettimet (yhdistelmäplasmidit) valmistettiin lovitranslatoinnilla (Rigby et al., 1977) käyttäen ^{32}p-dctp:tä (Amersham) ja hybridointi suoritettiin kuten Southern on kuvannut (1975). Autoradiografia suoritettiin -70 °C:ssa käyttäen vahvistavia seuloja.

l) Kompetenttien B. subtilis -solujen transformointi

Kompetenttien B. subtilis -solujen valmistamiseen käytettiin Dubnaun ja Davidoff-Adelsonin (1971) menetelmän modifioitua versiota. 10 ml LB-väliainetta inokuloitiin kannalla aamulla. 7 tunnin kuluttua tehtiin sarjalaimennokset KM-1-väliaineeseen ja inkuboitiin yön yli 37 °C:ssa. Seuraavana aamuna toiseksi eniten laimennettu ja kasvava viljelmä laimennettiin kymmenen kertaa KM-2-väliaineeseen. Solut kerättiin 45-60 minuutin kuluttua inkuboinnista sentrifugoimalla 3 min 7K:ssa. Ne suspendoitiin uudelleen supernatantin 1/10 tilavuuteen ja 1/50 tilavuutta 86-%:ista glyserolia lisättiin. 0,1 ml määrät jäädytettiin nestemäisessä typessä ja varastoitiin -80 °C:ssa.

Näiden kompetenttien solujen transformointiin käytettiin Ehrlichin menetelmää (1986), jossa oli joitakin modifikaatioita. PTF valmistetaan ja esikuumennetaan 42 °C:een. 0,01

ml DNA:ta sijoitetaan reaktioastiaan ja kompetentit solut sulatetaan 42 °C:ssa. BTF lisätään soluihin suhteessa 1:1 ja 0,1 ml seosta lisätään DNA:han. Soluja ja DNA:ta inkuboidaan ravistamalla 20 minuuttia 37 °C:ssa. 30 minuutin kuluttua geenieksprессiosta tarvitaan 0,1 ml NY-väliainetta, jos halutaan kanamysiiniresistenssi. Solut levitetään lopuksi relevanteille levyille. Varastoliuosten reseptit ovat seuraavat:

Varastoliuokset B. subtilis kompetenttien solujen valmistamiseksi

Suolaseos:	10 mM $CaCl_2$, 1 mM $FeCl_3$ ja 1 mM $MnCl_2$
10 X MM:	20 g $(NH_4)_2SO_4$, 60 g KH_2PO_4, 140 g $K_2HPO_4 \cdot 3H_2O$ 1 l:ssa
KM-varasto:	100 ml 10 X MM, 10 ml 10-%:ista Na-sitraattia, 2 ml 1M $MgSO_4$ 1 l:ssa
KM1-varasto:	960 ml KM-varastoa, 20 ml 20-%:ista glukoosia, 1 ml 20-%:isia kasaminohappoja, 20 ml 5-%:ista hiivauutetta, 30 l 1 mM $MnCl_2$ 1 l:ssa
KM2-varasto:	960 ml KM-varastoa, 20 ml 20-%:ista glukoosia, 1 ml 20-%:isia kasaminohappoja, 20 ml 5-%:ista hiivauutetta, 1 ml suolaseosta, 1 ml 0,5 M $CaCl_2$ ja 0,8 ml 1 mM $MgCl_2$ 1 l:ssa
BCG:	100 ml 10 X MM, 10 ml Na-sitraattia, 2 ml 1M $MgSO_4$, 1 ml suolaseosta, 20 ml 20-%:ista glukoosia 1 l:ssa
BTF:	800 ml BCG, 100 ml 10mM EGTA, 100 mM $MgCl_2$ 1 l:ssa (Valmistettu juuri ennen käyttöä).

Selluloosaa hajottavien kantojen eristäminen ja karakterisoiminen

Selluloosaa hajottavia entsyymejä koodaavien geenien kloonaamiseksi tehtiin seuraava seulonta-ohjelma sopivien selluloosaa hydrolysoivien luovuttajalajien löytämiseksi. Selluloosaa hydrolysoivien mikro-organismien lähteenä käytettiin erilaisia kompostinäytteitä.

Kompostinäytteiden sarjalaimennokset päällystettiin ASC-agarväliaineessa ja selluloosa hydrolysoiva aktiivisuus havaittiin pesäkkeiden ympärille muodostuvista kirkastuvista alueista. Useita selluloosaa hajottavia bakteereita eristettiin. Eräs aktiivisimmista näistä isolaateista, joka identifioitiin Bacillus spp. PL236:ksi (NCIMB 40250) valikoitiin luovuttajakannaksi kloonauskokeisiin.

Laji hajotti nopeasti sekä hapolla turvotetun selluloosan että mikrokiteisen selluloosan (Avicel, Merck) agarväliaineessa. Optimikasvulämpötilassa (42 °C) kirkastuvat alueet ilmestyivät 2-3 päivässä.

Pienten määrien Bacillus spp PL236 -viljelmää lisääminen mikrokiteisten selluloosien suspensioihin saa selluloosakiteet paakkuntumaan yhteen tiukasti ja laskeutumaan.

Tämä viittaa siihen, että joko Bacillus spp. PL236 -soluilla tai solunulkoisilla sellulaaseilla on vahva affiniteetti selluloosasubstraatille ja ne sitovat selluloosakiteet tiukasti yhteen.

Rekombinanttikloonien seulontatestit

Ensimmäiset yritykset Bacillus spp. PL236:sta peräisin olevien sellulaasien kloonaamiseksi suunnaattin Bacillus spp. PL236:sta peräisin oleviin endosellulaasigeeneihin. Endosellulaasipositiivisten kloonien levyillä seulomisen helpottamiseksi sovellettiin testiä, jossa käytettään väriä Congo-punainen (Teather ja Wood, 1982), E. coliin.

Rekombinantti-E. coli -soluihin saadun sellulaasiaktiivisuuden havaitsemiseksi solut lysoitiin lisäämällä SDS:ää peiteagarille. Tällä peiteagarin modifioinnilla ei ollut mitään mitattavaa vaikutusta Bacillus spp. PL236:n CMC-hajottaviin entsyymeihin, kun tähän organismiin käytetiin modifioitua testiä.

Endosellulaasien molekyylikloonaus

Useita endosellulaasigeenejä kloonattiin selluloosaa hajottavasta Bacillus spp. -kannasta (PL236). Endosellulaasi 1 -klooni (PL 517) tehtiin ligatoimalla Pst1:llä osittain hajotettuun PL236-kromosomaaliseen DNA:han Pst1:llä lohkaistu pBR322 ja transformoimalla seuraavaksi kompetentit E. coli -solut.

Endosellulaasi 2 -klooni (PL382) tehtiin ligatoimalla Hind III:lla osittain hajotettuun PL236-kromosomaaliseen DNA:han HindIII:lla lohkaistu pJKK3-1 (E. coli / B. subtilis -sukkulavektori) ja transformoimalla seuraavaksi kompetentit E. coli -solut.

Endosellulaasi 3 -klooni (PL591) tehtiin ligatoimalla EcoRI:lla osittain hajotettuun PL236 DNA:han EcoRI:lla lohkaistu pUN121 ja transformoimalla seuraavaksi kompetentit E. coli -solut.

Endosellulaasi 4 -klooni (PL592) tehtiin ligatoimalla HindIII:lla osittain hajotettuun PL236 DNA:han HindIII:lla lohkaistu pUN121 ja transformoimalla seuraavaksi kompetentit E. coli -solut.

Sekä pBR322:a ja pUN121-plasmidivektoria käyttämällä johdetut transformantit seulottiin niiden tetrasykliiniresistenssin suhteen (pBR322: 20 g/ml, pUN121: 7 µg/ml ja pJKK3-1: 10 µg/ml).

Transformantit replikoitiin toiseen levysarjaan ja päällystettiin modifioidulla CMC-peiteagarilla. Levyjä inkuboitiin yön yli ja ne värjättiin Congo-punaisella edellä kuvatusti.

Eristämismenetelmän aikana havaittiin, että oli mahdollista havaita positiiviset kloonit lisäämättä soluja lysoivaa ainetta SDS-peiteagarille, joskin halon halkaisija oli huomattavasti pienempi.

Kloonattujen sellulaasien ekspressio ja karakterisointi

Sellulaasipositiivisista transformanteista peräisin olevat solut eristettiin ja analysoitiin restriktioentsyymeillä. Kuviossa 1 on esitetty kaikista neljästä kloonauskokeesta peräisin olevien tyypillisten plasmidien restriktioentsyymikartat.

Kloonattujen endoglukanaasien molekyylipainon määrittämiseen käytettiin tsymogrammitekniikka (Bequin, 1983). Tyypillisistä endoglukanaasiklooneista peräisin olevat kokonaisproteiinivalmisteet erotettiin SDS-polyakryyliamidigeelillä. Geelissä olevat proteiinit renaturoitiin pesemällä SDS pois ja ne replikoitiin CMC:tä sisältävään agaroosigeeliin. Renaturoidut proteiinit hajaantuvat tähän toimintageeliin ja proteiinit, joissa on endoglukanaasiaktiivisuutta, hydrolysoivat CMC:n geelissä. Endoglukanaasikaistat tehtiin sitten näkyväksi värjäämällä toimintageeli Congopunaisella edellä kuvatusti.

Kuviosta 1 ilmenee, että kloonatun DNA:n restriktiokartat (samoin kuin kloonatulla DNA:lla koodattujen endoglukanaasien molekyylipaino) ovat erilaisia jokaisessa neljässä erilaisessa kloonissa. Tämä osoittaa, että nämä neljä kloonia ovat vähintään neljä erilaista Bacillus spp.:stä (PL236) peräisin olevaa endosellulaasigeeniä.

Seuraavassa selostetaan näissä neljässä kloonissa olevien erilaisten endosellulaasien yksityiskohtainen analyysi.

Esimerkki 1

Endosellulaasi 1 (Endo1)

Endosellulaasin fyysinen kartoitus

Edellä kuvatussa kloonauskokeessa saatiin useita sellulaasipositiivisia E. coli -klooneja, jotka sisälsivät erilai-

sia PstI:llä osittain lohkaistun Bacillus spp. PL236 DNA:n fragmentteja. Kuviossa 2 esitetään joidenkin näiden kloonien restriktiokartat. Plasmidit sisälsivät poikkeuksetta kaksi Pst1-fragmenttia (1000 ja 1350 bp), mikä osoitti, että kummatkin ovat tarpeellisia sellulaasiaktiivisuutta omaavan polypeptidin synteesiin; Nukleotidisekvennointi osoitti myöhemmin, että pPL217:ssä ja pPL517:ssä läsnäoleva Pst1-fragmentti (kuvio 2) sisältää täydellisen sellulaasigeenin C-terminaalisen osan (tiedot on esitetty seuraavaksi). B. spp.-kromosomin tämän osan rakenne vahvistettiin Southern-analyysillä (tiedot on esitetty seuraavaksi).

Aktiivisuusmittaukset

pPL212:ta, pPL216:ta, pPL517:ää ja pBR322:ta sisältävistä E. coli -klooneista peräisin olevat uutteet valmistettiin yön yli olleista viljelmistä, joita oli kasvatettu NY-väliaineessa, jota oli täydennetty tetrasykliinillä. Solut konsentroitiin 10-kertaisesti 100 mM fosfaattipuskurissa, pH 7,0. DNaasi lisättiin ja solut rikottiin viemällä ne kahdesti French Pressin (12000 lb/in) läpi. Uutteita sentrifugoitiin 60 min 40000 X g:ssä ja supernatanttien sellulaasiaktiivisuus testattiin. E. colin MC1000 (pPL212), E. colin MC1000 (pPl517) ja E. colin MC1000 (pPL216) soluton uute sisälsi vastaavasti 8 yksikköä, 7,2 yksikköä ja 0,2 yksikköä sellulaasia/ml alkuperäistä viljelytilavuutta. (Kuvio 2). Supernatantista löydettiin 1/7 näissä viljelmissä olevasta kokonaissellulaasiaktiivisuudesta. E. coli MC1000 (pBR322) -kloonissa ei ollut sellulaasiaktiivisuutta.

Lajien, joissa on plasmideja pPL212 ja pPL517, uutteissa oleva korkea sellulaasiaktiivisuustaso johtuu todennäköisimmin sellulaasigeenin lisääntyneestä transkriptiosta, joka on peräisin ß-laktamaasipromoottorista pBR322:ssa. Todettiin, että sellulaasi ilmenee kuviossa 2 esitetyssä suunnassa.

CMC-liuoksen viskoositeetin ja pelkistävien sokerien vapautumisen mittaukset osoittivat, että kloonattu sellulaasi on endo-(1,4-)-ß-glukanaasi.

Maksisolu- ja tsymogrammianalyysi

Endoglukanaasiproteiinin molekyylipaino analysoitiin maksisolutekniikalla. Plasmidit pPL212, pPl216 ja pBR322 transformoitiin maksisolukantaan CSR603 (Sancar et al., 1979) plasmidilla koodattujen proteiinien analysoimiseksi. Plasmidi pPL212 tuotti kolme polypeptidiä, 75000 D, 65000 D ja 58000 D pBR322:n koodaamisen proteiinien lisäksi. Lukuunottamatta pBR322 proteiineja ei voitu havaita plasmidin pPL216 koodaamia proteiineja johtuen sellulaasigeenin alhaisesta ekspressiosta. pPL212:sta peräisin olevat kolme polypeptidiä testattiin sellulaasiaktiivisuuden osalta käyttäen geelitoistotekniikkaa (Bequin, 1983).

Sellulaasiaktiivisuuskaistojen ja polyakryyliamidigeelistä peräisin olevien autoradiogrammissa olevien kaistojen vertailu osoitti, että 75000 D ja vähintään toisessa 58000 ja 65000 D proteiinin kaistoissa on sellulaasiaktiivisuutta. Havaittiin myös, että Bacillus spp. PL236 -viljelmän supernatantti sisälsi vähintään kolme erilaista proteiinia, joissa on sellulaasiaktiivisuutta. Eräs näistä proteiineista migratoitui yhdessä maksisolussa syntetoidun 58000 D proteiinin kanssa. E. colista MC1000 (pPL212) peräisin olevista uutteista paljastui vain yksi aktiivinen kaista, joka migratoitui yhdessä maksisolusta E. coli CSR603 (pPL212) peräisin olevan 58000 D proteiinin kanssa. Kuten aiemmin mainittiin, seuraava sekvenssitieto osoitti, että plasmidit pPL212 ja pPL216 eivät sisältäneet kokonainaista Endo1-glukanaasigeeniä. Maksisolussa E. coli CSR603 (pPL212) havaittu 75000 D aktiivinen proteiini on siten fuusioproteiini, jossa 105 C-terminaalista aminohappoa ovat pBR322-sekvenssien koodaamia (kuvio 2). Tämä fuusioproteiini on ilmeisesti posttranslationaalisesti käsitelty ja päättyy 58000 D endoglukanaasiin. Solu-uutteet, jotka ovat peräisin E. coli MC1000:sta (pPL217) ja E. coli MC1000:sta (pPL517) ja jotka sisältävät täydellisen endoglukanaasigeenin, tuottivat aktiivisuuskaistat M_r suunnilleen 58000 D ja 75000 D. Plasmideista pPL217 ja pPL517 peräisin oleva 75000 D aktiivinen proteiini on siten Endo1-glukanaasigeenistä peräisin oleva "oikea" alkutranslaatiotuote, joka on ilmeisesti käsitelty

hitaammin kuin pPL212:sta peräisin oleva fuusioproteiini. Kuitenkin sekä pPL212:sta syntetoitu 75000 D fuusioproteiini että pPL517:stä ja pPL217:stä syntetoitu "oikea" 75000 D proteiini vaimennuskäsitellään 58000 D proteiiniksi, jonka sellulaasiaktiivisuus on suuri.

Sekvenssitiedoista voidaan myös ennustaa, että pPL216:sta ilmennetty endoglukanaasi syntetoi fuusioproteiinin, jossa 26 C-terminaaliset aminohapot ovat pBR322-sekvenssien koodaamia (kuvio 2). Tätä proteiinia ei havaittu maksisolutekniikalla ja tsymogrammit käyttäen PL216:sta peräisin olevaa uutetta paljastivat vain yhden 58000 D aktiivisen kaistan, joka todennäköisimmin edustaa käsiteltyä proteiinia. Kahden pPL212:sta ja pPL216:sta peräisin olevan fuusioproteiinin käsittely E. colissa, joissa on C-terminaalinen "hännän" kaksi eripituutta, saa siten aikaan kummassakin tapauksessa suunnilleen 58000 D aktiivisen proteiinin. Tämä osoittaa, että endosellulaasia käsitellään C-terminaalista, koska N-terminaalin käsittely saisi aikaan kaksi proteiinia, joiden M_r:n ero on suunnilleen 9000 D, joka olisi havaittu helposti tsymogrammeissa.

Siten on todennäköisimpää, että endoglukanaasi 1 (Endo1) syntetoituu proentsyyminä, jonka modifioi ainakin E. colissa (ja mahdollisesti myös B.spp.:ssä) asteittainen suunnilleen 150 C-terminaalisten aminohappotähteiden ja suunnilleen 30 N-terminaalisten aminohappotähteiden poistuminen, mikä vastaa signaalipeptidin poistamista. Endoglukanaasi näyttää modifioituvan oikein, minkä osoittaa seikka, että lopullinen käsittelytuote, 58000 D toimintakaista, joka on läsnä E. coli MC1000 (pPL212) -uutteessa, migratoituu yhdessä yhden Bacillus spp -viljelmien supernatantissa olevista endoglukanaaseista kanssa.

Lämpötilaoptimi ja -stabiilius

Uutteiden sellulaasiaktiivisuus mitattiin erilaisissa lämpötiloissa ja suurin E. colissa tuotettu Endo1:n aktiivisuus havaittiin 60°C:ssa (kuvio 3). Endoglukanaasin lämpöstabiilius testattiin inkuboimalla uutteita 50 °C:ssa, 55

°C:ssa ja 60 °C:ssa vaihdellen aikoja ja jäämäaktiivisuus mitattiin, kuten kohdassa Materiaalit ja menetelmät selostetaan. Vaikka suurin aktiivisuus havaittiin 60 °C:ssa pysyvässä 30 min inkubointiajassa, entsyymi inaktivoituu tässä lämpötilassa, kun $t_{1/2}$ on 1m2 h. 50 °C:ssa ja 55 °C:ssa ei havaittu inaktivoitumista 5 tunnin inkuboinnin jälkeen.

DNA-sekvenssi

Endosellulaasigeenin 1 (Endo1) nukelotidisekvenssi dedusoitiin plasmidista pPL517, joka sisältää suunnilleen 2850 bp Bacillus spp. PL236 DNA:n.

Sekvenssi määritettiin kemiallisella modifiointimenetelmällä (Maxam ja Gilbert, 1980) käyttäen osittaista restriktiokarttaa ja kuviossa 4 esitettyä sekvennointistrategiaa.

Sekvenssiluettelossa ID#1 esitetään täydellinen nukleotidisekvenssi. Tämän sekvenssin tietokoneanalyysi paljasti vain yhden avoimen lukujärjestyksen, joka on riittävän pitkä koodaamaan suunnilleen 75000 D proteiinia, joka havaittiin sellulaasipositiivisesta E. colista MC1000 (pPL517) peräisin olevien uutteiden maksisolu- ja tsymogrammianalyysilla. Tämä sekvenssi, joka alkaa nukleotidissa 677 ja päättyy nukleotidissä 2776, koodaa 700 aminohapon entsyymiä. DNA-sekvenssistä laskettu M_r oli 77006 D.

Avoimessa lukujärjestyksessä oli kolme potentiaalista aloituskodonia (ATG asemissa 677, 737 ja 749) mutta vain asemassa 677 olevaa ATG-kodonia edelsi ribosomin sitomiskohta (AAGGAGG) (Mclaughlin et al., 1981). Siten pääteltiin, että asemassa 677 oleva ATG-kodoni on oikea aloituskodoni.

Aloituskodonia seuraa aminohapposekvenssi, joka muistuttaa gram-positiivisista organismeista löydettyjä signaalisekvenssejä. Tällaiset signaalisekvenssit koostuvat suhteellisen lyhyestä hydrofiilisestä alueesta N-terminaalissa, mitä seuraa pidempi hydrofobisten tähteiden sekvenssi.

Käyttämällä Heijnen (1983) esittämää signaalisekvenssilohkontakohtamallia voidaan lohkontakohdan ennustaa olevan kahden ensimmäisen alaniinitähteen välissä sekvenssissä Asn-Ala-Ala-Ala. Siten signaalisekvenssi on 31 aminohappoa pitkä.

Vastakkaiseen suuntaan ja myötäiseen suuntaan olevat alueet eivät sisältäneet mitään merkittävää homologiaa B. subtiliksen sigma 43 promoottorin konsensussekvenssin kanssa eikä mitään terminaattorin kaltaisia sekvenssejä.

Southern-analyysi

Bacillus spp.:n (PL236) kromosomaalinen DNA hajotettiin restriktioentsyymeillä HindIII, PstI, EcoRI ja XhoI ja plasmideita pPL212 ja pPL509 käytettiin hydridoinnin koettimina. Plasmidi pPL212 sisältää kaksi PstI-fragmenttia (1350 bp ja 1000 bp) ja plasmidi pPL509 sisältää vain kokonaisen Endo1-geenin 500 bp PstI-fragmentin, jota edustaa plasmidi pPL517, joka sisältää kolme Bacillus spp. PL236 DNA:n PstI-fragmenttia (1350 bp, 1000 bp ja 500 bp). PL212-koetin tunnisti odotetut kaksi PstI-fragmenttia (1350 bp, 1000 bp) ja PL509 500 bp PstI-fraggmentin Bacillus spp. PL236 PstI-hajotteessa. Sekä pPL212- ja pPL509-koetin tunnistivat saman limittyvän EcoRI-fragmentin ja saman limittyvän HindIII-fragmentin Bacillus spp. PL226, EcoRI- ja HindIII-hajotteessa. Nämä tulokset osoittavat, että Bacillus spp. PL236 DNA-insertti kloonattiin pPL517:ssä deletoitumattomassa muodossa ja että pPL517:ssä olevat kolme PstI-fragmenttia ovat jatkuvia Bacillus spp. PL236 -kromosomissa.

Endo1-geenin ekspressio B. subtiliksessa

B. subtiliksessa suoritettaviin kloonauskokeisiin käytettiin pPL517:ää Endo1-geenin luovuttajana ja pDN2801:tä, jossa on voimakas Bacillus-promoottori P_m, käytettiin Bacillus-vektorina (kuvio 5).

