(19)

SUOMI - FINLAND
(FI)

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

(10) **FI 945928 A7**

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

(21) Patenttihakemus - Patentansökan - Patent application 945928

(51) Kansainvälinen patenttiluokitus - Internationell patentklassifikation - International patent classification
C12N 9/24
A23K 1/165
A21D 8/04
D21C 9/10
A23K 1/16 (2006.01)

(22) Tekemispäivä - Ingivningsdag - Filing date **17.06.1993**

(23) Saapumispäivä - Ankomstdag - Reception date **16.12.1994**

(41) Tullut julkiseksi - Blivit offentlig - Available to the public **16.02.1995**

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

(86) Kansainvälinen hakemus - Internationell ansökan - International application **17.06.1993 PCT/GB1993/001283**

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

17.06.1992 AU 2985

(71) Hakija - Sökande - Applicant

1 • **Biotechnology and Biological Sciences,** Central Office, Polaris House North Star Avenue, Swindon SN2 1UH, ISO-BRITANNIA, (GB)
2 • **University of Newcastle-Upon-Tyne,** 6 Kensington Terrace Newcastle-upon-Tyne NE1 7RU, United Kingdom, ISO-BRITANNIA, (GB)

(72) Keksijä - Uppfinnare - Inventor

1 • **Hazlewood, Geoffrey Peter,** United Kingdom, ISO-BRITANNIA, (GB)
2 • **Gilbert, Harry John,** United Kingdom, ISO-BRITANNIA, (GB)

(74) Asiamies - Ombud - Agent

Berggren Oy Ab, Antinkatu 3 C, 00100 Helsinki

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

Rekombinantti-ksylanaaseja – Rekombinant-xylanaser

Tämä keksintö koskee anaerobisesta sienestä johdettavia rekombinantti-ksylanaaseja.

Ksylaani, eräs kasvihemiselluloosien pääkomponenteista, muodostuu 1,4-sidottujen ß-D-ksylopyranoosiyksikköjen, jotka on substituoitu pääasiassa asetyyli-, arabinosyyli- ja glukuronosyylitähteillä, polymeeristä. Lehtipuuksylaani on tyypillisesti O-asetyyli-4-O-metyyliglukuroniksylaania, jossa on suunnilleen 10 prosenttia ksyloosiyksikköjä, jotka ovat α,1,2-sidotut 4-O-metyyliglukuronihapposivuketjuun, ja 70 prosenttia C-2- tai C-3-asemissa asetyloituja ksyloositähteitä. Havupuuksylaanit ovat tavallisesti arabino-4-O-metyyliglukuroniksylaaneja, joissa enemmän kuin 10 prosenttia ksyloosiyksiköistä on substituoitu α-1,3-sidotuilla arabinofuranoosiradikaaleilla. Joukko mikrobientsyymejä toimii yhdessä ksylaanin muuttamiseksi sen rakenneosiksi, yksinkertaisiksi sokereiksi. Niitä ovat endo-ß-1,4-ksylanaasit (EC 3.2.1.8), ß-ksylosi-daasi (EC 3.2.1.37) ja sarja entsyymejä, jotka pilkkovat sivuketjusokereita (glykosi-daasit) tai poistavat asetyyliryhmiä ksylaanirangasta [Dekker R.F.H. ja Richards, G.N., Adv. Carbohydr. Chem. Biochem. 32: 277-352 (1976); Biely, Trends Biotechnol. 3: 286-290 (1985); Poutanen et al., "Acessory Enzymes Involved in the Hydrolysis of Xylans" julkaisussa Enzymes in Biomass Conversion. ACS Symposium Series 460, ss. 426-436. Toim. G.F. Letham. (1991)]. Ksylanolyyttiset mikro-organismit ilmentävät yleensä ksylanaasien isoentsyymisiä muotoja, joita monet geenit koodaavat [Hazlewood et al., FEMS Microbiol. Lett. 51: 231-236 (1988); Gilbert et al., J. Gen. Microbiol. 134: 3239-3247 (1988); Clarke et al., FEMS Microbiol. Lett. 83: 305-310 (1991)].

Eräät ksylanaasit hydrolysoivat ainoastaan ksylaania [Hall et al., Mol. Microbiol. 3: 1211-1219 (1989);Wong et al., Microbiol. Rev. 52: 305-317 (1988)]. Monet mikro-organismit, jotka hydrolysoivat ksylaania, hajoittavat myös selluloosaa. Pilkotun sidoksen (ß-1,4-glykosidisidokset) samankaltaisuutta ja sellulaasien ja ksylanaasien välillä toisinaan havaittua ristikkäisspesifisyyttä ajatellen näiden entsyymien fylogeneettiset suhteet ovat mielenkiintoinen kysymys. Äskettäin on käytetty sekvenssinkohdakkaisvertailua ja hydrofobista ryväsanalyysiä kasvin soluseinän hydrolaasien sijoittamiseksi kahdeksaan entsyymiryhmään [Henrissat et al., Gene 81: 83-95 (1989); Gilkes et al., Microbiol. Rev. 55: 303-315 (1991)]. Ksylanaaseilla ei ollut mitään vakuuttavaa sekvenssien yhdenmukaisuutta sellulaasien kanssa, mikä viittaa siihen, että nämä kaksi entsyymilajia ovat kehittyneet eri alkugeeneistä.

Monet kasvien soluseinän hydrolaasit muodostuvat kahdesta eri domainista; katalyyttisestä domainista (CD), joka on liittynyt hydroksiaminohappo/runsasproliinisilla linkkerisekvensseillä ei-katalyyttiseen selluloosaa sitovaan domainiin [CBD; Gilkes et al., Microbiol. Rev. 55: 303-315 (1991); Kellett et al., Biochem. J. 272: 369-376 (1990); Gilbert et al., Mol. Microbiol. 4: 759-767 (1990)]. CBD:n täsmällisestä roolista väitellään kovasti; aerobisissa sienisellulaaseissa CBD:llä on kriittinen rooli kiteisen selluloosan entsyymien hydrolyysissä [Tomme et al., Eur. J. Biochem. 170: 575-581 (1988)]. Tämän domainin rooli prokaryoottisissa sellulaaseissa ja ksylanaaseissa on vähemmän varma [Ferreira et al., Biochem. J. 269: 261-264 (1990)]. Modulaarisen rakenteensa lisäksi sellulaasit sisältävät usein pitkiä toistuvia sekvenssejä [Gilkes et al., Microbiol. Rev. 55: 303-315 (1991)]. Näiden tamdemtoistojen täsmällinen rooli on vielä kovasti ratkaisematon.

Monia selluloosaa ja hemiselluloosaa hajoittavia prokaryootteja on lehmien ja lampaiden pötsissä. Äskettäin on myös anaerobisten pötsissä esiintyvien sienien osoitettu hajoittavan tehokkaasti sekä selluloosaa että ksylaania [Orpin ja Letcher Curr. Microbiol. 3: 121-124 (1979); Lowe et al., Appln. Environ. Microbiol. 53: 1216-1223 (1987)] ja samanlaisia sieniä esiintyy suurikokoisten herbivorien ruoansulatuskanavassa [Orpin ja Joblin, "Anaerobic fungi". Julkaisussa The Rumen Microbial Ecosystem, P.N. Hobson (toim.), ss. 129-150, Elsevier, Lontoo (1988)]. Pötsissä esiintyvän sienen Neocallimastix frontalis selluloosakompleksin ovat karakterisoineet Wood et al., Biochemistry and Genetics of Cellulose Degradation: FEMS Symp. 43: 31-52 (1988). Alempi eukaryootti syntetisoi suuren monientsyymikompleksin, jonka M_r on 1-2 miljoonaa ja joka hydrolysoi nopeasti kiteistä selluloosaa. Kompleksi sisältää pääasiassa endoglukanaasiaktiivisuutta ja jonkin verran ß-glukosidaasiaktiivisuutta. Sieni syntetisoi myös erästä avicelaasia, oletettavasti erästä sellobiohydrolaasia. Eräs toinen pötsisieni, Neocallimastix patriciarum, tuottaa solunulkoisia entsyymejä, jotka hydrolysoivat suodatinpaperiselluloosaa, AVICELiä™ (eräs mikrokiteisen selluloosan tavaramerkki) ja ksylaania [Williams ja Orpin, Can J. Microbiol. 33: 418-426 (1987)]. Mitään näistä entsyymeistä ei ole karakterisoitu. Jonkin verran tietoa Neocallimastix-geeneistä, jotka koodaavat kasvien soluseinän hydrolaaseja, on kuvattu [Reymond et al., FEMS Microbiol. Lett. 77: 107-112 (1991)].

Ksylaaneja tavataan ligniiniin liittyneenä useimpien kasvien solujen primaari- ja sekundaariseinissä. Ksylaanin ja ligniinin välinen assosiaatio on avain ksylanaasien kaupalliseen potentiaaliin muun muassa paperimassan käsittelyssä. Sandoz Products Ltd USA:ssa on jo tehnyt käytännön kokeita, joissa on käytetty käsittelemätöntä sie-

niksylanaasia korvaamaan, ainakin osaksi, klooria ja kloorista johdettuja yhdisteitä, joita tavallisesti käytetään ei-toivottujen ruskeiden ligniiniperustaisten jäämien, joita syntyy puumassan käsittelyssä paperi- ja muiden puuperustaisten tuotteiden tuotannossa, valkaisemiseen. Nykyisten puunjalostustehtaiden kloorivaatimukset ovat sellaiset, että jokaisella tehtaalla tulee olla oma klooridioksidin tuotantoyksikkö.

Edut, joita kloorin käytön välttäminen tuottaa paperiteollisuudelle, ovat selviä: jätteiden käsittelyssä, työntekijöiden turvallisuudessa ja tehdaspääomassa voitaisiin saada aikaan parannuksia, mikäli pystyttäisiin löytämään sopiva korvaava aine kloorille. Paperiteollisuudessa käydään kiivasta kilpailua ja voittomarginaalit ovat kapeat, joten mahdollista kloorin korvaavaa ainetta tulee voida tuottaa kohtuullisen taloudellisesti ja sen tulee tietysti olla myös tarpeeksi tehokasta, jotta sellun ja paperin valmistajat saadaan vakuuttumaan sen käytön eduista.

Neocallimastix patriciarumista peräisin olevan ksylanaasin täysipitkä cDNA- ja proteiinisekvenssi oli saatavissa EMBL-tietopankista Heidelbergista, Saksasta, 5. maaliskuuta 1992 saantinumerolla X65526. Ksylanaasi nimettiin XYLAksi ja vastaava geeni *xyn*A:ksi.

Nyt on todettu, että anaerobisista sienistä peräisin olevista yksittäisistä ksylanaaseista johdetuilla muunnetuilla ksylanaaseilla, kuten N. patriciarumista peräisin olevalla XYLA-entsyymillä on ominaisuuksia, jotka tekevät ne sopiviksi teollisuuden käyttöön, erityisesti sellu- ja paperiteollisuuden käyttöön. Yllättävästi käy ilmi, että katkaiseminen saattaa parantaa entsyymin ilmentymistä.

Esillä olevan keksinnön ensimmäisen aspektin mukaan käyttöön saadaan ksylanaasi, jolla on ainakin yksi katalyyttinen domaini, joka on olennaisesti homologinen erään anaerobisen sienen ksylanaasin kanssa ja joka ei ole täyspitkä luonnon ksylanaasi.

Suositeltavat katalyyttiset domainit ovat identtisiä anaerobisista sienistä peräisin olevien luonnon ksylanaasien katalyyttisten domainien kanssa. Kuitenkin esillä olevan keksinnön tarkoituksiin ensimmäinen sekvenssi on olennaisesti homologinen toisen sekvenssin kanssa, jos se esimerkiksi jakaa sen biologisen aktiivisuuden ja aminohappotasolla on vähintään noin 40 prosentin homologia; keksinnön tämän aspektin mukaisen ksylanaasin katalyyttinen domaini on siis vähintään noin 40-prosenttisesti homologinen jostakin anaerobisesta sienestä peräisin olevan luonnon ksylanaasin katalyyttisen domainin kanssa. Yleensä voi olla suositeltavaa, että kahden vertailtavan aminohapposekvenssin välillä on vähintään 50 %:n, 60 %:n, 70 %:n,

80 %:n tai 90 %:n homologia (mieluiten nousevassa järjestyksessä). Homologia voidaan vaihtoehtoisesti tai lisäksi määrittää nukleiinihappotasolla. Ensimmäistä aminohapposekvenssiä koodaava DNA voi olla olennaisen homologinen toista aminohapposekvenssiä koodaavan DNA:n (joka voi olla cDNA tai genomi-DNA) kanssa tai hybridisoitua siksi tai se hybridisoituisi, mutta geneettisen koodin degeneroitumisen kustannuksella. Hybridisaatio-olosuhteet voivat olla tiukat, kuten 65 °C noin 0,9-molisessa suolaliuoksessa.

Esimerkkejä anaerobisista sienistä, jotka voivat olla ruoansulatuskanava-(erityisesti pötsi-)sieniä, ovat: Neocallimastix spp., kuten N. patriciarum, N. frontalis, N. hurleyensis ja N. stanthorpensis; Sphaeromonas spp., kuten S. communis, Caecomyces spp., kuten C. equi; Piromyces spp., kuten P. communis, P. equi, P. dumbonica, P. lethargicus ja P. mai; Ruminomyces spp., kuten P. elegans; Anaeromyces spp., kuten A. mucronatus, ja Orpinomyces spp., kuten O. bovis ja O. joyonii.

Caecomyces equi, Piromyces equi, Piromyces dumbonica ja Piromyces mai esiintyvät hevosissa eivätkä nautakarjan pötsissä kuten muut edellä luetellut sienet. Neocallimastix spp. ovat suositeltavimpia, erityisesti N. patriciarum.

Keksinnön mukaisilla ksylanaaseilla voi olla erittäin spesifinen aktiivisuus. Spesifinen aktiivisuus voi olla merkittävästi suurempi kuin bakteerien avulla johdettujen ksylanaasien ja se voi olla esimerkiksi vähintään 1000, 2000, 3000, 4000, 4500, 5000 tai jopa 5500 U/mg proteiinia, mieluiten suurenevassa järjestyksessä. (Ksylanaasiaktiivisuusyksikkö määritetään entsyymin määränä, joka vapauttaa 1 µmoolin tuotetta, ksyloosiekvivalentteina mitattuna, 1 minuutissa 37 °C:ssa). Tarkemmin sanottuna keksinnön tämän aspektin mukaiset ksylanaasit voivat ilmentyä merkittävästi paremmin kuin mitä N. patriciarum ilmentää luonnon XYLAa; ilmentyminen voi olla vähintään 10 kertaa parempi tai mielellään vähintään 100 kertaa parempi kuin villin tyypin entsyymillä.

Keksinnön mukaisilla ksylanaaseilla voi olla kyky hajoittaa ksylaania tehokkaasti. Yhtä grammaa ksylaania kohti voidaan tuottaa vähintään 0,1 ja mieluiten vähintään 0,5 tai jopa 0,75 g pelkistävää sokeria.

Keksinnön mukaisilla ksylanaaseilla ei ole mitään merkittävää jäännösaktiivisuutta selluloosaan päinvastoin kuin monilla tunnetuilla ksylanaaseilla. Tämä ominaisuus on erityisen hyödyllinen, kun keksintöä sovelletaan sellu- ja paperiteollisuuteen, koska entsyymi voi poistaa ksylaania ja hajoittaa ligniiniä kasvikuidusta vahingoittamat-

ta selluloosakuitua.

Keksinnön mukaisilla ksylanaaseilla voi olla vähintään kaksi katalyyttistä domainia. Katalyyttiset domainit voivat olla järjestyneinä kuten villin tyypin ksylanaasientsyymissä tai ne voivat olla järjestyneet keinotekoiseen konfiguraatioon entsyymin ksylanolyyttisen aktiivisuuden lisäämiseksi tai muulla tavoin parantamiseksi.

Eräs erityisen suositeltava ksylanaasi keksinnössä katalyyttisten domainien lähteenä käytettäväksi on ksylanaasi, joka on johdettu Neocallimastix patriciarumista ja joka on nimitetty XYLAksi; sillä on seuraavat ominaisuudet:

(i) spesifinen aktiivisuus 5980 U/mg proteiinia puhdistetulla entsyymillä, kun se on valmistettu seuraavan protokollan mukaisesti:
isäntäsoluja (E. coli XL1-Blue, joka sisältää entsyymiä ekspressoivan plasmidin) kerätään sentrifugoimalla ja suspendoimalla uudelleen 50 mM Tris-HCl-puskuriin, pH 8,0, ja sytoplasmifraktio valmistetaan, kuten ovat kuvanneet Clarke et al. [FEMS Microbiol. Letts. 83: 305-310 (1991)]. Ksylanaasi, joka on saostunut ammoniumsulfaattia (0,39 g/ml) lisättäessä, liuotetaan uudelleen 10 mM Tris-HCl-puskuriin, pH 8,0. Dialyysin jälkeen samaa kolmesti vaihdettua puskuria vastaan ksylanaasi puhdistetaan olennaisesti anioninvaihtokromatografialla DEAE-Triacryl M:llä, kuten ovat kuvanneet Hall et al. [Mol. Microbiol. 3: 1211-1219 (1989)];
(ii) kyky hajoittaa ksylaania tehokkaasti, vapauttaen 0,9 g pelkistävää sokeria/g substraattia;
(iii) ei merkittävää jäännösaktiivisuutta selluloosaan (määritetään ei havaittavana pelkistävän sokerin vapautumisena karboksimetyyliselluloosasta, ohran ß-glukaanista, laminariinista tai likenaanista); ja
(iv) kaksi katalyyttistä domainia.

Kypsän XYLAn rakenne voidaan esittää seuraavasti (N-päästä C-päähän):

CAT1-LINK1-CAT2-LINK2-CTR1-CTR2

jossa:
– CAT1 edustaa ensimmäistä katalyyttistä domainia, jolla on sekvenssi:

RLTVGN
GQTQHKGVADGYSYEIWLDNTGGSGSMTLGSGATFKAEWN
ASVNRGNFLARRGLDFGSQK KATDYSYIGLDYTATYRQTG

SASGNSRLCVYGWFQNRGVQ GVPLVEYYIIEDWVDWVPDA
QGRMVTIDGAQYKIFQMDHT GPTINGGSETFKQYFSVRQQ
KRTSGHITVSDHFKEWAKQG WGIGNLYEVALNAEGWQSSG
IADVTKLDVYTTQKGSNPAP;

CAT2 edustaa toista katalyyttistä domainia, jolla on sekvenssi:

K
FTVGNGQNQHKGVNDGFSYEIWLDNTGGNGSMTLGSGATF
KAEWNAAVNRGNFLARRGLDFGSQKKATDYDYIGLDYAAT
YKQTASASGNSRLCVYGWFQ NRGLNGVPLVEYYIIEDWVD
WVPDAQGKMVTIDGAQYKIF QMDHTGPTINGGSETFKQYF
SVRQQKRTSGHITVSDHFKE WAKQGWGIGNLYEVALNAEG
WQSSGVADVTLLDVYTTPKG SSPA;

LINK1 edustaa ensimmäistä linkkeriä, jolla on sekvenssi:

TSTGTVPSSSAGGSTANGK;

LINK2 edustaa toista linkkeriä, jolla on sekvenssi:

TSAAPRTTTRTTTRTKSLPTNYNK;

CTR1 edustaa ensimmäistä C-teerminaalista toistuvaa jaksoa, jolla on sekvenssi:

CSARITAQGYKCCSDPNCVVYYTDEDGTWGVENNDWCGCG;

ja

CTR2 edustaa toista C-terminaalista toistuvaa jaksoa, jolla on sekvenssi:

VEQCSSKITSQGYKCCSDPNCVVFYTDDDGKWGVENNDWC
GCGF.

Kaikki nämä sekvenssin osat ovat nähtävissä sekvenssin tunnistusnumerossa 1 (SEQ

ID NO: 1) ja sekvenssin tunnistusnumerossa 2 (SEQ ID NO: 2).

Muista anaerobisista sienistä peräisin olevien ksylanaasien rakenne saattaa olla kovin samanlainen, mutta tietysti komponenttien täsmälliset sekvenssit ovat yleensä erilaiset, jollei lähdeorganismi muistuta hyvin läheisesti N. patriciarumia. Jokaisen alueen (erityisesti katalyyttisten domainien) koko sekvenssin ei välttämättä tarvitse olla läsnä aktiivisuuden kannalta; vaikka koko katalyyttinen domaini voi olla läsnä, esillä olevassa keksinnössä riittää, että katalyyttisen domainin aktiivinen osa on läsnä (toisin sanoen katalyyttisen domainin tulee olla toiminnallisesti läsnä).

Näiden kahden katalyyttisen domainin voidaan havaita olevan kovin toistensa kaltaisia, mutta ei identtisiä. Niiden välinen eroavaisuus osoittaa luonnon sekvenssiin nähden homologian astetta, joka on erityisen suositeltava. Kahden C-terminaalisen toistuvan jakson voidaan myös havaita olevan toistensa kaltaisia (mutta vähemmän kuin kahden katalyyttisen domainin). Niiden välinen eroavaisuus osoittaa homologian asteen, mikä on vielä erittäin suositeltava. Kahden linkkerisekvenssin tarkka sekvenssi ei liene erityisen tärkeä; kaikki mitä tarvitaan on, että katalyyttisen domainin tai katalyyttisten domainien avaruudellinen järjestys on sellainen, että ne voivat toimia tehokkaasti (ja mieluiten optimaalisesti).

Keksinnön suositeltavimmat toteutusmuodot käsittävät katalyyttisen domainin, joka on olennaisen homologinen ainakin joko CAT1:n tai CAT2:n kanssa ja niistä puuttuu ainakin osa CAT2:n aminohapposekvenssistä alavirrassa (so. C-päähän päin). Ainakin osa CTR2:sta voi puuttua; vaihtoehtoisesti tai (mieluiten) lisäksi ainakin osa CRT1:stä voi puuttua.

Keksinnön mukaisten ksylanaasien erityiset toteutusmuodot sisältävät sellaiset, joissa on seuraavat alueet (ja jotka suositeltavimmin muodostuvat pääasiassa niistä):

A.	CAT1-LINK1-CAT2-LINK2-CTR1(katkaistu)	(eg pNX3);
B.	CAT1-LINK1-CAT2-LINK2(katkaistu)	(eg pNX4);
C.	LINK1(katkaistu)-CAT2-LINK2(katkaistu)	(eg pNX5);
D.	CAT1-LINK1(katkaistu)	(eg pNX6);
E.	CAT1(katkaistu)	(eg pNX7);
F.	LINK1(katkaistu)-CAT2-LINK2-CTR1-CTR	(eg pNX8);
G.	LINK1(katkaistu)-CAT2-LINK2-CTR1(katkaistu)	(eg pNX9);
H.	LINK1(katkaistu)-CAT2(katkaistu)	(eg pNX10).

(Sulkeissa olevat plasmidien tunnukset viittaavat esimerkeissä mainittuihin plasmideihin, joiden ilmentymistuotteita esitetyt ksylanaasit ovat). Signaalisekvenssejä voi aluksi olla läsnä, mutta lopullisesta molekyylistä ne mieluiten puuttuvat. Rakenteet C, F, G ja H ovat suositeltavia ja rakenteet C, G ja H ovat erityisen suositeltavia.

Keksinnön mukaiset entsyymit voivat sisältää yhden ainoan CAT1-domainin, yhden ainoan CAT2-domainin tai niissä voi olla kaksi tai useampia katalyyttisiä domaineja, joista kukin voidaan valita muista riippumatta CAT1:n ja CAT2:n joukosta. Saattaa olla, että läsnä on olennaisesti vain katalyyttisiä alueita; ja kuten edellä todettiin, voi olla, että kaikki luonnon katalyyttisen domainin sekvenssit eivät ole tarpeen sopivan aktiivisuuden kannalta.

Kypsymättömässä proteiinissa voi olla läsnä jokin signaalipeptidi; luonnossa esiintyvän signaalipeptidin sekvenssi on:

MRTIKFFFAVAIATVAKAQWGGGGASAGQ.

Tämä sekvenssi on jälleen nähtävissä SEQ ID NO: 1:ssä ja SEQ ID NO: 2:ssa.

Keksinnön mukaisia ksylanaaseja voidaan valmistaa millä tahansa sopivalla tavalla. Vaikka voidaan ryhtyä lähteenä olevan anaerobisen sienen bulkkifermentoimiseen ja voidaan yrittää polypeptidisynteesiä orgaanisen kemian menetelmillä, parhaassa valmistusmenetelmässä käytetään yleensä rekombinantti-DNA-tekniikkaa. Edellä kuvattu ksylanaasi on sen vuoksi mieluiten heterologisen ksylanaasia koodaavan DNA:n ilmentymistuote isäntäsolussa.

Keksinnön toisen aspektin mukaan kysymyksessä on eristetty tai rekombinantti-DNA-molekyyli, joka koodaa ksylanaasia, jossa on katalyyttinen domaini, joka on olennaisen homologinen jonkin anaerobisen sienen ksylanaasin kanssa, edellyttäen että DNA-molekyyli ei käsitä ksylanaasia koodaavan luonnossa esiintyvän mRNA:n täyspitkää kopiota.

Reymond et al., FEMS Microbiol. Lett. 77: 107-112 (1991) ovat kuvanneet cDNA:ta (joka ilmeisesti sisältää mRNA:n täyspitkän kopion), joka koodaa Neocallimastix frontalisin ksylanaasia, mutta ekspressiota ei ole raportoitu.

