



Patentdirektoratet
TAASTRUP

(21) Patentansøgning nr.: 0532/90

(51) Int.Cl.5

C 07 K 7/10

(22) Indleveringsdag: 01 mar 1990

C 12 N 15/81

(24) Løbedag: 06 sep 1988

/(C 12 N 15/81,

C 12 R 1:865)

(41) Alm. tilgængelig: 01 mar 1990

(44) Fremlagt: 31 aug 1992

(86) International ansøgning nr.: PCT/DK88/00147

(86) International indleveringsdag: 06 sep 1988

(85) Videreførelsesdag: 01 mar 1990

(30) Prioritet: 07 sep 1987 DK 4638/87

(71) Ansøger: *Novo Nordisk A/S; Novo Alle; 2880 Bagsværd, DK

(72) Opfinder: Lars *Christiansen; DK

(74) Fuldmægtig: -

(54) Syntetiske gær-leader-peptider, DNA-sekvens, der indkoder sådanne leader-peptider, vektorer og transformerede gærstammer samt en fremgangsmåde til fremstilling af proteiner i gær

(56) Fremdragne publikationer

(57) Sammendrag:

532-90

Syntetiske leaderpeptider til effektiv sekretion af proteiner i gær har den almene formel (I)

X_1 -Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-(Glu-Glu-Ser-Leu-Ile) $_n$ - X_2 -Ala-Glu-Asn-Thr-Thr-Leu-Ala- X_3 - X_4 - X_5

k e n d e t e g n e t v e d , a t

X_1 er Ala eller Gln;

X_2 er Ile, en peptidkæde med op til 10 aminosyrerester eller en peptidbinding;

X_3 er en peptidkæde med op til 5 aminosyrerester eller en peptidbinding,

X_4 og X_5 er Lys eller Arg; og

n er 1 eller 0.

0532-90

Fig. 13

```

      12      24      36      48      60
AATTCATTCAA GAATAGTTCAAA CAAGAAGATTAC AACTATCAATT TCATACACAATA

      72      84      96     108     120
TAAACGATTAAA AGAATGAAGGCT GTTTCCTGGTT TTGTCCTTGATC GGATTGTGCTGG
           M K A V F L V L S L I G L C W

      132     144     156     168     180
GCCCAACCAGTT ACTGGCGATGAA TCATCTGTTGAG ATTCCGGAAGAG TCTCTGATCATC
A Q P V T G D E S S V E I P E E S L I I

      192     204     216     228     240
GCTGAAACACC ACTTTGGCTAAC GTCGCCATGGCT AAGAGATTGTT AACCAACACTTG
A E N T T L A N V A M A K R F V N Q H L

      252     264     276     288     300
TGCGGTTCCCAC TTGGTTGAAGCT TTGTACTTGGTT TGCGGTGAAAGA GGTTTCTTCTAC
C G S H L V E A L Y L V C G E R G F F Y

      312     324     336     348     360
ACTCCTAAGGCT GCTAAGGGTATT GTCGAACAATGC TGTACCTCCATC TGCTCCTTGTAC
T P K A A K G I V E Q C C T S - I C S L Y

      372     384     396     408     420
CAATTGGAAAAC TACTGCACTAG ACGCAGCCCGCA GGCTCTAGA
Q L E N Y C N |

```

DNA-sekvens af EcoRI-XbaI fragmentet, der
 indkoder signal-leader-p-insulinen fra
 pLaC 212spx3
 Den tilsvarende proteinsekvens er angivet
 med étbogstavsnomenklaturen under DNA-
 -sekvensen.

Den foreliggende opfindelse angår syntetiske gær-leader-peptider, DNA sekvenser, der indkoder sådanne leader-peptider, vektorer og transformerede gærstammer samt en fremgangsmåde til fremstilling af proteiner i gær.

5

Gærorganismer danner en række proteiner, der syntetiseres intracellulært, men som har en funktion uden for cellen. Sådanne ekstracellulære proteiner kaldes secernerede proteiner. Disse secernerede proteiner udtrykkes til at begynde med inde
10 i cellen i en precursor eller i en præprotein form, der indeholder en præsekvens, som sikrer, at det udtrykte produkt ledes effektivt gennem membranen i det endoplasmatiske retikulum (ER). Præsekvensen, der sædvanligvis kaldes et signalpeptid, spaltes almindeligvis fra det ønskede produkt under translokation. Når proteinet er kommet ind i sekretionsbanen,
15 transporteres det til Golgi-apparatet. Fra Golgi-apparatet kan proteinet følge forskellige veje, som fører til rum såsom cellevakuolen eller cellemembranen, eller det kan ledes ud af cellen for at blive secerneret til det ydre medium (Pfeffer, S.R. og Rothman, J.E. Ann. Rev. Biochem. 56 (1987) 829-852).
20

Der er gjort mange forsøg på i gær at udtrykke og secernere proteiner, der er heterologe i forhold til gær. I EP offentliggørelsesskrift nr. 0088632A beskrives en fremgangsmåde, ved
25 hvilken proteiner, der er heterologe i forhold til gær, udtrykkes, spaltes og secerneret ved transformation af en gærorganisme med en ekspressionsvektor indeholdende DNA, som indeholder det ønskede protein, og et signalpeptid, ved fremstilling af en kultur af den transformerede organisme,
30 dyrkning af kulturen og udvinding af proteinet fra dyrkningsmediet. Signalpeptidet kan være det ønskede proteins eget signalpeptid, et heterologt signalpeptid eller en hybrid af et naturligt og heterologt signalpeptid.

35 Et problem, der kunne opstå i forbindelse med brugen af signalpeptider, der er heterologe i forhold til gær, er, at det heterologe signalpeptid ikke sikrer effektiv translokation

og/eller spaltning efter signalpeptidet.

S. cerevisiae MF α (α -faktor) syntetiseres som en præ-pro-form med 165 aminosyrer, der indeholder et 19 aminosyrer langt signal- eller præpeptid efterfulgt af et 64 aminosyrer langt "leader-" eller propeptid, der omfatter tre N-forbundne glykolyseringssteder efterfulgt af (LysArg(Asp/Glu,Ala)₂₋₃ α -faktor)₄ (Kurjan, J. og Herskowitz, I. Cell 30 (1982) 933-943). Signalleader-delen af præ-pro-MF α er blevet anvendt i vid
10 udstrækning for at opnå syntese og sekretion af heterologe proteiner i S. cerevisiae.

Anvendelsen af signal/leader-peptider, der er homologe i forhold til gær, er bl.a. beskrevet i US patent nr. 4.546.082,
15 EP offentliggørelsesskrifterne nr. 0116201A, 0123294A, 0123544A, 0163592A og 0123289A og DK patentansøgning nr. 3614/83.

I EP 0123289A beskrives anvendelsen af S. cerevisiae a-faktor precursoren, medens der i i DK 3614/83 beskrives anvendelse af Saccharomyces cerevisiae PH05 signalpeptid til sekretion af fremmede proteiner.

I US patentansøgning nr. 4.546.082, EP 0116201A, 0123294A, 25 0123544A og 0163529A beskrives fremgangsmåder, ved hvilke α -faktor signal-leaderen fra Saccharomyces cerevisiae (MF α 1 eller MF α 2) anvendes i sekretionsprocessen af udtrykte heterologe proteiner i gær. Ved at sammenføje en DNA sekvens, der indkoder S. cerevisiae MF α signal/leadersekvensen, med 5'-enden af genet
30 for det ønskede protein, påvistes sekretion og spaltning af det ønskede protein.

Et antal secernerede proteiner dirigeres således, at de udsættes for et proteolytisk spaltningssystem, som kan spalte
35 peptidbindingen på carboxyenden af to på hinanden følgende basiske aminosyrer. Denne enzymatiske aktivitet indkodes i S. cerevisiae af KEX 2 genet (Julius, D.A. et al., Cell 37 (1984b), 1075). Det er nødvendigt at spalte produktet med KEX

2 genproduktet for at secernere aktiv S. cerevisiae parringsfaktor α (MF α eller α -faktor), men denne spaltning er ikke nødvendig for at secernere aktiv S. cerevisia parringsfaktor a.

- 5 Der er stigende evidens for, at sekretion er en tilbagefalds-
rute ("default route") ved lokalisering af proteiner i
eukaryote celler, idet konformationen af det protein, der skal
secernerer, eller af dets precursor er den essentielle para-
meter for processens effektivitet (Pfeffer, S.R. og Rothman,
10 J.E., Ann. Rev. Biochem. 56 (1987) 829-852); dette har udløst
en søgen efter mere effektive leadere end MF α -leaderen til
sekretion i S. cerevisiae.

Det er derfor et formål med den foreliggende opfindelse at
15 tilvejebringe mere effektive leadere end α -faktor-leaderen til
sekretion i gær af små proteiner.

I kompetitive forsøg af den af M. Egel-Mitani et al., Gene
(1988) (under trykning) beskrevne type med insulinprecursorer
20 af den type, der er beskrevet i EP patentansøgning nr. 163.529,
blev der fundet en begrænsende faktor i ekspressionen og
sekretionen af insulinprecursoren, hvilken er forårsaget af den
N-terminale ende af precursoren, sandsynligvis forårsaget af
lysin-argininsekvensen, som adskiller signal-leaderen fra
25 insulin-precursoren.

Et andet formål med den foreliggende opfindelse er at til-
vejebringe et sekretionssystem til heterologe proteiner i gær,
som sikrer en meget effektiv spaltning af det modne protein ved
30 hjælp af KEX 2 indkodet endopeptidase på en dibasisk sekvens,
og derved forhøje udbyttet af det modne produkt.

I et første aspekt angår den foreliggende opfindelse et hidtil
ukendt syntetisk leader-peptid med nedenstående almene formel
35 (I)

X_1 -Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-Ser-Leu-Ile- X_2 -Ala-Glu-Asn-Thr-Thr-Leu-Ala- X_3 - X_4 - X_5

hvor X_1 er Ala eller Gln;

5 X_2 er Ile, en peptidkæde med op til 10 aminosyrerester eller en peptidbinding;

X_3 er en peptidkæde med op til 5 aminosyrerester eller en peptidbinding,

X_4 og X_5 er Lys eller Arg; og

10 hvor alle aminosyrerester er naturligt forekommende L-former.

