

(12) Ausschließungspatent

Erteilt gemäß § 17 Absatz 1 Patentgesetz

(19) DD (11) 267 997 A5

4(51) C 12 N 15/00

AMT FÜR ERFINDUNGS- UND PATENTWESEN

In der vom Anmelder eingereichten Fassung veröffentlicht

(21)	AP C 12 N / 313 474 6	(22)	17.10.86	(44)	17.05.89
(31)	789,206	(32)	18.10.85	(33)	US
	841,121		18.03.86		

(71) siehe (73)
 (72) Davidow, Lance St.; Dezeeuw, John R.; Franke, Arthur E., US
 (73) Pfizer Inc., New York, N.Y. 10017, US
 (74) Internationales Patentbüro Berlin, Wallstraße 23/24, Berlin, 1024, DD

(54) Verfahren zur Auffindung von transformierten *Yarrowia Lipolytica*-Spezies

(55) transformierter Microorganismus, *Yarrowia Lipolytica*, transformierte Vektor-DNA, gentechnologisches Herstellungsverfahren, alkalische Protease-Aktivität

(57) Die Erfindung betrifft ein Verfahren zur Auffindung von transformierten *Yarrowia Lipolytica*-Spezies, die am XPR2-Gen eingebaute Vektor-DNA enthalten. Die Aufgabe der Erfindung besteht darin, ein Verfahren zur Auffindung von transformierten *Yarrowia Lipolytica*-Spezies, die am XPR2-Gen transformierte Vektor-DNA enthalten, bereitzustellen. Erfindungsgemäß wird ein Verfahren zur Auffindung von transformierten *Yarrowia Lipolytica*-Spezies, die am XPR2-Gen eingebaute Vektor-DNA enthalten, bereitgestellt, welches

- (i) Transformieren eines XPR⁺-Stammes von *Y. lipolytica* mit einem XPR-Expressionsgebilde und
- (ii) Überprüfung der in (i) hergestellten Transformanten auf den Verlust der Aktivität der alkalischen Protease umfaßt.

ERFINDUNGSANSPRUCH

1. Verfahren zur Auffindung von transformierten Y. lipolytica-Spezies, die am XPR2-Gen eingebaute Vektor-DNA enthalten, dadurch gekennzeichnet, daß es

(i) Transformieren eines XPR⁺-Stammes von Y. lipolytica mit einem XPR-Expressionsgebilde und

(ii) Überprüfung der in (i) hergestellten Transformanten auf den Verlust der Aktivität der alkalischen Protease umfaßt.

2. Verfahren nach Punkt 1, dadurch gekennzeichnet, daß der XPR⁺-Stamm von Y. lipolytica mit ungeschnittener DNA des Plasmids pLS-3 oder mit durch SnaBI verdauter DNA von pLS-3 transformiert wird.

3. Verfahren nach Punkt 1, dadurch gekennzeichnet, daß die Y. lipolytica die Identifizierung-Charakteristika von Y. lipolytica ATCC 20688 besitzt.

Hierzu 17 Seiten Zeichnungen

267997

Verfahren zur Auffindung von transformierten Yarrowia Lipolytica-Spezies

Anwendungsgebiet der Erfindung

Die Erfindung bezieht sich auf ein Verfahren zur Auffindung von transformierten Yarrowia Lipolytica-Spezies, die am XPR2-Gen eingebaute Vektor-DNA enthalten.

Ganz allgemein bezieht sich die Erfindung auf die Technologie der Proteinsekretion aus Hefe und insbesondere auf rekombinante Yarrowia lipolytica-Klonierungsvehikel, die heterologe DNA enthalten, welche für die Expression und Sekretion von Säugerproteinen (z.B. Prochymosin) und anderen Polypeptiden codiert, und auf Expressionsvektoren, die einen Y. lipolytica-Genpromotor (z.B. XPR2 oder LEU2), eine Signalsequenz (oder Präsequenz) der alkalischen Protease, eine Proregion und eine XPR2-Terminatorregion enthalten, und Varianten oder funktionelle Äquivalente hiervon, die durch Degeneration des genetischen Codes oder durch die Verwendung von

anderen Genbestandteilen von Y. lipolytica entstehen. Speziell bezieht sie sich auf ein Verfahren zur Auffindung von transformierten Y. lipolytica-Spezies, die am XPR2-Gen eingebaute Vektor-DNA enthalten.

Charakteristik der bekannten technischen Lösungen

Die wirtschaftliche Attraktivität eines stetigen und ausreichenden Angebotes an einer Vielzahl von Proteinen oder Polypeptiden, die für einen Industriezweig (z.B. Prorennin, Rinderwachstumshormon) oder für medizinische Zwecke (z.B. Urogastron, Gewebefibrinogenaktivator, menschliches Anaphylatoxin C5a) von Wert sind und insbesondere an solchen, die aus einer Quelle stammen, welche hohe Qualitätsprodukte in leicht gewinnbarer, funktioneller Form liefert, veranlaßte viele Forscher dazu, die Technologie rekombinanter DNA auf Mikroorganismen als "Fabriken" für die Herstellung heterologer Proteine anzuwenden.

Ausgedehnte Studien sind auf die Proteinsekretion als mögliche Lösung für Schwierigkeiten gerichtet, die bei der Gewinnung exogener oder heterologer (fremder) Proteine in einer biologisch aktiven Form aus intrazellulären Anhäufungen in rekombinanten Wirtszellen, insbesondere aus Escherichia coli, auftreten. In E. coli wird das heterologe Protein häufig innerhalb der Zelle in Form von refraktilen Einschlusskörpern produziert. Solch ein Protein besitzt im allgemeinen nur geringe Wasserlöslichkeit und hat wenig oder keine biologische Aktivität. Extrahieren dieses Proteins aus den refraktilen Einschlusskörpern umfaßt im allgemeinen intensives chemisches Behandeln, das eventuell teuer ist und dazu führen kann, daß wenig oder gar kein Protein in der gewünschten, nativen, biologisch aktiven Form isoliert wird. Außerdem wird die Wahrscheinlichkeit der Kontamination des genannten Pro-

teins mit unerwünschten, von E. coli produzierten Substanzen dadurch vergrößert, daß es notwendig ist, die Zellen aufzubrechen, um die refraktilen Körper freizusetzen. Auch andere Organismen als E. coli produzieren heterologes Protein in unlöslicher, intrazellulärer Form. Beispielsweise offenbart das britische Patent 2,091,271, veröffentlicht am 28. Juli 1982, eine genetische Modifikation von S. cerevisiae mittels der Technologie rekombinanter DNA, um Rennin oder Chymosin des Kalbes - die Ausdrücke sind hier auswechselbar verwendet - zu exprimieren. Angesichts dieser Schwierigkeiten hat man sich bei dem Versuch, das Protein in einer nativen, aktiven Konfiguration herzustellen, der Sekretion des genannten Proteins aus dem Wirtsorganismus zugewandt.

Ob ein bestimmtes Protein, einschließlich heterologen Proteins, oder ein Polypeptid durch einen gegebenen Organismus sezerniert wird, scheint vom Protein abzuhängen. In den meisten eukariotischen Zellen ist ein Teil der Proteinsynthesemaschinerie mit der Membran des endoplasmatischen Retikulums assoziiert und die Sequenz von Aminosäuren (genannt die "Signalsequenz"), die sich nahe dem Aminoende der in Entstehung begriffenen Polypeptidkette befindet, dient dazu, das Protein zur Durchquerung der Membran zu veranlassen. Die Signalsequenz wird anschließend während des Sekretionsprozesses proteolytisch gespalten, was aktives, reifes Protein ergibt. Verschiedene Versuche sind gemacht worden, um Verfahren zur Sekretion von heterologen Proteinen unter Verwendung von Signalsequenzen in Mikroorganismen, einschließlich Bacillus subtilis, Saccharomyces cerevisiae, und in Säugierzellen in Kultur zu entwickeln. Die genannten Organismen haben sich jedoch nicht als ideal erwiesen.

Inhärente Eigenschaften von B. subtilis - er sezerniert beispielsweise viele Proteine, darunter zahlreiche Proteasen, die leicht das sekretierte heterologe Protein abbauen; Instabilität transformierter Stämme, die aus dem Verlust

heterologer DNA folgt - haben seine Nutzbarmachung verhindert.

Säugerzellen sind mit Erfolg gentechnologischen Manipulationen unterworfen worden, um heterologe Proteine zu exprimieren und zu sezernieren, aber diese Systeme sind technologisch anspruchsvoll und in der Handhabung teuer und bleiben für die kommerzielle Herstellung der meisten Proteine als Produkte unbrauchbar.

Obwohl Untersuchungen über Proteinsekretion mit S. cerevisiae erfolgreicher waren als solche mit B. subtilis, scheint selbst S. cerevisiae einigen inhärenten Beschränkungen in seiner Eigenschaft als System für die Proteinsekretion zu unterliegen. Die europäische Patentanmeldung 0123544, veröffentlicht am 31. Oktober 1984, beschreibt die Isolierung des alpha-Faktor-Genes von S. cerevisiae und die Verwendung von dessen Promotor- und/oder Signalpeptid-Teilen in Kombination mit DNA, die für Proteine codiert, die für Hefe heterolog sind, in einem Plasmid für die Transformation von Hefezellen, die in der Lage sind, in Zellkultur diskretes, reifes Protein zu produzieren. Die europäische Patentanmeldung 0088632, veröffentlicht am 14. September 1983, beschreibt ein Verfahren für die Expression und Sekretion von heterologem Protein in S. cerevisiae. Jedoch scheint die Größe des Proteins, das S. cerevisiae wirksam mit diesem und anderen Sekretionssystemen sezernieren kann, auf etwa 20,000 Dalton beschränkt zu sein. Die Überwindung dieser allgemeinen Ineffizienz von S. cerevisiae als Sekretionsorganismus machte vielfache Änderungen durch Mutationen nötig, wie von Smith et al., Science 229, 1219-1224 (1985) beschrieben. Eine Ausnahme dieses allgemeinen Trends ist die Beobachtung, daß Aspergillus-Enzyme, die größer als 20,000 sind, anscheinend durch S. cerevisiae sezerniert werden können, aber diese Enzyme werden von S. cerevisiae in hohem Maße glycosyliert und dies könnte die Wirksamkeit der Sekretion beeinflussen.

Besonderes Interesse ist auf Yarrowia lipolytica, eine industriell wichtige Hefespecies, die zur Herstellung von Citronensäure und Einzeller-Protein verwendet wird, gerichtet. Sie kann außerdem zur Herstellung von Erythritol, Mannit und Isopropyläpfelsäure verwendet werden. Im Gegensatz zu S. cerevisiae ist Y. lipolytica wegen ihrer Fähigkeit, Proteine mit höherem Molekulargewicht (alkalische Protease, saure Protease und RNase) wirkungsvoll in ihr Wachstumsmedium zu sezernieren, wodurch sie die Möglichkeit der Gewinnung heterologer Proteine in nativem Zustand ohne die Notwendigkeit der Zerstörung der produzierenden Zellen gestattet, von besonderem Interesse und Wert. Außerdem sezerniert Y. lipolytica quantitativ nur wenig Proteine, wodurch sie die Möglichkeit zur Herstellung eines gewünschten heterologen Proteins im Wachstumsmedium als vorherrschender Proteinspezies und die Vereinfachung der Gewinnung dieses heterologen Proteinproduktes bietet.

Y. lipolytica produziert große Mengen an extrazellulärer Protease. Dies ist das von Y. lipolytica vorwiegend sezernierte Protein. Die jeweilige Protease (alkalisch, sauer oder neutral) wird durch den verwendeten Stamm der Y. lipolytica bedingt (Ogrydziak et al., J. Gen. Microbiol. (1982), 128, 1225-1234). Über eine teilweise erfolgte Sequenzanalyse der N-terminalen Aminosäuresequenz alkalischer extrazellulärer Protease wurde von Ogrydziak et al., (loc. cit.) berichtet.

Die anhängige amerikanische Patentanmeldung mit dem Aktenzeichen 634,505, eingereicht am 25. Juli 1984, beschreibt Verfahren für die Transformierung von Y. lipolytica und für die Klonierung von Y. lipolytica-Genen durch Komplementierung von Mutationen. Sie offenbart die Klonierung des XPR2-Gens, das für eine sezernierte alkalische Protease codiert, durch Komplementierung einer xpr2-Mutation von Y. lipolytica. Die Methodologie umfaßt die Transformierung eines Wirtsstammes

von Y. lipolytica durch eine teilweise Bgl II-Verdauung einer Y. lipolytica-Genbank im Vektor pLD40, beschrieben in der europäischen Anmeldung 0138508, veröffentlicht am 24. April 1985, der korrespondierenden Anmeldung zu der obengenannten US-Anmeldung.

Ziel der Erfindung

Die vorliegende Erfindung stellt eine Methodologie für die Herstellung von Vektoren zur Verfügung, welche, wenn sie in Y. lipolytica eingeführt werden, den Wirten die Fähigkeit zur Herstellung spezifischer Proteine, die von heterologer DNA aus beliebiger Quelle, aber insbesondere von eukaryotischer und synthetischer DNA codiert werden, und zu ihrer Sekretion in das Medium vermitteln; rekombinante Y. lipolytica-Klonierungsvehikel, die heterologe DNA, welche für die Expression von Säugerprotein und anderen Polypeptiden codiert, enthalten, einschließlich von für die Transformation von Y. lipolytica-Wirten geeigneten Plasmiden, und besonders von integrierbaren Expressionsvektoren, die den LEU2-Genpromotor, den XPR2-Genpromotor, die Präpro-Region der alkalischen Protease und die XPR2-Terminator-Region enthalten; und Expressionsplasmide mit einer heterologen codierenden Sequenz mit XPR2-Sekretionssignalen strangabwärts vom LEU2-Promotor, die in der Lage sind, ein heterologes Protein in durch sie transformierter Y. lipolytica zu exprimieren und zu sezernieren.

Die Erfindung legt also die Expression und Sekretion von reifem heterologem Protein und insbesondere die von Prorennin und menschlichem Anaphylatoxin C5a von genetisch veränderten Y. lipolytica-Zellkulturen dar. Die Ermittlung der genauen Identität der Aminosäuresequenz wie auch der DNA-Sequenz für die extrazelluläre alkalische Protease von Y. lipolytica hat es möglich gemacht, festzustellen, daß heterologes Protein mittels der Technik rekombinanter DNA für die Produktion in Zellkultur exprimiert und sezerniert werden kann. Im Falle

von Prorennin wird die reife Form des Zymogens (Rennin-Vorläufers) exprimiert und sezerniert.

Darlegung des Wesens der Erfindung

Die Aufgabe der Erfindung besteht darin, ein Verfahren zur Auffindung von transformierten Y. lipolytica-Spezies, die am XPR2-Gen eingebaute Vektor-DNA enthalten, bereitzustellen.

Es ist nun gefunden worden, daß Y. lipolytica genetisch mittels der Technik der rekombinanten DNA modifiziert werden kann, um Transformanten zu produzieren, die heterologe Proteine in ihrer nativen Form exprimieren und sezernieren können. Dies wird durch die Konstruktion von Vektoren bewerkstelligt, die die Signal- oder die Signal- und die erste (Prol-) oder beide Prosequenzen (Prol- + Pro2-) des XPR2-Gens tragen, verknüpft mit der strukturellen Gensequenz des heterologen Proteins, welches sezerniert werden soll.

Transformierte Spezies, die durch Einbau von Vektor-DNA, welche ein Fragment des XPR2-Gens, dem regulatorische oder strukturelle Anteile an beiden Enden des Gens fehlen, am XPR2-Lokus hergestellt wurden, produzieren keine alkalische Protease mehr, eine Eigenschaft, die nicht nur für die Sekretion von heterologem Protein wünschenswert ist, sondern die auch für das Screening von vermuteten Transformanten benutzt werden kann.

Weiterhin sind Vektoren, die den XPR2-Promotor und Sequenzen für die sekretorische Signalsequenz der alkalischen Protease tragen, in der Lage, in einer transformierten Y. lipolytica-Zelle die Sekretion des reifen heterologen Proteins zu bewirken. Einige rekombinante DNA-Vektoren dieser Art können Expression/Sekretion in einem Hefegenom unabhängig vom Ort der Integration bewirken. Im allgemeinen liefern Vektoren, die genügend flankierende 5'- und 3'-DNA enthalten, unabhängig vom Ort der Integration die Expression des Produktes.

Außerdem ist überraschender- und unerwarteterweise entdeckt worden, daß die Integrierung eines von pBR322 abgeleiteten Plasmids in chromosomale DNA von Y. lipolytica eine Homologieregion ergibt, die in der Lage ist, weitere ortsgerichtete integrative Transformationen anzuregen. Die integrierte Kopie von pBR322 fungiert also als eine/*(docking platform) für dazukommende, transformierende DNA. Der Einbau einer bleibenden Kopie von pBR322 in chromosomale DNA von Y. lipolytica ergibt also trotz der Tatsache, daß pBR322 keine native DNA von Y. lipolytica ist, ein bekanntes Ziel für den Einbau. Transformierungsempfänger vom Stamm Y. lipolytica, die einen solchen Bereich enthalten, bieten zwei wesentliche Vorteile vor Empfängern, denen ein solcher Bereich fehlt, nämlich das Vorhandensein einer Region mit einer bekannten Sequenz und einer bekannten Restriktionskarte, die als Ziel für den ortsgerichteten Einbau dienen kann, und, falls pBR322 als Ziel für den Einbau verwendet wird, die Möglichkeit der Bestimmung, ob das eintretende Plasmid eine vollständige funktionelle Einheit oder ein solches Gen, im Gegensatz zu nur einem Teil des gesuchten Gens, enthält. Beispielsweise würde ein Plasmid, das nur ein 3'-Fragment des XPR2-Gens enthält, einen XPR2-1002-Empfänger transformieren, wenn es das Codon des Wildtyps enthielte und am XPR2-Locus eingebaut würde. Dasselbe Plasmid würde jedoch den XPR2-1002-Wirt nicht in den proteasepositiven Phänotyp transformieren, wenn es in pBR322 integriert würde, weil ihm die gesamte funktionelle Einheit fehlen würde.