Vaikka keksinnön tämän aspektin mukaisessa DNA:ssa ei ole luonnossa esiintyvän mRNA:n täyspitkää kopiota, lienee selvää, että keksintö ei rajoitu katkaistuihin

cDNA:ihin. Voidaan ajatella, että läsnä voi olla joitakin introneja tai kaikki intronit (jos niitä on), joita on luonnossa läsnä vastaavassa villin tyypin geenissä. Kuitenkin ainakin jokin sekvenssi, joka on läsnä täyspitkässä cDNA:ssa, puuttuu keksinnön tämän aspektin mukaisesta DNA:sta. Lienee myös selvää, että keksinnön tämä aspekti käsittää DNA:t, jotka koodaavat täyspitkiä ksylanaaseja; DNA:n poissa oleva osa voi olla (ja joissakin toteutusmuodoissa mieluiten on) 3'- ja 5'-translatoitumattomilla alueilla. Olennaisesti täyspitkiä tai katkaistuja ksylanaaseja voidaan sen vuoksi tuottaa keksinnön tämän aspektin mukaisesta DNA:sta, josta (a) olennaisesti puuttuu 3'-translatoitumaton alue tai josta (b) olennaisesti puuttuu 5'-translatoitumaton alue tai josta (c) puuttuvat sekä 3'- että 5'-translatoitumattomat alueet.

Täyspitkällä cDNA:lla, joka koodaa jonkin anaerobisen sienen ksylanaasia (ottaen N. patriciarumin *xyn*A-geenin prototyypiksi), voi olla seuraava rakenne:

*5'utr-sig-cat*1-*link*1-*cat*2-*link*2-*ctr*1-*ctr*2-*3'utr*,

jossa
- *5'utr* tarkoittaa 5'-translatoitumatonta aluetta;
- *sig* koodaa signaalipeptidiä;
- *cat*1 koodaa ensimmäistä katalyyttistä domainia;
- *link*1 koodaa ensimmäistä linkkerisekvenssiä;
- *cat*2 koodaa toista katalyyttistä domainia;
- *link*2 koodaa toista linkkerisekvenssiä;
- *ctr*1 koodaa ensimmäistä C-terminaalista toistuvaa jaksoa;
- *ctr*2 koodaa toista C-terminaalista toistuvaa jaksoa; ja
- *3'utr* tarkoittaa 3'-translatoitumatonta aluetta.

Genomisissa sekvensseissä voi olla yksi tai useampi introni sirottuneena edellä mainittuun rakenteeseen. N. patriciarumin XYLA-entsyymiä koodaavassa *xyn*A-geenissä eri DNA-segmenteillä on seuraavat sekvenssit:

3'utr:

TTTTATTATATCAATCTCTAATTTATTTTTTTAGGAAAAAAATAAAAAAATAAATATAAT
AAATATTAGAGAGTAATATTTAAAAACAAAGAAATTTAAAAACGTTTATTTAGTTATTTT
TTTTACTGGTTAAAAAAAAAATAAAAAACAAAATTAATAAAGATATTTTTGAAAAATATT
GAATTAGAAAAAAA;

sig:

ATGAGAACTATTAAATTCTTTTTCGCAGTAGCTATTGCAACTGTTG
CTAAGGCCCAATGGGGTGGAGGTGGTGCCTCTGCTGGTCAA;

cat1:

AGATTAACCGTCGGTAATG
GTCAAACCCAACATAAGGGTGTAGCTGATGGTTACAGTTATGAAATCTGGTTAGATAACA
CCGGTGGTAGTGGTTCTATGACTCTCGGTAGTGGTGCAACCTTCAAGGCTGAATGGAATG
CATCTGTTAACCGTGGTAACTTCCTTGCCCGTCGTGGTCTTGACTTCGGTTCTCAAAAGA
AGGCAACCGATTACAGCTACATTGGATTGGATTATACTGCAACTTACAGACAAACTGGTA
GCGCAAGTGGTAACTCCCGTCTCTGTGTATACGGTTGGTTCCAAAACCGTGGAGTTCAAG
GTGTTCCATTGGTAGAATACTACATCATTGAAGATTGGGTTGACTGGGTTCCAGATGCAC
AAGGTAGAATGGTAACCATTGATGGAGCTCAATATAAGATTTTCCAAATGGATCACACTG
GTCCAACTATCAATGGTGGTAGTGAAACCTTTAAGCAATACTTCAGTGTCCGTCAACAAA
AGAGAACTTCTGGTCATATTACTGTCTCAGATCACTTTAAGGAATGGGCCAAACAAGGTT
GGGGTATTGGTAACCTTTATGAAGTTGCTTTGAACGCCGAAGGTTGGCAAAGTAGTGGTA
TAGCTGATGTCACCAAGTTAGATGTTTACACAACCCAAAAAGGTTCTAATCCTGCCCCT;

link1:

ACCTCCACTGGTACTGTTCCAAGCAGTTCTGCTGGTGGAAGTACTGCCAATGGTAAA;

cat2:

AAGT
TTACTGTCGGTAATGGACAAAACCAACATAAGGGTGTCAACGATGGTTTCAGTTATGAAA
TCTGGTTAGATAACACTGGTGGTAACGGTTCTATGACTCTCGGTAGTGGTGCAACTTTCA
AGGCTGAATGGAATGCAGCTGTTAACCGTGGTAACTTCCTTGCCCGTCGTGGTCTTGACT
TCGGTTCTCAAAAGAAGGCAACCGATTACGACTACATTGGATTAGATTATGCTGCTACTT
ACAAACAAACTGCCAGTGCAAGTGGTAACTCCCGTCTCTGTGTATACGGATGGTTCCAAA
ACCGTGGACTTAATGGCGTTCCTTTAGTAGAATACTACATCATTGAAGATTGGGTTGACT
GGGTTCCAGATGCACAAGGAAAAATGGTAACCATTGATGGAGCTCAATATAAGATTTTCC
AAATGGATCACACTGGTCCAACTATCAATGGTGGTAGTGAAACCTTTAAGCAATACTTCA
GTGTCCGTCAACAAAAGAGAACTTCTGGTCATATTACTGTCTCAGATCACTTTAAGGAAT
GGGCCAAACAAGGTTGGGGTATTGGTAACCTTTATGAAGTTGCTTTGAACGCCGAAGGTT
GGCAAAGTAGTGGTGTTGCTGATGTCACCTTATTAGATGTTTACACAACTCCAAAGGGTT
CTAGTCCAGCC;

link2:

ACCTCTGCCGCTCCTCGTACTACTACCCGTACTACTACTCGTACCAAGTCTCTTCCAACC
AATTACAATAAG;

ctr1:

TGTTCTGCTAGAATTACTGCTCAAGGTTACAAGTGTTGTAGCGATCCAAATTGTGTTGTT
TACTACACTGATGAGGATGGTACCTGGGGTGTTGAAAACAACGACTGGTGTGGTTGTGGT;

ctr2:

GTTGAACAATGTTCTTCCAAGATCACTTCTCAAGGTTACAAGTGTTGTAGCGATCCAAAT
TGCGTTGTTTTCTACACTGATGACGATGGTAAATGGGGTGTTGAAAACAACGACTGGTGT
GGTTGTGGTTTC; ja

5'utr:

TAAGCAGTAAAATACTAATTAATAA
AAAATTAAAGAATTATGAAAAATTTAAATTTAAAAATTTAAAAGAATTATGAAAAATTTA
AATTTAAAAATTTAAAAAAAAACTAATTTAGTAAAAAATTAAAGAATTATTGAAAATTTTA
AATGTAAAAATTTAAAAAATACAAATTTGTAAAAAAAAATGAAAGAATTATGAAAAATTA
AAATGTAAAAGTTTAAAAAATACAAATTTGTAAGAAAAATAAAGAATTATAAAAAAAATA
AAGAATTATGAAAAACCCAAATGTAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

(Huomaa, että *5'utr*-segmentin kolme ensimmäistä nukleotidiä muodostavat aloituskodonin, joka yleensä on läsnä).

Näiden DNA-segmenttien (vähemmän kuin kaikki ne) tai niiden kanssa olennaisen homologisten segmenttien käyttö on suositeltavaa keksinnön tässä aspektissa. Suositeltavat toteutusmuodot vastaavat yleensä ksylanaasien *per se* suositeltavia toteutusmuotoja keksinnön ensimmäisen aspektin mukaan, mutta ottaen lisäksi huomioon, että (a) saattaa olla suositeltavaa, että läsnä on DNA-sekvenssi, joka koodaa jotakin peptidisignaalisekvenssiä, ja/tai (b) saattaa olla suositeltavaa, että yksi translatoitumattomista alueista tai molemmat translatoitumattomat alueet ovat katkaistuja tai poissa. Keksinnön tämän aspektin mukaiset erityistoteutusmuodot ovat sellaisia, jotka sisältävät seuraavat rakenteet (ja mieluiten muodostuvat olennaisesti niistä, vektorista johdettuja sekvenssejä lukuunottamatta):

a. *5'utr-sig-cat*1-*link*1-*cat*2-*link*2-*ctr*1(katkaistu) (eg pNX3);
b. *5'utr-sig-cat*1-*link*1-*cat*2-*link*2(katkaistu) (eg pNX4);
c. *link*1(katkaistu)-*cat*2-*link*2(katkaistu) (eg pNX5);
d. *5'utr-sig-cat*1-*link*1(katkaistu) (eg pNX6);
e. *5'utr-sig-cat*1(katkaistu) (eg pNX7);
f. *link*1(katkaistu)-*cat*2-*link*2-*ctr*1-*ctr*2-*3'utr* (eg pNX8);
g. *link*1(katkaistu)-*cat*2-*link*2-*ctr*1(katkaistu) (eg pNX9);
h. *link*1(katkaistu)-*cat*2(katkaistu) (eg pNX10).

(Sulkeissa olevat plasmidien tunnukset viittaavat esimerkeissä esiintyviin plasmideihin, jotka sisältävät esitetyt DNA-sekvenssit). Rakenteet *c, f, g* ja *h* ovat suositeltavia ja rakenteet *c, g* ja *h* ovat erityisen suositeltavia.

Rekombinantti-DNA voi keksinnön mukaan olla vektorin muodossa. Vektori voi olla esimerkiksi plasmidi, kosmidi tai faagi. Vektorit sisältävät usein yhden tai useamman markkerin niillä transfektoitujen (tai transformoitujen: termejä käytetään tässä selityksessä toistensa vaihtoehtoina) solujen valikoinnin ja, suositeltavimmin, niiden solujen, joissa heterologisen DNA:n sisältävät vektorit ovat, seulonnan mahdollistamiseksi. Läsnä on yleensä sopivia aloitus- ja lopetussignaaleja. Lisäksi jos vektori on tarkoitettu ekspressointiin, läsnä on riittävästi säätteleviä sekvenssejä ekspression ohjaamiseksi. Vektoreita, joissa ei ole säätteleviä sekvenssejä, voidaan käyttää kloonausvektoreina; ja ekspressiovektoreita voidaan tietysti myös käyttää kloonausvektoreina.

Kloonausvektorit voidaan viedä E. coliin tai johonkin muuhun sopivaan isäntään niiden käsittelyn helpottamiseksi. Keksinnön erään toisen aspektin mukaan se koskee sen vuoksi edellä kuvatulla DNA:lla transfektoitua tai transformoitua isäntäsolua.

Keksinnön mukainen DNA voidaan valmistaa millä tahansa sopivalla menetelmällä, jossa kytketään toisiinsa peräkkäisiä nukleotideja ja/tai ligatoivia oligo- ja/tai polynukleotideja, mukaanlukien in vitro -menetelmät, mutta rekombinantti-DNA-tekniikka on paras menetelmä.

Ksylanaasia koodaava DNA voidaan kloonata DNA-kirjastosta, joka voidaan valmistaa jostakin edellä mainitusta sienestä. Kirjasto voi olla genominen, mutta cDNA-kirjasto saattaa olla helpompi valmistaa ja käsitellä, erityisesti jos ryhdytään toimenpiteisiin ksylanaasia koodaavan cDNA:n läsnäolon cDNA-kirjastossa todennäköisyyden parantamiseksi.

Valitun sienen, kuten N. patriciarumin, viljely voidaan toteuttaa anaerobisesti jossakin sopivassa viljelyalustassa, joka sisältää pötsinestettä; ainoana tai hallitsevana hiilen lähteenä voi olla ksylaani ksylanaasin ilmentymisen edistämiseksi ja sen seurauksena ksylanaasia koodaavan RNA:n määrän lisäämiseksi. Viljeleminen ksylaanin läsnäollessa ei ole kuitenkaan ehdottoman välttämätöntä ja hiilen lähteenä voi sen sijaan olla jokin selluloosa, kuten mikrokiteinen selluloosa, jota myydään tavaramerkillä AVICEL.

Sienen viljelyn jälkeen kokonais-RNA voidaan erottaa millä tahansa sopivalla tavalla. Sienisoluja voidaan korjata talteen suodattamalla ja sen jälkeen ne voidaan hajoittaa sopivassa solunhajoituspuskurissa mekaanisesti rikkomalla. Jotakin sopivaa RNA:ta säilyttävää yhdistettä, kuten guanidiniumtiosyanaattia, voidaan myös lisätä sienisoluihin Rnaasivälitteisen digestion vähentämiseksi tai estämiseksi. Kokonais-RNA voidaan sitten eristää saadusta homogenaatista millä tahansa sopivalla menetelmällä, kuten ultrasentrifugoimalla $CsCl_2$-tyynyn läpi tai kuten kuvaavat Sambrook et al., Molecular Cloning: A Laboratory Manual, 2. painos. Cold Spring Harbor. New York: Cold Spring Harbor Laboratory Press (1989).

Eräs toinen menetelmä kokonaissieni-RNA:n valmistamiseksi edellä kuvatun lisäksi voi perustua menettelytapaan tai olla sovellettu menettelytavasta, jota kuvataan julkaisussa Puissant ja Houdebine Bio-Techniques 148-149 (1990). Tässä menetelmässä kokonaissieni-RNA voidaan eristää edellä mainitusta homogenaatista uuttamalla fenoli/kloroformilla pH 4:ssä DNA:n ja siihen liittyvän proteiinin poistamiseksi. Saatu puhdistamaton RNA puhdistettiin sitten pesemällä litiumkloridi-urealiuoksella.

Vielä eräs sopiva menetelmä sieni-RNA:n erottamiseksi on Teeren ja työtoverien menetelmä [Anal. Biochem. 164: 60-67 (1987)].

Kun kokonais-RNA on erotettu, millä tahansa menetelmällä, voidaan poly-A^+ mRNA sitten eristää kokonais-RNA:sta, esimerkiksi affiniteettikromatografialla yhdisteellä, joka sisältää paljon tymidiini- tai urasiilitähteitä, joihin mRNA:n poly-A-häntä voi sitoutua. Esimerkkejä sopivista yhdisteistä ovat oligo-dT-selluloosa ja poly-U SEPHADEX™. Poly-A^+ mRNA voidaan sitten eluoida sopivalla puskurilla.

cDNA-ekspressiokirjasto voidaan sitten rakentaa käyttämällä jotakin standardimenetelmää, joka perustuu poly-A^+ mRNA:n muuttamiseen cDNA:ksi käänteistranskriptaasin avulla. Vaikka on mahdollista rakentaa genominen kirjasto, cDNA-kirjasto on suositeltavampi, koska siinä vältytään vaikeuksilta, joita intronien läsnäolo sienipe-rustaisessa genomisessa DNA:ssa saattaa aiheuttaa. cDNA:n ensimmäinen nauha voidaan syntetisoida käyttämällä käänteistransskriptaasia ja toinen nauha voidaan syntetisoida käyttämällä mitä tahansa DNA-johtoista DNA-polymeraasia, kuten Escherichia coli DNA-polymeraasia I (E. coli pol I).

cDNA voidaan sen jälkeen fraktioida sopivaan kokoon ja se voidaan sitoa johonkin sopoivaan vektoriin, joka on suositeltavimmin faagivektori, kuten λZAP, λZAPII tai λgt 11. Tarkoitukseen sopivia kittejä on saatavissa Stratageneltä. Lisä- tai vaihtoeh-

toista neuvoa voidaan saada Reymondilta työtovereineen [FEMS Microbiol. Lett. 17: 107-112 (1991)], jotka esittävät yksityiskohtaisesti cDNA-kirjaston valmistuksen N. frontaliksesta. Saatu cDNA-kirjasto voidaan sitten monistaa pakkaamisen jälkeen in vitro, käyttäen mitä tahansa sopivaa isäntäbakteerisolua, kuten jotakin sopivaa E. coli-kantaa.

Ksylanaasille positiivisten rekombinantti-kloonien seulonta voidaan suorittaa millä tahansa sopivalla tekniikalla, joka voi perustua ksylaanin hydrolyysiin. Tässä menetelmässä klooneja voidaan kasvattaa viljelyalustassa, joka sisältää ksylaania, ja hydrolyysi voidaan detektoida ksylanaasipositiivisten plakkien esiintymisen avulla käyttäen jotakin sopivaa indikaattoriväriä. Menetelmiä ksylanaasi^{+}-kloonien seulomiseksi kuvataan kirjallisuudessa. Kaksi esimerkkiä ovat Clarke et al [FEMS Microbiol. Lett. 83 305-310 (1991)] ja Teather ja Wood [Appl. Environ. Microbiol. 43 777-800 (1982)].

Ksylanaasipositiiviset rekombinantti-kloonit voidaan sitten puhdistaa (toisin sanoen plakki voidaan muuttaa bakteeripesäkkeeksi) kauan käytössä olleilla menetelmillä. Sopivia menetelmiä löytyy julkaisusta Sambrook et al (1989) (loc. sit.), mutta käyttökelpoista olisi yksinkertaisesti noudattaa valmistajan ohjeita, olipa käytetty kitti sitten mikä tahansa, ja klooneissa oleva cDNA-insertti voidaan sitten leikata valittuun vektoriin, kuten pBLUESCRIPT$^{SK(-)}$:iin vain yhden esimerkin mainitaksemme. Muita sopivia plasmideja voidaan käyttää alakloonaukseen; esimerkkejä ovat pUC-plasmidit ja niistä johdetut plasmidit, kuten kuvaavat Sambrook et al., Molecular Cloning: A Laboratory Manual, 2. painos, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1989). Ekspressiovektoreita (erityisesti plasmideja), joissa ksylanaasia koodaava DNA on jonkin sopivan promoottorin kontrollissa, voidaan myös muodostaa ligatoimalla ja transformoida ja transfektoida johonkin sopivaan ekspressioisäntään. Esimerkkejä sopivista ekspressiovektoreista ovat pUC-sarja (jossa on *lacZp*-promoottori), pMTL-sarja (jossa myös on *lacZp*-promoottori) ja pBLUESCRIPT (jossa on sekä *lacZp*-promoottori että T_7-promoottori).

Promoottorin laadun ei yleensä uskota olevan erityisen kriittinen ja se riippuu ekspressio-isännästä ja olosuhteista, joissa ekspressiota halutaan. Kuten edellä mainittiin, eräs sopiva esimerkki bakteeri-ekspressioisännän, kuten E. colin, kohdalla on *lacZ*-promoottori. Vaihtoehtoisia promoottoreja bakteeri-isännille on bakteriofagi-T_7-promoottori.

Rekombinantti-ksylanaaseja ei ehkä ole tarpeen puhdistaa niiden ekspressioisännistä.

Vaikka E. coli isäntäsoluna saattaa olla sopiva, kun keksinnön mukaista ksylanaasia sovelletaan sellunvalmistukseen, on selvää, että muitakin isäntäsoluja voitaisiin käyttää, kuten gram-positiivisia bakteereita, mukaanlukien Bacillus subtilis, tai maitohappobakteereita. Vaihtoehtoisesti voidaan käyttää jotakin eukaryoottista ekspressioisäntää; eräs esimerkki olisi hiiva (kuten Saccharomyces cerevisiae).

Ksylanaaseja ekspressoivat isäntäsolut, jotka ovat edellä kuvatun kaltaisia ja/tai sisältävät edellä mainittuja DNA-sekvenssejä (joko ekspressiota varten tai muuten) muodostavat itse keksinnön vielä erään aspektin. Keksintö koskee myös isäntäsolujen valmistusmenetelmiä, joissa ksylanaasia koodaava DNA transformoidaan tai transfektoidaan johonkin soluun, ja ksylanaasin valmistusmenetelmiä, joissa ekspressioisäntiä viljellään ksylanaasia koodaavan DNA:n ilmentämistä varten.

Isäntäsolun laadusta riippuen saattaa keksinnön mukaisen rekombinantti-DNA:n kannalta olla suositeltavaa sisällyttää mukaan jokin signaalisekvenssi. Voidaan joko sisällyttää jokin isäntäspesifinen signaalisekvenssi tai eukaryooteissa tapahtuvaa ekspressiota varten käyttää entsyymin omaa signaalisekvenssiä. Voidaan käyttää jotakin translatoituvaa aloituskohtaa, joka on sovellettu ekspressioisäntään tai joka on sille suositeltavin; proteiinin oma translatoituva aloituskohta saattaa kuitenkin olla sopiva tai joissakin tapauksissa jopa suositeltavin.

Ekspressioisännästä saatu rekombinantti-ksylanaasientsyymi voidaan sitten karakterisoida. Tärkeimmät tunnusmerkit, joista on päästy varmuuteen keksinnön eräiden toteutusmuotojen kohdalla, ovat seuraavat:

(i) kloonatulla ksylanaasilla on erittäin suuri spesifinen aktiivisuus (5980 U/mg puhdistetun entsyymin proteenia); tämä on päinvastoin kuin monilla tähän mennessä raportoiduilla bakteereista johdetuilla ksylanaaseilla;
(ii) entsyymi pystyy hajoittamaan ksylaania epätavallisen suurella teholla, vapauttaen 0,9 g pelkistävää sokeria per g substraattia;
(iii) entsyymillä ei ole jäännösaktiivisuutta selluloosaan, kun sen sijaan monilla muilla ksylanaaseilla on jonkin verran aktiivisuutta selluloosaan; ja
(iv) entsyymi sisältää kaksi katalyyttistä domainia, joilla saattaa olla potentiaalia erittäin tehokkaan ksylanaasia tuottavan kloonin rakentamiseksi manipuloimalla geneettisesti edelleen ksylanaasi-cDNA:ta.

Täyspitkän kloonatun ksylanaasin (johon tästedes viitataan nimellä ksylanaasi A) erittäin spesifinen aktiivisuus (5980 U/mg puhdistetun entsyymin proteiinia) on tä-

män sieniksylanaasin sisäsyntyinen ominaisuus. *xyn*A-cDNA:n kysymyksessä olevan rakenteen ekspressiotaso pBluescript-vektorissa (pNX1) on suhteellisen matala E. colissa, 0,3 % E. coli-solujen syntetisoimasta liukoisesta proteiinista. Yleisesti ottaen kloonatun geenin ekspressiossa voidaan päästä tasolle >10 % E. colin kokonaissoluproteiinista.

*xyn*A-cDNA:n katkaistuja muotoja voidaan valmistaa käyttämällä rajoittavia entsyymejä. Jotkut katkaistut muodot, mukaanluettuna tunnuksella pNX5 nimetyssä plasmidissa oleva, tuottavat monta sataa kertaa voimakkaamman aktiivisuuden kuin pNX1. Eräs selitys tälle havainnolle on, että se on tulosta LacZ translaation aloitussekvenssin käytöstä katkaistun ksylanaasi A:n synteesissä. Eräs toinen selitys on, että vältettäessä AT-rikkaita alueita tuloksena saattavat olla korkeammat ekspressiotasot; eräs teoria on, että Rnaasi E:n mRNA:ta degradoiva aktiivisuus on nopeutta rajoittava vaihe proteiinisynteesissä ja että Rnaasi suosii mRNA:n AT-rikkaita alueita. On mahdollista edelleen nostaa sen ekspressiotasoa E. colissa käyttämällä jotakin voimakkaampaa promoottoria, kuten bakteriofagi-T_7-promoottoria.

Escerichia colista, joka antoi suojan *xyn*A:lle, puhdistetun rekombinantti-ksylanaasi A:n (XYLA) M_r oli 53 000 ja se hydrolysoi spelttikauran ??? ksylaania ksylobioosiksi ja ksyloosiksi. Entsyymi ei hydrolysoinut mitään selluloosasubstraateja. *xyn*A:n nukleotidisekvenssi paljasti yhden avoimen 1821 bp:n lukukehyksen, joka koodasi proteiinia, jonka M_r oli 66 192. XYLAn ennakoitu primaarirakenne muodostui N-pään signaalipeptidistä, jota seurasi 225 aminohapon toistuva sekvenssi, joka erotti peräkkäisestä 40 tähteen C-pään toistuvasta jaksosta treoniini/proliinikytkytsekvenssi. Laajat N-pään toistuvat alueet muodostuivat erillisistä katalyyttisistä domaineista, joilla oli samanlaisia substraattispesifisyyksiä täyspitkään entsyymiin nähden.

Keksinnön mukaisilla ksylanaaseilla on monia sovelluksia elintarvike-, rehu- ja puunjalostus- ja paperiteollisuudessa. Tässä kuvattujen ksylanaasien käyttö näillä teollisuudenaloilla sisältyy keksinnön suoja-alaan.

Käsittelemme ensiksi elintarviketeollisuutta, jossa taikinan ja siitä valmistettujen leivottujen tuotteiden tietyt ominaisuudet riippuvat käytetyn jauhon pentosaani- ja tärkkelyssisällöstä. Näitä ominaisuuksia ovat leivän tekstuuri, tilavuus ja vanhentuminen. Ksylaania käyttämällä voitaisiin leivontatuotteita muuntaa, jolloin saataisiin tavaroita, joilla oli potentiaalista kaupallista arvoja. Ominaisuuksia, joita ksylanaasikäsittelyllä voidaan muuntaa, on leivän ominaistilavuus. Ominaistilavuuden lisääntymistä parantaa vielä, kun lisätään amylaasia yhdessä ksylanaasin kanssa. Eräs tähän vaikut-

tavista tekijöistä on hiilihydraattien vedensidontakyky. Keksintö tuo käyttöön taikinan, joka sisältää tässä kuvattua ksylanaasia.

