

(19)



SUOMI - FINLAND

(FI)

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

(10) **FI 960401 A7**

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

(21) Patentihakemus - Patentansökan - Patent application **960401**

(51) Kansainvälinen patenttiluokitus - Internationell patentklassifikation -
International patent classification

C12N 15/53 (2006.01)

C12N 9/02 (2006.01)

C12P 19/62 (2006.01)

(22) Tekemispäivä - Ingivningsdag - Filing date **30.05.1994**

(23) Saapumispäivä - Ankomstdag - Reception date **29.01.1996**

(41) Tullut julkiseksi - Blivit offentlig - Available to the public **29.01.1996**

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

(86) Kansainvälinen hakemus - **30.05.1994 PCT/IB1994/000127**
Internationell ansökan - International
application

(32) (33) (31) Etu oikeus - Prioritet - Priority

30.07.1993 US 100518 P

(71) Hakija - Sökande - Applicant

1 • Pfizer Inc., 235 East 42nd Street, New York, NY 10017, AMERIKAN YHDYSVALLAT, (US)

(72) Keksijä - Uppfinnare - Inventor

1 • Denoya, Claudio D., Groton, CT 06340, AMERIKAN YHDYSVALLAT, (US)

(74) Asiamies - Ombud - Agent

Kolster Oy Ab, Salmisaarenaukio 1, 00180 Helsinki

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

**Streptomyces avermitiliksistä peräisin olevaa haarauneiden alfaketohappojen dehydrogenaasikompleksia koodaavia
geenejä**

Gener som kodar ett dehydrogenaskomplex av förgrenade alfaketosyror som härrör från Streptomyces avermitilis

(57) Tiivistelmä - Sammandrag - Abstract

Keksintö koskee DNA-sekvenssejä, jotka koodaavat Streptomyces-sukuun kuuluvan organismin haaroittuneiden alfa-ketohappojen dehydrogenaasikompleksia, ja näitä sekvenssejä ilmentämällä tuotettuihin polypeptideihin. Se koskee myös menetelmiä luonnollisen avermektiinin tuoton tehostamiseksi ja avermektiinin valmistamiseksi fermentoimalla.

Uppfinningen avser DNA sekvenser, vilka kodar förgrenade α -ketosyra dehydrogenas-komplex av organismen som hör till klassen Streptomyces, och polypeptider, vilka är framställda via expression av sådana sekvenser. Uppfinningen avser även förfaranden för att förbättra produktionen av naturlig avermektin och framställningen av avermektin via jäsnings.

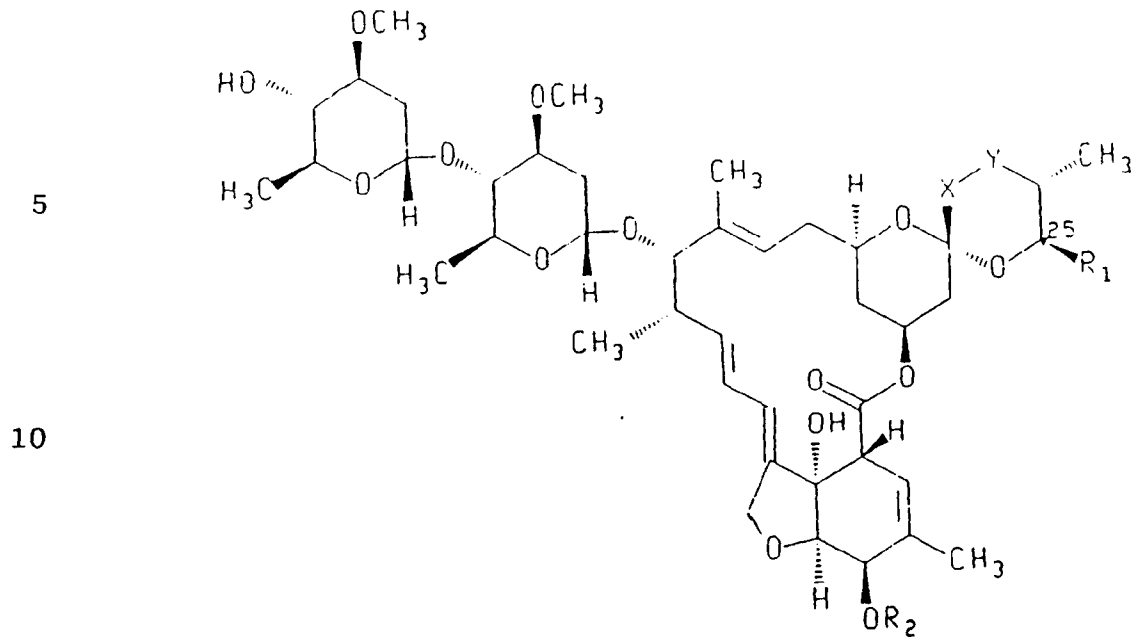
Streptomyces avermitiliksistä peräisin olevaa haarautuneiden alfa-ketohappojen dehydrogenaasikompleksia koodaavia geenejä

5 **Keksinnön tausta**

Keksintö liittyy uusiin DNA-sekvensseihin, jotka koodaavat Streptomyces-sukuun kuuluvasta organismista peräisin olevaa haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksia, ja näitä sekvenssejä ilmentämällä tuotettuihin uusiin polypeptideihin. Se liittyy myös uusiin menetelmiin luonnollisen avermektiinin tuotannon tehostamiseksi sekä uusien avermektiinien valmistamiseen fermentoimalla.

Mikro-organismit tuottavat useita farmaseuttisia tuotteita. Näistä mikro-organismeista on huomattavan suurta huomiota kiinnitetty Streptomyces-sukuun - tämä on ryhmä grampositiivisia maaperäbakteereja - siitä syystä, että yli 90 % terapeuttisesti käyttökelpoisista antibiooteista on peräisin tästä ryhmästä. Streptomykeettejä tutkitaan huomattavan voimakkaasti soveltaen yhdistelmä-DNA-kloonausmenetelmiä antibioottien biosynteettisten geenien eristämiseksi, uusien johdannaisten tai hybridiyhdisteiden aikaansaamiseksi, säätelygeenien eristämiseksi ja sekä primaari- että sekundaarimetaboliaan liittyvien säätelymekanismien tutkimiseksi.

S. avermitilis tuottaa kahdeksaa toisistaan poikkeavaa mutta toisiaan läheisesti muistuttavaa antiparaasiittista polyketidiyhdistettä, joita nimitetään avermektiineiksi. S. avermitiliksien tuottamassa avermektiinikompleksissa on neljä pääkomponenttia, A1a, A2a, B1a ja B2a ja neljä sivukomponenttia, Alb, A2b, B1b ja B2b. Seuraavaksi on esitetty useiden eri komponenttien rakenne.



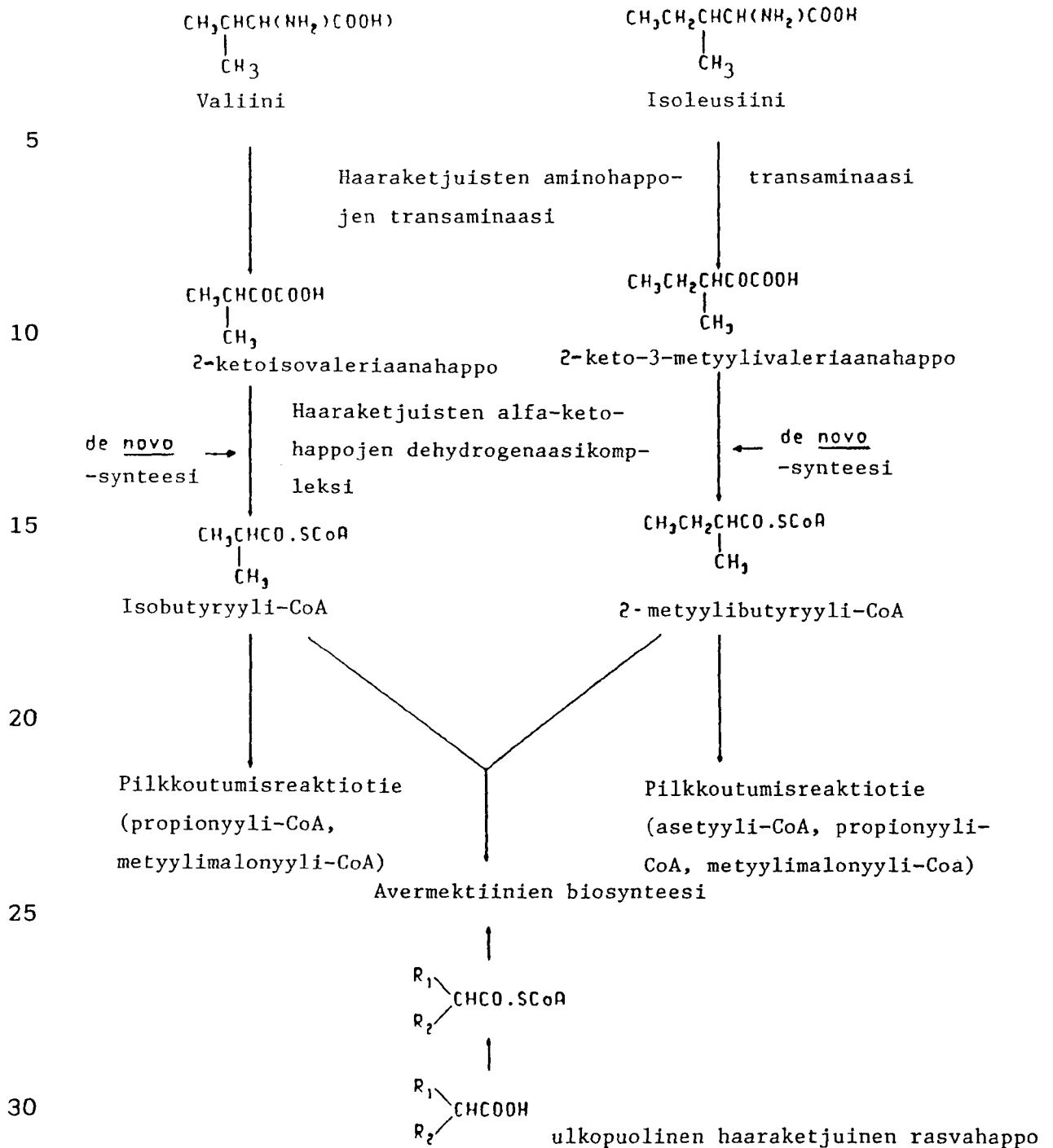
Avermektiini	R ¹	R ²	X-Y
Ala	sek-butyyl	Me	CH=CH
Alb	isopropyyl	Me	CH=CH
20 A2a	sek-butyyl	Me	CH ₂ -CH(OH)
A2b	isopropyyl	Me	CH ₂ -CH(OH)
B1a	sek-butyyl	H	CH=CH
B1b	isopropyyl	H	CH=CH
B2a	sek-butyyl	H	CH ₂ -CH(OH)
25 B2b	isopropyyl	H	CH ₂ -CH(OH)

Avermektiinin polyketidirakenne on peräisin seitsemästä asetaattimolekyylistä, viidestä propionaattimolekyylistä ja yhdestä haaraketjuisesta alfa-rasvahappomolekyylistä, joka on joko S(+)-2-metyyllivoihappo tai isovoihappo. Merkinnot "A" ja "B" tarkoittavat avermektiinejä, joissa 5-asemassa oleva substituentti on metoksi tai hydroksi, vastaavassa järjestyksessä ilmoitettuna. Numero "1" tarkoittaa avermektiinejä, joissa asemassa 22 - 23 on kaksoissidos ja numero "2" avermektiinejä, joissa asemassa 22 on vety ja asemassa 23 on hydroksi. Lopuksi on mainittava,

että C-25:ssa on kaksi mahdollista substituenttia: avermektiini-"a"-sarjassa esiintyy sek-butyylisubstituentti (tämä on peräisin L-isoleusiinista) ja avermektiini-"b"-sarjassa esiintyy isopropyylisubstituentti (tämä on peräisin L-valiinista) (näitä kysymyksiä on tarkasteltu katsauksessa Fisher, M.H. ja Mrozik, H., 1984, "Macrolide Antibiotics", Academic Press, chapter 14).

"Luonnollisilla" avermektiineillä tarkoitetaan sellaisia S. avermitiloksen tuottamia avermektiinejä, joissa 25-asemassa oleva substituentti on, kuten edellä mainittiin, joko isopropyyli tai sek-butyyli. Avermektiinien, joissa 25-asemassa oleva ryhmä on jokin muu kuin isopropyyli tai sek-butyyli, katsotaan tässä keksinnössä olevan uusia eli luonnon avermektiineistä poikkeavia avermektiinejä.

Yksi metabolinen reaktiotie, joka johtaa näihin CoA-muodossaan oleviin haaraketjuisiin alfa-rasvahappoihin, kulkee haaraketjuisten aminohappojen transaminaasin katalysoiman reaktion kautta, jonka jälkeen tapahtuu haaraketjuisten alfa-ketohappojen dehydrogenaasin katalysoima reaktio. (Vaihtoehtoisesti haaraketjuiset rasva-asyyli-CoA-johdannaiset voivat syntyä de novo -synteesin kautta muodostuneista haaraketjuisista alfa-ketohapoista). Nämä metaboliset reaktiotiet on kuvattu seuraavaksi.



Aikaisemmin on eristetty *S. avermitiliks*en mutantti, jossa täällä viimeksi mainitulla entsyymillä ei ole yhtään havaittavaa haaraketjuisten alfa-ketohappojen dehydrogenaasi (BCKDH) -aktiivisuutta (Hafner, et al., 1988, EP-patenttihakemus nro 88 300 353.5, julkaisu O 284 176). Mutantti oli eristetty *S. avermitilis* -kannalle ATCC 31272 tehdyn tavallisen kemiallisen mutageneesin jälkeen suoritetussa seulonnassa, jossa haettiin mutantteja, jotka eivät kyenneet muodostamaan $^{14}\text{CO}_2$:ta C-1-asemasta radioaktiivisella ^{14}C -leimalla varustetusta 2-oksoisokapriinihappo-substraatista (leusiinianalogi). Tämä mutantti ei kykene syntetisoimaan luonnollisia avermektiinejä muulloin kuin silloin, kun mediumiin, jossa mutantteja fermentoidaan, on lisätty haaraketjuista alfa-rasvahappoa tai tämän prekursoria, joka sisältää isopropyyli- tai sek-butyyliryhmän (S-muoto). Mutantti kykenee myös tuottamaan uusia eli luonnon avermektiineistä poikkeavia avermektiinejä fermentoitaessa sitä vesipitoisessa liuoksessa aerobisissa olosuhteissa ravintoliuoksessa, joka sisältää asiaankuuluvaa vaihtoehtoista karboksyylihappoa, kuten sykloheksaanikarboksyylihappoa (CHC) tai tämän prekursoria.

*S. avermitiliks*en BCKDH:n kloonaaminen on erittäin haluttua. Näiden geenien muokkaamisen yhdistelmä-DNA-menetelmillä pitäisi edistää luonnollisten ja uusien avermektiinien valmistusta. Tiettyjen kantojen kyseessä ollessa voitaisiin ajatella luonnollisten avermektiinien tiittereiden kohoavan BCKDH-geenien lukumäärää lisäämällä. Lisäksi sellaisen palautumattomasti inaktivoidun *bkd*-kannan valmistus, jossa BCKDH-aktiivisuus on pysyvästi deletoitu tai modifioitu käyttämällä geenien korvausta, olisi parempi vaihtoehto sille nykyisin käytössä olevalle *bkd*-mutantille, joka, kuten edellä mainittiin, on saatu kemiallista mutageneesiä käyttäen.

Monientsyymiset alfa-ketohappojen dehydrogenaasi-kompleksit - haaraketjuisten alfa-ketohappojen dehydroge-

naasi (BCKDH) -kompleksi, pyruvaattidehydrogenaasi (PDH) -kompleksi ja alfa-ketoglutaraattidehydrogenaasi (KGDH) -kompleksi - katalysoivat haaraketjuisten alfa-ketohappojen, pyruvaatin ja alfa-ketoglutaraatin hapettavia dekarboksylaatioreaktioita, tässä järjestyksessä mainittuna, ja niissä vapautuu CO₂:ta ja syntyy vastaava asyyli-CoA-molekyyli ja NADH:ta (Perham, R.N., Biochemistry 30 (1991) 8501 - 8512). Kukin kompleksi koostuu kolmesta erilaisesta katalyyttisestä entsyymistä: dekarboksylaasista (E1), dihydrolipoamidiasyyliitransferaasitransasylaasista (E2) ja dihydrolipoamididehydrogenaasista (E3).

Haaraketjuisten alfa-ketohappojen dehydrogenaasi (BCKDH) on monientsyymikompleksi, joka koostuu kolmesta toiminnallisesta komponentista, E1:stä, joka on dekarboksylaasi, E2:sta, joka on transasylaasi, ja E3:sta, joka on lipoamididehydrogenaasi. *Pseudomonas putidasta*, *Pseudomonas aeruginosasta* ja *Bacillus subtiliksesta* peräisin olevat puhdistetut kompleksit koostuvat neljästä polypeptidistä. Puhdistetut nisäkkäiden kompleksit koostuvat myös neljästä polypeptidistä, E1-alfasta, E1-beetasta, E2:sta ja E3:sta. *Bacillus subtiliksesta* on eristetty alfa-ketohappodehydrogenaasikompleksi, jossa on sekä pyruvaattidehydrogenaasiaktiivisuutta että haaraketjuisten alfa-ketohappojen dehydrogenaasiaktiivisuutta. Tämä kaksitoiminen kompleksi hapettaa pyruvaattia sekä muodostaa haaraketjuisia rasvahappoja membraanien fosfolipidejä varten.

Pseudomonasta ja *Bacillusta* koskevia prokaryoottien haaraketjuisten alfa-ketohappojen dehydrogenaasigeenien kloonausta käsitteleviä tutkimusraportteja on julkaistu, mutta näin ei ole asianlaita *Streptomyces*-suvun kohdalla. Näissä järjestelmissä on havaittu BCKDH:tä koodaavien geenien ryhmittyneen operoniksi. *Pseudomonas putidan* BCKDH-kompleksin geenit on kloonattu ja tämän alueen nukleotidisekvenssi on määritetty (Sykes, et al., J. Bacteriol. 169 (1987) 1619 - 1625 ja Burns, et al., Eur. J. Biochem.

176 (1988) 165 - 169, ja 176 (1988) 311 - 317). E1-alfan suhteellinen moolimassa on 45 289, E1-beetan on 37 138, E2:n on 45 134 ja E3:n on 48 164. Nämä neljä geeniä muodostavat ryhmittymän ja niiden järjestys on: E1-alfa, E1-beeta, E2 ja E3. Northern-blottianalyysi on osoittanut näiden neljän geenin ilmentymisen tapahtuvan yhdestä mRNA:sta ja näiden geenien muodostavan operonin. Välittömästi E1-alfan koodaavan alueen edellä on tyypillinen prokaryoottien konsensus-promoottori, joka mahdollistaa Pseudomonaksen bkd-geenien ilmentymisen konstitutiivisesti. E1-beetan koodaavan alueen aloituskodoni sijaitsee ainoastaan 40 nukleotidia myötäsuuntaan E1-alfan avoimen lukukehyksen (ORF) lopusta. Tästä poiketen E1-beetan ja E2:n ORF-sekvenssien välissä ei ole geenien välistä aluetta, koska E1-beetan ORF:n lopetuskodoni on E2:n ORF:n aloituskodonia välittömästi edeltävä tripletti. E2:n ja E3:n ORF-alueiden välinen geenien välitila on supistunut vain kahdeksi nukleotidiksi. Pseudomonaksen bkd-geenit ovat tämän vuoksi tiukasti kytkeytyneet. Bacillus subtiliksen BCKDH/PDH-kaksoiskompleksia koodaava operoni on samalla tavoin kloonattu (Hemila, et al., J. Bacteriol. 172 (1990) 5052 - 5063). Tämä operoni sisältää neljä ORF-aluetta, jotka koodaavat neljää 42, 36, 48 ja 50 kilodaltonin (kDa) suuruisia proteiinia, joilla on osoitettu olevan erittäin suurta homologiaa Pseudomonaksen bkd-ryhmittymän E1-alfa-, E1-beeta-, E2- ja E3-alayksikköjen suhteen. Bacillus stearothermophiluksesta peräisin olevan monientsyymisen BCKDH/PDH-kaksoiskompleksin E1-komponentin alfa- ja beeta-alayksiköjä koodaavat geenit on myös kloonattu ja sekvenssoitu (alfa- ja beeta-alayksikköjen arvioidut suhteelliset moolimassat ovat noin 41 000 ja 35 000, tässä järjestyksessä mainittuna) (Hawkins, et al., Eur. J. Biochem. 191 (1990) 337 - 346).

On myös lisäksi kuvattu useiden eukaryoottien BCKDH:n E1-alfa- ja -beeta-alayksikköjen sekvenssit (hu-

maani- ja boviinisekvenssit sekä rotan sekvenssi). Jonkin aikaa sitten on useista lajeista peräisin olevien PDH- ja BCKDH-kompleksien sekä EI-alfa- että EI-beeta-komponenttien kaikkia tunnettuja julkaistuja sekvenssejä analysoitu tietokoneella (Wexler, et al., FEBS Letters 282 (1991) 209 - 213). Oli kiinnostavaa, että tässä tunnistettiin useita alfa- ja beeta-alayksikköjen alueita, jotka eivät olleet säilyneet ainoastaan kaikissa tähän asti kuvatuissa PDH-komplekseissa vaan myös sekä prokaryoottisissa että eukaryoottisissa BCKDH-komplekseissa.