In transformierten Stämmen von Y. lipolytica, die eine Homologieregion zu heterologer Vektor-DNA enthalten, fungiert diese Region, die exogene DNA enthält, als Empfängerort während der integrativen Transformation der Y. lipolytica. Außer pBR322 und dessen Derivaten können Cosmide, Bakteriophagen wie M13 und Lambda, synthetisch derivatisierte DNA und übliche Plasmide wie pU13 zur Herstellung von transformierten Y. lipolytica-Spezies mit einer/*(docking platform) verwendet

/* Eingangsregion

werden.

Mit "LEU2"-Promotorsequenz ist die strangaufwärts gelegene, nicht translatierte Region oberhalb des ATG, die die meisten, wenn nicht alle für die Expression benötigten Merkmale enthält, gemeint.

Mit "XPR2"-Promotorsequenz ist die strangaufwärts nicht translatierte Region vor der Signal- (oder Prä-) Sequenz gemeint, die für die Expression notwendig ist. Außerdem kann die Signalsequenz des XPR2-Gens, mit oder ohne die Prosequenz, verwendet werden, um Proteine unter der Expressionskontrolle von anderen Promotoren von Y. lipolytica als denen des XPR2-Gens zu sezernieren. So sind Vektoren, die den LEU2-Promotor und Sequenzen für das Sekretionssignal der alkalischen Protease tragen, in der Lage, in einer transformierten Y. lipolytica-Zelle die Sekretion von reifem, heterologem Protein zu bewirken.

Menschliches Anaphylatoxin C5a, auch unter dem Namen "menschliches Komplement-Protein C5a" (menschliches C5a) bekannt, ist ein bioaktives Polypeptid-Fragment, das in vivo als Folge der Komplementaktivierung gebildet wird. Es wirkt als Immunmodulator bei der Regulierung bestimmter Teilbereiche der humoralen und zellulären Immunantwort. Seine Primärstruktur und die anderer Anaphylatoxine ist aufgeklärt. Eine Zusammenfassung der chemischen, physikalischen und biologischen Eigenschaften ist von Hugli in "Complement", herausgegeben von H.J. Müller-Eberhard und P.A. Miescher, S. 73-99, 1985, Springer-Verlag, New York, vorgelegt worden.

Es wird den Fachleuten bewußt sein, daß heterologe DNA, die für faktisch jede beliebige bekannte Aminosäuresequenz codiert, mutatis mutandi in der vorliegenden Erfindung Verwendung finden kann. Die hier offenbarte Methodologie ist mutatis mutandi auf die Produktion und Sekretion eines jeden

beliebigen bekannten, heterologen Proteins anwendbar, wobei repräsentative Vertreter im US Patent 4,532,207, erteilt am 30. Juli 1985, aufgelistet sind. Ferner kann ein beliebiges anderes Gen von durch Y. lipolytica sezernierten Proteinen wie z.B. das Ribonukleasegen und das Gen der sauren Protease anstelle des XPR2-Gens verwendet werden, wie auch Hybrid-Gene, die durch Vereinigen von Fragmenten von zwei oder mehreren dieser Gene, z.B. der Signalsequenz des XPR2-Gens und der Promotorsequenz des Ribonuklease-Gens, konstruiert werden.

Ebenfalls eingeschlossen in den Rahmen dieser Erfindung sind die funktionellen Äquivalente der hier beschriebenen DNA- oder Nukleotidsequenzen. Die Degeneration des genetischen Codes erlaubt die Substitution bestimmter Codons durch andere Codons, die für dieselbe Aminosäure codieren und demzufolge zu demselben Protein führen. Die DNA oder Nukleotidsequenz kann stark schwanken, da mit der Ausnahme von Methionin und Tryptophan die bekannten Aminosäuren durch mehr als ein Codon codiert werden können. Deshalb könnten das vollständige XPR2-Gen oder Teile davon unter Bildung einer DNA-Sequenz synthetisiert werden, die sich von der in Figur 3 gezeigten deutlich unterscheiden würde. Die dadurch codierte Aminosäuresequenz würde jedoch erhalten bleiben. Solche funktionellen Änderungen einer gegebenen DNA- oder Nukleotidsequenz liefern die Möglichkeit, Sekretion und/oder Prozessieren von heterologen Proteinen zu fördern, die durch daran geknüpfte, fremde DNA-Sequenzen codiert werden. Alle Abwandlungen der Nukleotidsequenz des XPR2-Gens und Fragmente hiervon, die der genetische Code gestattet, werden deshalb von dieser Erfindung umfaßt. Weiterhin ist es möglich, Codons zu deletieren oder ein oder mehrere Codons durch andere als die degenerierten Codons zu ersetzen, um ein strukturell modifiziertes Polypeptid herzustellen, das jedoch im wesentlichen dieselbe Anwendbarkeit oder Aktivität wie das von dem unmodifizierten DNA-Mo-

lekül produzierte Polypeptid besitzt. Diese zwei Polypeptide sind funktionell äquivalent, wie auch die zwei DNA-Moleküle, die ihre Produktion bewirken, auch wenn sich die Unterschiede zwischen den erwähnten DNA-Molekülen nicht von der Degeneriertheit des genetischen Codes ableiten. Das einfachste Beispiel hierfür findet man beim Prorennin A und Prorennin B, den zwei allelen Formen des Prorennins, die sich nur durch die Anwesenheit eines Aspartatrestes in Position 286 im Prorennin A und eines Glycinrestes in dieser Position im Prorennin B unterscheiden.

- Unter Anwendung dieser Methodologie ist die Expression und Sekretion der heterologen Säugerproteine Prorennin und menschliches Anaphylatoxin C5a in Y. lipolytica gelungen, wobei Expressions- und Sekretionssignale von Y. lipolytica-XPR2- und/oder LEU2-Genen verwendet wurden. Die DNA-Sequenzen für Prorennin und menschliches Anaphylatoxin wurden über synthetische Oligonukleotide mit der XPR2-Gensequenz an Orten verknüpft, von denen man annahm, daß sie für die Prozessierungsstellen des Signalpeptids der alkalischen Protease oder des Proteasevorläufer, die hier als Prol und Pro2 bezeichnet werden, codieren, und für die Herstellung von Genkonstruktionen verwendet, die anschließend durch integrative Transformation in Y. lipolytica eingebracht wurden. Die rekombinanten Kulturen exprimierten heterologe Proteine mit dem Molekulargewicht und den Immunreaktivitäten von Prorennin und menschlichem Anaphylatoxin C5a und sekretierten sie in das Wachstumsmedium. Das so hergestellte Prorennin wird wahrscheinlich zu einer nativen Konfiguration gefaltet, da es nach der Entfernung des Propeptides vollständige enzymatische Aktivität entwickelt.

Der Ausdruck "rekombinantes DNA-Material" umfaßt so, wie er hier verwendet wird, beliebig Material, das mindestens einen der folgenden Bestandteile enthält: das XPR2-Gen von Y. lipolytica, dessen Signal- (oder Prä-), dessen Prol- und

Pro2- (welche zusammen die Pro-Region bilden) , dessen Promotor- oder dessen Terminatorsequenz; den LEU2-Promotor; und funktionelle Äquivalente der vorgenannten Sequenzen, die wegen der Degeneration des genetischen Codes in Frage kommen.

Repräsentativ für das erwähnte rekombinante DNA-Material sind DNA-Fragmente, Plasmide oder Vektoren oder transformierte Spezies, die beliebige oder alle der vorgenannten Sequenzen enthalten.

Materialien. Restriktionsendonukleasen und T4-Ligase wurden von New England Biolabs bezogen, bakterielle alkalische Phosphatase von Bethesda Research Laboratories, T4-Polynukleotidkinase von PL-Biochemicals und $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ von New England Nuclear. Alle Enzyme wurden unter den von den Herstellern empfohlenen Bedingungen eingesetzt.

Medien.

GPP Medium - (Glycerin/Proteose-Pepton-Medium) enthielt (pro Liter): 6,7 g Glycerin, 1,6 g Difco Proteose-Pepton, 1,7 g Difco Hefe-Stickstoffbase ohne Aminosäuren und Ammoniumsulfat, 30 mg Uracil und 0,5 ml/l Polypropylenglycol mit MG 2000 (Polysciences) in 40 mM Phosphatpuffer (pH 6,8). (Das Polypropylenglycol wurde weggelassen, wenn das Medium für Kulturen benutzt wurde, die für die Verwendung von Rennin-Enzymbestimmungen gezüchtet wurden.) Proteose-Pepton wurde getrennt im Phosphatpuffer autoklaviert.

YEPD-Medium - (Hefeextrakt/Pepton/Dextrose-Medium) enthielt (pro Liter): 5 g Hefeextrakt, 10 g Pepton und 20 g Dextrose.

E. coli wurde in LB-Medium bei 37°C gezüchtet. LB-Medium enthielt (pro Liter): 10 g Bacto-Trypton, 10 g Bacto-Hefeextrakt, 10 g Natriumchlorid, mit Natriumhydroxid auf pH 7,5 eingestellt.

DNA-Sequenzanalyse. Die DNA-Fragmente der hier beschriebenen verschiedenen Plasmide wurden auf Polyacrylamidgelen isoliert und nach dem Verfahren von Maxam et al., Methods in Enzymology, 65, 499 (1980), sequenziert.

Verknüpfungsverfahren. DNA-Fragmente, gespaltene Vektorplasmide eingeschlossen, wurden durch Vermischen der gewünschten Bestandteile (DNA-Fragmente mit geeignet konstruierten Enden, um korrekte Verknüpfung zu liefern) mit T4-DNA-Ligase ligiert. Annähernd 10 Ligase-Einheiten wurden für μg -Mengen an Vektor und Insertionsfragmenten zugegeben. Die dabei entstandene Mischung der Ligationsreaktion wurde in kompetente Zellen von E. coli K 12 des Stammes MM294 (ATCC-33625) oder HB101 (ATCC-33694) transformiert.

Herstellung chemisch synthetisierter DNA. Um die Hybrid-Gene für die Expression und Sekretion von Prorennin zu konstruieren, wurden acht Oligonukleotide mit Hilfe eines modifizierten Phosphoramiditverfahrens (Sinha et al., Tetrahedron Letters 24, 5843 (1983) auf einem automatisierten Genetic Design 6500 DNA-Syntheseapparat (Watertown, MA) synthetisiert und über 5M Harnstoff-20% Polyacrylamid-Gele gereinigt. Aliquots komplementärer Oligonukleotide wurden vermischt und bei 4°C über Nacht in TE (10 mM Tris-HCl, pH 8,0, 1 mM NaEDTA) miteinander gepaart. Aliquots (etwa 2 μg) der doppelsträngigen Oligonukleotide wurden in einem 20 μg -Reaktionsansatz, der 70 mM Tris (pH 7,6), 10 mM MgCl_2 , 5 mM Dithiothreitol, und 5 mM ATP enthielt, bei 37°C unter Verwendung von T4-Polynukleotidkinase phosphoryliert.

Herstellung von Plasmid-DNA. Die Herstellung von Plasmid-DNA in großem Maßstab erfolgte nach dem Verfahren von Holmes et al., Anal. Biochem., 114, 195-197 (1981) und durch anschließende Ethidiumbromid-CsCl Gleichgewichts-Dichtegradienten-Zentrifugation. Mengen von Plasmid-

DNA im Miniprep-Maßstab wurden nach dem Verfahren von Birnboim et al., NAR 1, 1513 (1979), (alkalisches SDS) hergestellt.

Konstruktion der Expressions-/Sekretionsvektoren für Prorennin. Um die endgültigen Expressionsvektoren zu erhalten, wurde eine Reihe verschiedener Konstruktionen gemacht. Alle Schritte sind in den beigefügten Zeichnungen im Diagramm dargestellt. Im allgemeinen wurden DNA-Fragmente mittels Gelelektrophorese isoliert und mit anderen Fragmenten oder gespaltener Plasmid-DNA in 20 μ l 50 mM Tris-HCl (pH 7,5), 10 mM MgCl₂, 20 mM Dithiothreitol, 1mM ATP und 200 Einheiten T4-Ligase bei 4°C verknüpft. Wenn Teilverdauung der DNA mit Restriktionsendonuklease nötig war, wurden die optimalen Spaltungszeiten experimentell bestimmt.

Identifizierung von Prorennin in Kulturflüssigkeit. Man ließ transformierte Hefen, die die Expressionsvektoren enthielten, über Nacht in GPP-Medium (siehe oben) wachsen. Nach Zentrifugation zur Entfernung der Hefezellen wurden jeweils 1 ml 50 %ige TCA zu jeder 5 ml-Probe der Kulturflüssigkeit gegeben und 60 min lang bei 4°C gehalten. Durch Zentrifugieren erhielt man einen Rückstand in Tablettenform und wusch diesen zweimal mit 2 ml kaltem Aceton. Das ausgefällte Protein wurde in 100 μ l SDS-Probenpuffer gelöst und gleiche Mengen wurden auf 10 %igen SDS-Polyacrylamidgelen der Elektrophorese unterworfen (Laemmli, U.K., (1970), Nature 227, 680). Die durch das Gel getrennten Proteine wurden elektrophoretisch auf Nitrocellulose übertragen (Schleicher und Schuell, 0,22 μ m) und Prorennin wurde durch Immunoblot-Analyse von Plattengelen (Hawkes, R. et al., (1982) Anal. Biochem. 119, 142) identifiziert. Das Filter wurde mit Anti-Prorennin-Antikörper vom Kaninchen überschichtet und dann mit gegen Kaninchen-IgG gerichtetem, mit Peroxidase konjugiertem Ziegen-Antikörper inkubiert (Cappel, Malvern, Pa). Die gebundenen Antikörper wurden durch Anfärben mit 4-Chlor-1-naphthol und

Wasserstoffperoxid sichtbar gemacht.

Milchgerinnungsaktivität von Prorennin in Kulturflüssigkeit.
Die Kulturflüssigkeit von verschiedenen transformierten Y. lipolytica-Spezies wurde nach einem modifizierten Verfahren von Ernstrom, J. Dairy Sci. 41, 1664 (1958) auf ihre Milchgerinnungsaktivität untersucht. Die Bestimmung umfaßt, kurz umrissen, die Messung der Zeitspanne, die Rennin in aktivierten Kulturüberständen zur Gerinnung von gepufferter Magermilch benötigt, und die Korrelation dieser Werte mit einem gereinigten Rennin-Standard. Man ließ Hefekulturen (25 ml) über Nacht in GPP-Medium wachsen. Nach Zentrifugieren, um die Zellen zu entfernen, wurden jeweils 5 ml des Kulturüberstands im Vakuum gefriergetrocknet. Jeder der lyophilisierten Überstände wurde in 300 µl destilliertem Wasser resuspendiert. Außerdem wurde eine Verdünnungsreihe von gereinigtem Kalbs-Prorennin als Standard zum Kontrollvergleich hergestellt. Das Prorennin in den Medienkonzentraten und in den Kontrollen wurde durch Zugabe von etwa 5-10 µl konz. HCl bis zu einem pH von annähernd 2 und durch einstündiges Inkubieren bei 22°C aktiviert. Magermilch wurde durch Zusatz von 60 g trockenem Magermilchpulver (Difco) zu 500 ml 41.5 mM Natriumacetat (pH 6,3) und 13,6 mM CaCl₂ und 20minütiges Rühren bei 4°C hergestellt. Das Substrat wurde direkt nach seiner Herstellung für die Bestimmungen verwendet. Eine Probe von 60 µl (zu 1 ml des Kulturüberstandes äquivalent) jeder Enzympräparation wurde bei 37°C zu jeweils 1 ml Magermilch hinzugefügt und die Gerinnungszeit wurde aufgezeichnet.

Herstellung von synthetischen Oligonukleotiden für das C5a-Gen. Die für die Synthese des C5a-Strukturgens verwendeten Oligedesoxynukleotide wurden mit Hilfe eines modifizierten Phosphoramidit-Verfahrens (Sinha et al., loc cit.) unter Verwendung einer Glasfritte mit definierter Porengröße auf einer automatisierten Genetic Design 6500 (Watertown, MA) DNA.-Syntheseapparatur auf chemischem Wege hergestellt. Die

Vorschrift gibt die Verwendung von 3 %iger (w/v) Dichloressigsäure in Dichlormethan für die Detritylierung, "in line"-Aktivierung der Phosphoramidite mit gesättigtem Tetrazol in Acetonitril, Schützen durch "Capping-Reaktion" mit Di-Ethoxyphosphintetrazolid und Oxidation mit wäßrigem Iod/THF (Matteucci et al., (1981), J. Amer. Chem. Soc. 105, 3183) an. Die Gesamtzeit für einen Additionszyklus betrug 14 min. Die zehn 47-meren, die Segmente A-J in Fig. 9, erhielt man mit 98,8 %iger durchschnittlicher Ausbeute/Stufe (mittels Trityl-analyse), sie wurden nach dem Verfahren von Matteucci et al., loc. cit., deblockiert, mit Ethanol aus 0,3M Natriumacetat gefällt und mittels präparativer Gelelektrophorese auf 10 %igen, denaturierenden Polyacrylamid-Harnstoff-Gelen isoliert, bevor sie gepaart wurden.