Eläinrehuteollisuudessa entsyymilisäyksen käyttö kananpoikien rehun parantamiseksi raportoitiin jo 1957. Tuoreemmat tulokset viittaavat siihen, että eräissä viljalajeissa, kuten vehnässä ja erityisesti rukiissa, juuri pentosaanit, joita on sen siemenvalkuaisessa, ovat pääasiassa syynä ravinnon vähäiseen imeytymiseen ja tahmeisiin ulosteisiin kananpojilla. Molemmat ongelmat näyttävät johtuvan sulamattomien pentosaanien suuresta viskositeetista. Tämä estää ravinteiden diffusoitumista ja sitoo vettä ja tekee eritteet vetisiksi. Ongelmia voidaan lievittää käyttämällä ksylanaasivalmisteita. Ksylanaasin vaikutus saattaa parantaa sekä kananpoikien painonnousua että niiden rehun hyötysuhdetta. Näyttäisi, että ksylanaasilisää voitaisiin käyttää parantamaan rukiin ravintoarvoa tämän viljalajin käytön lisäämiseksi kananpoikien rehussa. Tämän käsittelyn tehokkuus voi riippua ruislajikkeesta. Keksintö koskee ksylanaasin käyttöä kananpoikien rehussa ja näihin tarkoitetuissa viljoissa.

Selluteollisuudessa liukenevat massat ovat puhdistettuja selluloosia, joita käytetään viskoosiraionien, selluloosaesterien ja selluloosaeetterien valmistukseen. Ne on johdettu esihydrolysoiduista sulfaattimassoista tai happosulfaattimassoista. Niiden prosessoinnille on tunnusomaista selluloosan derivatisointi yhdessä vaiheessa, johdannaisen ollessa tavallisiin liuotteisiin liukenevaa, jolloin siitä voidaan muodostaa kuituja, kalvoja ja muoveja. Selluloosassa olevat epäpuhtaudet haittaavat sen derivatisointia ja johtavat siten liukenemattomiin aineksiin, jotka tukkivat ruiskuttimien aukot tai aikaansaavat vikoja lopputuotteeseen. Lisäksi eräät ksylaaniepäpuhtaudet voivat johtaa värin, mittojen ja lämmön pysymättömyyteen asetaattituotteissa. Ksylanaasilla saattaa siten olla merkitystä epäpuhtauksien poistamisessa ja tässä kuvattujen ksylanaasien käyttö tähän tarkoitukseen sisältyy keksintöön.

Kraftmassan esivalkaiseminen sellulaasitonta ksylanaasia käyttäen on identifioitu yhdeksi biomenetelmistä, jotka todennäköisimmin hyväksytään sellu- ja paperiteollisuudessa lähitulevaisuudessa, mutta vain mikäli käyttöön tulee sopivia ksylanaaseja. Kraft-menetelmästä (tunnetaan myös alkali- tai sulfaattimenetelmänä) on tullut hallitseva sellunvalmistusmenetelmä Kanadassa, koska se tuottaa vahvoja puukuituja ja koska käytetyt kemikaalit otetaan talteen ja kierrätetään. Kraftmassoja, erityisesti havupuista johdettuja, on suhteellisen vaikea valkaista. Perinteisesti tarvitaan sarja vaiheita, joissa käytetään alkuainekloori a ja klooria sisältäviä yhdisteitä, näiden massojen valkaisemiseksi tehokkaasti haluttuun täyteen ~90 %:n valkoisuuteen. Valkaisuprosessissa syntyy, erityisesti alkuainekloori a käytettäessä, orgaanisia klooriyhdistei-

tä, jotka on perinteisesti johdettu valkaisulaitoksesta ulos jäteveden mukana. Sekä kansalaisten vaatimukset että lainsäädäntö painostavat nykyään sellutehtaita vähentämään näiden saasteiden päästöjä tai lopettamaan ne kokonaan. Sellu- ja paperiteollisuus harkitsee eri vaihtoehtoisten tekniikkojen käyttöönottoa vähentääkseen tehtaidensa ympäristökuormitusta. Näihin optioihin kuuluu kraftmassan esivalkaisu ksylanaasin avulla. Esillä olevan keksinnön mukaiset ksylanaasit sopivat erityisen hyvin tähän tarkoitukseen.

Esillä olevan keksinnön mukaisten ksylanaasien uskotaan sopivan erityisen hyvin paperi- ja selluteollisuuteen. Vaikka tiedetään, että entsyymien käyttö ei koskaan korvaa kemikaaleja kokonaan, ympäristöviranomaiset painostavat vähentämään kemikaalien käyttöä. On myös olemassa käytännön syitä kemikaalien vähentämiseksi paperi- ja selluteollisuudessa.

Sellutehtaat kehittävät yleensä omat kloorin ja klooridioksidin lähteensä paikan päällä ja tämä saattaa rajoittaa kapasiteettia sekä mahdollisesti olla vaarallista. Paperimassan (esim. kraftmassan) käsitteleminen ligniinin poistamiseksi vaatii kloorin, NaOH:n, H_2O_2:n ja klooridioksidin käyttöä. Sandoz USA:ssa on suorittanut käytännön kokeita käyttäen omaa CARTAZYME-tuotettaan, joka on sieniperustainen ksylanaasi (puhdistamaton), joka on aktiivinen 30 - 55 °C:ssa, pH 3 - 5, ja joka sisältää kahta ksylanaasia, ja on todennut, että 25 - 33 %:n vähennys kloorin käytössä on mahdollista käyttämällä 1U ksylanaasia/g massaa. Myös tuote on valkoisempi kuin mitä pelkästään kemikaaleja käytettäessä. Toinen ksylanaasin etu on, että se on spesifinen, kun taas kemikaalit voivat syövyttää vähäligniinistä selluloosaa, johtaen kuituvahvuuden heikkenemiseen ja muihin ei toivottuihin fysikaalisiin ominaisuuksiin. On sen vuoksi selvää, että ksylaaseista saattaisi tulla tärkeämpiä massanvalkaisussa ja erityisesti niiden rekombinanteista niiden spesifisyyden ja suuren saannon vuoksi. Ligniinin uskotaan sitoutuvan hemiselluloosaan ja mikäli hemiselluloosa (ksylaani) depolymeroidaan, ligniini voidaan osittain erottaa selluloosasta ja sitten pestä pois. Tällä hetkellä jonkinasteinen kemiallinen käsittely saattaa kuitenkin olla yhä tarpeen. Pääkohdat esillä olevan keksinnön mukaisessa ksylanaasissa kaupallisen käytön kannalta katsottuina ovat: (i) sen erittäin suuri spesifinen aktiivisuus ja sen voimakas ilmentymistaso tekisivät taloudelliseksi sen tuottamisen suuressa mittakaavassa ja (ii) sellulaasiaktiivisuuden puuttuminen siitä tekee sen erityisen käyttökelpoiseksi silloin, kun tarvitaan ksylaanin poistamista spesifisesti kuten paperinvalmistuksen ja tekstiiliteollisuuden sovelluksissa.

Uskotaan myös, että eräs keksinnön mukaisen ksylanaasin sovellus voisi olla sokeri-

teollisuudessa ja sokeriruokojätteen tai muiden ksylaania sisältävien tuotteiden käsittelyn yhteydessä jätteenkäsittelyn tehostamiseksi.

Aikaisemmin mainittiin, että XYLAn proteiinisekvenssi ja *xyn*A:n DNA-sekvenssi tulivat saataville toukokuun 5. päivänä 1992 EMBL-tietopankissa saantinumerolla X65526. Tämä saatavuus ei muodostane varsinaista uutuudenestettä kaikkien tässä hakemuksessa mainittujen valtioiden lainsäädännössä. Niiden lainsäädäntöjen valta, joissa EMBL-tietopankin tallennus ei muodosta varsinaista uutuudenestettä, huomautetaan täten, että keksintö on ja tullaan määrittelemään laajemmin kuin edellä on mainittu. Erityisesti keksinnön voidaan siten katsoa olevan seuraavissa lisäaspekteissa.:

– ksylanaasi, jossa on vähintään yksi katalyyttinen domaini, joka on olennaisesti homologinen jonkin anaerobisen sienen ksylanaasin kanssa; ksylanaasi voi olla jonkin anaerobisen sienen täyspitkä luonnon ksylanaasi; ja
– eristetty tai rekombinantti-DNA-molekyyli, joka koodaa ksylanaasia, jossa on katalyyttinen domaini, joka on olennaisesti homologinen jonkin anaerobisen sienen ksylanaasin kanssa, edellyttäen että jos DNA-molekyyli on cDNA-molekyyli, joka koodaa Neocallimastix frontaliksen ksylanaasia, DNA-molekyyli on toiminnallisesti kytketty johonkin promoottoriin; DNA-molekyyli voi sisältää ksylanaasia koodaavan luonnon mRNA:n täyspitkän kopion.

Edellä olevasta lienee selvää, että keksinnön suoja-alaan ei sisälly ainoastaan edellä kuvattu rekombinantti-ksylanaasi, vaan myös muista kuin edellä mainituista anaerobisista sienistä johdettuja ksylanaaseja, jotka voidaan valmistaa tässä kuvatuilla menetelmillä. Keksinnön suoja-alaan sisältyy myös mikä tahansa mutantti, joka on johdettu N. patriciarumista tai N. patriciarumista valikoimalla tai geeninsiirrolla johdetuista kannoista.

Keksinnön suoja-alaan kuuluvat myös:
(i) DNA-sekvenssit, jotka on johdettu pNX1:stä, pNX4:stä, pNX5:stä, pNX6:sta, pNX8:sta, pNX9:stä ja pNX10:stä ja DNA sekvensseistä, jotka hybridisoituvat siihen;
(ii) DNA-rakenne, joka sisältää kohdan (i) mukaisen DNA-sekvenssin, joka on toiminnallisesti kytketty sääteleviin alueisiin, jotka pystyvät ohjaamaan ksylanaasiaktiivisuutta omaavan polypeptidin ekspressiota tai yliekspressiota jossakin sopivassa ekspressioisännässä;
(iii) transformoitu mikrobi-isäntä, joka kykenee ekspressoimaan tai yliekspressoi-

maan sieniksylanaasia, joka sisältää kohdan (ii) mukaisen ekspressiorakenteen;
(iv) polypeptidi, jolla on ksylanaasiaktiivisuus ja joka on tuotettu ekspression avulla käyttämällä jotakin kohdan (iii) mukaista mikrobi-isäntää;
(v) kuvion 4 mukainen aminohapposekvenssi, joka sisältää komponentit A, B, C ja D, ja aminohapposekvenssit, jotka on johdettu tästä ksylanaasista; ja
(vi) kuviossa 1 kuvatut plasmidit.

Keksinnön suoja-alaan sisältyy myös menetelmä kuviossa 1 kuvatut rekombinanttiplasmidit käsittävästä E. colista johdetun ksylanaasin valmistamiseksi.

Jokainen suositeltavin tunnusmerkki, jota edellä kuvattiin keksinnön jonkin aspektin yhteydessä, on samalla tavoin suositeltavin, mutatis mutandius, jokaisen muunkin aspektin kohdalla.

Keksintöä kuvataan nyt seuraavilla esimerkeillä. Esimerkeissä viitataan liitteenä oleviin piirustuksiin, joissa:

kuvio 1 on *xyn*A:n sisältävien rekombinantti-plasmidien restriktiokartta. Kuviossa näkyvät rajaavien endonukleaasien *Eco*RI (R), *Sst*I (S), *Sca*I (Sc), *Hpa*I (Hp), *Kpn*I (K), *Xho*I (X), *Sma*I (Sm), *Pvu*II (Pv), *Nae*I (Na), *Nru*I (Nr), *Stu*I (St) ja *Hin*dIII (H) katkaisukohtien paikat. Vektoreiden multippelikloonausalueiden katkaisukohdat sulkeissa on tuhottu. Vektoreiden multippelikloonausalueet, jotka on merkitty tunnuksella *, on johdettu pSK(S):stä, pMTL20(20):stä ja pMTL22(2):sta, vastaavassa järjestyksessä. Nuolella varustettu yhtenäinen viiva osoittaa *xyn*A:n avoimen lukukehyksen pituuden ja suunnan. *xyn*A:n deleetiomutanttien rakenne esitään yksityiskohtaisesti tuonnempana. Kuviossa näkyvät rekombinantti-plasmidit käsittävien E. coli -kantojen fenotyypit.
Kuvio 2A ja 2B esittävät XYLA:n puhdistamista. pNX1:n (A) tai pNX5:n (B) käsittävästä soluttomasta E. coli XL1-Blue -ekstraatista puhdistetun XYLA:n SDS/PAGE. Kaista 1 sisälsi anioninvaihtokromatografialla puhdistetun XYLA:n, kaista 2 sisälsi pNX1:n tai pNX5:n käsittävän soluttoman E. coli -ekstraatin ja kaista 3 (vain B) sisälsi soluttoman ekstraktin pBluescript SK:n sisältävästä E. colista. Kuvioissa A ja B näkyvät geelit sisälsivät 10 % (w/v) tai 15 % (w/v) polyakryyliamidia, vastaavassa järjestyksessä. Proteiinien koot on esitetty kD:ina, vähennettyinä markkeriproteiineista, jotka ovat moolimassaltaan suuria (kuvio 2A) tai pieniä (kuvio 2B) markkereita Sigmalta.
Kuvio 3 kuvaa puhdistetun XYLA:n vaikutusta liukoisen ksylaanin (0,5 %) ominaisviskositeettiin PC-puskurissa, pH 6,5, 37 °C:ssa. Ominaisviskositeetti (■) ja

pelkistävät sokerit (●) mitattiin tuonnempana kuvatulla tavalla.

Kuvio 4 esittää XYLA:n primaarirakennetta. Kaksi homologista katalyyttistä domainia, jotka on merkitty A:lla ja B:llä, sekä kahdentuneet C-terminaaliset sekvenssit (C ja D) on laatikoitu.

Kuvio 5 esittää N. patriciarum XYLA:n ja prokaryoottiksylanaasien homologisten alueiden kohdakkaisuutta. Vertaillut entsyymit olivat seuraavat: B. pumilus -ksylanaasi A [XYLAB; Fukusaki et al., FEBS Lett. 171: 197-201 (1984)], B. circulans -ksylanaasi (XYLBC; Yang et al., Nucl. Acids Res. 16: 7178 (1988)] ja C. acetobutylicum -ksylanaasi B [XYLBCA; Zappe et al., Nucl. Acids Res.18 2179 (1990)]. Tähteet, joissa on samanlaisuutta tai samankaltaisuutta kaikissa vertailluissa primaarisekvensseissä, on laatikoitu. Kuviossa on osoitettu homologisten alueiden ensimmäisten ja viimeisten tähteiden asemat vastaavissa primaarisekvensseissä.

Kuvio 6 esittää plasmidi pNX1:n rakennetta.

Kuvio 7 esittää Neocallimastix patriciarum -ksylanaasi A:ta koodaavan cDNA:n kloonausta ja karakterisointia.

Esimerkki 1 - pNX1:n valmistus

1.1. Mikrobikannat, vektorit ja viljelyalusta

Anaerobinen sieni Neocallimastix particiarum (lajityyppi) eristettiin lampaan pötsistä menetelmällä Orpin, C.G. ja Munn, E.A., Trans. Br. Mycol. Soc. 86: 178-181 (1986). Isäntäkannat cDNA-kloonausta varten olivat E. coli PLK-F' ja XL1-Blue. E. coli -kantaa JM83 käytettiin ksylanaasi$^+$ cDNA-kloonien karakterisointiin.

Vektorit olivat λZAPII, pBluescript$^{SK(-)}$ (Stratagene), pMTL20, pMTL22 ja pMTL23 [Chambers et al., Gene 68: 139-149 (1988)]. N. patriciarum-viljelmää pidettiin viljelyalustassa, joka sisälsi 10 % pötsinestettä, kuten kuvaavat Kemp et al., J. Gen. Microbiol. 130: 27-37 (1984)]. E. coli -kantoja kasvatettiin L-lihaliemessä [Sambrook et al., Molecular Cloning: A Laboratory Manual, 2. painos, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1989)]. Rekombinantti-faagia kasvatettiin E. coli -kannoissa käyttäen NZY-viljelyalustaa Stratagenen ohjeiden mukaisesti.

1.2 Yleiset rekombinantti-DNA-menetelmät

Agaroosi-geelielektroforeesi, E. colin transformaatio ja DNA:n modifiointi käyttäen rajaavia entsyymejä ja T4 DNA -ligaasia olivat kuten kuvaavat Gilbert et al., J. Gen. Microbiol. 134 3239-3247 (1988). Suuria määriä plasmidi-DNA:ta uutettiin E. colista Brij'n lyyseillä ja sen jälkeen CsCl-tiheysgradienttisentrifugoinnilla [Clewell, D.B. ja Helinski, D.R., Proc. Natl. Acad. Sci. USA 62: 1159-1166 (1969)]. Nopeaa keit-

tomenetelmää Holmes, D.S. ja Quigley, M., Anal. Biochem. 114: 193-197 (1981) ja emäksistä hajoitusmenetelmää, Birnboim, H.L. ja Doly, J., Nucl. Acids Res. 7: 1513-1523 (1979), käytettiin plasmidin eristämiseksi nopeaa restriktioanalyysiä ja nukleotidisekvensointia varten, vastaavassa järjestyksessä. Northern-hybridisaatio suoritettiin kuten kuvaavat Gilbert et al., J. Bacteriol. 161: 314-320 (1985).

1.3. Pötsin anaerobisen sienen, N. patriciarumin, viljely

N. patriciarumia kasvatettiin pötsinestettä sisältävässä kasvualustassa [Kemp et al., J. Gen. Microbiol. 130: 27-37 (1984)] 1 % AVICELin ollessa läsnä 39 °C:ssa ja anaerobisissa olosuhteissa 48 tuntia. [Vaihtoehtoisia viljelyalustoja, kuten esimerkiksi tutkijoiden Philips, M.W. ja Gordon, G.L.R., Appln. Environ. Microbiol. 55: 1695-1702 (1989) ja Lowe et al., J. Gen. Microbiol. 131: 2225-2229 (1985) kuvaamia voidaan käyttää].

1.4. Kokonais-RNA:n eristäminen

Pakastettuja myseelejä jauhettiin hienojakoiseksi jauheeksi nestemäisen typen alla huhmaressa. 5 - 10 tilavuusosaa guanidiniumtiosyanaattiliuosta (4M guanidiniumtiosyanaattia, 0,5 % natriumlauryylisarkosiinia, 25 mM natriumsitraattia, pH 7,0, 1 mM EDTA:a ja 0,1 M ß-merkaptoetanolia) lisättiin pakastettuun sienirihmastojauheeseen ja seosta homogenisoitiin 5 minuuttia huhmaressa ja vielä 2 minuuttia täydellä nopeudella Polytron-homogenisaattoria käyttäen. Kokonais-RNA eristettiin homogenaatista ultrasentrifugoimalla CsCl-tyynyn läpi [Sambrook et al., Molecular Cloning: A Laboratory Manual, 2. painos, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1989)]. [Voidaan käyttää vaihtoehtoista menetelmää kokonaissieni-RNA:n valmistamiseksi, kuten sovellusta menettelytavasta, jota kuvaavat Puissant, C. ja Houdebine, L.M., Bio-Techniques 148-149 (1990)].

1.5. Poly A^+ mRNA:n puhdistus

Poly A^+ mRNA puhdistettiin kokonais-RNA:sta oligo(dT)-selluloosakromatografialla [Sambrook et al., Molecular Cloning: A Laboratory Manual, 2. painos, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1989)].

1.6. N. patriciarumin cDNA-ekspressiokirjaston rakentaminen

cDNA-kirjasto rakennettiin käyttäen Stratagenen λZAP cDNA synteesikittiä, periaatteessa valmistajan ohjeiden mukaisesti.

Menettelytapaa kuvataan lyhyesti seuraavasti: poly A^+ mRNA muutettiin cDNA:n ensimmäiseksi nauhaksi käänteistranskriptaasilla käyttäen *Xho*I-linkkeriä - oligo-

(dT)-primeria ja 5-metyyli dCTP:tä. Kaksinauhainen cDNA syntetisoitiin yksinauhaisesta cDNA:sta Rnaasi H:n ja DNA-polymeraasi I:n avulla. Kun cDNA:n päät oli katkaistu, cDNA sidottiin *Eco*RI-adaptorilla, fosforyloitiin ja digestoitiin *Xho*I:llä, jolloin saatiin cDNA, jossa oli *Eco*RI-kohta 5'-alueella ja *Xho*I-kohta 3'-alueella. cDNA fraktioitiin määrätyn kokoisiksi palasiksi 1 % alhaisen sulamispisteen agaroosigeelielektroforeesilla ja cDNA:n 1,2 - 8 kb:n kokoisia palasia otettiin talteen fenoliuutolla [Sambrook et al., Molecular Cloning: A Laboratory Manual, 2. painos, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1989)]. Sitten cDNA-palaset sidottiin *Eco*RI/*Xho*I-digestoituun λZAPII-vektoriin (muitakin ekspressiovektoreita voidaan käyttää).

cDNA-kirjasto pakattiin in vitro ja monistettiin käyttäen E. coli PLK-F':ä maljausso-luina.

1.7. Ksylanaasi-positiivisten rekombinantti-bakteriofagikloonien seulonta
Rekombinantti-faagia kasvatettiin E. coli XL1-Blue:ssa 0,7 % pinta-agarissa, joka sisälsi 0,1 % ksylaania ja 10 mM isopropyyli-ß-tiogalaktopyranosidia (IPTG, indusori LacZ-promoottorin säätelemää geenieksprossiota varten). Yön yli 37 °C:ssa inkuboinnin jälkeen lisättiin 0,5 % kongonpunaliuosta pinta-agarin päälle. Huoneenlämpötilassa 15 minuutin ajan inkuboimisen jälkeen sitoutumaton väriaine poistettiin pesemällä 1M NaCl:lla. Ksylanaasia tuottavia faagiplakkeja ympäröivät keltaiset kehät punaisella taustalla.

Ksylanaasi-positiiviset rekombinantti-faagit puhdistettiin homogeenisiksi maljaamalla ne uudelleen ja seulomalla faagi uudelleen kuten edellä 2 - 3 kertaa.

cDNA-insertit ksylanaasi-positiivisessa faagissa leikattiin pBluescript SK':hon käyttäen VCS-M13-auttajafaagia.

1.8. Ksylanaasi- ja sitä läheisesti muistuttavien entsyymien kokeet
Ksylanaasi-positiivisia rekombinantti-plasmideja sisältävästä E. colista uutetut entsyymit valmistettiin kuten kuvaavat Kellett et al., Biochem. J. 272: 369-376 (1990).

Entsyymejä kokeiltiin ksylaanin tai muiden substraattien hydrolyysin osalta 37 °C:ssa 50 mM potassiumfosfaatti/12 mM sitruunahappopuskurissa, pH 6,5, ja ksylaanista tai muista kasvipolysakkarideista (karboksimetyyliselluloosa, ohran ß-glukaani, laminariini, likenaani) vapautuneet pelkistävät sokerit mitattiin kuten kuvaavat Kellett et al., Biochem. J. 272: 369-376 (1990) ja Hazlewood et al., J. Gen. Microbiol. 136:

2089-2097 (1990). Aktiivisuuskokeita keinotekoisia substraatteja (metyyliumbelliferyyli-ß-D-sellobiosidi, metyyliumbelliferyyli-ß-D-glukosidi, metyyliumbelliferyyli-ß-D-ksylosidi ja p-nitrofenyyli-ß-D-ksylosidi) ovat kuvanneet Hazlewood et al., J. Gen. Microbiol. 136: 2089-2097 (1990).

1.9. DNA-sekvensointi

Alkalilla denaturoitu plasmidi-DNA neutraloitiin ja puhdistettiin edelleen spin-dialyysillä [Murphy, G. ja Kavanagh, T., Nucl. Acid Res. 16: 5198 (1988)]. Saadun DNA:n sekvensointi perustui protokollaan, jota Sequenase DNA -sekvensointikitin valmistaja (USA, Cleveland, Ohio) suosittelee. Päällekkäin meneviä sekvenssejä kehitettiin kloonaamalla sopivia resptriktiopalasia pMTL-perustaisiin vektoreihin. Sekvenssit pinottiin ja järjestettiin käyttämällä tietokone-ohjelmia, joita kuvaa Staden, R., Nucl. Acids Res. 16: 3673-3694 (1980). Plasmidin, jolla oli tunnus pNX1, sisältämä cDNA:n täydellinen sekvenssi määritettiin molemmista nauhoista. Plasmidin sisältämä ksylanaasia koodaava geeni nimettiin *xyn*A:ksi ja geenituote, itse ksylanaasientsyymi, nimettiin XYLA:ksi.

Esimerkki 2 - pNX4:n, pNX1:n (*xyn*A) erään deleetiomutantin, rakentaminen

pNX1 linearisoitiin *Xho*I:llä ja *xyn*A cDNA:n 3'-alue poistettiin Bal-31 -digestion avulla [Hall, J. ja Gilbert, H.J., Mol. Gen. Genet. 213: 112-117 (1988)]. Päiden katkaisun jälkeen katkaistu cDNA leikattiin pNX1:stä *Eco*RI-digestion avulla ja kloonattiin *Eco*RI/*Sma*I-leikattuun pMTL22-vektoriin.

Esimerkki 3 - pNX5:n, pNX1:n (*xyn*A) erään deleetiomutantin, rakentaminen

720 bp:n *Sca*U/*Nru*I-palanen leikattiin pNX4:stä ja kloonattiin pMTL20-vektoriin. Näin saatiin erittäin ekspressiivinen klooni, jossa entsyymin ekspressiotasot olivat muutama satakertaisia pNX1:een verrattuna.

Esimerkki 4 - pNX6:n, pNX1:n (*xyn*A) erään deleetiomutantin, rakentaminen

pNX6 rakennettiin pilkkomalla pNX1 *Eco*RI/*Sca*I:llä ja kloonaamalla saatu palanen *Eco*RI/*Sma*I:llä leikattuun pMTL22:een.