Endo1-geenin sisältävä EagI-fragmentti ligatoitiin EagI-lohkaistuun pDN2801:een ja transformoitiin seuraavaksi kom-

petentteihin B. subtilis-soluihin (DN1885), jolloin saatiin CH7. Sen testaamiseksi onko käsitelty C-terminaalinen osa tarpeellisen Endo1-geenin ekspressiolle B. subtilis -soluissa, tehtiin konstruktio, jossa Endo1-geeni oli fuusioitu sisäisessä BglII-paikassa oleviin vektorisekvensseihin. Tämä fuusio korvaa C-terminaalisia 94 aminohappoja koodaavaan alueen vektorisekvenssien koodaamilla 55 "satunnais"-aminohapoilla.

Samanlaiset konstruktiot tehtiin E. coli -vektoreihin, vaikkakin fuusioituna erilaisiin vektorisekvensseihin, mikä sai aikaan aktiivisen periplasmisen endoglukanaasin E. colissa, joka käsiteltiin "oikealla tavalla". pPL517:stä peräisin olevaan BglII-fragmenttiin sisältyvä Endo1-geenin osa ligatoitiin BamHI-lohkaistuun pDN2801:een ja sitä seuraava kompetentteihin B. subtilis-soluihin (DN1885) transformointi tuotti lajin C14.

Kummassakin tapauksessa seulottiin transformanttien kloramfenikoliresistenssi sekä halutut plasmidikonstruktiot kannoissa CH7 ja CH14 niiden plasmidien restriktioanalyysillä. Plasmideissa pCH7 ja pCH14 olevat kaksi Endo1-geenin versiota traskriboituvat täten samasta promoottorista P_m. Plasmidien restriktiokartta on esitetty kuviossa 6.

Näihin kokeisiin käytetty B. subtilis DN1885 tuottaa oman endoglukanaasinsa, joka tuotti ilman muuta hieman tausta-aktiivisuutta. Endo1-geenituote vietiin viljelyväliaineeseen rekombinanttilajista CH7 ja viljelysupernantista mitattu aktiivisuus oli suunnilleen 20 kertaa suurempi kuin tausta-aktiivisuus (kuvio 6). Taustatason yläpuolella ei havaittu mitään solunulkoista aktiivisuutta yhdistelmäkannasta CH14, joka sisältää Endo1-geenin, jossa on substituoitu C-terminaali.

Kannoista CH7 ja CH14 peräisin olevat supernatantit ja CH7-soluista peräisin olevat solu-uutteet analysoitiin tsymogrammitekniikalla. Tsymogrammi ilmaisi suunnilleen 75000 D ja 58000 D aktiiviset proteiinikaistat CH7-solu-uutteista

ja vain yhden aktiivisen noin 58000 D proteiinikaistan CH7-viljelysupernatantista. Nämä kaistat vastaavat E. colissa havaittuja ja suunnilleen 90 aminohapon C-terminaalista käsittely näyttää tapahtuvan myös B. subtiliksessa.

Plasmidi pCH7 transformoitiin PL1801:een, joka on DN1885:n johdannainen, josta puuttuu kaksi pääeksoproteaasia (apr⁻, npr⁻), mikä tuottaaa kannan CH14. CH14:sta tuotettu Endo1-sellulaasi käsiteltiin "normaalisti", mikä osoittaa sen, että B. subtiliksesta peräisin olevat kaksi pääeksoproteaasia eivät ole vastuussa Endo1-sellulaasin C-terminaalisesta käsittelystä.

CH14-supernatantista havaittiin hyvin heikko suunnilleen 58000 D aktiivinen kaista, joka osoittaa sen, että manipuloitu geeni ekspressoituu ja käsitellään vähintään lähes samalla tavalla kuin luonnollinen geenituote. Muun muassa erittäin alhainen pCH14:sta ekspressoituminen ja se seikka, että geenit ekspressoituvat samoista ekspressiosignaaleista, voi osoittaa sen, että suunnilleen 90 C-terminaalista aminohappoa ovat tarpeellisa Endo1-geenituotteen viemiseen B. subtiliksesta pois.

Endo1-geenin E. colissa ekspressoitumisen optimointi

Endo1-geenin ekpression optimoiseksi Endo1-geeniin yhdistettiin voimakkaat E. coli -promoottorit P_R ja P_L, jotka ovat alunperin peräisin faagi-lambdasta (Remaut et al., 1981). Kumpikin promoottori voidaan tukahduttaa lambda cI857 -repressorilla, joka on lämpöepävakaa, ja tekee siten P_R- ja P_L-promoottorit lämpöindusoitaviksi soluissa, jotka tuottavat lambda cI857 -geenituotteen (Ptashne et al., 1982).

P_R-promoottori sisältyy ekpressioplasmidiin pPL170 yhdessä lambda cI857 -geenin kanssa (Kuvio 7; Jørgensen, 1983).

P_R-promoottori sijoitettiin vastavirtaan Endo1-geenistä ligatoimalla pPL170:stä peräisin oleva P_R:n sisältävä PvuI-SalI-fragmentti Endo1-geeniin, joka sisältää PvuI-SalI-frag-

mentin. Kompetentteihin MC1000-soluihin transformointi sai aikaan lajin TL05, joka sisältää plasmidin pTL05. Plasmidissa pTL05 on ß-laktamaasipromoottori deletoitunut, mikä saattaa Endo1-geenin P_R-promoottorin transkriptionaaliseen kontrolliin. Tässä suhteessa Endo1-geenin uskottiin sisältyvän pPL2129:stä peräisin olevaan BglII-fragmenttiin. Endo1-geenifuusio pPL212:ssa oleviin vektorisekvensseihin transferroitiin siten pTL05:een, mikä tuotti fuusioproteiinin, jossa 105 C-terminaalista aminohappoa ovat vektorisekvenssien koodittamia. Tämä fuusioproteiini käsitellään kuitenkin oikein, kuten aiemmin osoitettiin kannan PL212 yhteydessä. Sellulaasin tuotanto TL05:stä vaimentuu täydellisesti 28 °C:ssa ja indusoituu 42 °C:ssa.

Plasmidissa pPLc28 olevaan P_L-promoottoriin (Remaut et al., 1981) yhdistettiin Endo1-geeni ligatoimalla pPL212:sta peräisin oleva BglI-fragmentti BamHI-lohkaistuun pPLc28:aan. Transformoiminen kompetentteihin PL248-soluihin, joissa on lambda cI857-geeni yhteensopivassa pACYC177:ään perustuvassa plasmidissa (pNF2690), sai aikaan kannan LA03, joka sisältää plasmidin pLA03 (kuvio 7). Endo1-geeni on siten fuusioitu vektorisekvensseihin mutta pPLc28:ssa olevista eispesifioiduista DNA-sekvensseistä johtuen on saadun fuusioproteiinin pituus ja luonne tuntematon. Sellulaasin tuotanto LA03:sta vaimentui täydellisesti 28 °C:ssa ja indusoitui 40,5 °C:ssa.

Sellulaasin tuotanto LA03:sta ja TL05:stä arvioitiin erilaisissa lämpötiloissa.

LA03, TL05 ja PL212 viljeltiin yön yli 28 °C:ssa NY-väliaineessa, jota oli täydennetty sopivilla antibiooteilla (AMP + KAN, KAN ja TET vastaavasti). Jokaisen kannan yön yli -viljelmät laimennettiin 100-kertaisesti NY-väliaineeseen (AMP + KAN, KAN ja TET vastaavasti) ja laimennettuja viljelmiä kasvatettiin erilaisissa lämpötiloissa välillä 28 °C ja 42 °C. Solut kerättiin jokaisesta viljelmästä OD_{450} :ssä = 1 ja lysoitiin French Pressissä ja uutteiden aktiivisuus määritettiin aiemmin kuvatusti.

LA03, jossa esiintyi suurin sellulaasin tuotanto, ei kyennyt kasvamaan yli 40,5 °C:n lämpötiloissa. Koe toistettiin ilman antibioottivalintapainetta laimennetuissa viljelmissä. Saatiin samanlaiset tulokset, mutta LA03 kasvoi hyvin hitaasti lämpötiloissa yli 40,5 °C. Tätä kasvua seurasi kuitenkin merkittävä plasmidin pLA03 kato. Mitään merkittävää pLA03:n katoa ei havaittu lämpötiloissa 40,5 °C saakka tai pPL212:n ja pTL05:n merktittävää katoa ei havaittu missään lämpötiloissa.

Endo1-geenin B. subtiliksessa ekspressoitumisen optimointi

Endo1-geenin ekpression optimoiseksi B. subtiliksessa Endo1-geeni fuusioitiin ekspressiosignaaleihin (promoottori, ribosomin sitomiskohta ja signaalisekvenssi) alfa-amylaasista B. licheniformiksesta, joka ekpressoituu suurina määrina B. subtiliksesta.

pPL1759 sisältää promoottorin, ribosomin sitomispaikan ja suurimman osan B. licheniformis alfa-amylaasin signaalipeptidistä (Stephens et al., 1984). Tämän alueen myötävirtainen puoli päättyy PstI-kohtaan, mitä seuraa jälleen polyliittäjä (kuvio 8). pPL1759:n polyliittäjässä olevan PstI- ja SalI-kohdan väliin insertoitiin synteettinen DNA-fragmentti, joka koostuu kahdesta täydentävästä oligonukleotidistä, jotka luovat PstI- ja SalI -"tahmaiset" päät. Saadussa plasmidissa pCH52 muodostaa uudelleen alfa-amylaasin signaalipeptidin puuttuvan osan ja koodaa lisäksi kypsän Endo1-sellulaasin ensimmäiset 14 N-terminaalista aminohappoa (kuvio 8). Täten liittäjä luo hybridisignaalin lohkaisukohdan alfa-amylaasin ja sellulaasin väliin. Odotettu lohkaisukohta on esitetty kuviossa 8. Endo1-geeni leikattiin pPL417:stä ilman EagI-fragmentissa olevaa promoottoria ja liitettiin uniikkiin EagI-kohtaan pCH52:ssa. Plasmidi, johon Endo1-geeni liitettiin oikeassa orientaatiossa, nimettiin pCH54:ksi.

pCH54 sisältää kaksi suoraa toistuvaa 45 bp sekvenssiä (esim. kypsän Endo1-geenin 45 terminaaliset emäsparia), jotka voivat rekombinoitua deletoiden kaksi niiden välissä

olevaa aluetta (Ehrlich et al., 1986). Tämä rekombinointitapahtuma ilmenee kuitenkin hyvin alhaisella taajuudella, kun toisto niin pieni kuin 45 bp. Rekombinoidun plasmidin määrän rikastamiseksi pCH54:n plasmidivalmiste leikattiin rikastusrestriktioentsyymillä KpnI. pCH54 sisältää uniikin KpnI-kohdan kahden suoran toiston välissä ja vain ei-rekombinanttiset plasmidit leikataan KpnI:llä, kun taas rekombinanttiplasmidit jäävät sirkulaarisiksi. Kun B. subtilis (DN1885) transformoitiin uudelleen tällä seoksella, transformantit sisälsivät pääasiassa (90 %) rekombinanttiplasmideja, koska B. subtilis-kompetentit solut eivät transformoidu linearisoidulla plasmidi-DNA:lla. Rekombinanttiplasmidi nimettiin pCH57:ksi ja se sisältyy kantaan CH57. Rakenne vahvistettiin DNA-sekvennoinnilla. Tässä konstruktiossa pCH57 on Endo1-geeni siten täydellisesti fuusioituneena alfa-amylaasiekspressiosignaaleihin.

Endoglukanaasi tuotetaan solunulkoisesti B. subtilis -kannasta CH57, mikä osoittaa, että hybridisignaalikohta toimii. Eritetty Endo1-glukanaasi on käsitellään odotettuun M_r:ään, nimittäin 58000 D.

Endo1-sellulaasin tuotanto CH57:stä arvioitiin kahdessa erilaisessa väliaineessa, NY (yön yli 37 °C:ssa) ja BPX (7 päivää 37 °C:ssa). BPX-väliaine on erittäin rikas väliaine, jossa ravintoaineet vapautuvat hitaasti, jolloin solut pidetään siten varhaisessa stationaarisessa faasissa useita päiviä fermentoinnin aikana. Alfa-amylaasiekspressiosignaalit toimivat erityisen hyvin tässä väliaineessa. Tulokset ilmenevät seuraavasta taulukosta 1.

Taulukko 1

Kanta	Plasmidi	U/ml NY	U/ml BPX
CH7	pCH7	6,5	55,0
CH57	pCH57	13,0	325,0
DN1885	---	0,3	40,0

Viljelysupernatantin analyysi (BPX-väliaine) PAGElla paljasti dominantin (90 %) endoglukanaasikaistan, joka vastaa suunnilleen 0,5 g/l endoglukanaasikonsentraatiota supernatantissa.

Esimerkki 2

Endosellulaasi 2 (Endo2)

DNA-sekvenssi

Endosellulaasin 2 (Endo2) nukleotidisekvenssi dedusoitiin plasmidista pL382, joka on kuvattu edellä. Plasmidi sisältää suunnilleen 2500 bp Bacillus spp. PL236 -DNA:n. Sekvenssi määritettiin kemiallisella modifiointimenetelmällä (Maxam ja Gilbert, 1980) käyttäen osittaista restriktiokarttaa ja sekvennointistrategiaa, joka on kuvattu kuviossa 9.

Täydellinen nukleotidisekvenssi on esitetty sekvenssiluettelossa ID#3. Tämän sekvenssin tietokoneanalyysi paljasti vain yhden avoimen lukujärjestyksen, joka on riittävän pitkä koodaamaan suunnilleen 56000 D proteiinia, joka havaittiin sellulaasipositiivisesta kloonista E. coli MC1000 (pPL382) peräisin olevan uutteen tsymogrammianalyysissä. Tämä sekvenssi alkaa asemassa 172 ja päättyy asemassa 1869 ja koodaa entsyymiä, joka koostuu 566 aminohaposta. Laskettu M_r on 62551 D, joka on hieman suurempi kuin tsymogrammianalyysillä määritetty 56000 D:n M_r. Tämä ero voi johtua tsymogrammmianalyysin epätarkkuudesta tai posttranslationaalisesta käsittelystä signaalipeptidin odotetun käsittelyn jälkeen. Asemassa 172 oleva ATG-kodoni valittiin, koska se oli ainoa aloituskodoni, jonka lukujärjestys on avoin ja jota edelsi ribosonin sitomiskohta AAGGAGG (Mclaughlin et al., 1981).

Tätä aloituskodonia seurasi signaalisekvenssin kaltainen sekvenssi ja käyttämällä Heijnen (1983) ehdottamaa signaalisekvenssin lohkaisumallia voitiin lohkaisukohdan ennustaa olevan sekvenssin Leu-Ala-Ala-Ala keskellä olevassa kahdessa alaniinitähteessä. Endo12:n signaalisekvenssi on täten 30 aminohappoa pitkä.

Alue, joka on vastavirtaan avoimesta lukujärjestyksesta, sisälsi sekvenssin, joka on homologinen B. subtiliksen sigma 43 -tyyppisten promoottorien kanssa (Johnson et al., 1983) asemassa 46-75.

Tämä sekvenssi koostuu TTTACA:sta -35-alueena ja TATTAT:sta -10-alueena; näitä kahta aluetta erottaa 18 nukleotidiä.

13 bp palindrominen toistuva sekvenssi löydettiin myötävirtaan lopetuskodonista asemassa 1956-1981, joka näyttää muistuttavan rho-itsenäistä terminaattoria (Rosenberg ja Court, 1979).

Southern-analyysi

Bacillus spp.:n (PL236) kromosomaalinen DNA hajotettiin HindIII:lla, PstI:llä, EcoRI:lla ja XhoI:llä ja plasmidia pPL382 käytettiin hydridoinnin koettimina. Saatu hybrodointikaavio vahvisti, että Bacillus spp. PL236 -DNA kloonattiin deletoitumattomassa muodossa, että kaksi pPL382:sta peräisin olevaa HindIII-fragmenttia oli jatkuvia Bacillus spp. PL236 -kromosomissa ja että Endo2-geeni erosi muista kloonatuista endoglukanaaseista.

Endo2-geenin ekspressio B. subtiliksessa

Plasmidi pPL382 transformoitiin B. subtilis DN1885:een kypsän Endo2-tuotteen saamiseksi. B. subtilis DN1885:tä (pPL382) viljeltiin aerobisesti 640 ml:ssa LB-väliainetta, joka sisältää 10 µg/ml tetrasykliiniä, 30 tuntia. Supernatantti konsentroitiin presipitoimalla 24 tuntia $(NH_4)_2SO_4$:lla 70 % kyllästyksessä. 5 tunnin 100 mM Tris-HCl, pH 7, vasten dialysointia jälkeen konsentroitu supernatantti kuumennettiin 55 °C:een 15 minuutiksi. Denaturoitu proteiini poistettiin sentrifugoimalla ja liukenevat proteiinit presipitoitiin seuraavaksi $(NH_4)_2SO_4$:llä 70 % kyllästyksessä. Uudelleen liukoisaksi tehdyt proteiinit dialysoitiin vasten 100 mM Tris-HCl, pH 7, lopullisella 2 ml tilavuudella ja ne levitettiin 80 cm x 1 cm -geelisuodatuspylvääseen, joka sisälsi Ultrogel AcA 44:ää (LKB). Aktiiviset fraktiot yhdistettiin, konsentroitiin ja levitettiin

SDS:ää sisältävälle polyakryyliamidigeelille, jossa endoglukanaasi esiintyi yhtenä 56 kDa:n kaistana. Menetelmän aktiivisuussaanto oli suunnilleen 10 %.

Endoglukanaasi migratoituu yhdessä MC1000 (pPL382) -uutteista saadun endoglukanaasin kanssa, mikä havaittiin tsymogrammianalyysillä DN1885 (pPL382):n supernatantissa. DN1885:n (pPL382) endoglukanaasiaktiivisuus on noin 25 kertaa kannan DN1885 aktiivisuudesta, kun kasvatetaan NY-väliaineessa.

Esimerkki 3
Endosellulaasi 3
Endosellulaasikloonin 3 fyysinen kartoitus

Endosellulaasiklooni 3 esitetään plasmidilla pPL591, joka sisältää Bacillus spp. PL236 -DNA:n 11000 bp EcoRI-fragmentin. Tämän plasmidin osittainen restriktiokartaa on esitetty kuviossa 10. E. coli MC1000 (pPL591):stä peräisin olevin uutteiden tsymogrammianalyysi osoittaa, että insertti tuottaa neljä proteiinia, joissa on sellulaasiaktiivisuutta. Näiden proteiinien M_r on suunnilleen 60000 D, 56000 D, 45000 D ja 30000 D (kuvio 10).

Southern-analyysi osoitti, että pPL591:stä peräisin oleva EcoRI-fragmentti kloonattiin deletoitumattomassa muodossa Bacillus spp. PL236 -kromosomista ja se osoittaa, että DNA ei sisältänyt Endo1- ja Endo2-geenejä koodaavia DNA-sekvenssejä.

Sen analysoimiseksi ovatko nämä proteiinit posttranslationaalisia käsittelytuotteita yhdestä tai useammasta sellulaasigeenistä, tehtiin deleetioplasmideita käyttämällä restriktioentsyymejä HindIII ja SmaI. 4800 bp SmaI-fragmentin deletointi, joka sai aikaan plasmidin pPL538, ei eliminoinut mitään neljästä tsymogrammissa olevista sellulaasikaistoista. pPL591:ssä oleva B. spp. -DNA-insertti sisälsi 5 HindIII-kohtaa. pPL591:n hajottaminen HindIII:lla tuotti täten 6 fragmenttia 6600 bp (vektorifragmentti), 3200 bp, 1700 bp, 1550 bp, 1350 bp ja 900 bp, jolloin 1700

bp fragmentin alkuperä oli pUN121-vektoriplasmidissa olevasta HindIII-kohdasta (kuvio 11).

Kaikkien HindIII-fragmenttien paitsi 1700 bp fragmentin (insertoitu uudelleen vastakkaiseen suuntaan) eliminoiminen sai aikaan plasmidin pPL540. Bacillus spp. PL 236 -DNA:n tähteen poistaminen eliminoimalla pPL540:stä EcoRI-fragmentti sai aikaan plasmidin pPL587. Sekä E. coli MC1000 (pPL540) että E. coli MC1000 (pPL587) olivat sellulaasipositiivisia ja näistä klooneista peräisin olevien uutteiden analysointi paljasti vain 45000 D ja 30000 D proteiinit. Bacillus spp. PL236 -DNA:n (suunnilleen 1500 bp) proteiinia koodaava kyky on suunnilleen 55000 D, joka on liian pieni sisältääkseen kaksi endoglukanaasigeenia 45000 D ja 30000 D. Tsymogrammissa oleva 30000 D proteiini on täten ilmeisemmin tulosta 45000 D proteiinin epäkypsästä posttranslationaalisesta käsittelystä. 45000 D proteiinia koodaava sellulaasigeeni nimettiin Endo3A:ksi. Endo3A-geeni kloonattiin kummassakin suunnassa plasmideihin pPL587 ja pPL538 ja tuotti samat kaksi proteiinia ja eliminoi siten proteiinin muutoksen ollen fuusioproteiini.

1550 bp, 1350 bp ja 900 bp fragmenttien eliminointi tuotti plasmidin pPL542. E. colista MC1000 (pL542) peräisin olevista uutteista paljastui suunnilleen 30000 D, 45000 D, 49000 D ja 56000 D sellulaasipositiivisia proteiineja (kuvio 10). Näistä alustavista tuloksista postuloitiin kahden lisäendoglukanaasigeenin läsnäolo alkuperäisessä insertissä pPL591:ssä ja pPL538:ssa. Lisäendoglukanaasigeenit nimettiin Endo3A:ksi ja Endo3C:ksi. Niiden oletettu asema Bacillus spp. PL236 -DNA:ssa on esitetty kuviossa 11. Postuloitu malli perustuu olettamukseen, että pPL538:sta ja pPL591:stä tehty 60000 D proteiini muuttuu 49000 D typistetyksi fuusioproteiiniksi, joka on tehty pPL542:sta, jolloin 1350 bp HindIII-fragmentti on deletoitu.

Endo3A:n DNA-sekvenssi

Endo3-geenin DNA-sekvenssi dedusoitiin plasmidista pPL450, joka sisältää suunnilleen 1500 bp Bacillus spp. PL236

-DNA:n, käyttämällä dideoksiketjun terminointimenetelmää. Geeni sijoitettiin pUC18:aan kummassakin orientaatiossa ja lukuisia deleetioita konstruoitiin. Käytettiin standardialoittajia paitsi yhtä synteettistä oligonukleotidia, jota käytettiin sekvennoimaan aluetta, jossa ei ole käytännöllisiä restriktiokohtia.