Tässä keksinnössä on kuvattu *Streptomyces avermiti-*liksesta peräisin olevien haaraketjuisten alfa-ketohappojen dehydrogenaasigeenien kloonaus. Nämä uudet geenit kloonattiin käyttäen kahden molekyyligeneettisen menetelmän, DNA-polymeraasiketjureaktion (PCR) ja koettimien käyttöön perustuvan homologian määrittämisen, yhdistelmää. Koettiin perustuva homologian määrittäminen sisältää cDNA-kirjastojen tai genomisten kirjastojen seulontaa radioaktiivisella leimalla varustetuilla synteettisillä oligonukleotidikoettimilla, jotka vastaavat proteiinin aminohapposekvenssejä. Tällä menetelmällä on valitettavasti tiettyjä rajoituksia, joista yksi on se, että oligonukleotidilla, jonka käyttö tässä menetelmässä on mahdollista, on degeneratiivisuutta rajoitettava melko voimakkaasti. Lisäksi seulontahybridisaatio suoritetaan heikosti rajoittavissa olosuhteissa, joten väärin positiivisten reaktioiden lukumäärä on usein korkea. Oligonukleotidihybridisaation joidenkin rajoitusten voittamiseksi on viime aikoina kehitetty koettimien käyttöön perustuvan homologian määrittämisen muunnelma, johon liittyy DNA-polymeraasiketjureaktio (PCR) ja jonka avulla on mahdollista käyttää koettimina erittäin degeneratiivisia oligonukleotideja. Tässä menetelmässä on vaatimuksena ainoastaan se, että sekvenssien koodaamasta proteiinista tunnetaan ainoastaan kaksi lyhyttä aminohapposekvenssialuetta (pituudeltaan

noin 7 - 10 aminohappoa). Reaktiossa käytetään alukkeen kahta oligonukleotidia, jotka vastaavat kumpaakin peptidisekvenssiä. Kutakin aluketta voidaan käyttää sellaisten täydellisten degeneratiivisten oligonukleotidien seoksena, joka sisältää kaikki sellaiset mahdolliset kodonien yhdistelmät, jotka saattaisivat koodata tunnettuja aminohapposekvenssejä. Mallijuosteena monistukseen voi olla mikä tahansa useista erilaisista DNA-lähtömateriaaleista, mukaan lukien genominen DNA ja plasmidikirjastojen superkierteiset muodot. Useissa tutkimusraporteissa, jotka on viime aikoina julkaistu alan ammattikirjallisuudessa, on osoitettu polymeraasiketjureaktion ja koettimien käyttöön perustuvan homologian määrittämisen yhdistelmän käyttökelpoisuus useista eri lajeista peräisin olevan geenin tunnistamiseen.

Sanasto

Tässä patenttihakemusjulkaisussa käytetyt tekniset termit ovat molekyylibiologian alan ammattikokemuksen perusteella tunnettuja. Näiden termien määritelmät on esitetty useissa molekyylibiologian alan oppikirjoissa, kuten "Genes", Second Edition, Dr. Benjamin Lewin, 1985, John Wiley & Sons, Inc. New York. Seuraavassa luettelossa on esitetty tässä keksinnössä usein käytetyt termit:

Antibiootti: kemiallinen aine, joka estää bakteerisolujen kasvun. Näitä käytetään yhdistelmä-bakteerisolujen selektioon.

Antibioottiresistenssigeeni: DNA-sekvenssi, joka saa aikaan resistenssin antibioottia kohtaan, silloin kun tämä sekvenssi on tuotu isäntäsoluun, joka on luontaisesti herkkä tätä tiettyä nimenomaista antibioottia kohtaan. Tästä käytetään myös nimitystä antibioottimarkkeri.

Bakteriofagit: bakteereja infektoivat virukset

cRNA: yksijuosteinen RNA, joka on komplementaarinen DNA:n suhteen ja joka syntetoidaan viimeksi mainitusta in vitro -olosuhteissa suoritettua transkription avulla.

Kromosomi: erillinen genomien yksikkö, joka sisältää useita geenejä.

Kloon: suuri määrä soluja tai molekyylejä, jotka ovat identtisiä yhden alkumuodon suhteen.

5 **Kloonausvektori:** mikä tahansa plasmidi, johon voidaan liittää kloonattavaa vierasta DNA:ta. Se vie transformaatiassa vieraan DNA:n isäntäbakteerisoluun.

CoA: koentsyymi A.

10 **Kohesiivinen päätesekvenssi (Cos):** bakteriofagi lambdasta peräisin oleva DNA-sekvenssi, joka mahdollistaa pakkaamisen in vitro -olosuhteissa.

Kosmidi: plasmidi, johon bakteriofagi lambdan Cos-kohdat on liitetty; tämän tuloksena plasmidi-DNA (joka sisältää vieraita DNA-liitossekvenssejä) voidaan pakata in vitro -olosuhteissa faagin vaippaan.

Dalton: massayksikkö, jota käytetään tavallisesti molekyylikokojen yhteydessä ja joka vastaa yhtä vetyatomeja.

20 **DNA:n ligaatio:** kahta DNA-fragmenttia yhdistävän kemiallisen sidoksen muodostuminen.

Eukaryoottisolut: korkeampien organismien soluja, jotka sisältävät membraanin ympäröimän tuman.

Geeniryhmittymä: sellaisten geenien ryhmä, jotka sijaitsevat fysikaalisesti toistensa lähellä kromosomissa.

25 **Genomi:** koko kromosomisto. Yksilön kaikkien geenien muodostama kokonaisuus.

Hybridisaatio, pesäkehybridisaatio: menetelmä, jota käytetään sellaisten kimeerisiä vektoreita sisältävien bakteeripesäkkeiden tunnistamiseen, joissa vektoriin liitetty DNA on samanlainen kuin jokin tietty nimenomainen sekvenssi.

ke: lyhenne, joka vastaa 1 000 emäsparia DNA:ta tai RNA:ta.

35 **NADH:** pelkistetty nikotiiniamidiadeniinidinukleotidi.

Kytkijäsekvenssi: lyhyt synteettinen oligodeoksinukleotididupleksi, joka sisältää yhden tai useamman restriktioentsyymin katkaisukohdan. Se lisätään vektoriin uuden polylinker-sekvenssin eli MCS-kloonauksen muodostamiseksi.

Nukleotidi: nukleiinihappojen rakenneyksikkö tai monomeeriyksikkö.

Oligonukleotidi: lyhyt nukleotidiketju.

Operoni: täydellinen bakteerien geenien ilmentymisen ja säätelyn yksikkö, johon sisältyvät DNA:ssa olevat rakennegeenit, säätelygeenit ja säätelygeeni(e)n tuotteen tunnistamat säätelyelementit.

Plasmidi: autonominen itsestään replikoituva kromosomin ulkopuolinen rengasmaisen DNA.

Plasmidien kappalemäärä: bakteerissa säilyvien plasmidimolekyylien lukumäärä jokaista isäntäkromosomia kohti.

Aluke: lyhyt DNA- tai RNA-sekvenssi, joka on liittynyt yhteen DNA-juosteeseen ja josta saadaan käyttöön vapaa 3'-OH-pää, josta DNA-polymeraasi aloittaa deoksiribonukleotidiketjun synteesin.

Prokaryoottisolut: pieniä, suhteellisen yksinkertaisia soluja, joihin suurin osa mikro-organismeista kuuluu.

Promoottori: DNA-alue, joka saa aikaan transkription alkamisen.

Restriktioentsyymi: entsyymi, joka tunnistaa spesifisen lyhyen DNA-sekvenssin ja pilkkoo sen.

Restriktioentsyymin tunnistussekvenssi: tietyn nimellisen restriktioentsyymin tunnistama DNA-sekvenssi. Tästä käytetään myös nimitystä vaikutuskohta.

Sukkulavektori: kaksitoiminen kloonauksvektori, joka kykenee replikoitumaan yhdessä tai useammassa vaihtoehdoissa isännässä (esimerkiksi E. colissa ja Streptomycesissä).

Southern-blottaus: menetelmä denaturoidun DNA:n siirtämiseksi agarosigeelistä nitroselluloosasuodattimelle, jossa se voidaan hybridisoida komplementaariseen nukleiinihappokoettiin.

5 **Jatkokloonaus:** kloonattujen DNA-fragmenttien siirto yhdentyypisistä vektorista toiseen, esimerkiksi yhdistelmäkosmidista plasmidiin. Uusi yhdistelmäplasmidi transformoidaan tämän jälkeen sopivaan isäntäsoluun jatkokloonikannan muodostamiseksi.

10 **Transkriptio:** RNA:n synteesi DNA-mallijuosteesta.

Bakteerisolujen transformaatio: tämä termi kuvaa uusien geneettisten markkerien saamista lisätyn DNA:n liittymisen avulla.

Keksinnön yhteenveto

15 Keksintö liittyy eristettyyn DNA-segmenttiin, joka koodaa Streptomyces-sukuun kuuluvan organismin haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksia.

20 Keksintö liittyy myös edellä kuvattua tyyppiä olevaan eristettyyn DNA-segmenttiin, joka lisäksi sisältää tällaisen haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksin ilmentymistä säätelevän DNA-alueen.

Keksintö liittyy myös Streptomyces avermitiloksen haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksia koodaavaan eristettyyn DNA-segmenttiin.

25 Keksintö liittyy myös DNA-segmenttiin, joka sisältää tuonnempana olevan kuvauksen mukaisen SEKVENSSI ID NO:1:n, SEKVENSSI ID NO:2:n, SEKVENSSI ID NO:3:n, SEKVENSSI ID NO:4:n tai SEKVENSSI ID NO:5:n mukaisen DNA-sekvenssin tai tällaisen sekvenssin alleelisen variantin. Se liittyy myös DNA-segmenttiin, joka on edellä olevan DNA-segmentin alajoukko ja joka vastaa sitä toiminnallisesti.

30

Tämä keksintö liittyy myös: (a) yhdistelmä-DNA:han, joka sisältää SEKVENSSI ID NO:1:n, SEKVENSSI ID NO:2:n, SEKVENSSI ID NO:3:n, SEKVENSSI ID NO:4:n tai SEKVENSSI ID NO:5:n mukaisen DNA-sekvenssin tai tällaisen sekvenssin

35

alleelisen variantin; (b) plasmidiin, joka sisältää tällaisen yhdistelmä-DNA:n; ja (c) isäntäsoluun, johon tällainen yhdistelmä-DNA on liitetty.

5 Tämä keksintö liittyy myös haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksia vastaaviin geeneihin, jotka sisältyvät DNA-segmenttiin, joka on pCD528, pCD545, pCD574, pCD550, pCD559 ja pCD577, jotka ovat tuonempana olevan kuvauksen mukaiset.

10 Tämä keksintö liittyy myös menetelmään luonnollisen avermektiinin valmistamiseksi, joka menetelmä käsittää sen, että fermentoidaan tämän luonnollisen avermektiinin valmistamiseen soveltuviissa olosuhteissa ja siihen soveltuvassa fermentaatiomediumissa *S. avermitilistä*, jossa haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksia
15 koodaavien geenien kappalemäärää on lisätty.

Tämä keksintö liittyy myös menetelmään luonnollisen avermektiinin valmistamiseksi, joka menetelmä käsittää sen, että fermentoidaan tämän luonnollisen avermektiinin valmistamiseen soveltuviissa olosuhteissa ja siihen soveltuvassa fermentaatiomediumissa *S. avermitilistä*, jossa
20 yhden tai useamman haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksia koodaavan geenin ilmentymistä on tehostettu muokkaamalla tai korvaamalla tällaisen ilmentymisen säätelystä vastaavia geenejä.

25 Tämä keksintö liittyy myös menetelmään uuden avermektiinin tuottamiseksi, joka menetelmä käsittää sen, että fermentoidaan tämän uuden avermektiinin valmistamiseen soveltuviissa olosuhteissa ja siihen soveltuvassa fermentaatiomediumissa *S. avermitilistä*, jossa haaraketjuisten
30 alfa-ketohappojen dehydrogenaasikompleksin ilmentymistä on heikennetty tai jossa tämä kompleksi ei ilmennyt, esimerkiksi siksi, että tällaisen ilmentymisen aikaansaavia geenejä on muokattu (esimerkiksi deletoimalla tai inaktivoimalla tai korvaamalla näitä geenejä muilla geeneillä).

Keksintö liittyy myös DNA-segmenttiin, joka sisältää SEKVENS SI ID NO:1:n, SEKVENS SI ID NO:2:n, SEKVENS SI ID NO:3:n, SEKVENS SI ID NO:4:n tai SEKVENS SI ID NO:5:n mukaisen DNA-sekvenssin tai tällaisen sekvenssin alleelisen variantin.

Keksintö liittyy myös DNA-segmenttiin, joka sisältää sellaisen DNA-sekvenssin, joka on SEKVENS SI ID NO:1:n, SEKVENS SI ID NO:2:n, SEKVENS SI ID NO:3:n, SEKVENS SI ID NO:4:n tai SEKVENS SI ID NO:5:n mukaisen DNA-sekvenssin alajoukko tai tällaisen sekvenssin alleelinen variaatio ja joka, mainitussa järjestyksessä ilmoitettuna, koettimena käytettynä kykenee hybridisoitumaan SEKVENS SI ID NO:1:een, SEKVENS SI ID NO:2:een, SEKVENS SI ID NO:3:een, SEKVENS SI ID NO:4:ään tai SEKVENS SI ID NO:5:een tai näiden tämän alleelisiin variaatioihin, tai polymeraasiketjureaktion alukkeena käytettynä monistamaan koko tällaisen sekvenssin tai tällaisen sekvenssin osan.

Tämä keksintö liittyy myös huomattavan puhtaaksi puhdistettuun polypeptidiin, joka sisältää SEKVENS SI ID NO:6:n, SEKVENS SI ID NO:7:n, SEKVENS SI ID NO:8:n tai SEKVENS SI ID NO:9:n mukaisen aminohapposekvenssin.

Piirustusten lyhyt kuvaus

Kuvio 1: Niiden polymeraasiketjureaktio (PCR) -alukkeiden nukleotidisekvenssi, joita käytettiin *S. avermitilix* BCKDH:n El-alfa-geenin fragmentin kloonamiseen. Kunkin oligodeoksinukleotidin koodaama päätelty aminohapposekvenssi on esitetty vastaavien DNA-sekvenssien yläpuolella. Nuolet osoittavat monistussuuntaa. Kuvio 1A on oikealle suuntautuva aluke ja kuvio 1B on vasemmalle suuntautuva aluke.

Kuvio 2: Haaraketjuisten alfa-ketohappojen dehydrogenaasin sekvenssien vertailu. *Streptomyces avermitilix* (Sa) PCR:n avulla kloonatun genomisen CD503-fragmentin, *Bacillus stearothermophilus* (Bs), *Pseudomonas putida* (Pp) ja *Homo sapiens* (Hs) pääteltyjen aminohapposekvens-

sien kohdistaminen keskenään. Pystysuorat merkit osoittavat sitä, että aminohapot ovat identtisiä. Kloonaukseen käytettyjä oikealle suuntautuneita ja vasemmalle suuntautuneita PCR-alukkeita vastaavien sekvenssien asema on merkitty (ylhäällä vasemmalla ja oikealla, mainitussa järjestyksessä ilmoitettuna).

Kuvio 3: *Streptomyces avermitiliks*en genomisen bkd-geeniryhmittymän restriktiokartta, sijainti ja jatkokloonit. Kartan alapuolella oleva musta suorakaide esittää sellaisen E1-alfa-spesifisen genomisen CD503-fragmentin sijaintia ja orientaatiota, jota käytettiin lähtömuotona ja joka oli kloonattu PCR:ää käyttäen. Kuviossa on esitetty genomiset jatkokloonit (pGEM-3Z:n johdannaiset). Kuviossa käy myös ilmi BCKDH:n E1-alfa- (E1a), E1-beeta- (E1b) ja E2-alayksikköjä koodaavien bkd-rakennegeenien sijainti ja organisaatio, vastaavassa järjestyksessä ilmoitettuna. Tunnistettujen avointen lukukehysten (näiden esittämiseen on käytetty suorakaiteita) polariteetti suuntautuu vasemmalta oikealle. Lyhenteet: B, BamHI; E, EcoRI; K, KpnI; Bg, BglII; ja S, SphI.

Kuviot 4a, 4b, 4c, 4d, 4e ja 4f: bkd:n avoimet E1-alfa-, E1-beeta- ja E2- (osittainen) lukukehykset (ORF:t) sisältävän 2728 ep:n suuruisen *S. avermitiliks*en genomisen DNA-fragmentin nukleotidisekvenssi ja päätelty translaatiotuotteet. E1-alfan ORF kattaa sekvenssin asemat 403 - 1548, E1-beetan ORF kattaa asemat 1622 - 2626 ja E2:n ORF alkaa asemasta 2626. Nukleotidit on numeroitu sekvenssirivien yläpuolelta. Lopetuskodonit on merkitty asteriskilla (*). Otaksutut ribosomia sitovat Shinen ja Dalgarnon sekvenssit on alleviivattu. BamHI-restriktioentsyymin tunnistussekvenssit on ympäröity suorakaiteilla.

Kuvio 5: 0,8 ke:n suuruisen *S. avermitiliks*en genomisen BglII-SphI-DNA-fragmentin (pCD509) nukleotidisekvenssi ja päätelty translaatiotuotteet, joka sekvenssi sisältää osan bkd:n E2-ORF:ää. BglII-kohdasta (tämä kohta

on ympäröity suorakulmiolla) lähtien sekvensoitiin 251 ep. Nukleotidien järjestysnumerot on esitetty sekvenssirivien yläpuolella.

Kuvio 6: Polymeraasiketjureaktion (PCR) mutageenis-
 5 sen (oikealle suuntautuneen) ja universaalisen (vasemmalle suuntautuneen) alukkeen nukleotidisekvenssi, joita alukkeita käytettiin sellaisen pT7-johdannaisten konstruoimiseen, jotka oli tarkoitettu *S. avermitiliks*en *bkd*-geenien heterologista ilmentämistä varten *E. coli*ssa. PCR-alukkei-
 10 ta käytettiin liittämään *NdeI*-restriktiokohta *S. avermitiliks*en avointen *bkd*-lukukehysten translaation aloituskodoniin (alukepari 55:31 ja 56:30, vastaavassa järjestyksessä ilmoitettuna). Kunkin mutageenisen oligodeoksinukleotidin koodaama päätelty aminohapposekvenssi on esitetty vastaa-
 15 van DNA-sekvenssin yläpuolella. Sekvensseissä on esitetty restriktioentsyymien tunnistussekvenssit. Nuolet osoittavat monistuksen suuntaa. Kuvio 6A esittää oikealle suuntautuneita mutageenisia alukkeita ja kuvio 6B esittää vasemmalle suuntautuneita mutageenisia alukkeita.

20 **Keksinnön yksityiskohtainen kuvaus**

Seuraavaksi on kuvattu uusia menetelmiä *S. avermitiliks*en *bkd*-geenien kloonaukseen ja *S. avermitiliks*en monientsyymistä BCKDH-kompleksia koodaavien geenien primaarirakenteen määrittämiseksi.

25 Ensimmäiseksi suunniteltiin kaksi PCR-aluketta, joille annettiin nimitykset "oikealle suuntautuva aluke" ja "vasemmalle suuntautuva aluke" (kuvio 1) ja jotka perustuivat alan ammattikirjallisuudesta saatuihin ja useista eri lajeista peräisin olevista pääteltyistä BCKDH:n E1-
 30 alfa-peptidisekvensseistä määritettyihin säilyviin alueisiin, jotka oli tunnistettu kohdistamalla useat sekvenssit toisiinsa nähden rinnakkain. Kun *S. avermitiliks*en genomien DNA:ta oli monistettu PCR:llä käyttäen sekä oikealle että vasemmalle suuntautuvia alukkeita, havaittiin noin 0,2
 35 ke:n pituinen PCR-tuote. Tämä PCR:llä monistettu DNA-frag-

mentti kloonattiin myöhemmin *E. coli* -vektoriin pGEM-3Z yhdistelmäplasmidin pCD503 aikaansaamiseksi. Tämän jälkeen plasmidi pCD503 transformoitiin kompetentteihin *E. coli* DH5-alfa -soluihin. Yhdelle transformantille annettiin

5 nimitys CD503-kanta. CD503:n kloonatun DNA-fragmentin DNA-sekvenssointi osoitti siinä esiintyvän avoimen lukukehyksen, joka sisälsi päätellyn peptidin, joka oli erittäin homologinen BCKDH:n E1-alfa-alayksikön suhteen (kuvio 2). Kloonattua CD503:n genomien DNA-fragmenttia käytettiin tä-

10 män jälkeen koettimena *S. avermitiliks*en kromosomikirjaston seulontaan pesäkehybridisaatiolla. Tunnistettiin neljä kosmidikloonaa, jotka olivat CD518, CD519, CD520 ja CD521. Restriktio- ja Southern-blottianalyysit osoittivat, että kaikissa neljässä kloonissa oli limittäisiä genomien frag-

15 mentteja. Jatkokloonatuista genomien DNA:n fragmenteista peräisin olevien sisäkkäisten deleetioiden DNA:n sekvenssointi (tämä on kuvattu yksityiskohtaisesti esimerkissä 8) osoitti, että CD503:n sekvenssi oli osa täydellistä bkd-geeniryhmittymää. Kloonattujen *S. avermitiliks*en bkd-geenien kattama alue kromosomista on noin 15 kiloemäksen pituinen (kuvio 3). DNA-sekvenssin analyysi osoitti sekvenssissä esiintyvien otaksuttujen transkription promoottorisekvenssien ja bkd-rakennegeenien järjestäytyneen ryhmittymäksi, jonka organisaatio oli seuraava: promoottorisek-

20 venssi sekä avoimet E1-alfa-, E1-beeta- ja E2-lukukehykset (kuviot 4 ja 5).

Lopuksi täydellinen *S. avermitiliks*en bkd-geeniryhmittymä kloonattiin myötäsuuntaan voimakkaasta *Escherichia coli* T7-promoottorista tämän ilmentämiseksi *E. coli* -isännässä. Samalla tavoin kloonattiin myös avoimet E1-alfa- ja E1-beeta-lukukehykset (ORF:t) joko erikseen tai yhdessä myötäsuuntaan T7-promoottorista ja kunkin konstruktion ilmentyminen tutkittiin. Mutageeniset uudet PCR-alukkeet, joita käytettiin liittämään ainutlaatuinen

30 NdeI-restriktiokohta E1-alfa- ja E1-beeta-ORF:n translaa-

35

tion ATG-aloituskohtaan, on kuvattu täydellisesti esimerkiksi 9 (ks. myös kuvio 6). Tutkimus, jossa käytettiin kaksois-plasmidi T7 -ilmentämisjärjestelmää, osoitti, että ainakin kaksi CD503:sta peräisin olevan *S. avermitiliks*en bkd-geeniryhmittymän avointa lukukehystä (El-alfa ja El-beeta) ovat transloitavissa täydellisesti *E. colissa* ilmenettynä. Lisäksi entsyymaattiset määritykset, joiden tarkoituksena oli analysoida spesifisesti BCKDH-kompleksin El-komponenttia, varmisti pitävästi sen, että kaksi *E. coli* -yhdistelmäklooneista - joista toinen sisälsi kokonaisen bkd-geeniryhmittymän ja toinen El-alfa-ORF:n ja El-beeta-ORF:n, sisälsi BCKDH:lle spesifistä El-entsyymiaktiivisuutta (taulukko 1).