Eksempler på leader-peptider ifølge den foreliggende opfindelse er følgende:

15

Ala-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-Ser-Leu-Ile-Gly-Phe-Leu-Asp-Leu-Ala-Gly-Glu-Glu-Ile-Ala-Glu-Asn-Thr-Thr-Leu-Ala-Lys-Arg

20 Ala-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-Ser-Leu-Ile-Ile-Ala-Glu-Asn-Thr-Thr-Leu-Ala-Lys-Arg

Ala-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-Ser-Leu-Ile-Ile-Ala-Glu-Asn-Thr-Thr-Leu-Ala-Asn-Val-Ala-Met-

25 Ala-Lys-Arg og

Gln-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-Ser-Leu-Ile-Ile-Ala-Glu-Asn-Thr-Thr-Leu-Ala-Asn-Val-Ala-Met-Ala-Lys-Arg

30

I et andet aspekt angår den foreliggende opfindelse DNA-sekvenser, der indkoder leader-peptider med den ovenfor angivne formel (I).

35 Eksempler på sådanne DNA sekvenser er vist i fig. 9-10 og 12-13.

I et tredje aspekt angår den foreliggende opfindelse en

replikerbar gærvektor, som indeholder en DNA sekvens, der indkoder et leader-peptid med den ovenfor angivne formel (I), hvilken DNA-sekvens er anbragt opstrøms for en DNA-sekvens, der indkoder det ønskede produkt, og som er operabelt forbundet med
 5 egnede promotor- og signalsekvenser.

Promotoren kan være enhver DNA-sekvens, som udviser transcriptionel aktivitet i gær, og som kan stamme fra gener, der indkoder proteiner, som enten er homologe eller heterologe i
 10 forhold til gær. Promotoren stammer fortrinsvis fra et gen, der indkoder et protein, som er homologt i forhold til gær. En egnet promotor kan eksempelvis være Saccharomyces cerevisiae MF α 1 promotor. Andre egnede promotorer kan være S. cerevisiae TPI, ADH eller PGK promotorer.

15

Vektoren ifølge den foreliggende opfindelse indeholder sædvanligvis en signalpeptidsekvens, som er anbragt opstrøms for DNA-sekvensen, der indkoder det ovenfor angivne leader-peptid (I). Signalpeptidet kan være ethvert signalpeptid, der sikrer, at
 20 det udtrykte produkt ledes effektivt ind i cellens sekretionsbane. Signalpeptidet kan være et naturligt forekommende signalpeptid eller funktionelle dele deraf eller et syntetisk peptid. Egnede signalpeptider har vist sig at være α -faktor signalpeptidet, signalpeptidet fra musespytkirtel-amylase eller
 25 et modificeret carboxypeptidasesignal. Musespytkirtel-amylase-signalsekvensen er beskrevet af O. Hagenbüchle et al., Nature (1981) 289, 643-646. Carboxypeptidasesignalet er beskrevet af L.A. Valls et al., Cell 48 (1987) 887-897.

30 Gener, der indkoder det ønskede produkt, og signal/leader-sekvensen kombineres med fragmenter, der koder for en promotor, f.eks. TPI promotor (P_{TPI}), og en terminator, f.eks. TPI-terminatoren (T_{TPI}) (T. Alber og G. Kawasaki: Nucleotide Sequence of the Triose Phosphate Isomerase Gene of Saccharomy-
 35 ces cerevisiae J. Mol. Applied Genet 1 (1982) 419-434).

Ekspressionsplasmiderne omfatter endvidere 2 μ gærreplikationsgenerne REP1-3 og replikationsorigin og én eller flere

selekterbare markører, f.eks. Schizosaccharomyces pombe TPI genet, som beskrevet i EP offentliggørelsesskrift nr. 171.142.

Fremgangsmåderne, der anvendes til ligering af DNA-sekvenserne, der indkoder signal- og/eller leader-peptidet ifølge opfindelsen, promotor- og terminatorsekvenserne og indsætning i egnede gærplasmider indeholdende den nødvendige information til gærreplikation, er alle velkendte fremgangsmåder inden for dette fagområde.

10

I et fjerde aspekt angår den foreliggende opfindelse gærstammer, som transformeres med en vektor ifølge den foreliggende opfindelse.

15 I et femte aspekt angår den foreliggende opfindelse en fremgangsmåde til fremstilling af proteiner i gær, ved hvilken en gærstamme, som er transformeret med en vektor indeholdende en DNA-sekvens, der indkoder et leader-peptid med den ovenfor angivne formel (I), som er anbragt opstrøms for en DNA-sekvens, der indkoder det ønskede produkt, og som er operabelt forbundet med egnede promotor- og signalsekvenser, dyrkes i et egnet medium, hvorefter det secernerede produkt isoleres fra dyrkningsmediet.

25 Den ved fremgangsmåden ifølge opfindelsen anvendte gærorganisme kan være enhver egnet gærorganisme, som producerer store mængder af det ønskede protein. Foretrukne gærorganismer er Saccharomyces cerevisiae og Saccharomyces kluyveri, men der kan dog også anvendes andre gærorganismer, f.eks. Schizosaccharomyces pombe og Saccharomyces uvarum.

Fremgangsmåden ifølge opfindelsen giver mulighed for fremstilling af mange forskellige proteiner eller polypeptider i gær. Eksempler på sådanne forbindelser er human insulin-precursorer, insulinanalogprecursorer, som beskrevet i EP patentansøgning nr. 163.529, 194.864 og 214.826 og glucagon og aprotinin.

I én udførelsesform af den foreliggende opfindelse forbindes en DNA sekvens, der indkoder et leader-peptid med den ovenfor angivne formel (I), direkte med genet for det ønskede produkt. Under sekretionen spaltes leader-peptidet fra det udtrykte produkt, hvilket muliggør isolering af det ønskede produkt (f.eks. glucagon eller aprotinin) fra dyrkningsmediet, uden at der kræves yderligere in vitro omdannelse.

For at sikre en mere effektiv spaltning af leader-peptidet ved de C-terminale Lys-Arg-rester under sekretionen kan det være fordelagtigt at udtrykke det ønskede produkt med yderligere en N-terminal aminosyresekvens. Denne yderligere aminosyresekvens skal senere spaltes fra den resterende del af molekylet ved in vitro omdannelse. Med henblik på sådan in vitro spaltning kan det yderligere peptid forsynes med et selektivt spaltningssted beliggende N-terminalt i forhold til det ønskede produkt. Hvis det ønskede produkt ikke indeholder methionin, er cyanogenbromid-spaltning ved methionin, som er nabostillet med det ønskede protein, effektiv. Dette omfatter også arginin- eller lysin-spaltning ved en arginin eller lysin, som er nabostillet med det ønskede protein, med en trypsin-lignende protease.

Endelig kan den yderligere peptidkæde fraspaltes in vitro ved andre egnede proteolytiske enzymer, forudsat at den indeholder et spaltningssted for sådanne enzymer.

I nærværende sammenhæng betegner udtrykket "proteiner, som er heterologe i forhold til gær" proteiner, som ikke naturligt dannes af gærorganismen. "Sekretion" betegner udledning af et udtrykt produkt gennem cellevæggen ud i mediet eller i det mindste gennem cellemembranen. "Spaltning" betegner cellulær spaltning af signal/leader-peptidet fra det modne protein til fremstilling af det heterologe protein uden ledsagelse af nogen del af signal/leader-peptidet. "Signal/leader-peptid" betegner en præ-pro-region, der sikrer, at det udtrykte produkt ledes ind i cellens sekretionsbane, og består af en N-terminal signalsekvens af hydrofob sammensætning efterfulgt af en hydrofob leader-region. "Modent protein" betegner proteiner,

der indkodes af DNA-sekvensen indsat i ekspressionsvektoren. Det modne protein kan være et aktivt protein eller polypeptid eller en precursor derfor, som yderligere omdannes til det aktive slutprodukt ved in vitro omdannelse.

5

Den foreliggende opfindelse er nærmere belyst under henvisning til den medfølgende tegning, hvor:

- Fig. 1 viser pMT743 plasmidet og konstruktionen af plasmid pT7 α MI3,
10
- Fig. 2 viser konstruktionen af plasmiderne pT7-178 og pLaC178,
- 15 Fig. 3 viser konstruktionen af plasmiderne pT7-193, pT7-194 og pT7-195,
- Fig. 4 viser konstruktionen af plasmiderne pT7-200 og pLaC200,
20
- Fig. 5 viser konstruktionen af plasmidet pT7-196,
- Fig. 6 viser konstruktionen af plasmidet pLaC212spx3,
- 25 Fig. 7 viser DNA-sekvensen af EcoRI-Xba fragmentet fra pT7 α MI3 og det tilsvarende protein angivet med étbogstavs-nomenklaturen nedenfor,
- Fig. 8 viser DNA-sekvensen af EcoRI-XbaI fragmentet fra pT7-178 og det tilsvarende protein angivet med étbogstavs-nomenklaturen nedenfor,
30
- Fig. 9 viser DNA-sekvensen af EcoRI-XbaI fragmentet fra pT7-193 og det tilsvarende protein angivet med étbogstavs-nomenklaturen nedenfor,
35
- Fig. 10 viser DNA-sekvensen af EcoRI-XbaI fragmentet fra pT7-194 og det tilsvarende protein angivet med étbog-

stavs-nomenklaturen nedenfor,

Fig. 11 viser DNA-sekvensen af EcoRI-XbaI fragmentet fra pT7-200 og det tilsvarende protein angivet med étbogstavs-nomenklaturen nedenfor,

Fig. 12 viser DNA-sekvensen af EcoRI-XbaI fragmentet fra pLaC212spx3 og det tilsvarende protein angivet med étbogstavnomenklaturen nedenfor,

10

Fig. 13 viser konstruktionen af plasmiderne pKFN371, pT7212spx-3017 og pT7 α -0174, og

15

Fig. 14 viser konstruktionen af plasmiderne pLaC212spx3-0174 og pKFN216.

Den foreliggende opfindelse belyses i nedenstående eksempler, der viser fremstillingen af visse insulinprecursorer.