Geenin C-terminaalinen osa dedusoitiin plasmidista pPL538. Osittainen DNA-sekvenssi on esitetty sekvenssiluettelossa ID#6. Sekvenssissä on avoin lukujärjestys, joka koodaa proteiinia, jonka M_r on noin 62000 D, joka on yhtäpitävä tsymogrammeissa havaitun 60 kD proteiinin kanssa. ATG-aloituskodonia (asema 30) edeltää tyypillinen ribosomin sitomiskohta (Mclaughlin et al., 1981). Aloituskodonia seuraa tyypillinen gram-positiivinen signaalisekvenssi ja käyttämällä signaalisekvenssin lohkaisumenetelmää (Heijne, 1983) tulee esiin 36 aminohapon signaalisekvenssi.

Esimerkki 4

Endosellulaasi 4

Tsymogrammianalyysi

Endosellulaasiklooni no 4:ää esittää plasmidi pPL592, joka sisältää suunnilleen 14000 bp Bacillus spp. PL236 -DNA:n. Osittainen restriktiokartta on esitetty kuviossa 1.

E. coli MC1000 :sta (pPL592) peräisin olevien uutteiden tsymogrammianalyysi paljasti kolme sellulaasiaktiivista proteiinia, joiden M_r-arvot ovat suunnilleen 92000 D, 74000 D ja 71000 D. Lisäanalyysi on tarpeen sen määrittämiseksi, ovatko nämä proteiinit yhden vai useamman sellulaasigeenin koodaamia.

Southern-analyysi vahvisti kloonatun DNA:n alkuperän Bacillus spp. PL236 -kromosomissa ja osoittaa että kloonattu DNA ei esiinny muissa endoglukanaasiklooneissa.

Esimerkki 5

A. Endo1, joka on kloonattu ja ekspressoitu Bacillus subtiliksessa

Agarviiste inokuloitiin B. subtilis -kannalla CH 57 ja inkuboitiin 20 tuntia 37 °C:ssa. 10 ml 0,9 % NaCl-liuosta lisättiin testiputkeen, jota ravistettiin solujen suspendoimiseksi. Solususpensiota käytettiin 2 l fermentterin inokuloimiseen.

Seuraavia parametrejä käytettiin fermennoinnin suorittamiseen.

Lämpötila: 37 °C
Ilmastus: 1,1 l/min
Sekoitus: 1100 rpm
Fermenttori: 2 litran malli, jonka työskentelytilavuus on 1,5 l.

pH pidettiin välillä 6,2-7,2 ensimmäiset 40 fermentointituntia. Sen jälkeen pH pidettiin välillä 6,7-7,2. pH:ta pidettiin tällä alueella annostamalla NH_3:a ja H_3PO_4:ä.

Glukoosiliuoksen annostaminen aloitettiin 40 tunnin kuluttua virtausnopeudessa 3,7 ml/tunti.

Substraatti

Perunatärkkelys, joka on hajotettu Termamyl*illä	50 g
Soijapapujauho	110 g
Maissinliotusneste	16,5 g
Alburex (perunaproteiini)	27,5 g
$(NH_4)_2SO_4$	2,2 g
KH_2PO_4	1,2 g
$Na_2HPO_4.2H_2O$	5,9 g

Vettä lisätään 1100 ml saakka

*Termanyl on kaupallinen B. licheniformis alfa-amylaasi, joka on saatavissa Novo Nordisk A/S:stä.

Glukoosiliuos

Glukoosi.H_2O	600 g
Sitruunahappo	0,6 g

Vettä lisätään 1000 ml saakka.

Fermentointi pysäytettiin 166 fermentointitunnin kuluttua OD_{650}-arvossa 122. Silloin fermenttoriin lisättiin 1200 ml fermentointilientä.

Fermentointiliemi sentrifugoitiin ja solunulkoinen tilavuus oli 400 ml, joka sisälsi 40 CMC-endoaasiyksikköä millilitraa kohti (160000 CMC-endoaasiyksikköä kaikkiaan). Viljelyväliainetta käsiteltiin edelleen suodattamalla ja laimentamalla, mitä seurasi konsentrointi Amicon ultrasuodatusmoduulissa leikkaus 10000 MW:ssä. Konsentroitu entsyymi jäädytettiin.

Puolet jäädytetystä nesteestä sulatettiin ja laimennettiin deionoidulla vedellä ja konsentroitiin sitten vielä kerran Amicon-ultrasuodatusmoduulissa. Kokonaissaanto oli 4471 CMC-endoaasiyksikköä (8000:sta endoaasiyksiköstä).

1050 ml:n kokonaistilavuus alistettiin ioninvaihtokromatografiaan pH:ssa 7. Entsyymi sidottiin DEAE-Sepharyl-anioninvaihtopylvääseen (300 ml tilavuus) pH:ssa 7 (50 mM tris-HCl). Endo1-entsyymi eluoitui pH:ssa 7 0,3 M NaCl:lla.

Puhdistetun entsyymin molekyylipaino on 58000 D SDS-PAGEssa. pH on 4,0. Sen aktiivisuus on 30 CMC-endoaasiyksikköä per mg proteiinia.

Proteiinimääritys perustuu DNA-sekvenssistä dedusoidun entsyymin aminohappokoostumukseen: 13 tryptofaani, 30 tyrosiini ja molekyylipaino 57566 D. Ekstinktiokerroin lasketaan seuraavan kaavan avulla:

$$(13 \times 5559 + 30 \times 1197) / 57566 = 1,88$$

Puhdistetun entsyymin endoglukanaasiaktiivisuus on 57 CMC-yksikköä per ml ja absorbanssi 280 nm:ssä 3,6. (57 x 1,88)/3,6 = 30 CMC-endoaasiyksikköä per mg proteiinia.

Endo1:n stabiilius pesuaineissa

Käytettiin seuraavaa 4 pesuainekoostumusta:

1. USA nestemäinen pesuaine: 2 grammaa per litra 6°-kovuista vettä (1 osa vesijohtovetttä 2 osaan deinoitua vettä). pH mitattiin 7,29:ksi.

2. USA Heavy Duty Powder -pesuaine: 0,9 grammaa per litra 6°-kovuista vettä. pH mitattiin 9,2:ksi.

3. Heavy Duty Powder -pesuaine (2), jossa on valkaisuainetta ja aktivaattoria: 0,12 grammaa/l Na-perboraatti-tetrahydraattia ja 0,088 grammaa/l NOBS. pH mitattiin 9,2:ksi.

4. European Heavy Duty Powder -pesuaine, jossa on valkaisuainetta ja aktivaattoria (erä DR 8806 Europe). 5 grammaa/litrassa 6°-kovuista vettä.

Celluzyme™ä (erä CAX 007 raaka entsyymi, jossa on sellulaasia ja muita entsyymeitä) 2353 CMC-endoaasiyksikköä grammaa kohti verrattiin Endo1 30000 CMC-endoaasiyksikköön grammaa kohti.

Entsyymit laimennettiin 3 CMC-endoaasiyksikköön per ml kaikissa 4 pesuaineliuoksessa: Laimentamisen jälkeinen endoglukanaasiaktiivisuus mitattiin edellä kuvatusti (määrittämällä CMC:n viskositeetin lasku). Endoglukanaasiaktiivisuus 60 min 40 °C:ssa inkuboinnin jälkeen mitattiin ja sitä verrattiin alkuaktiivisuuteen. Saatiin seuraavat tulokset:

Pesuaineliuos	1	2	3	4
Celluzyme™	89 %	75 %	66 %	75 %
Endo1	106 %	90 %	97 %	98 %

Standardipoikkeama on 10 %.

Taulukosta ilmenee, että Endo1 on vakaampi näissä pesuaineissa pH:ssa 9-10 verrattuna Celluzyme™:iin.

Esimerkki 6

Täysipituisen (~75 kD) Endo1:n valmistaminen B. subtiliksessa

10 ml:aan LB-väliainetta, joka sisältää 1 mM $CuCl_2$ ja 10 µg/ml kloramfenikolia, inokuloitiin 10 ml B. subtiliksen DN969 yön yli -viljelmää (Diderichsen et al., J. Bacteriol. 172(8), 1990, s. 4315-4321), joka sisältää plasmidin pCH7 (kuvattu edellä esimerkissä 1), jaettuna 10 steriiliin 2 l pulloon ja inkuboituna ravistamalla voimakkaasti 36 tuntia 37 °C:ssa. Viljelmää sentrifugoitiin 10 min 10000xg:ssä ja 4 °C:ssa, minkä jälkeen supernatanttiin lisättiin EDTAa, pH 8, loppukonsentraatioon 5 mM.

25 g Avicel PH-105, joka oli hydrattu etanolissa ja pesty tislatulla vedellä, lisättiin supernatanttiin, joka jätettiin seisomaan sekoittaen hieman 2 tuntia 4 °C:ssa. Supernatantti/Avicel-seosta sentrifugoitiin 10 min 10000xg:ssä ja 4 °C:ssa. Selkiytynyt supernatantti dekantoitiin pois välittömästi sen jälkeen, kun roottori oli pysähtynyt.

Avicel/entsyymikakku suspendoitiin uudelleen ja pestiin 200 ml:ssa 1 mM EDTAa ja seosta sentrifugoitiin 1 min 10000 xg:ssä. Tämä menettely toistettiin kahdesti. Sitten Avicel/entsyymikakku suspendoitiin uudelleen 150 ml:aan (1 % trietyyliamiinia ja 1 mM EDTAa) ja sen annettiin seisoa voimakkaasti sekoittaen 1 tunti 4 °C:ssa. Seosta sentrifugoitiin 1 min 10000 xg:ssä 1 ja 4 °C:ssa. Supernatantti säilytettiin. Tämä menettely toistettiin kahdesti.

Entsyymin ja trietyyliamiinin seos (noin 300 ml) haihdutettiin tyhjössä 100 ml:ksi. Liuoksen lämpötilan ei annettu ylittää 10 °C. pH säädettiin 7:ään lisäämällä 1 M HCl ja liuos jäädytettiin -70 °C:ssa.

Proteiinin määrä 100 ml:ssa entsyymiliuosta määritettiin 40 mg:ksi Bradfordin reagenssin avulla (saatavissa BioRadista) käytttäen naudan seerumin albumiinia standardina.

~75 kD Endo1:n karakterisointi

Edellä saadun entsyymin puhtaus oli noin 90 %. Entsyymin molekyylipainoksi havaittiin 75 kD SDS-PAGElla.

Immuunireaktiivisuuskokeissa (suoritettiin raketti-immunolektroforeesilla agaroosigeelissä, kuten N. Axelsen et al. ovat kuvanneet, kappaleessa 2, A Manual of Quantitative Immunoelectroforesis, Blackwell Scientific Publ. 1973) oli ~75 kD Endo1 immunoreaktiivinen monospesifisen kanin vasta-aineen kanssa, joka oli herätetty Endo1:n ydinaluetta (~58 kD muoto) vastaan.

Endo1:n ~75 kD ja ~58 kD (käsitellyt) muodot testattiin kyvyltään sitoutua selluloosaan (Avicel). 10 µg sekä ~75 kD Endo1:tä että ~58 kD Endo1:tä lisättiin 50 µl:aan selluloosasuspensiota (0,1 % paino/paino Avicel, 5 mM EDTA, pH 8,0). Suspensioita ravistettiin 30 minuuttia ja Avicelia kerättiin 10000 rpm:ssä 2 min. Supernatanttiin jääneeen Endo1-sellulaasin määrä analysoitiin SDS-PAGElla. Yli 95 % ~75 kD Endo1:stä oli sitoutunut selluloosaan, kun taas alle 5 % ~58 kD Endo1:stä oli sitoutunut samalla tavalla. Tämä osoittaa, että ~75 kD Endo1-sellulaasin C-terminaalinen osa käsittää selloloosaan sitoutuvan alueen.

Laaja aminohapposekvenssihomologia havaittiin Endo1-sellulaasin tämän alueen (aminohaposta 554 aminohappoon 700) ja muiden sellulaasien välillä, esim. Bacillus subtiliksesta peräisin oleva endosellulaasi (Nakamura et al., 1987), Caldocellum saccharolyticumista peräisin olevan bifunktionaalisen sellulaasin keskiosa (D.J. Saul et al., Nucl. Acids Res. 17, 1988, s. 439), Clostridium stercorariumista peräisin olevat kaksi endosellulaasia (W. Schwarz et al., Biotech. Lett. 11, 1989, s. 461-466.

~75 kD Endo1:n värejä kirkastava vaikutus

~75 kD Endo1:n värejä kirkastava vaikutus määritettiin altistamalla esipesty kulunut tesktiilipinta entsyymille ja määrittämällä sitten pintavärin kirkkaus verrattuna entsyymillä käsittelemättömien tekstiilien pintavärin kirkkauteen.

Mustat 100 % puuvillaa oleva tilkut (15 x10 cm) esipestiin ja rumpukuivattiin seuraavissa olosuhteissa:

Pesuaine:	Keminus (saatavissa Irma A/S:stä), 1,5 g/l
Lämpötila:	70 °C
Pesuaika:	60 minuuttia
Kuivausaika:	30 minuuttia
esipesu/kuivaus-käsittelyjen määrä:	15.

Tilkut esipestiin konventionaalisessa pesukoneessa (Miele Deluxe Electronic W761). Jokaisen pesun jälkeen tilkut kuivattiin rumpukuivaajassa. Esipesun/pesun visuaalinen vaikutus oli se, että pintaväri muuttui harmahtavaksi johtuen vahingoittuneista selluloosakuiduista, jotka aiheuttivat kuluneen näön.

Esipesun jälkeen tilkut pestiin Terg-O-tometerissä (päältä ladattava pienoispesukone) seuraavissa olosuhteissa:

Nestetilavuus:	800 ml
Sekoitus:	100 liikettä minuutissa
Pesuaine:	Standardipesuaine, 5 g/l
Pesuaika:	30 min
Pesulämpötila:	40 °C
Tilkkujen määrä:	2
~ 75 kD Endo1-annostus:	0-60 CMC-endoaasiyksikköä/l
pH:	7,0
Käsittelyjen määrä:	3

Standardipesuaine:

LAS NANSA 1169/P:	10 %
AE Berol 160:	15 %
Etanoli, 96 %:	10 %
TEA:	5 %
Vesi:	60 %

Tilkut huuhdeltiin jokaisen pesun jälkeen vesijohtovedessä ja kuivattiin huoneenlämpötilassa.

Tilkkujen pintaväri analysoitiin mittaamalla heijastunut valo. Pinnalle suunnattiin valkoinen valo ja heijastuminen/remissio mitattiin 16 aallonpituudessa (400 nm - 700 nm). Mittausten tulokset muokattiin ("Elrepro 2000" -laitteessa, joka on saatavissa Datacolorista, Sveitsi) Hunter-koordinaatteihin, joista L-koordinaatti on harmaan skaalan arvot. Jokainen tilkku analysoitiin kahdesti kummaltakin puolelta ja jäljempänä annetut tulokset ovat kahden tilkun mittausten keskiarvo samasta käsittelystä. Taulokossa valkoinen on L =100 ja musta on L = 0.

Annos	0 CMC-endoaasi/l	60 CMC-endoaasi/l
L	16,78	15,50
S.D.	0,08	0,03
Delta L	-	1,28

Vertailutulokset saatiin Celluzyme™:llä (erä PPC 2174, joka sisältää Humicola insolensistä, DSM 1800 peräisin olevien entsyymien seoksen)

Annos (CMC-endoaasi/l	0	15	30	60
Delta L	-	1,02	1,48	1,90

Viitteet:

Beguin, P.: Detection of cellulase activity in polyacrylamide gels using Congo red-stained agar replicas. Anal. Biochem. 131 (1983) 333-336.

Bolivar, F., Rodriguez, R.L., Green, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W. Crosa, J.H., and Falkow, S.: Construction and characterization of new cloning vehicles II. A multipurpose cloning system. Gene 2 (1977) 95-113.

Casadaban, J.M. and Cohen, S.N.: Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. J. Mol. Biol. 138 (1980) 179-207.

Chang, A.C.Y. and Cohen, S.N. J. Bacteriol., 134, 1141. (1978).

Diderichsen, B. and Christiansen, L. Cloning of a maltogenic alpha-amylase from Bacillus stearothermophilus. FEMS Microbiol. lett. 56, 53-60. (1988).

Dubnau, D. and Davidoff-Adelson, R. Fate of transforming DNA following uptake by competent B.subtilis 1. Formation and properties of the donor recipient complex. J. Mol. Biol. 56, 209-221. (1971).

Ehrlich, S.D., Riele, H., Petit, M.A., Janniere, L., Noirot, P. and Michael, B. DNA recombination in plasmids and the chromosome of B.subtilis. Bacillus Mol. Genet. and Biotech.Appl. 27-33. (1986).

Heijne, G.: Patterns of amino acids near signal-sequence cleaage sites. Eur. J. Biochem. 133, 17-21 (1983).

Holmes, D.S. and Quigley, M.: A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114 (1981) 193-197.

Horinouchi, S. and Weisblum, B.: Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J. Bact. 150, 2, 815-825.

Johnson, W.C., Moran, C.P. Jr. and Losich, R.: (1983) Two RNA polymerase sigma factors from Bacillus subtilis discriminate between overlapping promoters for a developmentally regulated gen.

Jørgensen, P.L.: Construction of vectors for cloning of promotorer. M.Sc. Thesis, The Technical University of Denmark (1983).

Keggins, K.M., Lovell, P.S. and Duvall, E.J. Molecular cloning of Bacillus DNA in Bacillus subtilis and properties of the vector plasmid pUB110. Proc.Natl.Acad.Sci.USA.75, 3, 1423-1427. (1978).

Kreft, J., Burger, J.K. and Goebel, W.: Expression of antibiotic resistance genes from E.coli in Bacillus subtilis. Mol.Gen.Genet. 190: 384-389 (1983).

Laemmli, U.K.: Cleavage of structural proteins during the assembly of head of bacteriophage T4. Nature 227 (1970) 680-685.

Loening, U.E.: The fractionation of high-molecular-weight ribonucleic acid polyacrylamide gel electrophoresis. Biochem. J. 102 (1967) 251-257.

Mandel, M. and Higa, A.: Calcium dependent bacteriophage DNA infection, J. Mol. Biol. 53 (1970) 154-162.

Maniatis, T., Fritsch, E.F., and Sambrook, J.: Molecular Cloning - A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.

Maxam, A.M. and Gilbert, W.: Sequencing end-labelled DNA with base-specific chemical cleavage. In: Grossman, L. and Moldave, K. (Eds., Methods in Enzymology, Vol. 65, Academic Press, New York, 1980, pp 499-560.

Mclaughlin, J.R., Murray, C.L. and Babinowitz, C.J. 1981. Unique features in the ribosome binding site sequence of the gram-positive Staphylococuccus aureus B-laclamase gene. J. Biol. Chem. 256: 11283-11291.

von Meyenburg, K., Jørgensen, B.B., Nielsen, J., and Hansen, F.G.: Promoters of the atp operon coding for the membrane bound ATP synthase of Escherichia coli mapped by Tn10 insertion mutations. Mol. Gen. Genet. 188 (1982) 240-248.

Miller, G.L.: Use of dinitrosalicyclic reagent for determination of reducing sugars, Anal. Chem. 31. (1959) 426-428.

Nakamura, A., Uozumi, T., Beppu, T. (1987). Nucleotide sequence of a cellulase gene of Bacillus subtilis. Eur. J. Biochem. 164, 317-320.

Nilsson, B., Uhlen, M., Josephson, S., Gatenbeck, S. and Philipson, L.: An improved positive selection plasmid vector constructed by oligonucleotide mediated mutagenesis. Nucleic Acids Res. 11: 8019-8030.

Ptashne, M., Johnson, A.D. and Pabo, C.O.: A genetic switch in a bacterial virus. Scientific American, November (1982).

Remaut et al., Gene 15, 1981, pp. 81-93.

Remaut, E., Stanssens, P. and Fiirs, W.: Plasmid vectors for high-efficiency expression controlled by the P_L promoter of coliphage lambda. Gene, 15, 81-93. (1981).

Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977): Labelling deoxyribonucleic acid to high specific activity in

vitro by nick translation with DNA polymerase I. I. Mol. Biol. 113, 237-251.

Rosenberg, M. and Court, D.: Regulatory sequences involved in the promotion and termination of RNA transcription. Ann. Rev. Genet. 1979 13: 319-53.

Sancar, A., Hack, A.M., and Rupp, W.D.: Simpel methods for identifications of plasmid-coded proteins, J. Bacteriol. 137 (1979) 692-693.

Sanger, F., Nichlen, S. and Coulson, A.R. (1977). Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Southern, E.: Detection of specific sequences among DNA fragments separated by gel electrophoresis. I. Mol. Biol. 98, 503-517 (1975).

Stephens, M.A., Ortlepp, S.A., Ollington, J.F. and McConnel, D.J. J.bact. Vol.158, No 1, p 369-372. (1984).

Teather, R.M. and Wood, P.J.: Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl. Environ. Microbiol. -3 (1982) 777-780.

Walseth, C.S.: Occurrence of cellulases in enzyme preparations from microorganisms. Tappi, Vol. 35, No. 5 (1952) 228-233.

Yanisch-Perron, C., Vieira, J. and Messing, J. Gene, 33, 103-119. (1985).

