(19) DANMARK (11) **DK 175910 B1**

(12) PATENTSKRIFT

Patent- og
Varemærkestyrelsen

(51) Int.Cl⁷.: **C 12 N 1/19** **C 12 N 15/81** **C 12 P 21/00**

(21) Patentansøgning nr: **PA 1988 07285**

(22) Indleveringsdag: **1988-12-29**

(24) Løbedag: **1988-12-29**

(41) Alm. tilgængelig: **1989-07-01**

(45) Patentets meddelelse bkg. den: **2005-06-13**

(30) Prioritet: **1987-12-30 US 139682**

(73) Patenthaver: **Chiron Corporation, 4560 Horton Street, Emeryville, Californien 94608, USA**

(72) Opfinder: **Patricia Tekamp-Olson, 1426 43rd Avenue, San Francisco, Californien 94608, USA**

(74) Fuldmægtig: **Zacco Denmark A/S, Hans Bekkevolds Allé 7, 2900 Hellerup, Danmark**

(54) Benævnelse: **Gærcelle omfattende en DNA-konstruktion, der tilvejebringer ekspression og secernering af et ikke-gærprotein, samt dobbeltstrenget DNA-molekyle**

(56) Fremdragne publikationer:
Ingen

(57) Sammendrag:

Et gær-α-faktor ekspressionssystem, som omfatter en trunkeret ledersekvens, der indeholder α-faktor signalpeptidet og et glycosyleringssted, og som gennem et processeringssted er koblet til en sekvens, der koder for et ikke-gærprotein.

fortsættes

Teknisk område

Den foreliggende opfindelse angår fremstillingen af rekombinerede proteiner i gær. Opfindelsen angår nærmere bestemt et forbedret α-faktor ekspressionssystem, som tilvejebringer secernering af heterologe proteiner fra gærværtsceller.

Baggrund

Kurjan et al. (1982) Cell, 30:933-943 omtaler den første kloning og sekventering af et gen, der koder for et gær α-faktor forstadiegen. Kurjan et al. omtaler også i U.S. patentskrift nr. 4.546.082 kloningen af dette gen og foreslår, at α-faktor ledersekvensen kan an vendes til at styre secerneringen af heterologe proteiner i gær. Patentskriftet indeholder imidlertid ikke data, som indicerer, at patenthaveren nogensinde med held har anvendt α-faktor lederen til at udtrykke og secernere et heterologt protein i gær.

I EPO publikation nr. 1.16.201 omtales den første vellykkede anvendelse af α-faktor lederen til at styre ekspressionen og secerneringen af et heterologt protein, epidermal vækstfaktor, fra en transformeret gær. Efter dette arbejde har der været yderligere rapporter om ekspression af heterologe proteiner i gær under anvendelse af α-faktor lederen. Se f.eks. Elliott et al. (1983) Proc. Natl. Acad. Sci. USA, 80:7080-7084; Bitter et al. (1984) Proc. Nat]. Acad.. Sci.. USA, 81:533O Smith et al. (1985) Science, 229:1219-1229; EPO publikationer nr. 114.695, 123.228, 123.294, 123.544, 128.773 og 206.783. Se også Brake et al. i Protein Transport and Secretion, s. 103 (J.M. Gething, udg. 1984).

Med de ovenfor omtalte ekspressionssystemer frembringes en α-faktor leder i fuld længde fusioneret til et heterologt protein. Om end der med det ovennævnte arbejde påvises, at ekspressionssystemet har en bred anvendelighed, er det forud for udførelse af forsøget almindeligvis ikke muligt at forudsi-

ge, om et bestemt heterologt protein med held vil blive secerneret, modnet og være biologisk aktivt. Se f.eks. EPO 206.783, ovenfor, s. 2-5, Rothblatt et al. (1987) EMBO 3., 6:3455-3463, V.L. MacKay, "Secretion of Heterologous Proteins in Yeast (under trykning).

På basis af ikke-publicerede resultater er det adskillige gange blevet omtalt, at deletioner fra "pro"-området i α-faktor lederen (mellem signalpeptidet og den første "spacer") medfører et væsentligt fald i mængden af ikke-gærprotein, som secerneres fra gær, der er transformeret med de heterologe konstruktioner. Sidu et al. (1987) omtaler i Gene 54:175-184, at sur phosphatase (PHO5) fra gær secerneres i mediet fra en heterolog konstruktion, hvortil der både er anvendt en α-faktor leder i fuld længde og en trunkeret α-faktor leder, men at ekspressionsniveauerne er 3-4 x mindre end for PHO5-genet under styring af dens homologe leder. Det er også blevet anført, at deletioner i præpro-α-faktor forstadiegenet resulterer i et væsentligt fald i secernering af det native α-faktor peptid. Se f.eks. V.L. MacKay, ovenfor, Rothblatt et al., ovenfor.

Der er derfor et behov for at forbedre α-faktor ekspressionssystemet, navnlig til brug for ikke-gærproteiner, som ikke produceres effektivt med de nuværende α-faktor ekspressionskonstruktioner.

Opsummering af opfindelsen

Det har overraskende vist sig, at en trunkeret form af α-faktor leder sekvensen effektivt kan styre ekspressionen og secerneringen af heterologe polypeptider i gær. Det er navnlig overraskende, at de trunkerede α-faktor ledersekvenser kan give en væsentlig forbedring af effektiviteten af ekspressionen af sådanne heterologe proteiner i forhold til. ekspressionssystemer, som udnytter α-faktor lederen i fuld længde; det vil sige højere niveauer af korrekt N-terminal processering, secernering af heterologe proteiner, hvoriblandt en

større procentdel af molekylerne er biologisk aktive osv. Disse resultater er navnlig overraskende på baggrund af rapporter om, at deletioner fra ledersekvensen i α-faktor forstadiet resulterer i faldende secerneringsniveauer af aktiv α-faktor.

Med den foreliggende opfindelse tilvejebringes derfor alternative α-faktorbaserede ekspressionskonstruktioner, som navnlig er anvendelige til ekspression af heterologe polypeptider, som enten ud trykkes ineffektivt af α-faktor lederkonstruktioner i fuld længde eller som ikke udtrykkes overhovedet af sådanne konstruktioner i fuld længde.

I én udførelsesform angår den foreliggende opfindelse en gærcelle omfattende en DNA-konstruktion, der tilvejebringer ekspression og secernering af et ikke-gærprotein, hvilken DNA-konstruktion omfatter en kodende sekvens under styring af gærgenkendte transkriptionsinitierings- og termineringssekvenser, hvilken kodende sekvens koder for et forstadiepolypeptid, der omfatter en ledersekvens og ikke gærproteinet, som er koblet sammen i et processeringssted, der tilvejebringer spaltning af ikke-gærproteinet fra forstadiepolypeptidet, hvorhos ledersekvensen udgør fra ca. 25 til ca. 50 N-terminale aminosyrerester af forstadiepolypeptidet og omfatter signalpeptidet af Saccharomyces eller Kluveromyces α-faktor forstadie og et enkelt glycosyleringssted.

I en anden udførelsesform angår den foreliggende opfindelse et dobbeltstrenget DNA-molekyle, der omfatter et område, som koder for et forstadiepolypeptid, der kan secerneres af en gærvært, hvilket område under henvisning til én af DNA-strengene har konstruktionen:

5'-AF-CHO-X_n-S-gen*-3'

hvor AF-CHO-X_n omfatter et enkelt funktionelt glycosyleringssted i hvilket:

AF koder for et Saccharomyces eller Kluveromyces α-faktor signalpeptid,

CHO koder for et glycosyleringssted,

X_n koder for et polypeptid med en længde på n aminosyrer, og som ikke indeholder et glycosyleringssted eller et processeringssted, som gør det muligt for gæren at spalte forstadiepolypeptidet in vivo,

n er et helt tal på fra 0 til ca. 30,

gen* koder for et ikke-gærprotein, og

S koder for et processeringssted, som muliggør spaltning af forstadiepolypeptidet.

Den foreliggende opfindelse angår også fremgangsmåder til at anvende de ovennævnte celler og DNA-konstruktioner til frembringelse af rekombinerede proteiner såvel som de sammensætninger af rekombinerede proteiner, som frembringes ved de ovennævnte fremgangsmåder. Andre udførelsesformer vil også være indlysende for en fagmand.

Beskrivelse af figurerne

Figur 1 er et arbejdsdiagram, som viser konstruktionen af pYBCA5, samt både nukleotid- og aminosyresekvensen af det i eksempel I anvendte syntetiske proinsulingen. De forskellige syntetiske oligonukleotider, som anvendes i konstruktionen, er markeret med sorte prikker. Pile over sekvensen viser begyndelsen og slutningen af B-, C- og A-proinsulinkæderne. De sorte kasser markerer de dibasiske processeringssteder.

Figur 2 viser nukleotid- og aminosyresekvensen af et syntetisk oligonukleotid, som koder for en modificeret α-faktor leder og de første 13 aminosyrer af proinsulingenet, som anvendes til konstruktion af pYGAI3 i eksempel I. Glycosyleringsstederne i den modificerede α-faktor leder er blevet fjernet ved at ændre codonerne for Asn til at kode for Gln (indrammet). Pilene angiver forbindelsesstedet mellem den sekvens, som koder for KEX2-endopeptidasestedet og den N-terminale ende af human-proinsulin.

Figur 3 viser det syntetiske gen af fragment pYGAIC3, som koder for proinsulinanalogen, hvor C-peptidet er blevet erstattet med et KEX2-endopeptidasested (indrammet). Det 133 bp lange syntetiske fragment, som der henvises til i eksempel II, er defineret af de vertikale og horisontale linier gennem nukleotidsekvensen.

Figur 4 er et restriktionskort af gærfærgevektoren pAB24.

Figur 5 viser DNA-sekvensen af det syntetiske gen, som koder for det i eksempel III beskrevne IGF1.

Figur 6 er et restriktionskort af pYLUIGF1-55, en ekspressions vektor, der er beskrevet i eksempel III, og som koder for IGF1 under styring af en trunkeret α-faktor leder.

Figur 7 er et restriktionskort af pYLUIGF1-24, en ekspressions vektor, som er beskrevet i eksempel III, og som koder for IGF1 under styring af en α-faktor leder i fuld længde og med tre glycosyleringssteder.

Detaljeret beskrivelse

Med mindre andet er anført, er der ved udøvelsen af den foreliggende opfindelse blevet anvendt traditionelle metoder inden for molekylær biologi, mikrobiologi og rekombinant DNA-teknikker, hvilke metoder vil være kendt af fagmanden. Metoderne er fuldt ud forklaret i litteraturen. Se f.eks. Maniatis, Fritsch & Sambrook, Molecular Cloning: A Laboratory Manual (1982); DNA Cloning, bind I & II (D.N. Glover, red., 1985); Oligonucleotide Synthesis (M.J. Gait, red., 1984); Transcription and Translation (B.D. Hames & S.J. Higgins, red., 1984); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide to Molecular Cloning (1984).

Ved beskrivelse af den foreliggende opfindelse vil følgende terminologi blive anvendt i overensstemmelse med de definitioner, som er anført nedenfor.

En "replicon" er et hvilket som helst genetisk element (f.eks. plasmid, cosmid, kromosom, virus), der fungerer som en autonom enhed ved DNA-replikation in vivo – dvs. er i stand til at replikere under egen styring.

En "vektor" er en replicon, såsom et plasmid, fag eller cosmid, til hvilken der kan knyttes et andet DNA-segment, hvorved det tilknyttede segment bringes til at replikere.