Zusammensetzen, Klonieren und Sequenzieren von menschlichem C5a-Gen. Fig. 9 zeigt die Aminosäuresequenz des gesuchten Proteins und die Anordnung der synthetischen Oligonukleotide, die zur Herstellung eines für menschliches C5a-Protein codierenden Gens benötigt werden. Alle Oligomeren, mit Ausnahme von A und F, wurden an ihren 5'-Enden mit T4-Polynukleotidkinase phosphoryliert. Das Zusammensetzen des Gens schloß zwei primäre Paarungs- /Verknüpfungs-Reaktionen ein, betreffend die Oligomere A, B, I und J und die Oligomere C, D, E, F, G und H. Die gebildeten 94 Bp und 141 Bp langen doppelsträngigen DNA-Fragmente wurden nach Elektrophorese auf einem 10 %igen Polyacrylamid-Gel isoliert und miteinander ligiert und das aus ihnen entstandene 235 Bp lange Produkt wurde mittels Gelelektrophorese isoliert. Das 235 Bp lange DNA-Fragment, das ein für C5a codierendes Strukturgen enthielt, wurde zwischen die EcoRI- und HindIII-Stellen der pBR 322-Vektor-DNA inseriert und in kompetente Zellen des Stammes HB101 von E. coli K-12 transformiert. Restriktionsanalyse von Plasmid-DNA, die aus 6 transformierten Spezies isoliert wurde, zeigte, daß 5 der 6 Klone ein EcoRI/HindIII-Fragment in der richtigen Größe enthielten. Die Nukleotid-

sequenz der C5a-Genregion eines jeden dieser Plasmide wurde nach dem Verfahren von Maxam et al., Methods Enzymol. 65, 499 (1980) bestimmt.

Konstruktion und Charakterisierung des C5a-Expressionsplasmids für E. coli. Verfahren für die Isolierung von DNA-Fragmenten und die Bedingungen für die Ligationsreaktionen waren identisch mit denen von Maniatis et al., (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, publizierte. Der trp-Promotor-Operator von E. coli wurde ursprünglich aus ptrpL1 gewonnen (Edman et al., (1981) Nature 291, 503). Das 360 Bp lange EcoRI-Fragment, das die im C5a-Expressionsplasmid (pC5a-48) verwendete trp Promotor-Operator-Sequenz enthielt, wurde aus dem Prorennin-Expressionsplasmid pPFZ-R2, beschrieben in der europäischen Patentanmeldung Nr. 0147178, veröffentlicht am 3. Juli 1985, isoliert.

Identifizierung von C5a in der Kulturflüssigkeit von Y. lipolytica. Das Verfahren war dasselbe wie dasjenige, das oben für Prorennin beschrieben wurde, mit der Ausnahme, daß Ziegen-Anti-C5a- und gegen Ziegen-IgG gerichtete Kaninchen-Antikörper (Cappel) beim Immunoblot verwendet wurden. Der Ziegen-Antikörper gegen menschliches C5a wurde nach dem Verfahren von Manderino et al., J. Immunol. Methods, 53, 41-50 (1982) hergestellt.

Die Vektoren

pLD40 - beschrieben in der europäischen Patentanmeldung
0133508, veröffentlicht am 24. April 1985

Die Mikroorganismen:

- ATCC 20774 Yarrowia lipolytica PC 30869
- ATCC 20781 mit XPR2 transformierte Yarrowia lipolytica
DL112- PC-30869
- ATCC 20776 Yarrowia lipolytica DL-148.
Transformierte Spezies von Y. lipolytica
ATCC 20688 mit durch SnaBI verdaulichem pLS-3
- ATCC 20775 Yarrowia lipolytica DL-144
Transformierte Spezies von Y. lipolytica
ATCC 20688 mit ungeschnittenem pLS-3
- ATCC 20777 Transformierte Spezies von Y. lipolytica
PC-30869 mit durch SnaBI geschnittenem
pC5aX-3
- ATCC 20778 Transformierte Spezies von Y. lipolytica
PC-30869 mit durch SnaBI geschnittenem pXX-11
- ATCC 20779 Transformierte Spezies von Y. lipolytica
PC-30869 mit durch SnaBI geschnittenem pXX-22
- ATCC 20780 Transformierte Spezies von Y. lipolytica
PC-30869 mit durch SnaBI geschnittenem pXX-33
- ATCC 20794 Transformierte Spezies von Y. lipolytica
PC-30869 mit pLD56

ATCC 20795 Transformierte Spezies von Y. lipolytica
ATCC 20794 mit durch NruI geschnittenem
pLX-34

Sie sind dem Budapester Abkommens entsprechend in der American Type Culture Collection, Rockville, Maryland, USA, hinterlegt worden, einer anerkannten Hinterlegungsstelle, die das Verbleiben des hinterlegten Gutes und dessen leichte Verfügbarkeit für die Öffentlichkeit für den Fall, daß ein Patent auf diese Anmeldung gewährt wird, garantiert. Die hinterlegten Stämme sind, solange die amerikanische Prioritätsanmeldung dieser Anmeldung anhängig ist, einer vom Präsidenten des amerikanischen Patent- und Warenzeichenamtes zu bestimmenden Person, die hierfür die Berechtigung erhält, unter den Bezeichnungen 37 CFR 1.14 und 35 USC 122 und in Übereinstimmung mit ausländischen Patentgesetzen in Ländern, in denen korrespondierende Anmeldungen dieser Anmeldung oder von deren Vorläufer eingereicht sind, zugänglich. Alle Beschränkungen der Verfügbarkeit der hinterlegten Mikroorganismen für die Öffentlichkeit werden bei Gewährung des Patentbesitzes unwiderruflich aufgehoben.

Die taxonomischen Studien von Y. lipolytica ATCC 20774 (identifiziert in der Kulturensammlung von Pfizer Inc. als PC 30869) wurden von Dr. J.R. De Zeeuw durchgeführt, der die nun folgende Beschreibung zur Verfügung stellte. Die angewandten Verfahren sind solche, die von J.L. Lodder in "The Yeasts", zweite Ausgabe, N. Holland Publishing Co., Amsterdam, 1970, empfohlen werden.

CBS 599, die Musterkultur für die Species Candida lipolytica ("The Yeasts", zweite Ausgabe, N. Holland Publishing Co., Amsterdam, 1970) und CBS 6124, die Musterkultur für Saccharomyces lipolytica in "The Yeasts, dritte Ausgabe, wurden zum Vergleich herangezogen. Früher wurde die Spezies auch mit Endomyces lipolytica bezeichnet. Ihr imperfektes

Stadium ist Candida lipolytica. Die taxonomische Stellung der Spezies wurde von van der Walt und von Arx, Antonie van Leeuwenhoek, 46, 517-521 (1980) festgelegt. Der bevorzugte Name ist nun Yarrowia lipolytica.

Die morphologischen und physiologischen Charakteristika des Stammes PC-30869 und die der Kulturhaltung stimmen mit der Standardbeschreibung für die Spezies überein, die als Saccharomycopsis lipolytica in "The Yeasts", dritte Ausgabe, herausgegeben von Kreger-van Rij, S. 406-408, Elsevier Science Publishers B.V., Amsterdam, 1984, aufgelistet ist.

Tabelle 1

Vergleich von Stämmen von Yarrowia lipolytica

<u>Pfizer-</u> <u>Zugriffs-</u> <u>Nummer</u>	<u>Quelle</u>	<u>Genotyp</u>
PC-30265	NRRL YB-423 (auch CBS 6124), Musterkultur in "The Yeasts", dritte Ausgabe	Wildtyp diploid
PC-30286	CBS 599, Musterkultur in "The Yeasts", zweite Ausgabe	MATA Wild- typ, haploid
PC-30869	Siehe unten	MATB bio-6- leu2-40 xpr2-1002

PC 30869 wurde durch genetische Rekombination geeigneter Mutanten von Y. lipolytica PC-22208, einer aus Boden isolierten Spezies von Pfizer, und Y. lipolytica PC-30026, einer Subkultur von NRRL Y-1094, konstruiert. PC-30869 unterscheidet sich im Phänotyp von ihren Wildtyp-Vorläufern dadurch, daß

sie (1) keine aktive exozelluläre alkalische Protease produziert, (2) Biotin für ihr Wachstum benötigt und (3) eine Quelle für L-Leucin benötigt.

Während des Wachstums von PC-30869 in der log-Phase in Hefeextrakt-Pepton-Glucose- (YEPD-) Kulturbrühe sind sprossende Zellen ovoid und besitzen eine durchschnittliche Größe von 2,6 x 5,5 Mikrometer. Auf YEPD-Agar sind Pseudo- und echtes Mycel hervorstechend. Blastosporen sind vorhanden, vorwiegend einzeln in pleuralen Positionen. Es sind keine Carotinoid-Pigmente sichtbar. Die Kultur verhält sich bei der Kreuzung mit authentischen Teststämmen für die Spezies als Haploid vom "B"-Typ ("mating type B") (Tabelle 5). Typische Ascosporeulierung wird auf V8-Agar beobachtet. Das Kohlenstoff-assimilierungsmuster ist in der beigefügten Tabelle 2 dargestellt. Fermentation findet nicht statt. Ammoniumionen und Harnstoff, aber nicht Nitrat, werden als einzige Stickstoffquellen verwertet (Tabelle 3). Der Stamm PC-30869 benötigt die Vitamine Thiamin und D-Biotin (Tabelle 4). Von den Wildtyp-Vorläufern der Kultur wird nur Thiamin benötigt. Kein Wachstum beobachtet man bei 37°C.

Tabelle 2

Kohlenstoffassimilation (a)

<u>Quelle</u>	Aüflistung der Ref.-Lit. ^(b) <u>Beschreibung</u>	Kultur		
		<u>30265</u>	<u>30286</u>	<u>30869</u>
1. L-Arabinose	-	-	-	-
2. Cellobiose	-	-	-	-
3. Erythritol	+	+++	+++	+++
4. D-Galactose	-	-	-	-
5. D-Glucose	+	+++	+++	+++
6. Inositol	-	-	-	-
7. Lactose	-	-	-	-
8. Maltose	-	-	-	-
9. D-Mannitol	+	+++	+++	+++
10. Raffinose	-	-	-	-
11. Ribitol	-	-	-	-
12. D-Ribose	-(+)	-	-	++
13. L-Rhamnose	-	-	-	-
14. Lösliche Stärke	-	-	-	-
15. Saccharose	-	-	-	-
16. Trehalose	-	-	-	-
17. D-Xylose	-	-	-	-
18. Bernsteinsäure	+	+++	+++	+++
19. Zitronensäure	+	+++	+++	+++

(a) Das Basalmedium war Bacto-Hefen-Stickstoffbase, ergänzt durch weitere 10 µg/l D-Biotin und 149 mg/l L-Leucinethylester x HCl, um 100 mg/l L-Leucin zur Verfügung zu stellen.

(b) Kreger-van Rij. (loc. cit.)

Tabelle 3

Stickstoffassimilation (a)

<u>Quelle</u>	Auflistung der Ref.-Lit.- (b) <u>Beschreibung</u>	Kultur		
		<u>30265</u>	<u>30286</u>	<u>30869</u>
1. $(\text{NH}_4)_2\text{SO}_4$	+	+++	+++	+++
2. KNO_3	-	-	-	-
3. Harnstoff	+	+++	+++	+++

(a) Das Basalmedium war Bacto-Hefen-Kohlenstoffbase, ergänzt durch 116 mg/l Natrium-Ketoisocaproat zur Bereitstellung eines Äquivalences von 100 mg/l L-Leucin und mit zusätzlichen 10 µg/l D-Biotin

(b) Kreger-van Rij. (loc. cit.)

Tabelle 4

Vitaminbedarf (a)

<u>Ergänzung des Basalmediums</u> (c)	Ergebnis bei d. zit. Lit. (b)	Kultur		
		<u>30628</u>	<u>30286</u>	<u>30869</u>
1. Keine	-	tr	tr	-
2. Thiamin x HCl	+	+++	+++	-
3. D-Biotin	-	tr	tr	tr
4. Thiamin plus Biotin	+	+++	+++	+++

- (a) Das Basalmedium war vitaminfreie Bacto-Hefebase plus 149 mg/l L-Leucinethylester x HCl, um 190 mg/l L-Leucin zur Verfügung zu stellen.
- (b) Kreger-van Rij (loc. cit.)
- (c) 200 µg/l Thiamin x HCl und/oder 10 µg/l D-Biotin wie angegeben

Tabelle 5

Ascosporulierung

<u>Test-Stamm</u>	gepaarte Kultur		
	<u>30265</u>	<u>30286</u>	<u>30869</u>
Keine (Kultur mit sich selbst gepaart)	++	-	-
30264 (ein A Paarungs-Typ)	++	-	+++
30267 (ein B Paarungs-Typ)	++	+++	-

(a) Die Kulturen 30264 und 30267 sind haploide Stämme vom entgegengesetztem Paarungs-Typ, von Dr. L.J. Wickerham freundlicherweise zur Verfügung gestellt. Sie sind in Science 167, 1141 (1970) formal beschrieben.

(b) 30264 /ist C. lipolytica YB-421

(c) 30267 /ist C. lipolytica YB-423-12

/* nach Wickerham

Tabelle 6

Andere Charakteristika

	In Ref. (a) an- gegebene Daten	30265	Kultur 30286	30869
Zellform	ovoid	ovoid	ovoid	ovoid
Durchschnittl. Zellgröße	(2-4,5) x (4-22)	3,3 x 9,1	3,0 x 8,2	2,6 x 5,5
Vegetative Vermehrung	Sprossung	Sprossung	Sprossung	Sprossung
Fermentation	fehlt	fehlt	fehlt	fehlt
Wachstum bei 37°C: Kolonie- Wachstum	keines Die drei Kulturen wuchsen auf ähnliche Weise und verhielten sich wie in der Literatur be- schrieben. Pseudo- und echtes Mycel vorherr- schend. Blastosporen vorhanden, meist einzeln in pleuralen Positionen. Kein Carotinoid-Pig- ment sichtbar.	keines	keines	keines

(a) Kreger-van Rij. (loc. cit.)

PC-30869 unterscheidet sich von anderen in der Patentliteratur beschriebenen Stämmen von Y. lipolytica, wie beim Vergleich ihrer Phänotypen (Tabellen 7 und 8) deutlich wird.

ATCC 20228 (Nubel et al., US Patent 4,155,811) zeigt wildtypartiges Verhalten bei der Nährstoffaufnahme wie die Mustertämme für die Species, CBS 599 und CBS 6124. Typischerweise benötigt er kein Uracil, Leucin oder Biotin zum Wachstum und er verflüssigt Gelatine.

ATCC 20628 (De Zeeuw et al., US Patent 4,407,953) benötigt im Gegensatz zu ATCC 20228 den Zusatz von Leucin für sein Wachstum. Wie ATCC 20228 benötigt er kein Uracil oder Biotin. Er verflüssigt außerdem Gelatine.

ATCC 20688 (Europäische Patentanmeldung 0138508) wächst nur, wenn das Medium sowohl durch Uracil als auch durch Leucin ergänzt wird. Dieser Bedarf an Uracil unterscheidet ATCC 20688 sowohl von ATCC 20228 als auch von ATCC 20628. ATCC 20688 benötigt kein Biotin und verflüssigt Gelatine.

Die Kultur PC-30869 unterscheidet sich von allen obengenannten Stämmen. Sie benötigt Biotin und Leucin, aber kein Uracil für ihr Wachstum. Sie verflüssigt Gelatine nicht.

Tabelle 7

<u>Kultur</u>	<u>Benötigte Nährstoffe</u>			
	<u>Aus dem angegebenen Medium weggelassene Nährstoffe</u>			
	<u>Keiner</u>	<u>Leucin</u>	<u>Uracil</u>	<u>Biotin</u>
CBS 599	+++	+++	+++	+++
CBS 6124	+++	+++	+++	+++
ATCC 20228	+++	+++	+++	+++
ATCC 20628	+++	-	+++	+++
ATCC 20688	+++	-	-	+++
PC-30869	+++	-	+++	-

Das vollständige Medium enthielt 16,7 g/l vitaminfreie Bacto-Hefebase plus 100 mg/l Uracil, 100 mg/l L-Leucin, 10 µg/l D-Biotin und 200 µl/l Thiamin x HCl.

Tabelle 8

<u>Gelatineverflüssigung</u>	
<u>Kultur</u>	<u>Verflüssigung</u>
CBS 599	+
CBS 6124	+
ATCC 20228	+
ATCC 20628	+
ATCC 20668	+
PC-30869	-

Das Medium enthielt 120 g/l Gelatine und 16,7 g/l vitaminfreie Bacto-Hefebase plus 100 mg/l Uracil, 100 mg/l L-Leucin, 10 µg/l D-Biotin und 200 µg/l Thiamin x HCl.

Kurze Beschreibung der Zeichnungen

- Fig. 1 - Partielle lineare Restriktionskarte der überlappenden Plasmide pLD 57, pLD 58 und pLD 62, isoliert aus Y. lipolytica, Stamm DL112.
- Fig. 2- Synthetische Oligonukleotid-Bruchstücke für das XPR2-Gen. Aus der veröffentlichten Sequenz für den größten Teil der ersten 25 Aminosäurereste der reifen Protease (Ogrydziak et al., loc. cit.) bieten zwei Regionen (markiert mit I und II) die Möglichkeit, 14-mere Oligonucleotid-Bruchstücke mit 32-facher oder geringerer Degeneration zu konstruieren. Die zwei Regionen beginnen bei den Aminosäuren 7 bzw. 18. Vier verschiedene, achtfach degenerierte Bruchstückgemische wurden für jede Region hergestellt und man ordnete ihnen Nummern zwischen 170 und 186 zu, wie dargestellt. In den gezeigten, abgeleiteten Nukleinsäuresequenzen bedeutet "X" alle 4 Basen, "U" steht für beide Purine und "Y" steht für beide Pyrimidine.
- Fig. 3 - Nukleotidsequenz des XPR2-Gens, wobei die Promo-

tor-, die Prä- (-157 bis -136), die Prol- (-135 bis -98), die Pro2- (-97 bis -1), die alkalische extrazelluläre Protease- und die Terminator-Sequenzen gezeigt sind.