Tässä keksinnössä kuvataan bkd-geenien eristys avermektiiniä tuottavasta *Streptomyces avermitiliks*estä. *Streptomyces*in suuri genomi (noin 10^4 ke) on yli kaksi kertaa niin suuri kuin *Escherichia coli*n genomi. *Streptomyces*in genomi koostuu DNA:sta, jonka guanosiini ynnä sytosiini (G+C) -sisältö on erittäin suuri (keskiarvo jopa 73 %, jotkin alueet > 90 %), mikä vie tämän lähelle luonnossa havaittua ylärajaa. Nämä selvästi poikkeavat ominaisuudet tekevät välttämättömäksi kehittää *streptomyces*iteille spesifisiä yhdistelmä-DNA-menetelmiä. Esimerkkejä näistä ponnisteluista on US-patenttihakemuksessa 08/048 719, joka on jätetty 16.4.1993 ja US-patenttihakemuksessa 08/032 925, joka on jätetty 18.3.1993. Nämä patenttihakemukset on liitetty tähän keksintöön kokonaisuudessaan siihen oikeuttavien säännösten nojalla.

Muita menetelmiä, jotka on erityisesti optimoitu tämän keksinnön tarkoituksia varten, kuten genomisen DNA:n PCR-monistus, sisäkkäisten deletioiden valmistus, DNA:n sekvenssointi ja *S. avermitiliks*en bkd-geenien heterologin ilmentäminen *E. colissa*, on kuvattu siten, että täydelliset yksityiskohdat on esitetty esimerkkiosassa.

S. avermitiloksen bkd-geenien kloonaukseen liittyneet kokeelliset vaiheet ja saadut tulokset ovat täydellisesti selitettyinä seuraavat:

- (a) Niiden BCKDH:n E1-alfa-peptidialayksikössä olevien säilyneiden alueiden tunnistaminen, jotka voisivat toimia ehdolla olevina PCR-alukkeiden sitoutumiskohtina
- Neljä BCKDH:n E1-alfa-peptidisekvenssiä, jotka olivat peräisin ihmisestä [Fisher, et al., J. Biol. Chem. 264 (1989) 3448 - 3453], rotasta (Zhang, et al., J. Biol. Chem. 262 (1987) 15220 - 15224), *Pseudomonas putidasta* [Sokatch, et al., Eur. J. Biochem. 176 (1988) 311 - 317]] ja *Bacillus stearothermophiluksesta* [Perham, et al., Eur. J. Biochem. 191 (1990) 337 - 346], kohdistettiin toistensa suhteen sellaista säilyneiden alueiden tunnistamiseksi, jotka voisivat toimia vastaavien PCR-alukkeiden suunnittelussa käytettävänä sekvensseinä. Tietokoneanalyysi sekä prokaryoottisissa että eukaryoottisissa BCKDH-komplekseissa erittäin hyvin säilyneiden E1-alfa-alayksikön alueiden tunnistamiseksi suoritettiin käyttäen GCG:n sekvenssianalyysitietokoneohjelmistosta (Madison, WI) saatavilla LineUp- ja Pretty-ohjelmilla. BCKDH:n neljän E1-alfa-peptidin kohdistaminen monenkeskisesti osoitti niissä esiintyvän useita runsasta homologiaa sisältäviä alueita [tätä on käsitelty julkaisussa Wexler, I.D., et al., FEBS Letters 282 (1991) 209 - 213]. Tiamiinipyrofosfaattia sitova ryhmitys (motif) [Perham, et al., FEBS Letters 255 (1989) 77 - 82], joka sijaitsee humaani-E1-alfa-aminohappojen 182 - 229 välisellä alueella, sekä fosforylaatiokohdat 1 ja 2 kattava alue, joka kattoi aminohappojen 245 - 289 välisen alueen, olivat säilyneet huomattavan hyvin kaikissa tutkituissa BCKDH:n neljässä E1-alfa-peptidissä. Peptideissä esiintyi myös aikaisemmin kuvattu aminohappojen 291 - 307 välinen suuren homologian alue. Tämä alue näyttää olevan ainutlaatuinen sekä alfa- että beeta-alayksiköt sisältäville alfa-ketohappodehydrogenaaseille eikä se ole

homologinen minkään niiden sekvenssien suhteen, joita esiintyy ainoastaan yhdestä E1-polypeptidistä muodostuvassa E. coliin PDC:n E1:ssä tai E. coliin ja hiivan alfa-ketoglutaratidehydrogenaasikompleksien E1-komponenteissa, jotka ovat ainoastaan yhdestä E1-polypeptidistä koostuvia dimeerejä. Edellä mainituista syistä on viimeksi mainitulla homologia-alueella ehdotettu olevan osuutta alayksikköjen vuorovaikutuksessa [Patel, et al., FEBS Letters 282 (1991) 209 - 213]. PCR-alukkeen suunnittelemiseen valitut säilyvät alueet koodasivat humaanin-BCKDH:n E1-alfa-proteiinin aminohapporyhmiä 192 - 200 ja 370 - 376.

(b) Kyseisiin BCKDH:n E1-alfan säilyneisiin alueisiin perustuvien PCR-alukkeina käytettävien uusien oligonukleotidien suunnittelu

Kuten jo aiemmin on mainittu, valittiin moninkeskisten kohdistustutkimusten perusteella käyttöön kaksi BCKDH:n E1-alfa-alayksikön säilynyttä aluetta. Oikealle suuntautuva PCR-alue (kuviot 1) suunniteltiin humaanin-BCKDH:n E1-alfa-alayksikön aminohapot 192 - 200 kattavan alueen perusteella - tätä käytettiin BCKDH:n E1-alfa-alayksikköjä edustavana mallina. Nämä aminohapot sijaitsevat tiamiinipyrofosfaattia sitovassa ryhmityksessä. Vasemmalle suuntautuva PCR-alue (kuviot 1) suunniteltiin humaanin-BCKDH:n E1-alfa-alayksikön aminohapot 370 - 376 kattavan alueen perusteella. Viimeksi mainittu aminohapposekvenssi sijaitsee peptidin C-terminaalisen alueen läheisyydessä. Käytetty kodonien nimitystapa oli Streptomycesin geenien mukainen [Wright, F. ja Bibb, M.J., Gene 113 (1992) 55 - 65]. Kunkin alukkeen 5'-puoleisessa päässä oli PCR-tuotteiden kloonauksen helpottamiseksi restriktioentsyymin tunnistuskohta (oikealle suuntautuvassa alukkeessa oleva EcoRI ja vasemmalle suuntautuvassa alukkeessa oleva XbaI).

Oikealle suuntautuvan PCR-alukkeen täydellinen sekvenssi on: 5'-GAATTCGGCGACGGCGCCACCTCCGAGGGCGAC-3'

Vasemmalle suuntautuvan PCR-alukkeen täydellinen sekvenssi on: 5'-~~TCTAGAC~~CGCAGGTGGTCCGGCATCTC-3'.

Sekvenssit, jotka eivät ole homologisia bkd:n El-alfa-geenien suhteen ja jotka on liitetty alukkeisiin
 5 kloonaustarkoituksia varten, on alleviivattu (ks. myös kuvio 1).

(c) S. avermitiloksen genomisten DNA-fragmenttien monistus PCR:n avulla

S. avermitiloksen genomien DNA monistettiin entsy-
 10 maattisesti käyttäen runsaasti GC:tä sisältävälle DNA:lle asianmukaisia reaktio-olosuhteita, joissa Streptomycesin DNA:ta kyetään tehokkaasti ja spesifisesti monistamaan (ks. esimerkki 2). PCR suoritettiin käyttäen edellä kuvattua alukeyhdistelmää (oikealle suuntautuva aluke,
 15 5'-GAATTCGGCGACGGCGCCACCTCCGAGGGCGAC-3', ja vasemmalle suuntautuva aluke, 5'-TCTAGACCGCAGGTGGTCCGGCATCTC-3'). Monistustuotteet fraktioitiin koon mukaan agarosigeeli-elektroforeesilla. Edellä kuvatuissa PCR:n olosuhteissa havaittiin tätä alukeyhdistelmää käyttäen yksittäinen DNA-
 20 vyöhyke (pituudeltaan noin 250 emäsparia).

(d) Monistetun genomisen DNA-fragmentin kloonaus Escherichia coli -kloonausvektoriin ja tämän jälkeen suoritettu transformaatio E. coli -isäntään

Kuten edellä mainittiin, liitettiin oikealle suun-
 25 tautuvaan PCR-alukkeeseen EcoRI-restriktiokohta kloonaamisen helpottamiseksi ja vasemmalle suuntautuvan alukkeen 5'-puoleinen pää sisälsi XbaI-restriktiokohdan. Yritykset 0,25 ke:n suuruisen PCR-fragmentin kloonaamiseksi käyttäen ligaatiomenetelmää, jossa sekä liitosjakso että kloonaus-
 30 vektori oli pilkottu EcoRI:llä ja XbaI:llä, eivät johtaneet suotuisiin tuloksiin. Tästä syystä tutkittiin sellaisen vaihtoehtoisen lähestymistavan käyttöä 0,25 ke:n suuruisen PCR-fragmentin kloonaamiseen, johon liittyi DNA-polymeraasi I:n Klenow' in fragmentin käyttö tasaisten päi-
 35 den valmistamiseen PCR-fragmenttiin. Kun päistään tasoi-

tettu fragmentti oli liitetty päistään tasoitettuun ja SmaI:llä linearisoituun E. coli -vektoriin (pGEM-3Z[f+]), saatiin talteen yksi yhdistelmäkloni, josta saatiin yhdistelmäplasmidi pCD503. Tämän jälkeen pCD503 liitettiin kompetentteihin E. coli DH5-alfa -soluihin transformoimalla. Valitulle transformantille annettiin nimitys CD503-kanta. Tämän kannan varmistamiseksi tehty restriktioanalyysi osoitti, että plasmidi pCD503 - tämä oli eristetty E. coli -kannasta CD503 - sisälsi todella 0,25 ke:n suuruisen S. avermitiloksen liitosjakson.

(e) PCR:llä monistetun 0,25 ke:n suuruisen DNA-liitosjakson jatkokloonaus M13-bakteriofagiin, kloonatun fragmentin DNA:n sekvensointi ja bkd-spesifisten geenien tunnistaminen

Plasmidissa pCD503 oleva 0,25 ke:n suuruinen liitosjakso jatkokloonattiin M13-bakteriofagiin. Tämän suorittamiseksi vapautettiin ensimmäinen liitosjakso E. coli -vektorista pilkkomalla pCD503 EcoRI:llä ja PstI:llä, jotka ovat kaksi restriktioentsyymiä, joiden tunnistussekvenssit esiintyivät pGEM-vektorin monikäyttöisessä kloonaukskohdassa kloonatun liitossekvenssin molemmiin puoliin. Spesifinen fragmentti kloonattiin sitten sekä M13mp18-vektoriin että M13mp19-vektoriin, jotka oli käsitelty EcoRI:lla ja PstI:lla. Kloonaus kumpaankin vektoriin varmistaa sen, että sekvenssointia varten on mahdollista tuottaa liitosjakson DNA:n kummankin juosteen yksijuosteista DNA:ta. Yhdelle spesifisen liitosjakson sisältävälle kloonille, joka oli valikoitu mp18-transfektiokokeesta, annettiin nimitys CD505. Niin ikään spesifisen liitosjakson sisältävälle toiselle kloonille, joka oli tässä tapauksessa valikoitu mp19-transfektiokokeesta, annettiin nimitys CD506. DNA-sekvensointi suoritettiin menetelmällä, jossa ketjun päätteeksi muodostetaan dideoksinukleotidi, käyttäen yksijuosteista DNA-mallijuostetta ja TaqTrack-reagenssipakkausta (Promega). Kaikissa tapauksissa sek-

venssoitiin molemmat DNA-juosteet - näistä toinen oli peräisin kloonista CD505 ja komplementaarinen juoste peräisin kloonista CD506. Klooneista CD505 ja CD506 saatujen DNA-sekvensointikoetulosten koodien käytön ensisijaisuusanalyysi (GCG-sekvenssianalyysi -tietokoneohjelmisto, Madison/WI), osoitti, että näissä esiintyi avoin lukukehys, jolla oli streptomykeettien geenien oikea koodien käyttötapa.

Otaksuttu avoin lukukehys käännettiin sitten aminohapposekvenssiksi käyttäen IntelliGenetics Suite -ohjelmiston Seq- ja Translate-ohjelmia (IntelliGenetics Inc., Mountain View, Kalifornia). Lopuksi suoritettiin tietokannasta samanlaisuushakuja käyttäen haettavana olevaa peptidisekvenssiä ja Intelligenetics-ohjelmiston FASTDB-ohjelmaa. Kaikki tietokantahaut, joissa haettiin joko DNA-tietokantoja (GenBank ja EMBL) tai proteiinitietokantoja (PIR ja Swiss-Prot), osoittivat kiistatta, että kloonista CD503 peräisin oleva sekvenssi oli erittäin homologinen BCKDH:n El-alfa-peptidien suhteen mutta se oli uusi ja erosi kaikista muista sekä prokaryoottisista että eukaryoottisista organismeista peräisin olevissa tietokannoissa luetelluista BCKDH:n El-alfa-peptideistä. Kuviossa 2 on esitetty ihmisestä, rotasta, *Pseudomonas putida*sta ja *Bacillus stearothermophilus*ksesta, peräisin olevien BCKDH:n El-alfa-peptidisekvenssien, mukaan luettuna uuden *Streptomyces avermitilis*ksen BCKDH:n El-alfan peptidisekvenssin, monenkeskinen kohdistus. Näistä koetuloksista voidaan tehdä se johtopäätös, että *E. coli* -kantaan CD503 kloonattu 250 ep:n suuruinen *S. avermitilis*ksen genomisen PCR-tuote edustaa todella uutta bkd:n El-alfa-geenin fragmenttia.

(f) Koko *S. avermitilis*ksen bkd-geeniryhmittymän kloonaus, restriktio- ja Southern-blottianalyysit ja kromosomaalisen kartan konstruointi

Noin 0,25 ke:n suuruista BamHI/EcoRI-DNA-fragmenttia, joka oli peräisin *S. avermitilis*ksen bkd-spesifisen

El-alfan sekvenssin sisältävästä pCD503:sta, käytettiin radioaktiivisella leimalla varustettuna koettimena *S. avermitilix*in genomisen DNA:n kosmidikirjaston seulontaan pesäkehybridisaation avulla. Tästä tunnistettiin ja
 5 otettiin talteen neljä kloonia (CD518, CD519, CD520 ja CD521). Restriktio- ja Southern-blottihybridisaatioanalyysit osoittivat, että nämä neljä kloonia sisältävät samalta kromosomialueelta peräisin olevia limittäisiä sekvenssejä. Tätä samaa koetinta käytettiin erittäin rajoittavissa olo-
 10 suhteissa suoritettussa tutkimuksessa, joka kohdistui *S. avermitilis* ATCC 31272:sta peräisin olevasta pilkotusta kromosomaalisesta DNA:sta valmistettuihin Southern-blotteihin. Viimeksi mainittu analyysi antoi varmuuden genomikirjastosta talteen saatujen kloonien identiteetistä.
 15 *S. avermitilis* CD503:n sekvenssin sisältävästä genomialueesta konstruoitu restriktiokartta on esitetty kuviossa 3).

(g) CD518- ja CD521-klooneista peräisin olevien genomisten DNA-fragmenttien jatkokloonaus ja *bkd*-geeniryhmittymän sisältävän *S. avermitilix*in kromosomaalisen alueen DNA-sekvenssointi

Genomiset fragmentit (pituudeltaan 1 - 2 ke), jotka kattoivat *S. avermitilix*in kromosomin koko CD503:n *bkd*-alueen, jatkokloonattiin DNA-kirjaston klooneista CD521 ja
 25 CD518 *E. coli* -vektoriin PGEM-3Z. Luettelo tässä tutkimuksessa konstruoiduista jatkoklooneista, joihin kuuluu lyhyt kuvaus kustakin plasmidista, on seuraava: 1. Plasmidi pCD528 sisältää 7 ke:n suuruisen BamHI-fragmentin, joka on jatkokloonattu pCD518:sta; 2. Plasmidi pCD545 sisältää
 30 2,3 ke:n suuruisen SphI-fragmentin, joka oli jatkokloonattu pCD528:sta; 3. Plasmidi pCD550 sisältää 6 ke:n suuruisen SphI-fragmentin, joka oli jatkokloonattu pCD521:stä; 4. Plasmidi pCD559 sisältää 7 ke:n suuruisen BamHI:n joka oli jatkokloonattu pCD521:stä; 5. Plasmidi pCD574 sisältää
 35 4,2 ke:n suuruisen SphI-BglII-fragmentin, joka oli jatko-

kloonattu pCD550:stä; ja 6. Plasmidi pCD577 sisältää noin 10,4 ke:n suuruisen liitosjakson. Tämä liitosjakso sisältää uudelleen yhteen kootut kaksi vierekkäistä genomista fragmenttia: 4,2 ke:n suuruisen SphI/BglII-fragmentin, joka oli jatkokloonattu pCD550:stä, ja 6,2 ke:n suuruisen BglII/BamHI-fragmentin, joka oli jatkokloonattu pCD559:stä. Kunkin jatkokloonin identtisyys varmistettiin plasmidien restriktiokartoituksella, Southern-hybridisatiolla ja PCR-analyysillä. Sangerin ketjun päättämismenetelmää käytettiin nukleotidisekvenssin määrittämiseksi. Tätä tarkoitusta varten jatkokloonattiin myöhemmin *S. avermitiliks*en genomisia fragmentteja M13mp18- ja M13mp19-bakteriofageihin molempien DNA-juosteiden sekvenssin määrittämiseksi. Edellä mainituista pGEM:sta peräisin olevista klooneista eristettiin useita DNA-restriktiofragmenteja ja nämä ligatoitiin M13mp18:aan ja M13mp19:ään, ja tulokseksi saatiin seuraavat yhdistelmäfragit:

CD535: 0,4 ke:n suuruinen *S. avermitiliks*en DNA-fragmentti, joka oli kloonattu PCR:n avulla käyttäen mallijuosteena pCD528:n DNA:ta, spesifistä 29-PCR-EX-aluketta (5'-AAGAATTCTCGAGCTGGCCCAACAAGCCGTCGGCTAC-3') ja yleisaluketta 31-PCRBP (ks. esimerkki 9 ja kuvio 6). Monistettu fragmentti käsiteltiin restriktioensymeillä käyttäen EcoRI:tä ja PstI:tä ja se kloonattiin EcoRI:n ja PstI:n seoksella linearisoituun M13mp18:n DNA:han.

CD536: edellä kuvattua fragmenttia muistuttava DNA-fragmentti, joka kloonattiin M13mp19-DNA:han.

CD537: 1,15 ke:n suuruinen SalI-DNA-fragmentti, joka sisältää CD503:n sekvenssin, joka on jatkokloonattu pCD528:sta M13mp18:aan.

CD538: 1,15 ke:n suuruinen pCD528:n DNA:n SalI-fragmentti (tämä sijaitsee vastasuuntaan edellä kuvattuun 1,15 ke:n suuruiseen SalI-fragmenttiin nähden), joka oli kloonattu M13mp18:aan.

CD539: 1,5 ke:n suuruinen BamHI/BglIII-DNA-fragmentti, joka oli jatkokloonattu pCD550:sta M13mp18:aan.

CD539: 1,5 ke:n suuruinen BamHI/BglIII-DNA-fragmentti, joka oli jatkokloonattu pCD550:sta M13mp18:aan.

5 CD540: edellä kuvatun DNA-fragmentin suhteen samanlainen fragmentti, joka oli kloonattu vastakkaiseen orientaatioon M13mp18:aan.

CD541: 0,35 ke:n suuruinen SalI/BamHI-DNA-fragmentti, joka oli jatkokloonattu pCD528:sta M13mp18:aan.

10 CD542: 0,35 ke:n suuruinen SalI/BamHI-DNA-fragmentti, joka oli jatkokloonattu pCD528:sta M13mp19:ään.

CD553: 0,8 ke:n suuruinen BamHI/BglIII-DNA-fragmentti, joka oli jatkokloonattu pCD550:sta M13mp18:aan.

15 CD554: 1,1 ke:n suuruinen BamHI-DNA-fragmentti, joka oli jatkokloonattu pCD550:sta M13mp18:aan.

CD555: edellä kuvatun DNA-fragmentin suhteen samanlainen fragmentti, joka oli kloonattu vastakkaiseen orientaatioon M13mp18:aan.

20 CD558: 0,8 ke:n suuruinen BamHI/HindIII-DNA-fragmentti, joka oli jatkokloonattu pCD553:sta M13mp19:ään.

CD561: 1,15 ke:n suuruinen SalI-DNA-fragmentti, joka oli jatkokloonattu pCD537:sta M13mp18:aan (orientaatioltaan vastakkainen konstruktiossa CD537 esiintyvään orientaatioon nähden).

25 CD565: 1,15 ke:n suuruinen SalI-DNA-fragmentti, joka oli jatkokloonattu pCD537:sta M13mp19:ään.

30 CD566: 1,15 ke:n suuruinen SalI-DNA-fragmentti, joka oli jatkokloonattu pCD537:sta M13mp19:ään (orientaatioltaan vastakkainen konstruktiossa CD565 esiintyvään orientaatioon nähden).

CD567: 1,15 ke:n suuruinen SalI-DNA-fragmentti, joka oli jatkokloonattu pCD538:sta M13mp19:ään.

CD582: 0,8 ke:n suuruinen BamHI/BglIII-DNA-fragmentti, joka oli jatkokloonattu pCD550:sta M13mp18:aan (orien-

taatioltaan vastakkainen konstruktiossa CD553 esiintyvään orientaatioon nähden).

Näiden kloonien sisältämiä *S. avermitiliks*en genomisia liitosjaksoja typistettiin myöhemmin käsittelemällä niitä eksonukleaasi III:lla jatkokloonisarjan aikaansaamiseksi ("sisäkkäiset deleetit", ks. esimerkki 8).