Et "dobbeltstrenget DNA-molekyle" refererer til den polymere form af deoxyribonukleotider (adenin, guanin, thymidin eller cytosin) i dens normale, dobbeltstrengede helix. Dette udtryk refererer kun til den primære og sekundære struktur af molekylet, og begrænser det ikke til nogen bestemte tertiære former. Dette udtryk omfatter således dobbeltstrenget DNA, som bla. findes i lineære DNA-molekyler (f.eks. restriktionsfragmenter), vira, plasmider og kromosomer. Ved diskussion af strukturen af et bestemt dobbeltstrenget DNA-molekyle, vil der her blive beskrevet sekvenser i overensstemmelse med den normale konvention ved kun at angive sekvensen i 5'- til 3'-retningen langs den ikke-transkriberede DNA-streng, dvs, den streng, som har en sekvens, der er homolog med det mRNA, som frembringes fra en bestemt kodende sekvens.

En DNA "kodende sekvens" er en DNA-sekvens, som in vivo kan transkriberes og translateres til et polypeptid, når det sættes under styring af hensigtsmæssige regulatoriske sekvenser. Grænserne for den kodende sekvens bestemmes af og omfatter translationsstart codonen ved 5'-enden (amino) og en translationsstopcodon ved 3'-enden (carboxy). En kodende sekvens kan omfatte, men er ikke begrænset til, prokaryotiske DNA-sekvenser, virale DNA-sekvenser, cDNA eller genomiske DNA-sekvenser fra eukaryotiske kilder (f.eks. pattedyr) eller syntetiske DNA-sekvenser.

"Gærgenkendte transkriptionsinitierings- og termineringssekvenser" refererer til DNA-regulatoriske områder, som flankerer en kodende sekvens, og som er ansvarlige for transkriptionen i gær af et mRNA, der er homolog med den kodende sekvens, som derefter kan translateres til det polypeptid, der er kodet af den kodende sekvens. Transkriptionsinitieringssekvenser omfatter gærpromotorsekvenser, hvilket er DNA regulatoriske sekvenser, som er i stand til at binde gær RNA-polymerase i en celle og initiere transkription af en kodende sekvens i nedenstrømsretningen (3'-retningen). I forbindelse med den foreliggende opfindelse er promotorsekvensen ved dens 3'-ende afgrænset af (og omfatter ikke) translationsstartcodonen for en kodende sekvens og strækker sig opstrøms (5'-retningen) til at omfatte det minimale antal nukleotider eller elementer, som er nødvendige for at initiere transkription ved niveauer, som kan påvises i forhold til baggrunden. Inden for promotorsekvensen vil der være et transkriptionsinitieringssted (kan bekvemt defineres ved kortlægning med nuclease SI) såvel som proteinbindende domæner (consensussekvenser), der er ansvarlige for bindingen af gær RNA polymerasen. Promotorer, som er brugbare ved den foreliggende opfindelse, omfatter vildtype α-faktor promotoren såvel som andre gær promotorer. Specielt foretrukne er gærpromotorer, som er involveret med enzymerne i den glycolytiske bane, f.eks. phosphoglucoisomerase, phosphofructokinase, phosphotrioseisomerase, phosphoglucomutase, enolase, pyrodruesyrekinase, glyceraldehyd-3-phosphatdehydrogenase, alkoholdehydrogenase såvel som hybrider af disse promotorer. Se f.eks. [publikationer nr. 120.551 og 164.556. Transkriptionsinitieringssekvenser kan også omfatte andre regulatoriske områder, som er ansvarlige for promotorregulering eller -forstærkning. På lignende måde kan en transkriptionstermineringssekvens, som er anbragt 3' for translationsstopcodonen enten være vildtype α-faktor transkriptionstermineringssekvensen eller en anden gærgenkendt termineringssekvens, således som sekvenserne fra generne for ovennævnte glycolytiske enzymer.

En kodende sekvens er "under styring" af transkriptionsinitierings og termineringssekvenser, når RNA-polymerase bindes til transkriptionsinitieringssekvensen og transkriberer den kodende sekvens til mRNA med en afslutning ved transkriptionstermineringssekvensen, og mRNA'et derefter translateres til det polypeptid, som er kodet af den kodende sekvens (dvs. "ekspression"). Det forstadiepolypeptid, som er kodet af de kodende sekvenser ifølge den foreliggende opfindelse, "secerneres", når i det mindste en del (sædvanligvis ikke gærproteinet uden Tedersekvensen) transporteres ekstracellulært, hvor det findes i cellevækstmediet. Sædvanligvis secerneres kun den del af forstadieproteinet, som er nedenstrøms ledersekvensen, og denne nedenstrøms del kan også udsættes for yderligere processering under secernering, såsom proteolytisk spaltning, glycosylering, foldning, dannelse af disulfidbinding osv.

En celle er blevet "transformeret" med exogent DNA, når et sådant exogent DNA er blevet ført inden for cellevæggen. Exogent DNA kan eller er ikke integreret (covalent koblet) med kromosomalt DNA, som udgør cellens genom. Det exogene DNA kan opretholdes ekstrakromosomalt på en replicon, såsom et plasmid. Når det exogene DNA er blevet integreret i kromosomet, nedarves det af døtrecellerne ved kromosomreplikering. En celle, der er blevet transformeret med exogent DNA, som er integreret i kromosomet, omtales som en "stabilt" transformeret celle. En "klon" eller en "klonpopulation" er en cellepopulation, som ved mitose hidrører fra en enkelt celle eller en fælles forfader.

To DNA-sekvenser er "i alt væsentligt homologe", når der er mindst ca. 60% (fortrinsvis mindst ca. 75% og mest fortrinsvis mindst ca. 90%) af nukleotiderne, som matcher sammen over en defineret længde af molekylerne. Sekvenser, som er i alt væsentligt homologe, kan identificeres ved et Southern hybridiserings forsøg under betingelser med en udvalgt stringenthed, hvilket er defineret for det specielle system. Fastlæggelse af passende hybridise-

ringsbetingelser kan foretages af en fagmand. Se f.eks. Maniatis et al., se ovenfor, DNA Cloning, se ovenfor, Nucleic Acid Hybridization, se ovenfor.

Et "heterologt område" af et DNA-molekyle er et identificerbart segment af DNA inden for et større DNA-molekyle, som ikke findes associeret med det større molekyle i naturen. Når det heterologe område således koder for et pattedyrprotein, vil det heterologe om råde sædvanligvis være flankeret af DNA, som ikke flankerer pattedyr DNA-sekvensen i genomet i den oprindelige organisme. Et andet eksempel på en heterologt kodende sekvens er en konstruktion, hvor den kodende sekvens i sig selv ikke findes i naturen (f.eks. et cDNA, hvor den genomiske, kodende sekvens indeholder intron-områder, eller syntetiske sekvenser, som har codoner, som adskiller sig fra organismer, der koder for det samme eller et lignende protein). Allele variationer eller naturligt forekommende muteringsbegivenheder giver ikke ophav til et "heterologt" område af DNA, sådan som udtrykket anvendes her.

Som anvendt her omfatter "gær" ascosporogene gærtyper (Endomyce tales), basidiosporogene gærtyper og gær, som tilhører Fungi imperfecti (Blastomycetes). De ascosporogene gærtyper inddeles i to familier, Spermophthoraceae og Saccharomycetaceae. Sidstnævnte består af fire underfamilier, Schizosaccharomycoideae (f.eks. slægten Schizosaccharomyces), Nadsonioidéae, Lipomycoideae og Saccharomycoideae (f.eks. slægterne Pichia, Kluyveromyces og Saccharomyces). De basidosporogene gærtyper omfatter slægterne Leucosporidium, Rhodosporidium, Sporidiobolus, Filobasidium og Filobasidiella. Gær, der tilhører Fungi Imperfecti, inddeles i to familier, Sporobolomycetaceae (f.eks. slægterne Sporobolomyces, Bullera) og Cryptococcaceae (f.eks. slægten Candida). Af speciel interesse for den foreliggende opfindelse er arterne inden for slægterne Pichia, Kluyveromyces, Saccharomyces, Schizosaccharomyces og Candida. Af speciel interesse er Saccharomycesarterne S. cerevisiae, S. carisbergensis, S. diastaticus, S. douglasii, S. kluyveri, S. norbensis og S. oviformis. Arter af speciel interesse inden for slægten Kluyveromyces omfatter K. lactis. Eftersom klassificeringen af gær

kan ændre sig i fremtiden, defineres gær i forbindelse med den foreliggende opfindelse som beskrevet i Biology and Activities of Yeast (F.A. Skinner, S.M. Passmore og R.R. Davenport, red., 1980) (Soc. App. Bacteriol. Symp. serie nr. 9). I forbindelse med ovennævnte er en fagmand sandsynligvis fortrolig med gærbiologien og håndteringen af gærgenetik. Se f.eks. Biochemistry and Genetics of Yeast (M. Bacila, B.L. Horecker & A.O.M. Stoppani, red., 1978), The Yeasts (A.H. Rose og J.S. Harrison, red., 2. udgave, 1987), The Molecular Biology of the Yeast Saccharoniyces (Strathern et al., red., 1981). Den fra ovennævnte referencer kendte teknik er medtaget her ved henvisning.

Ved den foreliggende opfindelse anvendes trunkerede ledersekvenser fra et gær α-faktor gen. α-Faktor er et parringspheromonoligopeptid med en længde på ca. 13 aminosyrer frembragt ud fra et stort forstadiepolypeptid med en længde på fra ca. 100 til 200 aminosyrer (typisk fra ca. 120 til 160) (præpro-α-faktor). Forstadiet omfatter en hydrofob "signalsekvens" med ca. 20 aminosyrer (f.eks. ca. 19-23, typisk ca. 26-22) efterfulgt af et yderligere lederområde på ca. 60 hydrofile aminosyrerester ("pro"-området), som derefter er koblet til adskillige tandemgentagelser af den modne pheromonsekvens (typisk ca. 2-6) adskilt af korte oligopeptidspaceområder (typisk ca. 6-8 aminosyrerester), som muliggør proteolytisk processering til det modne pheromon.

Kloningen af forskellige præpro-α-faktor gener er blevet beskrevet. Se f.eks. Kurjan et al., U.S. patent nr. 4.546.082, Singh et al. (1983), Nucleic Acids Res., 11:4049-4063, DK patentansøgning nr. 30 4222/88, indleveret 28. juli 1988, hvor den herfra kendte teknik er medtaget ved henvisning. Endvidere kan DNA-sekvenser, som koder for præpro-genet, identificeres ved hybridisering med prober fra kendte præpro-α-faktor sekvenser. Se f.eks. Brake et al. (1983), Molec. & Cell Biol. 3:1440-1450. α-Faktor kan også oprenses fra en gærart, sekventeres og der kan udformes prober til kloning af præpro-α-

faktor genet. Se f.eks. McCullough et al. (1979), J. Bacteriol. 138:146-154, Sato et al. (1981), Agric. Biol. Chem. 44:1451-1453, Singh et al. (1983), se ovenfor. Det har også vist sig, at α-faktor ledersekvensen fra én gærart kan fungere i en anden gærart. Se f.eks. U.S. ansøgning nr. 78.551, se ovenfor. Den foreliggende opfindelse angår således ikke kun anvendelsen af α-faktor ledersekvenser fra gær i almindelighed, men anvendelsen af sådanne ledersekvenser i heterologe gærarter. Af bekvemmeligheds grunde vil opfindelsen imidlertid blive diskuteret med hensyn til præpro-α-faktor genet MFcd fra S. cerevisiae. Se f.eks. Kurjan et al., U.S. patent nr. 4.546.083, Singh et al. (1983), se ovenfor.

Med den foreliggende opfindelse anvendes kimære DNA-konstruktioner, som koder for hybride forstadiepolypeptider, der omfatter en leder sekvens og et ikke-gærpolypeptid. I forbindelse med den foreliggende opfindelse afgrænses leder-DNA-sekvensen som begyndende med den N terminale startcodon (methionin) i forstadiepolypeptidet til den condon, som koder for den sidste aminosyrerest før processerings stedet, der ligger mellem ledersekvensen og den sekvens, som koder for ikke-gærproteinet. Ledersekvensen ifølge den foreliggende opfindelse er sammensat af én trunkeret form af en gær-α-faktor ledersekvens, som typisk har en længde på fra ca. 25 til ca. 50 aminosyrerester. Ledersekvensen ifølge den foreliggende opfindelse er således ca. 30 aminosyrerester kortere end den typiske α-faktor leder i fuld længde. MF indeholder f.eks. en ledersekvens på 83 aminosyrerester efterfulgt af en hexapeptidspacersekvens, som spaltes af gærprocesseringsenzymer. Ved frembringelse af deletioner fra ledersekvensen er det vigtigt, at mindst ét glycosyleringssted (-Asn-Y-Thr/Ser-) bevares til tilvejebringelse af effektiv secernering.