- Fig. 4 - Konstruktionssequenz für den Terminator-Vektor pterm 4.
- Fig. 5 - Konstruktionssequenz und Restriktionskarte des Plasmids pLS-3
- Fig. 6 - Konstruktionssequenz und Restriktionskarte des Plasmids pXX-33
- Fig. 7 - Konstruktionssequenz und Restriktionskarte des Plasmids pXX-22
- Fig. 8 - Konstruktionssequenz und Restriktionskarte des Plasmids pXX-11
- Fig. 9 - Aminosäuresequenz von menschlichem Anaphylatoxin C5a
- Fig.10 - Restriktionskarte des Plasmids pC5a-48
- Fig.11 - Konstruktionssequenz und Restriktionskarte des Plasmids pC5aX-3
- Fig.12 - Nukleotidsequenz des LEU2-Gens
- Fig.13 - Konstruktionssequenz und Restriktionskarte des Plasmids pLX-34

Sequenzanalyse des XPR2-Gens. Die DNA-Sequenzanalyse des klonierten XPR2-Gens wurde mit Hilfe des chemischen Abbauverfahrens (Maxam et al. 1980, Methods Enzymol. 65, 499) an überlappenden Restriktionsfragmenten durchgeführt, die aus den Plasmiden pLD57, pLD58, pLD62 (Fig. 1) und pLD84 und pLD86 (siehe unten) hergestellt worden waren. Die Ergebnisse zeigten, daß in der Tat die klonierten Hefegenom-DNA's das Gen für die exozelluläre alkalische Protease enthielten. Die Nukleotidsequenz des XPR2-Gens und die Aminosäuresequenz des Vorläufers der alkalischen Protease mit ihrer Signalsequenz, wie sie sich von der Nukleotidsequenz ableitet, sind in Fig. 3 dargestellt. Ein Großteil der Aminosäuresequenz der exozellulären Protease war unbekannt (Ogrydziak et al., loc. cit.)

und wird hier zum erstenmal vorgestellt. Darüberhinaus werden die für die Expression und die Sekretion der exozellulären Protease benötigten Sequenzen hier zum erstenmal beschrieben. Die DNA-Sequenz, die für die alkalische Protease, ihren Vorläufer und die Signalsequenzen codiert, besteht aus 1362 Basenpaaren (Fig. 3). Aus der Nukleotidsequenz ergab sich, daß die Primärstruktur dieser Polypeptidkette aus 454 Aminosäureresten besteht. Die alkalische Protease wird in der Zelle in Form eines Vorläufers synthetisiert, der proteolytisch in die sezernierte oder reife Form überführt wird. Die Analyse der N-terminalen Aminosäuresequenz, abgeleitet von der Nukleotidsequenz, offenbarte die Existenz eines mutmaßlichen Signalpeptids im Vorläufermolekül. Dieses Signalpeptid enthält 22 Aminosäurereste und seine strukturellen Eigenschaften sind denen höherer eukaryotischer und prokaryotischer Signalpeptide ähnlich (Perlman et al., 1983, J. Mol. Biol. 167, 391). Einer Region in der abgeleiteten Aminosäuresequenz, die im allgemeinen mit den bekannten 25 N-terminalen Aminosäuren der reifen alkalischen Protease übereinstimmt (Ogrydziak et al., 1982, J. Gen. Microbiol. 128, 1225), gehen 157 Aminosäurereste voraus, die das Signalpeptid und zwei Spaltstellen vom Trypsin-Typ (Lys-Arg) enthalten. Diese Spaltstellen wurden benutzt, um die Proregion in Pro1 (-135 bis -98) und Pro2 (-97 bis -1) zu teilen. Siehe Fig. 3. Die reife alkalische Protease besitzt 297 Aminosäuren, wie aus der Nukleotidsequenz abgeleitet wurde. Die für die verschiedenen Formen der Protease aus der Nukleotidsequenz abgeleiteten Aminosäuresequenzen stimmen mit den Größen der gereinigten Formen der Enzyme überein. Zusätzlich zu der Struktursequenz des Vorläufers der alkalischen Protease wurden annähernd 700 Bp der 5' flankierenden Sequenz und 600 Bp der 3'-flankierenden Sequenz bestimmt. Analysen dieser Regionen zeigten, daß sie Sequenzen umfassen, die zu anderen eukaryotischen Promotoren und Terminatoren analoge Sequenzen enthalten und wahrscheinlich für die Expression der alkalischen Protease wesentlich sind.

Wie oben erwähnt, ist die Methodologie für die Transformation von Y. lipolytica und für die Klonierung von Y. lipolytica-Genen durch Komplementation von Mutationen, einschließlich des Klonierens des XPR2-Gens, das für eine sezernierte alkalische Protease codiert, durch Komplementation einer xpr2-Mutation in der europäischen Patentanmeldung 0138508 beschrieben. Das dort beschriebene Verfahren umfaßt das Transformieren eines Y. lipolytica-Wirtsstammes mit einer Y. lipolytica-Genbank, unter teilweiser Verdauung durch BglII im Vektor pLD40 hergestellt, wobei dieser Vektor durch die Tatsache charakterisiert ist, daß er ein kleines Segment, das die LEU2-Region von Y. lipolytica enthält, und 3EcoRI, 4EcoRV, 6AvaI, 1BglII, 1NcoI, 1ApaI, 2XhoI und 1BstXI Endonukleasen-Restriktionsstellen beherbergt. Eine der transformierten Y. lipolytica-Spezies wurde verwendet, um das Wildtyp-Gen (pLD84 und pLD86) aus Y. lipolytica NRRL Y-1094 für die Verwendung in der Expressions-/Sekretions-Vektorkonstruktion wie in Beispiel 1 beschrieben zu gewinnen.

Sequenzanalyse des LEU2-Gens. Die DNA-Sequenzanalyse des klonierten LEU2-Gens in pLD25 (EP 0138508) wurde mit Hilfe der chemischen Abbaumethode (Maxam et al. 1980, Methods Enzymol. 65, 499) an überlappenden Restriktionsfragmenten durchgeführt. Um die für die beta-Isopropylmalat- (IPM-) Dehydrogenase codierende Region und den passenden Leserahmen zu lokalisieren, bediente man sich der abgeleiteten Aminosäuresequenz, die für das LEU2-Gen von S. cerevisiae bereits bestimmt wurde (Andreadis et al., 1984, J. Biol. Chem. 259, 8059). Die Region der Genomsequenz von Y. lipolytica, die für eine Aminosäuresequenz codiert, welche einer Region der S. cerevisiae-Proteinsequenz homolog ist, wurde identifiziert. Die Nukleotidsequenz des 2,8 Kilobasenpaare tragenden LEU2-Gens und die Aminosäuresequenz von beta-IPM-Dehydrogenase, wie sie sich von der Nukleotidsequenz ableitet, sind in Fig. 12 dargestellt. Darüberhinaus sind die Sequenzen, die

für die Expression der beta-IPM-Dehydrogenase von Y. lipolytica benötigt werden, hier zum erstenmal beschrieben. Die DNA-Sequenz, die für dieses 405 Aminosäuren enthaltende Protein codiert, besteht aus 1215 Basenpaaren (Fig. 12). Zusätzlich zu der für die beta-IPM-Dehydrogenase codierenden Sequenz wurden eine annähernd 798 Bp lange 5'-flankierende Sequenz und eine 797 Bp lange 3'-flankierende Sequenz (einschließlich des Translations-Terminations-Codons TAA) aufgeklärt. Die Analyse dieser Regionen zeigte, daß sie zu anderen eukaryotischen Promotoren und Terminatoren analoge Sequenzen enthalten und für die Expression unabdingbar sind.

Die in 5'-Richtung strangaufwärts liegende Region des LEU2-Gens von Y. lipolytica enthält 78 Bp vor dem Translationsstart und 30 Bp vor dem mutmaßlichen mRNA-Start eine TATATATA-Sequenz. Eine zweite Sequenz, die für den Beginn der Transkription in Eukaryoten wichtig ist, ist die CAAT-Box, die im LEU2-Gen 74 Bp vor der mutmaßlichen Transkriptions-Initiationsstelle liegt, die -48 Bp vom ATG entfernt liegt (Fig. 12).

Die in 3'-Richtung strangabwärts liegende Region besitzt eine Sequenz im Bereich der Nukleotide 72 bis 120 nach dem Stop-Codon (TAA), die der 5'-TAG....TA(T)GT.....TTT-3'-Sequenz homolog ist, welche nach einem Vorschlag von Zaret et al., Cell 28, 563 (1982) wesentlich für die Transkriptionsbeendigung in S. cerevisiae ist.

BEISPIEL 1

Der verwendete Wirtsstamm war ATCC 20774 (MATB leu2-40 bio-6 xpr2-1002). Die mit XPR2 transformierte Y. lipolytica ATCC 20781 wurde als Kolonie entdeckt, welche nach Replika-Plattierung von Leucin-Mangel-Platten eine Zone auf Magermilch-Indikatorplatten bildete. Chromosomale DNA wurde aus den transformierten Spezies nach dem Verfahren der europäischen Patentanmeldung 0138508 hergestellt und verwendet, um das Gen für die sekretierte Protease zu gewinnen. Die chromosomale DNA wurde mit BglII-Enzym teilweise verdaut, ligiert, um das Fragment, das sowohl das E. coli-Replikon als auch das ampicillinresistente Gen aus dem Vektor enthielt, zu zirkularisieren, und zur Transformierung von E. coli verwendet. Die chromosomale DNA wurde auch mit SalI-Enzym verdaut und für ein Southern-Experiment verwendet, welches ergab, daß die normale LEU2-Region der transformierten Spezies nicht gestört war. (Ein 520 Bp langes Segment von SalI bis Eco RI der LEU2-Region, direkt 5' zum Segment des in pLD40 enthaltenen LEU2 gelegen, wurde als Sonde verwendet.) Die XPR2-Region mußte also, da Homologie für den Einbau einer Plasmidbank in Y. lipolytica notwendig ist, der Integrationsort sein. Drei überlappende, jedoch unterschiedliche Plasmide, pLD57, pLD58 und pLD62, wurden als erstes aus Y. lipolytica ATCC 20781 isoliert. Sie sind in Fig. 1 dargestellt. Hybridisierungen mit synthetischen Oligonukleotid-Bruchstücken für das XPR2-Gen, basierend auf der bekannten Sequenz der ersten 25 Aminosäurereste des reifen, sekretierten Proteaseproteins (Fig. 2 und 3) zeigten, daß das Gen für die sekretierte Protease kloniert worden war. Um festzustellen, ob das isolierte Gen die Wildtyp-Kopie oder die Mutanten-Kopie verkörperte, wurde der empfangende Y. lipolytica-Stamm mit pLD58 transformiert. Da keine Protease-positiven, transformierten Stämme aus irgend einem der Leucin-unabhängigen Transformanten entstanden, wurde geschlossen, daß das pLD58 die mutierten Allele des Gens enthielt.

Die Form des XPR2-Gens, die im Wildtyp-Stamm NRRL Y-1094 vorliegt, wurde durch ein Hybridisierungsexperiment mit einer Kolonie von E. coli erhalten. Als Sonde wurde das 2 kB lange Fragment von PvuI bis EcoRI verwendet, von dem man aufgrund von Sequenzierungsdaten annimmt, daß es das gesamte Strukturgen enthält. Von der ursprünglichen Bank der Fragmente aus der Teilverdauung der NRRL Y-1094 DNA in pLD40 mit Sau3A, beschrieben in der europäischen Patentanmeldung 0138508, erhielt man einige Kolonien, die mit dem Bruchstück hybridisierten. Zwei dieser Kolonien enthielten die sehr ähnlichen Plasmide mit den Bezeichnungen pLD84 und pLD86, die zur Entwicklung des Expressionsvektors verwendet wurden. Beide Plasmide enthalten dasselbe 5'-Ende der XPR2-Region-- die Sau3A-Stelle (die mit der BamHI-Stelle des Vektors verknüpft war und sie regenerierte), bei der die Sequenz in Fig. 3 beginnt. Alle zwei enthalten das gesamte Strukturgen für die Protease und den vermutlichen Transkriptionsterminator und enthalten insgesamt annähernd 4 bis 5 insertierte kB von der XPR2-Region des Stammes NRRL Y-1094. Die Insertionssequenz in pLD86 enthält einige hundert zusätzliche Basenpaare am 3'-Ende. Da wir die 3'-Verlängerung bis zu der BglII-Stelle (Basenpaar 2655) für die Konstruktion des Expressionsvektors benutzten, lieferten die zwei Plasmide dieselbe DNA, die funktionell in der Sequenz identisch zur Fig. 3 war.

Konstruktion des Expressions-/Sekretionsvektoren. Der Plan, der zum Erreichen von Expression und Sekretion von Prorennin in Y. lipolytica entwickelt wurde, benutzt die Konstruktion von verschiedenen Hybridgenen in einem integrativen Klonierungsvektor. Eine solcher Ansatz bringt mehrere verschiedene Plasmide hervor, die extensive Regionen gemeinsamer DNA-Sequenzen teilen. Tatsächlich wurde ein modulares Konstruktionschema verwendet, um Vektoren zusammenzustellen, die das Prorennin-Gen in 3'-Richtung zur vermuteten Prozessierungsstelle für das XPR2-Signalpeptid, der vermuteten Prol-Prozessierungsstelle und der Spaltstelle, von der man

weiß, daß sie die reife alkalische Protease erzeugt, insertiert enthalten. Im allgemeinen ist für das heterologe Gen zu wünschen, daß es für die Expression zwischen die Hefe-Promotor- und Terminatorsequenzen eingefügt wird. Man hat erkannt, daß der N-terminale Teil der Hybridgensequenzen in den verschiedenen Plasmidkonstruktionen variiert, wogegen die Prorennin-Strukturgensequenz, die XPR2-Terminatorsequenz und die Shuttle-Vektor-DNA in jeder Expressionsplasmid-Konstruktion die gleiche ist. Es war vorgesehen, daß dasselbe Strukturgen-Fragment für Prorennin und das Terminator-/Vektorplasmid in jeder Expressionsplasmid-Konstruktion verwendet werden sollte, wie unten beschrieben. Die verschiedenen Expressions-/Sekretionsplasmid-Konstruktionen für Prorennin unterscheiden sich in der Region unmittelbar strangabwärts von der XPR2-Genpromotorsequenz in der Länge der N-terminalen Sequenz für den Vorläufer der alkalischen Protease, die sich vor der Gensequenz für das Prorennin befindet. Deshalb wurde vorherbestimmt, daß die Promotor-Fragmentkomponente jedes Expressionsplasmides die variable Sequenz in der Region der XPR2-Prorennin-Verknüpfung sein sollte. Alle Expressions-/Sekretionsvektoren wurden durch eine ähnliche Ligierungsreaktion, drei Fragment-Komponenten enthaltend, zusammengesetzt.

Die experimentellen Stufen, die für die Konstruktion des Terminatorvektors pterm 4 verwendet wurden, sind in Fig. 4 dargestellt. Zuerst wurde ein synthetischer Linker mit einem Fragment ligiert, das das 3'-Ende des XPR2-Gens, einschließlich der Transkriptionsterminierung und des Polyadenylierungssignals, enthält. Kurz gesagt wurde das Plasmid pLD84 mit der Endonuklease KpnI gespalten und mit der synthetischen, doppelsträngigen Linker-DNA wie in Fig. 4 abgebildet ligiert. Das Verknüpfungsprodukt wurde mit den Endonukleasen HindIII und BglII gespalten und ein Fragment mit 760 Basenpaaren wurde in das Plasmid pLD41, das mit den beiden selben Endonukleasen linearisiert worden war, unter Bildung von pterm 4 einge-

setzt. Plasmid pterm 4 wurde anhand seiner Restriktionskarte identifiziert. Die Ergebnisse einer Reihe von Verdauungen durch Restriktionsendonukleasen wurden unter Verwendung von EcoRV, EcoRI, KpnI, BglII-HindIII und BglII-BclI analysiert. Die Verdauungen ergeben geeignete Fragmente, die die Gegenwart des synthetischen Linkers und des "kompletten" 3'-Endes des XPR2-Gens im Shuttle-Plasmid pLD41, beschrieben in der europäischen Patentanmeldung 0138508, bestätigen. Eine partielle Karte dieses 7,3 kB Terminatorvektors ist in Fig. 4 dargestellt.

• Konstruktion des Expressions-/Sekretions-Plasmids pLS-3.