SEKVENSSILISTAUS

(1) Yleistä tietoa:

(i) Hakija: Novo Nordisk A/S

(ii) Keksinnön nimitys: Entsyymi, jossa on sellulaasiaktiivisuutta

(iii) Sekvenssien lukumäärä: 7

(iv) Kirjeenvaihto-osoite:

(A) Saaja: Novo Nordisk A/S, Patent Department

(B) Katu: Novo Alle

(C) Kaupunki: Bagsvaerd

(D) Maa: Tanska

(F) Postinumero: DK-2880

(v) Tietokoneella luettava muoto:

(A) Välinetyyppi

(B) Tietokone:

(C) Käyttöjärjestelmä:

(D) Software:

(vi) Hakemustiedot:

(A) Hakemusnumero:

(B) Jättöpäivä:

(C) Luokitus:

(vii) Prior. hakemustiedot:

(A) Hakemusnumero: DK 164/90

(B) Jättöpäivä: 19-01-1990

(viii) Asiamiestiedot:

(A) Nimi: Thalsø-Madsen, Kine Birgit

(B) Rekisteröintinumero:

(C) Viitenumero; 3425.204-WO

(ix) Yhteystiedot:

(A) Puhelin: +45 4444 8888

(B) Telefax: +45 4449 3256

(C) Telex 37204

(2) Sekvenssin ID NO 1: tiedot

(i) Sekvenssin piirteet:

(A) Pituus; 2977 emäsparia

(B) Tyyppi: nukleiinihappo

(C) Juosteisuus: Kaksoisjuoste
(D) Topologia; lineaarinen

(ii) Molekyylityyppi: DNA (genominen)

(vi) Alkuperälähde:
(A) Organismi: Bacillus lautus
(B) Kanta: NCIMB 40250

(ix) Piirre:
(A) Nimi/Avainsana: CDS
(B) Sijainti 677-2776
(D) Muuta tietoa:

(xi) Sekvenssin kuvaus: SEQ ID NO 1:

```
ATCTTGAAGC TTCGAGGTTC GCGTAACCAG GTCGATCGGC TGACTTTCCT CGCCAAGCTG      60

CACCATGGCT TCATAGATTA ATTGGTGAGG AGGATCATAG AAATCTTCCG TCCGCACCGC     120

TCCATGGCGG TAATCAGCGC TTCCGACTGC AGTACAGAGC GGCCGATACA AATGCAGCCG     180

ACAACCAGAT CAAGCCGTCC TTCAACATCA AAAACAACGG TACTTCGGCT GTTGATTTAA     240

GCACGCTCAA AATCCGCTAC TACTTCACCA AGGATGGTTC TGCGGCGGTG AACGGCTGGA     300

TCGACTGGGC GCAGCTCGGC GGCAGCAACA TTCAGATCTC GTTTGGCAAC CATACTGGCA     360

CGAATTCGGA TACGTACGTG GAGCTGAGCT TCTCGTCCGA GGCAGGCTCG ATTGCGGCGG     420

GCGGCCAATC CGGTGAAATC CAGCTGCGCA TGTCCAAGAC GGACTGGTCG AACTTTAACG     480

AGGCGAACGA CTACTCGTTC GATGGGACGA AGACGGCCTT TGCTGACTGG GATCGGGTCG     540

TATTGTACCA GAACGGCCAA ATAGTGTGGG GAACTGCTCC ATAAACCGAT ACAGGGGAAT     600

GTGCCGGAAC CGCTCTTTTG CAGGGCAGAC TGGCGGTATC CCTTGCTGAA ATGACTATTC     660

CTGGGAGGGA TCAAAA ATG AAG ACA AGA CAA AGA AAG CGG CTG TTC GTC           709
                  Met Lys Thr Arg Gln Arg Lys Arg Leu Phe Val
                    1               5                  10

AGT GCG GCG CTG GCA GTA TCC TTG ACA ATG ACC GTA CCG ATG CCC GCT        757
Ser Ala Ala Leu Ala Val Ser Leu Thr Met Thr Val Pro Met Pro Ala
             15                  20                  25
```

```
TCT GTA AAT GCA GCT GCG AGT GAT GTC ACT TTC ACG ATT AAT ACG CAG      805
Ser Val Asn Ala Ala Ala Ser Asp Val Thr Phe Thr Ile Asn Thr Gln
          30                  35                  40

TCG GAA CGT GCA GCG ATC AGC CCC AAT ATT TAC GGA ACC AAT CAG GAT      853
Ser Glu Arg Ala Ala Ile Ser Pro Asn Ile Tyr Gly Thr Asn Gln Asp
     45                  50                  55

CTG AGC GGG ACG GAG AAC TGG TCA TCC CGC AGG CTC GGA GGC AAC CGG      901
Leu Ser Gly Thr Glu Asn Trp Ser Ser Arg Arg Leu Gly Gly Asn Arg
 60                  65                  70                  75

CTG ACG GGT TAC AAC TGG GAG AAC AAC GCA TCC AGC GCC GGA AGG GAC      949
Leu Thr Gly Tyr Asn Trp Glu Asn Asn Ala Ser Ser Ala Gly Arg Asp
                 80                  85                  90

TGG CTT CAT TAC AGC GAT GAT TTT CTC TGC GGC AAC GGT GGT GTT CCA      997
Trp Leu His Tyr Ser Asp Asp Phe Leu Cys Gly Asn Gly Gly Val Pro
            95                 100                 105

GAC ACC GAC TGC GAC AAG CCG GGG GCG GTT GTT ACC GCT TTT CAC GAT     1045
Asp Thr Asp Cys Asp Lys Pro Gly Ala Val Val Thr Ala Phe His Asp
       110                 115                 120

AAA TCT TTG GAG AAT GGA GCT TAC TCC ATT GTA ACG CTG CAA ATG GCG     1093
Lys Ser Leu Glu Asn Gly Ala Tyr Ser Ile Val Thr Leu Gln Met Ala
  125                 130                 135

GGT TAT GTG TCC CGG GAT AAG AAC GGT CCA GTT GAC GAG AGT GAG ACG     1141
Gly Tyr Val Ser Arg Asp Lys Asn Gly Pro Val Asp Glu Ser Glu Thr
140                 145                 150                 155

GCT CCG TCA CCG CGT TGG GAT AAG GTC GAG TTT GCC AAA AAT GCG CCG     1189
Ala Pro Ser Pro Arg Trp Asp Lys Val Glu Phe Ala Lys Asn Ala Pro
                160                 165                 170

TTC TCC CTT CAG CCT GAT CTG AAC GAC GGA CAA GTG TAT ATG GAT GAA     1237
Phe Ser Leu Gln Pro Asp Leu Asn Asp Gly Gln Val Tyr Met Asp Glu
            175                 180                 185

GAA GTT AAC TTC CTG GTC AAC CGG TAT GGA AAC GCT TCA ACG TCA ACG     1285
Glu Val Asn Phe Leu Val Asn Arg Tyr Gly Asn Ala Ser Thr Ser Thr
       190                 195                 200

GGC ATC AAA GCG TAT TCG CTG GAT AAC GAG CCG GCG CTG TGG TCT GAG     1333
Gly Ile Lys Ala Tyr Ser Leu Asp Asn Glu Pro Ala Leu Trp Ser Glu
  205                 210                 215

ACG CAT CCA AGG ATT CAT CCG GAG CAG TTA CAA GCG GCA GAA CTC GTC     1381
Thr His Pro Arg Ile His Pro Glu Gln Leu Gln Ala Ala Glu Leu Val
220                 225                 230                 235

GCT AAG AGC ATC GAC TTG TCA AAG GCG GTG AAG AAC GTC GAT CCG CAT     1429
Ala Lys Ser Ile Asp Leu Ser Lys Ala Val Lys Asn Val Asp Pro His
                240                 245                 250
```

```
GCC GAA ATA TTC GGT CCT GCC CTT TAC GGT TTC GGC GCA TAT TTG TCT      1477
Ala Glu Ile Phe Gly Pro Ala Leu Tyr Gly Phe Gly Ala Tyr Leu Ser
            255                 260                 265

CTG CAG GAC GCA CCG GAT TGG CCG AGT TTG CAA GGC AAC TAC AGC TGG      1525
Leu Gln Asp Ala Pro Asp Trp Pro Ser Leu Gln Gly Asn Tyr Ser Trp
        270                 275                 280

TTT ATC GAT TAC TAT CTG GAT CAG ATG AAG AAT GCT CAT ACG CAG AAC      1573
Phe Ile Asp Tyr Tyr Leu Asp Gln Met Lys Asn Ala His Thr Gln Asn
    285                 290                 295

GGC AAA AGA TTG CTC GAT GTG CTG GAC GTC CAC TGG TAT CCG GAA GCA      1621
Gly Lys Arg Leu Leu Asp Val Leu Asp Val His Trp Tyr Pro Glu Ala
300                 305                 310                 315

CAG GGC GGA GGC CAG CGA ATC GTC TTT GGC GGG GCG GGC AAT ATC GAT      1669
Gln Gly Gly Gly Gln Arg Ile Val Phe Gly Gly Ala Gly Asn Ile Asp
                320                 325                 330

ACG CAG AAG GCT CGC GTA CAA GCG CCA AGA TCG CTA TGG GAT CCG GCT      1717
Thr Gln Lys Ala Arg Val Gln Ala Pro Arg Ser Leu Trp Asp Pro Ala
            335                 340                 345

TAC CAG GAA GAC AGC TGG ATC GGC ACA TGG TTT TCA AGC TAC TTG CCC      1765
Tyr Gln Glu Asp Ser Trp Ile Gly Thr Trp Phe Ser Ser Tyr Leu Pro
        350                 355                 360

TTA ATT CCG AAG CTG CAA TCT TCG ATT CAG ACG TAT TAT CCG GGT ACG      1813
Leu Ile Pro Lys Leu Gln Ser Ser Ile Gln Thr Tyr Tyr Pro Gly Thr
    365                 370                 375

AAG CTG GCG ATC ACA GAG TTC AGC TAC GGC GGA GAC AAT CAC ATT TCG      1861
Lys Leu Ala Ile Thr Glu Phe Ser Tyr Gly Gly Asp Asn His Ile Ser
380                 385                 390                 395

GGA GGC ATA GCT ACC GCG GAC GCG CTC GGC ATT TTT GGA AAA TAT GGC      1909
Gly Gly Ile Ala Thr Ala Asp Ala Leu Gly Ile Phe Gly Lys Tyr Gly
                400                 405                 410

GTT TAT GCC GCG AAT TAC TGG CAG ACG GAG GAC AAT ACC GAT TAT ACC      1957
Val Tyr Ala Ala Asn Tyr Trp Gln Thr Glu Asp Asn Thr Asp Tyr Thr
            415                 420                 425

AGC GCT GCT TAC AAG CTG TAT CGC AAC TAC GAC GGC AAT AAA TCG GGG      2005
Ser Ala Ala Tyr Lys Leu Tyr Arg Asn Tyr Asp Gly Asn Lys Ser Gly
        430                 435                 440

TTC GGC TCG ATC AAA GTG GAC GCC GCT ACG TCC GAT ACG GAG AAC AGC      2053
Phe Gly Ser Ile Lys Val Asp Ala Ala Thr Ser Asp Thr Glu Asn Ser
    445                 450                 455

TCG GTA TAC GCT TCG GTA ACT GAC GAG GAG AAT TCC GAA CTC CAC CTG      2101
Ser Val Tyr Ala Ser Val Thr Asp Glu Glu Asn Ser Glu Leu His Leu
460                 465                 470                 475
```

```
ATC GTG CTG AAT AAA AAT TTC GAC GAT CCG ATC AAC GCT ACT TTC CAG      2149
Ile Val Leu Asn Lys Asn Phe Asp Asp Pro Ile Asn Ala Thr Phe Gln
                480                 485                 490

CTG TCT GGT GAT AAA ACC TAC ACA TCC GGG AGA GTA TGG GGC TTC GAC      2197
Leu Ser Gly Asp Lys Thr Tyr Thr Ser Gly Arg Val Trp Gly Phe Asp
            495                 500                 505

CAA ACC GGA TCC GAC ATT ACG GAA CAA GCA GCT ATA ACG AAT ATT AAC      2245
Gln Thr Gly Ser Asp Ile Thr Glu Gln Ala Ala Ile Thr Asn Ile Asn
        510                 515                 520

AAC AAT CAA TTC ACG TAT ACG CTT CCT CCA TTG TCG GCT TAC CAC ATT      2293
Asn Asn Gln Phe Thr Tyr Thr Leu Pro Pro Leu Ser Ala Tyr His Ile
    525                 530                 535

GTT CTG AAA GCG GAT AGC ACC GAA CCG GTC AAC TCC GAT CTC GTC GTG      2341
Val Leu Lys Ala Asp Ser Thr Glu Pro Val Asn Ser Asp Leu Val Val
540                 545                 550                 555

CAG TAT AAG GAC GGT GAT CGC AAC AAT GCA ACC GAC AAT CAG ATC AAG      2389
Gln Tyr Lys Asp Gly Asp Arg Asn Asn Ala Thr Asp Asn Gln Ile Lys
                560                 565                 570

CCG CAT TTC AAT ATT CAA AAT AAA GGG ACC AGC CCG GTA GAT CTG AGT      2437
Pro His Phe Asn Ile Gln Asn Lys Gly Thr Ser Pro Val Asp Leu Ser
            575                 580                 585

TCC CTT ACC CTG CGC TAC TAT TTT ACC AAA GAC AGC TCT GCA GCG ATG      2485
Ser Leu Thr Leu Arg Tyr Tyr Phe Thr Lys Asp Ser Ser Ala Ala Met
        590                 595                 600

AAC GGC TGG ATC GAT TGG GCG AAG CTC GGC GGC AGC AAC ATT CAG ATT      2533
Asn Gly Trp Ile Asp Trp Ala Lys Leu Gly Gly Ser Asn Ile Gln Ile
    605                 610                 615

TCG TTC GGT AAT CAT AAT GGC GCG GAT TCG GAT ACG TAC GCG GAG CTG      2581
Ser Phe Gly Asn His Asn Gly Ala Asp Ser Asp Thr Tyr Ala Glu Leu
620                 625                 630                 635

GGC TTC TCG TCC GGC GCA GGC TCG ATT GCG GAG GGC GGT CAA TCC GGC      2629
Gly Phe Ser Ser Gly Ala Gly Ser Ile Ala Glu Gly Gly Gln Ser Gly
                640                 645                 650

GAA ATC CAG CTG CGC ATG TCG AAG GCG GAC TGG TCG AAC TTC AAC GAG      2677
Glu Ile Gln Leu Arg Met Ser Lys Ala Asp Trp Ser Asn Phe Asn Glu
            655                 660                 665

GCG AAC GAC TAC TCG TTC GAT GGG GCG AAG ACG GCC TAT ATA GAT TGG      2725
Ala Asn Asp Tyr Ser Phe Asp Gly Ala Lys Thr Ala Tyr Ile Asp Trp
        670                 675                 680
```

GAT CGC GTG ACG CTA TAC CAA GAC GGA CAA CTC GTA TGG GGA ATC GAG 2773
Asp Arg Val Thr Leu Tyr Gln Asp Gly Gln Leu Val Trp Gly Ile Glu
685 690 695

CCG TAGAAGATGA CTAGACAACA TTAGTGATGA GACGCGGCCG GCCATAACGG 2826
Pro
700

CTGTCTTGAC TCTGATTCGA TCAAAAAATC AAAGCAAAGG GGATGAAAGT AATGAATGTT 2886

GCGATTCAAA AGAGAATCGG ATCAATATTG ATGATTGCCT CACTAATTAT TAGCTTATTG 2946

CCGTTAGGGA GCAGCAGAAC GTATGCTGCA G 2977

(2) Sekvenssin ID NO 2 tiedot

(i) Sekvenssin karasteristiikka
(A) Pituus: 700 aminohappoa
(B) Tyyppi: aminohappo
(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO 2

Met Lys Thr Arg Gln Arg Lys Arg Leu Phe Val Ser Ala Ala Leu Ala
1 5 10 15

Val Ser Leu Thr Met Thr Val Pro Met Pro Ala Ser Val Asn Ala Ala
20 25 30

Ala Ser Asp Val Thr Phe Thr Ile Asn Thr Gln Ser Glu Arg Ala Ala
35 40 45

Ile Ser Pro Asn Ile Tyr Gly Thr Asn Gln Asp Leu Ser Gly Thr Glu
50 55 60

Asn Trp Ser Ser Arg Arg Leu Gly Gly Asn Arg Leu Thr Gly Tyr Asn
65 70 75 80

Trp Glu Asn Asn Ala Ser Ser Ala Gly Arg Asp Trp Leu His Tyr Ser
85 90 95

Asp Asp Phe Leu Cys Gly Asn Gly Gly Val Pro Asp Thr Asp Cys Asp
100 105 110

Lys Pro Gly Ala Val Val Thr Ala Phe His Asp Lys Ser Leu Glu Asn
115 120 125

Gly Ala Tyr Ser Ile Val Thr Leu Gln Met Ala Gly Tyr Val Ser Arg
130 135 140

Asp Lys Asn Gly Pro Val Asp Glu Ser Glu Thr Ala Pro Ser Pro Arg
145 150 155 160

Trp Asp Lys Val Glu Phe Ala Lys Asn Ala Pro Phe Ser Leu Gln Pro
165 170 175

Asp Leu Asn Asp Gly Gln Val Tyr Met Asp Glu Glu Val Asn Phe Leu
180 185 190

Val Asn Arg Tyr Gly Asn Ala Ser Thr Ser Thr Gly Ile Lys Ala Tyr
195 200 205

Ser Leu Asp Asn Glu Pro Ala Leu Trp Ser Glu Thr His Pro Arg Ile
210 215 220

His Pro Glu Gln Leu Gln Ala Ala Glu Leu Val Ala Lys Ser Ile Asp
225 230 235 240

Leu Ser Lys Ala Val Lys Asn Val Asp Pro His Ala Glu Ile Phe Gly
245 250 255

Pro Ala Leu Tyr Gly Phe Gly Ala Tyr Leu Ser Leu Gln Asp Ala Pro
260 265 270

Asp Trp Pro Ser Leu Gln Gly Asn Tyr Ser Trp Phe Ile Asp Tyr Tyr
275 280 285

Leu Asp Gln Met Lys Asn Ala His Thr Gln Asn Gly Lys Arg Leu Leu
290 295 300

Asp Val Leu Asp Val His Trp Tyr Pro Glu Ala Gln Gly Gly Gly Gln
305 310 315 320

Arg Ile Val Phe Gly Gly Ala Gly Asn Ile Asp Thr Gln Lys Ala Arg
325 330 335

Val Gln Ala Pro Arg Ser Leu Trp Asp Pro Ala Tyr Gln Glu Asp Ser
340 345 350

Trp Ile Gly Thr Trp Phe Ser Ser Tyr Leu Pro Leu Ile Pro Lys Leu
355 360 365

Gln Ser Ser Ile Gln Thr Tyr Tyr Pro Gly Thr Lys Leu Ala Ile Thr
370 375 380

Glu Phe Ser Tyr Gly Gly Asp Asn His Ile Ser Gly Gly Ile Ala Thr
385 390 395 400

Ala Asp Ala Leu Gly Ile Phe Gly Lys Tyr Gly Val Tyr Ala Ala Asn
405 410 415

Tyr Trp Gln Thr Glu Asp Asn Thr Asp Tyr Thr Ser Ala Ala Tyr Lys
420 425 430

Leu Tyr Arg Asn Tyr Asp Gly Asn Lys Ser Gly Phe Gly Ser Ile Lys
435 440 445

Val Asp Ala Ala Thr Ser Asp Thr Glu Asn Ser Ser Val Tyr Ala Ser
450 455 460

Val Thr Asp Glu Glu Asn Ser Glu Leu His Leu Ile Val Leu Asn Lys
465 470 475 480

Asn Phe Asp Asp Pro Ile Asn Ala Thr Phe Gln Leu Ser Gly Asp Lys
485 490 495

Thr Tyr Thr Ser Gly Arg Val Trp Gly Phe Asp Gln Thr Gly Ser Asp
500 505 510

Ile Thr Glu Gln Ala Ala Ile Thr Asn Ile Asn Asn Asn Gln Phe Thr
515 520 525

Tyr Thr Leu Pro Pro Leu Ser Ala Tyr His Ile Val Leu Lys Ala Asp
530 535 540

Ser Thr Glu Pro Val Asn Ser Asp Leu Val Val Gln Tyr Lys Asp Gly
545 550 555 560

Asp Arg Asn Asn Ala Thr Asp Asn Gln Ile Lys Pro His Phe Asn Ile
565 570 575

Gln Asn Lys Gly Thr Ser Pro Val Asp Leu Ser Ser Leu Thr Leu Arg
580 585 590

Tyr Tyr Phe Thr Lys Asp Ser Ser Ala Ala Met Asn Gly Trp Ile Asp
595 600 605

Trp Ala Lys Leu Gly Gly Ser Asn Ile Gln Ile Ser Phe Gly Asn His
610 615 620

Asn Gly Ala Asp Ser Asp Thr Tyr Ala Glu Leu Gly Phe Ser Ser Gly
625 630 635 640

Ala Gly Ser Ile Ala Glu Gly Gly Gln Ser Gly Glu Ile Gln Leu Arg
645 650 655

Met Ser Lys Ala Asp Trp Ser Asn Phe Asn Glu Ala Asn Asp Tyr Ser
660 665 670

Phe Asp Gly Ala Lys Thr Ala Tyr Ile Asp Trp Asp Arg Val Thr Leu
675 680 685

Tyr Gln Asp Gly Gln Leu Val Trp Gly Ile Glu Pro
690 695 700

(2) Sekvenssin ID NO 3 tiedot

(i) Sekvenssin karasteristiikka
(A) Pituus: 2323 emäsparia
(B) Tyyppi: nukleiinihappo
(C) Juosteisuus: kaksoisjuoste
(D) Topologia: lineaarinen

(ii) Molekyylityyppi: DNA (genominen)

(vi) Alkuperälähde:
(A) Organismi: Bacillus lautus
(B) Kanta: NCIMB 40250

Piirre:
(A) Nimi/Avainsana CDS
(B) Sijainti: 172-1869
(D) Muuta tietoa

(xi) Sekvenssin kuvaus: SEQ ID NO 3

GGGGTTTTGT TCATTTCAAG AAAGAGCACG AATAAACATC GCTAATTTAC ACATGATTTA 60

TTCTCTAATT ATTATGGTGC ATGCCAATTG GAAAATGTAG TAGATTAGTC ATCGTAACCT 120

GCTTTCATGC CAAGATATGT AATTTTTAAA AAAGAGTAAA GGAGAATTCA G ATG AAA 177
Met Lys
1

AAA CGT AGA AGC AGT AAA GTT ATT CTT TCG TTG GCC ATC GTT GTT GCA 225
Lys Arg Arg Ser Ser Lys Val Ile Leu Ser Leu Ala Ile Val Val Ala
5 10 15