(h) Kloonaatuista DNA-fragmenteista saatujen DNA:n sekvensointi koetulosten tietokoneanalyysi ja *S. avermitiliks*en bkd:n avointen E1-alfa-, E1-beeta- ja E2-lukukehysten tunnistaminen

Kuviossa 4 on esitetty bkd-geenit sisältävän 2,7 ke:n suuruisen *S. avermitiliks*en genomisen alueen nukleotidisekvenssi. *S. avermitiliks*en bkd:n avointen E1-alfa-, E1-beeta- ja E2 (osittainen) lukukehykset (ORF:t) sisältävän 2,7 ke:n suuruisen genomisen alueen liukuva emäskoostumusanalyysi suoritettiin käyttäen "DNA Inspector" -ohjelmistoa. Tästä analyysistä saatiin profiili, joka kuvasi G+C-sisällön liukuvaa keskiarvoa käyttäen 30 emäksen pitoista ikkunaa ja 20:n suuruista siirtymäarvoa. Tämän *S. avermitiliks*en kromosomin aluetta vastaava kokonais-G+C-koostumus oli 72 %. Heti bkd:n avointen lukukehysten jälkeen vastasuuntaan esiintyi pienen G+C-sisällön sisältävä alue (G+C-sisältö noin 50 %) - mikä osoitti kyseessä olevan promoottorialueen.

G+C-sisältö analysoitiin myös kodonin aseman funktiona. Avoimet lukukehykset havaittiin käyttäen ohjelmaa "CodonPreference" (Genetics Computer Group, Madison, WI) käyttäen suoritukseen 64 geenistä valmistettua *Streptomyces*in kodonien käyttötapataulukkoa [Wright, F. ja Bibb, M.J., Gene 113 (1992) 55 - 65]. CodonPreference-ohjelma on lukukehysspesifinen geenien hakuohjelma, joka pyrkii tunnistamaan proteiineja koodaavat sekvenssit sen perusteella, että ne muistuttavat kodonien esiintymistiheystaulukkoa, tai sen perusteella, että niiden koostumus (tavallisesti GC) kunkin kodonin kolmannessa asemassa on painottu-

nut tietyllä tavalla. ORF-alueet on esitetty suorakaiteina niitä vastaavia lukukehyksiä esittävän kaavion alapuolella. Havaittiin myös kaikki aloitus- (ATG) ja lopetuskodonit (pystysuorat viivat). Kussakin lukukehyksessä havaitut harvinaiset kodonit merkittiin myös kunkin ORF-kaavion alapuolelle. G+C-sisältö laskettiin käyttäen liukuvaa 25 kodonin ikkunaa, joten noin 25 kodonin viive oli odotettua ennen kuin proteiinia koodaavan alueen vaikutuksen havaittiin ilmenevän täysimääräisenä. Saatiin kolme profiilia, jotka olivat: 1, ensimmäinen asema tripletissä; 2, toinen asema tripletissä; 3, kolmas asema tripletissä. Tämän analyysin tuloksena paikannettiin kolme bkd:n ORF-aluetta, jotka vastasivat BCKDH-alayksikköjä El-alfa, El-beeta ja E2 (kuviot 4 ja 5).

(i) PCR:ään perustuvassa kohdemutageneesissä alukkeina käytettäväksi tarkoitettujen uusien oligonukleotidien suunnittelu

Kytkejäsekvenssiin tai PCR:ään perustuvaa kohdemutageneesiä käytettiin liittämään NdeI-restriktiokohta El-alfa- ja El-beeta-ORF-sekvenssien translaation ATG-alkukohtaan. Suunniteltiin seuraavat uudet oligonukleotidit (ks. myös esimerkki 9 ja kuvio 6):

Vasemmalle suuntautuvat universaalit (vektori) alukkeet:

30-PCR-BP: 5'-AAGGATCCTGCAGCCCAGTCACGACGTTGTAAAACGA-3'

31-PCR-BP: 5'-AAGGATCCTGCAGACAGCTATGACCATGATTACGCCA-3'

Oikealle suuntautuvat mutageeniset alukkeet:

55-PCR: 5'-AAGAGATCTCATATGACGGTCATGGAGCAGCGG-3'

56-PCR: 5'-AAGAGATCTCATATGACCACCGTTGCCCTGAAG-3'

(j) *S. avermitiliks*en bkd-geenien kohdemutageneesi uuden NdeI-restriktiokohdan muodostamiseksi vastasuuntaan avoimesta lukukehyksestä

Ilmentämisplasmidit olivat plasmidin pT7-7 (tätä on käsitelty julkaisussa Tabor, S., 1990, Current Protocols in Molecular Biology, s. 16.2.1 - 16.2.11. Greene Pub-

lishing and Wiley-Interscience, New York) johdannaisia, jotka sisälsivät El-alfa-, El-beeta- sekä El-alfa- ynnä El-beeta-ORF-sekvenssit, tai *S. avermitiliks*en täydellisen bkd-geeniryhmittymän. NdeI-restriktiokohta muodostettiin PCR:ään perustuvalla kohdemutageneesillä. Tässä tutkimuk-

5

Plasmidi pCD670: pT7-7:n johdannainen, joka sisältää *S. avermitiliks*en bkd:n avoimen El-alfa-lukukehyyksen (ORF1). *S. avermitiliks*en bkd:n El-alfa-geeniin lisättiin ATG-aloituskodonin kattava NdeI-restriktiokohta käyttämäl-

10

lä monistusta ja samanaikaista mutageneesiä käyttäen mutageenista PCR-aluketta 55-PCR (ks. esimerkki 9 ja kuvio 6). Plasmidi pCD666: pT7-7:n johdannainen, joka sisältää *S. avermitiliks*en bkd:n avoimen El-beeta-lukukehyyksen (ORF2). *S. avermitiliks*en bkd:n El-beeta-geeniin lisättiin ATG-aloituskodonin kattava NdeI-restriktiokohta käyttämäl-

15

lä monistusta ja samanaikaista mutageneesiä käyttäen mutageenista PCR-aluketta 55-PCR (ks. esimerkki 9 ja kuvio 6). Tämän ORF:n optimaalisen ilmentymisen aikaansaamiseksi muutettiin kodonin 7 kolmas asema C:sta G:ksi kodonisynonyymien muodostamiseksi, joka muistuttaa *E. coli*n kodonien käyttötapaa. Tämä muutos ei vaikuttanut El-beetan peptidisekvenssiin.

20

Plasmidi pCD736: pT7-7:n johdannainen, joka samanaikaisesti sisältää sekä El-alfa (ORF1) että El-beeta (ORF2) -ORF-sekvenssit T7-promoottorin ohjaamina.

25

Plasmidi pCD705: Tämä plasmidi muistuttaa pCD736:ta, mutta siinä on El-beeta-ORF:n 3'-puoleinen puolikas väärällä tavoin orientoituneena. Tätä konstruktia käytettiin negatiivisena kontrollina ilmentämiskokeessa.

30

Plasmidi CD685: PT7-7:n johdannainen, joka sisältää *S. avermitiliks*en täydellisen bkd-geeniryhmittymän.

(k) *S. avermitiliks*en bkd:n avointen lukukehysten ilmentäminen *E. coli*ssa käyttäen kaksois-T7-plasmidi-ilmentämisjärjestelmää

*S. avermitiliks*en bkd-geenit ilmennettiin *E. coli*ssa käyttäen T7:n RNA polymeraasi/promoottori -kaksoisplasmidijärjestelmää olennaisesti julkaisussa Tabor, S., (1990) Current Protocols in Molecular Biology, s. 16.2.1 - 16.2.11, Greene Publishing and Wiley-Interscience, New York), kuvatulla tavalla. Tutkittiin *E. coli* C600:n (pGP-1) johdannaisia, jotka sisälsivät eri pT7-7-konstruktioita. *S. avermitiliks*en ORF-sekvenssien ilmentymisen monitoroimiseksi *E. coli* -isännässä lämpöinduktion jälkeen käytettiin natriumdodekyylisulfaatti-polyakryyliamidigeelelektroforeesia (SDS-PAGE). Induktion jälkeen saadun proteiiniprofiilin SDS-PAGE-analyysi osoitti, että soluissa ilmentyivät normaalia voimakkaammin indusoituneet peptidit, jotka muistuttivat kooltaan E1-alfa- ja E1-beeta-ORF-sekvenssien ennustettua arvoa (pääteltyä vastaavasta DNA-sekvenssistä):

20

	Ennustettu koko (Da)	Havaittu koko (Da)
ORF1 (E1-alfa)	41 000	41 000
ORF2 (E1-beeta)	35 000	34 000

25

(l) *S. avermitiliks*en BCKDH:n E1:n aktiivisuuden havaitseminen käyttämällä spesifistä määritystä yhdistelmä-*E. coli* -kloonin raakauutteessa

Nämä tulokset on koottu yhteenvedona tuonnempana olevaan taulukkoon 1 (esimerkki 11). E1-spesifinen BCKDH-määritys suoritettiin käyttäen sellaisia *E. coli* -solujen raakauutteita, jotka sisälsivät pCD736:tta ja joissa esiintyi merkittävää E1-aktiivisuutta T7-promoottorin induktion jälkeen. Samanlaisessa viljelmässä, joka sisälsi

30

osan liitosjaksosta asetettuna väärään orientaatioon (pCD705), havaittiin taustatasoa oleva aktiivisuus.

Entsyymimääritykset osoittivat lisäksi, että plasmidia pCD685 sisältävästä *E. coli* -kannasta valmistetuissa raakauutteissa oli myös merkittävää El-BCKDH-aktiivisuutta (> 10-kertainen taustan taso). Tutkittiin myös tämän kloonin indusoimaton viljelmä, jonka sisältämä tausta-aktiivisuus osoittautui olevan kaksinkertainen taustaan verrattuna. Viimeksi mainittu tulos on odotettua, koska T7-järjestelmän tiedetään mahdollistavan kloonattujen geenien ilmentymisen konstitutiivisesti pienellä ilmentymistasolla jopa indusoimattomissa olosuhteissa.

Kloonattu *Streptomyces avermitiliks*en bkd-geeniryhmittymä on käyttökelpoinen luonnollisen avermektiinin tuoton parantamisessa sillä perusteella, että se lisää näiden geenien lukumäärää, tai sillä perusteella, että sillä saadaan optimoitua niiden ilmentyminen tuotantokannoissa. Yksi mahdollinen lähestymistapa saada kloonatut bkd-geenit ilmentymään tehokkaasti sisältää sen, että nämä geenit liitetään kappalemäärältään suureen *E. coli* / streptomykeetti-sukkulavektoriin (esimerkiksi plasmidiin pCD262, Denoya, C.D., (1993) "Novel Bacterial Plasmids Shuttle Vectors for Streptomyces and Escherichia coli", US-patenttihakemus 08/032 925, jätetty 18.3.1993) siten, että geenit transkriптоituvat voimakkaasta promoottorista. Tämä menetelmä varmistaa geenien tehokkaan transkriптоitumisen. Voidaan lisäksi suunnitella tiettyjä strategioita tehokkaan ilmentymisen takaamiseksi. Näitä ovat (a) promoottorin voimakkuus; (b) mRNA:n pysyvyys; (c) säätelytekijöiden käyttäminen tai näiden pois jättäminen; (d) indusoitavuus; ja (e) kohdemutageneesi ribosomin tunnistus-translaation aloitussignaalien parantamiseksi. bkd-geenien ilmentymistä voitaisiin myös optimoida käyttämällä villityyppisen promoottorin ja säätelyalueiden tilalla eri promoottoreja käyttämällä geenejä korvaavia menetelmiä. Alan ammattikir-

jallisuudessa on monia esimerkkejä käyttökelpoisista promootoreista, joita voitaisiin käyttää tässä keksinnössä kuvattujen uusien *S. avermitilix*en bkd-geenien ilmentämisen optimoimiseksi, esimerkiksi voimakas ermE-promoottori (Hopwood, et al., 1985, 'Genetic Manipulation of Streptomyces: A Laboratory Manual', The John Innes Foundation, Norwich, U.K.) ja tiostreptonilla indusoitavissa oleva tipA-promoottori (Murakami, et al., J. Bacteriol. 171 (1989) 1459 - 1466). Lisäksi bkd-geenien inaktivointi ja tästä johtuva BCKDH-aktiivisuuden puute käyttämällä deleetio- tai kohdemutageneesiä ja geenien korvausmenetelmiä käyttäen saadaan kehitettyä parannettuja palautumattomasti toimintakykynsä menettäneitä *Streptomyces avermitilix*en bkd-kantoja, jotka ovat käyttökelpoisia uusien avermektiinien valmistamisessa.

Esimerkit

Seuraavissa kappaleissa on esitetty yksityiskohtaisia esimerkkejä *S. avermitilix*estä peräisin olevien bkd-geenien kloonaukseen ja analysoimiseen käytetyistä kokeellisista toimenpiteistä, joita on myös kuvattu keksintöön liitetyissä kuvioissa. Muita yksityiskohtia vakiomenetelmistä, jotka ovat molekyylibiologian alan ammattikokemuksen perusteella tunnettuja, sekä tiettyjen nimenomaisten käytettyjen entsyymien nimitykset on kuvattu esimerkiksi laboriokäsikirjassa "Molecular Cloning", Maniatis, et al. (Cold Spring Harbor Laboratory, 1989).

Esimerkki 1

***S. avermitilix*en genomisen DNA:n valmistus**

S. avermitilis ATCC 31272 -myseeliä kasvatettiin kiinnikasvaneena mattona YDP-2-agaralustalla seitsemän päivää 29 °C:ssa. Mediumin koostumus oli:

	Difco Yeast Extract	10 g
	Difco Bacto-peptone	10 g
	Glukoosi	5 g
	Difco Bacto agar	20 g
5	Natriumasetaatti	2 g
	MOPS	10 g

pH säädettiin arvoon 7,0.

Lopullinen tilavuus: 1 l.

Mediumia autoklavoitiin 25 minuuttia 121 °C:ssa.

- 10 Tämän jälkeen kasvanutta myseeliä ympättiin 30 ml:aan AS-7-mediumia [tätä on käsitelty julkaisussa Hafner, et al. (1988), EP-patenttihakemus nro 88 300 353.5, julkaisu nro 0 284 176] 300 ml:n vetoisessa sisälaipallissessa pullossa, jota kasvatettiin ravistellen (230 kierrosta minuutissa) 29 °C:ssa 24 tuntia. Mediumin koostumus oli:

	Ohennettu tärkkelys ¹	20 g
	Ardamine pH ²	5 g
	Pharmamedia ³	15 g
20	CaCO ₃	2 g

pH säädettiin arvoon 7,2 natriumhydroksidilla (NaOH).

Lopullinen tilavuus: 1 l.

Medium autoklavoitiin 25 minuuttia 121 °C:ssä.

- 25 ¹ Valmistettu hydrolysoimalla tärkkelystä *Bacillus licheniformiksesta* peräisin olevalla alfa-amylaasilla ja menetellen niin, että dekstroosiekvivalenttien määräksi saatiin noin 40 %.

² Yhtiöstä Yeast Products, Inc., Clifton, NJ 07012.

³ Yhtiöstä Traders Protein, Memphis, TN 38108.

30

Noin 0,3 ml edellä olevaa viljelmää ympättiin toiseen 300 ml:n tilavuiseen sisälaipalliseen pulloon, joka sisälsi 30 ml modifioitua nestemäistä Yeast Extract Malt Extract (YEME) -mediumia [Bibb, M.J., Freeman, R.F. ja

Hopwood, D.A., Mol. Gen. Genetics 154 (1977) 155 - 166].

Modifioitu YEME-medium sisälsi litraa kohti:

	Difco Yeast extract	3 g
	Difco Bacto-peptone	5 g
5	Oxoid Malt extract	3 g
	Sakkaroosi	300 g
	Glukoosi	10 g

Medium autoklavoitiin 40 minuuttia 121 °C:ssä.

2 ml 2,5 M $\text{MgCl}_2 \cdot 6 \text{H}_2\text{O}$:ta lisättiin autoklavoinnin jälkeen.

10 Lopullinen tilavuus säädettiin 1 l:ksi.

Viljelmiä kasvatettiin 48 - 72 tuntia 29 °C:ssa.

Myseeli otettiin talteen sentrifugoimalla ja genomisen DNA valmistettiin noudattaen valmistusohjetta "Isolation of Streptomyces Total DNA by Cesium Chloride Gradient Centrifugation: Procedure 2", sellaisena kuin se on oppikirjassa "Genetic Manipulation of Streptomyces, A Laboratory Manual", The John Innes Foundation, Norwich, U.K., 1985, jonka tekijät ovat Hopwood, D.A., et al. DNA-napit suspendoitiin uudelleen 3 ml:aan TE- puskuria (10 mM Tris-HCl, pH 8,0, 1 mM EDTA).

20

Esimerkki 2

S. avermitiliksen genomisen DNA:n monistuksen polymeraasiketjureaktio

S. avermitiliksen genomisen DNA monistettiin entsyymaattisesti käyttäen Perkin-Elmer Cetus thermal cycler-laitetta. PCR-reaktio suoritettiin käyttäen Taq-polymeraasia (Perkin-Elmer Cetus) ja valmistajan toimittamaa puskuria, joka sisälsi dNTP-molekyylejä, joiden konsentraatio oli 200 μM , 15 % glyserolia, kutakin aluketta, joiden konsentraatio oli 0,5 μM , 50 ng mallijuoste-DNA:ta (tässä tapauksessa S. avermitiliksen genomista DNA:ta) ja 2,5 yksikköä entsyymiä, 100 mikrolitran lopputilavuudessa 30 syklin ajan. Ensimmäisen syklin lämpöprofiili oli: kolme minuuttia 95 °C:n lämpötilassa (denaturaatiovaihe), 35 kaksi minuuttia 55 °C:n lämpötilassa (pariutumisvaihe) ja

kaksi minuuttia 72 °C:n lämpötilassa (pidentämisvaihe). Myöhempien 29 syklin lämpöprofiili oli samanlainen lukuun ottamatta sitä, että denaturaatiovaihe lyhennettiin 1,5 minuutiksi. DNA-alukkeet oli hankittu yhtiöstä Genosys Biotechnologies, Inc. (Texas). Oikealle suuntautuva aluke (kuvio 1) oli 5'-GAATTCGGCGACGGCGCCACCTCCGAGGGCGAC-3', ja vasemmalle suuntautuva aluke (kuvio 1) oli 5'-TCTAGACCGCAGGTGGTCCGGCATCTC-3'. Monistustuotteet fraktioitiin koon mukaan agarosigeelielektroforeesilla. PCR-näytteelle suoritettiin elektroforeesi vaakasuuntaisessa 1,5-%:isessa agarosigeelissä 1 X TBE-puskurissa [(90 mM Tris-HCl, pH 8,5, 90 mM boorihappo, 2,5 mM etyleenidi-amiinitetraetikkahappo (EDTA)] 1,5 tunnin ajan julkaisussa Maniatis et al. kuvatulla tavalla. Erotetut PCR-DNA-tuotteet paikannettiin geelissä värjäämällä etidiumbromidilla ja visualisoimalla fluoresoivat vyöhykkeet 365 nm:n ultraviolettivalolla. Edellä kuvatuissa PCR-olosuhteissa havaittiin tätä alukeyhdistelmää käytettäessä yksi DNA-vyöhyke (kooltaan noin 250 emäsparia).

20 **Esimerkki 3**

PCR:llä monistetun 0,25 ke:n suuruisen *S. avermitil*liksen genomisen DNA-fragmentin kloonaukseen *E. coli* -vektoriin ja sen transformaatio tämän jälkeen *E. coli* -isäntään

A. 0,25 ke:n suuruisen PCR-tuotteen talteenotto

25 Kuten edellä mainittiin, monistettiin 0,25 ke:n suuruinen DNA-fragmentti PCR:llä käyttäen mallijuosteena *S. avermitil*liksen genomista DNA:ta ja oikealle suuntautuvan ynnä vasemmalle suuntautuvan alukkeen yhdistelmää. Kuten kuviosta 1 käy ilmi, on oikealle suuntautuvassa alukkeessa EcoRI-tunnistuskohda, joka sijaitsee 5'-puoleisessa päässä, ja vasemmalle suuntautuvassa alukkeessa on 30 5'-puoleisessa päässä sijaitseva XbaI-tunnistuskohda. 0,25 ke:n suuruisen PCR-fragmentin kloonauksyritykset käyttäen ligaatiomenetelmää, jossa sekä liitosjakso että kloonauksvektori pilkottiin EcoRI:lla ja XbaI:lla, eivät kui-

tenkaan johtaneet suotuisiin tuloksiin. Tämän vuoksi tutkittiin vaihtoehtoista lähestymistapaa 0,25 ke:n suuruisen PCR-fragmentin kloonaukseksi, johon liittyi DNA-polymeraasi I:n Klenow-fragmentin käyttö PCR-fragmentin päiden tasoitukseen. Monistuksen jälkeen (tämä tapahtui esimerkiksi 2 kuvatulla tavalla) uutettiin noin 80 µl PCR-reaktioseosta kaksi kertaa fenolin ja kloroformin seoksella ja kaksi kertaa eetterillä, minkä jälkeen PCR-DNA-tuote saostettiin etanolilla edellä kuvatulla tavalla. DNA suspendoitiin uudelleen 18,5 µl:aan H₂O:ta. Tämän jälkeen seokseen lisättiin 2,5 µl 10 x katkostranslaatiopuskuria (0,5 M Tris-HCl, pH 7,2, 0,1 M magnesiumsulfaatti (MgSO₄), 1 mM ditiotreitoli, 500 µg/ml naudan seerumialbumiinia (Maniatis, et al., 1989) ja 20 yksikköä E. colin DNA polymeraasi I:n Klenow-fragmenttia (Boehringer Mannheim Biochemicals), ja seosta inkuboitiin 37 °C:ssa 5 minuuttia. Tämän jälkeen seokseen lisättiin 1 µl 2 mM dNTP-molekyyliä (kutakin neljää dNTP-molekyyliä 2 mM) ja reaktioseosta inkuboitiin vielä 15 minuuttia huoneenlämmössä. Korjausreaktio lopetettiin lisäämällä seokseen 1 µl 0,5 M EDTA:a, jonka pH oli 8,0, ja koko reaktioseoksen sisältö pipetoitiin 0,5-%:iseen agarosigeeliin ja tälle suoritettiin elektroforeesi. 0,25 ke:n suuruinen DNA-fragmentti visualisoitiin edellä kuvatulla tavalla ja otettiin talteen elektroeluoimalla seuraavasti: 0,25 ke:n suuruinen DNA-vyöhyke poistettiin partaterällä ja DNA otettiin talteen agarosigeelistä elektroeluoimalla 35 minuutin ajan 80 V:n jännitteellä V:n muotoiseen kuoppaan, joka sisälsi 10 M ammoniumasetaattia käyttäen yksisuuntaista elektroelointilaitetta (International Biotechnology Inc., New Haven, CT). DNA saostettiin sitten etanolilla, sentrifugoitiin napiksi ja liuotettiin uudelleen 20 µl:aan DNA-puskuria (10 mM Tris-HCl, 4 mM natriumkloridi (NaCl), 0,1 mM EDTA; pH 7,5).