Fig. 5 umreißt die Konstruktion des anfänglichen Plasmids, das für die Sekretion von Prorennin in Y. lipolytica verwendet wurde. Seine Restriktionskarte wird in Fig 5 vorgestellt. Die Konstruktion des Sekretionsplasmids für Prorennin wurde mit der Herstellung eines Fragmentes begonnen, das den größten Teil der Prorennin-Strukturgenesequenz trägt. Das 1080 Basenpaare lange BclI-BamHI (partielle) DNA-Fragment, das die codierende Sequenz für die Prorenninreste 6 bis 365 enthält, wurde aus dem E. coli Prorennin-Expressionsplasmid pPFZ-84A isoliert. (Plasmid pPFZ-84A ist ein Derivat des Prorennin-Expressionsplasmids pPFZ-R2, dessen Konstruktion in der europäischen Anmeldung Nr. 0147178, veröffentlicht am 3. Juli 1985, beschrieben ist und wurde durch synthetische, oligonukleotid-gesteuerte Mutagenese gebildet, wobei ein Restriktionsfragment ersetzt wurde. pPFZ-84A unterscheidet sich im wesentlichen von pPFZ-R2 nur durch zwei Basenpaare an den Prorennin-Aminosäureresten 214 (Asn -- Asp) und 286 (Asp -- Gly), wobei es für das sogenannte Prorennin A Allel codiert, beide Plasmide enthalten jedoch die gesuchte Sequenz für Prorennin und sind in diesem Beispiel funktionell äquivalent.) Das Fragment der XPR2-Promotorkomponente, das die codierenden Sequenzen für den alkalischen Proteasevorläufer 1 bis 157 und für Prorennin 1 bis 5 enthält, wurde wie folgt hergestellt. Das 870 Basenpaare lange HindIII-AvaI DNA-Frag-

ment, das die Promotorregion und das 5'-Ende des alkalischen Protease-Gens enthält, wurde aus dem XPR2-Subklon-Plasmid pLD90 isoliert. Dieses Fragment wurde mit einem synthetischen Fragment mit der Struktur:

```

5'CCGAGATTCCTGCTTCTTCTAATGCCAAGCGAGCTGAGATCACTAG      3'
3'   CTAAGGACGAAGAAGATTACGGTTCGCTCGACTCTAGTGATCCTAG 5'
      Leserichtung  —————>

```

ligiert.

Diese Sequenz enthält ein kohäsives AvaI-Ende, gefolgt von Sequenzen, die für die letzten neun Codons des Propeptids der alkalischen Protease codieren, und daran anschließend Sequenzen, die für die ersten vier Aminosäuren von Prorennin codieren, und endet in einer BamHI-Stelle. Das Fragment der Promotorkomponente wurde durch eine Standard-Verknüpfungsreaktion mit T4-Ligase unter Verwendung des synthetischen Fragments und des 870 Basenpaare langen HindIII-AvaI-Fragments und anschließender Spaltung mit HindIII und BamHI gebildet. Die gebildeten, ligierten Sequenzen wurden mittels Polyacrylamidgel-Elektrophorese gereinigt, um das passende 916 Basenpaare lange HindIII-BamHI-DNA-Fragment zu selektieren. Das 3'-Ende des Hybrid-Gens wurde aus dem Terminator-/Vektorplasmid pterm 4, der oben beschrieben ist, gewonnen. Das Plasmid pterm 4 wurde mit HindIII und BclI verdaut und das annähernd 7,3 kB lange HindIII-BclI-Terminator-/Vektor-DNA-Fragment, das den XPR2-Terminator, einen selektionsfähigen LEU2-Marker und pBR322 enthielt, wurde aus einem Agarosegel isoliert.

Das Plasmid pLS-3 für die Expression/Sekretion von Prorennin wurde zusammengefügt, indem die drei DNA-Fragment-Komponenten (mit HindIII-BclI gespaltenes Plasmid pterm4, zusammen mit dem 916 Bp langen HindIII-BamHI-Promotor- und dem 1080 Bp langen BamHI-BclI-Fragment, das das Prorennin-Gen enthält) inkubiert

wurden, die wie oben beschrieben in Gegenwart von T4-Ligase konstruiert worden waren (siehe Fig. 5). Die Ligationsmischung wurde verwendet, um den Stamm MM294 von E. coli K12 mit Hilfe des CaCl_2 -Verfahrens von Dągert et al., Gene 6, 23-28 (1979) zu transformieren. Aus den ampicillin-resistenten, selektierten transformierten Spezies wurden Plasmide isoliert und das Plasmid pLS-3 wurde anhand seiner Restriktionskarte (Fig. 6A) identifiziert. Die XPR2-Promotor-Region dieses Plasmids wurde sequenziert, um die korrekte Sequenz der synthetischen DNA und die korrekte Verknüpfung der gewünschten Fragmente zu bestätigen.

Herstellung von pLD90.--Dieses Plasmid enthält einen Subklon aus pLD84. Eine DNA-Region von der PvuI-Stelle in der Promotorregion von XPR2 bis zur EcoRI-Stelle in der Terminatorregion wurde in die HindIII-Stelle von pBR322 auf folgende Weise subkloniert. Einige Mikrogramm pLD84 wurden mit den zwei oben erwähnten Restriktionsenzymen verdaut. Dann wurden die "vorstehenden" Enden der verdauten DNA-Moleküle mit dem Klenow-Fragment von DNA-Polymerase I aufgefüllt. Anschließend wurden mit Kinase behandelte HindIII-Linker (CAAGCTTG von New England Biolabs) mit T4-Ligase an die Enden geheftet. Überschüssige Linker wurden entfernt und "vorstehende" HindIII-Enden wurden durch anschließende Verdauung mit HindIII-Enzym gebildet. Die Mischung der DNA-Moleküle wurde auf einem präparativen Agarosegel aufgetrennt und die gewünschte 2-kB-Bande wurde ausgeschnitten, gereinigt und zu einer Ligations-Reaktionsmischung mit dem durch HindIII verdauten, mit bakterieller alkalischer Phosphatase behandelten Vektor pBR322 gegeben. Die Ligationsmischung wurde zur Transformation kompetenter E. coli verwendet. Die Anordnung, bei der die EcoRI-Stelle des XPR2-Terminators näher an der EcoRI-Stelle des pBR322 liegt, wurde pLD90 genannt und die umgekehrte Anordnung wurde pLD91 genannt.

Das äußerste 5'-Ende der XPR2-Promotorregion, die in das pLS-3 eingebaut wurde, ist die PvuI-Stelle, etwa 280 Bp nach dem Anfang des in Fig. 3 sequenzierten Bereichs. Es wurde gefunden, daß Plasmide, die das Wildtyp-Protease-Gen unter der Kontrolle von nur dieser Größe des Promotors enthielten, die transformierte Spezies nicht zur Herstellung größerer Proteasemengen befähigten, wenn sie an einer Stelle in größerer Entfernung vom jeweiligen Ort des xpr2 integriert waren (beurteilt durch klare Zonen auf Magermilchschalen).

Wir stellten fest, daß, wenn pLS-3 einen gekürzten und dadurch "unvollständigen" Promotor enthielt, ein "Integrand" der aus der Rekombination zwischen dem Plasmid und einem verbliebenen XPR2-Gen vom Wildtyp entstanden war, einen vollständigen Promotor für die Expression des Prochymosin-Fusionsproduktes, aber einen unvollständigen Promotor für die Steuerung der Protease-Expression ergab. Ein analoges Genstückelungs-Experiment wurde von Shortle et al. (Science 217, 371-373, 1982) mit dem Actin-Gen von S. cerevisiae durchgeführt. In Übereinstimmung mit unseren Erwartungen fehlte tatsächlich einigen Leucin-unabhängigen, mit pLS-3 transformierten Spezies nun die Protease. Die Transformanten, denen die Protease fehlte, schienen eher die gesuchten "Integranden" am XPR2-Locus zu sein als die unerwünschten Nebenprodukte wie z.B. die Genkonvertanten am leu2-Lokus. Wir entdeckten, daß ungeschnittenes pLS-3 mit dem Empfängerstamm ATCC 20688 6,5 % Transformanten mit Protease-Defizienz hervorbrachte, während mit SnaBI geschnittenes Plasmid annähernd 70% transformierte Spezies, denen die Protease fehlte, ergab. Das Auftreten von Genstückelung bei dieser Transformation wurde benutzt, um die Notwendigkeit einer großen Zahl von Southern-Blot-Experimenten zur Auffindung des korrekten integrierenden Plasmids unter allen Transformanten zu umgehen.

Plasmide, die das Protease-Strukturgen des Wildtyps unter Kontrolle des XPR2-Promotors (mit Beginn wie in Fig. 3 sequenziert) enthalten, gestatten die Expression bedeutender Mengen Protease, wenn sie an einer anderen Stelle als der des xpr2-Locus in Y. lipolytica eingebaut werden. Die wirksame Expression heterologer Gene aus diesen Arten von "Integranen" dürfte jedoch weitere Modifikation dieser Kontrollregion der DNA erfordern.

Sekretion von Prorennin. Der Stamm ATCC 20688 von Y. lipolytica wurde mit ungeschnittener pLS-3-DNA und mit durch SnaBI verdauter pLS-3-DNA transformiert, wobei man die $xpr^- leu^+$ -Transformanten ATCC 20775 (DL144) bzw. ATCC 20776 (DL148) erhielt. Dieser transformierten Stämme wurden in ein Teströhrchen mit YEPD-Medium inokuliert. Man ließ die Zellen über Nacht bei 28°C wachsen. Eine Probe (250 µl) dieser Kulturen wurde 1:100 in 25 ml GPP-Medium verdünnt. Man ließ die Zellen in Schüttelflaschen bei 28°C 16-18 Stunden lang bis zu einer resultierenden Absorption von 5,0 bis 7,0 bei 600 nm wachsen und erntete mittels Zentrifugation. Die gebildete Kulturflüssigkeit oder der Überstand wurde auf Gegenwart von Prorennin untersucht, indem der Überstand eingeeengt und das Konzentrat SDS-PAGE unterworfen wurde. Das Platten-Gel wurde elektrophoretisch in Gegenwart von 20 mM Tris-Base, 150 mM Glycin, 20 % Methanol bei 500 mA 2 Stunden lang bei 4°C auf Nitrozellulosepapier übertragen. Die Entfernung des Proteins aus dem Platten-Gel wurde durch Anfärben mit Coomassie-Blue bestätigt.

Das Nitrozellulosepapier wurde bei 37°C getrocknet und bei 65°C eine Stunde lang gebacken und daraufhin in TBS (200 mM NaCl, 50 mM Tris-HCl, pH 7,5) gewaschen. Dann wurde das Papier bei Raumtemperatur 30 min lang in TBS, das 10 % Pferdeserum (Gibco, Chagrin Falls, Ohio) enthielt, und anschließend 16 Stunden lang in TBS, das 10 % Pferdeserum und eine entsprechende Verdünnung von Prorennin-Antikörper

enthielt, inkubiert. Dann wurde das Papier dreimal 10 min lang in TBS gewaschen, gefolgt von Inkubation in TBS, das 10 % Pferdeserum enthielt, und darauf folgender zweistündiger Inkubation in TBS, das 10 % Pferdeserum und eine geeignete Verdünnung von Ziegen-Antikörper gegen Kaninchen-IgG, konjugiert an Meerrettich-Peroxidase, enthielt. Anschließend wurde das Papier dreimal 10 min lang in TBS gewaschen und in TBS, das 0,01 % Wasserstoffperoxid enthielt, in Gegenwart von 4-Chlor-1-Naphthol (3 mg/ml in Methanol) entwickelt, das bis zu einer Konzentration von 0,5 mg/ml zugegeben wurde. Das Vorliegen von Prorennin mit einem Molekulargewicht von 40,000 wurde in beiden Überständen gesichert.

Nach saurer Aktivierung der konzentrierten Kulturüberstände (siehe oben) war in den Proben, die aus den transformierten, pLS-3 enthaltenden Kulturen ATCC 20775 und 20776 hergestellt worden war, signifikante Milchgerinnungsaktivität vorhanden. Wie erwartet gab es keine Milchgerinnungsaktivität in den Kontroll-Kulturüberständen aus dem Empfängerstamm Y. lipolytica ATCC 20688.

Konstruktion des Expressions-/Sekretionsplasmids pXX-33.

Die Modifikation zur Umwandlung von pLS-3 in ein verbessertes Expressionsplasmid pXX-33 ist in Fig. 6 in groben Zügen dargestellt. Eine solche Modifizierung vergrößerte die XPR2-Promotorregion um 280 Bp. Wie im Fall des pLS-3 enthält das Expressionsplasmid pXX-33 ein Hybrid-Gen, das für das gesamte Präpro-Peptid (157 Aminosäurereste) der alkalischen Protease codiert und mit der gesamten Strukturgen-Sequenz des Prorennins verbunden ist.

Bevor das gegenüber dem pLS-3 in der XPR2-Promotorsequenz um 280 Bp vergrößerte Prorennin-Expressions-/Sekretions-Plasmid konstruiert werden konnte, war es notwendig, ein Restriktionsfragment, das das gesamte Gen der alkalischen Protease enthielt, in eine HindIII-Spaltstelle zu subklonieren. Dieser

Subklon wurde durch die Zugabe von synthetischen Linkern zu einem Restriktionsfragment zusammengesetzt, welches aus dem Klon pLD86 einer Genombank von XPR2 isoliert worden war. Die Konstruktion dieses XPR2-Subklons mit einer strangaufwärts gelegenen HindIII-Stelle wurde durch die Herstellung eines DNA-Fragments initiiert, das das gesamte Gen der alkalischen Protease enthielt. Das 2,3 kB lange EcoRI-BamHI-(Teil-)Fragment aus der Genom-Region des XPR2-Klons pLD86 wurde mittels Agarose-Gelelektrophorese gereinigt und mit einem synthetischen Fragment mit der Sequenz

```

5'  GATCGAAGCTTG      3'
3'          TTCGAACTTAA  5'

```

ligiert. Diese Linkersequenz enthält ein kohäsives Ende für BamHI (aber stellt die BamHI-Schnittstelle nicht wieder her), gefolgt von einer HindIII-Stelle und danach einem "vorstehenden" EcoRI-Ende. Das Verknüpfungsprodukt wurde mit HindIII verdaut und in die HindIII-Schnittstelle von pBR322 insertiert. Das Plasmid pXHP-24 wurde anhand seiner Restriktionskarte identifiziert und wurde zur Quelle der XPR2-Promotorfragmente für später folgende Expressionskonstruktionen.

Im Plasmid pXHP-24 ist das subklonierte XPR2-Gen annähernd um 280 Basenpaare der 5'-XPR2-Promotorsequenz länger als die XPR2-Promotorsequenz, die im pLS-3 enthalten ist. Das Fragment der Promotorkomponente wurde zuerst durch eine Standard-Verknüpfungsreaktion unter Beteiligung des synthetischen DNA-Fragments (das oben für pLS-3 beschrieben wurde) und des 1150 Basenpaare langen HindIII-AvaI-Fragments von pHP-24 mit Hilfe von T4-Ligase und nachfolgender Spaltung mit HindIII und BamHI gebildet. Die entstandenen ligierten Sequenzen wurden mittels Gelelektrophorese gereinigt, wobei das etwa 1196 Basenpaare lange HindIII-BamHI-Fragment selektiert wurde. Ein zweites Fragment, das für die Aminosäurereste 6 bis 151 des Prorennins codiert, wurde aus pLS-3 durch Spaltung mit BamHI und XmaI und Gelreinigung des

gebildeten 440 Basenpaare langen BamHI-XmaI-DNA-Fragmentes hergestellt. Ein drittes Fragment, das den Rest des Prorennin-Gens, den XPR2-Terminator und Vektorsequenzen enthält, wurde aus pLS-3 durch Spaltung mit HindIII und XmaI und durch Gelreinigung des annähernd 8,0 kB langen HindIII-XmaI-Vektorfragmentes hergestellt. Dann wurden die drei Fragmente unter Verwendung des oben beschriebenen Standardverfahrens ligiert. Die Ligations-Reaktionsmischung wurde verwendet, um den Stamm MM294 von E. coli K12 zu transformieren. Die Plasmide wurden aus den transformierten Stämmen isoliert, auf Basis der Ampicillin-Resistenz selektioniert und das Plasmid pXX-33 wurde mit Hilfe seiner Restriktionskarte identifiziert (Fig. 6). Die Protease-Prorennin-Region dieses Plasmids wurde sequenziert, um die korrekte Verbindung der gewünschten Fragmente zu bestätigen.

Y. lipolytica ATCC 20774 wurde anschließend mit durch SnaBI geschnittenem pXX-33 transformiert, was Y. lipolytica ATCC 20780 ergab, und das von den transformierten Kulturen in die Kulturbrühe sezernierte Prorennin wurde wie oben für den Fall des pLS-3 beschrieben bestimmt. Die Gegenwart von Prorennin im Kulturenüberstand wurde bestätigt.

Nach der Säureaktivierung von konzentrierten Kulturüberständen (siehe oben) beobachtete man signifikante Milchgerinnungsaktivität in den Proben, die aus den transformierten Kulturen des Stammes Y. lipolytica ATCC 20780 hergestellt worden waren.

Konstruktion des Expressions-/Sekretions-Plasmids pXX-22. Die experimentellen Stufen, die für die Konstruktion des Expressions-/Sekretionsplasmids pXX-22 verwendet wurden, sind in Fig. 7 dargestellt. Der Expressionsvektor unterscheidet sich von pLS-3 in doppelter Hinsicht. Er enthält, wie pXX-33, das zusätzliche, 280 Bp lange Segment der XPR2-Promotorsequenz. Außerdem enthält er die für das Signalpeptid der

alkalischen Protease codierende Sequenz und nur 38 Aminosäurereste des Propetids (Prol).

Die Konstruktionsmethode für pXX-22 war der, die für pXX-33 verwendet wurde, analog. Das Fragment der Promotorkomponente wurde zuerst durch eine Standard-Verknüpfungsreaktion unter Beteiligung des 890 Basenpaare langen HindIII-BglIII-Fragmentes aus pXHP-24 und des synthetischen DNA-Fragments mit der Sequenz

```

5' GATCTTGCTGAGATCACTAG      3'
3'   AACGACTCTAGTGATCCTAG    5'

```

mit Hilfe von T4-Ligase und nachfolgender Spaltung mit HindIII und BamHI gebildet. Die entstandenen ligierten Sequenzen wurden mittels Isolierung des 920 Basenpaare langen HindIII-BamHI-DNA-Fragmentes durch Gelelektrophorese gereinigt. Ein zweites Fragment, das für die Aminosäurereste 6 bis 151 von Prorennin codiert, wurde aus pLS-3 durch Spaltung mit BamHI und XmaI und Gelreinigung des gebildeten 440 Basenpaare langen BamHI-XmaI-DNA-Fragmentes hergestellt. Ein drittes Fragment, das den Rest des Prorennin-Gens, den XPR2-Terminator und Vektorsequenzen enthält, wurde aus pLS-3 durch Spaltung mit HindIII und XmaI und durch Gelreinigung des annähernd 8,0 kB langen HindIII-XmaI-Vektorfragmentes gereinigt. Dann wurden die drei Fragmente unter Verwendung des oben beschriebenen Standardverfahrens ligiert. Die Ligations-Reaktionsmischung wurde verwendet, um den Stamm MM294 von E. coli K12 zu transformieren. Die Plasmide wurden aus den transformierten Stämmen isoliert, und das Plasmid pXX-22 wurde mit Hilfe seiner Restriktionskarte identifiziert (Fig. 7).