20 1. Plasmider og DNA materialer

Alle ekspressionsplasmider er af C-POT-typen. Sådanne plasmider er beskrevet i EP offentliggørelsesskrift nr. 171.142 og er
25 ejendommelige ved at indeholde S. pombe triosefosfatisomerase-genet (POT) med henblik på plasmidstabilisering. Et plasmid, der indeholder POT genet, er tilgængeligt fra en deponeret E. coli stamme (ATCC 39685). Plasmiderne indeholder desuden S. cerevisiae triosefosfat-isomerasepromotor og -terminator (P_{TP1}
30 og T_{TP1}). De er identiske med pMT743 (M. Egel-Mitani et al. Gene (1988) (under trykning) (se fig. 1) bortset fra regionen defineret ved EcoRI-XbaI-restriktionsstedet, der indkoder en signal/leader/insulinprecursorsekvens.

35 Samlingen af de signal/leader/insulinprecursor-kodende fragmenter, som er beskrevet i denne sammenhæng, har af praktiske grunde fundet sted i et E. coli plasmid pT7 α MI3. Dette plasmid (se fig. 1) er et pBR322 derivat, hvori EcoRI stedet er

elimineret ved religering af plasmidet, der er skåret med EcoRI og gjort stump-endet, med Klenow polymerase og dNTP, og hvori BamHI-SphI fragmentet er udskiftet med 2,2 kb SphI-BamHI fragmentet fra ovennævnte pMT743. Sidstnævnte fragment indeholder en sekvens, der indkoder en insulinprecursor B(1-29)-Ala-Ala-Lys-A(1-21) forbundet med MF α leaderen og -signalet og S. cerevisiae triosefosfat-isomerasepromotor- og terminator-sekvenserne (se fig. 1). I dette SphI-Bam HI fragment er sekvensen, der indkoder signal/leader/-insulinprecursoren indeholdt i 0,5 kb EcoRI-XbaI fragmentet. Det almene princip for konstruktionsarbejdet beskrevet i de følgende eksempler er at udskifte MF α -leadersekvensen med en leader-sekvens ifølge den foreliggende opfindelse. Signalsekvensen samt sekvensen for det ønskede produkt kan også udskiftes. DNA-sekvensen af EcoRI-XbaI fragmentet af pT7 α MI3 er vist i fig. 7.

Adskillige af konstruktionerne omfatter endvidere fragmenter fra S. kluyveri α parringspheromon-genet (M. Egel-Mitani og M. Trier Hansen (1987) Nucl. Acid Res. 15, 6306).

Endelig er der anvendt en række syntetiske DNA fragmenter.

De syntetiske DNA fragmenter syntetiseres på et automatisk DNA-synteseapparat (Applied Biosystems Model 380A), idet der anvendes phosphoramiditkemi og kommercielt tilgængelige reagenser (S.L. Beaucage og M.H. Caruthers (1981) Tetrahedron Letters 22, 1859-1869). Oligonucleotiderne renses ved polyacrylamidgel-elektroforese under denaturerende betingelser.

Før sammenføjning blev de komplementære DNA-enkeltstrengene kinasebehandlet med T4 polynucleotidkinase og ATP.

Fragmenterne, som fremkommer ved denne fremgangsmåde, er følgende:

NOR243/244: ACCGGTGACGAAAGCTTCGTCTGA
 TGGCCACTGCTTTCGAAGCAGCTTTAA

NOR512/513: AATTCCTGAGGAATCTTT
GGACTCCTTAGAAACTAG

5 NOR308/309: AATTGATATCG
CTATAGC

NOR245/246: CTGAAAACTCCACTTTGGCTAAGAG
GACTTTTGGAGGTGAAACCGATTCTCTAA

10 NOR348/350: TAACGTCGCCATGGCTAAGAGATTCGTTAAC
GCAGCGGTACCGATTCTCTAAGCAATTG

NOR217/218: CAACCAGTCAC
CCGGGTTGGTCAGTGGC
15

Alle andre anvendte fremgangsmåder og materialer er gængse fra den kendte teknik (T. Maniatis et al. (1982) Molecular Cloning. Cold Spring Harbor Press).

20 Følgende eksempler 1-4 viser ekspressionen og sekretionen af insulinprecursorer af typen B(1-29)-Ala-Ala-Lys-A(1-21) ved forskellige signal-leader-sekvenser ifølge den foreliggende opfindelse, og eksempel 5 viser ekspressionen og sekretionen af en insulinprecursor af ovennævnte type, men hvori aminosyrerest-

25 sten i position B(27) er erstattet med Arg, og aminosyreresten i position A(21) er erstattet med Gly. Som anvendt i den følgende tekst betyder B(1-29) en forkortet B-kæde af human insulin fra B(1)Phe til B(29)Lys, og A(1-21) betyder A-kæden af human insulin.

Eksempel 1Konstruktion af pLAC178

- 5 4,5 kb SalI-XbaI fragmentet, 1317 bp SalI-HincII fragmentet og 333 bp EcoRI*-XbaI fragmentet fra pT7 α MI13 blev forbundet med det syntetiske fragment NOR243/244:

```

10          ACCGGTGACGAAAGCTTCGTCGA
          TGGCCACTGCTTTCGAAGCAGCTTTAA

```

hvilket resulterede i pT7-172.

- 15 5,6 kb EcoRI-XbaI, 172 bp EcoRI-PvuII og 278 bp HinfI-XbaI fragmenterne fra pT7-172 blev forbundet med det syntetiske fragment NOR 245/246:

```

20          CTGAAAACACCACTTTGGCTAAGAG
          GACTTTTGTGGTGAAACCGATTCTCTAA

```

hvilket resulterede i pT7-178.

- 25 Udskiftning af SphI-BamHI fragmentet fra pMT743 med 2,2 kb SphI-BamHI fragmentet fra pT7-178 resulterede i pLaC178. Konstruktionen af pLaC178 er vist i fig. 2.

DNA-sekvensen af EcoRI-XbaI fragmentet fra pT7-178 er vist i fig. 8.

30 Konstruktion af pLaC193-194

5,6 kb EcoRI-XbaI og 203 bp PvuII-XbaI fragmentet fra pT7 α MI3 blev forbundet med det syntetiske fragment NOR308/309:

```

35          AATTGATATCG
          CTATAGC

```

Fra det resulterende plasmid blev 210 bp EcoRV-XbaI fragmentet

samlet med 138 bp EcoRI-MaeIII og 5,6 bp EcoRI-XbaI fragmentet fra pT7-178, og de nedenfor angivne fragmenter fra S. kluyveri MF α -genet:

- 5 74 bp MaeIII-EcoRV resulterede i pT7-193
 47 bp MaeIII-Sau3A resulterede i pT7-194

De understregede restriktionsender blev udfyldt med Klenow polymerase og dNTP. Konstruktionen af disse plasmider er vist
 10 i fig. 3.

Udskiftning af SphI-BamHI fragmentet af pMT743 med SphI-BamHI fragmenter af pT7-193-195-plasmider resulterede i plasmiderne pLaC193-194.

15

Plasmidet pLaC193 indkoder følgende leader-sekvens:

Ala-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-
 Ser-Leu-Ile-Gly-Phe-Leu-Asp-Leu-Ala-Gly-Glu-Glu-Ile-Ala-Glu-
 20 Asn-Thr-Thr-Leu-Ala-Lys-Arg.

Signalsekvensen er en codon-optimeret S. cerevisiae signalsekvens (se Michi Egel Mitani et al., (1988) supra). DNA-sekvensen af EcoRI-XbaI fragmentet af pT7-193 er vist i fig. 9.

25

Plasmidet pLaC194 indkoder følgende leader-sekvens:

Ala-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-
 Ser-Leu-Ile-Ile-Ala-Glu-Asn-Thr-Thr-Leu-Ala-Lys-Arg.

30

Signalsekvensen er som ovenfor angivet. DNA-sekvensen af EcoRI-XbaI fragmentet af pT7-194 er vist i fig. 10.

35

Eksempel 2Konstruktion af pLaC200

- 5 213 bp EcoRI-DdeI og 5,9 kb HpaI-EcoRI fragmentet fra pT7-194 blev samlet med det syntetiske fragment NOR348/350:

TAACGTCGCCATGGCTAAGAGATTCGTTAAC
GCAGCGGTACCGATTCTCTAAGCAATTG

10

resulterede i pT7-200. Udskiftning af SphI-BamHI fragmentet i pMT743 med 2,2 kb SphI-BamHI fragmentet fra pT7-200 resulterede i pLaC200. Konstruktionen af pLaC200 er vist i fig. 4.

- 15 Plasmidet pLaC200 indkoder følgende leader-sekvens:

Ala-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-Ser-Leu-Ile-Ile-Ala-Glu-Asn-Thr-Thr-Leu-Ala-Asn-Val-Ala-Met-Ala-Lys-Arg.

20

Signalsekvensen er den samme som i eksempel 1. DNA-sekvensen af EcoRI-XbaI fragmentet fra pT7-200 er vist i fig. 11.

Eksempel 3

25

Konstruktion af pLaC212spx3

Plasmidet pT7-178 blev mutageniseret in vitro ved dobbelt-primermetoden, idet der anvendes det syntetiske DNA NOR544:

30

3' CTTACTTTCAGAAGATAAAAATGACG 5'

i alt væsentligt som beskrevet af K. Norris et al. (1983) Nucl. Acid Res. 11 5103-5112 (fig. 5).

35

Det resulterende plasmid indeholder nu et XmnI spaltningssted, hvilket muliggør isolering af et 82 bp EcoRI-XmnI fragment, som sammen med 37 bp XmnI-ApaI fragmentet fra mPSA104 (0.

Hagenbüchle et al. Nature (1981) 289, 643-646), der indkoder musespytkirtel-amylase-signalpeptid-aminosyreresterne tre til femten, blev samlet ved en ligering, som yderligere inkluderede 251 bp MspI-Xba og 5,6 bp EcoRI-Xba fragmenterne fra pT7-178 og
 5 det syntetiske DNA fragment NOR217/218:

CAACCAGTCAC
 CCGGGTTGGTCAGTGGC

10 hvilket resulterede i plasmidet pT7-196 (fig. 5), som indeholder et modificeret musespytkirtel-amylasesignal (aminosyrerester 3-15).

Udskiftning af SphI-BamHI fragmentet fra pMT743 med 2,2 kb
 15 SphI-BamHI fragmentet af pT7-196 resulterede i pLaC196.