Esimerkki 5 - pNX8:n, pNX1:n (*xyn*A) erään deleetiomutantin, rakentaminen

pNX1 leikattiin *Sca*I:llä ja *Xho*I:llä, jolloin saatiin 1,3 kb:n palanen, joka kloonattiin pMTL20:een niin, että *Xyn*A-sekvenssi oli faasissa vektorin sisältämän *LacZ* ATG:n kanssa. Näin saatiin erittäin ekspressiivinen klooni, jossa ekspressiotaso oli suunnilleen viisitoistakertainen pNX1:een verrattuna.

Esimerkki 6 - pNX9:n, pNX1:n (*xyn*A) erään deleetiomutantin, rakentaminen

pNX8 leikattiin *Kpn*I (1 kohta vektoripolylinkerissä) ja inserttifragmentti leikattiin sähköeluution jälkeen *Rsa*I:llä (leikkauskohdat geenin PT-linkkerialueella), jolloin saatiin -700 kb:n palanen, joka kloonattiin pMTL20:een, joka oli pilkottu *Kpn*I:llä ja *Stu*I:llä. Tällä tavoin saatiin erittäin ekspressiivinen klooni (paljon parempi kuin pNX8:n sisältävä klooni), jossa oli toinen katalyyttinen alue kehyksessä vektorin *LacZ* N-pään kanssa.

Esimerkki 7 - pNX4:n, pNX1:n (*xyn*A) erään deleetiomutantin, rakentaminen

pNX8 leikattiin *Kpn*I:llä ja palanen (-850 bp) sidottiin *Kpn*I:llä leikattuun pMTL20:een. Tämäkin klooni ekspressoi hyvin, mutta ekspressoitunut proteiini sisältää karboksipäässä joitakin tähteitä, jotka poistamalla saadaan korkeatasoinen ekspressio, joka todettiin pNX9:llä.

Esimerkki 8 - Ksylanaasi A:n N-pään puhdistaminen ja aminohapposekvensointi

pNX1:n tai pNX5:n sisältävää E. coli XL1-Blueta viljeltiin 16 tuntia LB-lihaliemessä, joka sisälsi ampisilliinia (100 µg/ml). Soluja, jotka korjattiin sentrifugoimalla, lietettiin uudelleen 50 mM Tris/HCl-puskuriin, pH 8,0, ja sytoplasminen fraktio valmistettiin edellä kuvatulla tavalla [Clarke et al., FEMS Microbiol. Lett. 83: 305-310 (1991)]. Ksylanaasi, joka saostettiin lisäämällä ammoniumsulfaattia (0,39 g/ml), liuotettiin uudelleen 10 mM Tris/HCl-puskuriin, pH 8,0. Dialyysin jälkeen saman puskurin 3 vaihtoa vastaan ksylanaasi puhdistettiin olennaisesti anioninvaihtokromatografialla DEAE-Trisacryl M:llä olennaisesti Poolen et al., Mol. Gen. Genet. 223: 217-223 (1990) kuvaamalla tavalla.

Ksylanaasi (nimettiin XYLA:ksi), joka puhdistettiin pNX1:n käsittävästä E. coli XL1-Bluen soluvapasta ekstraktista, fraktioitiin SDS/PAGE:lla ja elektroblotattiin PROBLOT"-kalvolle (Applied Biosystems Inc). N-pään sekvenssi määritettiin automaattisella Edman-sekvensoinnilla käyttäen 470 kaasufaasisekvenaattoria, joka oli varustettu 120 on-line fenyylitiohydantoiini-analysaattorilla [Applied Biosystems Inc: Hunkapillar et al., Methods Enzymol. 91: 399-413 (1983)].

Esimerkki 9 - Yhteenveto *xyn*A:n eristämisestä

cDNA-kirjasto, joka muodostui 10^6 kloonista, rakennettiin käyttämällä mRNA:ta, joka oli eristetty N. patriciarum -soluista, joita oli kasvatettu AVICEL ainoana hiilen lähteeenä. Kolmekymmentäyksi rekombinantti-bakteriofagia, jotka hydrolysoivat ksylaania, identifioitiin seulomalla 5 x 10^4 kloonia kirjastosta, ja 16 voimakkaasti ksylanaasi-positiivista faagia eristettiin lisäkarakterisointia varten. Restriktiokartoitus

ja hybridisaaatiotulokset osoittivat, että kaikki ksylanaasi-positiiviset rekombinantit sisälsivät cDNA-sekvenssejä, jotka oli johdettu samasta mRNA-lajista. Suurimman funktionaalista ksylanaasia, joka nimettiin *xyn*A:ksi, koodaavan cDNA-sekvenssin restriktiokartta on esitetty kuviossa 1. Nukleiinihappokoetin, joka muodostui 1*xyn*A:n 5'-alueen 7 kb:stä, hybridisoitui yhteen 2,5 kb:n Neocallimastix-RNA-lajiin. Tämä osoittaa, että pisin eristetty *xyn*A-DNA on lähes täyspitkä.

Esimerkki 10 - Ksylanaasi A:n karakterisointi

Neocallimastix-ksylanaaseja koodaavat cDNA-sekvenssit leikattiin λZAPII:stä ja otettiin talteen E. coli XL1-Blue'hin pBLUESCRIPT SK rekombinantteina. Ksylanaasiaktiivisuus, jonka plasmidi pNX1:n käsittämä rekombinanttikanta, joka sisälsi xynA:n pisimmän muodon, ekspressoi, todettiin hallitsevasti soluttomassa ekstraktissa, mikä osoitti, että E. coli ei erittänyt tehokkaasti entsyymiä. Ksylanaasi, joka nimettiin A:ksi (XYLA), puhdistettiin lähelle homogeenisuutta (>90 %:n puhtaus). Puhdistetun XYLA:n spesifinen aktiivisuus oli 5980 U/mg proteiinia, verrattuna soluttoman ekstraktin arvoon 16 U/mg proteiinia. Tämä osoittaa, että XYLA muodostuu 0,3 %:sta liukoista proteiinia, jonka pNX1:n käsittävät E. coli -solut syntetisoivat. Puhdistetun entsyymin M_r oli 53 000 (kuvio 2) ja sen eräs N-terminaalinen sekvenssi oli IATVAKAQWGGGGAS. XYLA hydrolysoi ksylaania, mutta ei osoittanut minkäänlaista aktiivisuutta karboksimetyyliselluloosaa, ohran ß-glukaania, likenaania tai keinotekoisia substraatteja 4-metyyli-umbelliferyyli-ß-D-ksylosidia ja p-nitrofenyyli-ß-D-ksylanopyranosidia kohtaan (taulukko 1).

Taulukko 1

pNX1 (*xyn*A-cDNA) plasmidin käsittämästä E. colista puhdistetun ksylanaasi A:n entsyymiaktiivisuus

Substraatti	Aktiivisuus[1] Yksikköä/mg proteiinia
Ohran ß-glukaani	0
Karboksimetyyliselluloosa	0
Ksylaani	5980
Ksylobioosi	0
p-Nitrofenyyli-ß-D-ksylobiosidi (PNX)	0
Metyyliumbelliferyyli-ß-D-sellobiosidi (MUC)	0
4-Metyylibelliferyyli-ß-D-glukosidi (MUG)	0
4-Metyylibelliferyyli-ß-D-ksylosidi (MUX)	0

[1] 1 yksikkö XYLA:a vapauttaa 1 µmoolia tuotetta/minuutti.

Entsyymi hajoitti liukoista ksylaania tavalla, joka on tyypillinen endo-ß-1,4-ksylanaasille (EC 3.1.2.8) edistäen viskositeetin nopeaa alenemista (kuvio 3) ja vapauttaen 893 mg pelkistävää sokeria per gramma substraattia. Hydrolyysituotteiden HPLC-analyysi osoitti, että XYLA vapautti suunnilleen samanlaisia määriä ksylobioosia ja ksyloosia. Reaktiotuotteiden joukosta ei löytynyt disakkarideja sisältävää arabinoosia, kauraspelttiksylaanin tärkeintä sivuketjusokeria, mikä osoittaa, että entsyymi ei hydrolysoi glykosidisidoksia, joissa on ksyloosiyksikköjä, jotka ovat sitoutuneet sivuketjusokereihin.

Esimerkki 11 - Nukleotidisekvenssi

pNX1:stä johdettu 2,3 kb:n Neocallimastix-cDNA sekvensoitiin molemmista nau-

hoista (saantinumero X65526 EMBL/Genbank/DDBJ Nucleotide Sequence Data Libraries). Nukleotidisekvenssin translaatio paljasti yhden yksinkertaisen 1821 kb:n lukukehyksen (ORF), joka koodasi M_r66192 polypeptidiä. Koodatun proteiinin johdettu primaarirakenne on esitetty kuviossa 4. Rekombinantti-XYLA:n, joka oli puhdistettu E. colista, 15 N-terminaalista tähdettä sopivat täydellisesti translatoidun sekvenssin aminohappoihin 12 - 26. Ehdotetun translaation aloituskodonin osoittaminen perustui seuraaviin huomioihin: (i) ORF:n ylävirrassa ei ole ATG-sekvenssejä; (ii) translatoituvia lopetuskodoneja on oletetun translatoituvan aloituskodonin kaikissa kolmessa ylävirran lukukehyksessä. Nukleotidisekvenssin tarkistus oletetun ATG-aloituskodonin läheisyydessä ei paljastanut vaihtoehtoisia sekvenssejä, jotka voisivat toimia translatoituvina aloituskodoneina E. colissa. Sen vuoksi on todennäköistä, että *xyn*A:n translation aloitus tapahtuu samalla kodonilla sekä enterobakteerissa että anaerobisessa sienessä. Näin siitä huolimatta, että alempien eukaryoottien mRNA:t eivät sisällä ribosomeja sitovia sekvenssejä, jotka ovat vastaavan E. coli -sekvenssin kanssa yhdenmukaisia. Oletettavasti sekvenssi AGA, 7 emäsparia ATG-aloituskodonin ylävirrassa, toimii heikkona ribosomeja sitovana sekvenssinä bakteerissa. *xyn*A:n transkription aloituskohta E. colissa on oletettavasti vektorin *lac*Zp:ssä, koska *xyn*A cDNA:n alakloonaus 2,3 kb:n *Eco*RI-*Xho*I-restriktiofragmentissa pMTL22:een synnytti rekombinantti-plasmidin (pNX2), joka ei ohjannut funktionaalista ksylanaasia. Vektorin *lac*Zp on *xyn*A:n 3'-alueella pNX2:ssa. Vaikka E. coli ei eritä XYLAa, ksylanaasin johdettu N-terminaalinen alue on yhdenmukainen erään signaalipeptidin vastaavan alueen kanssa: käsittää N-terminaalisen hydrofiilisen emäsalueen, jota seuraa 23 hallitsevasti hydrofobisen tai neutraalin aminohapon sekvenssi.

*xyn*A:n ORF:n G + C -pitoisuus oli 43,4 %, verrattuna 10,7 %:iin 5'- ja 3'-ei koodaavilla alueilla (3'-polyA-häntä poissuljettuna). Neocallimastix-DNA:n kokonais-G + C -pitoisuus on noin 15 % [Billon-Grand et al, FEMS Microbiol. Lett. 82: 267-270 (1991)], osoittaen, että genomin ei proteiinia koodaavat alueet ovat erittäin A + T-rikkaita. Poikkeama kodonin käytössä *xyn*A:ssa käy selville 61 aminohappokodonista 14:n puuttumisesta. Havaittavissa on korostunut preferenssi T:n esiintymiselle kolmannessa asemassa (~50 % kaikista kodoneista päättyy T:hen) ja G:n eksluusiolle huojunta-asemassa. Lukuunottamatta ATG:tä ja TGG:tä, jotka ovat ainoat kodonit Met:lle ja Trp:lle, vastaavassa järjestyksessä, vain 3 kodonia sisältää G:n kolmannessa asemassa: AAG, GAG ja TTG.

Kypsän XYLAn johdetun primaarirakenteen tarkastelu paljasti useita mielenkiintoisia piirteitä. Jäännösten 225-265 ja 491-519 välillä on alueita, joissa on runsaasti proliinia ja hydroksiaminohappoja. Monet sellulaasit ja ksylanaasit muodostuvat

monista alueista, jotka on sidottu sekvensseillä, joissa on runsaasti proliinia/hydroksiaminohappoja [Gilkes et al., Microbiol. Rev. 55: 303-315 (1991)]. Kahden sellaisen "linkkerisekvenssin" läsnäolo XYLAssa osoittaa, että entsyymi muodostuu vähintään 3 erillisestä domainista. Neocallimastix-ksylanaasi sisältää, linkkerialueiden lisäksi, myös 225 aminohapon toistuvan sekvenssin N-päässä ja C-pään 40 jäännöksen uudelleen alkavan domainin (kuvio 4). Mitään ilmiselvää sekvenssin konservaatiota ei ole laajojen ja pienten toistuvien alueiden välillä. Kahden N-terminaalisen toistuvan sekvenssin välillä oli 91,6 %:n ja 95,6 % identiteetti ja samankaltaisuus, vastaavassa järjestyksessä. Kahta toistuvaa aluetta koodaavalla DNA:lla oli myös sekvenssien samanlaisuutta, 699 emäsparin ja 120 emäsparin toistuvilla sekvensseillä oli 92,7 %:n ja 90,8 %:n identtisyys, vastaavassa järjestyksessä.

Esimerkki 12 - Homologia-tutkimukset

Hydrofobinen klusteranalyysi on osoittanut, että sellulaasit ja ksylanaasit voidaan ryhmittää yhdeksään entsyymiperheeseen. Yhden perheen proteiinit muistuttavat rakenteeltaan läheisesti toisiaan ja ne ovat oletettavasti kehittyneet yhteisestä esigeenistä [Henrissat et al., Gene 81: 83-95 (1989)]. XYLAn vertaaminen SWISSPROT-tietokannassa oleviin sekvensseihin paljasti homologian sienientsyymin ja Bacillus pumilis -ksylanaasi A:n [Fukusaki et al., FEBS Lett. 171: 197-201 (1984)], Bacillus circulans -ksylanaasin [Yang et al., Nucl. Acids Res. 16: 7178 (1988)], Clostridium acetobutylicum -ksylanaasi B:n [Zappe et al., Nucl. Acids Res. 18: 2179 (1990)] ja monidomainisen Ruminococcus flavefaciens -ksylanaasin N-terminaalisen alueen [Zhang & Flint, Mol. Microbiol. 6: 1013-1019 (1992)] välillä. Näiden entsyymien ja N. patriciarum XYLAn välinen homologisuusaste on esitetty kuviossa 5.

On mielenkiintoista havaita, että ainoastaan XYLAn suuri toistuva sekvenssi oli homologinen muiden hemisellulaasien kanssa; C-terminaalisella uudelleen alkavalla alueella ei ollut minkäänlaista samankaltaisuutta tietokannassa olevien proteiinien kanssa. Tämä osoittaa, että XYLAlla on modulaarinen rakenne, jossa N-terminaalinen alue muodostaa katalyyttisen domainin.

Esimerkki 13 - XYLAn rakenne ja toiminta

Sen väitteen tutkimiseksi, että N-terminaalinen toistuva sekvenssi muodosti XYLAn katalyyttisen domainin, *xyn*A:n 5'- ja 3'-alueet pilkottiin tai alakloonattiin sopiviin vektoreihin, ja saatujen *xyn*A-johdoksien kyky ekspressoida funktionaalista ksylanaasia määritettiin. *xyn*A:n katkaistu muoto, josta 3'-alueen 291 emäsparia, jotka koodasivat 40 aminohapon muodostamaa C-terminaalista toistuvaa jaksoa, oli poistettu, koodasi yhä funktionaalista ksylanaasia. Koodatun entsyymin ennakoitu M_r oli

53 000. Tämä on sama kuin XYLAlla, joka puhdistettiin E. colista, jossa oli pNX1. Siten rekombinantti-ksylanaasista, joka oli syntetisoitu täyspitkästä geenistä enterobakteerin avulla, saattoi myös puuttua C-terminaalinen toistuva sekvenssi. Tätä tuki se seikka, että useat monidomainiset sellulaasit ja ksylanaasit ovat erityisen herkkiä proteolyyttiselle pilkkomiselle linkkerisekvensseissä [Tomme et al., Eur. J. Biochem. 170: 575-581 (1988); Gilkes et al., J. Biol. Chem. 263: 10401-10407 (1988), mukaanlukien E. colin ekspressoima Pseudomonas-ksylanaasi, joka pilkottiin olennaisesti runsasseriinisten linkkerisekvenssien sisällä [Hall et al., Mol. Microbiol. 3: 1211-1219 (1989)]. Vielä olennaisempi 3'-deleetio (pNX6), joka ulottui 1011 emäsparin alueelle, ei vaikuttanut *xyn*A:n kykyyn ohjata funktionaalisen ksylanaasin synteesiä. 1324 emäsparin poistaminen *xyn*A:n 3'-alueelta johti kuitenkin XYLAn epäaktiivisen johdoksen synteesiin. Nämä tulokset osoittavat, että N. patriciarum -ksylanaasin 270 N-terminaalista tähdettä taittuvat katalyyttisesti aktiiviseen entsyymiin. Sen määrittämiseksi, taittuvatko molemmat N-terminaaliset uudelleen alkavat sekvenssit funktionaaliseen entsyymiin, 720 emäsparin *Sca*I-*Nru*I-restriktiofragmentti (*Nru*I pilkkoo pNX4:n multippelikloonausalueella) kloonattiin pMTL20:een pNX5:n synnyttämiseksi, jossa katkaistu *xyn*A oli faasissa vektoreiden *lacZ*' translaation aloituskodonin kanssa (kuvio 1). pNX5:n sisältävä E. coli ekspressoi 15 kertaa enemmän XYLAa verrattuna klooniin, joka sisälsi täyspitkän *xyn*A:n. Tämä sienientsyymin ilmentymisen voimistuminen johtuu oletettavasti E. coli translaation aloitussekvenssin käytöstä *xyn*A:ssa, jota pNX5 koodaa. XYLAn, joka puhdistettiin pNX5:n sisältävän E. colin soluttomasta ekstraktista, M_r oli 26 000 (kuvio 2B). Nämä tulokset vahvistavat sen, että toistuvat N-pään 225 jäännöstä muodostavat erillisiä katalyyttisiä domaineja. Ksylanaasiaktiivisuudessa saatiin mielenkiintoisesti lisänousu pilkkomalla muutamia aminohappoja toisen katalyyttisen domainin C-päästä, jolloin syntyi pNX9.

N- ja C-terminaalisten katalyyttisten domainien substraattispesifisyyksien tutkimiseksi määritettiin pNX6:n ja pNX5:n ekspressoimien ksylanaasien kyky pilkkoa kasvien rakennusaineina olevia polysakkarideja. Entsyymit pilkkoivat ainoastaan ksylaania, vapauttaen ksylobioosia ja ksyloosia samoja määriä kuin täyspitkä XYLA. Molemmilla katalyyttisillä domaineilla oli siten samat substraattispesifisyydet kuin täyspitkällä XYLAlla.

Vaikka monet sellulaasit ja ksylanaasit muodostuvat monista domaineista, Caldocellum saccharolyticumista johdettu *celB* [Saul et al., Appln. Environ. Microbiol. 56: 3117-3124 (1990)] on ainoa aikaisempi esimerkki entsyymistä, joka sisältää 2 erillistä katalyyttistä domainia. Tämä entsyymi muodostuu N-terminaalisesta eksogluka-

naasista ja C-terminaalisesta endoglukanaasista, jotka kuuluvat eri enstyymiperheisiin. Siten *celB*:tä koodaava geeni ilmeisesti syntyi kahden erillisen sellulaasigeenin fuusion kautta. Tämän keksinnön mukaan on selvää, että sieniperäiset ksylanaasit voivat nekin muodostua monista katalyyttisistä domaineista. Päinvastoin kuin *celB*, *xyn*A on selvästi tulosta esigeenin tandemduplikaatiosta. Ei ole itsestään selvää, mitä selektiivistä etua geeniduplikaatio antaa anaerobiselle sienelle. Onko se pelkästään mekanismi, joka lisää XYLAn katalyyttisten domainien ekspressiota? Koska tässä kuvataan ensimmäistä kertaa anaerobisesta sienestä peräisin olevaa ksylanaasia, on epäselvää, ovatko monet katalyyttiset domainit alempien eukaryootti-hemisellulaasien yhteinen piirre.

Sekvenssiluettelo

(1) Yleisinformaatio:

(i) Hakija:

(A) Nimi: Harry John Gilbert

(B) Katuosoite: 16 Kells Gardens, Low Fell

(C) Kaupunki: Gateshead

(D) Kreivikunta: Tyne and Wear

(E) Maa: Yhdistynyt Kuningaskunta

(F) Postinumero (ZIP): NE9 5XS

(A) Nimi: Geoffrey Peter Hazlewood

(B) Katuosoite: 109A Duchess Drive

(C) Kaupunki: Newmarket

(D) Kreivikunta: Suffolk

(E) Maa: Yhdistynyt Kuningaskunta

(F) Postinumero (ZIP): CB8 8AL

(ii) Keksinnön nimi: Rekombinantti-ksylanaasit

(iii) Sekvenssien lukumäärä: 18

(iv) Tietokoneella luettava muoto:
3.5" MS-DOS levyke, joka sisältää ASCII-tiedoston (93-01283.ASC)

(v) Voimassa olevat tiedot hakemuksesta:
Hakemuksen numero: WO PCT/GB93/01283

(2) SEQ ID NO: 1:tä koskevat tiedot:

(i) Sekvenssiominaisuudet:

(A) Pituus: 2338 emäsparia

(B) Tyyppi: nukleiinihappo

(C) Nauhaisuus: kaksinauhainen

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: cDNA

(iii) Hypoteettinen: ei

(iv) Anti-sense: ei

(iii) Alkuperä:
(A) Organismi: Neocallimastix patriciarum
(B) Kanta: (tyyppilaji)

(ix) Tunnusmerkki:
(A) Nimi/Avain: CDS
(B) Sijainti: 195...2018
(D) Muuta tietoa: /funktio = "Ksylanolyyttinen entsyymi"
/ tuote = "XYLA"
/standardinimi = "Ksylanaasi"

(ix) Tunnusmerkki:
(A) Nimi/Avain: signaalipeptidi
(B) Sijainti: 195...281

(ix) Tunnusmerkki:
(A) Nimi/Avain: kypsä peptidi
(B) Sijainti: 282...2018

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 282...959
(D) Muuta tietoa: /nimike = CAT1
/huomautus = "1. katalyyttinen domaini"

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 1017...1691
(D) Muuta tietoa: /nimike = CAT2
/huomautus = "2. katalyyttinen domaini"

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti:1764...1883
(D) Muuta tietoa: /nimike = CTR1
/huomautus = "1. C-terminaalinen toistuva jakso"

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti:1884...2015
(D) Muuta tietoa: /nimike = CTR2
/huomautus = "2. C-terminaalinen toistuva jakso"

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 1...2338
(D) Muuta tietoa: /nimike = pNX1-insertti

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti:1...2338
(D) Muuta tietoa: /nimike = pNX2-insertti
/huomautus = "pNX2-insertti on päinvastaisessa suunnassa kuin pNX1-insertti"

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 1...1847
(D) Muuta tietoa: /nimike = pNX3-insertti

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 1...1725
(D) Muuta tietoa: /nimike = pNX4-insertti

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 1002...1725
(D) Muuta tietoa: /nimike = pNX5-insertti

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 1...1001
(D) Muuta tietoa: /nimike = pNX6-insertti

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 1...690
(D) Muuta tietoa: /nimike = pNX7-insertti

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 1002...2338
(D) Muuta tietoa: /nimike = pNX8-insertti

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 1002...1847
(D) Muuta tietoa: /nimike = pNX9-insertti

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 1002...1709
(D) Muuta tietoa: /nimike = pNX10-insertti