Det er også nødvendigt, at lederen bevarer en funktionel α-faktor signalsekvens. Som ovenfor anført har signalpeptidet sædvanligvis en længde på ca. 20 aminosyrer og karakteristiske træk omfatter en hydrofob kerne. Se f.eks.

von Heijne (1984), J. Mol. Biol. 173:243- 251. Alle de præpro-α-faktor sekvenser, som er undersøgt indtil i dag, koder for et hydrofobt peptid med en længde på ca. 20 aminosyrer. Om end den nøjagtige længde af et signalpeptid, som er nødvendigt for at styre forstadiepolypeptidet til secerneringsbanen, ikke er defineret, vil det sædvanligvis kræve mellem ca. 19 og ca. 23 aminosyrerester, idet den minimale sekvens, som er nødvendig, let kan defineres ved at undersøge deletionsmutanter.

Med hensyn til MFα1 tænkes der således i forbindelse med den foreliggende opfindelse på deletioner i området fra ca. 30 til ca. 60 aminosyrerester, typisk fra mellem ca. 33 og ca. 58 aminosyrerester og mere typisk mellem ca. 48 og ca. 58 aminosyrerester. Disse deletioner vil almindeligvis forekomme i området mellem og omfattende aminosyreresterne 26 til 83. Det foretrækkes, at deletionerne omfatter glycosyleringsstederne ved aminosyreresterne 57-59 og 67-69. De deleterede α-faktor ledersekvenser kan om ønsket delvis erstattes med ikke-α-faktor ledersekvenser. Sekvenserne bør almindeligvis kode for hydrofile aminosyrerester, bør ikke kode for glycosylering eller processeringssteder og skal fortrinsvis udvælges til at opretholde den samlede længde af lederen til at være ca. 50 aminosyrerester eller mindre, fortrinsvis fra ca. 23 til 40 amino syrerester og mest fortrinsvis fra ca. 25 til 35 aminosyrerester.

Som ovenfor anført er der umiddelbart 3' for ledersekvensen ifølge den foreliggende opfindelse et processeringssted, som muliggør at lederen spaltes fra ikke-gærproteinsekvensen, til hvilken den er fusioneret i forstadiepolypeptidet. Det processeringssted, som anvendes ved den foreliggende opfindelse, defineres som de codoner, der definerer det minimale antal af aminosyrerester, som specifikt genkendes til spaltning ved deri udvalgte proces (f.eks. kemisk, enzymatisk osv.). Inden for fagområdet kendes forskellige processeringssteder, herunder både de steder, som er aktive in vivo og in vitro. Der kan f.eks. med processeringsstedet tilvejebringes in vitro processering ved

kodning af et spaltningssted for et proteolytisk enzym, som ikke forekommer i gærværten. Det indvundne forstadiepolypeptid vil derefter blive behandlet med enzymet til spaltning af ikke-gærproteinet fra forstadiet. Et andet in vitro processeringssted er en methionincodon, som kan spaltes ved post ekspressionsbehandling med cyanogenbromid. Se f.eks. U.S. patent nr. 4.366.246.

In vivo processeringssteder kan udvælges blandt hvilke som helst peptidsignaler, som genkendes af et proteolytisk gærenzym, som vil tilvejebringe ekspression af den ønskede ikke-gærproteinsekvens. Specielt foretrukne processeringssteder er stederne for de enzymer, som er involveret i processering af nativ præpro-α-faktor For eksempel fjerner dipeptidylaminopeptidase A (DPAPase A) terminale -X-Ala-sekvenser, hvor X er Glu eller Asp. Se f.eks. Julius et al. (1983), Cell 32:839-852. Den endopeptidase, som kodes af KEX2-genet, spalter basiske dipeptider omfattende Lys- og Arg-rester, dvs. Lys-Arg, Arg-Arg, Arg-Lys og Lys-Lys. Fuller et al., Microbiology 1986, s. 273-278 (1986). I gær spaltes α-faktor forstadiet først af KEX2-endopeptidasen og derefter trimmes den N-terminale ende af DPAPase A til tilvejebringelse af det modne α-faktor pheromon. Eftersom det ser ud til at sidstnævnte proteolytiske proces er et hastighedsbegrænsende trin, foretrækkes det at eliminere signalerne for DPAPase A, således at processeringsstedet kun omfatter signalet for KEX2-endopeptidasen. I en sådan udførelsesform vil ledersekvensen derfor være sammenføjet med ikke-gærproteinsekvensen ved det dibasiske peptidgenkendelsessted for KEX2-endopeptidase, såsom Lys-Arg eller Arg-Arg.

Den carboxyterminale del af forstadiepolypeptidet ifølge den foreliggende opfindelse er ét ikke-gærprotein. Den DNA-sekvens, som koder for denne del, defineres her som begyndende med den første codon nedenstrøms (3'-retningen) den sidste codon i processeringsstedet til translationsstopcodonen, som definerer den carboxyterminale ende af forstadiepolypeptidet. Denne DNA-sekvens vil betragtes som kodende for et "ikke-gærprotein", når det

set over hele sekvensen definerer et polypeptid, som ikke i det væsentlige er homologt med et polypeptid, som udtrykkes af gær. De foretrukne ikke-gærproteiner vil almindeligvis være pattedyrproteinsekvenser (herunder deres analoger, dvs. "muteiner", fragmenter osv.). Som her defineret kan ikke-gærproteiner derfor omfatte et fusionsprotein, som omfatter pattedyr- og gærsekvenser, såvel som "pro"-former af det modne pattedyrprotein.

DNA-sekvenser, som koder for ikke-gærproteinerne, kan være sekven ser, som er klonet fra ikke-gærorganismer, eller de kan være syntetiske sekvenser, der sædvanligvis fremstilles under anvendelse af gærforetrukne codoner. Sædvanligvis vil ikke-gærproteinerne have en længde på mindst ca. 8 aminosyrer og kan omfatte polypeptider på op til ca. 100.000 dalton eller mere. Sædvanligvis vil ikke polypeptidsekvensen være mindre end ca. 300.000 dalton og mere sædvanligvis mindre end ca. 150.000 dalton. Af speciel interesse er polypeptider på fra ca. 5000 til ca. 150.000 dalton, mere foretrukket fra ca. 5000 til ca. 100.000 dalton. Illustrative ikke-gærproteiner, som har interesse, omfatter hormoner og faktorer, såsom væksthormon, somatomediner, epidermal vækstfaktor, luteiniserende hormon, thyroidstimulerende hormon, oxytocin, insulin, vasopressin, renin, caicitonin, follikelstimulerende hormon, prolactin, erythropoietin, kolonistimulerende faktorer, lymfokiner, såsom interleukin-2, globiner, immunoglobuliner, interferoner (f.eks. a, b eller q), enzymer, b-endorphin, enkephalin, dynorphin, insulinlignende vækstfaktorer mv.

I en foretrukken udførelsesform for den foreliggende opfindelse har DNA-konstruktioner, som koder for de ovenfor beskrevne forstadie polypeptider, strukturen:

$$5'\text{-AF-CHO-}X_n\text{-S-gen*-}3'$$

hvor AF koder for et gær-α-faktor signal, CHO koder for et glycosyleringssted, X_n koder for et polypeptid med en længde på n aminosyrer, og som ikke

indeholder et glycosyleringssted eller et processeringssted, som gør det muligt for gæren at spalte forstadiepolypeptidet in vivo, n er et helt tal på fra 0 til ca. 30, gen* koder for et ikke-gærprotein, og S koder for et processeringssted, som muliggør spaltning af forstadiepolypeptidet.

Det signalpeptid, som kodes af AF, er det samme α-faktor signalpeptid, som er beskrevet ovenfor. Det har en længde på ca. 20 aminosyrerester (f.eks. ca. 19-23) og er tilstrækkeligt langt til at styre forstadiepolypeptidet ind i gærens sekretoriske bane. Den præcise minimale eller maksimale længde kan bestemmes for en bestemt α-faktor ved at screene en serie af deletionskonstruktioner.

Den DNA-sekvens, der defineres af CHO, koder for et glycosyleringssted. Den vil sædvanligvis have en længde på 9 nukleotider inkluderende 3 codoner for aminosyrerne Asn-Y-Y'-, hvor Y er en hvilken som helst aminosyrerest, og Y' er Thr eller Ser.

X_n koder, hvis den er til stede, for f.eks. dele af α-faktor lede ren, som ikke er deleteret, eller ikke-beslægtede aminosyre sekvenser. X_n vil almindeligvis have en maksimal længde på ca. 30 aminosyrerester, mere fortrinsvis en maksimal længde på ca. 20 aminosyrerester og mest fortrinsvis en maksimal længde på ca. 10 aminosyrerester. Om end det ikke er nødvendigt, at X_n for nogen som helst polypeptider (dvs. n = 0), kan det være ønskeligt at tilvejebringe nogen afstand mellem glycosyleringsstedet CHO og processeringsstedet S i det tilfælde, at kulhydrattilføjelser på glycosyleringsstedet sterisk hindrer, at det middel, som spalter processeringsstederne, får adgang. I sådanne tilfælde vil n sædvanligvis have et minimum på ca. 1, mere fortrinsvis et minimum på ca. 2, og mest fortrinsvis et minimum på ca. 3.

Hvis X er til stede, foretrækkes det, at det ikke indeholder nogen funktionelle glycosyleringssteder eller processeringssteder, som genkendes og spaltes af

gærværten. Når den afviger fra sekvenser, som findes i en α-faktor leder, foretrækkes det ydermere at vælge hydrofile aminosyrerester. Det er muligt, at længden af X_n vil påvirke effektiviteten af ekspressionen og secerneringen af ikke-gær proteinet. Udvælgelse af en passende længde af X til optimering af ekspression kan udføres ved screening af konstruktioner med for skellige størrelser.

Ikke-gærproteinet, som kodes af gen*, og processeringsstedet, der kodes af 5, er som ovenfor beskrevet.

DNA-konstruktionerne ifølge den foreliggende opfindelse vil sædvanligvis opretholdes i en replicon, som kan opretholdes stabilt i en vært, navnlig en gærvært. Repliconerne, sædvanligvis plasmider, vil omfatte et eller flere replikationssystemer, ønskeligt to replikationssystemer, hvilket muliggør, at repliconen bevares i både en gærvært til ekspression og i en prokaryotisk vært ved kloning. Eksempler på sådanne gær-bakteriefærgevektorer omfatter YEp24 (Botstein et al. (1979), Gene 8:17-24), pCl/1 (Brake et al. (1984), Proc. Nat]. Acad. Sci. USA, 81:4642-4646) og YRp17 (Stinchomb et al. (1982), J. Mol. Biol. 158:157). Endvidere kan plasmidekspressionsvektoren være et plasmid med et højt eller lavt kopiantal, hvor kopiantallet sædvanligvis ligger på fra ca. 1 til ca. 200. Med gærvektorer med et højt kopiantal vil der almindeligvis være mindst 10, fortrinsvis mindst 20 og sædvanligvis ikke over ca. 150 kopier i en enkelt vært. Afhængigt af det udvalgte ikke-gærprotein kan en vektor med et højt eller lavt kopiantal være ønskelig, hvilket afhænger af den virkning, vektoren og det fremmede protein har på værten. Se f.eks. Brake et al., se ovenfor. DNA-konstruktioner ifølge den foreliggende opfindelse kan også integreres i gærgenomet med en integrerende vektor. Eksempler på sådanne vektorer kendes inden for fagområdet. Se f.eks. Botstein et al., se ovenfor.