Det modificerede musespyt-amylasesignalpeptid blev yderligere mutageniseret ved udskiftning af XmnI-ApaI fragmentet af pLaC196 med de syntetiske adaptorer NOR520T eller 521T (fig.
 20 6a) fremstillet ud fra

NOR520: tcccatctgttttcttgttattgtccttgatcggattctgctGGGCCCA

NOR521: tcttcttgttattgtccttgatcggattctgctGGGCCCA

25 hvor nucleotiderne angivet med små bogstaver blev inkorporeret fra puljer sammensat af 91% af det angivne nukleotid og 3% af hver af de andre nukleotider. NOR520 og 521 blev autosammenføjet (understregede 3'-ende), forlænget med Klenow-polymerase
 30 og dNTP, hvilket giver dobbeltstrengede DNA fragmenter på henholdsvis 90 og 72 bp. Disse fragmenter blev spaltet med ApaI, og de fremkomne fragmenter NOR521T på 47 bp og NOR521T på 38 bp blev oprenset ved polyacrylamid-gelelektroforese.

35 Plasmider fremkommet ved denne fremgangsmåde blev isoleret fra transformanter af en E. coli stamme og hver især transformeret ind i S. cerevisiae stammen MT663. De fremkomne gærtransformanter blev screenet for insulinprecursorsekretion ved koloni-

immunblotting med en monoklonal muse-antiinsulinprecursor og
 detekteret med peberrodsperoxidase-konjugeret kanin-antimus
 IgG. De transformanter, som giver den stærkeste respons, blev
 analyseret yderligere. Den bedste af disse indeholdt plasmidet
 5 pLaC196spx3, hvori signalpeptidet var ændret til:

MKAVFLVLSLIGFCWA

Den nøjagtige nucleotidsekvens bestemt ved Maxam-Gilbert
 10 sekventering fra EcoRI stedet mod XbaI er angivet i fig. 12.

spx3 signalpeptidet er i ovennævnte sammenhæng omtrent en
 faktor to bedre end det oprindelige signalpeptid i pLaC196,
 bestemt ud fra de opnåede secernerede insulinprecursor-niveauer.
 15

Overraskende nok er spx3, som stammer fra NOR520T mutagenese,
 to kodoner kortere end forventet, og den eneste forklaring på
 dette må være længdeheterogenitet i de oprindelige enkeltstren-
 20 ge af NOR520 eller et polymerasespring i forlængelsesreaktionen
 undervejs til NOR520T.

Udskiftning af MaeIII-XbaI fragmentet, der indkoder leader-
 insulinprecursordelen af pLaC196spx3 med det analoge MaeIII-
 25 XbaI fragment af pLaC200, resulterede i pLaC212spx3 (fig. 6b).

Plasmid pLaC212spx3 indkoder følgende leadersekvens:

Gln-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-
 30 Ser-Leu-Ile-Ile-Ala-Glu-Asn-Thr-Thr-Leu-Ala-Asn-Val-Ala-Met-
 Ala-Lys-Arg.

Signalsekvensen er det muterede musespytkirtelamylasesignal.
 DNA-sekvensen af EcoRI-XbaI fragmentet af pLaC212spx3 er vist
 35 i fig. 12.

Eksempel 4

DNA, der indkoder en precursor af en insulinanalog med henholdsvis Gln i position B(12), Arg i position B(27) og Gly i position A(21), blev afledt af plasmidet pKFN126 (Markussen et al. (1987) Protein Engineering: 1, 215-223) ved at udskifte 90 bp Stul-XbaI fragmentet med følgende syntetiske DNA:

```

10  CCCAAGGCTGCTAAGGGTATTGTCGAACAATGCTGTACCTCCATCTGCTCCTTGTA
    GGGTTCCGACGATTCCCATAACAGCTTGTTACGACATGGAGGTAGACGAGGAACAT

```

```

    CCAATTGGAAACTACTGCGGTTAAT
    GGTTAACCTTTTGATGACGCCAATTAGATC

```

15 Fra det resulterende plasmid pKFN371 blev 124 bp HindIII-XbaI fragmentet isoleret og anvendt til udskiftning af det analoge HindIII-XbaI fragment fra pT7212spx3 og pT7 α MI3 (se fig. 13). Herved blev DNA-sekvensen, der indkoder ovennævnte insulinanalog, ændret til en sekvens, der koder for en insulinanalog, hvori B(27) er Arg, og A(21) er Gly. Ved at udskifte SphI-BamHI fragmentet fra pMT743 med 2,2 kb SphI-BamHI fragmentet fra plasmiderne pT7212spx3-0174 og pT7 α -0174, som opstod ved disse konstruktioner, konstrueredes henholdsvis pLaC212spx3-0174 og pKFN216 (se fig. 14). I disse plasmider medierer henholdsvis 25 212spx3 og S. cerevisiae MF α leaderen sekretionen af insulinanalog.

Eksempel 5

30 Plasmider fremstillet som beskrevet ovenfor blev transformeret til S. cerevisiae stammer, som bærer deletioner i TPI genet, ved udvælgelse til vækst på glucose.

De transformerede gærstammer blev dyrket på YPD-medium (Sherman, F. et al., Methods in Yeast Genetics, Cold Spring Harbor Laboratory 1981). For hver stamme blev 6 individuelle 5 ml kulturvæsker rystet ved 30°C, indtil de nåede en OD₆₀₀ på ca. 15 (ca. 48 timer) Efter centrifugering blev supernatanten fjernet

med henblik på HPLC-analyse, ved hvilken koncentrationen af secerneret insulinprecursor blev målt ved en fremgangsmåde beskrevet af Leo Snel et al. (1987) *Chromatographia* 24, 329-332.

- 5 Ekspressionsniveauerne for forskellige insulinprecursorer er opsummeret i følgende tabeller I og II.

I tabel 1 og 2 er ekspressionsniveauet af insulinprecursorerne under anvendelse af en leadersekvens ifølge opfindelsen angivet
 10 i procent af ekspressionsniveauet af den samme insulinprecursor udtrykt og secerneret i den samme gærstamme, men under anvendelse af MF α signal/leadersekvensen.

Tabel I

15

Ekspressionsniveauer for insulinprecursor B(1-29)-Ala-Ala-Lys-A(1-21) i gærtransformanter.

	Plasmid	Ekspressionsniveau*
20	pMT743	100%
	pLaC193	160%
	pLaC194	200%
	pLaC200	170%
	pLaC212spx3	120%

25

Ekspressionsniveauerne er angivet i procent af ekspressionsniveauet for pMT743, som arbitrært fastsættes til 100%.

Tabel II

Ekspressionsniveauer for insulinprecursorer, hvori B(27) er
5 Arg, og A(21) er Gly.

Plasmid	Ekspressionsniveau**
pKFN216	100%
pLaC212spx3-0174	150%

10

**Ekspressionsniveauerne er angivet i procent af ekspressions-
niveauet for pKFN216, som arbitrært fastsættes til 100%.

PATENTKRAV

1. Syntetisk leaderpeptid til sekretion i gær, k e n -
d e t e g n e t ved, at det har følgende almene formel (I)

5

X_1 -Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-Ser-
Leu-Ile- X_2 -Ala-Glu-Asn-Thr-Thr-Leu-Ala- X_3 - X_4 - X_5

k e n d e t e g n e t ved, at

10

X_1 er Ala eller Gln;

X_2 er Ile, en peptidkæde med op til 10 aminosyrerester
eller en peptidbinding;

X_3 er en peptidkæde med op til 5 aminosyrerester eller en
peptidbinding,

15

X_4 og X_5 er Lys eller Arg; og hvor alle aminosyrerester
er naturligt forekommende L-former.

2. Leaderpeptid ifølge krav 1, k e n d e t e g n e t
ved, at X_2 er Ile, n er 1, og X_1 , X_3 , X_4 og X_5 har de i krav 1
20 angivne betydninger.

3. Leaderpeptid ifølge krav 1, k e n d e t e g n e t
ved, at X_2 er peptidkæden Gly-Phe-Leu-Asp-Leu-Ala-Gly-Asp-Asp-
Ile, n er 1, og X_1 , X_3 , X_4 og X_5 har de i krav 1 angivne
25 betydninger.

4. Leaderpeptid ifølge krav 1, k e n d e t e g n e t
ved, at det har følgende peptid-sekvens:

30 Ala-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-
Ser-Leu-Ile-Gly-Phe-Leu-Asp-Leu-Ala-Gly-Glu-Glu-Ile-Ala-Glu-
Asn-Thr-Thr-Leu-Ala-Lys-Arg.

5. Leaderpeptid ifølge krav 1, k e n d e t e g n e t
35 ved, at det har følgende peptidsekvens:

Ala-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-
Ser-Leu-Ile-Ile-Ala-Glu-Asn-Thr-Thr-Leu-Ala-Lys-Arg

6. Leaderpeptid ifølge krav 1, kendetegnet ved, at det har følgende peptidsekvens:

Ala-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-Glu-
 5 Ser-Leu-Ile-Ile-Ala-Glu-Asn-Thr-Thr-Leu-Ala-Asn-Val-Ala-Met-
 Ala-Lys-Arg.

7. Leaderpeptid ifølge krav 1, kendetegnet ved, at det har følgende peptidsekvens:

10

Gln-Pro-Val-Thr-Gly-Asp-Glu-Ser-Ser-Val-Glu-Ile-Pro-Glu-
 Glu-Ser-leu-Ile-Ile-Ala-Glu-Asn-Thr-Thr-Leu-Ala-Asn-Val-
 Ala-Met-Ala-Lys-Arg.

15 8. DNA-sekvens, kendetegnet ved, at den indkoder et leaderpeptid med formlen (I) som beskrevet i krav 1.

9. DNA-sekvens ifølge krav 8, kendetegnet ved, at den har nucleotidsekvensen fra nucleotid 133 til 246 i
 20 fig. 9.

10. DNA-sekvens ifølge krav 8, kendetegnet ved, at den har nucleotidsekvensen fra nucleotid 133 til 218 i fig.
 25 10.

11. DNA-sekvens ifølge krav 8, kendetegnet ved, at den har nucleotidsekvensen fra nucleotid 133 til 233 i fig.
 11.

30

12. DNA-sekvens ifølge krav 8, kendetegnet ved, at den har nucleotidsekvensen fra nucleotid 124 til 223 i fig.
 12.

35 13. Replikerbar gærvektor, kendetegnet ved, at den indeholder en DNA-sekvens, der indkoder et leaderpeptid med formlen (I) ifølge krav 1, anbragt opstrøms for en DNA-sekvens, der indkoder det ønskede produkt, og som er operabelt

forbundet med promotor- og signalsekvenser.