(xi) Sekvenssin kuvaus: SEQ ID NO: 1:

```
TTTTATTATA TCAATCTCTA ATTTATTTTT TTAGGAAAAA AATAAAAAAA TAAATATAAT      60

AAATATTAGA GAGTAATATT TAAAAACAAA GAAATTTAAA AACGTTTATT TAGTTATTTT     120

TTTTACTGGT TAAAAAAAAA ATAAAAAACA AAATTAATAA AGATATTTTT GAAAAATATT     180

GAATTAGAAA AAAA ATG AGA ACT ATT AAA TTC TTT TTC GCA GTA GCT ATT        230
                Met Arg Thr Ile Lys Phe Phe Phe Ala Val Ala Ile
                -29             -25                 -20

GCA ACT GTT GCT AAG GCC CAA TGG GGT GGA GGT GGT GCC TCT GCT GGT       278
Ala Thr Val Ala Lys Ala Gln Trp Gly Gly Gly Gly Ala Ser Ala Gly
        -15                 -10                 -5

CAA AGA TTA ACC GTC GGT AAT GGT CAA ACC CAA CAT AAG GGT GTA GCT       326
Gln Arg Leu Thr Val Gly Asn Gly Gln Thr Gln His Lys Gly Val Ala
      1               5                  10                  15

GAT GGT TAC AGT TAT GAA ATC TGG TTA GAT AAC ACC GGT GGT AGT GGT       374
Asp Gly Tyr Ser Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Ser Gly
                  20                  25                  30

TCT ATG ACT CTC GGT AGT GGT GCA ACC TTC AAG GCT GAA TGG AAT GCA       422
Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn Ala
              35                  40                  45

TCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT GGT CTT GAC TTC GGT       470
Ser Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe Gly
          50                  55                  60
```

```
TCT CAA AAG AAG GCA ACC GAT TAC AGC TAC ATT GGA TTG GAT TAT ACT      518
Ser Gln Lys Lys Ala Thr Asp Tyr Ser Tyr Ile Gly Leu Asp Tyr Thr
     65                  70                  75

GCA ACT TAC AGA CAA ACT GGT AGC GCA AGT GGT AAC TCC CGT CTC TGT      566
Ala Thr Tyr Arg Gln Thr Gly Ser Ala Ser Gly Asn Ser Arg Leu Cys
 80                  85                  90                  95

GTA TAC GGT TGG TTC CAA AAC CGT GGA GTT CAA GGT GTT CCA TTG GTA      614
Val Tyr Gly Trp Phe Gln Asn Arg Gly Val Gln Gly Val Pro Leu Val
                100                 105                 110

GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG GTT CCA GAT GCA CAA      662
Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala Gln
            115                 120                 125

GGT AGA ATG GTA ACC ATT GAT GGA GCT CAA TAT AAG ATT TTC CAA ATG      710
Gly Arg Met Val Thr Ile Asp Gly Ala Gln Tyr Lys Ile Phe Gln Met
        130                 135                 140

GAT CAC ACT GGT CCA ACT ATC AAT GGT GGT AGT GAA ACC TTT AAG CAA      758
Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu Thr Phe Lys Gln
    145                 150                 155

TAC TTC AGT GTC CGT CAA CAA AAG AGA ACT TCT GGT CAT ATT ACT GTC      806
Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly His Ile Thr Val
160                 165                 170                 175

TCA GAT CAC TTT AAG GAA TGG GCC AAA CAA GGT TGG GGT ATT GGT AAC      854

Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp Gly Ile Gly Asn
                180                 185                 190

CTT TAT GAA GTT GCT TTG AAC GCC GAA GGT TGG CAA AGT AGT GGT ATA      902
Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln Ser Ser Gly Ile
            195                 200                 205

GCT GAT GTC ACC AAG TTA GAT GTT TAC ACA ACC CAA AAA GGT TCT AAT      950
Ala Asp Val Thr Lys Leu Asp Val Tyr Thr Thr Gln Lys Gly Ser Asn
        210                 215                 220

CCT GCC CCT ACC TCC ACT GGT ACT GTT CCA AGC AGT TCT GCT GGT GGA      998
Pro Ala Pro Thr Ser Thr Gly Thr Val Pro Ser Ser Ser Ala Gly Gly
    225                 230                 235

AGT ACT GCC AAT GGT AAA AAG TTT ACT GTC GGT AAT GGA CAA AAC CAA     1046
Ser Thr Ala Asn Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln
240                 245                 250                 255

CAT AAG GGT GTC AAC GAT GGT TTC AGT TAT GAA ATC TGG TTA GAT AAC     1094
His Lys Gly Val Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn
                260                 265                 270

ACT GGT GGT AAC GGT TCT ATG ACT CTC GGT AGT GGT GCA ACT TTC AAG     1142
Thr Gly Gly Asn Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys
            275                 280                 285

GCT GAA TGG AAT GCA GCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT     1190
Ala Glu Trp Asn Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg
        290                 295                 300

GGT CTT GAC TTC GGT TCT CAA AAG AAG GCA ACC GAT TAC GAC TAC ATT     1238
Gly Leu Asp Phe Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile
    305                 310                 315
```

```
GGA TTA GAT TAT GCT GCT ACT TAC AAA CAA ACT GCC AGT GCA AGT GGT      1286
Gly Leu Asp Tyr Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly
320                 325                 330                 335

AAC TCC CGT CTC TGT GTA TAC GGA TGG TTC CAA AAC CGT GGA CTT AAT      1334
Asn Ser Arg Leu Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn
                340                 345                 350

GGC GTT CCT TTA GTA GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG      1382
Gly Val Pro Leu Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp
            355                 360                 365

GTT CCA GAT GCA CAA GGA AAA ATG GTA ACC ATT GAT GGA GCT CAA TAT      1430
Val Pro Asp Ala Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr
        370                 375                 380

AAG ATT TTC CAA ATG GAT CAC ACT GGT CCA ACT ATC AAT GGT GGT AGT      1478
Lys Ile Phe Gln Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser
    385                 390                 395

GAA ACC TTT AAG CAA TAC TTC AGT GTC CGT CAA CAA AAG AGA ACT TCT      1526
Glu Thr Phe Lys Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser
400                 405                 410                 415

GGT CAT ATT ACT GTC TCA GAT CAC TTT AAG GAA TGG GCC AAA CAA GGT      1574
Gly His Ile Thr Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly
                420                 425                 430

TGG GGT ATT GGT AAC CTT TAT GAA GTT GCT TTG AAC GCC GAA GGT TGG      1622
Trp Gly Ile Gly Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp
            435                 440                 445

CAA AGT AGT GGT GTT GCT GAT GTC ACC TTA TTA GAT GTT TAC ACA ACT      1670
Gln Ser Ser Gly Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr
        450                 455                 460

CCA AAG GGT TCT AGT CCA GCC ACC TCT GCC GCT CCT CGT ACT ACT ACC      1718
Pro Lys Gly Ser Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr
    465                 470                 475

CGT ACT ACT ACT CGT ACC AAG TCT CTT CCA ACC AAT TAC AAT AAG TGT      1766
Arg Thr Thr Thr Arg Thr Lys Ser Leu Pro Thr Asn Tyr Asn Lys Cys
480                 485                 490                 495

TCT GCT AGA ATT ACT GCT CAA GGT TAC AAG TGT TGT AGC GAT CCA AAT      1814
Ser Ala Arg Ile Thr Ala Gln Gly Tyr Lys Cys Cys Ser Asp Pro Asn
                500                 505                 510

TGT GTT GTT TAC TAC ACT GAT GAG GAT GGT ACC TGG GGT GTT GAA AAC      1862
Cys Val Val Tyr Tyr Thr Asp Glu Asp Gly Thr Trp Gly Val Glu Asn
            515                 520                 525

AAC GAC TGG TGT GGT TGT GGT GTT GAA CAA TGT TCT TCC AAG ATC ACT      1910
Asn Asp Trp Cys Gly Cys Gly Val Glu Gln Cys Ser Ser Lys Ile Thr
        530                 535                 540

TCT CAA GGT TAC AAG TGT TGT AGC GAT CCA AAT TGC GTT GTT TTC TAC      1958
Ser Gln Gly Tyr Lys Cys Cys Ser Asp Pro Asn Cys Val Val Phe Tyr
    545                 550                 555

ACT GAT GAC GAT GGT AAA TGG GGT GTT GAA AAC AAC GAC TGG TGT GGT      2006
Thr Asp Asp Asp Gly Lys Trp Gly Val Glu Asn Asn Asp Trp Cys Gly
560                 565                 570                 575

TGT GGT TTC TAAGCAGTAA AATACTAATT AATAAAAAAT TAAAGAATTA              2055
Cys Gly Phe
```

```
TGAAAAATTT AAATTTAAAA ATTTAAAAGA ATTATGAAAA ATTTAAATTT AAAAATTTAA   2115
AAAAAACTAA TTTAGTAAAA AATTAAAGAA TTATTGAAAA TTTTAAATGT AAAAATTTAA   2175
AAAATACAAA TTTGTAAAAA AAAATGAAAG AATTATGAAA AATTAAAATG TAAAAGTTTA   2235
AAAAATACAA ATTTGTAAGA AAAATAAAGA ATTATAAAAA AAATAAAGAA TTATGAAAAA   2295
CCCAAATGTA AAGAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA                     2338
```

(2) Tietoa SEQ ID NO: 2:sta

(i) Sekvenssin ominaisuudet:

(A) Pituus: 607 aminohappoa

(B) Tyyppi: aminohappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO: 2:

```
Met Arg Thr Ile Lys Phe Phe Phe Ala Val Ala Ile Ala Thr Val Ala
-29              -25                 -20                 -15

Lys Ala Gln Trp Gly Gly Gly Gly Ala Ser Ala Gly Gln Arg Leu Thr
            -10                 -5                  1

Val Gly Asn Gly Gln Thr Gln His Lys Gly Val Ala Asp Gly Tyr Ser
      5                  10                  15

Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Ser Gly Ser Met Thr Leu
 20                  25                  30                  35

Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn Ala Ser Val Asn Arg
                40                  45                  50

Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe Gly Ser Gln Lys Lys
            55                  60                  65

Ala Thr Asp Tyr Ser Tyr Ile Gly Leu Asp Tyr Thr Ala Thr Tyr Arg
         70                  75                  80

Gln Thr Gly Ser Ala Ser Gly Asn Ser Arg Leu Cys Val Tyr Gly Trp
      85                  90                  95

Phe Gln Asn Arg Gly Val Gln Gly Val Pro Leu Val Glu Tyr Tyr Ile
100                 105                 110                 115

Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala Gln Gly Arg Met Val
                120                 125                 130

Thr Ile Asp Gly Ala Gln Tyr Lys Ile Phe Gln Met Asp His Thr Gly
            135                 140                 145

Pro Thr Ile Asn Gly Gly Ser Glu Thr Phe Lys Gln Tyr Phe Ser Val
        150                 155                 160
```

```
Arg Gln Gln Lys Arg Thr Ser Gly His Ile Thr Val Ser Asp His Phe
    165                 170                 175

Lys Glu Trp Ala Lys Gln Gly Trp Gly Ile Gly Asn Leu Tyr Glu Val
180                 185                 190                 195

Ala Leu Asn Ala Glu Gly Trp Gln Ser Ser Gly Ile Ala Asp Val Thr
                200                 205                 210

Lys Leu Asp Val Tyr Thr Thr Gln Lys Gly Ser Asn Pro Ala Pro Thr
            215                 220                 225

Ser Thr Gly Thr Val Pro Ser Ser Ser Ala Gly Gly Ser Thr Ala Asn
        230                 235                 240

Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln His Lys Gly Val
    245                 250                 255

Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Asn
260                 265                 270                 275

Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn
                280                 285                 290

Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe
            295                 300                 305

Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile Gly Leu Asp Tyr
        310                 315                 320

Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly Asn Ser Arg Leu
    325                 330                 335

Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn Gly Val Pro Leu
340                 345                 350                 355

Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala
                360                 365                 370

Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr Lys Ile Phe Gln
            375                 380                 385

Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu Thr Phe Lys
        390                 395                 400

Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly His Ile Thr
    405                 410                 415

Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp Gly Ile Gly
420                 425                 430                 435

Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln Ser Ser Gly
                440                 445                 450

Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr Pro Lys Gly Ser
            455                 460                 465

Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr Arg Thr Thr Thr
        470                 475                 480

Arg Thr Lys Ser Leu Pro Thr Asn Tyr Asn Lys Cys Ser Ala Arg Ile
    485                 490                 495

Thr Ala Gln Gly Tyr Lys Cys Cys Ser Asp Pro Asn Cys Val Val Tyr
500                 505                 510                 515

Tyr Thr Asp Glu Asp Gly Thr Trp Gly Val Glu Asn Asn Asp Trp Cys
                520                 525                 530
```

```
Gly Cys Gly Val Glu Gln Cys Ser Ser Lys Ile Thr Ser Gln Gly Tyr
          535                 540                 545

Lys Cys Cys Ser Asp Pro Asn Cys Val Val Phe Tyr Thr Asp Asp Asp
      550                 555                 560

Gly Lys Trp Gly Val Glu Asn Asn Asp Trp Cys Gly Cys Gly Phe
  565                 570                 575
```

(2) Tietoa SEQ ID NO: 3:sta

(i) SekvenssiTunnusmerkki:

(A) Pituus: 1847 emäsparia

(B) Tyyppi: nukleiinihappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: cDNA

(ix) Tunnusmerkki:

(A) Nimi/Avain: CDS

(B) Sijainti: 195...1847

(ix) Tunnusmerkki:

(A) Nimi/Avain: signaalipeptidi

(B) Sijainti: 195...281

(ix) Tunnusmerkki:

(A) Nimi/Avain: sekal. tunnusmerkki

(B) Sijainti: 1...1847

(D) Muuta tietoa: /nimike = pNX3-insertti

(xi) Sekvenssin kuvaus: SEQ ID NO: 3:

```
TTTTATTATA TCAATCTCTA ATTTATTTTT TTAGGAAAAA AATAAAAAAA TAAATATAAT      60

AAATATTAGA GAGTAATATT TAAAAACAAA GAAATTTAAA AACGTTTATT TAGTTATTTT     120

TTTTACTGGT TAAAAAAAAA ATAAAAAACA AAATTAATAA AGATATTTTT GAAAAATATT     180

GAATTAGAAA AAAA ATG AGA ACT ATT AAA TTC TTT TTC GCA GTA GCT ATT        230
                Met Arg Thr Ile Lys Phe Phe Phe Ala Val Ala Ile
                  1               5                  10

GCA ACT GTT GCT AAG GCC CAA TGG GGT GGA GGT GGT GCC TCT GCT GGT       278
Ala Thr Val Ala Lys Ala Gln Trp Gly Gly Gly Gly Ala Ser Ala Gly
          15                  20                  25
```

```
CAA AGA TTA ACC GTC GGT AAT GGT CAA ACC CAA CAT AAG GGT GTA GCT        326
Gln Arg Leu Thr Val Gly Asn Gly Gln Thr Gln His Lys Gly Val Ala
     30                  35                  40

GAT GGT TAC AGT TAT GAA ATC TGG TTA GAT AAC ACC GGT GGT AGT GGT        374
Asp Gly Tyr Ser Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Ser Gly
 45                  50                  55                  60

TCT ATG ACT CTC GGT AGT GGT GCA ACC TTC AAG GCT GAA TGG AAT GCA        422
Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn Ala
                 65                  70                  75

TCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT GGT CTT GAC TTC GGT        470
Ser Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe Gly
             80                  85                  90

TCT CAA AAG AAG GCA ACC GAT TAC AGC TAC ATT GGA TTG GAT TAT ACT        518
Ser Gln Lys Lys Ala Thr Asp Tyr Ser Tyr Ile Gly Leu Asp Tyr Thr
         95                  100                 105

GCA ACT TAC AGA CAA ACT GGT AGC GCA AGT GGT AAC TCC CGT CTC TGT        566
Ala Thr Tyr Arg Gln Thr Gly Ser Ala Ser Gly Asn Ser Arg Leu Cys
     110                 115                 120

GTA TAC GGT TGG TTC CAA AAC CGT GGA GTT CAA GGT GTT CCA TTG GTA        614
Val Tyr Gly Trp Phe Gln Asn Arg Gly Val Gln Gly Val Pro Leu Val
 125                 130                 135                 140

GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG GTT CCA GAT GCA CAA        662
Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala Gln
                 145                 150                 155

GGT AGA ATG GTA ACC ATT GAT GGA GCT CAA TAT AAG ATT TTC CAA ATG        710
Gly Arg Met Val Thr Ile Asp Gly Ala Gln Tyr Lys Ile Phe Gln Met
             160                 165                 170

GAT CAC ACT GGT CCA ACT ATC AAT GGT GGT AGT GAA ACC TTT AAG CAA        758
Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu Thr Phe Lys Gln
         175                 180                 185

TAC TTC AGT GTC CGT CAA CAA AAG AGA ACT TCT GGT CAT ATT ACT GTC        806
Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly His Ile Thr Val
     190                 195                 200

TCA GAT CAC TTT AAG GAA TGG GCC AAA CAA GGT TGG GGT ATT GGT AAC        854
Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp Gly Ile Gly Asn
 205                 210                 215                 220

CTT TAT GAA GTT GCT TTG AAC GCC GAA GGT TGG CAA AGT AGT GGT ATA        902
Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln Ser Ser Gly Ile
                 225                 230                 235

GCT GAT GTC ACC AAG TTA GAT GTT TAC ACA ACC CAA AAA GGT TCT AAT        950
Ala Asp Val Thr Lys Leu Asp Val Tyr Thr Thr Gln Lys Gly Ser Asn
             240                 245                 250

CCT GCC CCT ACC TCC ACT GGT ACT GTT CCA AGC AGT TCT GCT GGT GGA        998
Pro Ala Pro Thr Ser Thr Gly Thr Val Pro Ser Ser Ser Ala Gly Gly
         255                 260                 265

AGT ACT GCC AAT GGT AAA AAG TTT ACT GTC GGT AAT GGA CAA AAC CAA       1046
Ser Thr Ala Asn Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln
     270                 275                 280

CAT AAG GGT GTC AAC GAT GGT TTC AGT TAT GAA ATC TGG TTA GAT AAC       1094
His Lys Gly Val Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn
285                 290                 295                 300
```

```
ACT GGT GGT AAC GGT TCT ATG ACT CTC GGT AGT GGT GCA ACT TTC AAG      1142
Thr Gly Gly Asn Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys
                305                 310                 315

GCT GAA TGG AAT GCA GCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT      1190
Ala Glu Trp Asn Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg
            320                 325                 330

GGT CTT GAC TTC GGT TCT CAA AAG AAG GCA ACC GAT TAC GAC TAC ATT      1238
Gly Leu Asp Phe Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile
        335                 340                 345

GGA TTA GAT TAT GCT GCT ACT TAC AAA CAA ACT GCC AGT GCA AGT GGT      1286
Gly Leu Asp Tyr Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly
    350                 355                 360

AAC TCC CGT CTC TGT GTA TAC GGA TGG TTC CAA AAC CGT GGA CTT AAT      1334
Asn Ser Arg Leu Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn
365                 370                 375                 380

GGC GTT CCT TTA GTA GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG      1382
Gly Val Pro Leu Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp
                385                 390                 395

GTT CCA GAT GCA CAA GGA AAA ATG GTA ACC ATT GAT GGA GCT CAA TAT      1430
Val Pro Asp Ala Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr
            400                 405                 410

AAG ATT TTC CAA ATG GAT CAC ACT GGT CCA ACT ATC AAT GGT GGT AGT      1478
Lys Ile Phe Gln Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser
        415                 420                 425

GAA ACC TTT AAG CAA TAC TTC AGT GTC CGT CAA CAA AAG AGA ACT TCT      1526
Glu Thr Phe Lys Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser
    430                 435                 440

GGT CAT ATT ACT GTC TCA GAT CAC TTT AAG GAA TGG GCC AAA CAA GGT      1574
Gly His Ile Thr Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly
445                 450                 455                 460

TGG GGT ATT GGT AAC CTT TAT GAA GTT GCT TTG AAC GCC GAA GGT TGG      1622
Trp Gly Ile Gly Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp
                465                 470                 475

CAA AGT AGT GGT GTT GCT GAT GTC ACC TTA TTA GAT GTT TAC ACA ACT      1670
Gln Ser Ser Gly Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr
            480                 485                 490

CCA AAG GGT TCT AGT CCA GCC ACC TCT GCC GCT CCT CGT ACT ACT ACC      1718
Pro Lys Gly Ser Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr
        495                 500                 505

CGT ACT ACT ACT CGT ACC AAG TCT CTT CCA ACC AAT TAC AAT AAG TGT      1766
Arg Thr Thr Thr Arg Thr Lys Ser Leu Pro Thr Asn Tyr Asn Lys Cys
    510                 515                 520

TCT GCT AGA ATT ACT GCT CAA GGT TAC AAG TGT TGT AGC GAT CCA AAT      1814
Ser Ala Arg Ile Thr Ala Gln Gly Tyr Lys Cys Cys Ser Asp Pro Asn
525                 530                 535                 540

TGT GTT GTT TAC TAC ACT GAT GAG GAT GGT ACC                          1847
Cys Val Val Tyr Tyr Thr Asp Glu Asp Gly Thr
                545                 550
```

(2) Tietoa SEQ ID NO: 4:stä:

(i) Sekvenssiominaisuudet:

(A) Pituus: 551 aminohappoa

(B) Tyyppi: aminohappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO: 4:

```
Met Arg Thr Ile Lys Phe Phe Phe Ala Val Ala Ile Ala Thr Val Ala
 1               5                  10                  15

Lys Ala Gln Trp Gly Gly Gly Gly Ala Ser Ala Gly Gln Arg Leu Thr
            20                  25                  30

Val Gly Asn Gly Gln Thr Gln His Lys Gly Val Ala Asp Gly Tyr Ser
        35                  40                  45

Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Ser Gly Ser Met Thr Leu
    50                  55                  60

Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn Ala Ser Val Asn Arg
65                  70                  75                  80

Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe Gly Ser Gln Lys Lys
                85                  90                  95

Ala Thr Asp Tyr Ser Tyr Ile Gly Leu Asp Tyr Thr Ala Thr Tyr Arg
            100                 105                 110

Gln Thr Gly Ser Ala Ser Gly Asn Ser Arg Leu Cys Val Tyr Gly Trp
        115                 120                 125

Phe Gln Asn Arg Gly Val Gln Gly Val Pro Leu Val Glu Tyr Tyr Ile
    130                 135                 140

Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala Gln Gly Arg Met Val
145                 150                 155                 160

Thr Ile Asp Gly Ala Gln Tyr Lys Ile Phe Gln Met Asp His Thr Gly
                165                 170                 175

Pro Thr Ile Asn Gly Gly Ser Glu Thr Phe Lys Gln Tyr Phe Ser Val
            180                 185                 190

Arg Gln Gln Lys Arg Thr Ser Gly His Ile Thr Val Ser Asp His Phe
        195                 200                 205

Lys Glu Trp Ala Lys Gln Gly Trp Gly Ile Gly Asn Leu Tyr Glu Val
    210                 215                 220

Ala Leu Asn Ala Glu Gly Trp Gln Ser Ser Gly Ile Ala Asp Val Thr
225                 230                 235                 240

Lys Leu Asp Val Tyr Thr Thr Gln Lys Gly Ser Asn Pro Ala Pro Thr
                245                 250                 255
```

```
Ser Thr Gly Thr Val Pro Ser Ser Ser Ala Gly Gly Ser Thr Ala Asn
            260                 265                 270

Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln His Lys Gly Val
        275                 280                 285

Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Asn
    290                 295                 300

Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn
305                 310                 315                 320

Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe
                325                 330                 335

Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile Gly Leu Asp Tyr
            340                 345                 350

Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly Asn Ser Arg Leu
        355                 360                 365

Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn Gly Val Pro Leu
    370                 375                 380

Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala
385                 390                 395                 400

Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr Lys Ile Phe Gln
                405                 410                 415

Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu Thr Phe Lys
            420                 425                 430

Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly His Ile Thr
        435                 440                 445

Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp Gly Ile Gly
    450                 455                 460

Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln Ser Ser Gly
465                 470                 475                 480

Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr Pro Lys Gly Ser
                485                 490                 495

Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr Arg Thr Thr Thr
            500                 505                 510

Arg Thr Lys Ser Leu Pro Thr Asn Tyr Asn Lys Cys Ser Ala Arg Ile
        515                 520                 525

Thr Ala Gln Gly Tyr Lys Cys Cys Ser Asp Pro Asn Cys Val Val Tyr
    530                 535                 540

Tyr Thr Asp Glu Asp Gly Thr
545                 550
```

(2) Tietoa SEQ ID NO: 5:stä:

(i) Sekvenssiominaisuudet:

(A) Pituus: 1725 emäsparia

(B) Tyyppi: nukleiinihappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: cDNA

(ix) Tunnusmerkki:

(A) Nimi/Avain: CDS

(B) Sijainti: 195...1724

(ix) Tunnusmerkki:

(A) Nimi/Avain: signaalipeptidi

(B) Sijainti: 195...281

(ix) Tunnusmerkki:

(A) Nimi/Avain: sekal. tunnusmerkki

(B) Sijainti: 1...1725

(D) Muuta tietoa: /nimike = pNX4-insertti

(xi) Sekvenssin kuvaus: SEQ ID NO: 5:

```
TTTTATTATA TCAATCTCTA ATTTATTTTT TTAGGAAAAA AATAAAAAAA TAAATATAAT      60

AAATATTAGA GAGTAATATT TAAAAACAAA GAAATTTAAA AACGTTTATT TAGTTATTTT     120

TTTTACTGGT TAAAAAAAAA ATAAAAAACA AAATTAATAA AGATATTTTT GAAAAATATT     180

GAATTAGAAA AAAA ATG AGA ACT ATT AAA TTC TTT TTC GCA GTA GCT ATT        230
                Met Arg Thr Ile Lys Phe Phe Phe Ala Val Ala Ile
                  1               5                  10

GCA ACT GTT GCT AAG GCC CAA TGG GGT GGA GGT GGT GCC TCT GCT GGT        278
Ala Thr Val Ala Lys Ala Gln Trp Gly Gly Gly Gly Ala Ser Ala Gly
          15                  20                  25

CAA AGA TTA ACC GTC GGT AAT GGT CAA ACC CAA CAT AAG GGT GTA GCT        326
Gln Arg Leu Thr Val Gly Asn Gly Gln Thr Gln His Lys Gly Val Ala
     30                  35                  40

GAT GGT TAC AGT TAT GAA ATC TGG TTA GAT AAC ACC GGT GGT AGT GGT        374
Asp Gly Tyr Ser Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Ser Gly
 45                  50                  55                  60

TCT ATG ACT CTC GGT AGT GGT GCA ACC TTC AAG GCT GAA TGG AAT GCA        422
Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn Ala
                  65                  70                  75

TCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT GGT CTT GAC TTC GGT        470
Ser Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe Gly
             80                  85                  90
```

```
TCT CAA AAG AAG GCA ACC GAT TAC AGC TAC ATT GGA TTG GAT TAT ACT      518
Ser Gln Lys Lys Ala Thr Asp Tyr Ser Tyr Ile Gly Leu Asp Tyr Thr
         95                  100                 105

GCA ACT TAC AGA CAA ACT GGT AGC GCA AGT GGT AAC TCC CGT CTC TGT      566
Ala Thr Tyr Arg Gln Thr Gly Ser Ala Ser Gly Asn Ser Arg Leu Cys
    110                 115                 120

GTA TAC GGT TGG TTC CAA AAC CGT GGA GTT CAA GGT GTT CCA TTG GTA      614
Val Tyr Gly Trp Phe Gln Asn Arg Gly Val Gln Gly Val Pro Leu Val
125                 130                 135                 140

GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG GTT CCA GAT GCA CAA      662
Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala Gln
                145                 150                 155

GGT AGA ATG GTA ACC ATT GAT GGA GCT CAA TAT AAG ATT TTC CAA ATG      710
Gly Arg Met Val Thr Ile Asp Gly Ala Gln Tyr Lys Ile Phe Gln Met
            160                 165                 170

GAT CAC ACT GGT CCA ACT ATC AAT GGT GGT AGT GAA ACC TTT AAG CAA      758
Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu Thr Phe Lys Gln
        175                 180                 185

TAC TTC AGT GTC CGT CAA CAA AAG AGA ACT TCT GGT CAT ATT ACT GTC      806
Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly His Ile Thr Val
    190                 195                 200

TCA GAT CAC TTT AAG GAA TGG GCC AAA CAA GGT TGG GGT ATT GGT AAC      854
Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp Gly Ile Gly Asn
205                 210                 215                 220

CTT TAT GAA GTT GCT TTG AAC GCC GAA GGT TGG CAA AGT AGT GGT ATA      902
Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln Ser Ser Gly Ile
                225                 230                 235

GCT GAT GTC ACC AAG TTA GAT GTT TAC ACA ACC CAA AAA GGT TCT AAT      950
Ala Asp Val Thr Lys Leu Asp Val Tyr Thr Thr Gln Lys Gly Ser Asn
            240                 245                 250

CCT GCC CCT ACC TCC ACT GGT ACT GTT CCA AGC AGT TCT GCT GGT GGA      998
Pro Ala Pro Thr Ser Thr Gly Thr Val Pro Ser Ser Ser Ala Gly Gly
        255                 260                 265

AGT ACT GCC AAT GGT AAA AAG TTT ACT GTC GGT AAT GGA CAA AAC CAA     1046
Ser Thr Ala Asn Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln
    270                 275                 280

CAT AAG GGT GTC AAC GAT GGT TTC AGT TAT GAA ATC TGG TTA GAT AAC     1094
His Lys Gly Val Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn
285                 290                 295                 300

ACT GGT GGT AAC GGT TCT ATG ACT CTC GGT AGT GGT GCA ACT TTC AAG     1142
Thr Gly Gly Asn Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys
                305                 310                 315

GCT GAA TGG AAT GCA GCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT     1190
Ala Glu Trp Asn Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg
            320                 325                 330

GGT CTT GAC TTC GGT TCT CAA AAG AAG GCA ACC GAT TAC GAC TAC ATT     1238
Gly Leu Asp Phe Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile
        335                 340                 345

GGA TTA GAT TAT GCT GCT ACT TAC AAA CAA ACT GCC AGT GCA AGT GGT     1286
Gly Leu Asp Tyr Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly
    350                 355                 360
```

```
AAC TCC CGT CTC TGT GTA TAC GGA TGG TTC CAA AAC CGT GGA CTT AAT      1334
Asn Ser Arg Leu Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn
365                 370                 375                 380

GGC GTT CCT TTA GTA GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG      1382
Gly Val Pro Leu Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp
                385                 390                 395

GTT CCA GAT GCA CAA GGA AAA ATG GTA ACC ATT GAT GGA GCT CAA TAT      1430
Val Pro Asp Ala Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr
            400                 405                 410

AAG ATT TTC CAA ATG GAT CAC ACT GGT CCA ACT ATC AAT GGT GGT AGT      1478
Lys Ile Phe Gln Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser
        415                 420                 425

GAA ACC TTT AAG CAA TAC TTC AGT GTC CGT CAA CAA AAG AGA ACT TCT      1526
Glu Thr Phe Lys Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser
    430                 435                 440

GGT CAT ATT ACT GTC TCA GAT CAC TTT AAG GAA TGG GCC AAA CAA GGT      1574
Gly His Ile Thr Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly
445                 450                 455                 460

TGG GGT ATT GGT AAC CTT TAT GAA GTT GCT TTG AAC GCC GAA GGT TGG      1622
Trp Gly Ile Gly Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp
                465                 470                 475

CAA AGT AGT GGT GTT GCT GAT GTC ACC TTA TTA GAT GTT TAC ACA ACT      1670
Gln Ser Ser Gly Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr
            480                 485                 490

CCA AAG GGT TCT AGT CCA GCC ACC TCT GCC GCT CCT CGT ACT ACT ACC      1718
Pro Lys Gly Ser Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr
        495                 500                 505

CGT ACT A                                                           1725
Arg Thr
    510
```

(2) Tietoa SEQ ID NO: 6:sta:

(i) Sekvenssiominaisuudet:

(A) Pituus: 510 aminohappoa

(B) Tyyppi: aminohappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO: 6:

```
Met Arg Thr Ile Lys Phe Phe Phe Ala Val Ala Ile Ala Thr Val Ala
 1               5                  10                  15

Lys Ala Gln Trp Gly Gly Gly Gly Ala Ser Ala Gly Gln Arg Leu Thr
            20                  25                  30
```

```
Val Gly Asn Gly Gln Thr Gln His Lys Gly Val Ala Asp Gly Tyr Ser
         35                  40                  45

Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Ser Gly Ser Met Thr Leu
     50                  55                  60

Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn Ala Ser Val Asn Arg
 65                  70                  75                  80

Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe Gly Ser Gln Lys Lys
                 85                  90                  95

Ala Thr Asp Tyr Ser Tyr Ile Gly Leu Asp Tyr Thr Ala Thr Tyr Arg
            100                 105                 110

Gln Thr Gly Ser Ala Ser Gly Asn Ser Arg Leu Cys Val Tyr Gly Trp
        115                 120                 125

Phe Gln Asn Arg Gly Val Gln Gly Val Pro Leu Val Glu Tyr Tyr Ile
    130                 135                 140

Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala Gln Gly Arg Met Val
145                 150                 155                 160

Thr Ile Asp Gly Ala Gln Tyr Lys Ile Phe Gln Met Asp His Thr Gly
                165                 170                 175

Pro Thr Ile Asn Gly Gly Ser Glu Thr Phe Lys Gln Tyr Phe Ser Val
            180                 185                 190

Arg Gln Gln Lys Arg Thr Ser Gly His Ile Thr Val Ser Asp His Phe
        195                 200                 205

Lys Glu Trp Ala Lys Gln Gly Trp Gly Ile Gly Asn Leu Tyr Glu Val
    210                 215                 220

Ala Leu Asn Ala Glu Gly Trp Gln Ser Ser Gly Ile Ala Asp Val Thr
225                 230                 235                 240

Lys Leu Asp Val Tyr Thr Thr Gln Lys Gly Ser Asn Pro Ala Pro Thr
                245                 250                 255

Ser Thr Gly Thr Val Pro Ser Ser Ser Ala Gly Gly Ser Thr Ala Asn
            260                 265                 270

Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln His Lys Gly Val
        275                 280                 285

Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Asn
    290                 295                 300

Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn
305                 310                 315                 320

Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe
                325                 330                 335

Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile Gly Leu Asp Tyr
            340                 345                 350

Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly Asn Ser Arg Leu
        355                 360                 365

Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn Gly Val Pro Leu
    370                 375                 380
```

```
Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala
385                 390                 395                 400

Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr Lys Ile Phe Gln
                405                 410                 415

Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu Thr Phe Lys
            420                 425                 430

Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly His Ile Thr
        435                 440                 445

Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp Gly Ile Gly
    450                 455                 460

Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln Ser Ser Gly
465                 470                 475                 480

Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr Pro Lys Gly Ser
                485                 490                 495

Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr Arg Thr
            500                 505                 510
```

(2) Tietoa SEQ ID NO: 7:stä:

(i) Sekvenssiominaisuudet:

(A) Pituus: 724 emäsparia

(B) Tyyppi: nukleiinihappo

(C) Nauhaisuus: kaksinauhainen

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(ix) Tunnusmerkki:

(A) Nimi/Avain: CDS

(B) Sijainti: 1...723

(ix) Tunnusmerkki:

(A) Nimi/Avain: sekal. tunnusmerkki

(B) Sijainti: 1...724

(D) Muuta tietoa: /nimike = pNX5-insertti

(xi) Sekvenssin kuvaus: SEQ ID NO: 7:

```
ACT GCC AAT GGT AAA AAG TTT ACT GTC GGT AAT GGA CAA AAC CAA CAT      48
Thr Ala Asn Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln His
 1               5                  10                  15
```

```
AAG GGT GTC AAC GAT GGT TTC AGT TAT GAA ATC TGG TTA GAT AAC ACT          96
Lys Gly Val Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn Thr
                20                  25                  30

GGT GGT AAC GGT TCT ATG ACT CTC GGT AGT GGT GCA ACT TTC AAG GCT         144
Gly Gly Asn Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala
            35                  40                  45

GAA TGG AAT GCA GCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT GGT         192
Glu Trp Asn Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly
        50                  55                  60

CTT GAC TTC GGT TCT CAA AAG AAG GCA ACC GAT TAC GAC TAC ATT GGA         240
Leu Asp Phe Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile Gly
 65                  70                  75                  80

TTA GAT TAT GCT GCT ACT TAC AAA CAA ACT GCC AGT GCA AGT GGT AAC         288
Leu Asp Tyr Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly Asn
                85                  90                  95

TCC CGT CTC TGT GTA TAC GGA TGG TTC CAA AAC CGT GGA CTT AAT GGC         336
Ser Arg Leu Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn Gly
            100                 105                 110

GTT CCT TTA GTA GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG GTT         384
Val Pro Leu Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val
        115                 120                 125

CCA GAT GCA CAA GGA AAA ATG GTA ACC ATT GAT GGA GCT CAA TAT AAG         432
Pro Asp Ala Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr Lys
    130                 135                 140

ATT TTC CAA ATG GAT CAC ACT GGT CCA ACT ATC AAT GGT GGT AGT GAA         480
Ile Phe Gln Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu
145                 150                 155                 160

ACC TTT AAG CAA TAC TTC AGT GTC CGT CAA CAA AAG AGA ACT TCT GGT         528
Thr Phe Lys Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly
                165                 170                 175

CAT ATT ACT GTC TCA GAT CAC TTT AAG GAA TGG GCC AAA CAA GGT TGG         576
His Ile Thr Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp
            180                 185                 190

GGT ATT GGT AAC CTT TAT GAA GTT GCT TTG AAC GCC GAA GGT TGG CAA         624
Gly Ile Gly Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln
        195                 200                 205

AGT AGT GGT GTT GCT GAT GTC ACC TTA TTA GAT GTT TAC ACA ACT CCA         672
Ser Ser Gly Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr Pro
    210                 215                 220

AAG GGT TCT AGT CCA GCC ACC TCT GCC GCT CCT CGT ACT ACT ACC CGT         720
Lys Gly Ser Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr Arg
225                 230                 235                 240

ACT A                                                                   724
Thr
```

(2) Tietoa SEQ ID NO: 8:sta:

(i) Sekvenssiominaisuudet:

(A) Pituus: 241 aminohappoa

(B) Tyyppi: aminohappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO: 8:

```
Thr Ala Asn Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln His
 1               5                  10                  15

Lys Gly Val Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn Thr
            20                  25                  30

Gly Gly Asn Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala
        35                  40                  45

Glu Trp Asn Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly
    50                  55                  60

Leu Asp Phe Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile Gly
65                  70                  75                  80

Leu Asp Tyr Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly Asn
                85                  90                  95

Ser Arg Leu Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn Gly
            100                 105                 110

Val Pro Leu Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val
        115                 120                 125

Pro Asp Ala Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr Lys
    130                 135                 140

Ile Phe Gln Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu
145                 150                 155                 160

Thr Phe Lys Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly
                165                 170                 175

His Ile Thr Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp
            180                 185                 190

Gly Ile Gly Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln
        195                 200                 205

Ser Ser Gly Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr Pro
    210                 215                 220

Lys Gly Ser Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr Arg
225                 230                 235                 240

Thr
```

(2) Tietoa SEQ ID NO: 9:stä:

(i) Sekvenssiominaisuudet:

(A) Pituus: 1001 emäsparia

(B) Tyyppi: nukleiinihappo

(C) Nauhaisuus: kaksinauhainen

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: cDNA

(ix) Tunnusmerkki:

(A) Nimi/Avain: CDS

(B) Sijainti: 195...1001

(ix) Tunnusmerkki:

(A) Nimi/Avain: signaalipeptidi

(B) Sijainti: 195...281

(ix) Tunnusmerkki:

(A) Nimi/Avain: sekal. tunnusmerkki

(B) Sijainti: 1...1001

(D) Muuta tietoa: /nimike = pNX6-insertti

(xi) Sekvenssin kuvaus: SEQ ID NO: 9:

```
TTTTATTATA TCAATCTCTA ATTTATTTTT TTAGGAAAAA AATAAAAAAA TAAATATAAT      60

AAATATTAGA GAGTAATATT TAAAAACAAA GAAATTTAAA AACGTTTATT TAGTTATTTT     120

TTTTACTGGT TAAAAAAAAA ATAAAAAACA AAATTAATAA AGATATTTTT GAAAAATATT     180

GAATTAGAAA AAAA ATG AGA ACT ATT AAA TTC TTT TTC GCA GTA GCT ATT        230
                Met Arg Thr Ile Lys Phe Phe Phe Ala Val Ala Ile
                  1               5                  10

GCA ACT GTT GCT AAG GCC CAA TGG GGT GGA GGT GGT GCC TCT GCT GGT        278
Ala Thr Val Ala Lys Ala Gln Trp Gly Gly Gly Gly Ala Ser Ala Gly
          15                  20                  25

CAA AGA TTA ACC GTC GGT AAT GGT CAA ACC CAA CAT AAG GGT GTA GCT        326
Gln Arg Leu Thr Val Gly Asn Gly Gln Thr Gln His Lys Gly Val Ala
      30                  35                  40

GAT GGT TAC AGT TAT GAA ATC TGG TTA GAT AAC ACC GGT GGT AGT GGT        374
Asp Gly Tyr Ser Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Ser Gly
  45                  50                  55                  60
```

```
TCT ATG ACT CTC GGT AGT GGT GCA ACC TTC AAG GCT GAA TGG AAT GCA        422
Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn Ala
                65                  70                  75

TCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT GGT CTT GAC TTC GGT        470
Ser Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe Gly
            80                  85                  90

TCT CAA AAG AAG GCA ACC GAT TAC AGC TAC ATT GGA TTG GAT TAT ACT        518
Ser Gln Lys Lys Ala Thr Asp Tyr Ser Tyr Ile Gly Leu Asp Tyr Thr
        95                  100                 105

GCA ACT TAC AGA CAA ACT GGT AGC GCA AGT GGT AAC TCC CGT CTC TGT        566
Ala Thr Tyr Arg Gln Thr Gly Ser Ala Ser Gly Asn Ser Arg Leu Cys
    110                 115                 120

GTA TAC GGT TGG TTC CAA AAC CGT GGA GTT CAA GGT GTT CCA TTG GTA        614
Val Tyr Gly Trp Phe Gln Asn Arg Gly Val Gln Gly Val Pro Leu Val
125                 130                 135                 140

GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG GTT CCA GAT GCA CAA        662
Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala Gln
                145                 150                 155

GGT AGA ATG GTA ACC ATT GAT GGA GCT CAA TAT AAG ATT TTC CAA ATG        710
Gly Arg Met Val Thr Ile Asp Gly Ala Gln Tyr Lys Ile Phe Gln Met
            160                 165                 170

GAT CAC ACT GGT CCA ACT ATC AAT GGT GGT AGT GAA ACC TTT AAG CAA        758
Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu Thr Phe Lys Gln
        175                 180                 185

TAC TTC AGT GTC CGT CAA CAA AAG AGA ACT TCT GGT CAT ATT ACT GTC        806
Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly His Ile Thr Val
    190                 195                 200

TCA GAT CAC TTT AAG GAA TGG GCC AAA CAA GGT TGG GGT ATT GGT AAC        854
Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp Gly Ile Gly Asn
205                 210                 215                 220

CTT TAT GAA GTT GCT TTG AAC GCC GAA GGT TGG CAA AGT AGT GGT ATA        902
Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln Ser Ser Gly Ile
                225                 230                 235

GCT GAT GTC ACC AAG TTA GAT GTT TAC ACA ACC CAA AAA GGT TCT AAT        950
Ala Asp Val Thr Lys Leu Asp Val Tyr Thr Thr Gln Lys Gly Ser Asn
            240                 245                 250

CCT GCC CCT ACC TCC ACT GGT ACT GTT CCA AGC AGT TCT GCT GGT GGA        998
Pro Ala Pro Thr Ser Thr Gly Thr Val Pro Ser Ser Ser Ala Gly Gly
        255                 260                 265

AGT                                                                   1001
Ser
```

(2) Tietoa SEQ ID NO: 10:stä:

(i) Sekvenssiominaisuudet:

(A) Pituus: 269 aminohappoa

(B) Tyyppi: aminohappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO: 10:

```
Met Arg Thr Ile Lys Phe Phe Phe Ala Val Ala Ile Ala Thr Val Ala
 1               5                  10                  15

Lys Ala Gln Trp Gly Gly Gly Gly Ala Ser Ala Gly Gln Arg Leu Thr
            20                  25                  30

Val Gly Asn Gly Gln Thr Gln His Lys Gly Val Ala Asp Gly Tyr Ser
        35                  40                  45

Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Ser Gly Ser Met Thr Leu
    50                  55                  60

Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn Ala Ser Val Asn Arg
65                  70                  75                  80

Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe Gly Ser Gln Lys Lys
                85                  90                  95

Ala Thr Asp Tyr Ser Tyr Ile Gly Leu Asp Tyr Thr Ala Thr Tyr Arg
            100                 105                 110

Gln Thr Gly Ser Ala Ser Gly Asn Ser Arg Leu Cys Val Tyr Gly Trp
        115                 120                 125

Phe Gln Asn Arg Gly Val Gln Gly Val Pro Leu Val Glu Tyr Tyr Ile
    130                 135                 140

Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala Gln Gly Arg Met Val
145                 150                 155                 160

Thr Ile Asp Gly Ala Gln Tyr Lys Ile Phe Gln Met Asp His Thr Gly
                165                 170                 175

Pro Thr Ile Asn Gly Gly Ser Glu Thr Phe Lys Gln Tyr Phe Ser Val
            180                 185                 190

Arg Gln Gln Lys Arg Thr Ser Gly His Ile Thr Val Ser Asp His Phe
        195                 200                 205

Lys Glu Trp Ala Lys Gln Gly Trp Gly Ile Gly Asn Leu Tyr Glu Val
    210                 215                 220

Ala Leu Asn Ala Glu Gly Trp Gln Ser Ser Gly Ile Ala Asp Val Thr
225                 230                 235                 240

Lys Leu Asp Val Tyr Thr Thr Gln Lys Gly Ser Asn Pro Ala Pro Thr
                245                 250                 255

Ser Thr Gly Thr Val Pro Ser Ser Ser Ala Gly Gly Ser
            260                 265
```

(2) Tietoa SEQ ID NO: 11:stä:

(i) Sekvenssiominaisuudet:

(A) Pituus: 690 emäsparia

(B) Tyyppi: nukleiinihappo

(C) Nauhaisuus: kaksinauhainen

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: cDNA

(ix) Tunnusmerkki:
(A) Nimi/Avain: CDS
(B) Sijainti: 195...689

(ix) Tunnusmerkki:
(A) Nimi/Avain: signaalipeptidi
(B) Sijainti: 195...281

(ix) Tunnusmerkki:
(A) Nimi/Avain: sekal. tunnusmerkki
(B) Sijainti: 1...690
(D) Muuta tietoa: /nimike = pNX7-insertti

(xi) Sekvenssin kuvaus: SEQ ID NO: 11:

```
TTTTATTATA TCAATCTCTA ATTTATTTTT TTAGGAAAAA AATAAAAAAA TAAATATAAT      60

AAATATTAGA GAGTAATATT TAAAAACAAA GAAATTTAAA AACGTTTATT TAGTTATTTT     120

TTTTACTGGT TAAAAAAAAA ATAAAAAACA AAATTAATAA AGATATTTTT GAAAAATATT     180

GAATTAGAAA AAAA ATG AGA ACT ATT AAA TTC TTT TTC GCA GTA GCT ATT        230
                Met Arg Thr Ile Lys Phe Phe Phe Ala Val Ala Ile
                  1               5                  10

GCA ACT GTT GCT AAG GCC CAA TGG GGT GGA GGT GGT GCC TCT GCT GGT        278
Ala Thr Val Ala Lys Ala Gln Trp Gly Gly Gly Gly Ala Ser Ala Gly
          15                  20                  25

CAA AGA TTA ACC GTC GGT AAT GGT CAA ACC CAA CAT AAG GGT GTA GCT        326
Gln Arg Leu Thr Val Gly Asn Gly Gln Thr Gln His Lys Gly Val Ala
     30                  35                  40

GAT GGT TAC AGT TAT GAA ATC TGG TTA GAT AAC ACC GGT GGT AGT GGT        374
Asp Gly Tyr Ser Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Ser Gly
 45                  50                  55                  60

TCT ATG ACT CTC GGT AGT GGT GCA ACC TTC AAG GCT GAA TGG AAT GCA        422
Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn Ala
                 65                  70                  75

TCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT GGT CTT GAC TTC GGT        470
Ser Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe Gly
            80                  85                  90

TCT CAA AAG AAG GCA ACC GAT TAC AGC TAC ATT GGA TTG GAT TAT ACT        518
Ser Gln Lys Lys Ala Thr Asp Tyr Ser Tyr Ile Gly Leu Asp Tyr Thr
          95                  100                 105

GCA ACT TAC AGA CAA ACT GGT AGC GCA AGT GGT AAC TCC CGT CTC TGT        566
Ala Thr Tyr Arg Gln Thr Gly Ser Ala Ser Gly Asn Ser Arg Leu Cys
     110                 115                 120
```

```
GTA TAC GGT TGG TTC CAA AAC CGT GGA GTT CAA GGT GTT CCA TTG GTA      614
Val Tyr Gly Trp Phe Gln Asn Arg Gly Val Gln Gly Val Pro Leu Val
125                 130                 135                 140

GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG GTT CCA GAT GCA CAA      662
Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala Gln
                145                 150                 155

GGT AGA ATG GTA ACC ATT GAT GGA GCT C                               690
Gly Arg Met Val Thr Ile Asp Gly Ala
            160                 165
```

(2) Tietoa SEQ ID NO: 12:sta:

(i) Sekvenssiominaisuudet:

(A) Pituus: 165 aminohappoa

(B) Tyyppi: aminohappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO: 12:

```
Met Arg Thr Ile Lys Phe Phe Phe Ala Val Ala Ile Ala Thr Val Ala
 1               5                  10                  15

Lys Ala Gln Trp Gly Gly Gly Gly Ala Ser Ala Gly Gln Arg Leu Thr
            20                  25                  30

Val Gly Asn Gly Gln Thr Gln His Lys Gly Val Ala Asp Gly Tyr Ser
        35                  40                  45

Tyr Glu Ile Trp Leu Asp Asn Thr Gly Gly Ser Gly Ser Met Thr Leu
    50                  55                  60

Gly Ser Gly Ala Thr Phe Lys Ala Glu Trp Asn Ala Ser Val Asn Arg
65                  70                  75                  80

Gly Asn Phe Leu Ala Arg Arg Gly Leu Asp Phe Gly Ser Gln Lys Lys
                85                  90                  95

Ala Thr Asp Tyr Ser Tyr Ile Gly Leu Asp Tyr Thr Ala Thr Tyr Arg
            100                 105                 110

Gln Thr Gly Ser Ala Ser Gly Asn Ser Arg Leu Cys Val Tyr Gly Trp
        115                 120                 125

Phe Gln Asn Arg Gly Val Gln Gly Val Pro Leu Val Glu Tyr Tyr Ile
    130                 135                 140

Ile Glu Asp Trp Val Asp Trp Val Pro Asp Ala Gln Gly Arg Met Val
145                 150                 155                 160

Thr Ile Asp Gly Ala
                165
```

(2) Tietoa SEQ ID NO: 13:sta:

(i) Sekvenssiominaisuudet:

(A) Pituus: 1337 emäsparia

(B) Tyyppi: nukleiinihappo

(C) Nauhaisuus: kaksinauhainen

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: cDNA

(ix) Tunnusmerkki:

(A) Nimi/Avain: CDS

(B) Sijainti: 1...1014

(ix) Tunnusmerkki:

(A) Nimi/Avain: sekal. tunnusmerkki

(B) Sijainti: 1...1337

(D) Muuta tietoa: /nimike = pNX8-insertti

(xi) Sekvenssin kuvaus: SEQ ID NO: 13:

```
ACT GCC AAT GGT AAA AAG TTT ACT GTC GGT AAT GGA CAA AAC CAA CAT        48
Thr Ala Asn Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln His
  1               5                  10                  15

AAG GGT GTC AAC GAT GGT TTC AGT TAT GAA ATC TGG TTA GAT AAC ACT        96
Lys Gly Val Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn Thr
             20                  25                  30

GGT GGT AAC GGT TCT ATG ACT CTC GGT AGT GGT GCA ACT TTC AAG GCT       144
Gly Gly Asn Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala
         35                  40                  45

GAA TGG AAT GCA GCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT GGT       192
Glu Trp Asn Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly
     50                  55                  60

CTT GAC TTC GGT TCT CAA AAG AAG GCA ACC GAT TAC GAC TAC ATT GGA       240
Leu Asp Phe Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile Gly
 65                  70                  75                  80

TTA GAT TAT GCT GCT ACT TAC AAA CAA ACT GCC AGT GCA AGT GGT AAC       288
Leu Asp Tyr Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly Asn
                 85                  90                  95

TCC CGT CTC TGT GTA TAC GGA TGG TTC CAA AAC CGT GGA CTT AAT GGC       336
Ser Arg Leu Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn Gly
            100                 105                 110
```

```
GTT CCT TTA GTA GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG GTT      384
Val Pro Leu Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val
        115                 120                 125

CCA GAT GCA CAA GGA AAA ATG GTA ACC ATT GAT GGA GCT CAA TAT AAG      432
Pro Asp Ala Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr Lys
    130                 135                 140

ATT TTC CAA ATG GAT CAC ACT GGT CCA ACT ATC AAT GGT GGT AGT GAA      480
Ile Phe Gln Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu
145                 150                 155                 160

ACC TTT AAG CAA TAC TTC AGT GTC CGT CAA CAA AAG AGA ACT TCT GGT      528
Thr Phe Lys Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly
                165                 170                 175

CAT ATT ACT GTC TCA GAT CAC TTT AAG GAA TGG GCC AAA CAA GGT TGG      576
His Ile Thr Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp
            180                 185                 190

GGT ATT GGT AAC CTT TAT GAA GTT GCT TTG AAC GCC GAA GGT TGG CAA      624
Gly Ile Gly Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln
        195                 200                 205

AGT AGT GGT GTT GCT GAT GTC ACC TTA TTA GAT GTT TAC ACA ACT CCA      672
Ser Ser Gly Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr Pro
    210                 215                 220

AAG GGT TCT AGT CCA GCC ACC TCT GCC GCT CCT CGT ACT ACT ACC CGT      720
Lys Gly Ser Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr Arg
225                 230                 235                 240

ACT ACT ACT CGT ACC AAG TCT CTT CCA ACC AAT TAC AAT AAG TGT TCT      768
Thr Thr Thr Arg Thr Lys Ser Leu Pro Thr Asn Tyr Asn Lys Cys Ser
                245                 250                 255

GCT AGA ATT ACT GCT CAA GGT TAC AAG TGT TGT AGC GAT CCA AAT TGT      816
Ala Arg Ile Thr Ala Gln Gly Tyr Lys Cys Cys Ser Asp Pro Asn Cys
            260                 265                 270

GTT GTT TAC TAC ACT GAT GAG GAT GGT ACC TGG GGT GTT GAA AAC AAC      864
Val Val Tyr Tyr Thr Asp Glu Asp Gly Thr Trp Gly Val Glu Asn Asn
        275                 280                 285

GAC TGG TGT GGT TGT GGT GTT GAA CAA TGT TCT TCC AAG ATC ACT TCT      912
Asp Trp Cys Gly Cys Gly Val Glu Gln Cys Ser Ser Lys Ile Thr Ser
    290                 295                 300

CAA GGT TAC AAG TGT TGT AGC GAT CCA AAT TGC GTT GTT TTC TAC ACT      960
Gln Gly Tyr Lys Cys Cys Ser Asp Pro Asn Cys Val Val Phe Tyr Thr
305                 310                 315                 320

GAT GAC GAT GGT AAA TGG GGT GTT GAA AAC AAC GAC TGG TGT GGT TGT     1008
Asp Asp Asp Gly Lys Trp Gly Val Glu Asn Asn Asp Trp Cys Gly Cys
                325                 330                 335

GGT TTC TAAGCAGTAA AATACTAATT AATAAAAAAT TAAAGAATTA TGAAAAATTT     1064
Gly Phe

AAATTTAAAA ATTTAAAAGA ATTATGAAAA ATTTAAATTT AAAAATTTAA AAAAAACTAA  1124

TTTAGTAAAA AATTAAAGAA TTATTGAAAA TTTTAAATGT AAAAATTTAA AAAATACAAA  1184

TTTGTAAAAA AAAATGAAAG AATTATGAAA AATTAAAATG TAAAAGTTTA AAAAATACAA  1244

ATTTGTAAGA AAAATAAAGA ATTATAAAAA AAATAAAGAA TTATGAAAAA CCCAAATGTA  1304

AAGAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA                               1337
```

(2) Tietoa SEQ ID NO: 14:sta:

(i) Sekvenssiominaisuudet:

(A) Pituus: 338 aminohappoa

(B) Tyyppi: aminohappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO: 14:

```
Thr Ala Asn Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln His
 1               5                  10                  15

Lys Gly Val Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn Thr
            20                  25                  30

Gly Gly Asn Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala
        35                  40                  45

Glu Trp Asn Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly
    50                  55                  60

Leu Asp Phe Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile Gly
65                  70                  75                  80

Leu Asp Tyr Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly Asn
                85                  90                  95

Ser Arg Leu Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn Gly
            100                 105                 110

Val Pro Leu Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val
        115                 120                 125

Pro Asp Ala Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr Lys
    130                 135                 140

Ile Phe Gln Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu
145                 150                 155                 160

Thr Phe Lys Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly
                165                 170                 175

His Ile Thr Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp
            180                 185                 190

Gly Ile Gly Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln
        195                 200                 205

Ser Ser Gly Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr Pro
    210                 215                 220

Lys Gly Ser Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr Arg
225                 230                 235                 240

Thr Thr Thr Arg Thr Lys Ser Leu Pro Thr Asn Tyr Asn Lys Cys Ser
                245                 250                 255

Ala Arg Ile Thr Ala Gln Gly Tyr Lys Cys Cys Ser Asp Pro Asn Cys
            260                 265                 270
```

```
Val Val Tyr Tyr Thr Asp Glu Asp Gly Thr Trp Gly Val Glu Asn Asn
        275                 280                 285

Asp Trp Cys Gly Cys Gly Val Glu Gln Cys Ser Ser Lys Ile Thr Ser
    290                 295                 300

Gln Gly Tyr Lys Cys Cys Ser Asp Pro Asn Cys Val Val Phe Tyr Thr
305                 310                 315                 320

Asp Asp Asp Gly Lys Trp Gly Val Glu Asn Asn Asp Trp Cys Gly Cys
                325                 330                 335

Gly Phe
```

(2) Tietoa SEQ ID NO: 15:sta:

(i) Sekvenssiominaisuudet:

(A) Pituus: 846 emäsparia

(B) Tyyppi: nukleiinihappo

(C) Nauhaisuus: kaksinauhainen

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: cDNA

(ix) Tunnusmerkki:

(A) Nimi/Avain: CDS

(B) Sijainti: 1...846

(ix) Tunnusmerkki:

(A) Nimi/Avain: sekal. tunnusmerkki

(B) Sijainti: 1...846

(D) Muuta tietoa: /nimike = pNX9-insertti

(xi) Sekvenssin kuvaus: SEQ ID NO: 15:

```
ACT GCC AAT GGT AAA AAG TTT ACT GTC GGT AAT GGA CAA AAC CAA CAT      48
Thr Ala Asn Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln His
  1               5                  10                  15

AAG GGT GTC AAC GAT GGT TTC AGT TAT GAA ATC TGG TTA GAT AAC ACT      96
Lys Gly Val Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn Thr
            20                  25                  30
```

```
GGT GGT AAC GGT TCT ATG ACT CTC GGT AGT GGT GCA ACT TTC AAG GCT        144
Gly Gly Asn Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala
          35                  40                  45

GAA TGG AAT GCA GCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT GGT        192
Glu Trp Asn Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly
      50                  55                  60

CTT GAC TTC GGT TCT CAA AAG AAG GCA ACC GAT TAC GAC TAC ATT GGA        240
Leu Asp Phe Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile Gly
 65                  70                  75                  80

TTA GAT TAT GCT GCT ACT TAC AAA CAA ACT GCC AGT GCA AGT GGT AAC        288
Leu Asp Tyr Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly Asn
                 85                  90                  95

TCC CGT CTC TGT GTA TAC GGA TGG TTC CAA AAC CGT GGA CTT AAT GGC        336
Ser Arg Leu Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn Gly
            100                 105                 110

GTT CCT TTA GTA GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG GTT        384
Val Pro Leu Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val
        115                 120                 125

CCA GAT GCA CAA GGA AAA ATG GTA ACC ATT GAT GGA GCT CAA TAT AAG        432
Pro Asp Ala Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr Lys
    130                 135                 140

ATT TTC CAA ATG GAT CAC ACT GGT CCA ACT ATC AAT GGT GGT AGT GAA        480
Ile Phe Gln Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu
145                 150                 155                 160

ACC TTT AAG CAA TAC TTC AGT GTC CGT CAA CAA AAG AGA ACT TCT GGT        528
Thr Phe Lys Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly
                165                 170                 175

CAT ATT ACT GTC TCA GAT CAC TTT AAG GAA TGG GCC AAA CAA GGT TGG        576
His Ile Thr Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp
            180                 185                 190

GGT ATT GGT AAC CTT TAT GAA GTT GCT TTG AAC GCC GAA GGT TGG CAA        624
Gly Ile Gly Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln
        195                 200                 205

AGT AGT GGT GTT GCT GAT GTC ACC TTA TTA GAT GTT TAC ACA ACT CCA        672
Ser Ser Gly Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr Pro
    210                 215                 220

AAG GGT TCT AGT CCA GCC ACC TCT GCC GCT CCT CGT ACT ACT ACC CGT        720
Lys Gly Ser Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr Arg
225                 230                 235                 240

ACT ACT ACT CGT ACC AAG TCT CTT CCA ACC AAT TAC AAT AAG TGT TCT        768
Thr Thr Thr Arg Thr Lys Ser Leu Pro Thr Asn Tyr Asn Lys Cys Ser
                245                 250                 255

GCT AGA ATT ACT GCT CAA GGT TAC AAG TGT TGT AGC GAT CCA AAT TGT        816
Ala Arg Ile Thr Ala Gln Gly Tyr Lys Cys Cys Ser Asp Pro Asn Cys
            260                 265                 270

GTT GTT TAC TAC ACT GAT GAG GAT GGT ACC                               846
Val Val Tyr Tyr Thr Asp Glu Asp Gly Thr
        275                 280
```

(2) Tietoa SEQ ID NO: 16:sta:

(i) Sekvenssiominaisuudet:

(A) Pituus: 282 aminohappoa

(B) Tyyppi: aminohappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO: 16:

```
Thr Ala Asn Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln His
 1               5                  10                  15

Lys Gly Val Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn Thr
            20                  25                  30

Gly Gly Asn Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala
        35                  40                  45

Glu Trp Asn Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly
    50                  55                  60

Leu Asp Phe Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile Gly
65                  70                  75                  80

Leu Asp Tyr Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly Asn
                85                  90                  95

Ser Arg Leu Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn Gly
            100                 105                 110

Val Pro Leu Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val
        115                 120                 125

Pro Asp Ala Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr Lys
    130                 135                 140

Ile Phe Gln Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu
145                 150                 155                 160

Thr Phe Lys Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly
                165                 170                 175

His Ile Thr Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp
            180                 185                 190

Gly Ile Gly Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln
        195                 200                 205

Ser Ser Gly Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr Pro
    210                 215                 220

Lys Gly Ser Ser Pro Ala Thr Ser Ala Ala Pro Arg Thr Thr Thr Arg
225                 230                 235                 240

Thr Thr Thr Arg Thr Lys Ser Leu Pro Thr Asn Tyr Asn Lys Cys Ser
                245                 250                 255

Ala Arg Ile Thr Ala Gln Gly Tyr Lys Cys Cys Ser Asp Pro Asn Cys
            260                 265                 270

Val Val Tyr Tyr Thr Asp Glu Asp Gly Thr
        275                 280
```

(2) Tietoa SEQ ID NO: 17:sta:

(i) Sekvenssiominaisuudet:

(A) Pituus: 708 emäsparia

(B) Tyyppi: nukleiinihappo

(C) Nauhaisuus: kaksinauhainen

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: cDNA

(ix) Tunnusmerkki:

(A) Nimi/Avain: CDS

(B) Sijainti: 1...708

(ix) Tunnusmerkki:

(A) Nimi/Avain: sekal. tunnusmerkki

(B) Sijainti: 1...708

(D) Muuta tietoa: /nimike = pNX10-insertti

(xi) Sekvenssin kuvaus: SEQ ID NO: 17:

```
ACT GCC AAT GGT AAA AAG TTT ACT GTC GGT AAT GGA CAA AAC CAA CAT        48
Thr Ala Asn Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln His
  1               5                  10                  15

AAG GGT GTC AAC GAT GGT TTC AGT TAT GAA ATC TGG TTA GAT AAC ACT        96
Lys Gly Val Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn Thr
            20                  25                  30

GGT GGT AAC GGT TCT ATG ACT CTC GGT AGT GGT GCA ACT TTC AAG GCT       144
Gly Gly Asn Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala
        35                  40                  45

GAA TGG AAT GCA GCT GTT AAC CGT GGT AAC TTC CTT GCC CGT CGT GGT       192
Glu Trp Asn Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly
    50                  55                  60

CTT GAC TTC GGT TCT CAA AAG AAG GCA ACC GAT TAC GAC TAC ATT GGA       240
Leu Asp Phe Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile Gly
65                  70                  75                  80

TTA GAT TAT GCT GCT ACT TAC AAA CAA ACT GCC AGT GCA AGT GGT AAC       288
Leu Asp Tyr Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly Asn
                85                  90                  95

TCC CGT CTC TGT GTA TAC GGA TGG TTC CAA AAC CGT GGA CTT AAT GGC       336
Ser Arg Leu Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn Gly
            100                 105                 110

GTT CCT TTA GTA GAA TAC TAC ATC ATT GAA GAT TGG GTT GAC TGG GTT       384
Val Pro Leu Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val
        115                 120                 125
```

```
CCA GAT GCA CAA GGA AAA ATG GTA ACC ATT GAT GGA GCT CAA TAT AAG      432
Pro Asp Ala Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr Lys
    130                 135                 140

ATT TTC CAA ATG GAT CAC ACT GGT CCA ACT ATC AAT GGT GGT AGT GAA      480
Ile Phe Gln Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu
145                 150                 155                 160

ACC TTT AAG CAA TAC TTC AGT GTC CGT CAA CAA AAG AGA ACT TCT GGT      528
Thr Phe Lys Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly
                165                 170                 175

CAT ATT ACT GTC TCA GAT CAC TTT AAG GAA TGG GCC AAA CAA GGT TGG      576
His Ile Thr Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp
            180                 185                 190

GGT ATT GGT AAC CTT TAT GAA GTT GCT TTG AAC GCC GAA GGT TGG CAA      624
Gly Ile Gly Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln
        195                 200                 205

AGT AGT GGT GTT GCT GAT GTC ACC TTA TTA GAT GTT TAC ACA ACT CCA      672
Ser Ser Gly Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr Pro
    210                 215                 220

AAG GGT TCT AGT CCA GCC ACC TCT GCC GCT CCT CGT                      708
Lys Gly Ser Ser Pro Ala Thr Ser Ala Ala Pro Arg
225                 230                 235
```

(2) Tietoa SEQ ID NO: 18:sta:

(i) Sekvenssiominaisuudet:

(A) Pituus: 236 aminohappoa

(B) Tyyppi: aminohappo

(D) Topologia: lineaarinen

(ii) Molekyylityyppi: proteiini

(xi) Sekvenssin kuvaus: SEQ ID NO: 18:

```
Thr Ala Asn Gly Lys Lys Phe Thr Val Gly Asn Gly Gln Asn Gln His
 1               5                  10                  15

Lys Gly Val Asn Asp Gly Phe Ser Tyr Glu Ile Trp Leu Asp Asn Thr
            20                  25                  30

Gly Gly Asn Gly Ser Met Thr Leu Gly Ser Gly Ala Thr Phe Lys Ala
        35                  40                  45

Glu Trp Asn Ala Ala Val Asn Arg Gly Asn Phe Leu Ala Arg Arg Gly
    50                  55                  60

Leu Asp Phe Gly Ser Gln Lys Lys Ala Thr Asp Tyr Asp Tyr Ile Gly
65                  70                  75                  80

Leu Asp Tyr Ala Ala Thr Tyr Lys Gln Thr Ala Ser Ala Ser Gly Asn
                85                  90                  95

Ser Arg Leu Cys Val Tyr Gly Trp Phe Gln Asn Arg Gly Leu Asn Gly
            100                 105                 110
```

```
Val Pro Leu Val Glu Tyr Tyr Ile Ile Glu Asp Trp Val Asp Trp Val
        115                 120                 125

Pro Asp Ala Gln Gly Lys Met Val Thr Ile Asp Gly Ala Gln Tyr Lys
    130                 135                 140

Ile Phe Gln Met Asp His Thr Gly Pro Thr Ile Asn Gly Gly Ser Glu
145                 150                 155                 160

Thr Phe Lys Gln Tyr Phe Ser Val Arg Gln Gln Lys Arg Thr Ser Gly
                165                 170                 175

His Ile Thr Val Ser Asp His Phe Lys Glu Trp Ala Lys Gln Gly Trp
            180                 185                 190

Gly Ile Gly Asn Leu Tyr Glu Val Ala Leu Asn Ala Glu Gly Trp Gln
        195                 200                 205

Ser Ser Gly Val Ala Asp Val Thr Leu Leu Asp Val Tyr Thr Thr Pro
    210                 215                 220

Lys Gly Ser Ser Pro Ala Thr Ser Ala Ala Pro Arg
225                 230                 235
```

Yhteenveto sekvenssiluettelosta

SEQ ID NO: 1	pNX1 DNA ja koodaava alue
SEQ ID NO: 2	SEQ ID NO: 1:n proteiinisekvenssi
SEQ ID NO: 3	pNX3 DNA ja koodaava alue
SEQ ID NO: 4	SEQ ID NO: 3:n proteiinisekvenssi
SEQ ID NO: 5	pNX4 DNA ja koodaava alue
SEQ ID NO: 6	SEQ ID NO: 5:n proteiinisekvenssi
SEQ ID NO: 7	pNX5 DNA ja koodaava alue
SEQ ID NO: 8	SEQ ID NO: 7:n proteiinisekvenssi
SEQ ID NO: 9	pNX6 DNA ja koodaava alue
SEQ ID NO: 10	SEQ ID NO: 9:n proteiinisekvenssi
SEQ ID NO: 11	pNX7 DNA ja koodaava alue
SEQ ID NO: 12	SEQ ID NO: 11:n proteiinisekvenssi
SEQ ID NO: 13	pNX8 DNA ja koodaava alue
SEQ ID NO: 14	SEQ ID NO: 13:n proteiinisekvenssi
SEQ ID NO: 15	pNX9 DNA ja koodaava alue
SEQ ID NO: 16	SEQ ID NO: 15:n proteiinisekvenssi
SEQ ID NO: 17	pNX10 DNA ja koodaava alue
SEQ ID NO: 18	SEQ ID NO: 17:n proteiinisekvenssi

Patenttivaatimukset

1. Katkaistu anaerobinen fungaalinen ksylanaasi, **tunnettu** siitä, että se käsittää katkaistun ksylanaasikonstruktin, jossa on ainakin aktiivinen osa villityypin anaerobisen fungaalisen ksylanaasin katalyyttisestä domainista, jolloin katkaistu ksylanaasikonstrukti ilmentyy korkeammalla tasolla kuin ekvivalentti villityypin konstrukti.

2. Patenttivaatimuksen 1 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että anaerobinen sieni on pötsisieni.

3. Patenttivaatimuksen 2 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että pötsisieni on Neocallimastix-sukua.

4. Patenttivaatimuksen 3 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että anaerobinen sieni on Neocallimastix particiarum.

5. Jonkin patenttivaatimuksen 1-4 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että katkaistun ksylanaasikonstruktin ilmentyminen on ainakin kymmenkertaisesti parempi kuin ekvivalentin villityypin konstruktin.

6. Jonkin patenttivaatimuksen 1-5 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että katkaistun ksylanaasikonstruktin ilmentyminen on ainakin satakertaisesti parempi kuin ekvivalentin villityypin konstruktin.

7. Minkä tahansa patenttivaatimuksista 1 - 6 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että se on johdettu ksylanaasista, jolla on rakenne (N-päästä C-päähän päin):

CAT1-LINK1-CAT2-LINK2-CTR1-CTR2

jossa:
- CAT1 edustaa ensimmäistä katalyyttistä domainia,
- CAT2 edustaa toista katalyyttistä domainia,
- LINK1 edustaa ensimmäistä linkkeriä,
- LINK2 edustaa toista linkkeriä,
- CTR1 edustaa ensimmäistä C-terminaalista toistuvaa jaksoa ja
- CTR2 edustaa toista C-terminaalista toistuvaa jaksoa.

8. Patenttivaatimuksen 7 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että katalyyttinen domaini on olennaisen homologinen ainakin CAT1:n tai CAT2:n kanssa ja siitä puutuu ainakin osa CAT2:n alavirran (so. C-päähän päin) aminohapposekvenssistä

9. Patenttivaatimuksen 8 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että vähintään osa CTR2:sta puuttuu.

10. Patenttivaatimuksen 8 tai 9 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että vähintään osa CTR1:sta puuttuu.

11. Jonkin patenttivaatimuksen 7-10 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että sillä on rakenne:
LINK1(katkaistu)-CAT2-LINK2-CTR1-CTR2.

12. Jonkin patenttivaatimuksen 7-10 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että sillä on rakenne:
LINK1(katkaistu)-CAT2-LINK2-CTR1(katkaistu).

13. Jonkin patenttivaatimuksen 7-10 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että sillä on rakenne:
LINK1(katkaistu)-CAT2(katkaistu).

14. Jonkin patenttivaatimuksen 7-10 mukainen katkaistu ksylanaasi, **tunnettu** siitä, että sillä on jokin seuraavista rakenteista:

CAT1-LINK1-CAT2-LINK2-CTR1(katkaistu)
CAT1-LINK1-CAT2-LINK2(katkaistu)
LINK1(katkaistu)-CAT2-LINK2(katkaistu)
CAT1-LINK1(katkaistu)
CAT1(katkaistu)
LINK1(katkaistu)-CAT2-LINK2-CTR1-CTR2
LINK1(katkaistu)-CAT2-LINK2-CTR1(katkaistu) tai
LINK1(katkaistu)-CAT2(katkaistu).

15. DNA-molekyyli, **tunnettu** siitä, että se koodaa jonkin patenttivaatimuksen 1-14 mukaista katkaistua anaerobista fungaalista ksylanaasia.

16. Patenttivaatimuksen 15 mukainen DNA-molekyyli, **tunnettu** siitä, että villityypin ksylanaasilla on seuraava DNA-rakenne:

*5'utr-sig-cat*1-*link*1-*cat*2-*link*2-*ctr*1-*ctr*2-*3'utr*,

jossa
– *5'utr* tarkoittaa 5'-translatoitumatonta aluetta;
– *sig* koodaa signaalipeptidiä;
– *cat*1 koodaa ensimmäistä katalyyttistä domainia;
– *link*1 koodaa ensimmäistä linkkerisekvenssiä;
– *cat*2 koodaa toista katalyyttistä domainia;
– *link*2 koodaa toista linkkerisekvenssiä;
– *ctr*1 koodaa ensimmäistä C-terminaalista toistuvaa jaksoa;

– *ctr*2 koodaa toista C-terminaalista toistuvaa jaksoa; ja
– *3'utr* tarkoittaa 3'-translatoitumatonta aluetta.

17. Patenttivaatimuksen 15 tai 16 mukainen DNA-sekvenssi, **tunnettu** siitä, että se käsittää seuraavan segmentin:
*link*1(katkaistu)-*cat*2-*link*2-*ctr*1-*ctr*2-*3'utr*.

18. Patenttivaatimuksen 15 tai 16 mukainen DNA-sekvenssi, **tunnettu** siitä, että se käsittää seuraavan segmentin:
*link*1(katkaistu)-*cat*2-*link*2-*ctr*1(katkaistu).

19. Patenttivaatimuksen 15 tai 16 mukainen DNA-sekvenssi, **tunnettu** siitä, että se käsittää seuraavan segmentin:
*link*1(katkaistu)-*cat*2(katkaistu).

20. Patenttivaatimuksen 15 tai 16 mukainen DNA-sekvenssi, **tunnettu** siitä, että se käsittää seuraavat segmentit:

*5'utr-sig-cat*1-*link*1-*cat*2-*link*2-*ctr*1(katkaistu);
*5'utr-sig-cat*1-*link*1-*cat*2-*link*2(katkaistu);
*link*1(katkaistu)-*cat*2-*link*2(katkaistu) ;
*5'utr-sig-cat*1-*link*1(katkaistu);
*5'utr-sig-cat*1(katkaistu) ;
*link*1(katkaistu)-*cat*2-*link*2-*ctr*1-*ctr*2-*3'utr*;
*link*1(katkaistu)-*cat*2-*link*2-*ctr*1(katkaistu);
*link*1(katkaistu)-*cat*2(katkaistu).

21. Ekspressiovektori, **tunnettu** siitä, että se käsittää minkä tahansa patenttivaatimuksen 15-20 mukaisen DNA-molekyylin.

22. Patenttivaatimuksen 21 mukainen ekspressiovektori, **tunnettu** siitä, että vektori on plasmidi.

23. DNA-molekyyli, **tunnettu** siitä, että se käsittää pNX8:n, pNX9:n tai pNX10:n, jotka on kuvattu SEQ ID numeroissa 13, 15 ja 17, vastaavasti.

24. Isäntäsolu, **tunnettu** siitä, että se on transfektoitu tai transformoitu minkä tahansa patenttivaatimuksen 15 – 20 tai 23 mukaisella DNA-molekyylillä.

25. Katkaistu ksylanaasi, **tunnettu** siitä, että sitä koodaa minkä tahansa patenttivaatimuksen 15 – 20 tai 23 mukainen DNA-molekyyli.

26. Minkä tahansa patenttivaatimuksen 1 – 14 ja 25 mukaisen ksylanaasin käyttö epäpuhtauden poistajana sellusta.

27. Minkä tahansa patenttivaatimuksen 1 – 14 ja 25 mukaisen ksylanaasin käyttö kraftmassan esivalkaisussa.

28. Minkä tahansa patenttivaatimuksen 1 – 14 ja 25 mukaisen ksylanaasin käyttö eläinten rehuun tarkoitettuna entsyymilisänä.