TTA TTG GCA GCC GTC GAA CCT AAT GCC GCT TTA GCA GCG GCT CCA CCA 273
Leu Leu Ala Ala Val Glu Pro Asn Ala Ala Leu Ala Ala Ala Pro Pro
20 25 30

```
AGT GCC ATG CAG TCC TAT GTT GAA GCG ATG CAG CCT GGC TGG AAC CTT      321
Ser Ala Met Gln Ser Tyr Val Glu Ala Met Gln Pro Gly Trp Asn Leu
 35                  40                  45                  50

GGC AAT TCT CTG GAT GCT GTC GGT GCG GAT GAG ACG CTG GCA CGG GGC      369
Gly Asn Ser Leu Asp Ala Val Gly Ala Asp Glu Thr Leu Ala Arg Gly
                    55                  60                  65

AAT CCG CGG ATC ACG AAA GAG CTC ATT CAG AAC ATC GCT GCG CAA GGC      417
Asn Pro Arg Ile Thr Lys Glu Leu Ile Gln Asn Ile Ala Ala Gln Gly
                70                  75                  80

TAT AAG AGC ATA CGG ATT CCT GTT ACC TGG GAT TCC CAT ATC GGC GCG      465
Tyr Lys Ser Ile Arg Ile Pro Val Thr Trp Asp Ser His Ile Gly Ala
        85                  90                  95

GCC CCA AAT TAT CAA ATT GAA GCT GCG TAC CTC AAT CGA GTG CAG GAG      513
Ala Pro Asn Tyr Gln Ile Glu Ala Ala Tyr Leu Asn Arg Val Gln Glu
    100                 105                 110

GTC GTA CAG TGG GCT TTG GAC GCG AAC CTC TAT GTG ATG ATT AAT GTC      561
Val Val Gln Trp Ala Leu Asp Ala Asn Leu Tyr Val Met Ile Asn Val
115                 120                 125                 130

CAT CAT GAT TCC TGG CTA TGG ATC AGC AAA ATG GAG TCG CAG CAC GAT      609
His His Asp Ser Trp Leu Trp Ile Ser Lys Met Glu Ser Gln His Asp
                135                 140                 145

CAA GTA CTG GCC CGT TAT AAT GCG ATT TGG ACG CAA ATT GCC AAC AAG      657
Gln Val Leu Ala Arg Tyr Asn Ala Ile Trp Thr Gln Ile Ala Asn Lys
            150                 155                 160

TTC AAG AAC AGC CCG AGC AAG CTG ATG TTC GAG AGC GTG AAT GAG CCT      705
Phe Lys Asn Ser Pro Ser Lys Leu Met Phe Glu Ser Val Asn Glu Pro
        165                 170                 175

CGC TTT ACG GAT GGC GGA ACT ACG GAT GAA GCC AAG CAG CAA AAA ATG      753
Arg Phe Thr Asp Gly Gly Thr Thr Asp Glu Ala Lys Gln Gln Lys Met
    180                 185                 190

CTG GAC GAG CTG AAC GTA TCC TTT TTC AAC ATC GTC AGA AAT TCC GGC      801
Leu Asp Glu Leu Asn Val Ser Phe Phe Asn Ile Val Arg Asn Ser Gly
195                 200                 205                 210

GGC CAG AAC GCG ACT CGC CCG CTA GTT CTT TCT ACG TTG GAG GCC TCT      849
Gly Gln Asn Ala Thr Arg Pro Leu Val Leu Ser Thr Leu Glu Ala Ser
                215                 220                 225

CCC ACC CAA GAG AGA ATG ACG GCG CTT TAT AAT ACG ATG ACC AAA CTG      897
Pro Thr Gln Glu Arg Met Thr Ala Leu Tyr Asn Thr Met Thr Lys Leu
            230                 235                 240

AAC GAC AAG AAT CTG ATC GCA ACC GTT CAT TTT TAT GGA TTC TGG CCG      945
Asn Asp Lys Asn Leu Ile Ala Thr Val His Phe Tyr Gly Phe Trp Pro
        245                 250                 255
```

```
TTT AGC GTA AAT ATC GCA GGA TAT ACG AAA TTT GAT GCG GAG ACG CAA      993
Phe Ser Val Asn Ile Ala Gly Tyr Thr Lys Phe Asp Ala Glu Thr Gln
    260                 265                 270

AAT GAT ATT ATA ACG ACC TTC GAT AAC GTG TAT AAC ACA TTT GTA GCA     1041
Asn Asp Ile Ile Thr Thr Phe Asp Asn Val Tyr Asn Thr Phe Val Ala
275                 280                 285                 290

AAG GGA ATC CCG GTG GTA GTC GGC GAA TAT GGC CTT CTT GGA TTC GAT     1089
Lys Gly Ile Pro Val Val Val Gly Glu Tyr Gly Leu Leu Gly Phe Asp
                295                 300                 305

AAG AAT ACC GGC GTC ATT GAA CAG GGT GAG AAA TTG AAA TTT TTC GAG     1137
Lys Asn Thr Gly Val Ile Glu Gln Gly Glu Lys Leu Lys Phe Phe Glu
            310                 315                 320

TTT TTT GCC CAG TAT GTG AAG CAA AAA AGC ATT TCC ACT ATG CTA TGG     1185
Phe Phe Ala Gln Tyr Val Lys Gln Lys Ser Ile Ser Thr Met Leu Trp
        325                 330                 335

GAT AAC GGA CAG CAC TTC AAC CGC ACG AGC TTC AAG TGG TCT GAC CCG     1233
Asp Asn Gly Gln His Phe Asn Arg Thr Ser Phe Lys Trp Ser Asp Pro
    340                 345                 350

GAT TTA TTC AAT ATG ATC AAG GCC AGT TGG ACC GGA CGT TCA TCC ACG     1281
Asp Leu Phe Asn Met Ile Lys Ala Ser Trp Thr Gly Arg Ser Ser Thr
355                 360                 365                 370

GCT TCC AGC GAC CTG ATC CAT GTC AAG CAG GGC ACG GCG GTA AAA GAT     1329
Ala Ser Ser Asp Leu Ile His Val Lys Gln Gly Thr Ala Val Lys Asp
                375                 380                 385

ACT TCG GTT CAG CTC AAT CTT AAC GGG AAT ACG CTA ACT TCC CTT TCC     1377
Thr Ser Val Gln Leu Asn Leu Asn Gly Asn Thr Leu Thr Ser Leu Ser
            390                 395                 400

GTA AAT GGA ACG ACA CTG AAA TCA GGC ACA GAT TAC ACT TTA AAC AGC     1425
Val Asn Gly Thr Thr Leu Lys Ser Gly Thr Asp Tyr Thr Leu Asn Ser
        405                 410                 415

AGC AGA TTA ACT TTT AAA GCG AGC CAG TTG ACC AAG CTG ACC TCC TTG     1473
Ser Arg Leu Thr Phe Lys Ala Ser Gln Leu Thr Lys Leu Thr Ser Leu
    420                 425                 430

GGC AAA TTG GGG GTC AAC GCG ACG ATC GTG ACT AAA TTC AAT AGA GGC     1521
Gly Lys Leu Gly Val Asn Ala Thr Ile Val Thr Lys Phe Asn Arg Gly
435                 440                 445                 450

GCC GAC TGG AAG TTC AAC GTA GTC CTG TAC AAT ACG CCT AAG CTT AGC     1569
Ala Asp Trp Lys Phe Asn Val Val Leu Tyr Asn Thr Pro Lys Leu Ser
                455                 460                 465
```

```
AGT ACG ACG GGG ACT ACT TCT TCC TTT GCG ATT CCA ACG GCT TTC AAC      1617
Ser Thr Thr Gly Thr Thr Ser Ser Phe Ala Ile Pro Thr Ala Phe Asn
            470                 475                 480

GGG GAT CAG CTT GCT ACG ATG GAA GCG GTC TAT GTA AAC GGC GGC AAT      1665
Gly Asp Gln Leu Ala Thr Met Glu Ala Val Tyr Val Asn Gly Gly Asn
        485                 490                 495

GCC GGT CCG CAT AAC TGG ACT TCC TTT AAG GAA TTC GAA ACG ACG TTC      1713
Ala Gly Pro His Asn Trp Thr Ser Phe Lys Glu Phe Glu Thr Thr Phe
    500                 505                 510

AGC CCC GCT TAT AGC GAG GGG AAA ATC AAA CTG CAG CAG GCG TTC TTT      1761
Ser Pro Ala Tyr Ser Glu Gly Lys Ile Lys Leu Gln Gln Ala Phe Phe
515                 520                 525                 530

AAT GAA GTG AAT GAT ACC ACA GTC ACG CTC AAG TTC CAA TTC TGG AGC      1809
Asn Glu Val Asn Asp Thr Thr Val Thr Leu Lys Phe Gln Phe Trp Ser
                535                 540                 545

GGG GAG ATC GTC AAC TAC ACG ATT AAA AAG AGC GGT TCG ACG GTG ACG      1857
Gly Glu Ile Val Asn Tyr Thr Ile Lys Lys Ser Gly Ser Thr Val Thr
            550                 555                 560

GGT ACG GCT TCA TAAGCGAGTT TGGCAAAAAA GGACCGATAT ACTGCCTAAT          1909
Gly Thr Ala Ser
        565

TTGGTATTGC CTTAGTTGAA AGCAATTGCT CCGAATAAAC AGAATGAAGC CCCGGCCAGC    1969

TGGCCGGGAC TTATGCGTTT AGGAAGTATA AACGAATCAT CAGCAATTTA TTTAGCTCGT    2029

CTCAGTTCAG CAATATCGGC TTCATGTGAA ACGGAGCGGA TGAACAATCT TTCGAGCAAT    2089

TTCTCATGCT CCTGCTGGGT TTGGAGAACG GTTTGCTGAT TAGTTTTAAG TACAGATATA    2149

TCCTCACGGA CTTGATTGAT TCATGTGGTC CGTTAGTTCT TCTACCTTTG TATTTGTGGC    2209

AGCAACGATA TGAATTAATT GTTGAATGTG CCCGCCATGA CTGTTTAGCT GCTCATTGTG    2269

GCTTTGTAAC TGTTCTCGGA TTTCTTTGAA TTCTTGGTCG TGCTCATTAA GCTT          2323
```

(2) Sekvenssin ID NO 4 tiedot

(i) Sekvenssin karasteristiikka

(A) Pituus: 566 aminohappoa

(B) Tyyppi: aminohappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO 4

```
Met Lys Lys Arg Arg Ser Ser Lys Val Ile Leu Ser Leu Ala Ile Val
 1               5                  10                  15

Val Ala Leu Leu Ala Ala Val Glu Pro Asn Ala Ala Leu Ala Ala Ala
            20                  25                  30

Pro Pro Ser Ala Met Gln Ser Tyr Val Glu Ala Met Gln Pro Gly Trp
        35                  40                  45

Asn Leu Gly Asn Ser Leu Asp Ala Val Gly Ala Asp Glu Thr Leu Ala
    50                  55                  60

Arg Gly Asn Pro Arg Ile Thr Lys Glu Leu Ile Gln Asn Ile Ala Ala
65                  70                  75                  80

Gln Gly Tyr Lys Ser Ile Arg Ile Pro Val Thr Trp Asp Ser His Ile
                85                  90                  95

Gly Ala Ala Pro Asn Tyr Gln Ile Glu Ala Ala Tyr Leu Asn Arg Val
            100                 105                 110

Gln Glu Val Val Gln Trp Ala Leu Asp Ala Asn Leu Tyr Val Met Ile
        115                 120                 125

Asn Val His His Asp Ser Trp Leu Trp Ile Ser Lys Met Glu Ser Gln
    130                 135                 140

His Asp Gln Val Leu Ala Arg Tyr Asn Ala Ile Trp Thr Gln Ile Ala
145                 150                 155                 160

Asn Lys Phe Lys Asn Ser Pro Ser Lys Leu Met Phe Glu Ser Val Asn
                165                 170                 175

Glu Pro Arg Phe Thr Asp Gly Gly Thr Thr Asp Glu Ala Lys Gln Gln
            180                 185                 190

Lys Met Leu Asp Glu Leu Asn Val Ser Phe Phe Asn Ile Val Arg Asn
        195                 200                 205

Ser Gly Gly Gln Asn Ala Thr Arg Pro Leu Val Leu Ser Thr Leu Glu
    210                 215                 220

Ala Ser Pro Thr Gln Glu Arg Met Thr Ala Leu Tyr Asn Thr Met Thr
225                 230                 235                 240

Lys Leu Asn Asp Lys Asn Leu Ile Ala Thr Val His Phe Tyr Gly Phe
                245                 250                 255

Trp Pro Phe Ser Val Asn Ile Ala Gly Tyr Thr Lys Phe Asp Ala Glu
            260                 265                 270
```

Thr Gln Asn Asp Ile Ile Thr Thr Phe Asp Asn Val Tyr Asn Thr Phe
275 280 285

Val Ala Lys Gly Ile Pro Val Val Val Gly Glu Tyr Gly Leu Leu Gly
290 295 300

Phe Asp Lys Asn Thr Gly Val Ile Glu Gln Gly Glu Lys Leu Lys Phe
305 310 315 320

Phe Glu Phe Phe Ala Gln Tyr Val Lys Gln Lys Ser Ile Ser Thr Met
325 330 335

Leu Trp Asp Asn Gly Gln His Phe Asn Arg Thr Ser Phe Lys Trp Ser
340 345 350

Asp Pro Asp Leu Phe Asn Met Ile Lys Ala Ser Trp Thr Gly Arg Ser
355 360 365

Ser Thr Ala Ser Ser Asp Leu Ile His Val Lys Gln Gly Thr Ala Val
370 375 380

Lys Asp Thr Ser Val Gln Leu Asn Leu Asn Gly Asn Thr Leu Thr Ser
385 390 395 400

Leu Ser Val Asn Gly Thr Thr Leu Lys Ser Gly Thr Asp Tyr Thr Leu
405 410 415

Asn Ser Ser Arg Leu Thr Phe Lys Ala Ser Gln Leu Thr Lys Leu Thr
420 425 430

Ser Leu Gly Lys Leu Gly Val Asn Ala Thr Ile Val Thr Lys Phe Asn
435 440 445

Arg Gly Ala Asp Trp Lys Phe Asn Val Val Leu Tyr Asn Thr Pro Lys
450 455 460

Leu Ser Ser Thr Thr Gly Thr Thr Ser Ser Phe Ala Ile Pro Thr Ala
465 470 475 480

Phe Asn Gly Asp Gln Leu Ala Thr Met Glu Ala Val Tyr Val Asn Gly
485 490 495

Gly Asn Ala Gly Pro His Asn Trp Thr Ser Phe Lys Glu Phe Glu Thr
500 505 510

Thr Phe Ser Pro Ala Tyr Ser Glu Gly Lys Ile Lys Leu Gln Gln Ala
515 520 525

Phe Phe Asn Glu Val Asn Asp Thr Thr Val Thr Leu Lys Phe Gln Phe
530 535 540

Trp Ser Gly Glu Ile Val Asn Tyr Thr Ile Lys Lys Ser Gly Ser Thr
545 550 555 560

Val Thr Gly Thr Ala Ser
565

(2) Sekvenssin ID NO 5 tiedot

(i) Sekvenssin karasteristiikka
(A) Pituus: 1775 emäsparia
(B) Tyyppi: nukleiinihappo
(C) Juosteisuus: kaksoisjuoste
(D) Topologia: lineaarinen

(ii) Molekyylityyppi: DNA (genominen)

(v) Fragmenttityyppi: N-terminaalinen
(vi) Alkuperälähde:
(A) Organismi: Bacillus lautus
(B) Kanta: NCIMB 40250

Piirre:
(A) Nimi/Avainsana CDS
(B) Sijainti: 30-(1625.1775)
(C) Identifiointimenetelmä: kokeellinen
(D) Muuta tietoa: /osittainen
/todiste = kokeellinen
/transl-paitsi =(asemat 1446-1458,
aa:OTR)

(xi) Sekvenssin kuvaus: SEQ ID NO 5

TTGAAGCGCT GAATTCAGGA GGTTAAATAA TGCGTATTCA TGCAATTCGG CAATCTTGCC 60
GTTTGGTATT GACGATGGTT TTGATGCTTG GCTTATTGCT GCCTGTGGGC GCCCCCAAAG 120
GCTATGCCGC TCCGGCTGTT CCTTTTGGCC AATTAAAAGT TCAGGGCAAT CAATTGGTAG 180
GACAGTCCGG GCAAGCTGTT CAACTGGTTG GCATGAGCTC ACATGGATTG CAGTGGTATG 240
GCAATTTCGT CAACAAATCG TCGTTGCAGT GGATGAGAGA TAACTGGGGC ATCAACGTCT 300
TCCGTGCCGC TATGTATACT TCCGAAGATG GTTACATTAC GGATCCTTCC GTCAAGAACA 360
AGGTGAAGGA GGCGGTTCAG GCATCCATCG ATCTGGCCTT GTACGTTATT ATTGACTGGC 420

ATATCTTGTC TGATGGGAAT CCGAATACGT ACAAAGCGCA ATCGAAAGCG TTCTTCCAAG 480

AGATGGCCAC ATTGTACGGC AACACGCCGA ATGTAATCTA TGAAATCGCG ACGAGCCCAA 540

CGGAATGTGT CCTGGGCAGA TGTCAGTCGT CGGAAGAAGT GATCACGGCC ATTCGTTCGA 600

TTGACCCCGA CGGAGTGGTT ATCGTTGGCA GTCCAACCTG GAGCCAGGAT ATTCATCTGG 660

CGGCCGATAA CCCGGTATCA CATAGCAATG TCATGTATGC GCTTCATTTC TATTCAGGCA 720

CGCATGGACA GTTTTTGAGA GACCGAATTA CCTATGCGAT GAACAAAGGA GCAGCGATCT 780

TCGTTACCGA GTGGGGCACC AGTGATGCAT CCGGGAACGG CGGGCCGTAT TTGCCTCAGT 840

CCAAAGAGTG GATCGATTTC TTGAATGCTC GCAAGATCAG CTGGGTGAAC TGGTCGCTCG 900

CTGATAAAGT AGAAACGTCT GCTGCTCTTA TGCCAGGTGC ATCGCCTACC GGCGCTGGAC 960

CGATGCCCAA TTGTCGAATG GGCAAATCGG GTTCGCGATC AAATCCGGCA AGCAACTGGA 1020

GGCGGCAGGG CAATCCAACT GCACCGGCTG CCCCTACTAA CCTCTCGGCA AACGGCGGCA 1080

ACGCCCAGGT ATCATTAACC TGGAACGCAG TTAGCGGGGC GACGAGCTAT ACCGTAAAGC 1140

GAGCAACGAC GAGCGGCGGT CCGTACACGA ATGTGGACCG GGGTGTCACG GCGACGAGCT 1200

ACACGAACAC CGGGCTGACG AATGGCACGA CGTATTATTA TGTCGTGAGG GCATCCAATA 1260

GCGCGGGCAG CAGCGCGAAC TCCGCGCAAG CGAGCGCAAC GCCGGCTAGC GGCGGCGCCA 1320

GTACGGGGAA CCTTGTTGTC CAATACAAAG TTGGCGACAC TAGCGCCACG GATAACCAAA 1380

TGAAGCCTTC CTTTAACATC AAGAACAACG GTACAACCCC TGTTAACCTG AGCGGCCTCA 1440

AGCTTNNNNN NNNNNNNNAA AAAGACGGAC CTGCGGATAT GAGCTGCTCG ATCGACTGGG 1500

CGCAAATCGG CCGAACGAAT GTTCTGCTGG CATTCGCTAA CTTTACCGGG AGTAATACGG 1560

ATACTTACTG TTGTGAGCTA AGCTTTAGCT GCACTGCAGG TTCGTATCCC GGCTATGCGT 1620

GGACNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN 1680

NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN 1740

NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNN 1775

(2) Sekvenssin ID NO 6 tiedot

(i) Sekvenssin karasteristiikka
(A) Pituus: 1609 emäsparia
(B) Tyyppi: nukleiinihappo

(C) Juosteisuus: kaksoisjuoste
(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(v) Fragmenttityyppi: N-terminaalinen
(vi) Alkuperälähde:
(A) Organismi: Bacillus lautus
(B) Kanta: NCIMB 40250