Y. lipolytica ATCC 20774 wurde anschließend mit durch SnaBI geschnittenen pXX-22 transformiert, was Y. lipolytica ATCC 20779 ergab, und das von den transformierten Kulturen in die Kulturbrühe sezernierte Prorennin wurde wie oben für den Fall

des pLS-3 beschrieben bestimmt. Die Gegenwart von Prorennin im Kulturenüberstand wurde nach dem oben beschriebenen Verfahren bestätigt. Nach der sauren Aktivierung von konzentrierten Kulturüberständen (siehe oben) beobachtete man signifikante Milchgerinnungsaktivität in den Proben, die aus den transformierten Kulturen des Stammes Y. lipolytica ATCC 20779 hergestellt worden waren.

Konstruktion des Expressions-/Sekretions-Plasmids pXX-11. Die experimentellen Stufen für die Konstruktion des Expressions-/Sekretionsplasmids pXX-11 sind in Fig. 8 in groben Zügen dargestellt. Dieses Plasmid enthält die Sequenz für den XPR2-Promotor und das 22 Aminosäurereste lange Signalpeptid in Verknüpfung mit der Sequenz, die für Prorennin codiert. Die Konstruktionsmethode, die für pXX-11 angewandt wurde, war der für pXX-22 und pXX-33 benutzten ähnlich. Kurz gesagt wurde das Fragment der Promotorkomponente durch eine Standard-Verknüpfungsreaktion unter Beteiligung des etwa 750 Basenpaare langen HindIII-BglII-DNA-Fragmentes aus pXHP-24 und des synthetischen DNA-Fragments mit der Sequenz

```

5'      TGGCCGCTCCCCTGGCCGCCCTGCCGCTGAGATCACTAG      3'
3'AAGACCGGCGAGGGGACCGCGGGGACGGCGACTCTAGTGATCCTAG 5'———>

```

mit Hilfe von T4-Ligase und nachfolgender Spaltung mit HindIII und BamHI gebildet. Die entstandenen ligierten Sequenzen wurden mittels Gelelektrophorese gereinigt, wobei das etwa 790 Basenpaare lange HindIII-BamHI-DNA-Fragment selektiert wurde. Ein zweites Fragment, das für die Aminosäurereste 6 bis 151 des Prorennins codiert, wurde aus pLS-3 durch Spaltung mit BamHI und XmaI und Gelreinigung des gebildeten 440 Basenpaare langen BamHI-XmaI-DNA-Fragmentes hergestellt. Ein drittes Fragment, das den Rest des Prorennin-Strukturgens, den XPR2-Terminator und Shuttle-Vektor-Sequenzen enthält, wurde aus pLS-3 durch Spaltung mit HindIII und XmaI und durch Gelreinigung des annähernd 8,0 kB langen

Vektorfragments gereinigt. Dann wurden die drei DNA-Fragmente unter Verwendung des oben beschriebenen Standardverfahrens ligiert. Die Ligations-Reaktionsmischung wurde verwendet, um den Stamm MM294 von E. coli K12 zu transformieren. Die Plasmide wurden aus den selektierten Transformanten isoliert und das Plasmid pXX-11 wurde mit Hilfe seiner Restriktionskarte identifiziert (Fig. 8). Der XPR2-Prorennin-Teil dieses Plasmids wurde sequenziert, um die richtige Sequenz der synthetischen DNA und die korrekte Verbindung der gewünschten Fragmente zu bestätigen.

- . Y. lipolytica ATCC 20774 wurde anschließend mit durch SnaBI geschnittenem pXX-11 transformiert, was Y. lipolytica ATCC 20778 ergab, und das von den transformierten Kulturen in die Kulturbrühe sezernierte Prorennin wurde wie oben für den Fall des pLS-3 beschrieben bestimmt. Die Gegenwart von Prorennin im Kulturenüberstand wurde dem oben beschriebenen Verfahren entsprechend bestätigt.

Bestimmungen der Milchgerinnung (siehe oben) zeigten die Gegenwart von signifikanter Milchgerinnungsaktivität in den Kulturüberständen der pXX-11-haltigen Transformanten ATCC 20778.

BEISPIEL 2Konstruktion der Eingangsregion (Docking Platform)

Das Wildtyp-BIO-Gen, das dem bio-6-Allel in ATCC 20774 entspricht, wurde mittels Komplementierung wie folgt kloniert. Eine Genbank einer partiell mit Sau3A verdauten, chromosomalen Y. lipolytica-DNA, eingefügt in die BamHI-Schnittstelle von pLD40 (welches pBR322 plus LEU2 an der EcoRI-Stelle ist) wurde konstruiert und eine große Menge der Genbank-DNA wurde als Plasmidpräparation von verschiedenen E. coli Kulturen hergestellt. (Diese ist mit der Genbank identisch, die für die Klonierung des XPR2-Gens verwendet wurde.) Einige Mikrogramm der Genbank-DNA wurden mit dem Enzym ApaI (das einmal in der LEU2-Region schneidet) verdaut. Dann wurde diese DNA benutzt, um ATCC 20774 (leu2 xpr2 bio) zu transformieren, wobei die Transformationsmischung auf synthetischem Medium, das kein Leucin enthielt, ausplattiert wurde. Man erhielt Zehntausende von leucin-unabhängigen, transformierten Spezies. Um herauszufinden, welche Kolonien, wenn überhaupt, Genbankplasmide enthielten, die das BIO-Gen umfaßten, wurden die leucinunabhängigen transformierten Spezies auf Agar-Platten abklatschplattiert, die Biotin-Selektionsmedium enthielten (Rp. pro Liter: 25 mg Desthiobiotin, 20 g Glucose, 5 g Ammoniumsulfat, 1 g KH_2PO_4 , 0,5 g $\text{MgSO}_4 \times 7 \text{H}_2\text{O}$, 0,1 g CaCl_2 , 0,1 g NaCl, 500 μg Borsäure, 400 μg Thiamin x HCl, 400 μg $\text{ZnSO}_4 \times 7 \text{H}_2\text{O}$, 400 μg MnSO_4 , 200 μg $\text{Na}_2\text{MoO}_4 \times 2 \text{H}_2\text{O}$, 200 μg $\text{FeCl}_3 \times 6 \text{H}_2\text{O}$, 100 μg KI und 40 μg $\text{CuSO}_4 \times 5 \text{H}_2\text{O}$).

Eine von mehreren BIO⁺-transformierten Y. lipolytica-Spezies, die auf dem Biotin-Selektionsmedium wuchsen, wurde mit DL31 bezeichnet. Es gelang uns im folgenden, das Genbankplasmid, das das BIO-Gen enthielt, aus dem Y. lipolytica-Stamm DL31 zu gewinnen. Chromosomale DNA wurde aus einer Kultur des Stammes DL31 hergestellt. Einige Mikrogramm dieser chromosomalen DNA wurden mit dem Restriktionsenzym ApaI verdaut, um das

Genbankplasmid herauszuschneiden. Eine Probe der verdauten DNA wurde in einer Ligationsreaktion eingesetzt, um das unbekannte Genbankplasmid zu zirkularisieren. Die Ligationsmischung wurde daraufhin benutzt, um eine Kultur von E. coli auf Ampicillinresistenz zu transformieren, um das unbekannte, BIO-haltige Plasmid in E. coli zu gewinnen. Man erhielt ein paar transformierte, ampicillinresistente E. coli-Spezies. Plasmidpräparationen wurden in kleinem Maßstab an den transformierten E. coli-Spezies ausgeführt. Die so erhaltenen Restriktionsbruchstücke der Plasmid-DNA zeigten, daß das unbekannte, BIO-haltige Plasmid, wie erwartet, dem pLD40 mit einem in die BamHI-Schnittstelle insertierten Bruchstück äquivalent war. Dieses Plasmid muß ursprünglich aus unserer Genbank stammen und wurde mit pLD51 bezeichnet.

Das Plasmid pLD56 wurde als Subklon von pLD51 durch die Abtrennung des LEU2-Gens aus pLD51 wie folgt gebildet. Eine Probe des Plasmids pLD51 wurde mit dem Enzym EcoRI verdaut, um die LEU2-Region zu entfernen. Die verdaute DNA wurde in einer DNA-Ligationsreaktion benutzt, um das Plasmid zu rezirkularisieren. Anschließend wurde eine Transformation von E. coli durchgeführt, um das kleinere, BIO-haltige Plasmid zu klonieren. Es ließ sich zeigen, daß einer der ampicillinresistenten, transformierten E. coli-Stämme das erwartete kleinere Plasmid enthielt, das mit pLD56 bezeichnet wurde. Einige Restriktionsbruchstücke von pLD56 wurden hergestellt. Das BIO-haltige Segment von pLD56, das als Insertion an der BamHI-Stelle von pBR322 erscheint, war annähernd 3,6 kB lang.

Eine sehr ungenaue Restriktionskarte der 3,6 kB langen, in die BamHI-Stelle von pBR322 insertierten Y. lipolytica-DNA (die pLD56 enthält) ist unten mit den ungefähren Abständen der Basenpaare vom Anfangspunkt der Insertion an gerechnet in Klammern beschrieben. Die Größenabschätzungen wurden von nur wenigen Agarosegelen her vorgenommen und unterliegen relativ

großen quantitativen Fehlern: PvuII (800), PvuII (1200), PstI (1800), MluI (2000), PstI (2000), EcoRV (2700), NcoI (3200) (Für die Orientierung: die SalI-Stelle von pBR322 sollte vor den beschriebenen Schnittstellen liegen und die HindIII-Stelle sollte ihnen folgen).

Der Stamm ATCC 20774 (MATB leu2-40 bio-6 xpr2-1002) wurde mit intakter pLD56 (pBR322 plus annähernd 3,6 kB aus der chromosomalen DNA von Y. lipolytica, die das BIO-Gen enthalten) transformiert. Drei verschiedene Biotin-unabhängige transformierte Spezies wurden auf im Vergleich zum Ausgangsstamm hochfrequente Transformation von durch NruI geschnittenes (auf pBR322 gerichtet) pLD40 (LEU2 auf pBR322) getestet, um festzustellen, welche ein in die BIO-Region eingebautes ortsfestes pBR322 enthielt. Alle drei zeigten wegen der Integration des pLD40 in das ortsfeste pBR322 eine hochfrequente Transformation. Dies wurde durch Hybridisierungsexperimente nach dem Southern-Blot-Verfahren bestätigt. Eine der drei ursprünglichen BIO-transformierten Y. lipolytica-Spezies wurde DL118 genannt und weiter als DNA-Empfänger benutzt. Die obige Restriktionskarte wurde benötigt, um zu bestimmen, i) was als BIO-spezifische Hybridisierungssonde verwendet werden sollte (ein NcoI-PvuII-Stück), ii) welches Enzym benötigt würde, um das pLD56-Plasmid korrekt herauszuschneiden (MluI), iii) welches Enzym nur einmal im pBR322-Teil schneiden würde (ClaI) und iv) welches Enzym im gesamten Plasmid nicht schneiden würde (ApaI). Southern-Hybridisierungen von durch ClaI und ApaI verdauten Bruchstücken von DNA aus ATCC 20774 und DL118 (mit einem BIO-Fragment als Sonde) zeigten, daß wie erwartet die Biotin-Region von DL118 (im Vergleich zu der BIO-Region von ATCC 20774) durch einen zusätzlichen Einbau von DNA mit der ungefähren Länge von pLD56 unterbrochen war. MluI-Verdauungen an DL118-DNA (getestet mit pBR322) zeigten außerdem, daß dieser Zusatz die gleiche Größe wie intaktes pLD56 besaß.

Konstruktion des Expressions-/Sekretions-Plasmids pLX-34. Ein Expressionsplasmid wurde konstruiert, welches die für Prorennin codierende Sequenz mit den XPR2-Sekretions-Signalen (157 Aminosäuren Präpro-Sequenz) stangabwärts vom LEU2-Promotor aus gesehen placiert. Dieses Expressionsplasmid zeigt auf, daß auch ein anderer Promotor als der XPR2-Promotor benutzt werden kann, um die Sekretion von heterologem Protein zu bewirken. Weiterhin ist dieser Expressionsvektor in der Lage, unabhängig von der Einbaustelle im Y. lipolytica-Genom Expression/Sekretion zu bewirken. Die gelungene Sekretion von Prorennin mit einem anderen Promotor als dem XPR-Promotor zeigt die Durchführbarkeit einer Expressions-Vektor-Konstruktion zur Identifizierung alternativer neuer, starker Promotoren in Y. lipolytica. Zusätzlich kann dieser Ansatz benutzt werden, um eine Expressionskultur mit zwei getrennten Prorennin-Hybridgenen zu gewinnen, wobei eines vom LEU2-Promotor und das andere vom XPR2-Promotor exprimiert wird, die an verschiedenen Stellen im Genom des Wirtes eingebaut sind.

Die experimentellen Stufen, die für die Konstruktion eines Expressionsvektors angewendet werden, welcher das Prorennin-Gen mit Sekretions-Signalen für die alkalische Protease (157 Aminosäuren lange XPR2-Präpro-Sequenz), exprimiert durch LEU2-Promotor-Sequenzen, enthält, sind in Fig. 13 dargestellt. Die Konstruktion dieses Plasmids wurde mit der Herstellung eines LEU2-Promotor-Fragmentes begonnen, welches etwa 300 Basenpaare der 5'-untranslatierten Sequenz enthält, die dem ATG-Codon, das die Translation des beta-Isopropylmalat-Dehydrogenase-Gens (Fig. 12) initiiert, vorangehen. Das 300 Bp lange HindIII-FokI-DNA-Fragment, das für einen 270 Bp langen Teil der LEU2-Promotorsequenz codiert, wurde aus dem Shuttle-Vektor pLD40 isoliert. Dieses Fragment wurde mit einem 54 Bp langen Linker mit der Sequenz

```

-----Leu2-----
      PstI
5'- ATACAACCACACACATCCACAATG
3'-      TTGGTGTGTGTAGGTGTTAC

-----Xpr2-----
                                     BglI
AAGCTCGCTACCGCCTTACTATTCTCACTGCCGTTTC-3'
TTCGAGCGATGGCGGAAATGATAAGAGTGACGGC   -5'

```

mit T4-Ligase verknüpft und anschließend mit HindIII verdaut. Die entstandenen ligierten Sequenzen wurden durch Isolierung des etwa 360 Basenpaare langen HindIII-BglI-DNA-Fragmentes mittels Gelelektrophorese gereinigt. Ein zweites, als Bestandteil nötiges Fragment, das für den Rest der XPR2-Präpro-Sequenz und die ersten 152 Aminosäuren des Prorennins codiert, wurde durch Spaltung mit BglI und XmaI aus dem Expressionsplasmid pXX-33 (Fig. 6) und Gelreinigung des gebildeten 887 Basenpaare langen DNA-Fragmentes isoliert. Ein drittes Fragment, das den Rest des Prorennin-Gens, den XPR2-Terminator und Vektorsequenzen enthält, wurde durch Spaltung von pXX-33 mit HindIII und XmaI und Gelreinigung (durch Spaltung) des etwa 8,0 kB langen Vektorfragmentes hergestellt. Die drei DNA-Fragmente wurde unter Verwendung der oben beschriebenen Standardverfahren miteinander ligiert. Die Ligations-Reaktionsmischung wurde verwendet, um den Stamm HB101 von E. coli Kl² zu transformieren. Plasmide wurden aus den transformierten Spezies auf der Basis der Ampicillin-resistenz isoliert und das Plasmid pLX-34 wurde durch seine Restriktionskarte identifiziert (Fig. 13).

Y. lipolytica ATCC 20794 (DL118) wurde mit durch NruI gespaltene pLX-34-DNA behandelt, wobei man Y. lipolytica ATCC 20795 (DL251) erhielt, und das von der leucin-unabhängigen, transformierten Kultur in die Kulturflüssigkeit sezernierte Prorennin wurde wie oben für das pLS-3 beschrieben bestimmt. Diese Transformationstechnik lenkte den Einbau von pLX-34 in eine pBR322-Sequenz, die vorher in den bio-Locus im Wirts-

chromosom (oben beschrieben) eingebracht worden war. Der Einbau von pLX-34 an dieser Stelle wurde durch Southern-Analyse bestätigt.

Die Verwendung von DL118 als Empfänger. Southern-Hybridisierungsexperimente wurden wie folgt ausgeführt:

NruI-Verdauung von DNA aus transformierten DL118-Spezies (hybridisiert mit z.B. einem Prochymosin-Teststück, wenn das eingefügte Plasmid ein Prochymosin-Expressionsplasmid war) schnitt genau das eingefügte Plasmid heraus. Ein paar Nanogramm des durch NruI verdauten transformierenden Plasmids diente zur Überprüfung der korrekten Größe der Hybridisierungsbande. Auch die Verdauung von DNA dieser transformierten Spezies (getestet mit ^{32}p -markiertem pBR322) durch MluI (MluI schnitt das transformierende Plasmid nicht) zeigte, daß die ortsansässige pBR322-Sequenz von DL118 durch Zugabe von einem oder mehreren Molekülen des transformierenden Plasmids unterbrochen wurde. Dies zeigte, daß der Einbau an der gewünschten Stelle stattfand.