B. Plasmidivektorin pGEM3Z:n pilkkominen SmaI:llä ja sen defosforylointi

Noin 1:tä µg plasmidia pGEM-3Zf(+) (Promega Corp., Madison, WI) ja kahta yksikköä SmaI-restriktioentsyymiä (5 (kaikki restriktioentsyymit oli hankittu yhtiöstä Boehringer Mannheim Biochemicals) inkuboitiin toimittajan antamien ohjeiden mukaisessa määrityspuskurissa 25 °C:ssa 3,5 tunnin ajan 40 mikrolitran (µl) kokonaisreaktiotilavuudessa lineaaristen päistään tasoitettujen molekyylien 10 aikaansaamiseksi. Tämän jälkeen SmaI:lla linearisoitu vektori defosforyloitiin käyttäen vasikan suolen alkalista fosfataasia (CIAP) (tämä oli hankittu yhtiöstä Promega Corp., Madison, WI) toimittajalta saatuja ohjeita noudattaen. Reaktioseosta inkuboitiin 35 minuuttia 37 °C:ssa ja 15 DNA uutettiin sitten seoksen kaksinkertaista tilavuutta vastaavalla määrällä fenolin ja kloroformin seosta, seoksen kaksinkertaista tilavuutta vastaavalla määrällä eetteriä ja DNA saostettiin lopuksi lisäämällä seokseen sen kaksinkertaista tilavuutta vastaava määrä absoluuttista 20 etanolia. Saostunut DNA otettiin talteen sentrifugoimalla 10 000 x g:n voimalla 10 minuutin ajan ja se kuivattiin vakuuissa. Lopullinen nappi liuotettiin uudelleen 20 µl:aan DNA-puskuria.

C. pCD503:n aikaansaamiseksi tehty ligaatio

25 Noin 9 µl Klenow-fragmentilla käsiteltyä 0,25 ke:n suuruista PCR-DNA-tuotetta ja noin 1 µl SmaI:llä linearisoitua ja CIAP:llä defosforyloitua päistään tasoitettua PGEM-3Zf(+):aa inkuboitiin yön yli käyttäen 1 yksikköä liigaasia (New England Biolabs, IC, Beverly, MA) toimittajan 30 kuvaamissa olosuhteissa 14 °C:ssa 20 µl:n kokonaisreaktiotilavuudessa. Reaktio lopetettiin siirtämällä määritysmikrokoeputki jäihin, minkä jälkeen kompetentit E. coli JM109-solut transformoitiin 15 µl:lla reaktioseosta noudattaen vakiomenetelmää, joka on kuvattu teoksessa Maniatis, et al., 1989. Saatiin talteen useita ampisilliinille resis-

tenttejä transformantteja. Plasmidivektori pGEM-3Zf(+) sisältää *Escherichia coli* lac-operonista peräisin olevan DNA-segmentin, joka koodaa beeta-galaktosidaasin aminotermi-
 5 naalista fragmenttia [Yanisch-Perron, C., Vieira, J. ja Messing, J., *Gene* 33 (1985) 103]. Tämä fragmentti, jonka synteesi voidaan indusoida isopropyyli-tio-beeta-D-galak-
 tosidillä (IPTG), kykenee intra-alleeliseen (alfa-) kom-
 plementaatioon isännän koodaaman vaillinaisen beeta-galak-
 tosideasin muodon kanssa. IPTG-indusorilla käsiteltyt *E.*
 10 *coli* -solut syntetisoivat molempia entsyymifragmentteja ja muodostavat sinisiä pesäkkeitä maljattuina ravinto-
 alustoille, jotka sisältävät kromogeenista substraattia, 5-bromi-4-kloori-3-indolyyli-beeta-D-galaktosidia (X-gal).
 Vieraan DNA:n liittäminen plasmidin monikäyttöiseen kloo-
 15 nauskohtaan inaktivoi beeta-galaktosidaasin aminotermi-
 nialisen fragmentin ja hävittää kyvyn alfa-komplementaatioon. Tämän vuoksi yhdistelmäplasmideja sisältävät bakteerit
 muodostavat valkeita pesäkkeitä. Tästä transformaatioko-
 keesta saatiin talteen useita valkeita pesäkkeitä. Näiden
 20 pesäkkeiden pitäisi sisältää plasmidi pCD503. Tämä varmis-
 tettiin valikoimalla yksi pesäke, jolle annettiin nimitys CD503-kanta, ja tutkimalla se edelleen. Yksi *E. coli* -kan-
 nan CD503 bakteeripesäke ympättiin tavallisia mikrobiolo-
 gisia menetelmiä noudattaen Luria-Bertani (LB) -mediumiin,
 25 joka sisälsi 50 µg ampicilliinia millilitrassa. LB-mediumin koostumus oli

Bacto-tryptone	10 g
Bacto-yeast extract	5 g
NaCl	10 g

30 pH säädettiin arvoon 7,0 5 N natriumhydroksidilla (NaOH).
 Liuoksen lopullinen tilavuus säädettiin 1 l:ksi.
 Liuos steriloidiin autoklavoimalla 20 minuuttia 121 °C:n
 lämpötilassa.

Viljelmää inkuboitiin 35 °C:ssa yön yli. Seuraavana
 35 aamuna bakteerisolut koottiin talteen sentrifugoimalla no-

peudella 10 000 kierrosta minuutissa 5 minuutta 4 °C:ssa. Plasmidivektori eristettiin juuri talteen otetuista *Escherichia coli* CD503 -soluista käyttäen Birnboimin ja Dölyn menetelmästä [Nucleic Acids Res. 7 (1979) 1513] tehtyä muunnosta, joka on kuvattu julkaisussa Denoya, et al., Microbios Lett. 29 (1985) 87. Eristettyä plasmidi-DNA:ta liuotettiin lopuksi DNA-puskuriin (10 mM Tris-HCl, 4 mM NaCl, 0,1 mM EDTA; pH 7,5) sellainen määrä, että sen konsentraatioksi saatiin noin 1 µg pCD503:a 10 µl:ssa puskuria. Varmistukseksi suoritettu restriktioanalyysi käyttäen EcoRI:tä ja PstI:tä osoitti odotetusti, että pCD503 sisälsi 0,25 ke:n suuruisen DNA-liitosjakson.

Esimerkki 4

Radioaktiivisella leimalla varustettujen DNA- ja RNA-koettimien valmistus

A. Yhtenäisellä radioaktiivisella leimalla varustettujen kaksijuosteisten DNA-koettimien valmistus

Kaksijuosteiset DNA-koettimet valmistettiin katkos-translaatiolla (tätä on käsitelty yleisesti julkaisussa Maniatis, et al., 1989). Ensimmäiseksi valmistettiin kohdesekvenssin sisältävä spesifinen DNA-fragmentti käyttäen asianmukaisella restriktioentsyymillä suoritettua pilkkomista ja puhdistusta elektroeluoimalla olennaisesti esimerkissä 1 kuvatulla tavalla. Noin 1 µg DNA:ta varustettiin kussakin tapauksessa leimalla käyttäen [alfa-³²P]-dCTP:tä (deoksisytidiini-5'-trifosfaatti, tetra(trietyyliammonium)suola, [alfa-³²P]-), joka oli hankittu yhtiöstä NEN-Dupont, ja BRL:n Nick Translation System -reagenssipakkausta, joka oli hankittu yhtiöstä BRL Life Technologies, Inc., noudattaen toimittajalta saatuja ohjeita. Tyypillinen reaktio suoritettiin 50 µl:n tilavuudessa. Kun seokseen oli lisätty 5 µl lopetuspuskuria (BRL:n suosittelemassa suoritushjeessa kuvatulla tavalla), erotettiin leimalla varustettu DNA koettimeen liittymättömistä nukleotideista käyttäen Stratagenen ulosajo (Push) -pylvästä

toimittajalta saatuja ohjeita noudattaen. Näitä suoritushojeita noudattaen saatiin rutiininomaisesti ^{32}P -leimalla varustettua DNA:ta, jonka spesifinen aktiivisuus oli huomattavasti yli 10^8 cpm/ μg .

5 **B. Leimalla varustettujen yksijuosteisten RNA-koettimien valmistus**

^{32}P -leimalla varustetut RNA-koettimet valmistettiin käyttämällä in vitro -transkriptiota Riboprobe Gemini-transkriptiojärjestelmän (Promega) avulla. Puhdistettu kohde-DNA-fragmentti kloonattiin transkriptiovektoriin pGEM-3Z käyttäen vakiomenetelmiä. Mallijuosteplasmidin DNA:n valmistus in vitro olosuhteissa suoritettavaa transkriptioreaktioita varten suoritettiin esimerkissä 3 kuvatulla tavalla, mutta suoritukseen lisättiin polyetyleeniglykoli (PEG) -saostusvaihe sellaisten pienten nukleotidien poistamiseksi selektiivisesti, jotka voivat kontaminoida näitä preparaatteja, menetellen seuraavasti: etanolisaostusvaiheen jälkeen nappi suspendoitiin uudelleen 520 μl :aan vettä. Tämän jälkeen seokseen lisättiin 100 μl 5 M NaCl:ää ja 620 μl 13-%:ista PEG:tä (mp. 6 000 - 8 000). Kun putkea oli ensin sekoitettu, niin sitä inkuboitiin sitten jäällä yhden tunnin ajan ja DNA sentrifugoitiin napiksi 4 °C:ssa 10 000 x g:n voimalla 15 minuutin ajan. Nappi pestiin kerran 500 μl :lla 80-%:ista kylmää etanolia ja suspendoitiin uudelleen tavanomaiseen tapaan. Noin 1 μg mallijuosteena käytettyä plasmidi-DNA:ta lineaarisoiitiin käyttäen joko SacI- tai HindIII-restriktioentsyymejä ja tämän jälkeen transkriptoititiin in vitro -olosuhteissa käyttäen SP6- tai T7-bakteriofagin DNA:sta riippuvaista RNA-polymeraasia, mainitussa järjestyksessä ilmoitettuna. Tässä reaktiossa käytettiin sytidiini-5'-trifosfaatti-tetra(trietyyliammonium)suolaa [α - ^{32}P] (CTP), joka oli hankittu yhtiöstä NEN-dupont. Suorituksessa noudatettiin toimittajan suosittelemia reaktio-olosuhteita.

35 Tämän inkuboinnin jälkeen reaktioseosta käsiteltiin 1 yk-

siköllä RQ1-DNasea (Promega) DNA-templaatin pilkkomiseksi, seos uutettiin kaksi kertaa fenolin ja kloroformin seoksella ja saostettiin tämän jälkeen etanolilla vakiomenetelmiä noudattaen. Nappi kuivattiin ja suspendoitiin uudelleen 20 µl:aan vettä, joka ei sisältänyt RNAasia (Promega). Pieni määrä leimalla varustettua RNA-transkriptiotuotetta analysoitiin polyakryyliamidiagarosigeelielektroforeesilla julkaisussa Denoya, et al., J. Bacteriol. 169 (1987) 3857 - 3860 kuvatulla tavalla. Kuvatuissa olosuhteissa saatiin rutiininomaisesti leimalla varustettuja täysimittaisia transkriptiotuotteita.

Esimerkki 5

S. avermitiloksen genomisen DNA:n analyysi Southern-hybridisaatiolla

Noin 10 µg puhdistettua S. avermitiloksen genomista DNA:ta pilkottiin 2 yksiköllä BamHI-restriktioentsyymiä 37 °C:ssa vähintään 2 tunnin ajan. Pilkkomisvaiheen lopulla DNA-fragmentit erotettiin elektroforeesilla 1-prosenttisen agarosigeelin läpi (ks. esimerkki 1A) ja ne siirrettiin yön yli nailonmembraanille (huokoskoko 0,45 µm) (Schleicher and Schuell -yhtiön Nytran-membraanit) käyttäen kapillaarisiirtomenetelmää [Southern, E.M., J. Mol. Biol. 98 (1975) 503]. Seuraavana päivänä nailonmembraanit käärittiin muovikääreeseen ja kunkin membraanin DNA-puoli käsiteltiin ultraviolettisäteilylähteellä (302 nm) DNA:n kiinnittämiseksi membraaniin. Radioaktiivisella leimalla varustettujen RNA- tai DNA-koettimien hybridisaatio nailonmembraaneille immobilisoituun DNA:han suoritettiin noudattaen suoritussvaiheita, jotka on kuvattu teoksessa Maniatis et al. (1989). Esihybridisaatio ja hybridisaatio suoritettiin 42 °C:ssa. Hybridisaatioliuoksen koostumus oli 6 x SSC:tä (1 x: 0,15 M natriumkloridi (NaCl), 15 mM natriumsitraatti, pH 7,0), 10 x Denhardtin reagenssi [1 x: 0,02 % ficoll, 0,02 % polyvinyylipyrrolidoni, 0,02 % naudan seerumialbumiini], 1 % SDS:ää (natrium dodekyylisul-

faatti), 100 µg/ml denaturoitua ja fragmentoitua lohen maidin DNA:ta, 100 µg/ml E. colin tRNA:ta ja 50 % formamidia (Fluka). Yön yli suoritettun hybridisaation jälkeen membraanit pestiin seuraavasti: kaksi pesua seoksella, joka sisälsi 1 x SSC:tä ja 0,1 % SDS:ää, huoneenlämmössä 15 minuutin ajan, kaksi pesua seoksella, joka sisälsi 0,1 x SSC:tä ja 0,1 % SDS:ää, 42 °C:ssa 15 minuutin ajan. Joissakin kokeissa hybridisaatio suoritettiin 65 °C:ssa ilman formamidia ja SSC:n tilalla käytettiin SSPE-seosta (1 x: 0,18 M NaCl, 10 mM natriumfosfaattia (NaPO₄), pH 7,7, 1 mM EDTA). Lopuksi röntgenfilmi valotettiin membraaneilla autoradiografisen kuvan aikaansaamiseksi.

Esimerkki 6

S. avermitiliksen genomisen 0,25 ke:n suuruisen CD503-fragmentin kloonaukseen M13-bakteriofagiin ja DNA:n sekvenssointi

0,25 ke:n suuruinen S. avermitiliksen CD503-DNA-fragmentti kloonattiin bakteriofageihin M13mp18 ja M13mp19 sellaisen yksijuosteisen yhdistelmä-DNA:n valmistamiseksi, jota oli määrä käyttää mallijuosteina Sangerin dideoksi-sekvenssointimenetelmässä [Sanger, et al., Proc. Nat. Acad. Sci. USA 74 (1977) 5463 - 5467]. Noin 2 µg plasmidia pCD503, joka oli valmistettu edellä kuvattua minipreparaattimenetelmää käyttäen, pilkottiin reskriktioentsyymeillä EcoRI ja PstI 0,25 ke:n suuruisen S. avermitiliksen genomisen liitosjakson vapauttamiseksi. Aikaisemmin suoritettu restriktioanalyysi oli osoittanut, että EcoRI:lla tai PstI:lla pilkottu pCD503 oli lineaarinen. Tämä analyysi osoitti, että 0,25 ke:n suuruinen liitosjakso ei sisältänyt EcoRI:n eikä PstI:n tunnustuskohtia. Pilkkomiseseokselle suoritettiin elektroforeesi 1,2-%:isessa agarosegeelissä, minkä jälkeen 0,25 ke:n suuruinen fragmentti elektroeluoitiin ja saostettiin edellä kuvatulla tavalla. Tämän lisäksi noin 1 µg kutakin M13mp18:n ja M13mp19:n DNA-molekyylien puhdistettua kaksijuosteista replikaati-

vista muotoa (RF) pilkottiin kahdella restriktioentsyymillä käyttäen EcoRI:tä ja PstI:tä, defosforyloitiin vasikan suolen alkalisella fosfataasilla (CIAP) (tämä oli hankittu yhtiöstä Promega Corp., Madison, WI), ja ligatoitiin lopuksi 0,25 ke:n suuruiseen DNA-fragmenttiin aikaisemmin kuvatulla tavalla. Puhdistetut RF M13-kloonausvektorit oli hankittu yhtiöstä New England Biolabs. Ligaatioseosta käytettiin kompetenttien *E. coli* JM109 -solujen transfektioon. mp18-transfektioista valikoitiin yksi valkea plakki ja mp19-transfektioista yksi plakki, faagit kasvatettiin ja yksijuosteinen DNA valmistettiin aiemmin kuvatulla tavalla (Maniatis, et al., 1989). Kunkin yksijuosteisen DNA-mallijuosteen DNA-sekvenssointi suoritettiin käyttäen M13-spesifistä -40-sekvenssointialuketta (New England Biolabs, tuoteluettelon nro 1212), deoksiadenosiini-5'-[alfa-tio]trifosfaattia, [³⁵S] (NEN-Dupont), ja TaqTrack-sekvenssointireagenssipakkausta (Promega) noudattaen toimittajalta (Promega) saatuja ohjeita. *S. avermitiliks*en genomisen pCD503-fragmentin DNA-sekvenssointikoetulokset on esitetty kuviossa 4.

Esimerkki 7

Koko *S. avermitiliks*en bkd-geeniryhmittymän kloonauus ja kromosomikartan konstruktio

Noin 5 µg puhdistettua pCD503:a käsiteltiin kahdella restriktioentsyymillä käyttäen sekä BamHI- että EcoRI-restriktioentsyymejä, DNA-fragmentit erotettiin elektrofooresilla 1,2-%:isessa agarosigeelissä, noin 0,25 ke:n suuruinen DNA-fragmentti, jossa oli *S. avermitiliks*en bkd El-alfa-geenille spesifinen sekvenssi, otettiin talteen elektroeluoimalla ja varustettiin leimalla käyttäen katkostranslaatiota olennaisesti edellä kuvatulla tavalla. Tämän jälkeen ³²P-leimalla varustettua DNA-fragmenttia käytettiin sitten koettimena *S. avermitiliks*en genomikosmidikirjaston seulomiseksi. Yksityiskohtainen kuvaus genomisten kirjastojen yleisestä valmistustavasta on esitetty

teoksessa *Molecular Cloning, A Laboratory Manual*, Maniatis, et al. (1989). Täydellinen kuvaus streptomykeettien kromosomaalisen kirjaston valmistuksesta on esitetty teoksessa *Genetic Manipulation of Streptomyces - A Laboratory Manual*, Hopwood, et al. (1985). Kosmidivektoria koskeva kuvaus on julkaisussa "Cosmid Vectors for Streptomyces Genomic DNA Cloning", Denoya C.D., US-patenttihakemus 08/048 719, jätetty 16.4.1993. Neljä kloonia tunnistettiin yli 2 200 yhdistelmäkirjastokloonin seulonnan jälkeen. Näitä neljää hybridisoituvaa kloonia (jotka oli nimetty E. coli -klooneiksi CD518, CD519, CD520 ja CD521) kasvatettiin LB-mediumissa ampisilliiniselektiopaineessa. Kustakin viljelmästä valmistettiin plasmidit edellä kuvatulla tavalla. Restriktio- ja Southern-blottihybridisaatioanalyysissä kävi ilmi, että nämä neljä kloonia olivat toisiaan muistuttavia ja että niiden kromosomialueet menivät limittein toistensa suhteen. S. avermitiloksen genomisen restriktiokartta, joka kattaa koko kromosomaalisen alueen mukaan lukien CD503-sekvenssin, saatiin tavanomaisia menetelmiä noudattaen ja se on esitetty kuviossa 3.

Esimerkki 8

Sisäkkäisten deleetiomutanttiryhmien valmistaminen bkd-geeniryhmittymän sisältävän S. avermitiloksen kromosomaalisen alueen tiettyyn suuntaan suunnattua DNA-sekvensointia varten

Eksonukleaasi III:a käyttäen muodostettiin sisäkkäisiä deleetiomutanttiryhmiä, joilta puuttuu tasaisesti kasvava määrä nukleotideja S. avermitiloksen bkd-kohde-DNA-molekyylien toisesta tai toisesta päästä, noudattaen menetelmiä, jotka olivat olennaisesti samanlaisia kuin julkaisussa Henikoff, S., *Methods Enzymol.* 155 (1987) 156, kuvatut menetelmät. Yksisuuntaisten deleetiomutaatioiden aikaansaamiseksi pilkottiin kunkin yhdistelmä-bakteriofagi M13:n replikatiivisen muodon DNA:n kaksijuosteista DNA:ta kahdella restriktioentsyymillä, joiden kummankin pilkko-

miskohdat sijaitsevat kohde-DNA:n toisen pään ja aluketta sitovan kohdan välissä. Restriktioentsyymi, joka pilkkoi lähimpänä kohdesekvenssejä, sai aikaan tasaisen pään tai sisään vetäytyneen 3'-puoleisen pään; toinen entsyymi muodosti vapaaksi jäävän 3'-pään. Eksonukleaasi III katalysoi 5'-puoleisten mononukleotidien vaiheittaista irtoamista kaksijuosteisen DNA:n sisäänvetäytyneistä tai tasaisista 3'-hydroksyyllipäistä. Vapaaksi jäävät 3'-puoleiset päät ovat kuitenkin täysin resistenttejä tämän entsyymin aktiivisuutta kohtaan. Tästä syystä ainoastaan toinen tulokseksi saadusta lineaarisesta DNA:sta oli herkkä eksonukleaasi III:lle ja pilkkoutuminen eteni yksisuuntaisesti pilkkoutumiskohdasta kohde-DNA sekvenssien suuntaan.