Piirre:
(A) Nimi/Avainsana CDS
(B) Sijainti: 172-1869
(D) Muuta tietoa

(xi) Sekvenssin kuvaus: SEQ ID NO 6

```
TTGAAGCGCT GAATTCAGGA GGTTAAATA ATG CGT ATT CAT GCA ATT CGG CAA        53
                                Met Arg Ile His Ala Ile Arg Gln
                                  1               5

TCT TGC CGT TTG GTA TTG ACG ATG GTT TTG ATG CTT GGC TTA TTG CTG       101
Ser Cys Arg Leu Val Leu Thr Met Val Leu Met Leu Gly Leu Leu Leu
     10                  15                  20

CCT GTG GGC GCC CCC AAA GGC TAT GCC GCT CCG GCT GTT CCT TTT GGC       149
Pro Val Gly Ala Pro Lys Gly Tyr Ala Ala Pro Ala Val Pro Phe Gly
 25                  30                  35                  40

CAA TTA AAA GTT CAG GGC AAT CAA TTG GTA GGA CAG TCC GGG CAA GCT       197
Gln Leu Lys Val Gln Gly Asn Gln Leu Val Gly Gln Ser Gly Gln Ala
                 45                  50                  55

GTT CAA CTG GTT GGC ATG AGC TCA CAT GGA TTG CAG TGG TAT GGC AAT       245
Val Gln Leu Val Gly Met Ser Ser His Gly Leu Gln Trp Tyr Gly Asn
             60                  65                  70

TTC GTC AAC AAA TCG TCG TTG CAG TGG ATG AGA GAT AAC TGG GGC ATC       293
Phe Val Asn Lys Ser Ser Leu Gln Trp Met Arg Asp Asn Trp Gly Ile
         75                  80                  85

AAC GTC TTC CGT GCC GCT ATG TAT ACT TCC GAA GAT GGT TAC ATT ACG       341
Asn Val Phe Arg Ala Ala Met Tyr Thr Ser Glu Asp Gly Tyr Ile Thr
     90                  95                  100

GAT CCT TCC GTC AAG AAC AAG GTG AAG GAG GCG GTT CAG GCA TCC ATC       389
Asp Pro Ser Val Lys Asn Lys Val Lys Glu Ala Val Gln Ala Ser Ile
105                 110                 115                 120
```

```
GAT CTG GCC TTG TAC GTT ATT ATT GAC TGG CAT ATC TTG TCT GAT GGG      437
Asp Leu Ala Leu Tyr Val Ile Ile Asp Trp His Ile Leu Ser Asp Gly
                125                 130                 135

AAT CCG AAT ACG TAC AAA GCG CAA TCG AAA GCG TTC TTC CAA GAG ATG      485
Asn Pro Asn Thr Tyr Lys Ala Gln Ser Lys Ala Phe Phe Gln Glu Met
            140                 145                 150

GCC ACA TTG TAC GGC AAC ACG CCG AAT GTA ATC TAT GAA ATC GCG ACG      533
Ala Thr Leu Tyr Gly Asn Thr Pro Asn Val Ile Tyr Glu Ile Ala Thr
        155                 160                 165

AGC CCA ACG GAA TGT GTC CTG GGC AGA TGT CAG TCG TCG GAA GAA GTG      581
Ser Pro Thr Glu Cys Val Leu Gly Arg Cys Gln Ser Ser Glu Glu Val
    170                 175                 180

ATC ACG GCC ATT CGT TCG ATT GAC CCC GAC GGA GTG GTT ATC GTT GGC      629
Ile Thr Ala Ile Arg Ser Ile Asp Pro Asp Gly Val Val Ile Val Gly
185                 190                 195                 200

AGT CCA ACC TGG AGC CAG GAT ATT CAT CTG GCG GCC GAT AAC CCG GTA      677
Ser Pro Thr Trp Ser Gln Asp Ile His Leu Ala Ala Asp Asn Pro Val
                205                 210                 215

TCA CAT AGC AAT GTC ATG TAT GCG CTT CAT TTC TAT TCA GGC ACG CAT      725
Ser His Ser Asn Val Met Tyr Ala Leu His Phe Tyr Ser Gly Thr His
            220                 225                 230

GGA CAG TTT TTG AGA GAC CGA ATT ACC TAT GCG ATG AAC AAA GGA GCA      773
Gly Gln Phe Leu Arg Asp Arg Ile Thr Tyr Ala Met Asn Lys Gly Ala
        235                 240                 245

GCG ATC TTC GTT ACC GAG TGG GGC ACC AGT GAT GCA TCC GGG AAC GGC      821
Ala Ile Phe Val Thr Glu Trp Gly Thr Ser Asp Ala Ser Gly Asn Gly
    250                 255                 260

GGG CCG TAT TTG CCT CAG TCC AAA GAG TGG ATC GAT TTC TTG AAT GCT      869
Gly Pro Tyr Leu Pro Gln Ser Lys Glu Trp Ile Asp Phe Leu Asn Ala
265                 270                 275                 280

CGC AAG ATC AGC TGG GTG AAC TGG TCG CTC GCT GAT AAA GTA GAA ACG      917
Arg Lys Ile Ser Trp Val Asn Trp Ser Leu Ala Asp Lys Val Glu Thr
                285                 290                 295

TCT GCT GCT CTT ATG CCA GGT GCA TCG CCT ACC GGC GCT GGA CCG ATG      965
Ser Ala Ala Leu Met Pro Gly Ala Ser Pro Thr Gly Ala Gly Pro Met
            300                 305                 310

CCC AAT TGT CGA ATG GGC AAA TCG GGT TCG CGA TCA AAT CCG GCA AGC     1013
Pro Asn Cys Arg Met Gly Lys Ser Gly Ser Arg Ser Asn Pro Ala Ser
        315                 320                 325

AAC TGG AGG CGG CAG GGC AAT CCA ACT GCA CCG GCT GCC CCT ACT AAC     1061
Asn Trp Arg Arg Gln Gly Asn Pro Thr Ala Pro Ala Ala Pro Thr Asn
    330                 335                 340
```

```
CTC TCG GCA AAC GGC GGC AAC GCC CAG GTA TCA TTA ACC TGG AAC GCA      1109
Leu Ser Ala Asn Gly Gly Asn Ala Gln Val Ser Leu Thr Trp Asn Ala
345                 350                 355                 360

GTT AGC GGG GCG ACG AGC TAT ACC GTA AAG CGA GCA ACG ACG AGC GGC      1157
Val Ser Gly Ala Thr Ser Tyr Thr Val Lys Arg Ala Thr Thr Ser Gly
                365                 370                 375

GGT CCG TAC ACG AAT GTG GAC CGG GGT GTC ACG GCG ACG AGC TAC ACG      1205
Gly Pro Tyr Thr Asn Val Asp Arg Gly Val Thr Ala Thr Ser Tyr Thr
            380                 385                 390

AAC ACC GGG CTG ACG AAT GGC ACG ACG TAT TAT TAT GTC GTG AGG GCA      1253
Asn Thr Gly Leu Thr Asn Gly Thr Thr Tyr Tyr Tyr Val Val Arg Ala
        395                 400                 405

TCC AAT AGC GCG GGC AGC AGC GCG AAC TCC GCG CAA GCG AGC GCA ACG      1301
Ser Asn Ser Ala Gly Ser Ser Ala Asn Ser Ala Gln Ala Ser Ala Thr
    410                 415                 420

CCG GCT AGC GGC GGC GCC AGT ACG GGG AAC CTT GTT GTC CAA TAC AAA      1349
Pro Ala Ser Gly Gly Ala Ser Thr Gly Asn Leu Val Val Gln Tyr Lys
425                 430                 435                 440

GTT GGC GAC ACT AGC GCC ACG GAT AAC CAA ATG AAG CCT TCC TTT AAC      1397
Val Gly Asp Thr Ser Ala Thr Asp Asn Gln Met Lys Pro Ser Phe Asn
                445                 450                 455

ATC AAG AAC AAC GGT ACA ACC CCT GTT AAC CTG AGC GGC CTC AAG CTT      1445
Ile Lys Asn Asn Gly Thr Thr Pro Val Asn Leu Ser Gly Leu Lys Leu
            460                 465                 470

NNN NNN NNN NNN NAA AAA GAC GGA CCT GCG GAT ATG AGC TGC TCG ATC      1493
Xaa Xaa Xaa Xaa Xaa Lys Asp Gly Pro Ala Asp Met Ser Cys Ser Ile
        475                 480                 485

GAC TGG GCG CAA ATC GGC CGA ACG AAT GTT CTG CTG GCA TTC GCT AAC      1541
Asp Trp Ala Gln Ile Gly Arg Thr Asn Val Leu Leu Ala Phe Ala Asn
    490                 495                 500

TTT ACC GGG AGT AAT ACG GAT ACT TAC TGT TGT GAG CTA AGC TTT AGC      1589
Phe Thr Gly Ser Asn Thr Asp Thr Tyr Cys Cys Glu Leu Ser Phe Ser
505                 510                 515                 520

TGC ACT GCA GGT TCG TAT CCC GGC TAT GCG TGG AC                       1624
Cys Thr Ala Gly Ser Tyr Pro Gly Tyr Ala Trp
                525
```

(2) Sekvenssin ID NO 7 tiedot

(i) Sekvenssin karasteristiikka
(A) Pituus: 526 emäsparia
(B) Tyyppi: aminohappo
(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO 7

Met Arg Ile His Ala Ile Arg Gln Ser Cys Arg Leu Val Leu Thr Met
1 5 10 15

Val Leu Met Leu Gly Leu Leu Leu Pro Val Gly Ala Pro Lys Gly Tyr
20 25 30

Ala Ala Pro Ala Val Pro Phe Gly Gln Leu Lys Val Gln Gly Asn Gln
35 40 45

Leu Val Gly Gln Ser Gly Gln Ala Val Gln Leu Val Gly Met Ser Ser
50 55 60

His Gly Leu Gln Trp Tyr Gly Asn Phe Val Asn Lys Ser Ser Leu Gln
65 70 75 80

Trp Met Arg Asp Asn Trp Gly Ile Asn Val Phe Arg Ala Ala Met Tyr
85 90 95

Thr Ser Glu Asp Gly Tyr Ile Thr Asp Pro Ser Val Lys Asn Lys Val
100 105 110

Lys Glu Ala Val Gln Ala Ser Ile Asp Leu Ala Leu Tyr Val Ile Ile
115 120 125

Asp Trp His Ile Leu Ser Asp Gly Asn Pro Asn Thr Tyr Lys Ala Gln
130 135 140

Ser Lys Ala Phe Phe Gln Glu Met Ala Thr Leu Tyr Gly Asn Thr Pro
145 150 155 160

Asn Val Ile Tyr Glu Ile Ala Thr Ser Pro Thr Glu Cys Val Leu Gly
165 170 175

Arg Cys Gln Ser Ser Glu Glu Val Ile Thr Ala Ile Arg Ser Ile Asp
180 185 190

Pro Asp Gly Val Val Ile Val Gly Ser Pro Thr Trp Ser Gln Asp Ile
195 200 205

His Leu Ala Ala Asp Asn Pro Val Ser His Ser Asn Val Met Tyr Ala
210 215 220

Leu His Phe Tyr Ser Gly Thr His Gly Gln Phe Leu Arg Asp Arg Ile
225 230 235 240

Thr Tyr Ala Met Asn Lys Gly Ala Ala Ile Phe Val Thr Glu Trp Gly
245 250 255

Thr Ser Asp Ala Ser Gly Asn Gly Gly Pro Tyr Leu Pro Gln Ser Lys
260 265 270

Glu Trp Ile Asp Phe Leu Asn Ala Arg Lys Ile Ser Trp Val Asn Trp
275 280 285

Ser Leu Ala Asp Lys Val Glu Thr Ser Ala Ala Leu Met Pro Gly Ala
290 295 300

Ser Pro Thr Gly Ala Gly Pro Met Pro Asn Cys Arg Met Gly Lys Ser
305 310 315 320

Gly Ser Arg Ser Asn Pro Ala Ser Asn Trp Arg Arg Gln Gly Asn Pro
325 330 335

Thr Ala Pro Ala Ala Pro Thr Asn Leu Ser Ala Asn Gly Gly Asn Ala
340 345 350

Gln Val Ser Leu Thr Trp Asn Ala Val Ser Gly Ala Thr Ser Tyr Thr
355 360 365

Val Lys Arg Ala Thr Thr Ser Gly Gly Pro Tyr Thr Asn Val Asp Arg
370 375 380

Gly Val Thr Ala Thr Ser Tyr Thr Asn Thr Gly Leu Thr Asn Gly Thr
385 390 395 400

Thr Tyr Tyr Tyr Val Val Arg Ala Ser Asn Ser Ala Gly Ser Ser Ala
405 410 415

Asn Ser Ala Gln Ala Ser Ala Thr Pro Ala Ser Gly Gly Ala Ser Thr
420 425 430

Gly Asn Leu Val Val Gln Tyr Lys Val Gly Asp Thr Ser Ala Thr Asp
435 440 445

Asn Gln Met Lys Pro Ser Phe Asn Ile Lys Asn Asn Gly Thr Thr Pro
450 455 460

Val Asn Leu Ser Gly Leu Lys Leu Xaa Xaa Xaa Xaa Xaa Lys Asp Gly
465 470 475 480

Pro Ala Asp Met Ser Cys Ser Ile Asp Trp Ala Gln Ile Gly Arg Thr
485 490 495

Asn Val Leu Leu Ala Phe Ala Asn Phe Thr Gly Ser Asn Thr Asp Thr
500 505 510

Tyr Cys Cys Glu Leu Ser Phe Ser Cys Thr Ala Gly Ser Tyr Pro Gly
515 520 525

Tyr Ala Trp
530

International Application No: PCT/ /

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 3, line 20 of the description

A. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet ☐

Name of depositary institution

NATIONAL COLLECTION OF INDUSTRIAL & MARINE BACTERIA LTD.

Address of depositary institution (including postal code and country)

Torry Research Station, P.O. Box 31, 23 St Machar Drive, Aberdeen AB9 8DG, Scotland

Date of deposit	Accession Number
18 January 1990	NCIMB 40250

B. ADDITIONAL INDICATIONS (leave blank if not applicable). This information is continued on a separate attached sheet ☐

In respect of those designations in which a European patent is sought, a sample of the deposited micro-organism will be made available only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC) until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or is deemed to be withdrawn.

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☒ This sheet was received with the international application when filed (to be checked by the receiving Office)

Yvonne Jakobsen
(Authorized Officer)

Yvonne Jakobsen
Head Clerk

☐ The date of receipt (from the applicant) by the International Bureau was

(Authorized Officer)

Form PCT/RO/134 (January 1981)

Patenttivaatimukset

1. Entsyymi, jolla on sellulaasiaktiivisuutta, joka entsyymi on tuotettavissa Bacillus spp., NCIMB 40250, -kannalla tai sen kaltaisella Bacillus spp. -kannalla tai sanotun sellulaasin johdannaisella.

2. Patenttivaatimuksen 1 mukainen entsyymi, **tunnettu** siitä, että sen endoglukanaasiaktiivisuus on vähintään noin 10, edullisemmin vähintään noin 20, edullisimmin vähintäin noin 25, kuten noin 30, CMC-yksikköä (kuten tässä määritetään) kokonaisproteiinin milligrammaa kohti alkalisissa olosuhteissa.

3. Patenttivaatimuksen 1 mukainen entsyymi, **tunnettu** siitä, että se on aktiivinen lämpötilassa noin 65 °C:een saakka.

4. Jonkin patenttivaatimuksen 1-3 mukainen entsyymi, **tunnettu** siitä, että se on endoglukanaasi, jonka näennäismolekyylipaino on 75 kD, tai sen lohkontatuote, jolla on endoglukanaasiaktiivisuutta.

5. Patenttivaatimuksen 4 mukainen entsyymi, **tunnettu** siitä, että se on endoglukanaasi, joka on koodattu DNA-sekvenssillä, joka on esitetty oheistetussa sekvenssiluettelossa ID#1, tai sen muunnelmalla, joka koodaa sanotun endoglukanaasin johdannaista.

6. Jonkin patenttivaatimuksen 1-3 mukainen entsyymi, **tunnettu** siitä, että se on endoglukanaasi, jonka näennäismolekyylipaino on 56 kD, tai sen lohkontatuote, jolla on endoglukanaasiaktiivisuutta.

7. Patenttivaatimuksen 6 mukainen entsyymi, **tunnettu** siitä, että se on endoglukanaasi, joka on koodattu DNA-sekvenssillä, joka on esitetty oheistetussa sekvenssiluettelossa ID##, tai sen muunnelmalla, joka koodaa sanotun endoglukanaasin johdannaista.

8. Jonkin patenttivaatimuksen 1-3 mukainen entsyymi, **tunnettu** siitä, että se on endoglukanaasi, jonka näennäismolekyylipaino on 45 kD, tai sen lohkontatuote, jolla on endoglukanaasiaktiivisuutta.

9. Patenttivaatimuksen 8 mukainen entsyymi, **tunnettu** siitä, että se on endoglukanaasi, joka on koodattu DNA-sekvenssillä, joka on esitetty oheistetussa sekvenssiluettelossa ID#6, tai sen muunnelmalla, joka koodaa sanotun endoglukanaasin johdannaista.

10. Jonkin patenttivaatimuksen 1-3 mukainen entsyymi, **tunnettu** siitä, että se on endoglukanaasi, jonka näennäismolekyylipaino on 60 kD, tai sen lohkontatuote, jolla on endoglukanaasiaktiivisuutta, tai se on endoglukanaasi, jonka näennäismolekyylipaino on 56 kD, tai sen lohkontatuote, jolla on endoglukanaasiaktiivisuutta.

11. Jonkin patenttivaatimuksen 1-3 mukainen entsyymi, **tunnettu** siitä, että se on endoglukanaasi, jonka näennäismolekyylipaino on 92 kD, tai sen lohkontatuote, jolla on endoglukanaasiaktiivisuutta.

12. Entsyymi, **tunnettu** siitä, että se käsittää minkä tahansa patenttivaatimuksen 1-11 mukaisesta endoglukanaasista johdetun ydinalueen yhdistettynä selluloosaa sitovaan domeeniin, joka on johdettu toisesta sellulaasientsyymistä, tai ydinalueen, joka on johdettu toisesta sellulaasientsyymistä, yhdistettynä minkä tahansa patenttivaatimuksen 1-11 mukaisesta endoglukanaasista johdettuun selluloosaa sitovaan domeeniin.

13. Patenttivaatimuksen 12 mukainen entsyymi, **tunnettu** siitä, että ydinalue on johdettu sellulaasientsyymistä, joka ei käsitä luonnossa selluloosaa sitovaa aluetta.

14. DNA-rakenne, **tunnettu** siitä, että se käsittää DNA-sekvenssin, joka koodaa entsyymiä, jossa on sellulaasiaktiivisuutta, ja joka entsyymi on johdettavissa Bacillus spp.

kannasta NCIMB 40250 tai sen kaltaisesta Bacillus spp. -kannasta tai sanotun sellulaasin johdannaisesta.

15. Patenttivaatimuksen 14 mukainen DNA-rakenne, **tunnettu** siitä, että DNA-sekvenssi koodaa endoglukanaasia, jonka näennäismolekyylipaino on 75 kD.

16. Patenttivaatimuksen 15 mukainen DNA-rakenne, **tunnettu** siitä, että DNA-sekvenssi on jokin oheistetussa sekvenssiluettelossa ID#1 esitetty sekvenssi tai sen muunnelma, joka koodaa sanotun endoglukanaasin johdannaista.

17. Patenttivaatimuksen 14 mukainen DNA-rakenne, **tunnettu** siitä, että DNA-sekvenssi koodaa endoglukanaasia, jonka näennäismolekyylipaino on 56 kD.

18. Patenttivaatimuksen 17 mukainen DNA-rakenne, **tunnettu** siitä, että DNA-sekvenssi on jokin oheistetussa sekvenssiluettelossa ID#3 esitetty sekvenssi tai sen muunnelma, joka koodaa sanotun endoglukanaasin johdannaista.

19. Patenttivaatimuksen 14 mukainen DNA-rakenne, **tunnettu** siitä, että DNA-sekvenssi koodaa endoglukanaasia, jonka näennäismolekyylipaino on 45 kD.

20. Patenttivaatimuksen 19 mukainen DNA-rakenne, **tunnettu** siitä, että DNA-sekvenssi on jokin oheistetussa sekvenssiluettelossa ID#6 esitetty sekvenssi tai sen muunnelma, joka koodaa sanotun endoglukanaasin johdannaista.

21. Patenttivaatimuksen 14 mukainen DNA-rakenne, **tunnettu** siitä, että DNA-sekvenssi koodaa endoglukanaasia, jonka näennäismolekyylipaino on 92 kD.

22. Ekspressiovektori, **tunnettu** siitä, että siinä on minkä tahansa patenttivaatimuksen 14-21 mukainen insertoitu DNA-rakenne.

23. Solu, **tunnettu** siitä, että se on transformoitu minkä tahansa patenttivaatimuksen 14-21 mukaisella DNA-rakenteella tai patenttivaatimuksen 22 mukaisella ekspressiovektorilla.

24. Patenttivaatimuksen 23 mukainen solu, **tunnettu** siitä, että se on bakteeri.

25. Patenttivaatimuksen 24 mukainen solu, **tunnettu** siitä, että se on gram-positiivinen bakteeri.

26. Patenttivaatimuksen 25 mukainen solu, **tunnettu** siitä, että gram-positiivinen bakteeri on valittu ryhmästä, jonka muodostaa Bacillus subtilis, Bacillus licheniformis, Bacillus lentus, Bacillus brevis, Bacillus stearothermophilus, Bacillus alkalophilus, Bacillus amyloliquefaciens, Bacillus coagulans, Bacillus circulans tai Bacillus lautus.

27. Patenttivaatimuksen 24 mukainen solu, **tunnettu** siitä, että se on gram-negatiivinen bakteeri, esim. Escherichia coli.

28. Menetelmä jonkin patenttivaatimuksen 1-13 mukaisen entsyymin tuottamiseksi, **tunnettu** siitä, että jonkin patenttivaatimuksen 23-27 mukaista isäntäsolua viljellään olosuhteissa, jotka johtavat endoglukanaasin tai sen johdannaisen tuotantoon, ja endoglukanaasi tai sen johdannainen otetaan sen jälkeen talteen viljelmästä.

29. Patenttivaatimuksen 28 mukainen menetelmä, **tunnettu** siitä, että endoglukanaasi otetaan talteen kypsässä muodossa.

30. Selluloosaa hydrolysoiva aine, **tunnettu** siitä, että se kykenee hajottamaan selluloosakuitujen amorfiset alueet aineen käsittäessä minkä tahansa patenttivaatimuksen 1-3 mukaista entsyymiä.

31. Patenttivaatimuksen 30 mukainen aine, **tunnettu** siitä, että se käsittää kahden tai useamman minkä tahansa patenttivaatimuksen 1-13 mukaisen sellulaasin yhdistelmän, tai yhden tai useamman minkä tahansa patenttivaatimuksen 1-13 mukaisen sellulaasin yhdistelmän yhdessä yhden tai useamman muun entsyymin kanssa, jolla on sellulaasiaktiivisuutta.

32. Patenttivaatimuksen 30 tai 31 mukainen aine, **tunnettu** siitä, että se on pölyämättömän granulaatin, stabiloidun nestemäiseen tai suojatun entsyymin muodossa.

33. Jonkin patenttivaatimuksista 30-32 mukainen aine, **tunnettu** siitä, että sen endoglukanaasiaktiivisuus on 500-10000 CMC-endoaasiyksikköä grammaa ainetta kohti.

34. Jonkin patenttivaatimuksista 30-33 mukainen aine, **tunnettu** siitä, että se on pesuaineen lisäaine.

35. Patenttivaatimuksen 34 mukainen aine, **tunnettu** siitä, että se käsittää lisäksi toista entsyymiä, kuten proteaasia, lipaasia ja/tai amylaasia.

36. Pesuainekoostumus, **tunnettu** siitä, että se käsittää minkä tahansa patenttivaatimuksen 30-35 mukaista selluloosaa hydrolysoivaa ainetta.

37. Patenttivaatimuksen 36 mukainen pesuainekoostumus, **tunnettu** siitä, että sen endoglukanaasiaktiivisuus on 0,3-400 CMC-yksikköä grammaa pesuainetta kohti.

38. Menetelmä määrän vähentämiseksi, jona selluloosaa sisältävät tekstiilit tulevat karheiksi, tai selluloosaa sisältävien tektiilien karheuden vähentämiseksi, **tunnettu** siitä, että menetelmä käsittää selluloosaa sisältävien tekstiilien käsittelemisen minkä tahansa patenttivaatimuksen 30-35 mukaisella selluloosaa hydrolysoivalla aineella.

39. Patenttivaatimuksen 38 mukainen menetelmä, **tunnettu** siitä, että tekstiilien käsitteleminen selluloosaa hydroly-

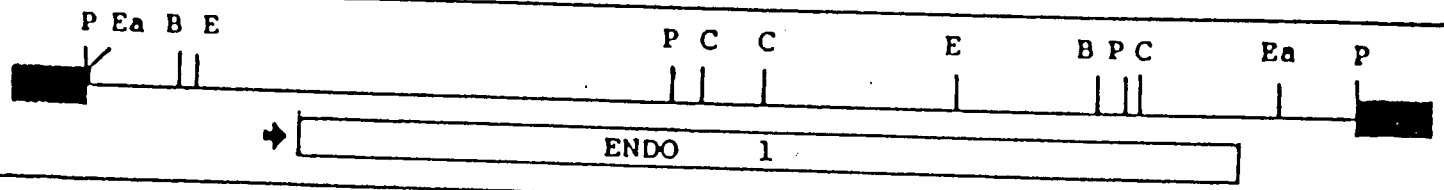
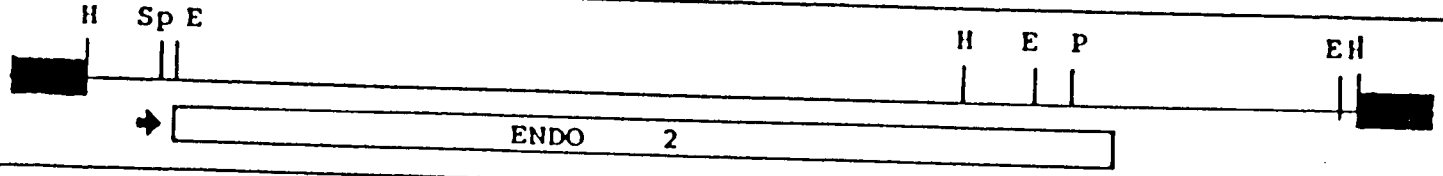
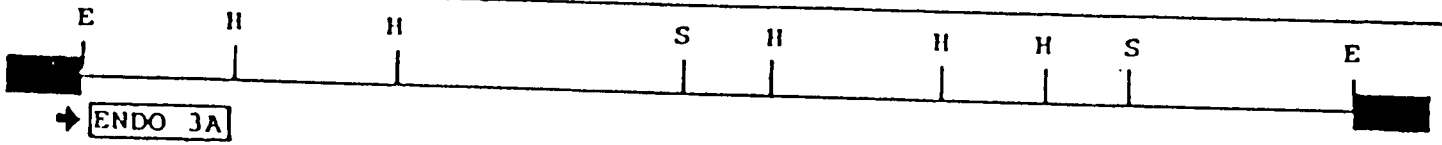
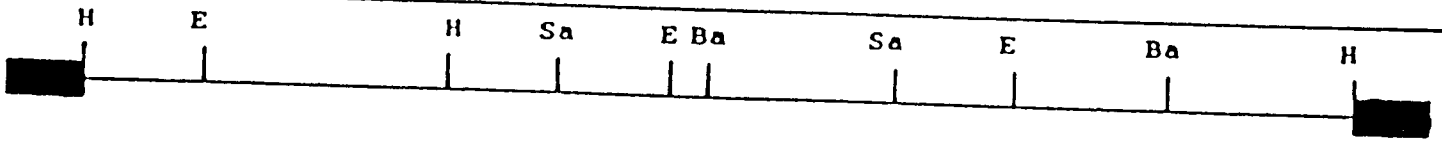
soivalla aineella suoritetaan tekstiilien liottamisen, pesemisen tai huuhtelemisen aikana.