Die transformierte Kultur Y. lipolytica ATCC 20795 (DL251) wurde in YEPD-Medium bei 22°C gezüchtet, um die Expression durch den LEU2-Promotor zu begünstigen. Das Vorliegen von Prorennin in den Kulturüberständen wurde mit Hilfe des Milchgerinnungstests (siehe oben) an sauer aktivierten Kulturüberständen gesichert und durch Immuno-Blot-Analyse (siehe oben) bestätigt. Diese Ergebnisse zeigen, daß dieses Hybrid-Gen eine unabhängige Expressionseinheit ist, die zur Expression/Sekretion befähigt ist, wenn sie an einem anderen Ort als XPR2 oder LEU2 eingebaut wird. Diese Eigenschaft erlaubt die Konstruktion einer Expressionskultur mit mehrfachen Hybrid-Genen, die möglicherweise in der Lage sind, gesteigerte Mengen an extrazellulärem Prorennin zu erzielen.

BEISPIEL 3

Sequenz eines synthetischen Gens für menschliches C5a. Der Plan, der zum Erreichen der bakteriellen Produktion von menschlichem Anaphylatoxin C5a entwickelt wurde, war vorausgehenden Verfahren analog, die für die Synthese und Expression von EGF verwendet wurden, wie in der europäischen Patentanmeldung Nr. 0147178 beschrieben. Er sieht die Konstruktion eines Gens vor, in dem die codierende Sequenz für die aktivierte Komplementkomponente C5a synthetisch hergestellt wurde. Da die bekannte Aminosäuresequenz von menschlichem C5a festgelegt ist, entwarfen wir ein DNA-Fragment, das für die Information ihrer 74 Aminosäuren codiert (Fig. 9). Die synthetische Gensequenz war so ausgewählt, daß sie die von E. coli und S. cerevisiae bevorzugte Codon-Benutzung maximiert und Stellen für mehrere Restriktionsendonukleasen anbietet, um die Charakterisierung zu erleichtern. Dieser Ansatz erlaubte die direkte Expression von Anaphylatoxin in E. coli, indem man ein ATG-Initiationscodon für die Proteinsynthese vor das für die erste Aminosäure des C5a Polypeptids codierende Triplet einführt. Um seine Insertion in der gewünschten Anordnung in das Plasmid pBR322 zu erleichtern, wurde das synthetische C5a-Gen so konstruiert, daß es an seinen Enden Erkennungsstellen für die Restriktionsendonukleasen EcoRI und HindIII enthielt. Um die zu bildende C5a-Gensequenz herzustellen, wurden zehn 47-merer mittels des Phosphoramidit-Verfahrens synthetisiert und zu einem 235 Basenpaare langen, doppelsträngigen DNA-Fragment zusammengefügt. Das C5a-Genfragment wurde in an der geeigneten Stelle gespaltenes pBR322 insertiert und das klonierte Gen wurde mittels Analyse durch Restriktionsspaltung von Plasmid-DNA aus willkürlich ausgewählten transformierten Spezies identifiziert. Mehrere C5a-Klone wurden dann durch Sequenzieren der DNA analysiert, um einen Klon mit der richtigen Sequenz zu identifizieren. Die gewünschte Nukleotidsequenz für die C5a-Gen-Region wurde

in 2 der 5 untersuchten Klone gefunden.

Bakterielle Expression von menschlichem C5a. Die Konstruktion des C5a-Expressions-Plasmids wurde durch Spaltung des C5a-Subklons mit der Restriktionsendonuklease EcoRI begonnen, gefolgt von Dephosphorylierung durch Behandeln mit bakterieller alkalischer Phosphatase. Durch Verwenden eines 360 Bp langen EcoRI-DNA-Fragments aus pPFZ-R2, das den trp-Promotor-Operator und Ribosomenbindungsstellen-Sequenzen enthält, wurde ein C5a-Expressionsplasmid konstruiert. Kompetente E. coli-Zellen vom Stamm HB101 wurden mit der Ligationsreaktionsmischung transformiert. Einige Antibiotika-resistente Kolonien aus jeder Transformation wurden gereinigt und ihre Plasmid-DNA's wurden der Analyse durch Restriktionsendonukleasen-Kartierung unterworfen, um diejenigen zu identifizieren, in denen der trp-Promotor so angeordnet ist, daß Transkription des C5a-Gens resultiert. Mehrere isolierte Spezies aus dieser Verknüpfungsreaktion wurden als Plasmide identifiziert, die das Anaphylatoxin-Gen in Nachbarschaft zur bakteriellen Promotorsequenz in der Anordnung enthielten, die für die direkte Expression von C5a benötigt wird. Eine Restriktionskarte des C5a-Expressionsplasmids pC5a-48 ist in Fig. 10 dargestellt.

Expression und Sekretion von menschlichem Anaphylatoxin in *Y. lipolytica*

Der Expressions-/Sekretionsvektor pC5aX-3, der für die Sekretion von menschlichem Anaphylatoxin C5a codiert, wurde nach Techniken hergestellt, wie sie in Beispiel 1 für pXX-33 beschrieben wurden. *Y. lipolytica* ATCC 20774 wurde anschließend durch diesen Sekretionsvektor transformiert und das durch die transformierten Kulturen produzierte menschliche C5a wurde wie oben beschrieben bestimmt, mit der Ausnahme, daß Anti-C5a der Ziege und Anti-Ziegen-IgG des Kaninchens im Immuno-Blot-Verfahren verwendet wurden. Für das in diesem

Beispiel beschriebene Plasmid wurde die Anwesenheit von C5a im Kulturüberstand bestätigt.

Konstruktion des Expressions-/Sekretions-Plasmids pC5aX-3.

Die experimentellen Stufen für die Konstruktion des Anaphylatoxin-Expressions-/Sekretions-Plasmids pC5aX-3 sind in Fig. 11 dargestellt. Dieses Plasmid enthält die Sequenz für den "vollständigen" XPR2-Promotor und für das 157 Aminosäurereste lange Signal- und Propeptid, die an eine synthetische, für die 74 Aminosäurereste von C5a codierende Sequenz geknüpft sind. Der Konstruktionsplan, der für das pC5aX-3 angewendet wurde, war dem für pXX-33 angewandten ähnlich. Zuerst wurde das Plasmid pXHP-24 (oder ein anderes Plasmid, das die gewünschte Sequenz enthält) mit HindIII und AvaI gespalten und das 1150 Basenpaare lange Fragment, das den XPR2-Promotor enthält, wurde mittels Gel gereinigt. Ein zweites Fragment, das das 3'-Ende der XPR2-Propeptid- und der C5a-Strukturgen-Sequenz enthält, wurde mit Hilfe einer Standard-Ligationsreaktion gebildet, wobei das annähernd 220 Basenpaare lange HinfI-HindIII-DNA-Fragment aus dem E. coli-Expressionsplasmid pC5a-48 und das synthetische Fragment

```

5'   CCGAGATTCCTGCTTCTTCTAATGCCAAGCGA   3'
3'           CTAAGGACGAAGAAGATTACGGTTCGCTTGA   5'

```

mit T4-Ligase eingesetzt wurden, gefolgt von Spaltung mit AvaI und HindIII. Die entstandenen ligierten Sequenzen wurden durch Gelelektrophorese gereinigt, wobei man das etwa 250 Basenpaare lange AvaI-HindIII-Fragment selektierte. Das den Promotor enthaltende HindIII-AvaI-Fragment und das für C5a codierende AvaI-HindIII-Fragment wurden darauf mit T4-Ligase ligiert und anschließend mit HindIII verdaut. Das etwa 1,4 kB lange Fragment wurde mit Hilfe von Gel gereinigt und für eine Verknüpfung mit durch HindIII gespaltenem ptem 4 (oben beschrieben) verwendet. Die Ligations-Reaktionsmischung wurde verwendet, um den Stamm MM294 von E. coli K12 zu transformieren. Auf Ampicillinresistenz selektierte Plasmide wurden

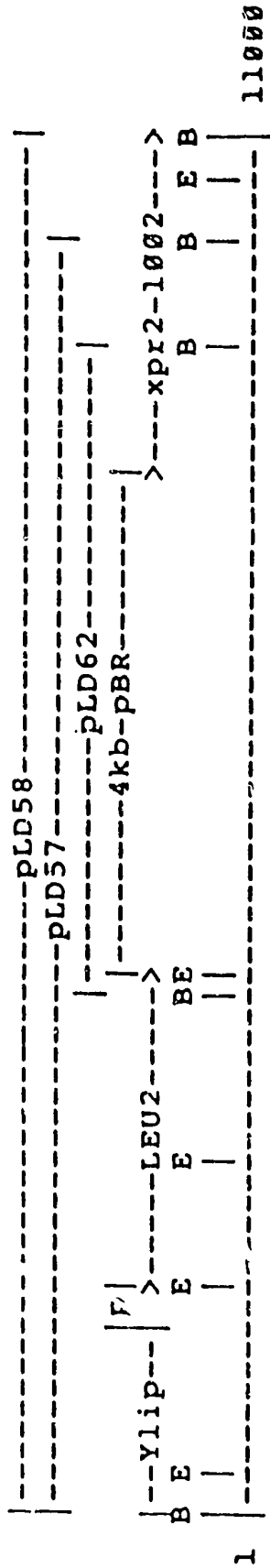
aus den selektierten transformierten Spezies isoliert und das Plasmid pC5aX-3 wurde anhand seiner Restriktionskarte identifiziert. Y. lipolytica, Stamm PC-30869, ATCC 20774, wurde dann mit durch SnaBI gespaltenem pC5aX-3 transformiert und das von den transformierten Kulturen in das Kulturmedium sezernierte Anaphylatoxin wurde wie oben beschrieben bestimmt. Die Anwesenheit von C5a im Kulturüberstand wurde durch das oben beschriebene Verfahren bestätigt.

Es ist bekannt, daß viele Proteine, die durch an das endoplasmatische Reticulum gebundene Ribosomen synthetisiert werden, als Glycoproteine produziert werden. In der Tat kann die Glycosylierung die Sekretion eines gegebenen Proteins beeinflussen. N-verknüpfte Glycosylierung von eukaryotischen Proteinen tritt an den Tripeptid-Sequenzen Asparagin-X-Threonin und Asparagin-X-Serin auf, wobei X eine beliebige Aminosäure außer möglicherweise Aspartat (Hubbard, S. et al. (1981), Ann. Rev. Biochem. 50, 555) sein kann. Die Aminosäuresequenz von Prorennin enthält zwei solche Tripeptid-Sequenzen, eine Analyse von in Y. lipolytica sezerniertem Prorennin mittels Gelelektrophorese ergab jedoch keinen Hinweis auf die Glycosylierung. In anderen sezernierten eukaryotischen Proteinen sind nicht alle Asparagin-X-Threonin/-Serin-Stellen glycosyliert. Es ist wahrscheinlich, daß bestimmte Asparagine innerhalb von Tripeptid-Sequenzen nicht glycosyliert werden, da sie für die Glycosylierungsenzyme unerreichbar sind.

Im Fall von menschlichem C5a enthält die Aminosäuresequenz eine einzige Glycosylierungsstelle oder Tripeptid-Sequenz (Asn-Ile-Ser), die normalerweise ein komplexes Oligosaccharid besitzt, das an Asparagin gebunden ist (Fernandez, H., et al. (1976), J. Immunol. 117, 1688). Ein Teil der in das Kulturmedium von Y. lipolytica sezernierten C5a-Moleküle scheint glycosyliert zu sein, da ein breiter Bereich antigener Aktivität im Bereich der hohen Molekulargewichte

des Immunoblots zu sehen war. Diese heterogene elektrophoretische Mobilität ist der analog, die man an anderen sezernierten Proteinen beobachtet hat, und sie ist wahrscheinlich eine Folge des schwankenden Grades an Kohlehydrataddition. Für die vorliegende Erfindung läßt die scheinbare Glycosylierung bestimmter sezernierter heterologer Proteine vermuten, daß die Expression und Sekretion in Y. lipolytica für die Herstellung von vielen, normal glycosylierten eukaryotischen Proteinen nützlich sein wird.

FIG. 1



Schematische Darstellung der Plasmide wie aus der mit XPR2 transformierten Spezies DL112 gewonnen wurden. Die markierten Regionen umfassen (von rechts nach links) Y. lipolytica-Material, das nicht sequenziert wurde, das kleine Segment von pBR322 zwischen der Eco_{RI}-Stelle und der Verknüpfung mit der Y. lipolytica-DNA (vorher eine BamHI-Stelle), das LEU2-Gen von Y. lipolytica, das große pBR322-Segment und das mutierte Allel von XPR2, das sequenziert wurde. BglII-Schnittstellen sind mit "B" und EcoRI-Stellen mit "E" bezeichnet. Die drei in inearisierter Form gezeigten Plasmide wurde als zirkulare Moleküle, verknüpft an ihren terminalen BglII-Stellen, gewonnen.

FIG. 2

I	Proteinsequenz	val thr gln trp gly	
	mRNA 5'	GTX ACX CAU TGG GG	3'
	Bruch- stück 3'	CAX TGX GTY ACC CC	5'
	170 3'	CAA TGX GTY ACC CC	5'
	172 3'	CAT TGX GTY ACC CC	5'
	174 3'	CAG TGX GTY ACC CC	5'
	176 3'	CAC TGX GTY ACC CC	5'
II	Proteinsequenz	lys lys ala gln thr	
	mRNA 5'	AAU AAU GCX CAU AC	3'
	Bruch- stück 3'	TTY TTY CGX GTY TG	5'
	180 3'	TTC TTY CGX GTC TG	5'
	182 3'	TTT TTY CGX GTC TG	5'
	184 3'	TTT TTY CGX GTT TG	5'
	186 3'	TTC TTY CGX GTT TG	5'

GATCCGGGGTTCTATGCATCCCTGAAACATTGATTGAAATTAACATATGAGCTGCGTG
 CTTTTTGCATTCAAGGGCGCAGCTTATCTTGTATCCTTAATTACACATGACCTCTTGAG
 CGCCACGGTACATTCTGGCGTCAGTTCGGTGGAGCGGACACTTTTCTCTCCTTTGTCT
 GACATGTTGGTTAAGTTGTAGTCCAGGGACACAAGGGGTCCAACGGCAGTGGCAGCCT
 ACCCCACGCTACCCACCACTGGCCCTGGTCTAACTTCGACGATCGGCATCAGGGTTCAT
 GGATAGCGGGTGTGATTTACGATGTGATGGACAATGTTAGAGAGATCCCACTACTTGTA
 GTCAGGCCATCTTTTACGTACGCACTGTACCATGATGTCAATGGAGTATGATGAACCGA
 CTTTGAGAGACTCACATCTGCACAACACCATGTTTCAGCGGAATCCGACTTCCAACCGA
 AACCCAAGCCCCTGTCAGATATCGTGAGAAGGCACGGCACCAACTAATGCACACACTCC
 ACCTGTATTGCACCAAGATAATGAGGGCATCGTCTTGGCGGTCTTGGCGAGAGCCGTG
 TTTCGTGACGCAATCAGAGCAGTTTCTGGATAGTATCTTGTCCAGAAACACGATATAAA
 CCCCATCGACGGGCCCCTTGAAGAGCACCAACCCACTATCCAATCCTCCAATCCAACA

-157

met lys leu ala thr ala phe thr ile leu thr ala val leu ala
 ATG AAG CTC GCT ACC GCC TTT ACT ATT CTC ACT GCC GTT CTG GCC

-150

-140

ala pro leu ala ala pro ala pro ala pro asp ala ala pro ala
 GCT CCC CTG GCC GCC CCT GCC CCT GCT CCT GAT GCT GCC CCT GCT

-130

-120

ala val pro glu gly pro ala ala ala ala tyr ser ser ile leu
 GCT GTG CCT GAG GGC CCT GCC GCC GCT GCC TAC TCA TCT ATT CTG

-110

ser val val ala lys gln ser lys lys phe lys his his lys arg
 TCC GTG GTC GCT AAG CAG TCC AAG AAG TTT AAG CAC CAC AAG CGA

-100

-90

asp leu asp glu lys asp gln phe ile val val phe asp ser ser
 GAT CTT GAT GAG AAG GAT CAG TTC ATC GTT GTC TTT GAC AGT AGC

-80

ala thr val asp gln ile ala ser glu ile gln lys leu asp ser
 GCT ACT GTT GAC CAG ATC GCC TCC GAA ATC CAG AAG CTG GAC TCT

-70

-60

leu val asp glu asp ser ser asn gly ile thr ser ala leu asp
 CTG GTC GAC GAG GAC TCG TCC AAC GGT ATC ACC TCT GCT CTT GAT

-50

leu pro val tyr thr asp gly ser gly phe leu gly phe val gly
 CTT CCT GTC TAC ACG GAT GGA TCT GGC TTT CTC GGA TTT GTT GGA

-40

-30

lys phe asn ser thr ile val asp lys leu lys glu ser ser val
 AAG TTC AAC TCC ACT ATC GTT GAC AAG CTC AAG GAG TCG TCT GTT

-20

leu thr val glu pro asp thr ile val ser leu pro glu ile pro
 CTG ACC GTC GAG CCC GAT ACC ATT GTG TCT CTC CCC GAG ATT CCT

-10

1

ala ser ser asn ala lys arg ala ile gln thr thr pro val thr
 GCT TCT TCT AAT GCC AAG CGA GCT ATC CAG ACT ACT CCC GTC ACT