En fagmand kan udvælge egnede gærværter og andre mikroorganismeværter til udøvelse af den foreliggende opfindelse. Når gærværter udvælges til ekspression, kan egnede værter omfatte de værter, som bl.a. har vist god secerneringskapacitet, lav proteolytisk aktivitet og en samlet robusthed. Gær og andre mikroorganismer kan sædvanligvis fås fra en række kilder, herunder Yeast Genetic Stock Center, Department of Biophysics and Medical Physics, University of California, Berkeley, California samt American Type Culture Collection, Rockville, Maryland.

Metoder til indføring af exogent DNA i gærværter er velkendte inden for fagområdet. Der er en lang række forskellige måder at transformere gær på. Belæringer om spheroplasttransformation findes f.eks. i Hinnen et al. (1978), Proc. Nati . Acad. Sci. USA, 75:1919-1933, og Stinchomb et al., EPO publikation nr. 45.573. Transformanter dyrkes i et hensigtsmæssigt næringsmedium, når det hensigtsmæssigt opretholdes under et selektivt tryk for at sikre bevarelse af endogent DNA. Når ekspressionen er inducibel, kan vækst af gærværten tillades til opnåelse af en høj tæthed af celler, og ekspressionen kan derefter induceres. Det secernerede, modnede ikke-gærprotein kan høstes ved hjælp af en hvilken som helst traditionel måde og oprenses ved kromatografi, elektroforese, dialyse, opløsningsmiddel-opløsningsmiddel ekstraktion og lignende.

Eksempler

De efterfølgende eksempler gives alene til illustration og har ikke til hensigt at begrænse den foreliggende opfindelses rækkevidde. Det antages, at deponering af de biologiske udgangsmaterialer ikke er nødvendig til udøvelse af den foreliggende opfindelse, eftersom enten det samme eller ækvivalent materiale er offentligt tilgængeligt.

Eksempel I

Med det efterfølgende eksempel tilvejebringes en sammenligning af ekspressions- og secerneringsniveauer opnået med modificerede α-faktor konstruktioner anvendt til at udtrykke human-proinsulin. Tre konstruktioner udnytter α-faktor ledere i fuld længde; én har en α-faktor leder med de tre native glycosyleringssteder, hos én er alle tre glycosyleringssteder elimineret og hos én er alle stederne, med undtagelse af ét ved Asn_{23} fjernet. Den fjerde konstruktion er en trunkeret α-faktor leder, som bevarer et enkelt glycosyleringssted ved Asn_{23}.

A. pYGAI1

Dette plasmid koder for en α-faktor leder (Brake et al. (1984), Proc. Natl. Acad. Sci. USA, 81:4642-4646, EPO publikation nr. 116.201) koblet til human-proinsulin. Proinsulinet er kodet af et syntetisk gen, der er fremstillet med gærforetrukne codoner (figur 30 1). α-Faktor ledersekvensen, det syntetiske proinsulingen og α-faktor terminatorsekvensen er fra pYBCA5, hvis konstruktion er vist i figur 1. Transskriptionen medieres af det 404 bp lange Bam HI-Nco I-GAPDH-promotorfragment. Travis et al. (1985), J. Biol. Chem., 260:4384-4389. Den 1206 bp lange Bam HI-ekspressionskassette omfattende GAPDH-promotoren, den sekvens, som koder for α-faktor lederen koblet til proinsulin og α-faktor terminatoren, blev klonet i det unikke Bam HI-sted i gærfærgevektoren pAB24 (nedenfor) eller pCI/1, således at GAPDH-promotorsekvensen var proximal til Sal I-stedet i vektoren til frembringelse af plasmiderne pYGAI1-AB24 henholdsvis pYGAI1-C1/1. Den 1206 bp lange Ban HJ-ekspressionskassette blev også subklonet i det unikke Bam HI-sted i et derivat af pBR322 (pBR322(WEco RI-Sal I)Bam HI; Travis et al., se ovenfor). Dette plasmid blev kaldt pGAI1.

Plasmid pAB24 (figur 4) er en gærfærgevektor, som indeholder den komplette 2m-sekvens (Broach, i Molecular Biology of the Yeast Saccharomyces,

bind 1, s. 445 (1 og pBR322-sekvenser. Det indeholder også gær URA3-genet hidrørende fra plasmid YEp24 (Botstein et al. (1979), Gene 8:17) og gær LEU2d-genet hidrørende fra plasmid pCl/1. EPO publikation nr. 116.201. Plasmid pAB24 blev konstrueret ved at fordøje YEp24 med Eco RI og genligere vektoren til fjernelse af de partielle 2m-sekvenser. Det fremkomne plasmid, YEp24WRI, blev lineariseret ved fordøjelse med Cia I og ligeret med det komplette 2m-plasmid, som var blevet lineariseret med Cla 1. Det fremkomne plasmid, pCBou, blev derefter fordøjet med Xba I, og det 8605 bp lange vektorfragment blev gelisoleret. Det isolerede Xba I-fragment blev ligeret med et 4460 bp langt Xba I-fragment, som indeholdt LEU2d-genet isoleret fra pCl/1; orienteringen af LEU2d-genet er i samme retning som URA3-genet.

B. pYGAI3

Plasmid pYGAI3 adskiller sig fra pYGAI1 ved, at det koder for en 25 modificeret α-faktor leder, hvor codonerne for Asn ved resterne 23, 57 og 67 er blevet udskiftet til at kode for Gin, hvorved alle tre signaler for N-koblet glycosylering er elimineret.

α-faktor lederen og de N-terminale 13 aminosyrer af proinsulin, der 30 kodes af dette plasmid, blev konstrueret ved at ligere syntetiske oligoriukieotider til frembringelse af et 294 bp langt fragment med et 5' Nco I-fremspring og et 3' Hind III-fremspring, hvis sekvens er vist i figur 2. Sekvensen af hensigtsmæssige oligonukleotider blev ændret under syntesen, således at codoner, som specificerede Asn ved positionerne 23, 57 og 67 i den naturlige α-faktor leder, nu specificerer Gin ved de samme positioner. DNA-sekvensen, som specificerer de N-terminale 13 aminosyrer af proinsulin, var identisk med sekvensen i pYGAI1. Det 294 bp lange syntetiske DNA (Nco 1-Hind III) fragment blev indsat i stedet for det sammenlignelige fragment i pGAI1 og pYGAI1-C1/1, hvilket gav plasmiderne pGAI3 henholdsvis pYGAI3.

C. pYGAI8

Plasmid pYGAI8 indeholder DNA, som koder for en α-faktor leder, der eliminerer to af de tre glycosyleringssteder. $Asn_{57,67}$ er blevet modificeret til $Gln_{57,67}$. Det fremkomne plasmid har kun et enkelt glycosyleringssted ved position Asn_{23} pYGAI8 blev fremstillet på følgende måde.

Der blev først isoleret et 5'-fragment fra ekspressionskassetten i pGAII ved at skære med Hpa II, efterfulgt af skæring med Bam HI og derefter gelisolering af et 504 bp langt fragment indeholdende GAPDH-promotoren og den sekvens, som koder for aminosyreresterne 1-33 i α-faktor lederen. Derefter blev plasmid pYGAI3, som koder for en α-faktor leder, der mangler glycosyleringssteder, også sekventielt skåret med Hpa II og Bam HI, og der blev isoleret et 702 bp langt fragment indeholdende sekvenser, som koder for modificerede α-faktor lederrester 34-83, LysArg-processeringsstedet, proinsulin sekvensen og α-faktor termineringssekvensen. Dette fragment blev derefter ligeret til det 504 bp lange fragment fra pGAII, blev skåret med Bam HI, og et 1206 bp langt fragment blev isoleret.

Det ovenfor nævnte 1,2 kb lange Bam HI-fragment, som indeholdt en fuldstændig GAPDH-promotor/α-faktor leder/proinsulin/α-faktor termineringsekspressionskassette, blev derefter ligeret til Bam HI skåret og phosphatasebehandlet pBR322(WEco RI-Sal I)Bam HI til frembringelse af plasmid pGAI8, som blev klonet i E. coli.

Den 1,2 kb lange ekspressionskassette fra pGAI8 blev fjernet ved skæring med Bam HI, og derefter blev fragmentet gelisoleret. Det blev ligeret til Bam HI-skåret og phosphatasebehandlet gærfærgevektor pAB24. Indsættelsen af ekspressionskassetten var i det unikke Bam HI-sted i pBR322-sekvenserne, således at GAPDH-promotoren var proximalt det unikke Sal I-sted i vektoren. Dette plasmid var pYGAI8.

D. pYGAI7

Plasmid pYGAI7 indeholdt DNA, som koder for en trunkeret α-faktor leder og det syntetiske gen for human-proinsulin. α-Faktor lederen er blevet trunkeret således, at den kun koder for aminosyrer 1-35 i α-faktor lederen og derfor indeholder et enkelt sted for glycosylering ved Asn Denne gærekspressionsvektor blev konstrueret på følgende måde.

Først blev pGAII skåret med Hind III. En Hpa II-Hind III-linker med den efterfølgende struktur blev tilføjet:

5'-CGGCTAAAAGATTCGTTAACCAACACTTGTGTGGTTCTCACTTGGTTGA
CGATTTTCTAAGCAATTGGTTGTGAACACACCAAGAGTGAACCAACTTCGA-5'

Efter tilføjelsen af linkeren blev det lineariserede plasmid skåret med Bam HI, og et 558 bp langt Hpa II-Bam HI-fragment blev gelisoleret. Dette fragment indeholdt codonerne for aminosyreresterne 34-35 i α-faktor lederen koblet direkte til et Lys-Arg-processeringssted og proinsulinsekvensen. Der er ingen intervenerende sekvenser mellem codonen for aminosyrerest 35 i α-faktor lederen og processeringsstedet umiddelbart stødende op til proinsulinsekvensen;

Dernæst blev pGAII skåret med HpaII og Bam HI, og et 504 bp langt fragment blev gelisoleret. Dette fragment indeholdt GAPDH-promotoren og nukleotider, som koder for aminosyrerne 1-33 i α-faktor lederen, 3'-enden, som afsluttes i et Hpa II-fremspring, der er komplementært til 5'-enden i det ovenfor beskrevne 558 bp lange Hpa 11-Bam HI- fragment. Disse to fragmenter blev ligeret sammen og skåret med Bam HI til tilvejebringelse af en ekspressionskassette, som indeholdt GAPDU-promotoren, sekvenser, som koder for en modificeret α-faktor leder, der indeholder aminosyreresterne 1-35 direkte koblet til et Lys-Arg-processeringssted, proinsulingenet og α-faktor terminatoren. Kassetten blev derefter ligeret ind i Bam HI-stedet i pBR322(WEco 35

RI-Sal I)Bam HI, som ovenfor beskrevet, til frembringelse af plasmid pGAJ7 og blev klonet i E. coli.

pGAI7 blev derefter skåret med Bam HI, og den 1062 bp lange ekspressionskassette blev gelisoleret. Ekspressionskassetten blev derefter ligeret ind i Bam HI-stedet i pAB24 til frembringelse af plasmid pYGAI7.

E. Sammenlignende ekspression

Saccharomyces cerevisiae stamme AB1O3.1 (Mata, leu2-3,112,ura3--52, his4-580, pep4-3(cir°)) blev transformeret med plasmiderne pYGAI1-C1/1, pYGAI3, pYGAI1-AB24, pYGAI7 og pYGAI8 stort set som beskrevet af Hinnen et al. (1978), Proc. Natl. Acad. Sd. USA, 75:1929-1933. Transformanter med pYGAII-C1/1 og pYGAI3 blev udvalgt for leucinprototrofi, transformanter med de andre plasmider blev ud valgt for ura-prototrofi.