29. Minkä tahansa patenttivaatimuksen 1 – 14 ja 25 mukaisen ksylanaasin käyttö leivottujen tuotteiden modifioinnissa.

FIG. 1

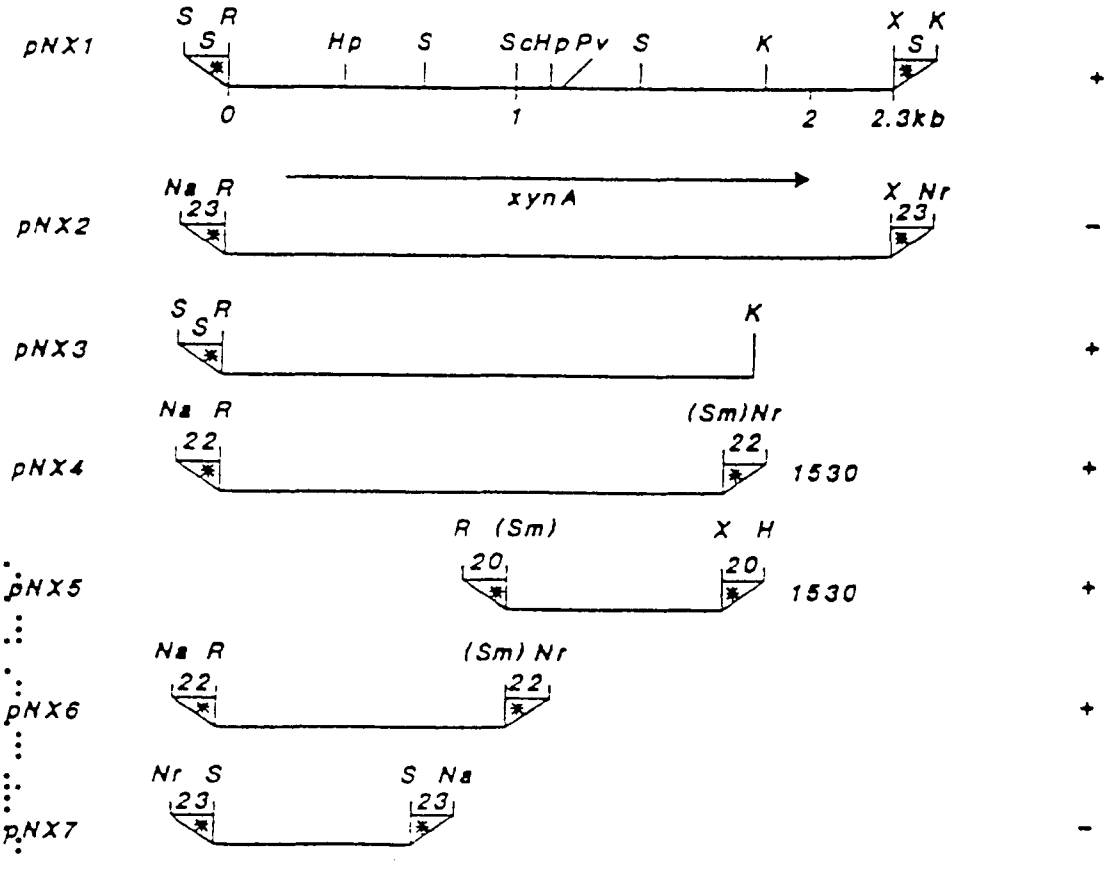

Ksynalaasiaktiivisuus
pNX1
S R
S
Hp
S
ScHpPv
S
K
X K
S
0
1
2
2.3kb
+
pNX2
Na R
23
xynA
X Nr
23
-
pNX3
S S R
K
+
pNX4
Na R
22
(Sm)Nr
22
1530
+
pNX5
R (Sm)
20
X H
20
1530
+
pNX6
Na R
22
(Sm) Nr
22
+
pNX7
Nr S
23
S Na
23
-

FIG. 2

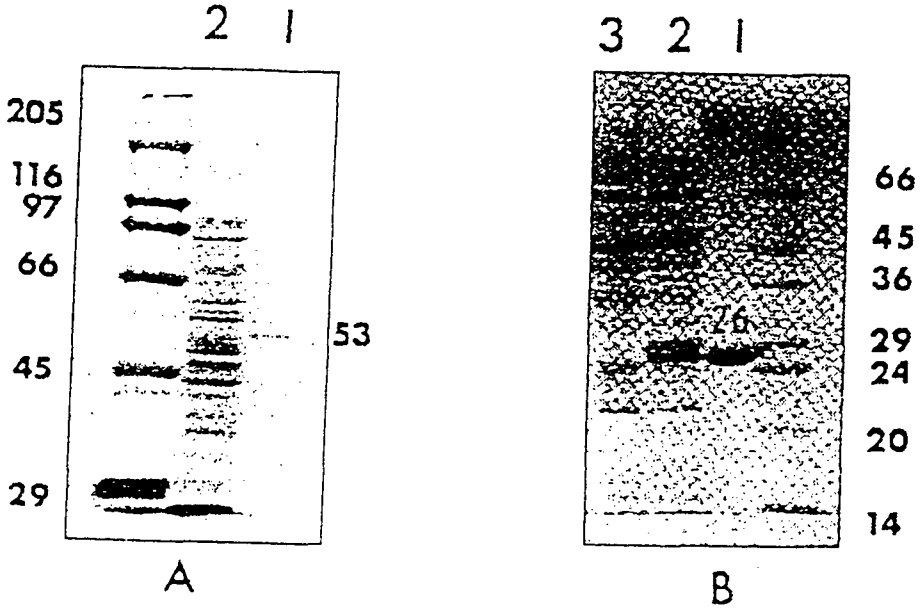
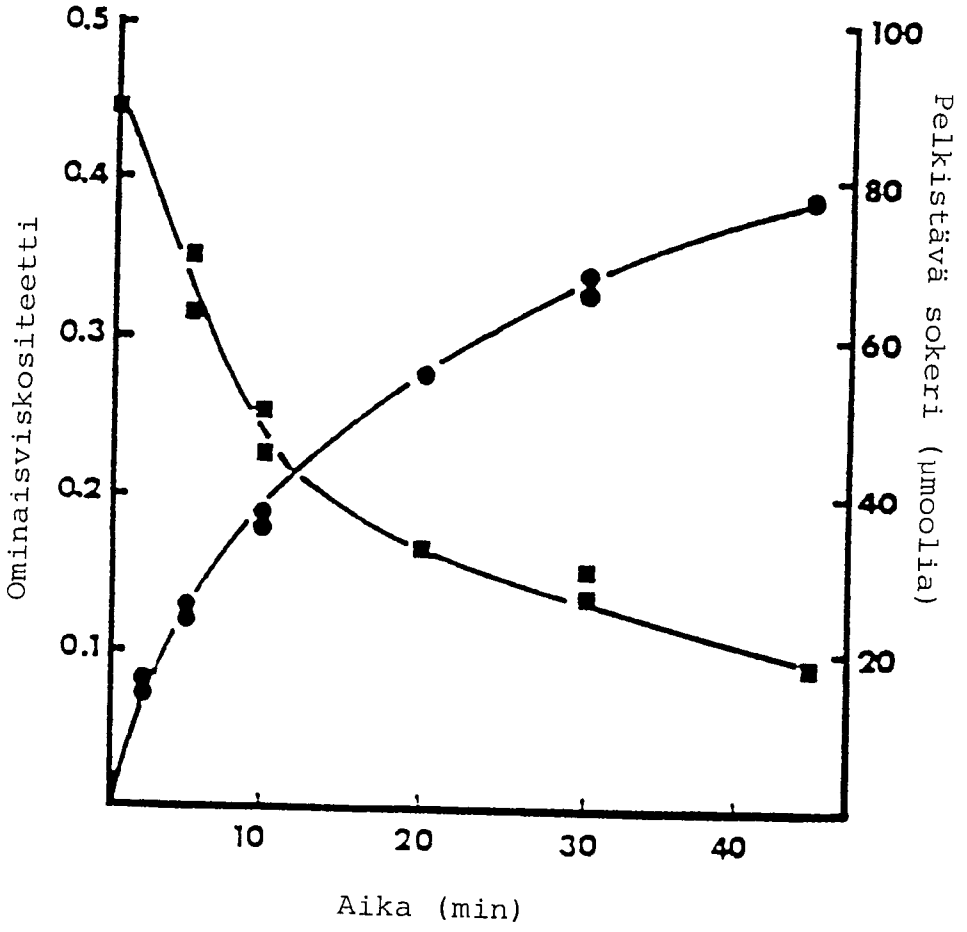


FIG. 3


0.5
0.4
0.3
0.2
0.1
Ominaisviskositeetti
100
80
60
40
20
Pelkistävä sokeri (µmoolia)
10
20
30
40
Aika (min)

16-02-95 94503

FIG.4

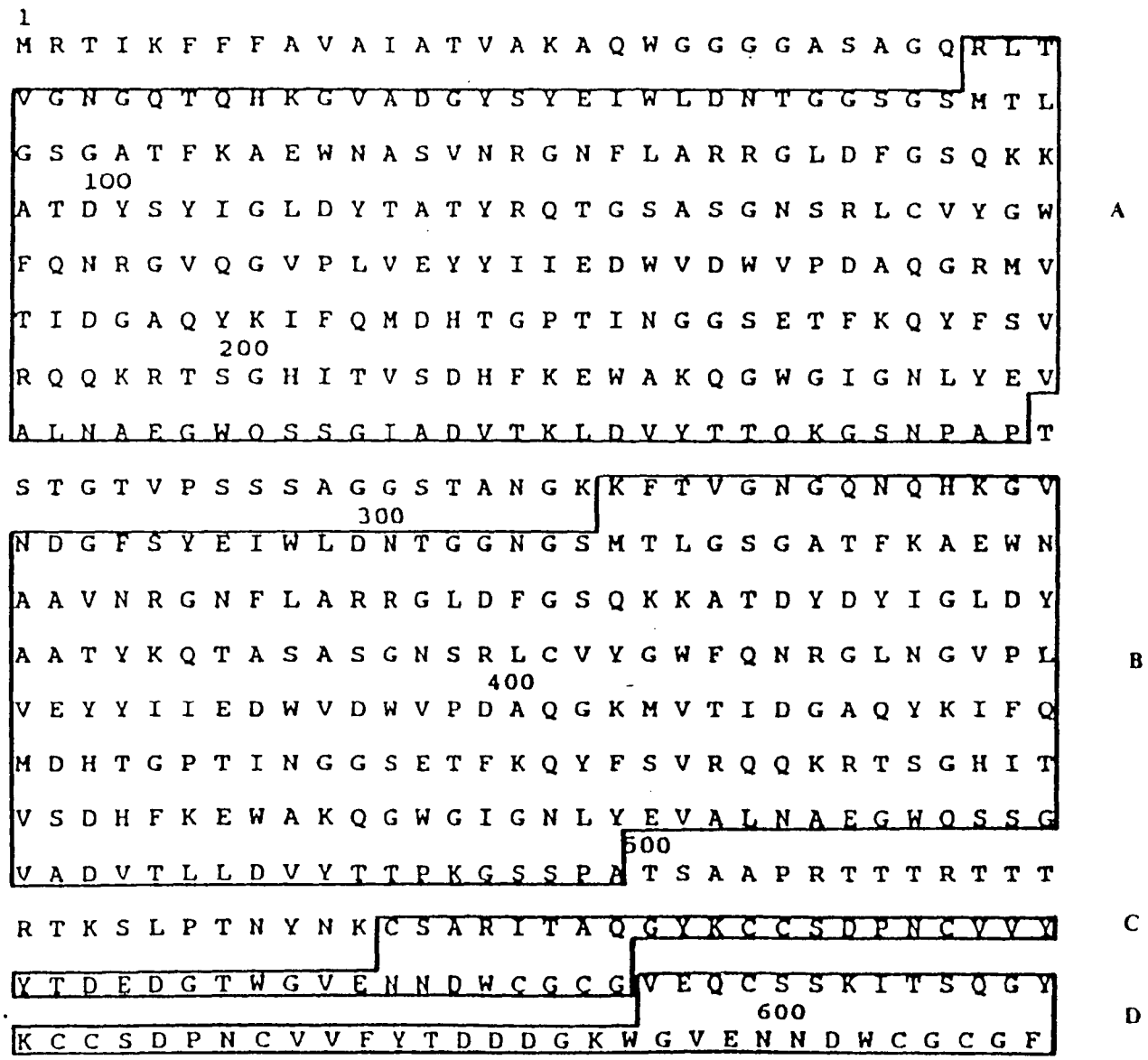


FIG. 5 (1/3)

```
XYLA   285 H K   G V N D G F S Y E I W L D N T G G N G S M T L G
XYLA    40 H K   G V A D G Y S Y E I W L D N T G G S G S M T L G
XYLBCA  66 E I   G V N G G Y D Y E L W K D   Y G N T S   M T L K
XYLAB   33 N E M G N H S G Y D Y E L W K D   Y G N T S   M T L N

XYLA       S G A T F K A E W N A A V N R G N F L A R R G L D F G
XYLA       S G A T F K A E W N A S V N R G N F L A R R G L D F G
XYLBCA     N G G A F S C Q W S N   I G N A L F     R K G K K F N
XYLAB      N G G A F S A G W N N   I G N A L F     R K G K K F D

XYLA       S Q K K A T D Y D Y I G L D Y A A T   Y K Q T A S A S
XYLA       S Q K K A T D Y S Y I G L D Y T A T   Y R Q T G S A S
XYLBCA     D T Q T Y K Q L G N I S V N Y D C N   Y   Q   P Y
XYLAB      S T R T H H Q L G N I S I N Y N A S   F   N   P S
XYLBC 72               T G S P F R T I N Y N A G V W   A   P N
```

16-02-95 9458128

FIG. 5(2/3)

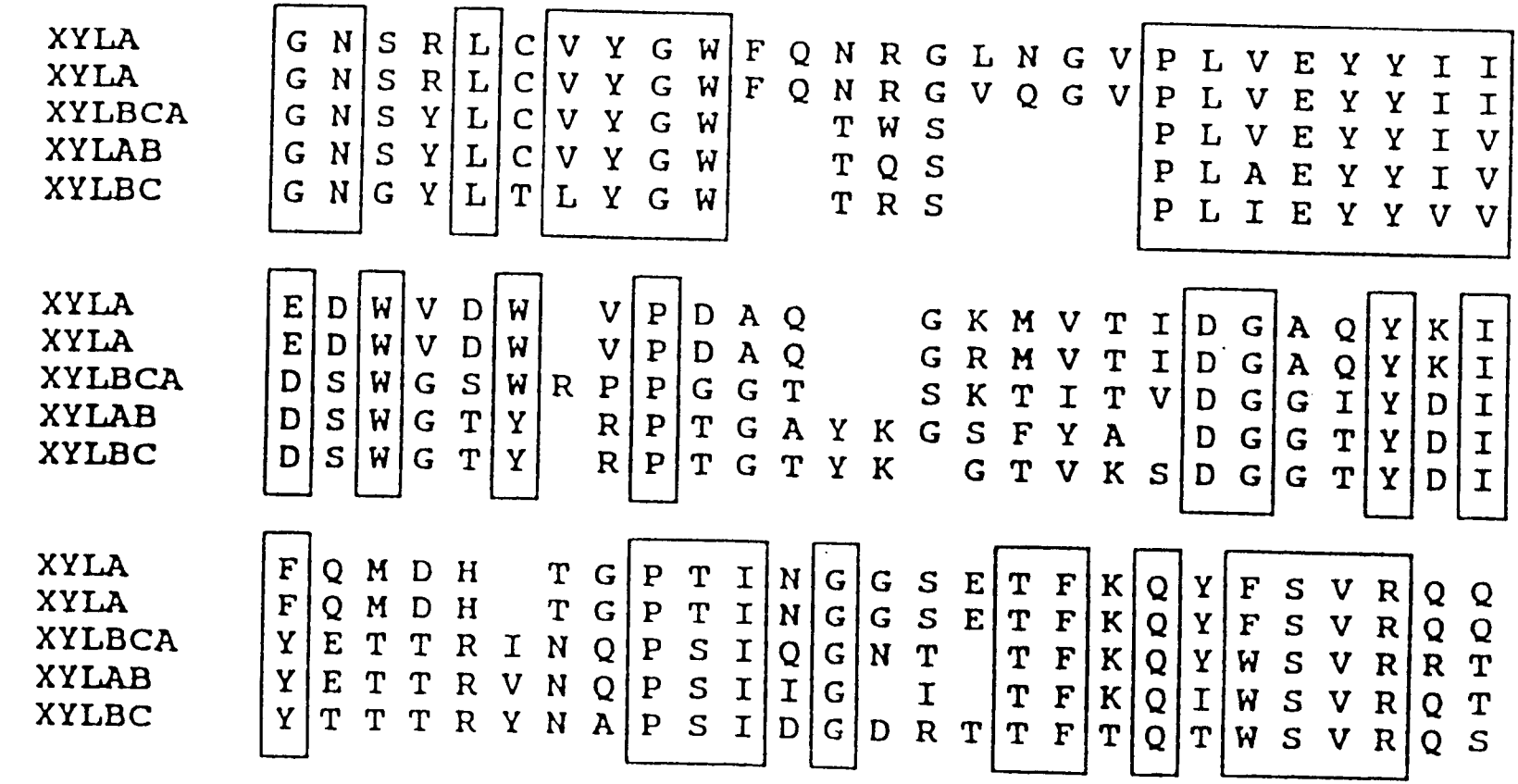


FIG. 5(3/3)

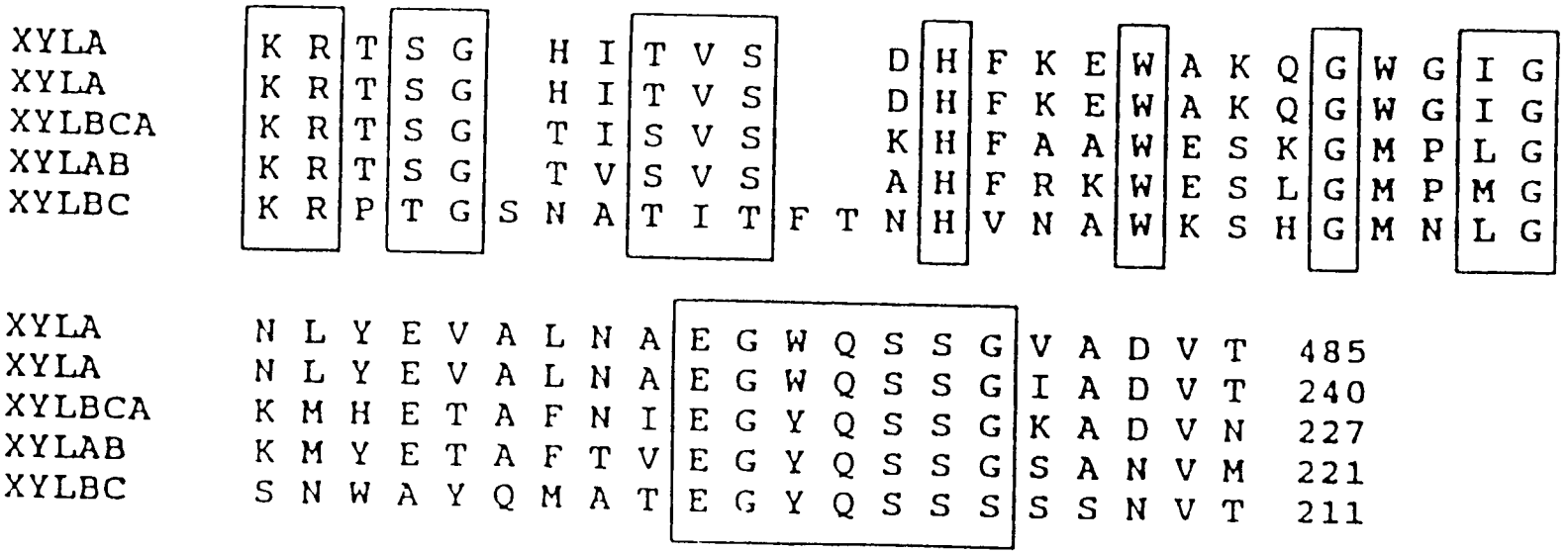
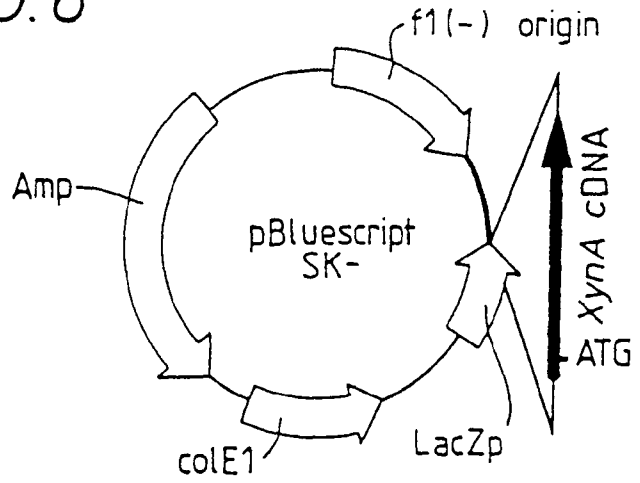

XYLA K R T S G H I T V S D H F K E W A K Q G W G I G
XYLA K R T S G H I T V S D H F K E W A K Q G W G I G
XYLBCA K R T S G T I S V S K H F A A W E S K G M P L G
XYLAB K R T S G T V S V S A H F R K W E S L G M P M G
XYLBC K R P T G S N A T I T F T N H V N A W K S H G M N L G
XYLA N L Y E V A L N A E G W Q S S G V A D V T 485
XYLA N L Y E V A L N A E G W Q S S G I A D V T 240
XYLBCA K M H E T A F N I E G Y Q S S G K A D V N 227
XYLAB K M Y E T A F T V E G Y Q S S G S A N V M 221
XYLBC S N W A Y Q M A T E G Y Q S S S S S N V T 211

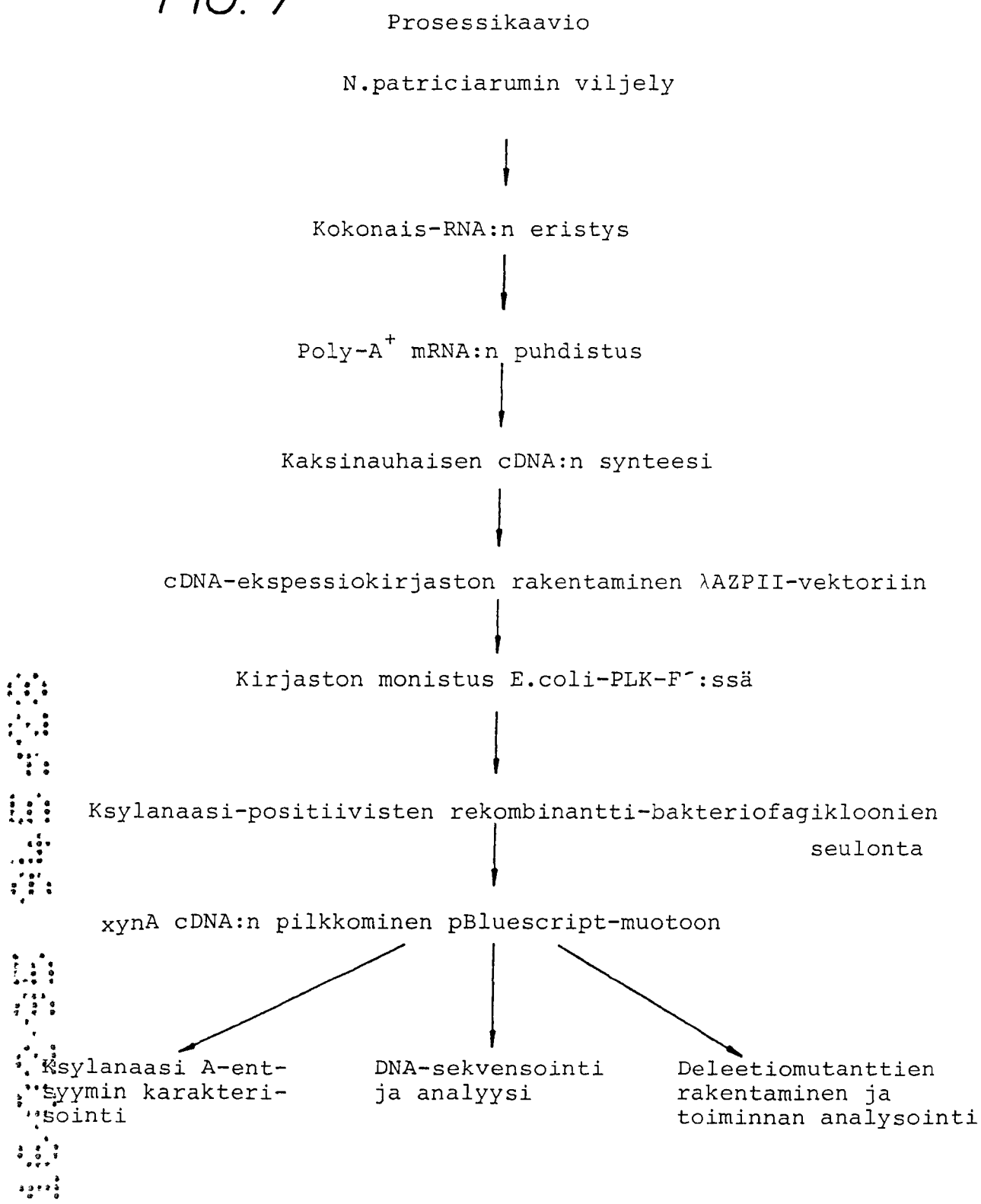
FIG. 6



pNX1Plasmidi

Vektori: pBluescript SK-
Insertti: 2-3 KB:n xynA cDNA N.Patriciarumista
XynA cDNA:n ekspressiota ohjaa LacZ-promoottori.
Ksylanaasi A:n synteesi käynnistyy sen omasta
translaation aloituskohdasta (ATG).

FIG. 7


Prosessikaavio
N.patriciarumin viljely
Kokonais-RNA:n eristys
Poly-A+ mRNA:n puhdistus
Kaksinauhaisen cDNA:n synteesi
cDNA-ekspessiokirjaston rakentaminen λAZPII-vektoriin
Kirjaston monistus E.coli-PLK-F´:ssä
Ksylanaasi-positiivisten rekombinantti-bakteriofagikloonien seulonta
xynA cDNA:n pilkkominen pBluescript-muotoon
Ksylanaasi A-entsyymin karakterisointi
DNA-sekvensointi ja analyysi
Deleetiomutanttien rakentaminen ja toiminnan analysointi