Esimerkkinä kuvataan seuraavaksi pCD565:n sisäkkäisten deletioiden valmistus. Plasmidi pCD565 on M13mp19 RF:n johdannainen, joka sisältää 1,15 ke:n suuruisen Sallifragmentin, joka sisältää osan *S. avermitilix*en bkd:n avointa E1-alfa-lukukehystä. Plasmidi pCD565 puhdistettiin tasapainosentrifugointia käyttäen cesiumkloridi- ja etidiumbromidigradien-teissa teoksessa Maniatis, et al. (1989) kuvatulla tavalla. Eksonukleaasi III kykenee aloittamaan pilkkomisensa yksijuosteisista katkoskohdista, joten on tärkeää käyttää preparaattia, joka sisältää alle 10 % relaxoituneita rengasmaisia molekyylejä. Noin 10 µg plasmidia pCD565 (ks. kappale "keksinnön yksityiskohtainen kuvaus") pilkottiin kahdella restriktioentsyymillä käyttäen entsyymejä SacI ja XbaI 37 °C:ssa 4 tunnin ajan, minkä jälkeen seos uutettiin fenolin ja kloroformin seoksella ja eetterillä, minkä jälkeen molekyylit saostettiin etanolilla aikaisemmin kuvatulla tavalla. Nappi suspendoitiin uudelleen 60 µl:aan eksonukleaasi III-reaktiopuskuriin (10 x eksonukleaasi III -puskuri; 0,66 M Tris-HCl, pH 8,0, 66 mM magnesiumkloridi (MgCl₂)). DNA-liuosta inkuboitiin sitten 37 °C:ssa käyttäen seoksessa 300 yksikköä eksonukleaasi III:a (Ambion Inc.), ja seoksesta poistettiin 30 sekunnin

välein 2,5 µl:n suuruisia näytteitä. Tämän jälkeen näytteitä inkuboitiin S1-nukleaasin kanssa ja kustakin näytteestä otetut näyte-erät analysoitiin agarosigeelielektroforeesilla. Halutun suuruisia DNA-fragmentteja sisältävät näytteet yhdistettiin, DNA korjattiin käyttäen DNA-polymeraasi I:n Klenow-fragmenttia, ligatoitiin yön yli ja transfektoitiin kompetentteihin *E. coli* JM109 -soluihin. Talteen saatujen kloonien sisältämän liitosjakson koko analysoitiin käyttämällä EcoRI/HindIII-restriktioentsyymikäsittelyä ja agarosigeelielektroforeesia. Sekvenssointiin valittiin 5 kloonia 565-D19 (1,1 ke), 565-D7 (0,88 ke), 565-d24 (0,77 ke), 565-D1 (0,51 ke) ja 565-D16 (0,36 ke). Yksijuosteinen DNA valmistettiin kustakin näistä klooneista ja se sekvenssoitiin edellä kuvatulla tavalla.

15 **Esimerkki 9**

Plasmidien pCD670, pCD666, pCD736 ja pCD685 konstruktio käytettäväksi *S. avermitiliks*en bkd-geenien ilmentämisessä *E. coli*ssa

*S. avermitiliks*en bkd-geenien ilmentäminen *E. coli*ssa suoritettiin käyttäen T7-RNA polymeraasi/promoottori-kaksoisplasmidijärjestelmää olennaisesti julkaisussa Tabor, S. (1990, teoksessa *Current Protocols in Molecular Biology*, pp. 16.2.1 - 16.2.11. Greene Publishing and Wiley-Interscience, New York) kuvatulla tavalla.

25 **A. *S. avermitiliks*en bkd:n E1-alfa-ORF:n sisältävän pCD670:n konstruktio**

*S. avermitiliks*en bkd:n E1-alfa-geeniin liitettiin ATG-aloituskodonin kattava NdeI-restriktiokohta. PCR-mallijuoste oli plasmidi pCD528, joka on sellainen pGEM-3Z:n johdannainen, joka sisältää 7 ke:n suuruisen *S. avermitiliks*en genomisen liitosjakson, joka sisältää E1-alfa-ORF:n aminotermiinaalisen puoliskon. PCR-reaktiossa käytettiin alukkeina kahta oligonukleotidia (ks. kuvio 6):

1. Vasempaan suuntautunut universaalinen (vektori)
35 aluke 31-PCR-BP

(5'-AAGGATCCTGCAGACAGCTATGACCATGATTACGCCA-3'), joka kartoittuu pGEM-3Z:n MCS-kohdan HindIII-kohdasta myötäsuntaan (asema 91 - 114). Tämän alukkeen 5'-päässä on kaksi A:ta ja kaksi restriktiokohtaa (BamHI ja PstI) PCR-tuotteiden kloonauksen edistämiseksi.

2. Mutageeninen alue 55-PCR

(5'-AAGAGATCTCATATGACGGTCATGGAGCAGCGG-3'.).

Tämän alukkeen 5'-päässä on kaksi A:ta, yksi G ja kaksi restriktiokohtaa (BglII ja NdeI). NdeI-kohta menee osittain päällekkäin E1-alfan avoimen lukukehyyksen ATG-aloituskodonin kanssa. Suoritettiin polymeraasiketjureaktio edellä kuvatulla tavalla. Reaktiotuotteen analysoitiin elektroforeesilla 0,8-%:isessa agarosigeelissä. PCR:lla monistettu DNA-fragmentti, joka oli kooltaan oikea (noin 1,1 ke:n pituinen), elektroluoitiin, pilkottiin restriktioensyymeillä NdeI ja BamHI ja jatkokloonattiin NdeI/BamHI-seoksella linearisoituun plasmidiin pT7-7 plasmidin pCD663 muodostamiseksi ligaation ja E. coli DH5-alfa -soluihin suoritettun transformaaion jälkeen. Noin 1 µg plasmidia pCD663 (valmistettu E. coli- kannasta CD663 käyttäen miniprep-plasmidin valmistusmenetelmää) linearisoitiin BamHI:lla, defosforyloitiin ja ligatoitiin lopuksi käyttämällä seoksessa noin 0,5 µg elektroluoitua puhdistettua 1,1 ke:n suuruista BamHI-fragmenttia, joka oli eristetty plasmidista pCD550 pilkkomalla tämä BamHI:lla, plasmidin pCD670 muodostamiseksi. 1,1 ke:n suuruisen BamHI-fragmentin oikea orientaatio viimeksi mainitussa konstruktiossa määritettiin liitosjaksossa olevien SalI-kohtien kartoituksella. Lopuksi plasmidi pCD670 liitettiin E. coli -kantaan C600, joka sisälsi plasmidin pGP1-2 (tämä plasmidi sisältää T7:n RNA-polymeraasigeenin) (tätä on käsitelty julkaisussa Tabor, 1990). Jatkotutkimuksia varten valittiin yksi transformantti, ja sille annettiin nimitys CD676.

B. S. avermitiliksen bkd:n El-beeta-ORF:n sisältävän pCD666:n konstruktio

S. avermitiliksen bkd:n El-beeta-geeniin liitettiin ATG-aloituskodonin kattava NdeI-restriktiokohta käyttäen PCR:ään perustuvaa menetelmää. PCR-mallijuoste oli plasmidi pCD574, joka on sellainen pGEM-3Z:n johdannainen, joka sisältää 4,5 ke:n suuruisen S. avermitiliksen genomisen liitosjakson, joka sisältää El-beetan ORF:n. Kaksi PCR-reaktiossa alukkeina käytettyä oligonukleotidia olivat (ks. myös kuvio 6):

1. Vasemmalle suuntautuvaa universaali (vektori-) aluke 30-PCR-BP:

(5'-AAGGATCCTGCAGCCCAGTCACGACGTTGTAAAACGA-3') kartoittuu pGEM3Z MCS:n EcoRI-kohdan suhteen vastasuuntaan (paikkaan 2689-2712). Tämän alukkeen 5'-päässä on kaksi A:ta ja kaksi restriktiokohtaa (BamHI ja PstI) PCR-tuotteiden kloonauksen edistämiseksi.

2. Oikealle suuntautuva mutageeninen aluke 56-PCR:

(5'-AAGAGATCTCATATGACCACCGTTGCCCTGAAG-3'). Tämän alukkeen 5'-päässä on kaksi A:ta, yksi G ja kaksi restriktiokohtaa (BglII ja NdeI). NdeI-kohta menee päällekkäin El-beetan avoimen lukukehyksen ATG-aloituskodonin suhteen. Polymeeraasiketjureaktio suoritettiin edellä kuvatulla tavalla. Reaktiotuotteet analysoitiin käyttämällä elektroforeesia 0,8-%:isessa agarosigeelissä. PCR:lla monistettu DNA-fragmentti, joka oli kooltaan oikea (noin 1,9 ke:n pituinen), elektroeluoitiin, pilkottiin restriktioensyymeillä NdeI ja EcoRI ja jatkokloonattiin NdeI/EcoRI-seoksella linearisoituun plasmidiin pT7-7 plasmidin pCD666 muodostamiseksi ligaation ja E. coli DH5-alfa-soluihin suoritettun transformatian jälkeen. Lopuksi plasmidi pCD666 liitettiin E. coli -kantaan C600, joka sisälsi plasmidin pGP1-2 (plasmidi, joka sisältää T7:n RNA-polymeraasigeenin) (ks. Tabor, 1990). Jatkotutkimuksiin valittiin yksi transformantti ja sille annettiin nimitys CD673-kanta.

C. S. avermitiloksen bkd:n E1-alfa- ja E1-beeta-ORF:t sisältävän pCD736:n konstruktio

Noin 2 µg plasmidia pCD670 linearisoitiin pilkkomalla se osittain BamHI:lla. Plasmidin pCD670 lineaarisen muodon aikaansaamiseksi otettiin näytteitä BamHI-pilkkomisseksestä ajankohtina, jotka olivat 1, 3, 5, 10 ja 20 minuuttia. Näytteille suoritettiin ajo 0,8-%:isessa agarosigeelissä. Lineaarinen muoto (noin 4,3 ke:n pituinen) otettiin talteen elektroeluoimalla ja tämä defosforyloitiin käyttäen CIAP:tä (menetellen edellä kuvatulla tavalla). Tämän jälkeen puolet plasmidin pCD670 defosforyloidusta lineaarisesta muodosta ligatoitiin plasmidista pCD577 eristettyyn 0,8 ke:n suuruiseen BamHI/BglII-fragmenttiin. Ligaatioseosta käytettiin kompetenttien E. coli DH5-alfa -solujen transformointiin. Otettiin talteen 10 kloonia, jotka analysoitiin miniprep-menetelmällä valmistetun plasmidi-DNA:n restriktioanalyysillä. Yksi kloon, jolle annettiin nimitys CD736-kanta, sisälsi oikealla tavalla kokoonpannun plasmidin pCD736. Lopuksi plasmidi pCD736 liitettiin E. coli C600 -kantaan, joka sisälsi plasmidin pGP1-2. Jatkotutkimuksia varten valittiin yksi transformantti ja sille annettiin nimitys CD737-kanta. Vielä yksi kloon, jolle annettiin nimitys CD705-kanta, sisälsi plasmidin pCD705, jossa oli 0,8 ke:n suuruinen BamHI/BglII-fragmentti väärässä orientaatioissa. Konstruktiota pCD705 käytettiin negatiivisena kontrollina ilmentämiskokeessa.

D. S. avermitiloksen bkd-geeniryhmittymän sisältävän pCD685:n konstruktio

Jäljellä oleva toinen puoli plasmidin pCD670 defosforyloitua lineaarista muotoa, joka oli saatu aikaan edellä kuvatulla tavalla pilkkomalla osittain restriktioentsyymillä BamHI, ligatoitiin plasmidista pCD577 eristettyyn 7 ke:n suuruiseen BamHI-fragmenttiin. Ligaatioseosta käytettiin kompetenttien E. coli DH5-alfa -solujen transfor-

mointiin. Saatiin talteen useita klooneja, joista valittiin jatkotutkimukseen 16 kloonia. Plasmidi-DNA uutettiin ja se analysoitiin restriktioanalyysillä. Yksi kloon, jolle annettiin nimitys CD685, sisälsi oikealla tavalla kokoonpannun plasmidin (pCD685). Lopuksi pCD685-plasmidi liitettiin *E. coli* C600 -kantaan, joka sisälsi pGP1-2-plasmidin. Jatkotutkimuksia varten valittiin yksi transformantti ja tälle annettiin nimitys CD687-kanta.

Esimerkki 10

10 **S. avermitiloksen bkd-geenien ilmentäminen *Escherichia colissa* käyttäen T7-kaksoisplasmidijärjestelmää**

E. coli C600:n (pGP-1) johdannaisia, jotka sisälsivät eri pT7-7-konstruktioita (kannat CD676, CD673, CD737 ja CD687), kasvatettiin yön yli 30 °C:ssa 5 ml:ssa LB-mediumia, joka sisälsi sekä kanamysiiniä (60 µg/ml) että ampisilliiniä (60 µg/ml). Yön yli kasvatetut viljelmät laimennettiin sitten suhteessa 1:40 (0,25:10,00 ml) viljeltäväksi koeputkessa (25 x 150 mm), joka sisälsi tuoretta LB/ampisilliini/kanamysiini-mediumia, ja sitä kasvatettiin ravistelijassa ilmastaen 30 °C:n lämpötilassa siihen asti, kunnes määritetty optinen tiheys (OD₅₉₀) oli noin 0,4. T4-RNA-polymeraasigeeni indusoitiin kohottamalla lämpötila 42 °C:seen 30 minuutin ajaksi, joka puolestaan indusoi T7-promoottorin ohjaaman(t) geenin(t) (julkaisussa Tabor, S., 1990, kuvatulla tavalla). Lämpötila alennettiin lopuksi 37 °C:seen ja soluja kasvatettiin vielä 90 minuuttia ravistellen. Indusoimattomia kontrolliviljelmiä pidettiin koko ajan 30 °C:ssa. Proteiinit analysoitiin natriumdodekyyli sulfaatti (SDS) - polyakryyliamidigeelielektroforeesilla julkaisussa Denoya, C.D., et al., J. Bacteriol. 168 (1986) 1133 - 1141 kuvatulla tavalla. Entsyymiaktiivisuus analysoitiin esimerkissä 11 kuvatulla tavalla.

Esimerkki 11

S. avermitiloksen BCKDH:n E1-aktiivisuuden määrittäminen yhdistelmä-E. coli -kantojen raakauutteissa

A. Solulysaatin valmistus

5 Solut (nämä olivat peräisin 8 ml:n suuruisista viljelmistä) koottiin talteen sentrifugoimalla (5 minuuttia nopeudella 5 000 kierrosta minuutissa - 3 000 x g - käyttäen Sorvall SS-34 -roottoria, joka oli jäädytetty lämpötilaan 4 °C), ja solut suspendoitiin uudelleen 5 ml:aan

10 "hajotuspuskuria" (0,05 M kaliumfosfaattipuskuri, pH 7,0, joka sisältää 3 % Triton X-100:aa, 15 % glyserolia, 3 mM ditiotreitolia, 1 mg/ml kalkkunan munanvalkuaisen trypsiini-inhibiittoria, 5 mM EDTA:a, ja 0,04 mM TPP:tä [tiamiini-pyrofosfaattia]). Uudelleen suspendoidut solut siirrettiin

15 French-puristimeen ja solut hajotettiin johtamalla ne laitteen läpi kerran 34 mPa:n (5 000 x psi:n) paineessa. French-puristimesta saatu 1,5 ml:n suuruinen näyte siirrettiin tämän jälkeen mikrosentrifugiputkeen ja se selkeytettiin 30 sekunnin mittaisella sentrifugoinnilla nopeudella 14 000 kierrosta minuutissa. Kuhunkin entsyymimääritykseen käytettiin kustakin supernatantista otettuja 100 µl:n suuruisia näytteitä. Proteiinikonsentraatio määritettiin Bio-Rad-proteiinimäärityksellä (Bio-Rad Laboratories, Richmond, Kalifornia), joka perustuu Bradfordin

20 kehittämään väriaineen sitoutumismenetelmään [Bradford, M., Anal. Biochem. 72 (1976) 248].

B. S. avermitiloksen haaraketjuisten alfa-ketohappojen dehydrogenaasi (BCKDH) -kompleksin E1-komponentin määrittäminen

30 BCKDH:n E1-aktiivisuus määritettiin modifioidulla muunnoksella radiokemiallisesta määrittäyksestä, joka on kuvattu aikaisemmin [Chuang, D.T., Methods Enzymol. 166 (1988) 146 - 154; ja Hafner, E.W., et al., J. Antibiotics 44 (1991) 349]. 15 ml:n vetoisen lasista valmistetun tuikepullon pohjalle lisättiin: 0,148 ml 0,25 M kaliumfos-

35

faattipuskuria, pH 6,5; 0,002 ml 0,1 M etyleenidiamiini-tetraetikkahappoa (EDTA, dinatriumsuola); 0,004 ml 0,1 M MgCl_2 :ta; 0,02 ml 3,7 mM tiamiinipyrofosfaattia (TPP); 0,02 ml 37 mM Na_2SO_4 :ää, 0,01 ml 37 mM 2,6-dikloorifenoli-indofenolia (natriumsuola, Sigma D-1878); 0,008 ml alfa-[1- ^{14}C]-ketoisokaproaatin kantaliuosta (valmistettu edellä kuvatulla tavalla); 0,058 ml vettä; ja 0,1 ml selkeytettyä solutonta uutetta. Pullon suu peitettiin välittömästi Whatman 4CHR-paperilla (Whatmanin tuoteluettelon numero 3004614), johon oli imeytetty Solvablea (kudos- ja geelisolubilisointireagenssi, joka oli hankittu yhtiöstä NEN-Dupont). Pullo suljettiin muovikorkilla, joka asetettiin tiukasti paikoilleen, ja sekä korkki että pullon ylempi puolisko käärittiin parafilmiin ja inkuboitiin varovasti ravistellen 2 tuntia 30 °C:ssa. Inkuboinnin päätyttyä suodatinpaperi siirrettiin lasista valmistettuun 7 ml:n tuikepulloon, joka sisälsi 4 ml radioaktiivisuuden määrittämiseen käytettävää "Ready Safe" (Beckman) -tuikenesteseosta. Alfa-[1- ^{14}C]-ketoisokaproaatin kantaliuos valmistettiin lisäämällä 5,6 mikrolitraa 20 mM alfa-ketoisokaproaattia (natriumsuola, Sigma, K-0629), 50 mikrolitraa alfa-[1- ^{14}C]-ketoisokaproaattia (55 mCi/mmol, 50 $\mu\text{Ci/ml}$, Amersham), ja sellainen määrä vettä, että seoksen lopulliseksi tilavuudeksi saatiin 1 ml. Haaraketjuisten alfa-ketohappojen dehydrogenaasin EI-komponentin ominaisaktiivisuus on ilmoitettu pikomooleina vapautunutta CO_2 :ta minuutissa kutakin proteiinimilligrammaa kohti, kuten seuraavasta taulukosta I käy ilmi.

Taulukko I

Haaraketjuisten alfa-ketohappojen dehydrogenaasin E1-aktiivisuus yhdistelmä-E. coli -solujen raakauutteissa

	Konstruktio	Plasmidi	Kanta	Induktio	BCKDH:n E1:n spesifinen aktiivisuus ^{1,2}
5	Ei liitos-jaksoa	pT7-7	CD677	+	0,9
	E1-a	pCD670	CD676	-	0,6
				+	0,8
10	E1-b	pCD666	CD673	-	0,5
				+	0,7
	E1-[a+b]	pCD736	CD737	-	2,0
				+	13,7
	E1-[a+b] ³	pCD705	CD705	-	0,9
15				+	0,5
	E1-[a+b]-E2-E3	pCD685	CD687	-	2,9
				+	6,0

¹ Haaraketjuisten alfa-ketohappojen dehydrogenaasin E1-komponentin spesifinen aktiivisuus on pikomooleja kehittyntä CO₂:ta minuutissa yhtä proteiinimilligrammaa kohti.

² Tulokset ovat keskiarvoja kahdesta rinnakkaisesta määrittäyksestä.

³ Tässä konstruktiossa on avoimen E1-beetan lukukehyksen c-terminaalinen osa väärässä orientaatiossa ja sitä käytettiin negatiivisena kontrollina.

30

SEKVENSSI ID:iden kuvaus

SEKVENSSI ID NO:1 on S. avermitiloksen BCKDH:n E1-alfa-alayksikköä koodaava DNA-sekvenssi. Tämä sekvenssi on myös kuvattu kuviossa 4 emäksinä 403 - 1548.

SEKVENSSI ID NO:2 on *S. avermitilix* BCKDH:n E1-beeta-alayksikköä koodaava DNA-sekvenssi. Tämä sekvenssi on myös kuvattu kuviossa 4 emäksinä 1622 - 2626.

5 SEKVENSSI ID NO:3 on DNA-sekvenssi, joka aloittaa sen avoimen lukukehyksen, joka koodaa *S. avermitilix* BCKDH:n E2-alayksikön aminoterminalista osaa. Tämä sekvenssi on myös esitetty kuviossa 4 emäksinä 2626 - 2727.

10 SEKVENSSI ID NO:4 on DNA-sekvenssi, joka edustaa pCD539:n emäksiä 3 - 251. Tämä on *S. avermitilix* BCKDH:n E2-alayksikköä koodaavan geenin osittainen sisäinen sekvenssi. Tämä sekvenssi on myös esitetty kuviossa 5.

15 SEKVENSSI ID NO:5 edustaa kuviossa 4 kuvatun *S. avermitilix* genomisen DNA-fragmentin 2728 emäsparia ja sisältää *S. avermitilix* BCKDH:n E1-alfa-, E1-beeta- ja E2 (osittainen) -alayksikköjen avoimet lukukehykset.

20 SEKVENSSI ID NO:6 on *S. avermitilix* BCKDH:n E1-alfa-alayksikön aminohapposekvenssi. Tämä aminohapposekvenssi on SEKVENSSI ID NO:1:n mukaisen DNA-sekvenssin koodaama.

20 SEKVENSSI ID NO:7 on *S. avermitilix* BCKDH:n E1-beeta-alayksikön aminohapposekvenssi. Tämä aminohapposekvenssi on SEKVENSSI ID NO:2:n mukaisen DNA-sekvenssin koodaama.

25 SEKVENSSI ID NO:8 on *S. avermitilix* BCKDH:n E2-alayksikön aminoterminalisen osan aminohapposekvenssi. Tämä aminohapposekvenssi on SEKVENSSI ID NO:3:n mukaisen DNA-sekvenssin koodaama.

30 SEKVENSSI ID NO:9 on sellainen aminohapposekvenssi, jota koodaa pCD539:n emästen 3 - 251 edustama DNA-sekvenssi (SEKVENSSI ID NO:4). Tämä aminohapposekvenssi edustaa *S. avermitilix* BCKDH:n E2-alayksikön sisäistä peptidi-fragmenttia.