Resultaterne vist, i tabel 1 sammenligner secernering af proinsulin 15 medieret af den naturlige α-faktor leder (pYGAI1-C1/1) eller α-faktor lederen med Gin substitueret for Asn ved positionerne 23, 57 og 67 (pYGAI3). Inoculumkulturer (0,2 ml af individuelle transformanter) blev dyrket i 48 timer i komplet syntetisk medium uden leucin (SD-leu, Sherman et al., Methods in Yeast Genetics, s. 62 (Cold Spring Harbor Laboratory, 1982)) og blev fortyndet 20 gange i samme, medium. Kulturer blev dyrket. 48-72 timer, kultursupernatanterne blev fremstillet ved centrifugering og blev analyseret for immunoreaktiv krydsreagerende insulinlignende materiale (ILM) i en kompetitiv radioimmunanalyse med ^{125}I-mærket insulin. Som det fremgår af tabel 1, resulterede fjernelsen af de tre glycosyleringssteder fra α-faktor lederen i stort set ingen secernering af insulinlignende materiale sammenlignet med den secernering, som blev medieret af den native α-faktor leder.

Resultater i tabel 2 sammenligner transformanter af pYGAI1-pAB24 (nativ α-faktor leder i fuld længde), pYGAI8 (α-faktor leder i fuld længde med kun et

glycosyleringssted ved Asn og pYGAI7 (trunkeret α-faktor leder indeholdende et enkelt glycosyleringssted ved Asn for deres evne til at secernere insulinlignende materiale. Inoculumkulturer af de anførte transformanter (0,2 ml) i SD-Leu dyrket i 0,48 timer ved 30°C blev pelletteret ved centrifugering, vasket og fortyndet 20 til 50 gange i ura medium. Dette medium indeholder 0,67% gærnitrogenbase, 1% ravsyre, 0,35% NaOH, 0,5% casaminosyre, 2% glucose, 0,005% adenin, 0,01% tryptophan og 0,02% threonin. Kulturerne blev dyrket ved 30°C i 48-72 timer, og kultursupernatanterne blev fremstillet og analyseret som ovenfor beskrevet. Resultater anført i tabel 2 viser, at de transformanter, som bærer den konstruktion, der indeholder den trunkerede α-faktor leder under bevarelse af et enkelt glycosyleringssted ved Asn almindeligvis secernerede mere immunoreaktivt insulinlignende materiale end de transformanter, der bar konstruktionen med en nativ α-faktor leder i fuld længde. Transformanter, som bar konstruktionen med α-faktor lederen i fuld længde og med det samme enkelte glycosyleringssted (Asn_{23}), secernerede langt mindre insulinkrydsreagerende materiale end de transformanter, som bar den native α-faktor leder i fuld længde eller den trunkerede α-faktor leder.

TABEL 1

Effekten af eliminering af α-faktor lederglycosyleringssteder på secerneringen af insulinlignende materiale

Transformant	OD_{650}	ILM[1)]	
		mg/ml	mg/ml,OD_{650}
AB103.1(pYGAI1-C1/1)-1	5,9	0,24	0,04
-2	5,9	0,24	0,04
-3	2,1	0,08	0,04
AB103.1(pYGAI3) -1	ND	0,003	-
-2	ND	0,007	-

1) Krydsreagerende insulinlignende materiale (ILM) bestemt ved en kompetitiv radioimmunanalyse med ^{125}I-mærket insulin og insulin standarder. Resultaterne er angivet som ILM secerneret pr. ml kultur og i nogle tilfælde som ILM secerneret pr. ml normaliseret til en kulturcelletæthed med en absorbans på 1 ved bølgelængden 650 nm.

TABEL 2

Effekt af trunkeret α-faktor leder eller α-faktor leder i fuld længde med et enkelt glycosyleringssted ved Asn_{23} på secernering af insulinlignende materiale

Transformant	Antal forsøg[1)]	ILM[2)]					
		mg/ml			mg/ml, OD_{650}		
		Område	Gen.	Std. afv.	Område	Gen.	Std. afv.
AB103.1 (pYGAI1-AB24)	7	0,22-0,55	0,37	0,12	0,02-0,04	0,026	0,01
AB103.1 (pYGAI7)	8	0,38-1,0	0,62	0,24	0,03-0,06	0,043	0,01
AB103.1 (pYGAI8)	8	0,03-0,18	0,09	0,05	0,003-0,01	0,007	0,003

1) Der blev undersøgt et minimum på tre uafhængige transformanter.

2) Secerneret, krydsreaktivt insulinlignende materiale (ILM) blev bestemt ved en kompetitiv radioimmunanalyse med ^{125}I-mærket insulin og insulinstandarder. Resultaterne er angivet som ILM secerneret pr. ml kultur og som ILM secerneret pr. ml normaliseret til en kultur celletæthed med en absorbans på 1 ved bølgelængden 650 nm.

Eksempel II

I dette eksempel sammenlignes ekspressionen af en α-faktor leder konstruktion i fuld længde og med alle glycosyleringssteder med en ekspressionskonstruktion, der indeholder en trunkeret α-faktor sekvens, som kun bevarer et enkelt glycosyleringssted ved Asn_{23}. Det i dette eksempel anvendte ikke-gærprotein er en human-proinsulinana-log, hvor det forbindende "C"-peptid er blevet erstattet med et gær KEX2-endopeptidase-spaltningssted.

A. pYGAIC3

Plasmidet pGAIC3 blev fremstillet ved at udskifte det 231 bp lange Hind III-Sal I-fragment i pGAII, der koder for aminosyrerne 14 til 30 i B-kæden, C-peptidet, A-kæden og 2-translationsstopcodoner, med et 132 bp langt syntetisk Hind III-Sal I-genfragment (vist i figur 3), som koder for aminosyrerne 14 til 30 i B-kæden, et Lys-Arg KEX2-endopeptidase-spaltningssted, A-kæden og translationsstop codonerne. Plasmidet pYGAIC3 blev fremstillet ud fra pGAIC3 på følgende måde.

Plasmid pGAIC3 blev fordøjet med Bam HI, og den 1107 bp lange Bam HI-ekspressionskassette, som indeholdt GAPDH-promotoren, den sekvens, som koder for α-faktor lederen koblet til proinsulinanalogen, og α-faktortranskriptionsterminatoren, blev isoleret og ligeret ind i Bam HI-fordøjet og phosphatasebehandlet pAB24, og blev derefter klonet i E. coli. Der blev opnået et plasmid pYGAIC3, i hvilket ekspressionskassetten var således orienteret, at GAPDH-promotoren var proximalt det unikke Sal I-sted i vektoren.

B. pYαF_L7C3

Plasmid pYαF_L7C3 indeholder DNA, som koder for den for pYGAI7 oven for beskrevne trunkerede α-faktor leder koblet til sekvensen for proinsulinanalogen, som også er beskrevet ovenfor (pYGAIC3). Dette plasmid blev konstrueret på følgende måde.

Først blev pGAIC3 skåret med Hind III og Sal I, og der blev gelisoleret et 132 bp langt fragment. Dette fragment indeholdt sekvenser, som koder for alle med undtagelse af de første 12 codoner i proinsulinanalogen. Det blev ligeret ind i et gelisoleret 4640 bp langt fragment fra Hind III- og Sal I-fordøjet pGAI7 til tilvejebringelse af plasmid pYaFL7C3. Efter kloning i E. coli blev dette plasmid skåret med Bam HI, og der blev gelisoleret et 1062 bp langt 35 Bam HI-fragment. Denne ekspressionskassette indeholdt den trunkerede α-faktor lederkonstruktion pGAI7 med proinsulinanalogen i stedet for den normale

proinsulinsekvens. Ekspressionskassetten blev derefter ligeret ind i Bam HI-stedet i pAB24, som ovenfor beskrevet, til frembringelse af pYαF$_L$7C3.

Sammenlignende ekspression

Der blev bestemt ekspressionsniveauer for pYGAIC3 og pYαF$_L$7C3 i S. cerevisiae. Stamme AB1O3.1 er blevet beskrevet i eksempel I. Stamme ABJ1O-4 er et derivat af Saccharomyces cerevisiae stamme AB11O (Mata, leu2, ura3-52, pep4-3, his4-580(cir°)), i hvilket der er blevet indført en deletion i pep4-genet. Disse stammer blev transformeret som ovenfor beskrevet med plasmiderne pYGAIC3 og pYαF$_L$7C3, og der blev selekteret ura-prototrofe stammer. Inoculumkulturer blev dyrket i SD-leu (Sherman et al., se ovenfor) ved 30°C i 24-48 timer, blev derefter pelleteret ved centrifugering, blev vasket og fortyndet 20 gange i ura medium (beskrevet ovenfor) og blev dyrket i 48-72 timer ved 30°C. Cellefrit konditioneret kulturmedium blev fremstillet ved centrifugering til analyse i en kompetitiv insulin-radioimmunanalyse.

Resultaterne er vist i tabel 3. Som det fremgår, medierer den trunkerede α-faktor konstruktion forøget secernering af immunoreaktiv proinsulinanalog sammenlignet med den naturlige α-faktor ledersekvens.

TABEL 3

Secernering af ILM fra en proinsulinanalogkonstruktion medieret af en trunkeret α-faktor leder eller en naturlig α-faktor leder

Transformant	Antal forsøg[1)]	ILM[2)] mg/ml Område	Gen.	Std. afv.	ILM[2)] mg/ml,OD_{650} Område	Gen.	Std. afv.
AB103.1 (pYGAIC3)	6	1-2,75	1,66	0,65	0,11-0,20	0,14	0,04
AB103.1 (pYαF_L7C3)	6	1,5-6,63	4,46	2,15	0,14-0,60	0,40	0,19
AB110.4 (pYGAIC3)	3	1,15-1,4	1,28	0,13	0,10-0,12	0,11	0,01
AB110.4 (pYαF_L7C3)	3	2,15-4,12	3,46	1,14	0,19-0,38	0,31	0,10

1) Der blev undersøgt et minimum på tre uafhængige transformanter.

2) Secerneret, krydsreaktivt insulinlignende materiale (ILM) blev bestemt ved en kompetitiv radioimmunanalyse med ^{125}I-mærket insulin og insulinstandarder. Resultaterne er angivet som mængder af ILM pr. ml kultur og som mængde secerneret pr. ml normaliseret til en celletæthed med en absorbans på 1 ved bølgelængden 650 nm.

Eksempel III

I dette eksempel beskrives konstruktionen af en trunkeret α-faktor ekspressionsvektor, som medierer forøgede ekspressionsniveauer af aktiv insulinlignende vækstfaktor-1.

Der blev først fremstillet en DNA-sekvens, som koder for en trunkeret α-faktor leder og en kodende sekvens for IGF1. En syntetisk sekvens blev

fremstillet ved standardmetoder under anvendelse af en "Applied Biosystems 380A DNA-syntesemaskine i overensstemmelse med leverandørens retningslinier. Der blev syntetiseret 14 DNA-sekvenser med en længde på fra 22 til 57 baser, de blev oprenset ved PAGE og individuelt phosphoryleret med T4-kinase i nærværelse af ATP. Sekvenserne blev derefter annelleret (eng.: annealed) og ligeret ved standardprocedurer.

Sekvensen af det syntetiske gen er vist i figur 5. Det oprensede, syntetiske genfragment blev klonet i Nco I/Sal 1-fordøjet pBS100 (beskrevet nedenfor). Det fremkomne plasmid blev kaldt pBS100 TαF$_L$IGF1.

Plasmid pBS100 indeholder en gærekspressionskassette klonet i et pBR322-derivat, pAB12. Ekspressionskasseten indeholder hybrid ADH-2/GAPDH-promotoren og GAPDH-terminatoren, som flankerer et ikke-essentielt gensegment. ADH-2/GADPH-promotoren er et 1200 bp langt Bam HI-Nco I-fragment isoleret fra pJS103 (se nedenfor), og GAPDH terminatoren er et 900 bp langt Sal I-Bam 111-fragment isoleret fra plasmid pPAGi. EPO publikation nr. 164.556. Plasmid pBS100 indeholder også et ikke-essentielt fragment mellem Nco I- og Sal I- stederne, hvilket fragment erstattes med de genfragmenter, som har interesse. Ekspressionskassetten kan fjernes fra pBS100 ved fordøjelse med Bam HI og klones i gærfærgevektorer til indføring i gærceller.