40. Menetelmä värillisen, selluloosaa sisältävän tekstiilin käsittelemiseksi värien kirkastumisen aikaansaamiseksi, **tunnettu** siitä, että menetelmä käsittää selluloosaa sisältävän tekstiilin käsittelemisen minkä tahansa patenttivaatimuksen 30-35 selluloosaa hajottavalla aineella.

41. Patenttivaatimuksen 40 mukainen menetelmä, **tunnettu** siitä, että tekstiilin käsitteleminen selluloosaa hydrolysoivalla aineella suoritetaan vesiväliaineessa tekstiilin liottamisen, pesemisen tai huuhtelemisen aikana.

42. Patenttivaatimuksen 41 mukainen menetelmä, **tunnettu** siitä, että vesipitoisen väliaineen endoglukanaasiaktiivisuus on yli noin 250 CMC-endoaasiyksikköä litraa vesipitoista väliainetta kohti.

Fig. 1

Plasmidi	Insertti	
pPL 517	2750bp	
pPL 382	2500bp	
pPL 591	11000bp	
pPL 592	14000bp	

17·07·92 923275

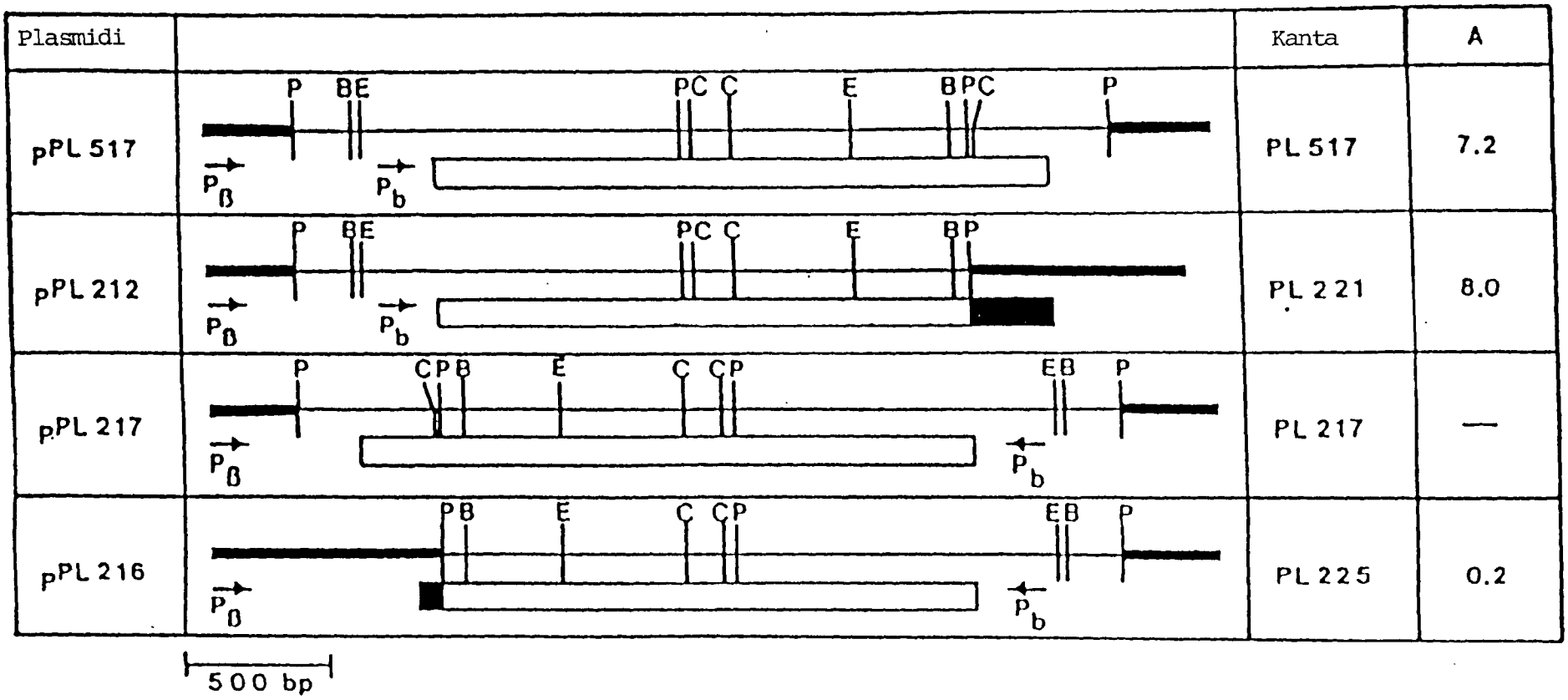


Fig. 2

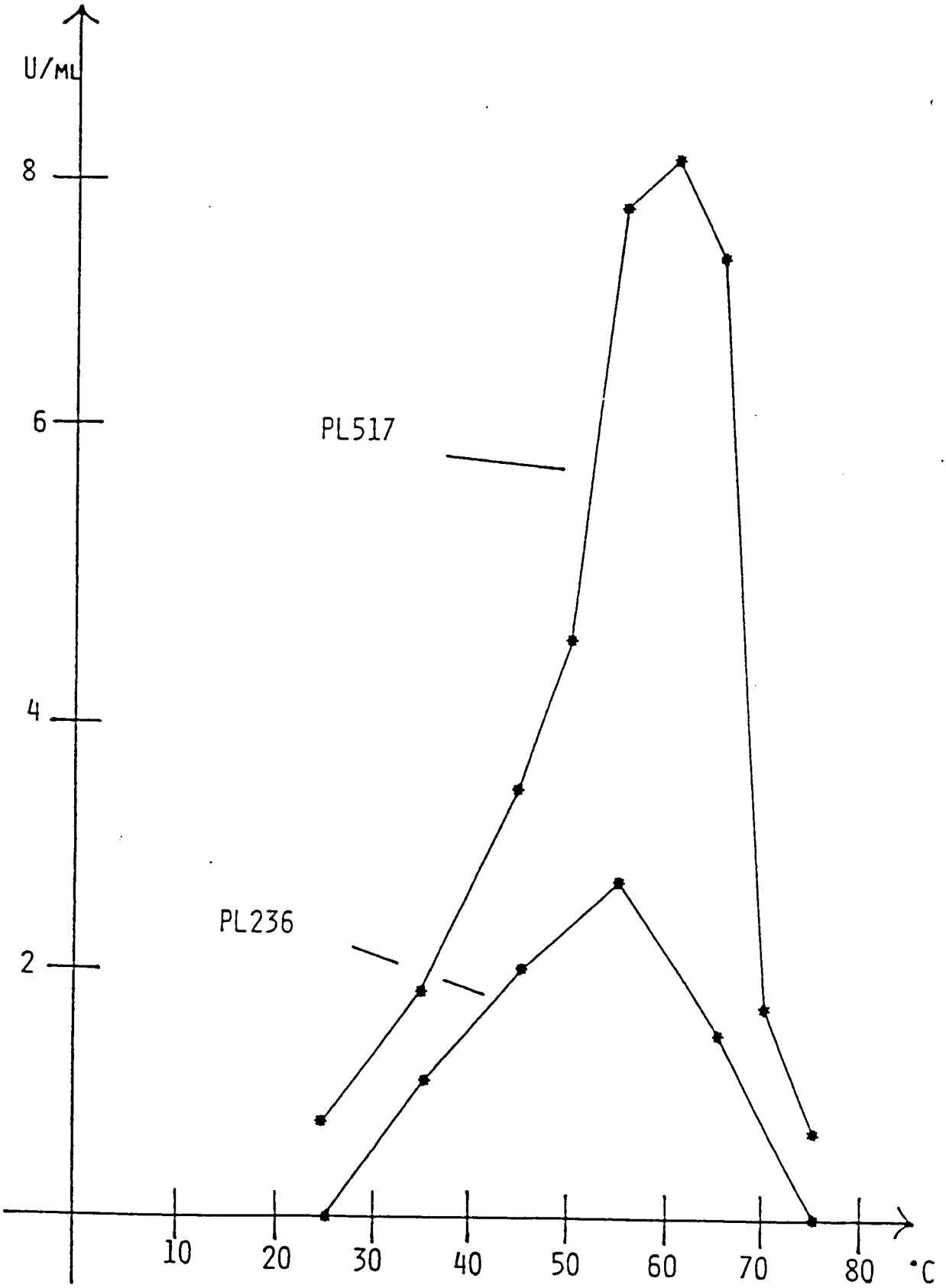


Fig. 3

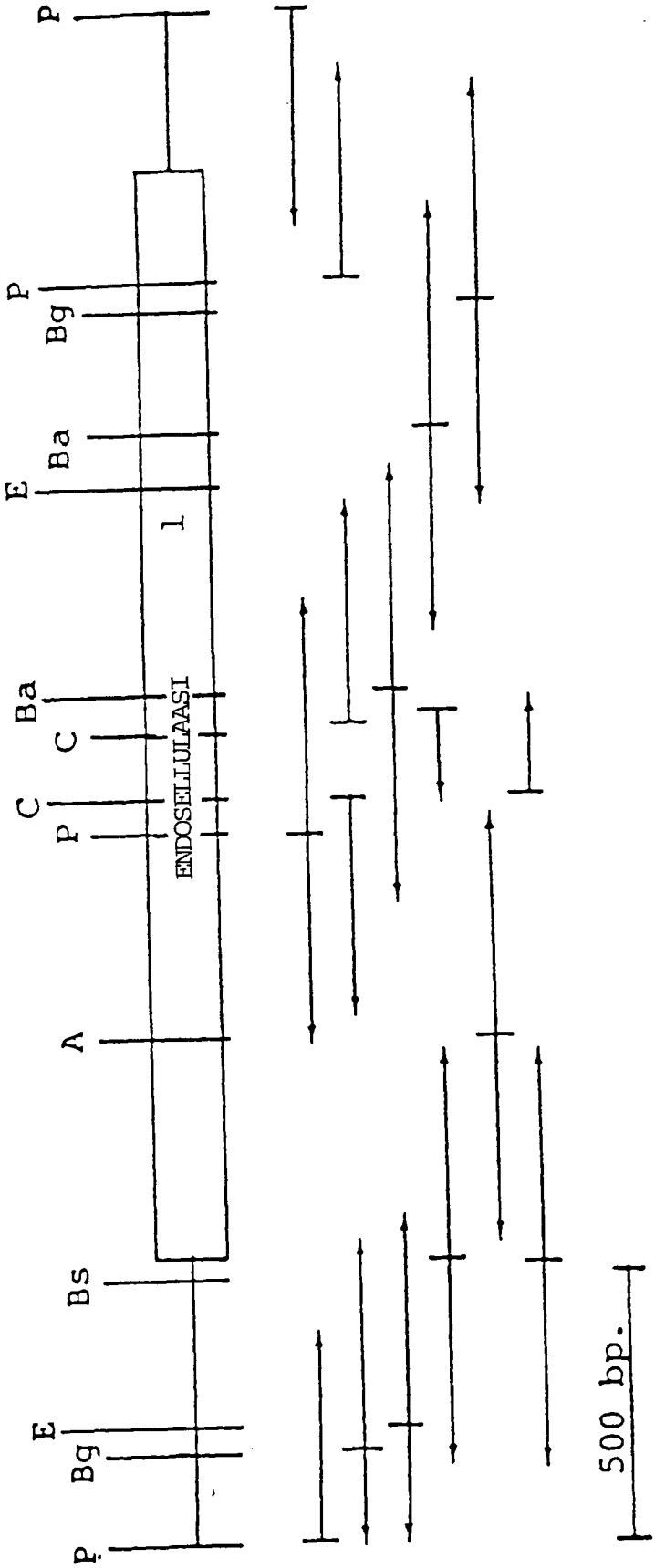


Fig. 4

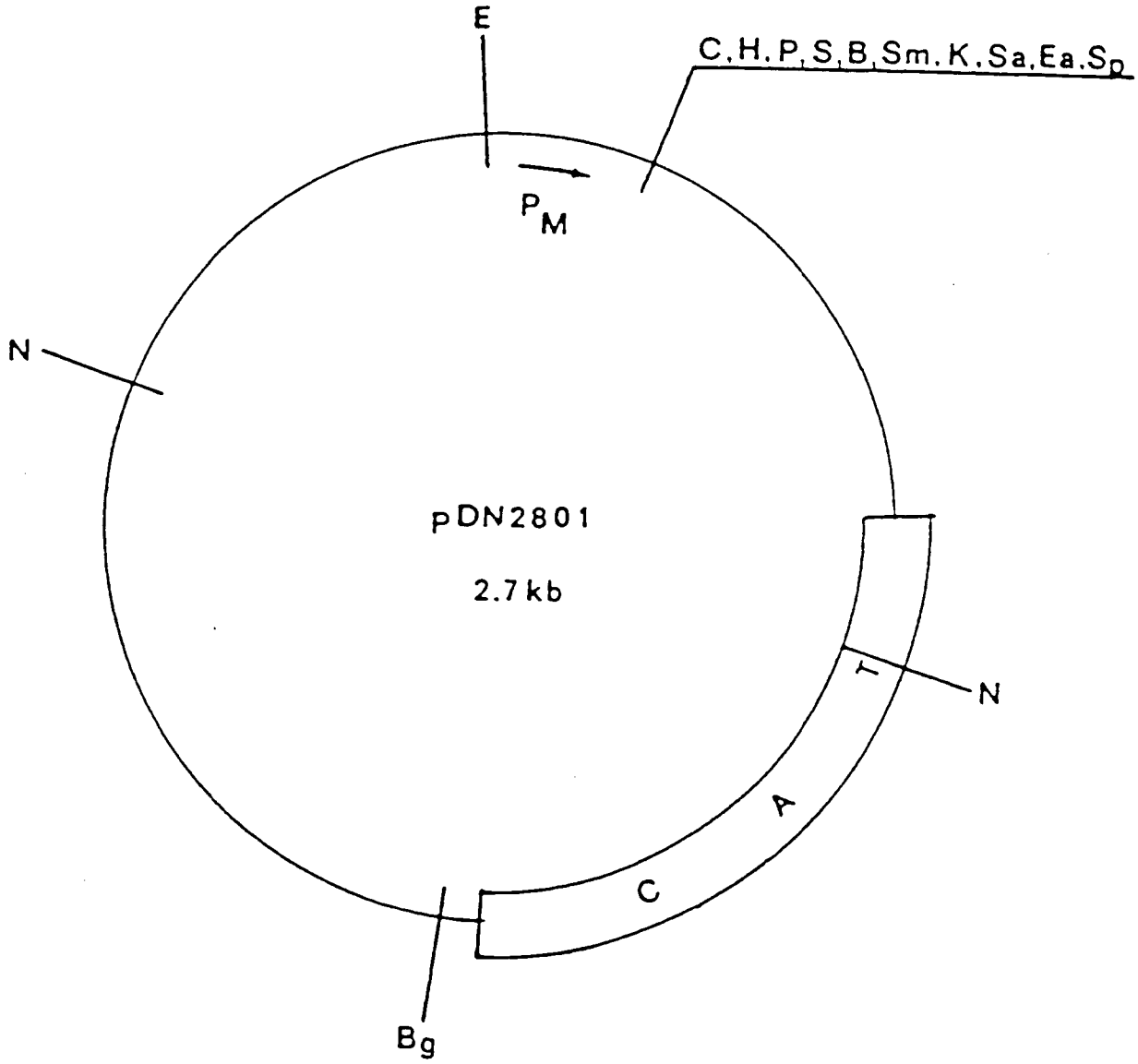
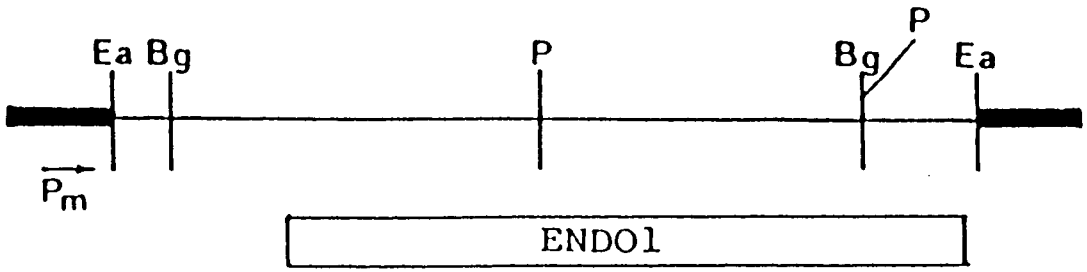
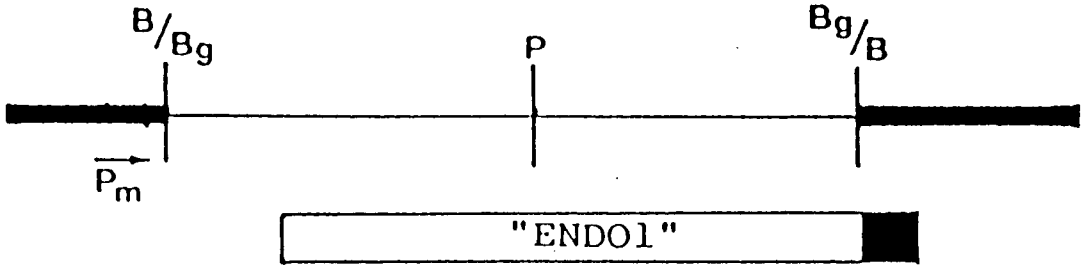


Fig. 5

17.07.92 92325

Plasmidi		Kanta	A
pCH 7		CH7	6,5
pCH14		CH14	0,3
—		DN1885	0,3

500 bp

Fig. 6

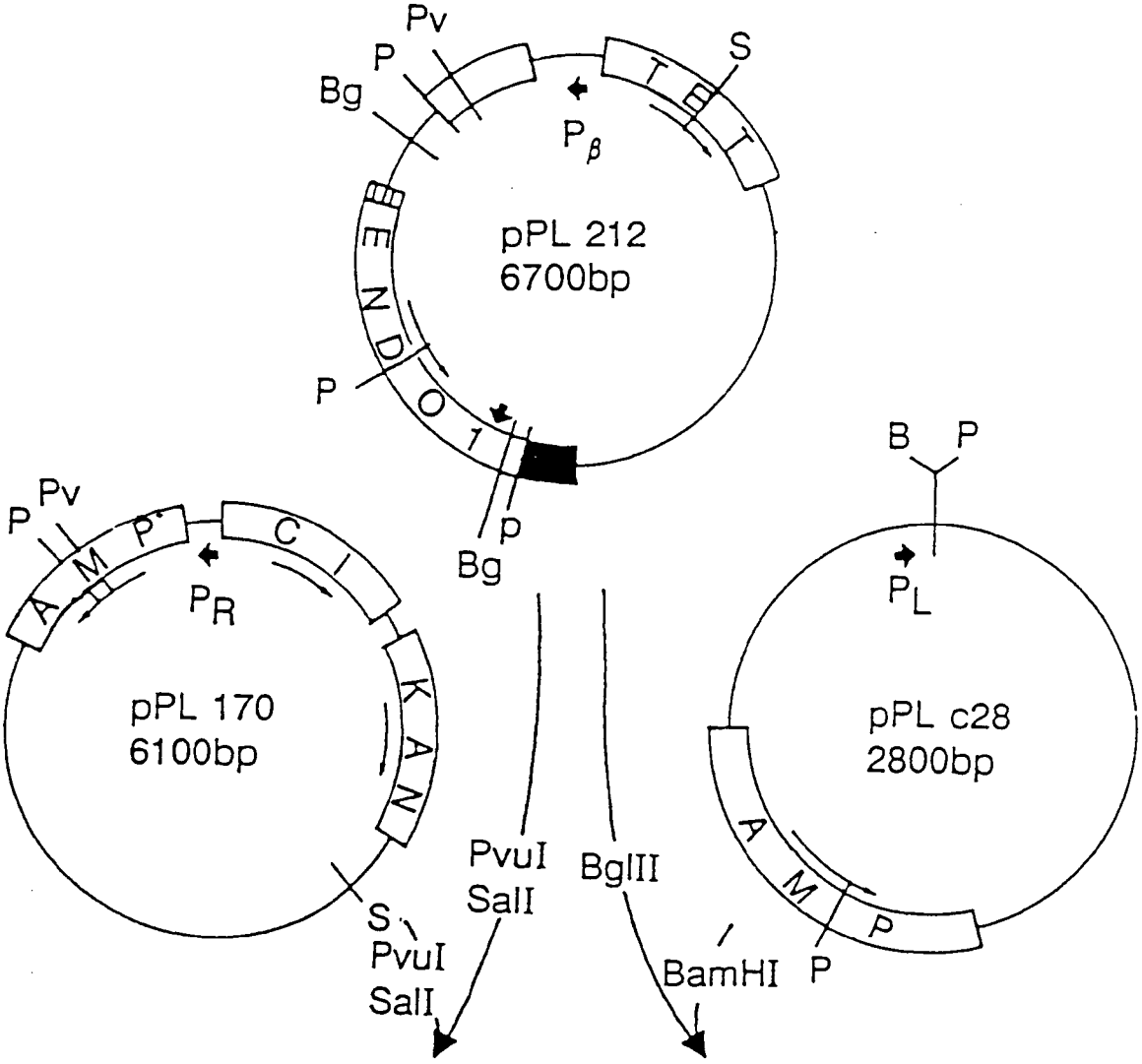

Pv
P
Bg
S
T
B
T
P_β
pPL 212
6700bp
B
E
N
D
O
1
P
Bg
p
P
Pv
P
A
M
P'
C
I
P_R
pPL 170
6100bp
K
A
N
S
PvuI
SalI
PvuI
SalI
BglII
BamHI
B
P
P_L
pPL c28
2800bp
A
M
P
P