14. Gærstamme, k e n d e t e g n e t ved, at den er transformeret med en vektor ifølge krav 13.

5

15. Gærstamme ifølge krav 14, k e n d e t e g n e t ved, at den tilhører arten Saccharomyces cerevisiae.

16. Fremgangsmåde til fremstilling af proteiner i gær,
10 k e n d e t e g n e t ved, at en gærstamme, som er transformeret med en vektor ifølge krav 13, dyrkes i et egnet dyrkningsmedium, og det ønskede proteinprodukt isoleres fra dyrkningsmediet.

15 17. Fremgangsmåde ifølge krav 16, k e n d e t e g n e t ved, at gærstammen tilhører arten Saccharomyces cerevisiae.

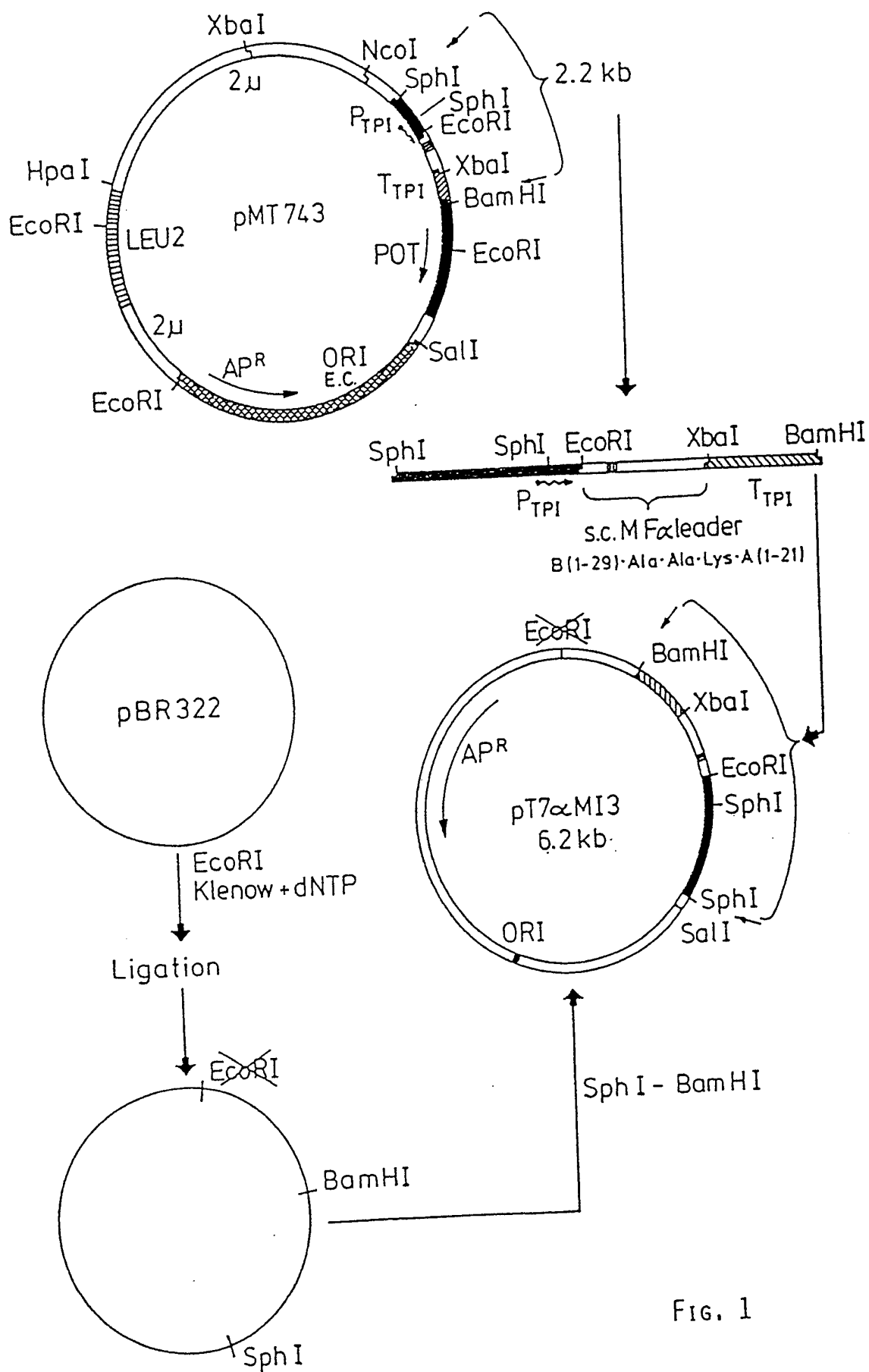


FIG. 1

FIG. 2

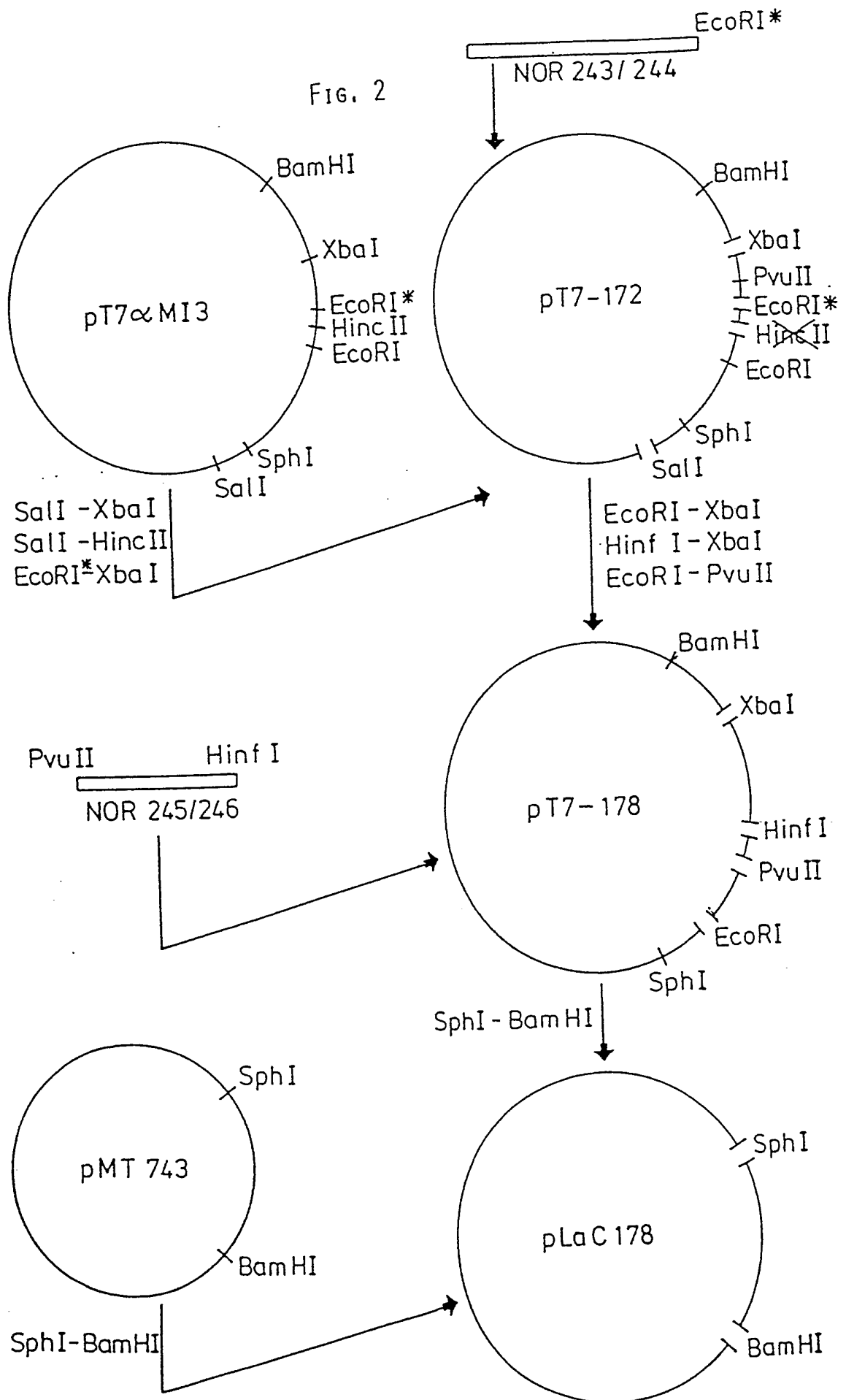


FIG. 3

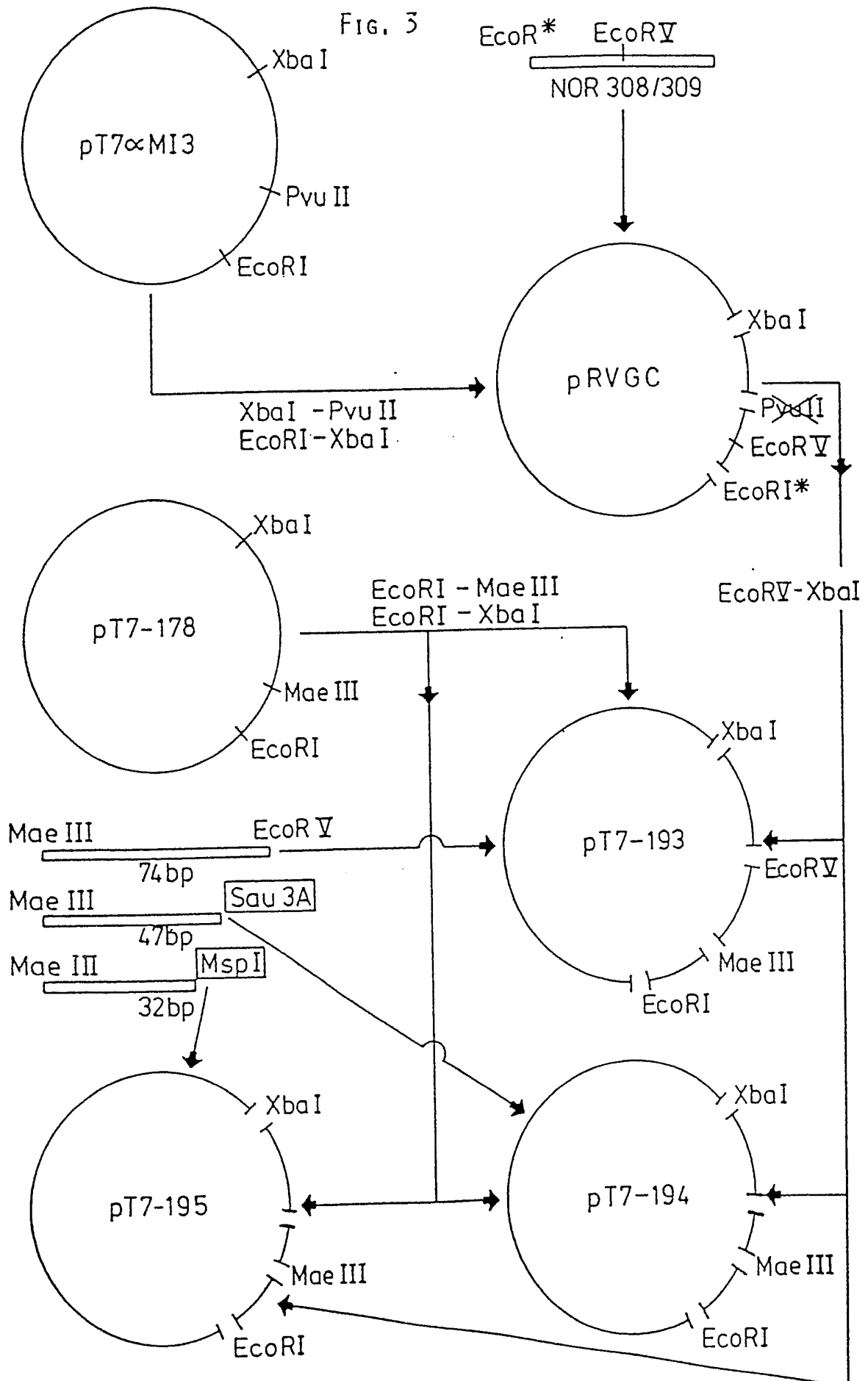


FIG. 4

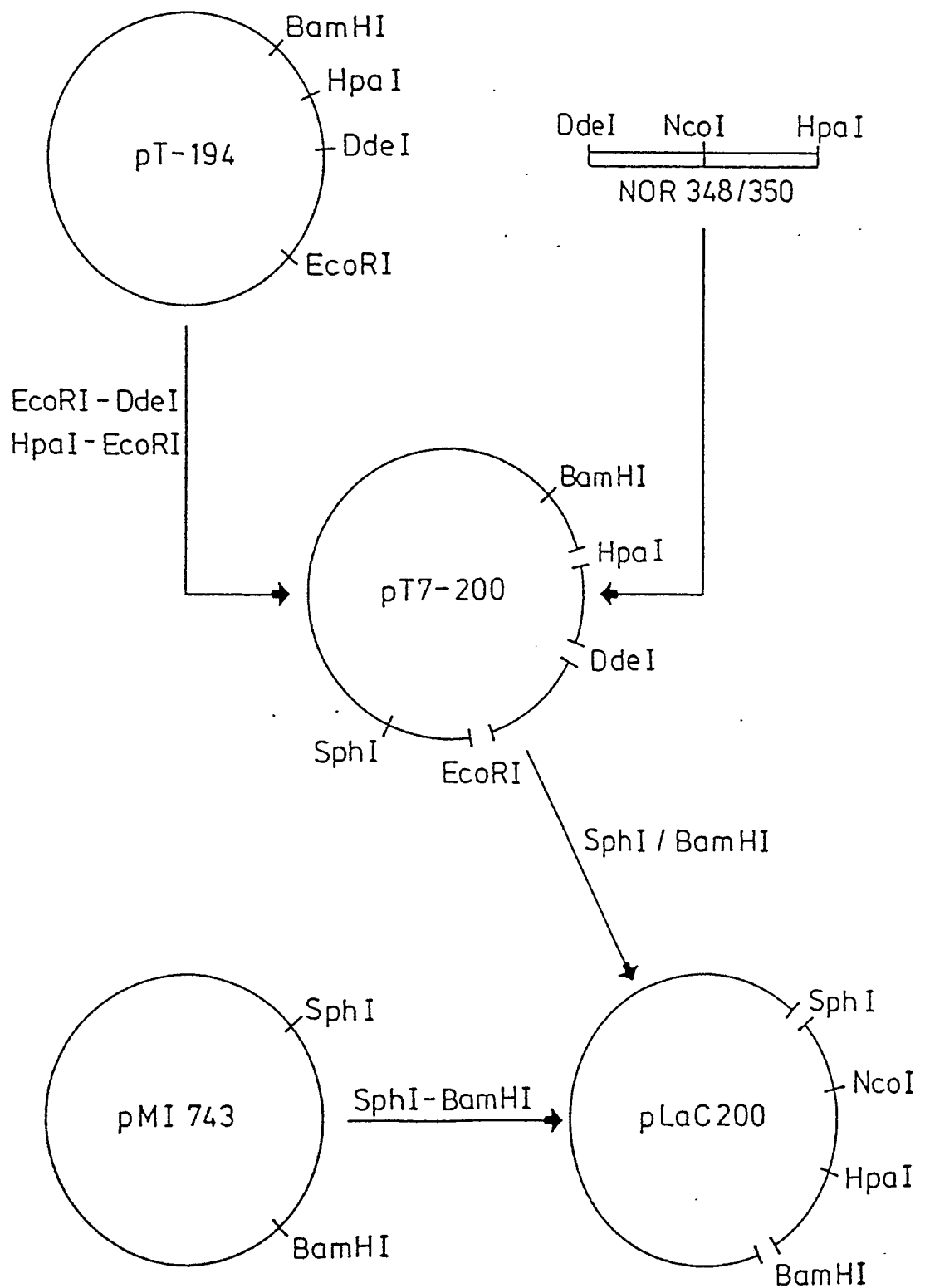


FIG. 5

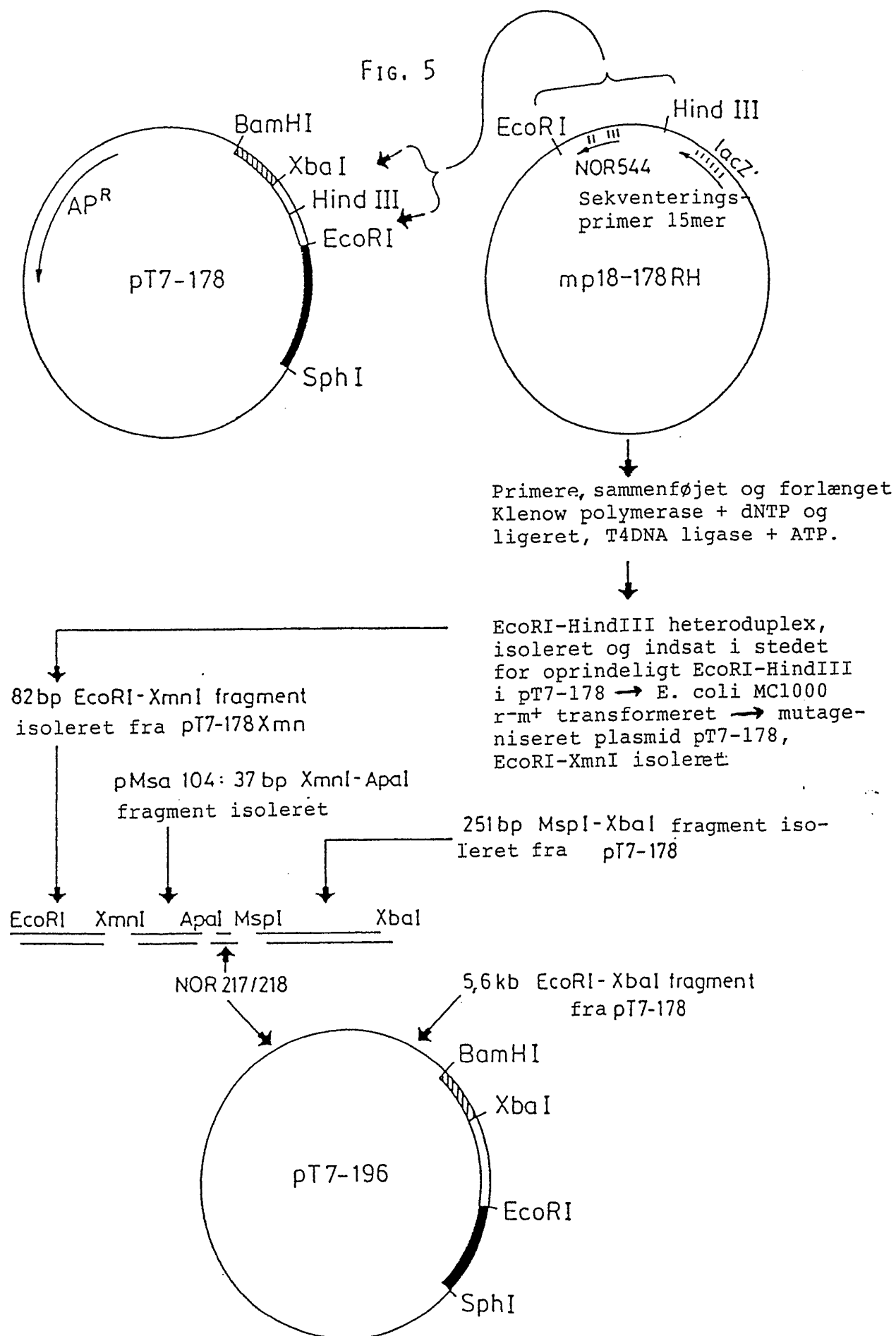


FIG. 6

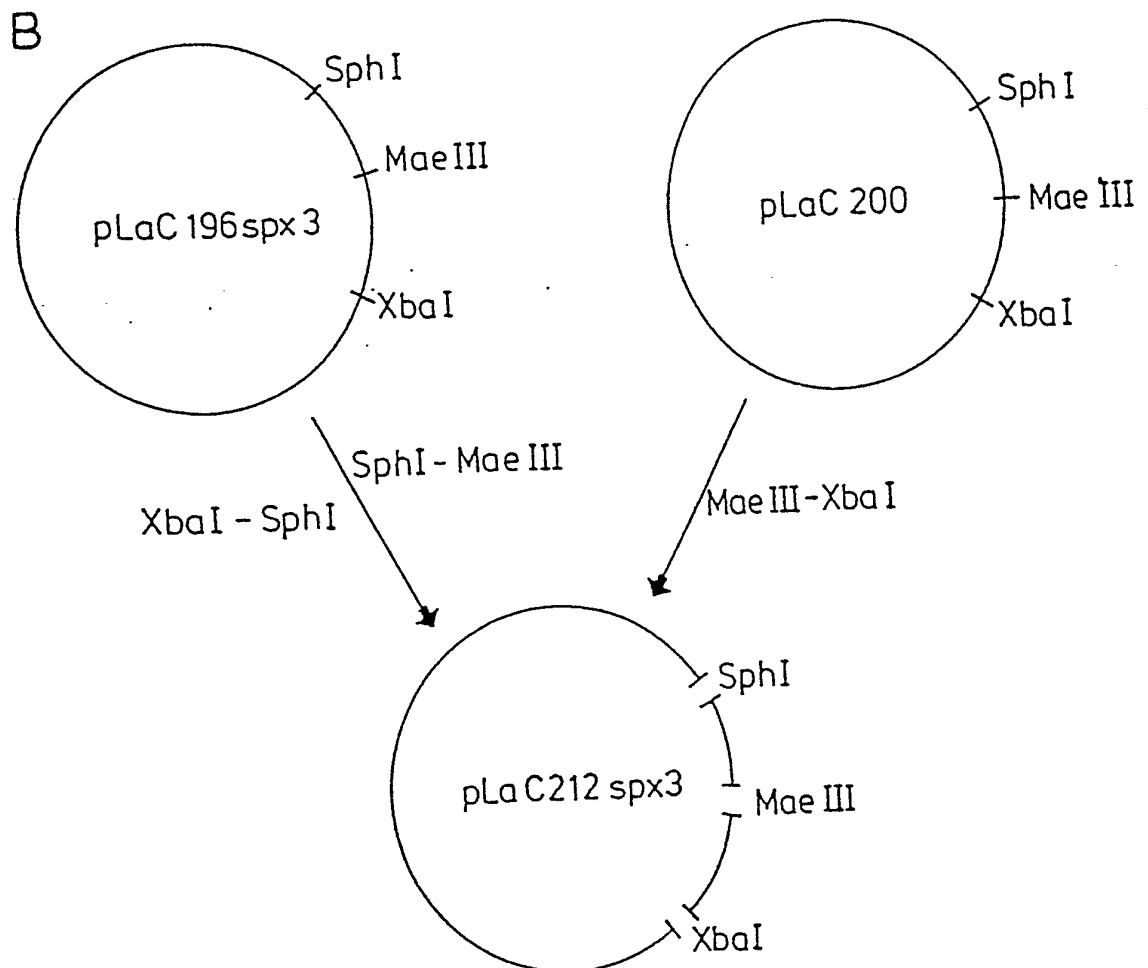
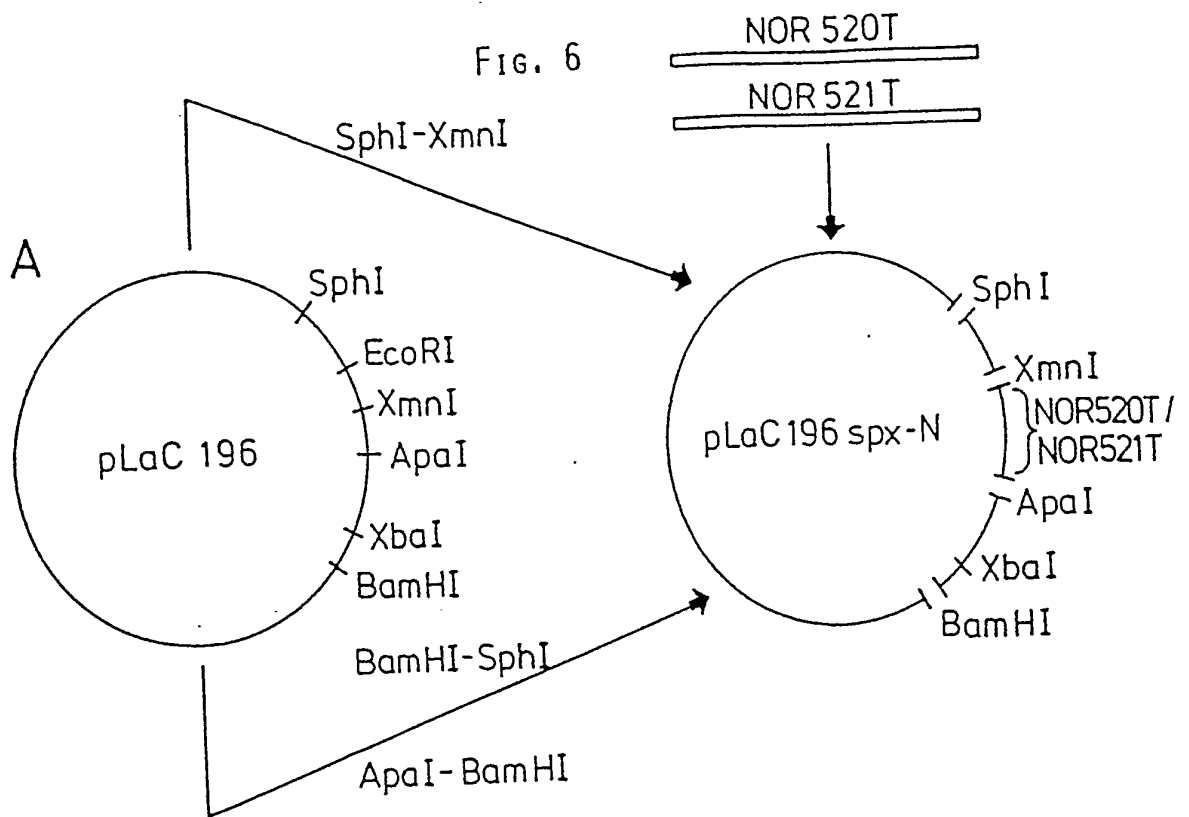


FIG. 7

```

      12      24      36      48      60
AATTCATTCAA GAATAGTTCAAA CAAGAAGATTAC AACTATCAATT TCATACACAATA

      72      84      96      108      120
TAAACGATTAAA AGAATGAGATTT CCTTCTATTTTT ACTGCTGTTTTA TTCGCTGCTTCC
      M R F P S I F T A V L F A A S