Plasmid pJSIO3, der indeholder hybrid-ADH-2/GAPDG-promotoren, som er anvendt ovenfor, blev konstrueret på følgende måde. ADH-2-delen af promotoren blev konstrueret ved at skære et plasmid, som indeholder vildtype ADH2-genet fra plasmid pADR2 (Beier et al. (1982), Nature 300:724-728) med restriktionsenzymet Eco R5, som skærer ved position +66 i forhold til ATG-startcodonen, såvel som i to andre steder i ADR2, uden for ADH2-området. Den fremkomne blanding af et vektor fragment og to mindre fragmenter blev omsat med Bal3l-exonuclease til fjernelse af ca. 300 bp. Synteti-

ske Xho 1-linkere blev ligeret med det Bal3l-behandlede DNA. Det fremkomne DNA-linker-vektor fragment (ca. 5 kb) blev adskilt fra linkerne ved søjlekromatografi, blev skåret med restriktionsenzymet Xho I, blev genligeret og anvendt til at transformere E. coli til ampicillinresistens. Positionerne af Xho 1-linkeren blev bestemt ved DNA-sekventering. Et plasmid, som indeholdt en Xho 1-linker i det ikke-transkriberede 5 5'-område af ADH2-genet (position -232 fra ATG) blev skåret med restriktionsenzymet Xho I, blev behandlet med nuclease Sl og blev derefter behandlet med restriktionsenzymet Eco RI til frembringelse af et lineært vektormolekyle med i stump ende ved Xho 1-linkeren og en Eco RI-ende. GAPDH-delen af promotoren blev konstrueret ved at skære plasmid pPGAP (publikation nr. 164.556) med enzymerne Bam HI og Eco RI efterfulgt af isolering af et 0,4 kbp langt DNA-fragment. Dette oprensede fragment blev derefter fuldstændigt fordøjet med enzymet Alu I, og der blev isoleret et ca. 200 bp langt fragment. Dette GAPDH-promotorfragment blev ligeret med det ADH-2-fragment, som var til stede på den ovenfor beskrevne lineære vektor, til frembringelse af plasmidet pJS1O3.

Der blev isoleret et Bam HI-fragment fra pBS100TaFL IGF1. Dette fragment indeholdt ADH2/GAPDH-promotoren, en trunkeret α-faktor leder (AA 1-25, 81-83), et LysArg-processeringssted, en kodende sekvens for JGF1 og GAPDH-terminatorsekvensen. Dette Bam HI-fragment blev derefter klonet i pAB24, som tidligere var blevet fordøjet med Bam HI. Der blev udvalgt en positiv klon, og medens den i starten blev kaldt plasmid 18.5, blev den derefter kaldt pYLUJGFI-55. (Se figur 6).

Der blev også fremstillet en anden ekspressionsvektor, pYLUIGFI-24, ved hjælp af analoge metoder. Et restriktionskort er vist i figur 7. Denne vektor ligner pYLUIGF1-55 med undtagelse af, at den til styring af secernering har en α-faktor i fuld længde med tre glycosyleringssteder (sammenlign eksempel I.A.) og α-faktor terminatoren.

Gærstammen AB110 (EPO publikation nr. 164.556) blev transformeret med pYLUIGF1-55 og pYLUIGF1-24 ved traditionelle spheroplastdannende metoder (Hinnen et al. (1978), Proc. Nati. Acad. Sd. USA, 75:1919-1933), og ekspression blev sammenlignet.

Ekspressionen af IGFI fra AB110 (pYLUIGF1-55) og AB110 (pYLUIGFI-25) er ikke-konstitutiv. Induktion af IGF1-ekspression blev opnået ved at tilvejebringe en lav koncentration af glukose i vækstmediet. Under standardbetingelser udnytter rysteflaskekulturer (25 ml) fuldt ud glukosen i mediet 18-24 timer efter inokulering. 25 ml Kultur af AB110 (pYLUIGF1-55) og AB110 (pXLUIGF1-24) blev således dyrket under standardbetingelser i 72 timer. Supernatantprøver blev taget 49 og 72 timer efter inokulering og blev analyseret for biologisk IGF1-aktivitet (RRA) og for immunoreaktivitet (RIA) med antistoffer mod AGF1. Som det fremgår secernerede pYLUIGF1-55 med et trunkeret α-faktor lederprotein, hvoraf en væsentligt større fraktion var biologisk aktiv. Skønt pYLUIGF1-24 secernerede mere protein, som udviste reaktivitet med IFG1-antistoffer, var relativt lidt af dette protein biologisk aktivt.

Resultaterne er vist i tabel 4. Radioreceptoranalysen (RRA) måler IGF-1's evne til at bindes til dens receptor. Dette er et mål for det rekombinerede polypeptids biologiske aktivitet, eftersom det formodes, at IGF-1 udviser hele sin aktivitet via dens receptor. Receptoranalysen er beskrevet i Marshall et al. (1974), J.Clin. Endocrinol. Metab. 19:283-292. Radioimmunoanalysen (RIA) er en kompetitiv analyse, der måler mængden af protein, som antigent krydsreagerer med nativ IFG-1, hvad enten det er biologisk aktivt eller ej. Analysen er beskrevet i Zapf et al. (1981), 3. Clin. Invest. 68:1321-1330.

TABEL 4

Secernering af IFG1 medieret af en tunkeret α-faktor leder eller en naturlig α-faktor leder

Transformant	49 timer		72 timer	
	RRA[1)]	RIA[2)]	RRA	RIA
AB110 (pYLUIGF1-55)	1,0	14	1,3	14
AB110 (pYLUIGF1-24)	1,3	54	2,7	66

1) mg/ml
2) mg/ml

Deponering af biologiske materialer

Følgende ekspressionsvektorer blev deponeret hos American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockwille, Maryland, USA, og vil blive opretholdt under retningslinierne ifølge Budapesttraktaten. Deponeringsnumrene og datoerne for deponeringen er angivet nedenfor.

Deponeret materiale	ATCC-nr.	Deponeringsdato
E. coli (pYGAI7)	67597	29/12/87
E. coli (pYαF$_L$7C3)	67596	29/12/87
E. coli (pYLUIGF1-55)	67595	29/12/87

Disse deponeringer er foretaget af hensyn til en fagmand. Disse deponeringer er hverken en indrømmelse af, at sådanne deponeringer er nødvendige til udøvelse af den foreliggende opfindelse, ej heller at ækvivalente udførelsesformer ikke ligger inden for en fagmands område på baggrund af den foreliggende beskrivelse. Den offentlige tilgængelighed til disse deponeringer er ikke en tildeling af licens til fremstilling, anvendelse eller salg af de under

dette patent eller et andet patent deponerede materialer. Nukleinsyresekvenserne i de deponerede materialer er inkorporeret i den foreliggende beskrivelse ved henvisning og er de styrende, hvis der er konflikt med nogen som helst af de her beskrevne sekvenser.

Om end ovennævnte opfindelse er blevet beskrevet i nogen detalje for at illustrere opfindelsen, er det indlysende, at ændringer og modifikationer kan foretages af en fagmand inden for de tilhørende kravs rækkevidde.

Patentkrav

1. Gærcelle omfattende en DNA-konstruktion, der tilvejebringer ekspression og secernering af et ikke-gærprotein, hvilken DNA-konstruktion omfatter en kodende sekvens under styring af gærgenkendte transkriptionsinitierings- og terminerings sekvenser, hvilken kodende sekvens koder for et forstadiepolypeptid, som omfatter en ledersekvens og ikke-gærproteinet koblet sammen af et processeringssted, som muliggør spaltning af ikke-gærproteinet fra forstadiepolypeptidet, hvorhos ledersekvensen udgør fra ca. 25 til 50 N-terminale aminosyrerester af forstadiepolypeptidet og omfatter signalpeptidet af et Saccharomyces eller Kluveromyces alfa-faktor forstadie og et enkelt glycosyleringssted.

2. Celle ifølge krav 1, k e n d e t e g n e t ved, at ikke-gærproteinet er et pattedyrprotein, og at alfa-faktoren er en Saccharomyces alfa-faktor.

3. Celle ifølge krav 2, k e n d e t e g n e t ved, at pattedyrproteinet er en human insulin-lignende vækstfaktor I eller et forstadie til human-insulin.

4. Celle ifølge krav 3, k e n d e t e g n e t ved, at forstadiet til human-insulin er human-proinsulin eller forstadiet til human-insulin omfatter insulin a-kæde og insulin b-kæde koblet sammen af et gærgenkendt processeringssted, som spaltes in vivo.

5. Celle ifølge krav 4, k e n d e t e g n e t ved, at processeringsstedet spaltes af KEX2-genproduktet af Saccharomyces.

6. Celle ifølge ethvert af de foregående krav, k e n d e t e g n e t ved, at gær cellen er af slægten Saccharomyces.

7. Celle ifølge krav 6, k e n d e t e g n e t ved, at gær cellen er S. cerevisiae.

8. Celle ifølge krav 7, k e n d e t e g n e t ved, at gær-α-faktor forstadiet er S. cerevisiae MFα1.

9. Dobbeltstrenget DNA-molekyle, k e n d e t e g n e t ved, at det omfatter et område, der koder for et forstadiepolypeptid, som kan secerneres af en gærvært, hvilket område, under henvisning til en af strengene, har strukturen:

$5'$-AF-CHO-X_n-S-gen*-$3'$

hvor AF-CHO-X_n omfatter et enkelt funktionelt glycosyleringssted i hvilket:

AF koder for et Saccharomyces eller Kluveromyces α-faktor signalpeptid,

CHO koder for et glycosyleringssted,

X_n koder for et polypeptid med en længde på n aminosyrer, og som ikke indeholder et glycosyleringssted eller et processeringssted, som gør det muligt for gæren at spalte forstadiepolypeptidet in vivo,

.. er et helt tal på fra 0 til ca. 30,

gen* koder for et ikke-gærprotein, og

S koder for et processeringssted, som muliggør spaltning af forstadiepolypeptidet.

10. Celle ifølge krav 9, k e n d e t e g n e t ved, at AF koder for et polypeptid med en længde på fra ca. 19 til 23 aminosyrer.

11. DNA-molekyle ifølge krav 9 eller 10, k e n d e t e g n e t ved, at n er et helt tal på fra ca. 0 til ca. 20

12. DNA-molekyle ifølge krav 11, k e n d e t e g n e t ved, at n er et helt tal på fra ca. 0 til ca. 10.

13. DNA-molekyle ifølge krav 12, k e n d e t e g n e t ved, at n er et helt tal på fra ca. 3 til ca. 10.

14. DNA-molekyle ifølge ethvert af kravene 9-13, k e n d e t e g n e t ved, at gær værten er en Saccharomyces.

15. DNA-molekyle ifølge ethvert af kravene 9-14, k e n d e t e g n e t ved, at gær-α-faktor signalpeptidet er et Saccharomyces signalpeptid.

16. DNA-molekyle ifølge ethvert af kravene 9-15, k e n d e t e g n e t ved at S koder for et processeringssted, der genkendes in vivo af gærværten.

17. DNA-molekyle ifølge krav 16, k e n d e t e g n e t ved, at S koder for et dipeptid, som genkendes af KEX2-endopeptidasen.

18. DNA-molekyle ifølge krav 17, k e n d e t e g n e t ved, at dipeptidet er 5'-Lys-Arg-3' eller 5'-Arg-Arg-3'.

19. DNA ifølge ethvert af kravene 9-18, k e n d e t e g n e t ved, at det omfatter en replicon.

20. DNA-molekyle ifølge krav 19, k e n d e t e g n e t ved, at det område, som koder for forstadiepolypeptidet, er under styring af gærgenkendte transkriptionsinitierings- og termineringssekvenser, og at repliconen er en gærreplicon.