(1) GENERAL INFORMATION:

- (i) APPLICANT (all designated states except the United States of America):
 - (A) NAME: Pfizer Inc.
 - (B) STREET: 235 East 42nd Street, 20th Floor
 - (C) CITY: New York
 - (D) STATE: New York
 - (E) COUNTRY: U.S.A.
 - (F) POSTAL CODE: 10017-5755
- (ii) APPLICANT (for the purposes of the United States of America only):
 - (A) CLAUDIO D. DENOYA
 - (B) STREET: 80 Spyglass Circle
 - (C) CITY: Groton
 - (D) STATE: Connecticut
 - (E) COUNTRY: U.S.A.
 - (F) POSTAL CODE: 06340
- (iii) TITLE OF INVENTION: Genes Encoding Branched-Chain Alpha-Ketoacid Dehydrogenase Complex from Streptomyces Avermitilis
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Peter C. Richardson
 - (B) STREET: Pfizer Inc., 235 East 42nd Street, 20th Floor
 - (C) CITY: New York
 - (D) STATE: New York
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 10017-5755
- (v) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/100,518
 - (B) FILING DATE: July 30, 1993
 - (C) CLASSIFICATION:
- (vi) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/100,518
 - (B) FILING DATE: JULY 30, 1993
- (vii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Sheyka, Robert F.
 - (B) REGISTRATION NUMBER: 31,304
 - (C) REFERENCE/DOCKET NUMBER: PC8346
- (viii) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (212)573-1189
 - (B) TELEFAX: (212)573-1939
 - (C) TELEX: N/A
- (ix) NUMBER OF SEQUENCES: 9
- (x) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1146 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

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

ATGACGGTCA TGGAGCAGCG GGGCGCTTAC CGGCCACAC CGCGCCCGC CTGGCAGCCC	60
CGCACCGACC CCGCGCCACT GCTGCCCGAC GCGCTGCCCC ACCGCGTCCT GGGCACCGAG	120
GCGGCCGCGG AGGCCGACCC GCTACTGCTG CGCCGCTGT ACGCGGAGCT GGTGCGCGGC	180
CGCCGCTACA ACACGCAGGC CACGGCTCTC ACCAAGCAGG GCCGGCTCGC CGTCTACCCG	240
TCGAGCACGG GCCAGGAGGC CTGCGAGGTC GCCGCCGCGC TCGTGCTGGA GGAGCGCGAC	300
TGGCTCTTCC CCAGCTACCG GGACACCTC GCCGCCGTCG CCCGCGGCCT CGATCCCGTC	360
CAGGCGCTCA CCCTCCTGCG CGGCGACTGG CACACCGGT ACGACCCCCG TGAGCACCGC	420
ATCGCGCCCC TGTGCACCCC TCTCGCGACC CAGCTCCCGC ACGCCGTCGG CCTCGCGCAC	480
GCCGCCCGCC TCAAGGGCGA CGACGTGGTC GCGCTCGCCC TGGTCGGCGA CGGCGGCACC	540
AGCGAGGGCG ACTTCCACGA GGCCTGAAC TTCGCCCGC TCTGGCAGGC GCCGGTCGTC	600
TTCCTCGTGC AGAACAACGG CTTCGCCATC TCGTCCCGC TCGCCAAGCA GACCGCCGCC	660
CCGTCGCTGG CCCACAAGGC CGTCGGCTAC GGGATGCCGG GCCGCCTGGT CGACGGCAAC	720
GACGCGGCGG CCGTGCACGA GGTCTCAGC GACGCCGTGG CCCACGCGCG CGCGGGAGGG	780
GGGCCGACGC TCGTGGAGGC GGTGACCTAC CGCATCGACG CCCACACCAA CGCCGACGAC	840
GCGACGCGCT ACCGGGGGGA CTCCGAGGTG GAGGCCTGGC GCGCGCACGA CCCGATCGCG	900
CTCCTGGAGC ACGAGTTGAC CGAACGCGGG CTGCTCGACG AGGACGGCAT CCGGGCCGCC	960
CGCGAGGACG CCGAGGCGAT GGCCGCGGAC CTGCGCGCAC GCATGAACCA GGATCCGGCC	1020
CTGGACCCCA TGGACCTGTT CGCCCATGTG TATGCCGAGC CCACCCCCCA GCTGCGGGAG	1080
CAGGAAGCCC AGTTGCGGGC CGAGCTGGCA GCGGAGGCCG ACGGGCCCCA AGGAGTCGGC	1140
CGATGA	1146

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1005 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

```

ATGACCACCG TTGCCCTCAA GCCGGCCACC ATGGCGCAGG CACTCACACG CGCGTTGCGT      60
GACGCCATGG CCGCCGACCC CGCCGTCCAC GTGATGGGCG AGGACGTCGG CACGCTCGGC      120
GGGGTCTTCC GGGTCACCGA CGGGCTCGCC AAGGAGTTCG GCGAGGACCG CTGCACGGAC      180
ACGCCGCTCG CCGAGGCAGG CATCCTCGGC ACGGCCGTCG GCATGGCGAT GTACGGGCTG      240
CGGCCGGTTCG TCGAGATGCA GTTCGACGCG TTCGCGTACC CGGCGTTCGA GCAGCTCATC      300
AGCCATGTCG CGCGGGATGC GCAACGCACC CGCGGGGCGA TGCCGCTGCC GATCACCATC      360
CGTGTCCCTT ACGGCGGCGG AATCGGCGGA GTCGAACACC ACAGCGACTC CTCCGAGGCG      420
TACTACATGG CGACTCCGGG GCTCCATGTC GTCACGCCCC CCACGGTCGC CGACGCGTAC      480
GGGCTGCTGC GCGCCGCCAT CGCCTCCGAC GACCCGGTCG TCTTCCTGGA GCCCAAGCGG      540
CTGTACTGGT CGAAGGACTC CTGGAACCCG GACGAGCCCG GGACCGTTGA ACCGATAGGC      600
CGCGCGGTGG TCGGGCGCTC GGGCCGGAGC GCCACGCTCA TCACGTACGG GCCTTCCCTG      660
CCCGTCTGCC TGGAGGCGGC CGAGGCGGCC CGGSCCGAGG GCTGGGACCT CGAAGTCGTC      720
GATCTCGGCT CCCTGGTGCC CTTCGACGAC GAGACGGTTG TGCGCGTCGG TGCGCGGACC      780
GGACGCGCCG TCGTCGTGCA CGAGTCGGGT GGTACGGCG GCCCGGGCGG GGAGATCGCC      840
GCGGGCATCA CCGAGCGCTG CTTCCACCAT CTGGAGGCGC CGGTGCTGCG CGTCGCCGGG      900
TTCGACATCC CGTATCCGCC GCCGATGCTG GAGCGCCATC ATCTGCCCGG TGTCGACCGG      960
ATCCTGGACG CGGTGGGGCG GCTTCAGTGG GAGGCGGGGA GCTGA '      1005

```

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 102 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

```

ATGGCCCAGG TGCTCGAGTT CAAGCTCCCC GACCTCGGGG AGGGCCTGAC CGAGGCCGAG      60
ATCGTCCGCT GGCTGGTGCA GGTCGGCGAC GTCGTGCGCA TC      102

```

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 249 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

ATCTCCCTCA TCGCGCTGCT CGCCAGGATC TGCACCGCCG CACTGGCCCCG CTCCCCCGAG 60
 CTCAACTCCA CCGTCGACAT GGACGCCCCG GAGGTCGTAC GGCTCGACCA GGTGCACCTG 120
 GGCTTCGCCG CGCAGACCGA ACGGGGGCTC GTCGTCCCGG TCGTGGGGGA CGCGCACGCG 180
 CGGGACGCCG AGTCGCTCAG CGCCGAGTTC GCGCGGCTGA CCGAGGCCGC CCGGACCGGC 240
 ACCCTCACA 249

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2728 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GTGACGCGG GCTCCGAAAC CGCGGATCAC GCCGTCGTCG ATGAGCCGGT TGATTGCGCG 60
 GTAGGCATTG GCACGCGACA CGTGGGCGCG CTCGGCCACG GACCGTATCG AGGCGCGGCC 120
 GTCCGCCTGG AGCATCTGGA GGATGTCCTG ATCGATGGCG TCCAGCGGGC GGGCGGGCGG 180
 CAGGGGTACC CCGGGCTCCG CTCCTCGGC CATTTGTTCA GGTGCCATGT CCTCCGGCCT 240
 CCTTACCATG GACGTAGTGC GTTCATTCCA GGCTGTGGAG AACCGTTTGT CCACAGCCTG 300
 ACGGTGCCTG TAGCCAAAAT GTCCGACGA CCGAACAATC GGTAGGTGAG GCGCCTCACA 360
 CCCGTGGCGC GCCCAAAGCC GCTCCCACGA GGAGGTGCCG TCATGACGGT CATGGAGCAG 420
 CGGGGCGCTT ACCGGCCCAC ACCGCGCCCC GCCTGGCAGC CCCGCACCGA CCCCAGGCCA 480
 CTGCTGCCCC ACGCGCTGCC CCACCGCGTC CTGGGCACCG AGGCGGCCGC GGAGGCCGAC 540
 CCGCTACTGC TGCGCCGCCT GTACGCGGAG CTGGTGCGCG GCCGCCGCTA CAACACGCAG 600
 GCCACGGCTC TCACCAAGCA GGGCCGGCTC GCCGTCTACC CGTCGAGCAC GGGCCAGGAG 660
 GCCTGCGAGG TCGCCGCCGC GCTCGTGCTG GAGGAGCGCG ACTGGCTCTT CCCCAGCTAC 720
 CGGGACACCC TCGCCCGCGT CGCCCGCGGC CTCGATCCCC TCCAGGCGCT CACCCTCCTG 780
 CGCGGCGACT GGCACACCGG GTACGACCCC CGTGAGCACC GCATCGCGCC CCTGTGCACC 840
 CCTCTCGCGA CCCAGCTCCC GCACGCCGTC GGCCTCGCGC ACGCCGCCCG CCTCAAGGGC 900
 GACGACGTGG TCGCGCTCGC CCTGGTCGGC GACGGCGGCA CCAGCGAGGG CCACTTCCAC 960
 GAGGCACTGA ACTTCGCCGC CGTCTGGCAG GCGCCGGTCG TCTTCCTCGT GCAGAACAAC 1020
 GGCTTCGCCA TCTCCGTCCC GCTCGCCAAG CAGACCGCCG CCCCCTCGCT GGCCCAACAG 1080
 GCCGTGGGCT ACGGGATGCC GGGCCGCCTG GTCGACGGCA ACGACGGGC GGGCGTGCAC 1140
 GAGGTCCTCA GCGACGCCGT GGCCACGCG CGCGCGGAG GGGGGCCGAC GCTCGTGGAG 1200
 GCGGTGACCT ACCGCATCGA CGCCACACC AACGCCGACG ACGCGACGCG CTACCGGGGG 1260

GACTCCGAGG TGGAGGCCTG GCGCGCGCAC GACCCGATCG CGCTCCTGGA GCACGAGTTG 1320
 ACCGAACGCG GGCTGCTCGA CGAGGACGGC ATCCGGGCGG CCGCGAGGA CGCCGAGGCG 1380
 ATGGCCGCGG ACCTGCGCGC ACGCATGAAC CAGGATCCGG CCCTGGACCC CATGGACCTG 1440
 TTCGCCCATG TGTATGCCGA GCCCAGCCCC CAGCTGCGGG AGCAGGAAGC CCAGTTGCGG 1500
 GCGGAGCTGG CAGCGGAGGC CGACGGGCCC CAAGGAGTCG GCCGATGAAG AGAGTTGACC 1560
 ATCGGGCCCC GAGAAGCGGG CCGATGACCT CCGTTGGCCT TTGGCCGGAA GGAGCCGGGC 1620
 GATGACCACC GTTGCCCTCA AGCCGGCCAC CATGGCGCAG GCACTCACAC GCGCGTTGCG 1680
 TGACGCCATG GCCGCCGACC CCGCCGTCCA CGTGATGGGC GAGGACGTCG GCACGCTCGG 1740
 CGGGGTCTTC CGGGTCACCG ACGGGCTCGC CAAGGAGTTC GGCGAGGACC GCTGCACGGA 1800
 CACGCCGCTC GCCGAGGCAG GCATCCTCGG CACGGCCGTC GGCATGGCGA TGTACGGGCT 1860
 GCGGCCGGTC GTCGAGATGC AGTTCGACGC GTTCGCGTAC CCGGCGTTTC AGCAGCTCAT 1920
 CAGCCATGTC GCGCGGGATG CGCAACGCAC CCGCGGGGCG ATGCCGCTGC CGATCACCAT 1980
 CCGTGTCCCC TACGGCGGCG GAATCGGCGG AGTCGAACAC CACAGCGACT CCTCCGAGGC 2040
 GTACTACATG GCGACTCCGG GGCTCCATGT CGTCACGCCC GCCACGGTCG CCGACGCGTA 2100
 CGGGCTGCTG CGCGCCGCCA TCGCCTCCGA CGACCCGGTC GTCTTCCTGG AGCCCAAGCG 2160
 GCTGTACTGG TCGAAGGACT CCTGGAACCC GGACGAGCCG GGGACCGTTG AACCAGATAGG 2220
 CCGCGCGGTC GTGCGGCGCT CGGGCCGGAG CGCCACGCTC ATCACGTACG GGCCTTCCCT 2280
 GCCCCGTCTG CTGGAGGCGG CCGAGGCGGC CCGGGCCGAG GGCTGGGACC TCGAAGTCGT 2340
 CGATCTGCGC TCCCTGGTGC CCTTCGACGA CGAGACGGTT GTGCGCGTCG GTGCGCGGAC 2400
 CGGACGCGCC GTCGTCGTGC ACGAGTCGGG TGGTTACGGC GGCCCGGGCG GGGAGATCGC 2460
 CGCGGGCATC ACCGAGCGCT GCTTCCACCA TCTGGAGGCG CCGGTGCTGC GCGTCGCCGG 2520
 GTTCGACATC CCGTATCCGC CGCCGATGCT GGAGCGCCAT CATCTGCCCC GTGTCGACCG 2580
 GATCCTGGAC GCGGTGGGGC GGCTTCAGTG GGAGCGGGG AGCTGATGGC CCAGGTGCTC 2640
 GAGTTCAAGC TCCCCGACCT CGGGGAGGGC CTGACCGAGG CCGAGATCGT CCGCTGGCTG 2700
 GTGCAGGTCG GCGACGTCGT GGCGATCG 2728

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 381 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

Met Thr Val Met Glu Gln Arg Gly Ala Tyr Arg Pro Thr Pro Pro Pro
 1 5 10 15

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

355	360	365
Leu Ala Ala Glu Ala Asp Gly	Pro Gln Gly Val Gly	Arg
370	375	380

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 334 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

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

Gly Ala Arg Thr Gly Arg Ala Val Val Val His Glu Ser Gly Gly Tyr
 260 265 270

Gly Gly Pro Gly Gly Glu Ile Ala Ala Gly Ile Thr Glu Arg Cys Phe
 275 280 285

His His Leu Glu Ala Pro Val Leu Arg Val Ala Gly Phe Asp Ile Pro
 290 295 300

Tyr Pro Pro Pro Met Leu Glu Arg His His Leu Pro Gly Val Asp Arg
 305 310 315 320

Ile Leu Asp Ala Val Gly Arg Leu Gln Trp Glu Ala Gly Ser
 325 330

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 34 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

Met Ala Gln Val Leu Glu Phe Lys Leu Pro Asp Leu Gly Glu Gly Leu
 1 5 10 15

Thr Glu Ala Glu Ile Val Arg Trp Leu Val Gln Val Gly Asp Val Val
 20 25 30

Ala Ile

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 83 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

Ile Ser Leu Ile Ala Leu Leu Ala Arg Ile Cys Thr Ala Ala Leu Ala
 1 5 10 15

Arg Phe Pro Glu Leu Asn Ser Thr Val Asp Met Asp Ala Arg Glu Val
 20 25 30

Val Arg Leu Asp Gln Val His Leu Gly Phe Ala Ala Gln Thr Glu Arg
 35 40 45

Gly Leu Val Val Pro Val Val Arg Asp Ala His Ala Arg Asp Ala Glu
 50 55 60

Ser Leu Ser Ala Glu Phe Ala Arg Leu Thr Glu Ala Ala Arg Thr Gly
 65 70 75 80

Thr Leu Thr

Patenttivaatimukset

1. Eristetty DNA-segmentti, joka koodaa Streptomyces-sukuun kuuluvan organismin haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksia, t u n n e t t u siitä, että mainittu DNA-segmentti sisältää SEKVENSSSI ID NO:1:n, SEKVENSSSI ID NO:2:n, SEKVENSSSI ID NO:3:n, SEKVENSSSI ID NO:4:n tai SEKVENSSSI ID NO:5:n mukaisen DNA-sekvenssin tai näiden sekvenssien alleelisen variantin.
2. Patenttivaatimuksen 1 mukainen eristetty DNA-segmentti, t u n n e t t u siitä, että se koodaa S. avermitiloksen haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksia.
3. Patenttivaatimuksen 1 mukainen eristetty DNA-segmentti, t u n n e t t u siitä, että se lisäksi sisältää DNA-alueen, joka säätelee mainitun haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksin ilmentymistä.
4. Patenttivaatimuksen 2 mukainen eristetty DNA-segmentti, t u n n e t t u siitä, että se lisäksi sisältää DNA-alueen, joka säätelee mainitun haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksin ilmentymistä.
5. DNA-segmentti, t u n n e t t u siitä, että se on patenttivaatimuksen 1 mukaisen DNA-segmentin osajoukko ja että se on tätä toiminnallisesti vastaava segmentti.
6. Yhdistelmä-DNA, t u n n e t t u siitä, että se sisältää patenttivaatimuksen 1 mukaisen DNA-segmentin.
7. Isäntäsolu, t u n n e t t u siitä, että siihen on liitetty patenttivaatimuksen 6 mukainen yhdistelmä-DNA.
8. Plasmidi, t u n n e t t u siitä, että se sisältää patenttivaatimuksen 6 mukaista yhdistelmä-DNA:ta.
9. DNA-segmentti, t u n n e t t u siitä, että se on pCD528, pCD545, pCD550, pCD559, pCD574 tai pCD577, ja mainittujen DNA-segmenttien genomiset kartat ovat kuviossa 3 esitettyjen genomisten karttojen mukaisia.

10. Isäntäsolu, t u n n e t t u siitä, että siihen on liitetty patenttivaatimuksen 9 mukainen DNA-segmentti.

5 11. Patenttivaatimuksen 1 mukainen DNA-segmentti, t u n n e t t u siitä, että se sisältää DNA-alueen, joka säätelee SEKVENSSI ID NO:1:n, SEKVENSSI ID NO:2:n, SEKVENSSI ID NO:3:n, SEKVENSSI ID NO:4:n tai SEKVENSSI ID NO:5:n mukaisen DNA-sekvenssin tai näiden sekvenssien alleelisen variantin transkriptiota tai translaatiota.

10 12. Haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksia vastaavat geenit, t u n n e t t u siitä, että ne sisältyvät patenttivaatimuksen 9 mukaiseen DNA-segmenttiin.

15 13. Menetelmä luonnollisen avermektiinin valmistamiseksi, t u n n e t t u siitä, että menetelmä käsittää sen, että fermentoidaan sellaisissa olosuhteissa ja sellaisessa fermentaatiomediumissa, jotka soveltuvat tällaisen luonnollisen avermektiinin valmistamiseen, S. avermitilistä, jossa haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksia koodaavien geenien kappalemäärää on lisätty.

20 14. Menetelmä luonnollisen avermektiinin valmistamiseksi, t u n n e t t u siitä, että menetelmä käsittää sen, että fermentoidaan sellaisissa olosuhteissa ja sellaisessa fermentaatiomediumissa, jotka soveltuvat tällaisen luonnollisen avermektiinin valmistamiseen, S. avermitilistä, jossa haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksia koodaavien geenien ilmentymistä on tehostettu tällaisen ilmentymisen säätelystä vastaavien geenien muokkaus- tai korvaustoimenpiteillä.

30 15. Menetelmä uuden avermektiinin valmistamiseksi, t u n n e t t u siitä, että menetelmä käsittää sen, että fermentoidaan sellaisissa olosuhteissa ja sellaisessa fermentaatiomediumissa, jotka soveltuvat tällaisen uuden
35 avermektiinin valmistamiseen, S. avermitilistä, jossa haa-

Patl
375

raketjuisten alfa-ketohappojen dehydrogenaasikompleksin ilmentymistä on rajoitettu tai joka ilmentyminen on tehty mahdottomaksi tällaisesta ilmentymisestä vastaavien geenien deletoinnilla, inaktivoinnilla, korvaamisella tai
5 muilla tavoin suoritettulla muokkaustoimenpiteellä.

16. Patenttivaatimuksen 15 mukainen menetelmä, t u n n e t t u siitä, että deletoidut, inaktivoituneet, korvatut tai muulla tavoin muokatut geenit ovat käynnistäjä (enhancer), inhibiittori-, aktivaattori- tai represso-
10 rigeenejä.

17. Patenttivaatimuksen 15 mukainen menetelmä uuden avermektiinin valmistamiseksi, t u n n e t t u siitä, että menetelmä käsittää sen, että fermentoidaan sellaisissa olosuhteissa ja sellaisessa fermentaatiomediumissa,
15 jotka soveltuvat tällaisen uuden avermektiinin valmistamiseen, S. avermitilistä, jossa haaraketjuisten alfa-ketohappojen dehydrogenaasikompleksin ilmentymisestä vastaavat geenit on deletoitu tai inaktivoitu.

18. DNA-segmentti, t u n n e t t u siitä, että se
20 sisältää DNA-sekvenssin, joka on SEKVENSSI ID NO:1:n, SEKVENSSI ID NO:2:n, SEKVENSSI ID NO:3:n, SEKVENSSI ID NO:4:n tai SEKVENSSI ID NO:5:n mukaisen DNA-sekvenssin alajoukko, tai tällaisen sekvenssin alleelinen variantti, ja joka kykenee hybridisoitumaan SEKVENSSI ID NO:1:n, SEKVENSSI ID
25 NO:2:n, SEKVENSSI ID NO:3:n, SEKVENSSI ID NO:4:n tai SEKVENSSI ID NO:5:n mukaiseen DNA-sekvenssiin tai tämän alleeliseen varianttiin silloin, kun tätä käytetään koettimena, tai monistamaan tällaisen sekvenssin kokonaan tai osittain, silloin kun tätä käytetään polymeerasiketjureak-
30 tion alukkeena.

19. Olennaisesti puhtaaksi puhdistettu polypeptidi, t u n n e t t u siitä, että se sisältää SEKVENSSI ID NO:6:n, SEKVENSSI ID NO:7:n SEKVENSSI ID NO:8:n tai SEKVENSSI ID NO:9:n mukaisen aminohapposekvenssin.