      132      144      156      168      180
TCCGCTTTAGCT GCTCCAGTCAAC ACTACCACTGAA GATGAAACGGCT CAAATTCCAGCT
S A L A A P V N T T T E D E T A Q I P A

      192      204      216      228      240
GAAGCTGTCATC GGTACTCTGAT TTAGAAGGTGAT TTCGATGTTGCT GTTTTGCCATTT
E A V I G Y S D L E G D F D V A V L P F

      252      264      276      288      300
TCCAACCTCCACC AATAACGGTTTA TTGTTTATCAAT ACTACTATTGCC TCCATTGCTGCT
S N S T N N G L L F I N T T I A S I A A

      312      324      336      348      360
AAAGAAGAAGGT GTTTCCTTGGAT AAAAGATTCGTT AACCAACACTTG TCGGGTCCCAC
K E E G V S L D K R F V N Q H L C G S H

      372      384      396      408      420
TTGGTTGAAGCT TTGTA CTGGTT TGCGGTGAAAGA GGTTTCTTCTAC ACTCCTAAGGCT
L V E A L Y L V C G E R G F F Y T P K A

      432      444      456      468      480
GCTAAGGGTATT GTCGAACAATGC TGTACCTCCATC TGCTCCTTGAC CAATTGAAAAC
A K G I V E Q C C T S I C S L Y Q L E N

      492      504      516      528      540
TACTGCAACTAG ACGCAGCCCGCA GGCTCTAGA
Y C N |

```

DNA-sekvens af EcoRI-Xba fragmentet fra
 pT~~7~~MI3 og det tilsvarende protein angiv-
 et med étbogstavsnomenklaturen nedenfor.

Fig. 8

```