21. DNA-molekyle ifølge krav 20, k e n d e t e g n e t ved, at repliconen er et plasmid eller et kromosom.

22. Fremgangsmåde til fremstilling af et ikk-gær protein, hvilket omfatter vækst af gær celle ifølge ethvert af kravene 1-8 under betingelser hvorved ikke-gæren er udtrykt.

23. Fremgangsmåde ifølge krav 22, hvor gær cellen omfatter et DNA molekyle ifølge ethvert af kravene 9-21.

24. Fremgangsmåde ifølge krav 22 eller 23, som endvidere omfatter indvinding af ikke-gær proteinet i oprenset form.

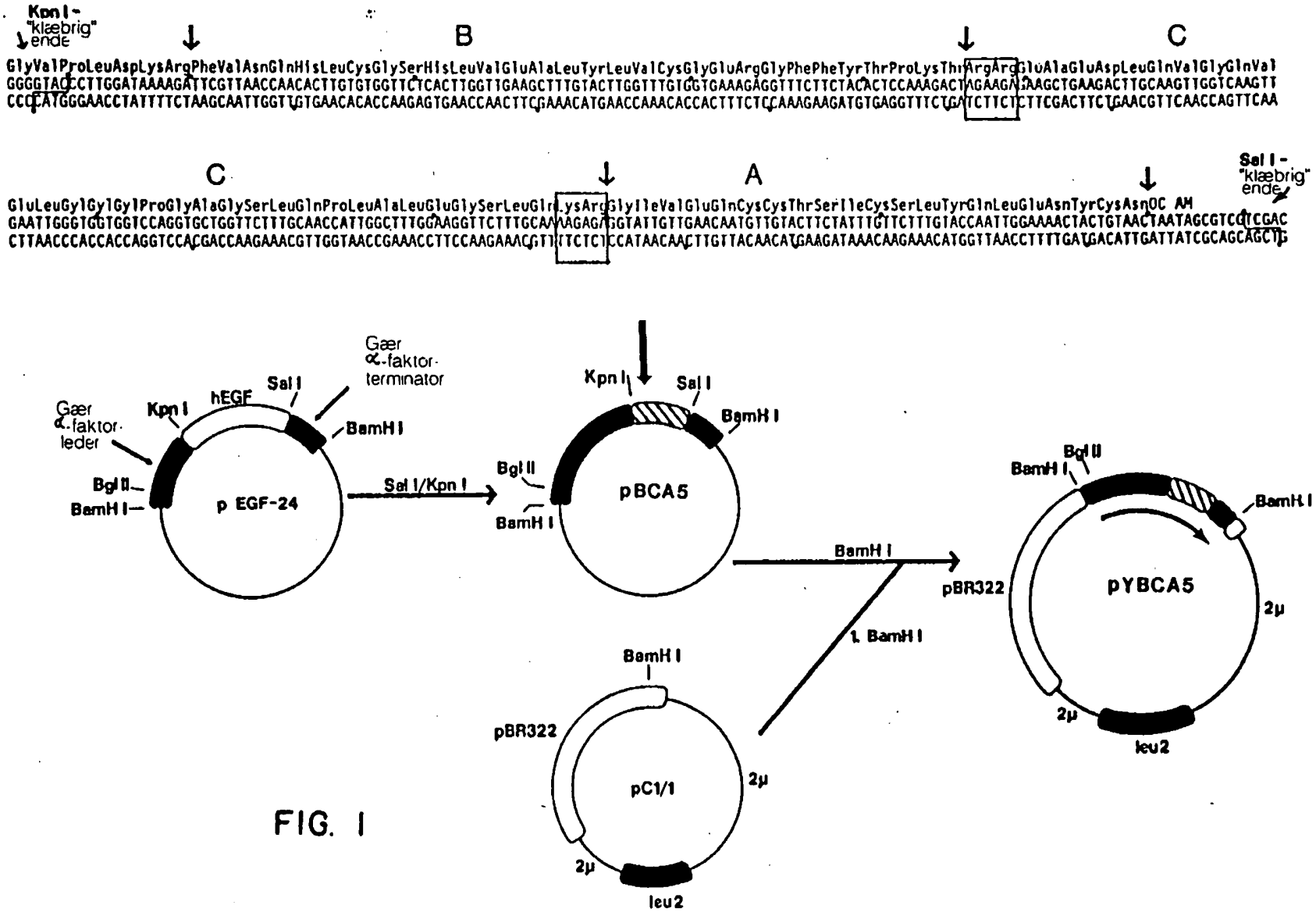
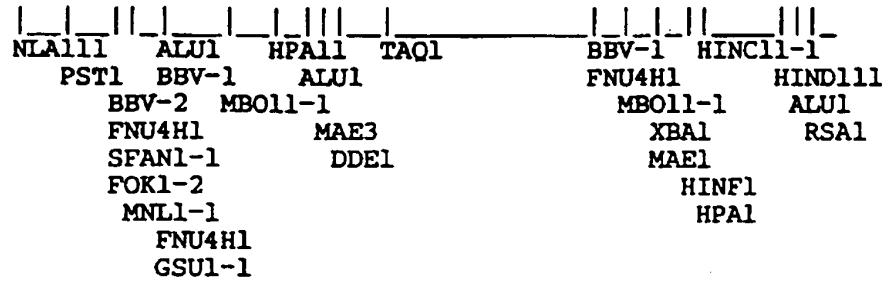


FIG. 1

DK 175910 B1

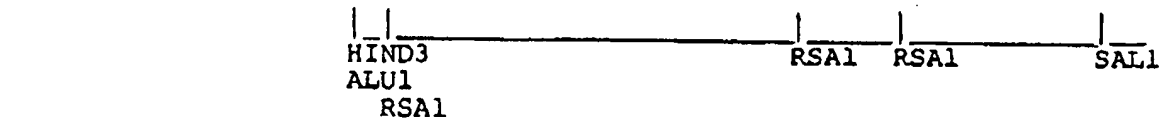
Sekvens af syntetisk genfragment af pYGAI3




```
   MetArgPheProSerIlePheThrAlaValLeuPheAlaAlaSerSerAlaLeuAlaAla
CATGAGATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCATTAGCTGCT
TACTCTAAAGGAAGTTAAAAATGACGTCAAAATAAGCGTCGTAGGAGGCGTAATCGACGA

1 NLA111, 24 PST1, 38 BBV FNU4H1, 41 SFAN1, 42 FOK1, 45 MNL1
, 55 ALU1, 56 BBV FNU4H1, 60 GSU1,

   ProValGlnThrThrThrGluAspGluThrAlaGlnIleProAlaGluAlaValIleGly
62 CCAGTCCAAACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT
   GGTCAGGTTTGATGTTGTCTTCTACTTTGCCGTGTTTAAGGCCGACTTCGACAGTAGCCA

   80 MBO11, 101 HPA11, 109 ALU1, 120 MAE3,

   TyrLeuAspLeuGluGlyAspPheAspValAlaValLeuProPheSerGlnSerThrAsn
122 TACTTAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATTTTCCCAAAGCACAAAT
   ATGAATCTAAATCTTCCCCTAAAGCTACAACGACAAAACGGTAAAAGGGTTTCGTGTTTA

   124 DDE1, 144 TAQ1,

   AsnGlyLeuLeuPheIleGlnThrThrIleAlaSerIleAlaAlaLysGluGluGlyVal
182 AACGGGTTATTGTTTATACAAACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
   TTGCCCAATAACAAATATGTTTGATGATAACGGTCGTAACGACGATTTCTTCTTCCCCAT

   221 BBV FNU4H1, 230 MBO11,

   SerLeuAspLysArgPheValAsnGlnHisLeuCysGlySerHisLeuValGluAlaLeu
242 TCTCTAGATAAAAGATTCGTTAACCAACACTTGTGTGGTTCTCACTTGGTTGAAGCTTTG
   AGAGATCTATTTTCTAAGCAATTGGTTGTGAACACACCAAGAGTGAACCAACTTCGAAAC

   244 XBA1, 245 MAE1, 255 HINF1, 260 HINC11 HPA1, 294 HIND111,
    295 ALU1, 301 RSA1,

   Tyr
302 TACTT
   ATGAA
```

FIG. 2

Sekvens af syntetisk genfragment af pYGAIC3




```
    GluAlaLeuTyrLeuValCysGlyGluArgGlyPhePheTyrThrProLysThrLysArg
1   GAAGCTTTGTACTTGGTTTGTGGTGAAAGAGGTTTCTTCTACACTCCAAAGACTAAGAGA
    CTTCGAAACATGAACCAAACACCACTTTCTCCAAAGAAGATGTGAGGTTTCTGATTCTCT

    2 HIND3, 3 ALU1, 9 RSA1,

    GlyIleValGluGlnCysCysThrSerIleCysSerLeuTyrGlnLeuGluAsnTyrCys
61  GGTATTGTTGAACAATGTTGTACTTCTATTTGTTCTTTGTACCAATTGGAAAACTACTGT
    CCATAACAACTTGTTACAACATGAAGATAAACAAGAAACATGGTTAACCTTTTGATGACA

    80 RSA1, 99 RSA1,

    AsnOC AM
121 AACTAATAGCGTCGTCGAC
    TTGATTATCGCAGCAGCTG

    134 SAL1,
```

FIG. 3

181

DK 175910 B1

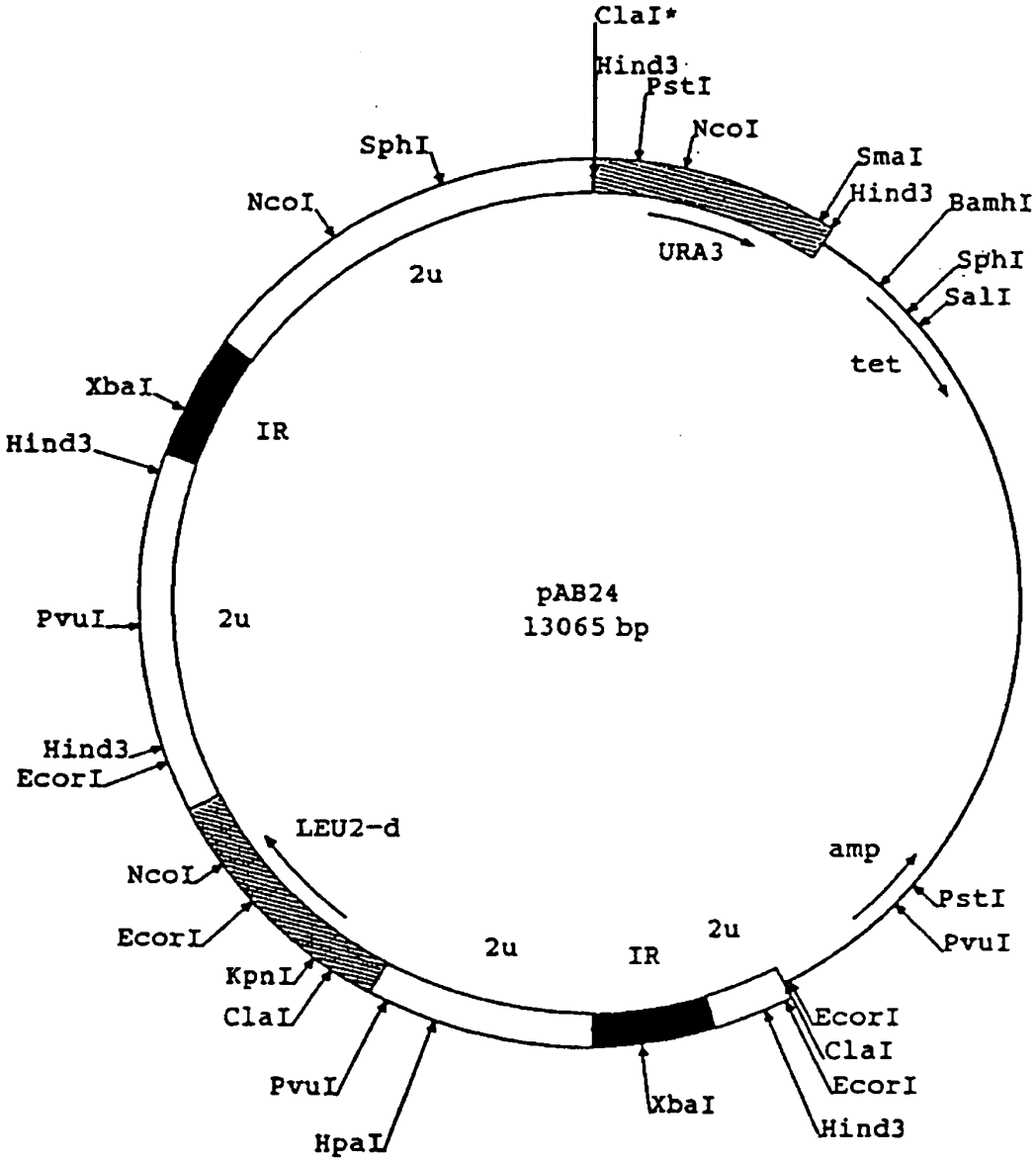

ClaI*
Hind3
PstI
NcoI
SmaI
Hind3
BamhI
SphI
SalI
tet
URA3
2u
SphI
NcoI
XbaI
IR
Hind3
pAB24
13065 bp
PvuI
2u
Hind3
EcorI
LEU2-d
NcoI
amp
PstI
PvuI
EcorI
2u
IR
2u
KpnI
ClaI
EcorI
ClaI
EcorI
PvuI
XbaI
HpaI
Hind3