FIG. 1A

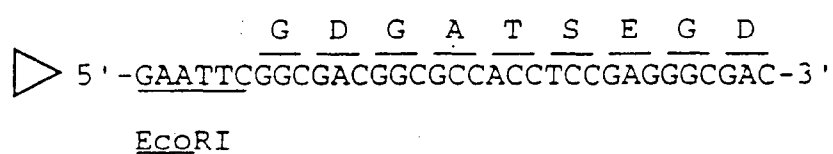


FIG. 1B

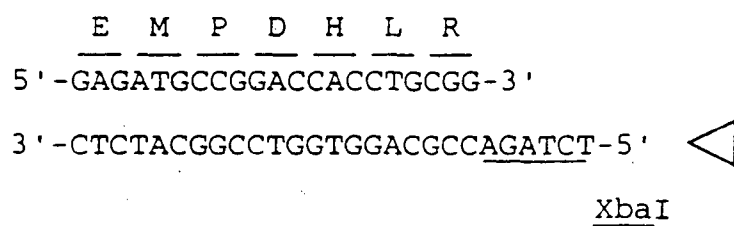
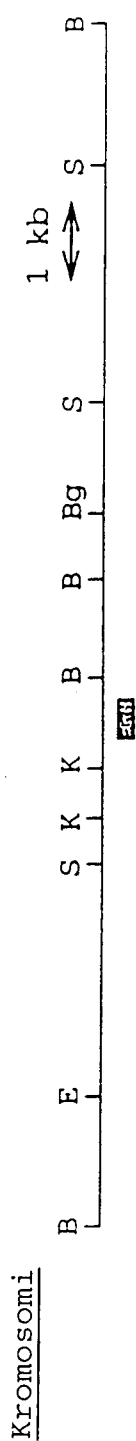


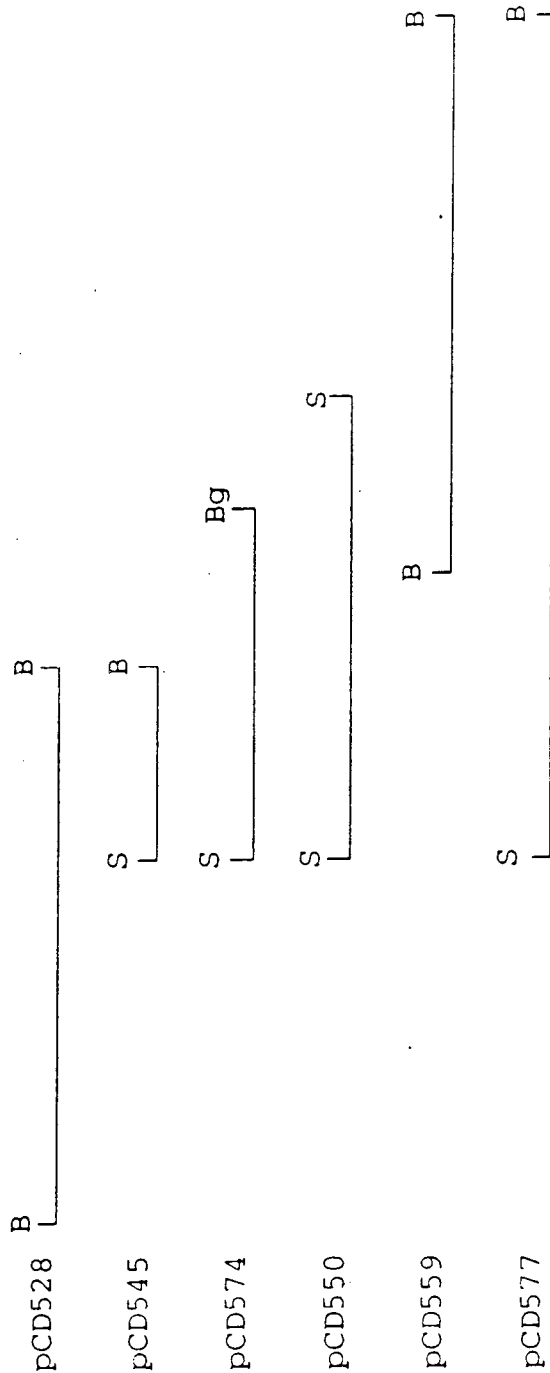
FIG. 2

	↑		↓
Sa	DGATSEGD FHEALNFAA VWQAPVVFL VQNNGFAISV .PLAKQTAAP SLAQKAVGY EMPDHLR		
Bs	DGGTSQGD FYEGINFAG AFKAPAFV VQNNRFAIST .PVEKQTVAK TLAQKAVAAG IPG.IQ		
Pp	DGATAESD FHTALTFA HVYRAPVIL NVNNQWAIST FFQAIAGGES TTFAGRGVGC GGIAS.LR		
Hs	EGAASEGD AHAGFNFAA TLECPIIFFC RNNNGYAIST .PTSEQYRGD GIAARGPGY GIMS.IR		

FIG. 3



Jatkokloonit



Ela E1b E2
 24157231 24157231 24157231

FIG. 4A

10	20	30	40	50	60	70											
GTCGACGCGG GCTCCGAAAC CGCGGATCAC GCCGTCGTCG ATGAGCCGGT TGATTCGCGC GTAGGCATTG																	
80	90	100	110	120	130	140											
GCACGCGACA CGTGGGCGCG CTCGGCCACG GACCGTATCG AGGCGCGGCC GTCCGCCTGG AGCATCTGGA																	
150	160	170	180	190	200	210											
GGATGTCCTG ATCGATGGCG TCCAGCGGGC GGGCGGGCGG CAGGGGTACC CCGGGCTCCG CTCCCTCGGC																	
220	230	240	250	260	270	280											
CATTTGTTCA GGTGCCATGT CCTCCGGCCT CTTACCATG GACGTAGTGC GTTCATTCCA GGCTGTGGAG																	
290	300	310	320	330	340	350											
AACCGTTTGT CCACAGCCTG ACGGTGCCTG TAGCCAAAAT GTGCCGACGA CCGAACAATC GGTAGGTGAG																	
360	370	380	390	400	411												
GCGCCTCACA CCCGTGGCGC GCCCAAAGCC GCTCCCACGA <u>GGAGGTGCCG</u> TC						$\begin{matrix} > \\ \text{ATG} & \text{ACG} & \text{GTC} \\ \text{M} & \text{T} & \text{V} \end{matrix}$											
420	429	438	447	456	465												
ATG	GAG	CAG	CGG	GGC	GCT	TAC	CGG	CCC	ACA	CCG	CCG	CCC	GCC	TGG	CAG	CCC	CGC
M	E	Q	R	G	A	Y	R	P	T	P	P	P	A	W	Q	P	R
474	483	492	501	510	519												
ACC	GAC	CCC	GCG	CCA	CTG	CTG	CCC	GAC	CCG	CTG	CCC	CAC	CGC	GTC	CTG	GGC	ACC
T	D	P	A	P	L	L	P	D	A	L	P	H	R	V	L	G	T
528	537	546	555	564	573												
GAG	GCG	GCC	GCG	GAG	GCC	GAC	CCG	CTA	CTG	CTG	CGC	CGC	CTG	TAC	GCG	GAG	CTG
E	A	A	A	E	A	D	P	L	L	L	R	R	L	Y	A	E	L
582	591	600	609	618	627												
GTG	CGC	GGC	CGC	CGC	TAC	AAC	ACG	CAG	GCC	ACG	GCT	CTC	ACC	AAG	CAG	GGC	CGG
V	R	G	R	R	Y	N	T	Q	A	T	A	L	T	K	Q	G	R

FIG. 4B

636				645				654				663				672				681			
CTC	GCC	GTC	TAC	CCG	TCG	AGC	ACG	GGC	CAG	GAG	GCC	TGC	GAG	GTC	GCC	GCC	GCG						
L	A	V	Y	P	S	S	T	G	Q	E	A	C	E	V	A	A	A						
690				699				708				717				726				735			
CTC	GTG	CTG	GAG	GAG	CGC	GAC	TGG	CTC	TTC	CCC	AGC	TAC	CGG	GAC	ACC	CTC	GCC						
L	V	L	E	E	R	D	W	L	F	P	S	Y	R	D	T	L	A						
744				753				762				771				780				789			
GCC	GTC	GCC	CGC	GGC	CTC	GAT	CCC	GTC	CAG	GCG	CTC	ACC	CTC	CTG	CGC	GGC	GAC						
A	V	A	R	G	L	D	P	V	Q	A	L	T	L	L	R	G	D						
798				807				816				825				834				843			
TGG	CAC	ACC	GGG	TAC	GAC	CCC	CGT	GAG	CAC	CGC	ATC	GCG	CCC	CTG	TGC	ACC	CCT						
W	H	T	G	Y	D	P	R	E	H	R	I	A	P	L	C	T	P						
852				861				870				879				888				897			
CTC	GCG	ACC	CAG	CTC	CCG	CAC	GCC	GTC	GGC	CTC	GCG	CAC	GCC	GCC	CGC	CTC	AAG						
L	A	T	Q	L	P	H	A	V	G	L	A	H	A	A	R	L	K						
906				915				924				933				942				951			
GGC	GAC	GAC	GTG	GTC	GCG	CTC	GCC	CTG	GTC	GGC	GAC	GGC	GGC	ACC	AGC	GAG	GGC						
G	D	D	V	V	A	L	A	L	V	G	D	G	G	T	S	E	G						
960				969				978				987				996				1005			
GAC	TTC	CAC	GAG	GCA	CTG	AAC	TTC	GCC	GCC	GTC	TGG	CAG	GCG	CCG	GTC	GTC	TTC						
D	F	H	E	A	L	N	F	A	A	V	W	Q	A	P	V	V	F						
1014				1023				1032				1041				1050				1059			
CTC	GTG	CAG	AAC	AAC	GGC	TTC	GCC	ATC	TCC	GTC	CCG	CTC	GCC	AAG	CAG	ACC	GCC						
L	V	Q	N	N	G	F	A	I	S	V	P	L	A	K	Q	T	A						

FIG. 4C

1068	1077	1086	1095	1104	1113
GCC CCG TCG CTG	GCC CAC AAG	GCC GTC GGC	TAC GGG ATG	CCG GGC CGC	CTG GTC
A P S L	A H K	A V G	Y G M	P G R	L V
1122	1131	1140	1149	1158	1167
GAC GGC AAC	GAC GCG GCG	GCC GTG CAC	GAG GTC CTC	AGC GAC GCC	GTG GCC CAC
D G N D	A A A	V H E	V L S	D A V	A H
1176	1185	1194	1203	1212	1221
GCG CGC GCG	GGA GGG GGG	CCG ACG CTC	GTG GAG GCG	GTG ACC TAC	CGC ATC GAC
A R A G	G G P T	L V E	A V T	Y R I	D
1230	1239	1248	1257	1266	1275
GCC CAC ACC	AAC GCC GAC	GAC GCG ACG	CGC TAC CGG	GGG GAC TCC	GAG GTG GAG
A H T N	A D D	A T R	Y R G	D S E	V E
1284	1293	1302	1311	1320	1329
GCC TGG CGC	GCG CAC GAC	CCG ATC GCG	CTC CTG GAG	CAC GAG TTG	ACC GAA CGC
A W R A	H D P I	A L L	E H E	L T E	R
1338	1347	1356	1365	1374	1383
GGG CTG CTC	GAC GAG GAC	GGC ATC CGG	GCC GCC CGC	GAG GAC GCC	GAG GCG ATG
G L L D	E D G I	R A A	R E D	A E A	M
1392	1401	1410	1419	1428	1437
GCC GCG GAC	CTG CGC GCA	CGC ATG AAC	CAG GAT CCG	GCC CTG GAC	CCC ATG GAC
A A D L	R A R	M N Q	D P A	L D P	M D
1446	1455	1464	1473	1482	1491
CTG TTC GCC	CAT GTG TAT	GCC GAG CCC	ACC ACC CCC	CAG CTG CGG	GAG CAG GAA
L F A H	V Y A	E P T	P Q L	R E Q	E A

FIG. 4D

1500	1509	1518	1527	1536	1545
<u>CAG</u> <u>TTG</u> <u>CGG</u> <u>GCC</u> <u>GAG</u> <u>CTG</u> <u>GCA</u> <u>GCG</u> <u>GAG</u> <u>GCC</u> <u>GAC</u> <u>GGG</u> <u>CCC</u> <u>CAA</u> <u>GGA</u> <u>GTC</u> <u>GGC</u> <u>CGA</u>					
Q L R A E L A A E A D G P Q G V G R					
1558	1568	1578	1588	1598	1608
<u>TGA</u> <u>AGAGAGTTGA</u> <u>CCATCGGGCC</u> <u>CCGAGAAGCG</u> <u>GGCCGATGAC</u> <u>CTCCGTTGGC</u> <u>CTTTGGCCGG</u>					
1618	1627	1636	1645	1654	1663
<u>AAGGAGCCCGG</u> <u>GCG</u> <u>ATG</u> <u>ACC</u> <u>ACC</u> <u>GTT</u> <u>GCC</u> <u>CTC</u> <u>AAG</u> <u>CCG</u> <u>GCC</u> <u>ACC</u> <u>ATG</u> <u>GCG</u> <u>CAG</u> <u>GCA</u>					
M T T V A L K P A T M A Q A					
1672	1681	1690	1699	1708	1717
<u>CTC</u> <u>ACA</u> <u>CGC</u> <u>GCG</u> <u>TTG</u> <u>CGT</u> <u>GAC</u> <u>GCC</u> <u>ATG</u> <u>GCC</u> <u>GCC</u> <u>GAC</u> <u>CCC</u> <u>GCC</u> <u>GTC</u> <u>CAC</u> <u>GTG</u> <u>ATG</u>					
L T R A L R D A M A A D P A V H V M					
1726	1735	1744	1753	1762	1771
<u>GGC</u> <u>GAG</u> <u>GAC</u> <u>GTC</u> <u>GGC</u> <u>ACG</u> <u>CTC</u> <u>GGC</u> <u>GGG</u> <u>GTC</u> <u>TTC</u> <u>CGG</u> <u>GTC</u> <u>ACC</u> <u>GAC</u> <u>GGG</u> <u>CTC</u> <u>GCC</u>					
G E D V G T L G G V F R V T D G L A					
1780	1789	1798	1807	1816	1825
<u>AAG</u> <u>GAG</u> <u>TTC</u> <u>GGC</u> <u>GAG</u> <u>GAC</u> <u>CGC</u> <u>TGC</u> <u>ACG</u> <u>GAC</u> <u>ACG</u> <u>CCG</u> <u>CTC</u> <u>GCC</u> <u>GAG</u> <u>GCA</u> <u>GGC</u> <u>ATC</u>					
K E F G E D R C T D T P L A E A G I					
1834	1843	1852	1861	1870	1879
<u>CTC</u> <u>GGC</u> <u>ACG</u> <u>GCC</u> <u>GTC</u> <u>GGC</u> <u>ATG</u> <u>GCG</u> <u>ATG</u> <u>TAC</u> <u>GGG</u> <u>CTG</u> <u>CGG</u> <u>CCG</u> <u>GTC</u> <u>GTC</u> <u>GAG</u> <u>ATG</u>					
L G T S V G M A M Y G L R P V V E M					
1888	1897	1906	1915	1924	1933
<u>CAG</u> <u>TTC</u> <u>GAC</u> <u>GCG</u> <u>TTC</u> <u>GCG</u> <u>TAC</u> <u>CCG</u> <u>GCG</u> <u>TTC</u> <u>GAG</u> <u>CAG</u> <u>CTC</u> <u>ATC</u> <u>AGC</u> <u>CAT</u> <u>GTC</u> <u>GCG</u>					
Q F D A F A Y P A F E Q L I S H V A					

FIG. 4E

1942	1951	1960	1969	1978	1987
CGG GAT GCG CAA CGC ACC CGC GGG GCG ATG CCG CTG CCG ATC ACC ATC CGT GTC					
R D A Q R T R G A M P L P I T I R V					
1996	2005	2014	2023	2032	2041
CCC TAC GGC GGC GGA ATC GGC GGA GTC GAA CAC CAC AGC GAC TCC TCC GAG GCG					
P Y G G G I G G V E H H S D S S E A					
2050	2059	2068	2077	2086	2095
TAC TAC ATG GCG ACT CCG GGG CTC CAT GTC GTC ACG CCC GCC ACG GTC GCC GAC					
Y Y M A T P G L H V V T P A T V A D					
2104	2113	2122	2131	2140	2149
GCG TAC GGG CTG CTG CGC GCC GCC ATC GCC TCC GAC GAC CCG GTC GTC TTC CTG					
A Y G L L R A A I A S D D P V V F L					
2158	2167	2176	2185	2194	2203
GAG CCC AAG CCG CTG TAC TGG TCG AAG GAC TCC TGG AAC CCG GAC GAG CCG GGG					
E P K R L Y W S K D S W N P D E P G					
2212	2221	2230	2239	2248	2257
ACC GTT GAA CCG ATA GGC CGC GCG GTG GTG CCG CCG TCG GGC CCG AGC GCC ACG					
T V E P I G R A V V R R S G R S A T					
2266	2275	2284	2293	2302	2311
CTC ATC ACG TAC GGG CCT TCC CTG CCC GTC TGC CTG GAG GCG GCC GAG GCG GCC					
L I T Y G P S L P V C L E A A E A A					
2320	2329	2338	2347	2356	2365
CGG GCC GAG GGC TGG GAC CTC GAA GTC GTC GAT CTG GCG TCC CTG GTG CCC TTC					
R A E G W D L E V V D L R S L V P F					

FIG. 4F

2374	2383	2392	2401	2410	2419												
<u>GAC</u> D	<u>GAC</u> D	<u>GAG</u> E	<u>ACG</u> T	<u>GTT</u> V	<u>GTG</u> V	<u>CGC</u> R	<u>GTC</u> V	<u>GGT</u> G	<u>GCG</u> A	<u>CGG</u> R	<u>ACC</u> T	<u>GGA</u> G	<u>CGC</u> R	<u>GCC</u> A	<u>GTC</u> V	<u>GTC</u> V	<u>GTG</u> V
2428	2437	2446	2455	2464	2473												
<u>CAC</u> H	<u>GAG</u> E	<u>TCG</u> S	<u>GGT</u> G	<u>GGT</u> G	<u>TAC</u> Y	<u>GGC</u> G	<u>GGC</u> G	<u>CCG</u> P	<u>GGC</u> G	<u>GGG</u> G	<u>GAG</u> E	<u>ATC</u> I	<u>GCC</u> A	<u>GCG</u> A	<u>GGC</u> G	<u>ATC</u> I	<u>ACC</u> T
2482	2491	2500	2509	2518	2527												
<u>GAG</u> E	<u>CGC</u> R	<u>TGC</u> C	<u>TTC</u> F	<u>CAC</u> H	<u>CAT</u> H	<u>CTG</u> L	<u>GAG</u> E	<u>GCG</u> A	<u>CCG</u> P	<u>GTG</u> V	<u>CTG</u> L	<u>CGC</u> R	<u>GTC</u> V	<u>GCC</u> A	<u>GGG</u> G	<u>TTC</u> F	<u>GAC</u> D
2536	2545	2554	2563	2572	2581												
<u>ATC</u> I	<u>CCG</u> P	<u>TAT</u> Y	<u>CCG</u> P	<u>CCG</u> P	<u>CCG</u> P	<u>ATG</u> M	<u>CTG</u> L	<u>GAG</u> E	<u>CGC</u> R	<u>CAT</u> H	<u>CAT</u> H	<u>CTG</u> L	<u>CCC</u> P	<u>GGT</u> G	<u>GTC</u> V	<u>GAC</u> D	<u>GGG</u> R
2590	2599	2608	2617	2628													
<u>ATC</u> I	<u>CTG</u> L	<u>GAC</u> D	<u>GCG</u> A	<u>GTG</u> V	<u>GGG</u> G	<u>CGG</u> R	<u>CTT</u> L	<u>CAG</u> Q	<u>TGG</u> W	<u>GAG</u> E	<u>GCG</u> A	<u>GGG</u> G	<u>AGC</u> S	<u>TG</u> *	<u>ATG</u> M	<u>GCC</u> A	<u>CAG</u> Q
2637	2646	2655	2664	2673	2682												
<u>GTG</u> V	<u>CTC</u> L	<u>GAG</u> E	<u>TTC</u> F	<u>AAG</u> K	<u>CTC</u> L	<u>CCC</u> P	<u>GAC</u> D	<u>CTC</u> L	<u>GGG</u> G	<u>GAG</u> E	<u>GGC</u> G	<u>CTC</u> L	<u>ACC</u> T	<u>GAG</u> E	<u>GCC</u> A	<u>GAG</u> E	<u>ATC</u> I
2691	2700	2709	2718	2727													
<u>GTC</u> V	<u>CGC</u> R	<u>TGG</u> W	<u>CTG</u> L	<u>GTG</u> V	<u>CAG</u> Q	<u>GTC</u> V	<u>GGC</u> G	<u>GAC</u> D	<u>GTC</u> V	<u>GTG</u> V	<u>GCG</u> A	<u>ATC</u> I	^{>} G				

FIG. 5

		11		20		29		38		47								
AG		ATC	TCC	CTC	ATC	GCG	CTG	CTC	GCC	AGG	ATC	TGC	ACC	GCC	GCA	CTG	GCC	CGC
I	S	L	I	A	L	L	A	R	I	C	T	A	A	L	A	R		
56		65		74		83		92		101								
TTC	CCC	GAG	CTC	AAC	TCC	ACC	GTC	GAC	ATG	GAC	GCC	CGC	GAG	GTC	GTA	CGG	CTC	
F	P	E	L	N	S	T	V	D	M	D	A	R	E	V	V	R	L	
110		119		128		137		146		155								
GAC	CAG	GTG	CAC	CTG	GGC	TTC	GCC	GCG	CAG	ACC	GAA	CGG	GGG	CTC	GTC	GTC	CCG	
D	Q	V	H	L	G	F	A	A	Q	T	E	R	G	L	V	V	P	
164		173		182		191		200		209								
GTC	GTG	CGG	GAC	GCG	CAC	GCG	CGG	GAC	GCC	GAG	TCG	CTC	AGC	GCC	GAG	TTC	GCG	
V	V	R	D	A	H	A	R	D	A	E	S	L	S	A	E	F	A	
218		227		236		245												
CGG	CTG	ACC	GAG	GCC	GCC	CGG	ACC	GGC	ACC	CTC	ACA							
R	L	T	E	A	A	R	T	G	T	L	T							

FIG. 6A

#55-PCR ▷ 5'-AAGAGATCTCATATGACGGTCATGGAGCAGCGG-3'
 BglII NdeI

#56-PCR ▷ 5'-AAGAGATCTCATATGACCACCGTTGCCCTGAAG-3'
 BglII NdeI

FIG. 6B

#30-BP-PCR 3'-AGCAAAATGTTGCAGCACTGACCGACGTCCTAGGA-5' ◁
 PstI BamHI

#31-BP-PCR 3'-ACCGCATTAGTACCAGTATCGACAGACGTCCTAGGA-5' ◁
 PstI BamHI