      12           24           36           48           60
AATTCCATTCAA GAATAGTTCAAA CAAGAAGATTAC AACTATCAATT TCATACACAATA

      72           84           96           108          120
TAAACGATTAAA AGAATGAGATTT CCTTCTATTTT ACTGCTGTTTTA TTCGCTGCTTCC
           M R F P S I F T A V L F A A S

      132          144          156          168          180
TCCGCTTTAGCT GCTCCAGTCACC GGTGACGAAAGC TTCGTCGAAATT CCAGCTGAAAAC
S A L A A P V T G D E S F V E I P A E N

      192          204          216          228          240
ACCACTTTGGCT AAGAGATTCGTT AACCAACACTTG TGCGGTTCCAC TTGGTTGAAGCT
T T L A K R F V N Q H L C G S - H L V E A

      252          264          276          288          300
TTGTACTTGTTT TGCGGTGAAAGA GGTTTCTTCTAC ACTCCTAAGGCT GCTAAGGGTATT
L Y L V C G E R G F F Y T P K A A K G I

      312          324          336          348          360
GTCGAACAATGC TGTACCTCCATC TGCTCCTTGTAC CAATTGGAAAAC TACTGCAACTAG
V E Q C C T S I C S L Y Q L E N Y C N |

      372          384          396          408          420
ACGCAGCCCCGA GGCTCTAGA

```

DNA-sekvens af EcoRI-XbaI fragmentet fra pT7-178 og det tilsvarende protein angivet med étbogstavsnomenklaturen nedenfor.