FIG. 4

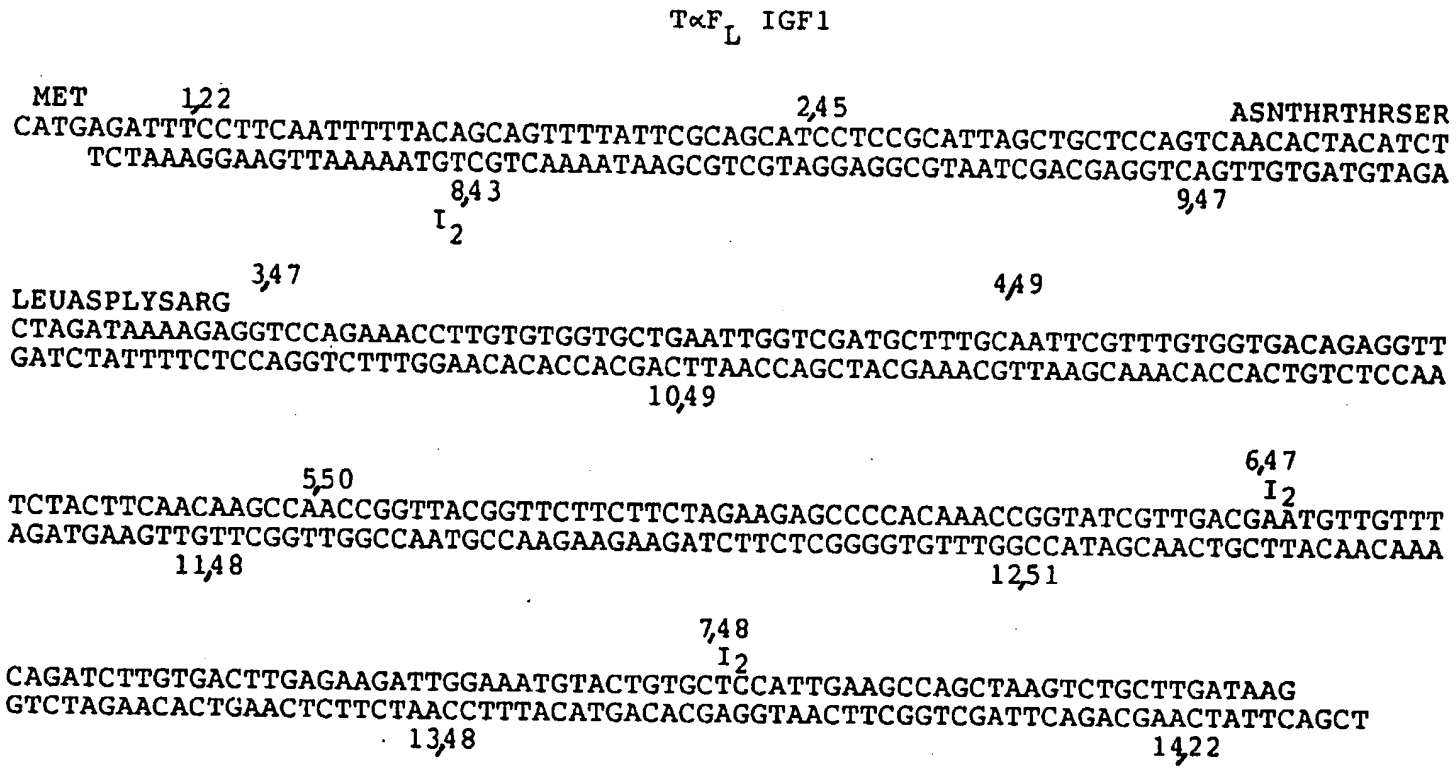


FIG. 5

DK 175910 B1

pYLUIGF1-55

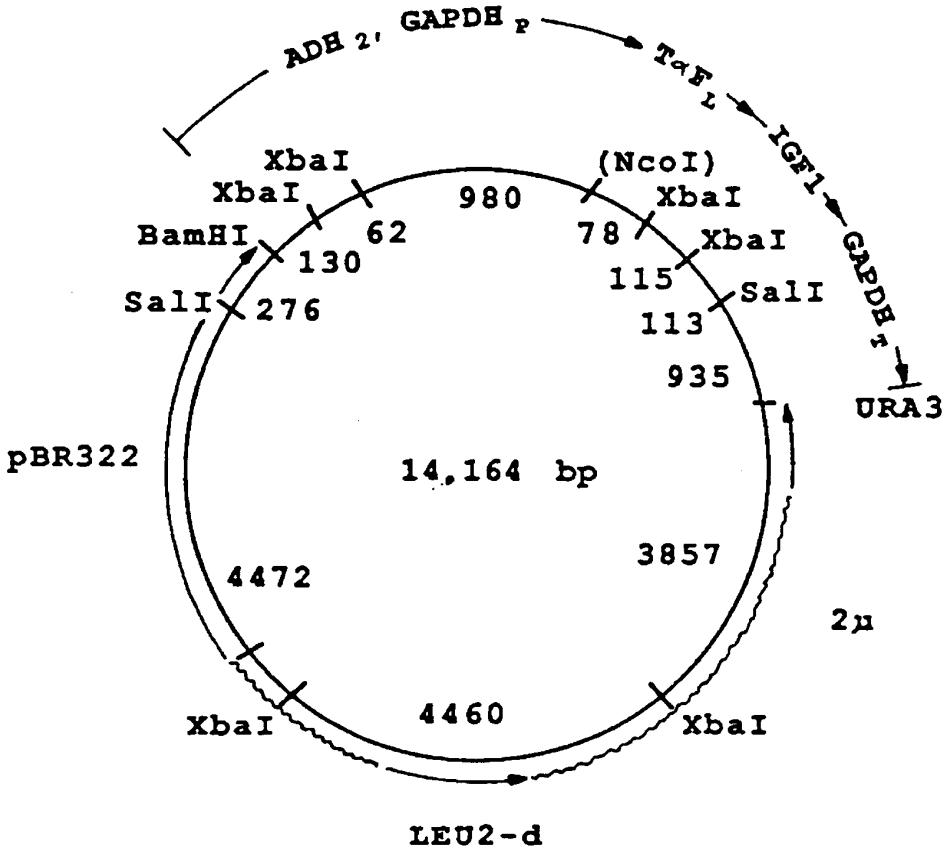


FIG. 6

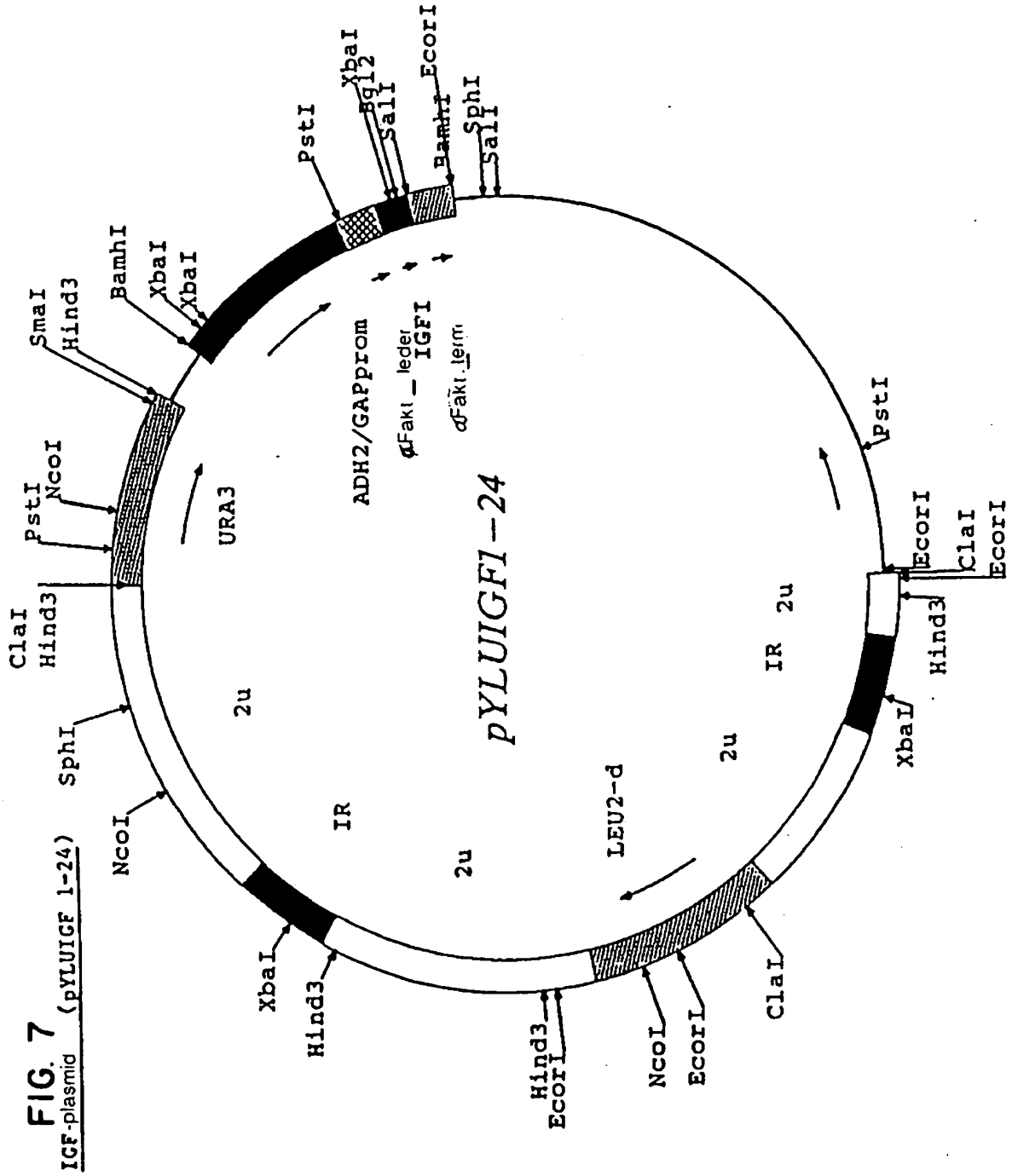


FIG. 7
IGF-plasmid (pYLUIGF 1-24)