Fig. 7a

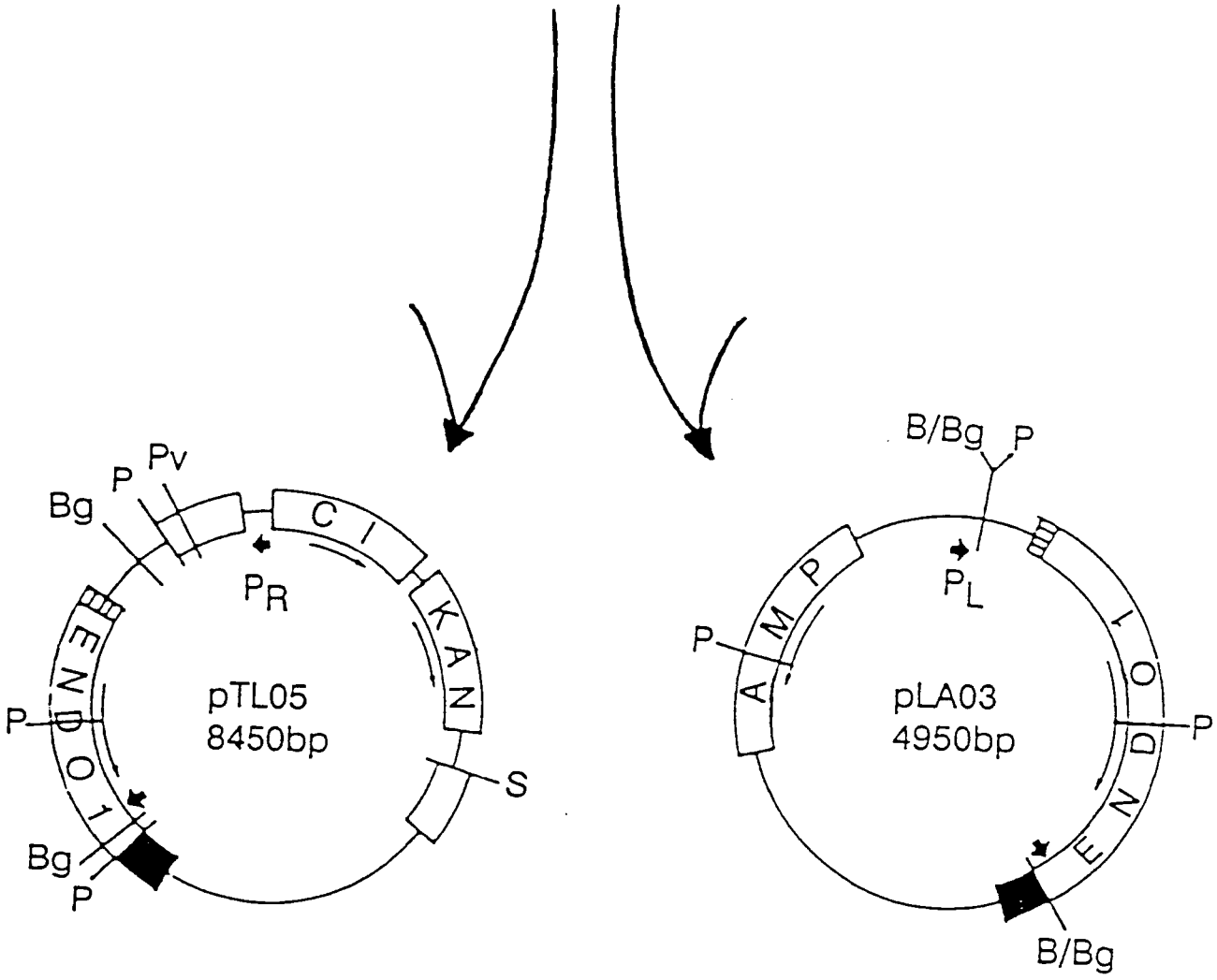

Pv
P
Bg
C I
KAN
PR
ENDO 1
P
pTL05
8450bp
S
Bg
P
B/Bg
P
AMP
PL
P
ENDO 1
pLA03
4950bp
P
B/Bg

Fig. 7b

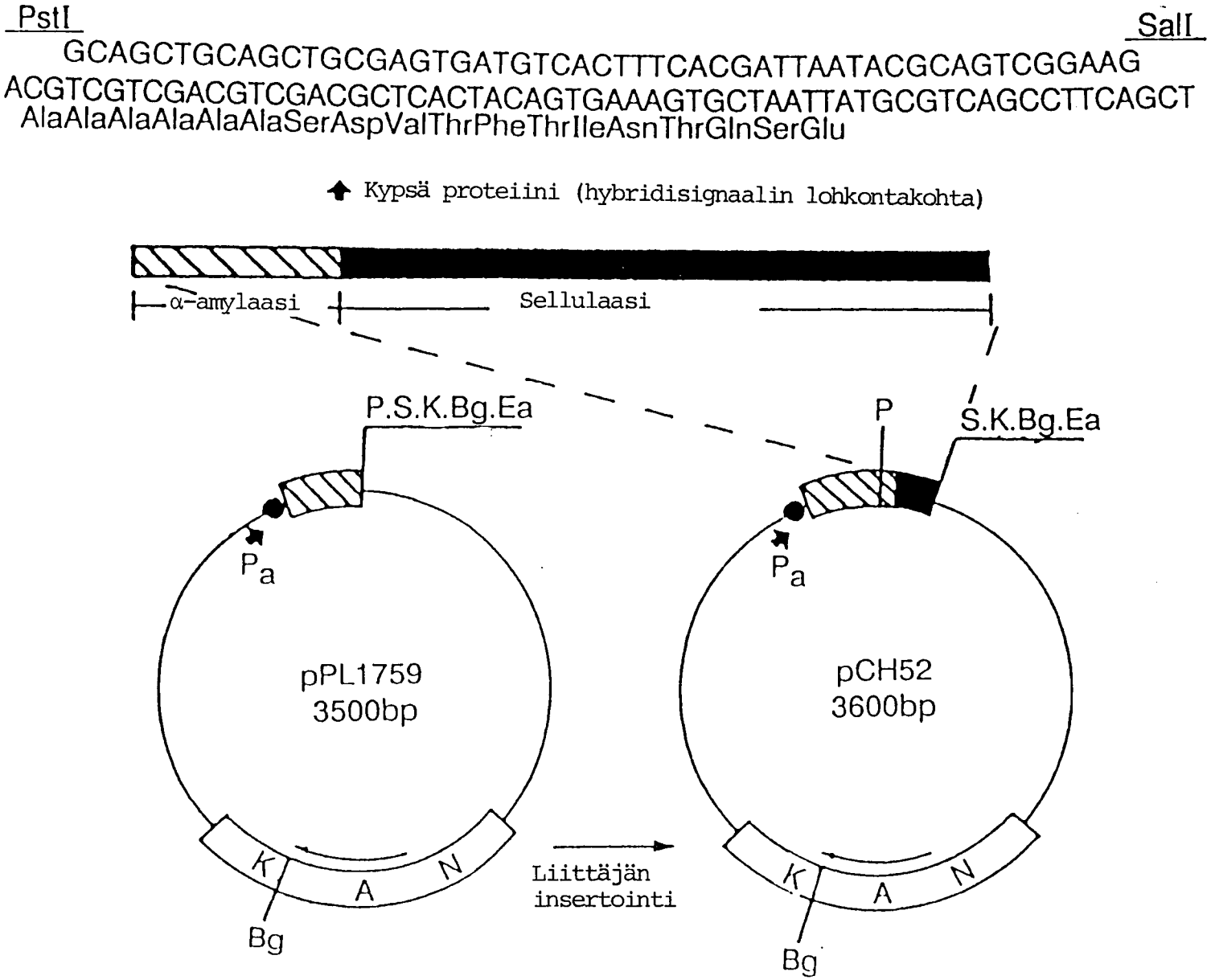
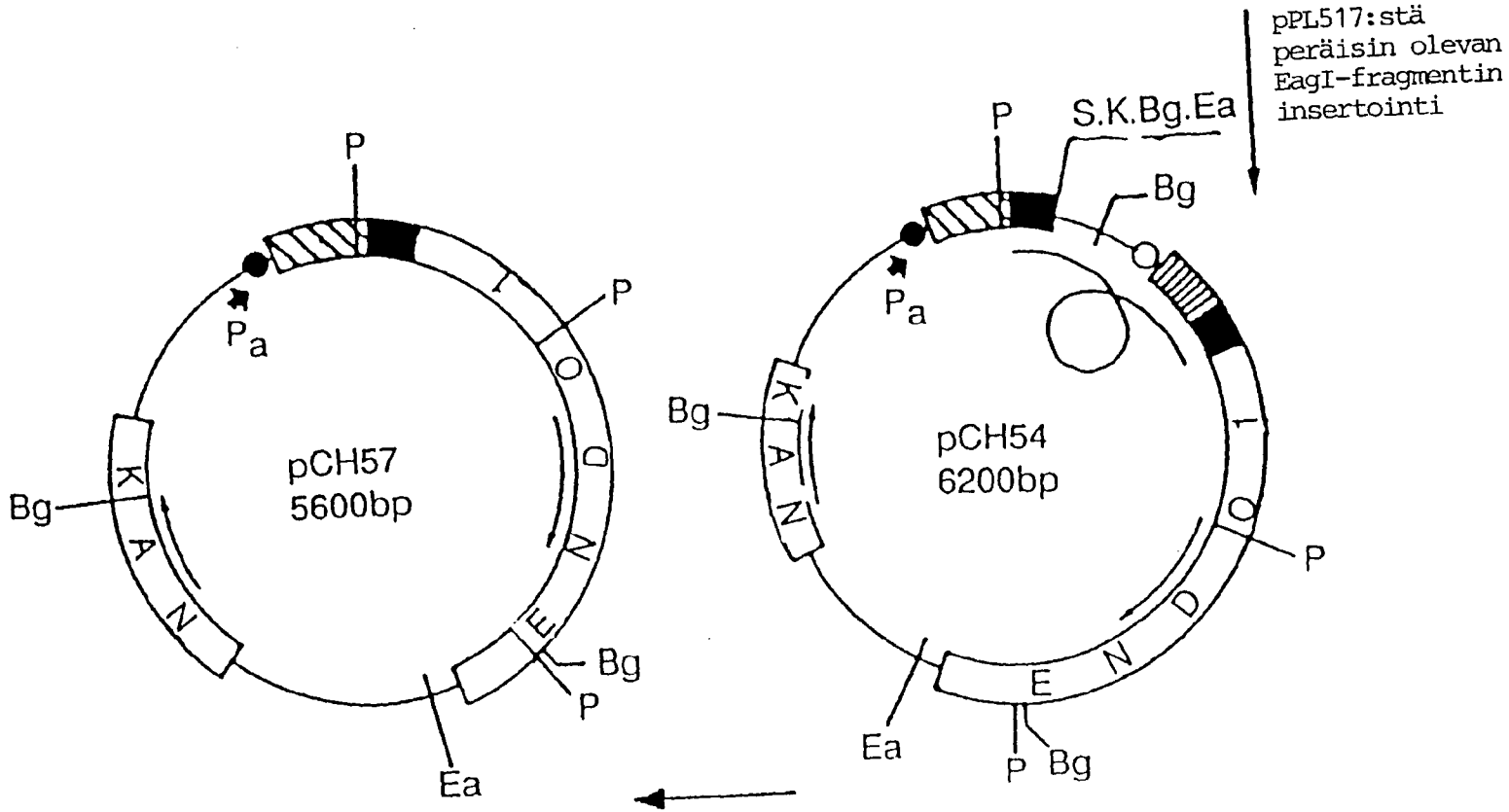

PstI
SalI
GCAGCTGCAGCTGCGAGTGATGTCACTTTCACGATTAATACGCAGTCGGAAG
ACGTCGTCGACGTCGACGCTCACTACAGTGAAAGTGCTAATTATGCGTCAGCCTTCAGCT
AlaAlaAlaAlaAlaAlaSerAspValThrPheThrIleAsnThrGlnSerGlu
Kypsä proteiini (hybridisignaalin lohkontakohta)
α-amylaasi
Sellulaasi
P.S.K.Bg.Ea
P
S.K.Bg.Ea
Pa
Pa
pPL1759
3500bp
pCH52
3600bp
K
A
N
Bg
Liittäjän
insertointi
K
A
N
Bg

Fig. 8a

Fig. 8b



1: In vivo-rekombinaatio

2: Rikastusleikkaus KpnI:llä

3: Transformointi

FI 92375

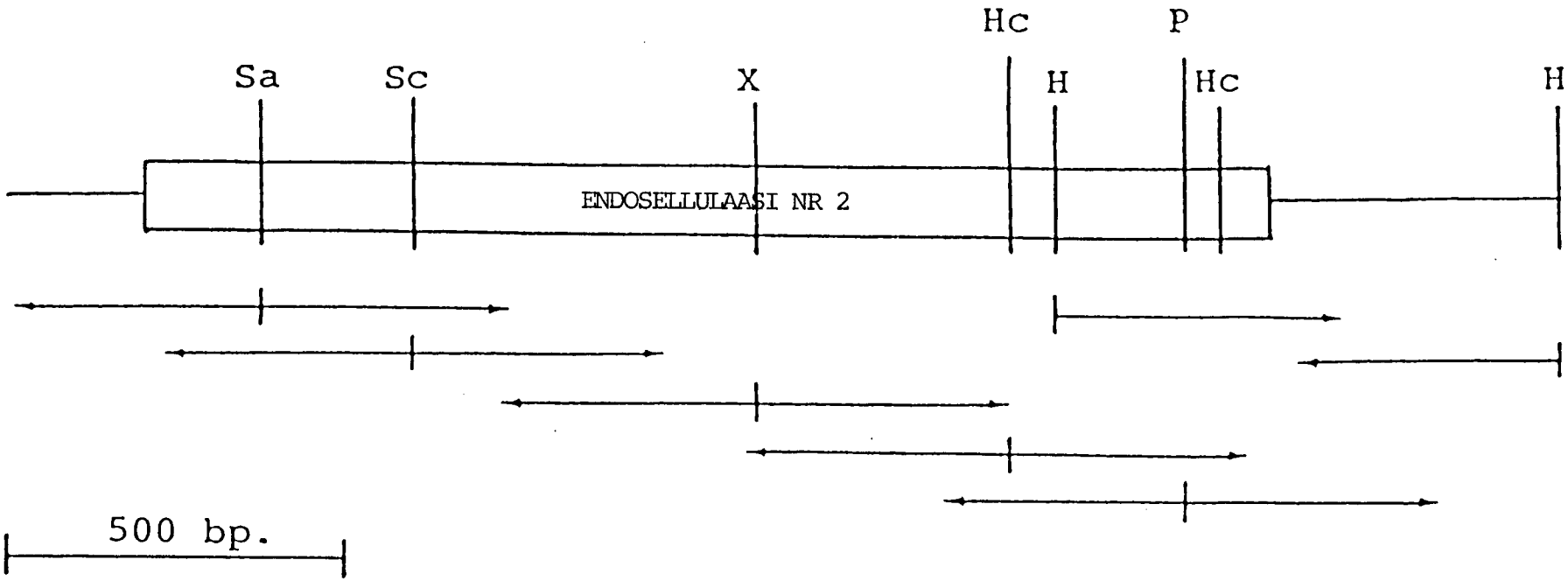


Fig. 9

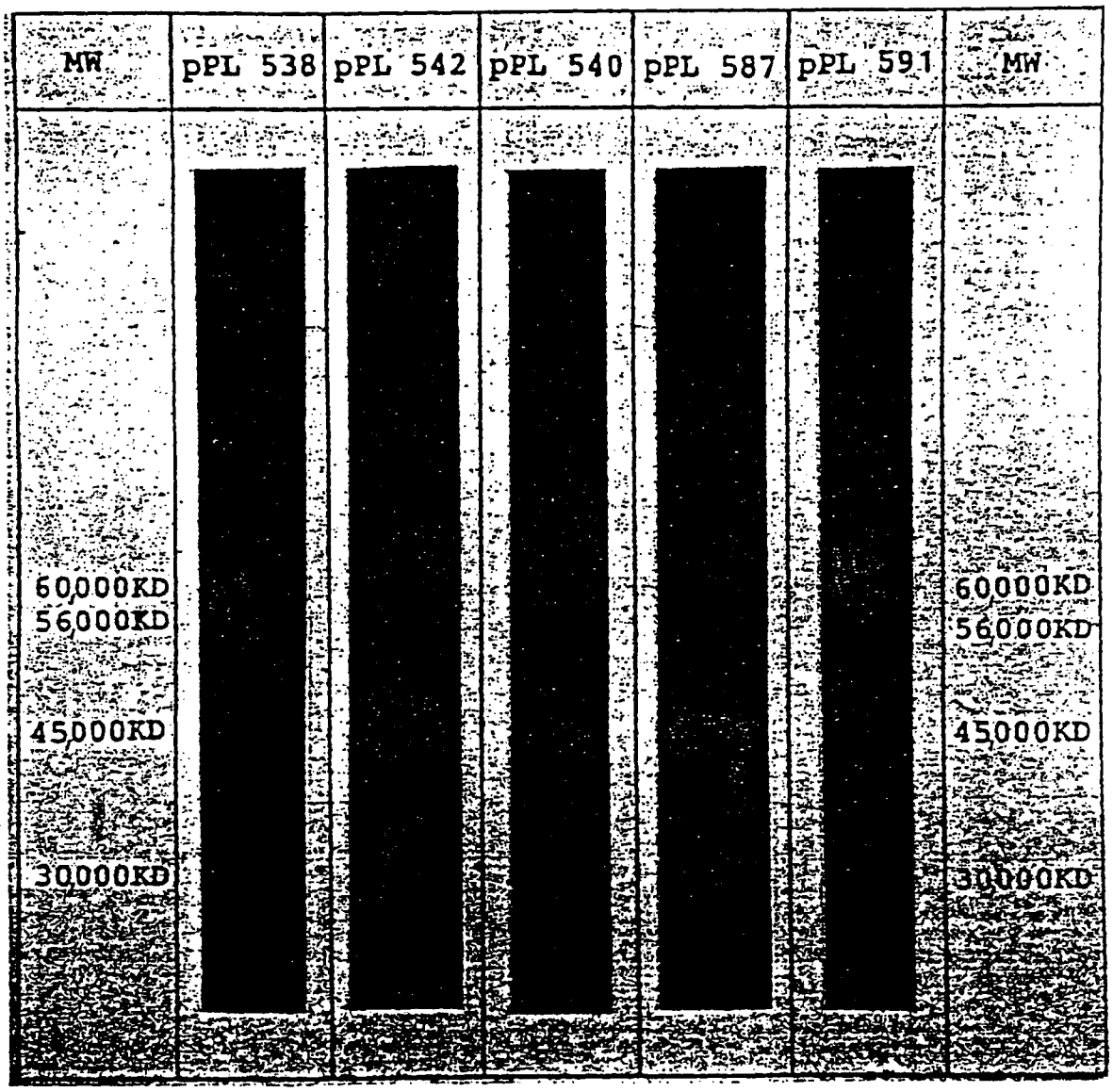

MW
pPL 538
pPL 542
pPL 540
pPL 587
pPL 591
MW
60,000KD
56,000KD
45,000KD
30,000KD
60000KD
56000KD
45000KD
30000KD

Fig. 10

17·07·92 923275

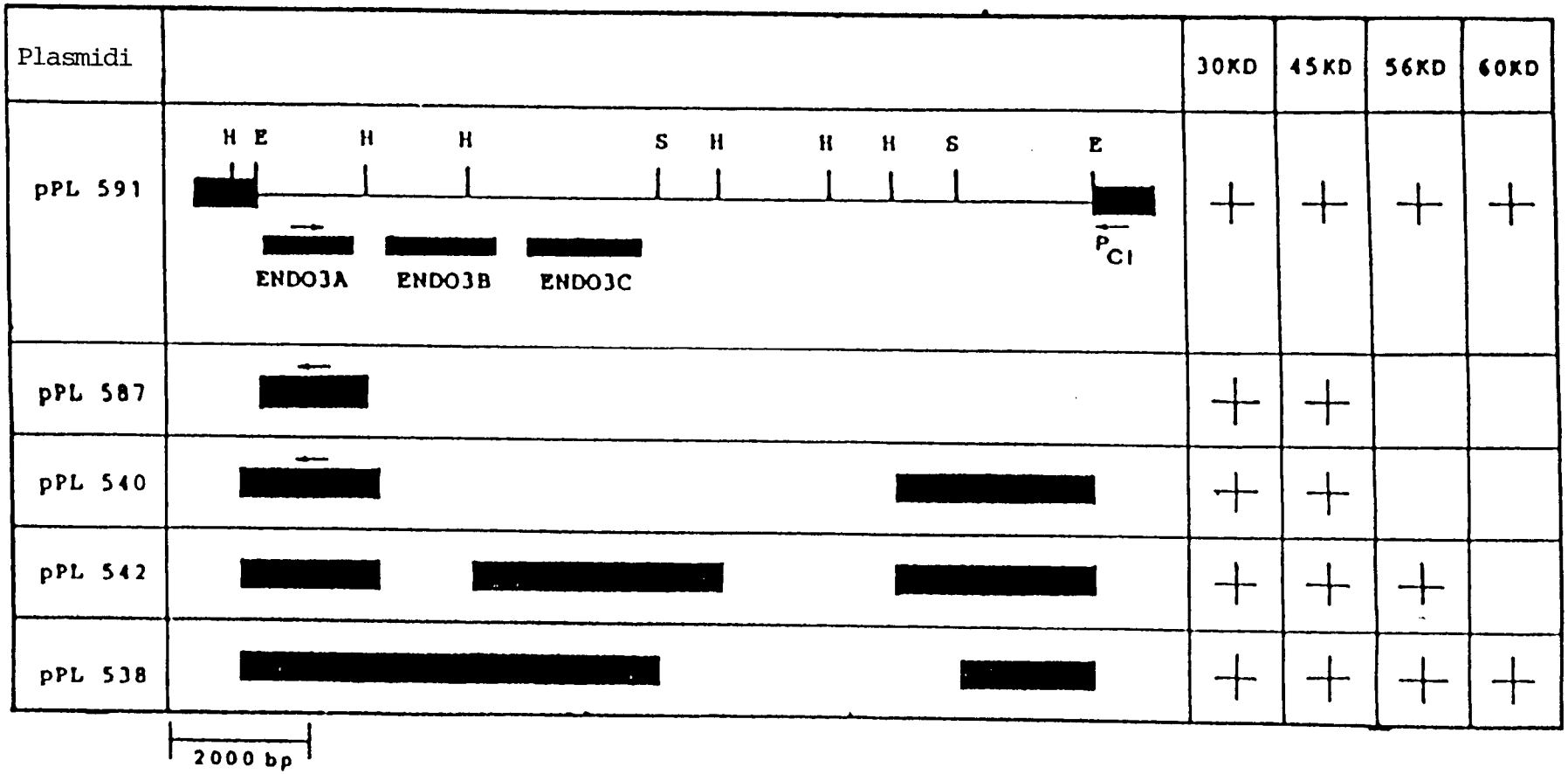


Fig. 11