Fig. 9

12	24	36	48	60
AATTCATTCAA	GAATAGTTCAA	CAAGAAGATTAC	AACTATCAATT	TCATACACAATA
72	84	96	108	120
TAAACGATTAAA	AGAATGAGATTT	CCTTCTATTTTT	ACTGCTGTTTTA	TTCGCTGCTTCC
	M R F	P S I F	T A V L	F A A S
132	144	156	168	180
TCCGCTTTAGCT	GCTCCAGTTACT	GGCGATGAATCA	TCTGTTGAGATT	CCGGAAGAGTCT
S A L A	A P V T	G D E S	S V E I	P E E S
192	204	216	228	240
CTGATCGGTTTC	TTGGATTTAGCA	GGTGATGATATC	GCTGAAAACACC	ACTTTGGCTAAG
L I G F	L D L A	G D D I	A E N T	T L A K
252	264	276	288	300
AGATTCGTTAAC	CAACACTTGTGC	GGTTCCCACTTG	GTTGAAGCTTTG	TACTTGGTTTGC
R F V N	Q H L C	G S H L	V E A L	Y L V C
312	324	336	348	360
GGTGAAAGAGGT	TTCTTCTACACT	CCTAAGGCTGCT	AAGGGTATTGTC	GAACAATGCTGT
G E R G	F F Y T	P K A A	K G I V	E Q C C
372	384	396	408	420
ACCTCCATCTGC	TCCTTGTACCAA	TTGGAAAACACTAC	TGCAACTAGACG	CAGCCCCGAGGC
T S I C	S L Y Q	L E N Y	C N	
432	444	456	468	480
TCTAGA				

DNA-sekvens af EcoRI-XbaI fragmentet fra pT7-193 og det tilsvarende protein angivet med étbogstavsnomenklaturen nedenfor.

Fig. 10

12	24	36	48	60
AATTCCATTCAA	GAATAGTTCAAA	CAAGAAGATTAC	AAACTATCAATT	TCATACACAATA
72	84	96	108	120
TAAACGATTAAA	AGAATGAGATTT	CCTTCTATTTTT	ACTGCTGTTTTA	TTCGCTGCTTCC
	M R F	P S I F	T A V L	F A A S
132	144	156	168	180
TCCGCTTTAGCT	GCTCCAGTTACT	GGCGATGAATCA	TCTGTTGAGATT	CCGGAAGAGTCT
S A L A	A P V T	G D E S	S V E I	P E E S
192	204	216	228	240
CTGATCATCGCT	GAAAACACCACT	TTGGCTAAGAGA	TTCGTTAACCAA	CACTTGTGCGGT
L I I A	E N T T	L A K R	F V N Q	H L C G
252	264	276	288	300
TCCCACTTGGTT	GAAGCTTTGTAC	TTGGTTTGCGGT	GAAAGAGGTTTC	TTCTACACTCCT
S H L V	E A L Y	L V C G	E R G F	F Y T P
312	324	336	348	360
AAGGCTGCTAAG	GGTATTGTGCGAA	CAATGCTGTACC	TCCATCTGCTCC	TTGTACCAATTG
K A A K	G I V E	Q C C T	S I C S	L Y Q L
372	384	396	408	420
GAAAACTACTGC	AACTAGACGCAG	CCCGCAGGCTCT	AGA	
E N Y C	N			

DNA-sekvens af EcoRI-XbaI fragmentet fra pT7-194 og det tilsvarende protein angivet med étbogstavsnomenklaturen nedenfor.

Fig. 11

```

      12          24          36          48          60
AATTCATTCAA GAATAGTTCAAA CAAGAAGATTAC AAATATCAATT TCATACACAATA

      72          84          96          108          120
TAAACGATTAAA AGAATGAGATTT CCTTCTATTTTT ACTGCTGTTTTA TTCGCTGCTTCC
              M R F P S I F T A V L F A A S

      132          144          156          168          180
TCCGCTTTAGCT GCTCCAGTTACT GGCGATGAATCA TCTGTTGAGATT CCGGAAGAGTCT
S A L A A P V T G D E S S V E I P E E S

      192          204          216          228          240
CTGATCATCGCT GAAAACACCACT TTGGCTAACGTC GCCATGGCTAAG AGATTCGTTAAC
L I I A E N T T L A N V A M A K R F V N

      252          264          276          288          300
CAACACTTGTGC GGTTCCCACCTTG GTTGAAGCTTTG TACTTGGTTTGC GGTGAAAGAGGT
Q H L C G S H L V E A L Y L V C G E R G

      312          324          336          348          360
TTCTTCTACACT CCTAAGGCTGCT AAGGGTATTGTC GAACAATGCTGT ACCTCCATCTGC
F F Y T P K A A K G I V E Q C C T S I C

      372          384          396          408          420
TCCTTGTACCAA TTGGAAAACCTAC TGCAACTAGACG CAGCCCGCAGGC TCTAGA
S L Y Q L E N Y C N |

```

DNA-sekvens af EcoRI-XbaI fragmentet fra pT7-200 og det tilsvarende protein angivet med étbogstavsnomenklaturen nedenfor.

Fig. 12

```

      12           24           36           48           60
AATTCATTCAA GAATAGTTCAA CAAGAAGATTAC AACTATCAATT TCATACACAATA

      72           84           96          108          120
TAAACGATTAAA AGAATGAAGGCT GTTTTCTTGGTT TTGTCCTTGATC GGATTGTGCTGG
      M K A   V F L V   L S L I   G L C W

      132          144          156          168          180
GCCCAACCAGTT ACTGGCGATGAA TCATCTGTTGAG ATTCCGGAAGAG TCTCTGATCATC
A Q P V   T G D E   S S V E   I P E E   S L I I

      192          204          216          228          240
GCTGAAAACACC ACTTTGGCTAAC GTCGCCATGGCT AAGAGATTCGTT AACCAACACTTG
A E N T   T L A N   V A M A   K R F V   N Q H L

      252          264          276          288          300
TGCGGTCCCCAC TTGGTTGAAGCT TTGTACTTGGTT TGCGGTGAAAGA GGTTCCTTCTAC
C G S H   L V E A   L Y L V   C G E R   G F F Y

      312          324          336          348          360
ACTCCTAAGGCT GCTAAGGGTATT GTCGAACAATGC TGTACCTCCATC TGCTCCTTGTAC
T P K A   A K G I   V E Q C   C T S I   C S L Y

      372          384          396          408          420
CAATTGGAAAAC TACTGCAACTAG ACGCAGCCCGCA GGCTCTAGA
Q L E N   Y C N |

```

DNA-sekvens af EcoRI-XbaI fragmentet, der
 indkoder signal-leader-p-insulinen fra
 pLaC 212spx3
 Den tilsvarende proteinsekvens er angivet
 med étbogstavsnomenklaturen under DNA-
 -sekvensen.

Fig. 13

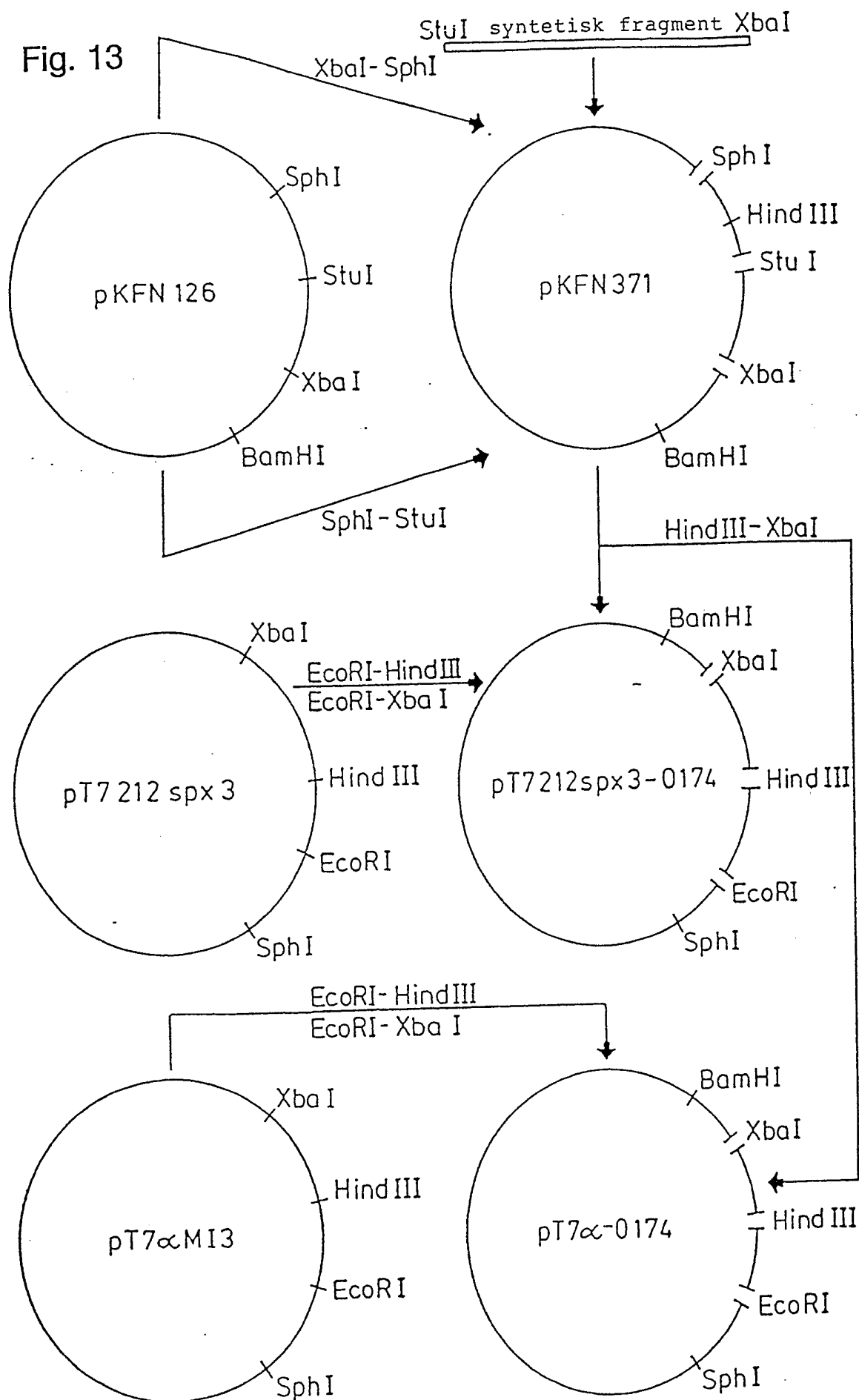


Fig. 14