10

gln trp gly leu ser arg ile ser his lys lys ala gln thr gly
 CAA TGG GGC CTG TCT AGA ATC TCT CAT AAG AAG GCC CAG ACT GGA

20

30

asn tyr ala tyr val arg glu thr val gly lys his pro thr val
 AAC TAC GCC TAC GTT CGA GAG ACA GTT GGC AAG CAC CCC ACC GTT

40

ser tyr val val asp ser gly ile arg thr thr his ser glu phe
 TCT TAC GTT GTT GAC TCT GGT ATC CGA ACC ACC CAC TCC GAG TTC

50

60

gly gly arg ala val trp gly ala asn phe ala asp thr gln asn
 GGA GGC CGA GCT GTC TGG GGA GCC AAC TTC GCT GAC ACA CAG AAC

70

ala asp leu leu gly his gly thr his val ala gly thr val gly
 GCT GAI CTT CTC GGT CAC GGC ACT CAC GTT GCA GGT ACC GTG GGA

80

90

gly lys thr tyr gly val asp ala asn thr lys leu val ala val
 GGA AAG ACA TAC GGA GTC GAC GCC AAC ACC AAG CTG GTG GCC GTC

100

lys val phe ala gly arg ser ala ala leu ser val ile asn gln
 AAG GTG TTT GCA GGC CGA TCC GCA GCT CTC TCC GTC ATC AAC CAG

110

120

gly phe thr trp ala leu asn asp tyr ile ser lys arg asp thr
 GGC TTC ACC TGG GCT CTC AAC GAC TAC ATC TCC AAG CGA GAC ACT

130

leu pro arg gly val leu asn phe ser gly gly gly pro lys ser
 CTG CCT CGA GGA GTG CTG AAC TTC TCT GGA GGA GGA CCC AAG TCC

140

150

ala ser gln asp ala leu trp ser arg ala thr gln glu gly leu
 GCT TCC CAG GAC GCC CTA TGG TCT CGA GCT ACC CAG GAG GGT CTG

160

leu val ala ile ala ala gly asn asp ala val asp ala cys asn
 CTT GTC GCC ATC GCT GCG GGA AAC GAT GCC GTG GAC GCC TGT AAC

170

180
 asp ser pro gly asn ile gly gly ser thr ser gly ile ile thr
 GAC TCT CCC GGT AAC ATT GGA GGC TCC ACC TCT GGT ATC ATC ACT

190 200
 val gly ser ile asp ser ser asp lys ile ser val trp ser gly
 GTG GGT TCC ATT GAC TCT AGC CAT AAG ATC TCC GTC TGG TCC GGT

210
 gly glu gly ser asn tyr gly thr cys val asp val phe ala pro
 GGA CAG GGA TCC AAC TAC GGA ACT TGT GTT GAT GTC TTT GCC CCC

220 230
 gly ser asp ile ile ser ala ser tyr gln ser asp ser gly thr
 GGC TCC GAT ATC ATC TCT GCC TCT TAC CAG TCC GAC TCT GGT ACT

240
 leu val tyr ser gly thr ser met ala cys pro his val ala gly
 TTG GTC TAC TCC GGT ACC TCC ATG GCC TGT CCC CAC GTT GCC GGT

250 260
 leu ala ser tyr tyr leu ser ile asn asp glu val leu thr pro
 CTT GCC TCC TAC TAC CTG TCC ATC AAT GAC GAG GTT CTC ACC CCT

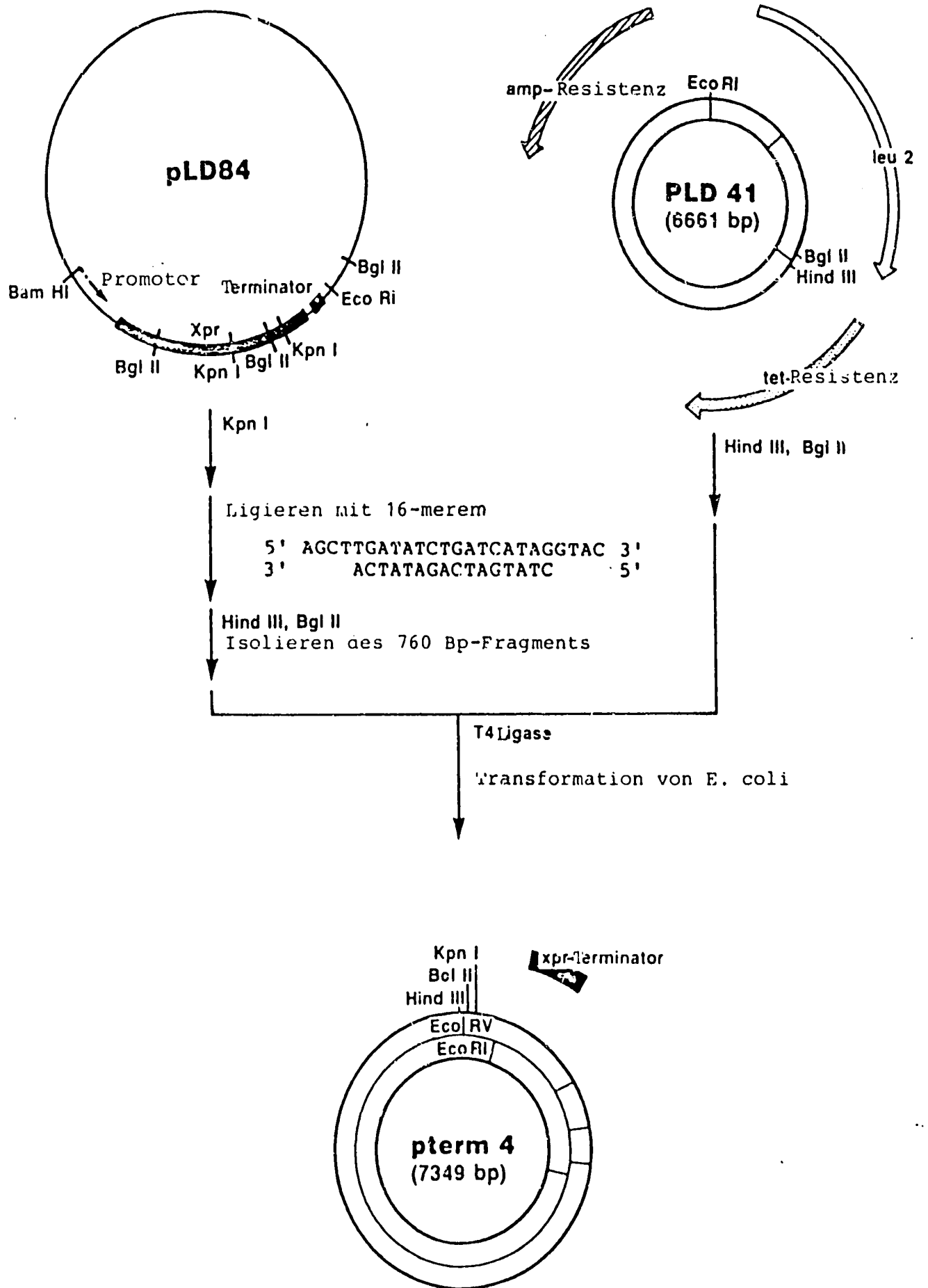
270
 ala gln val glu ala leu ile thr glu ser asn thr gly val leu
 GCC CAG GTC GAG GCT CTT ATT ACT GAG TCC AAC ACC GGT GTT CTT

280 290
 pro thr thr asn leu lys gly ser pro asn ala val ala tyr asn
 CCC ACC ACC AAC CTC AAG GGC TCT CCC AAC GCT GTT GCC TAC AAC

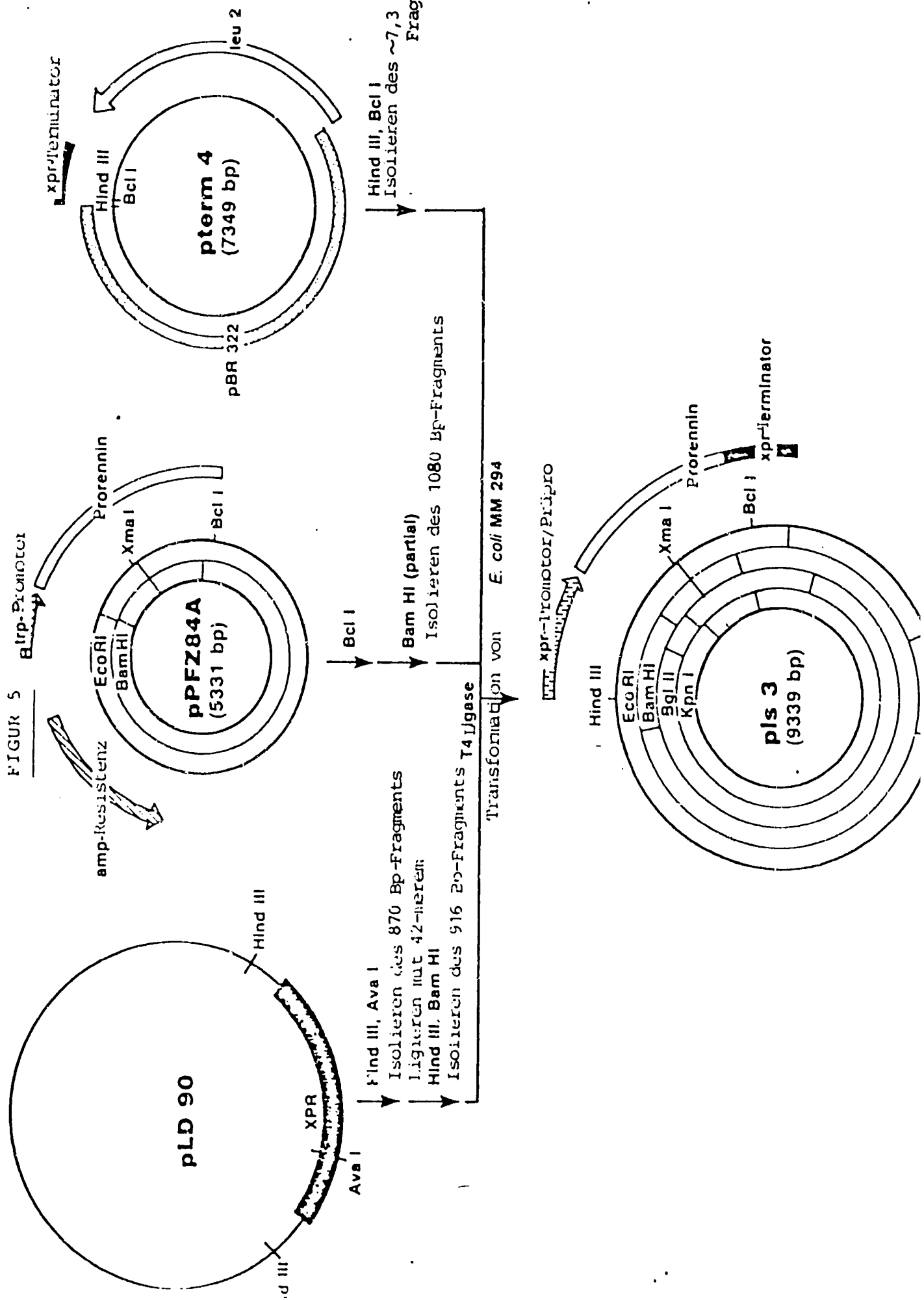
297
 gly val gly ile
 GGT GTT GGC ATT TAG GCAATTAACAGATAGTTTGCCGGTGATAATTCTCTTAACCTC

CCACACTCCTTTGACATAACGATTTATGTAACGAAACTGAAATTTGACCAGATATTGTTG
 TAAATAGAAAATCTGGCTTGTAGGTGGCAAAATCCCGTCTTTGTTTCGTCGGTTCCTCTG
 TGACTGCTCGTCGTCCCTTTGTGTTCCGACTGTCGTGTTTTGTTTTCCGTGCGTGCCGAAG
 TGAQATGCCCGTGTTTGAATTCGGTAGTCGCACGGACCATCGGTTGCTCTGCACACACAC
 ACACGGGAGGCTGGAACCTACATCAGAGCACTACTTGCAGGGTTGATGCAACATTCAAGA
 AAAGCGCAAGCAGTGGTGCATGTATAGCAGCTAACAGCAACTACTGCTCAACATGAAAAA
 GGAGGGTGTAAAGACGGCCAAGACTGCTTTCTGTCTACGCCCTGAGCAACGTGCTCTGCAA
 CAGAGCAACAGATAATCGCCTACGGAGACAGAGACAGAGACAGAAACAGAAACAAAAGCA
 ACAGAAACTGCTGTAGTGTGTTGCAGTGAGGCGGAGATTTAACCGTATAATTACCGCTCA
 GATCT

FIG. 4



FIGUR 5 Btrp-Promotor



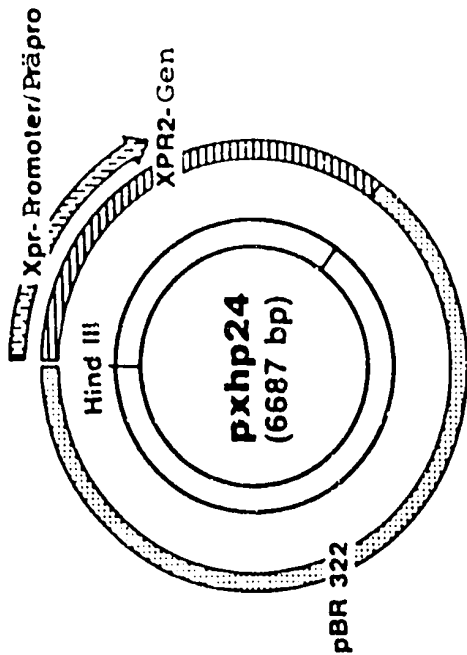
Hind III, Bcl I
 Isolieren des ~7,3 kB-
 Fragments

Bcl I
 Bam HI (partial)
 Isolieren des 1080 Bp-Fragments

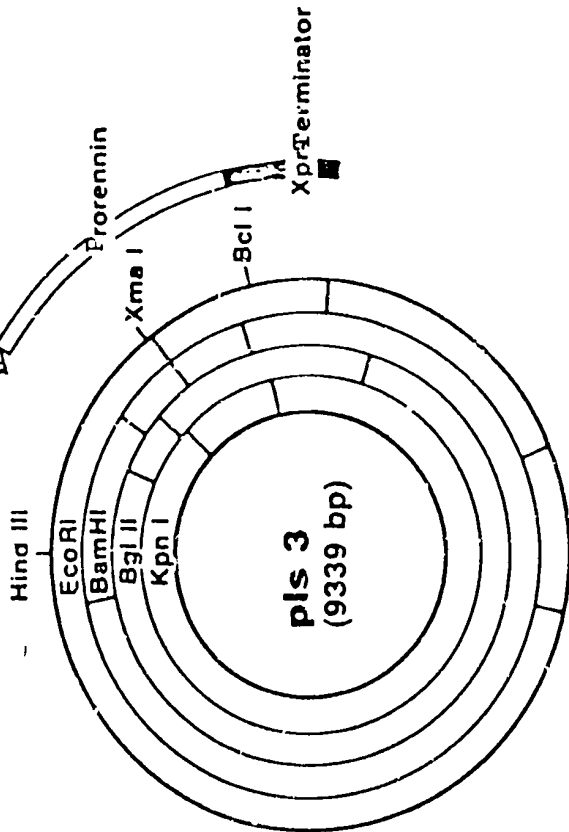
Hind III, Ava I
 Isolieren des 870 Bp-Fragments
 Ligation mit 42-meren
 Hind III, Bam HI
 Isolieren des 916 Bp-Fragments T4 Ligase

Transformation von E. coli MM 294

xpr-promotor/Prüpro



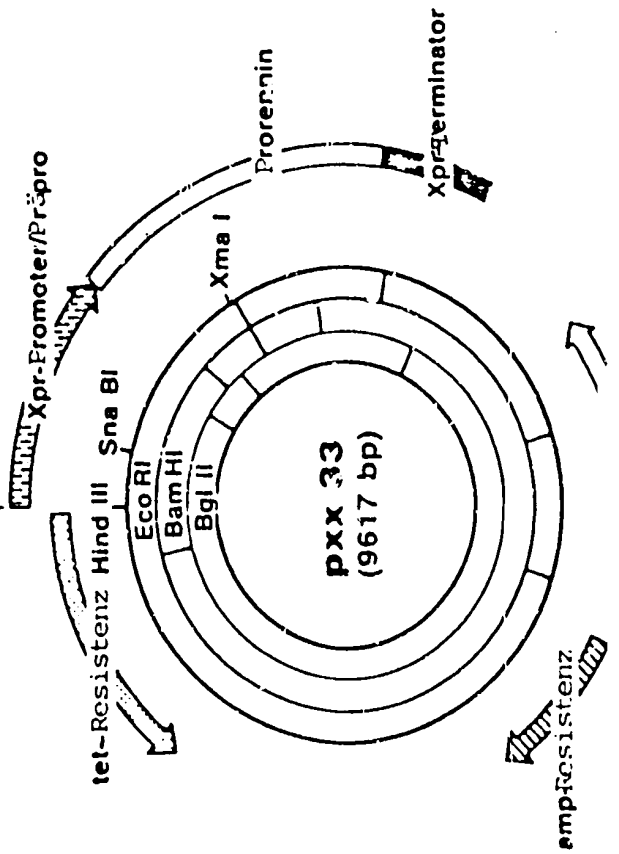
Hind III, Ava I
 Isolieren des 1150 Bp-Fragments
 Ligieren mit 42-merem
 Hind III, Bam HI
 Isolieren des 1196 Bp-Fragments



Hind III, Xma I
 Isolieren des ~8,0 kR-Fragments

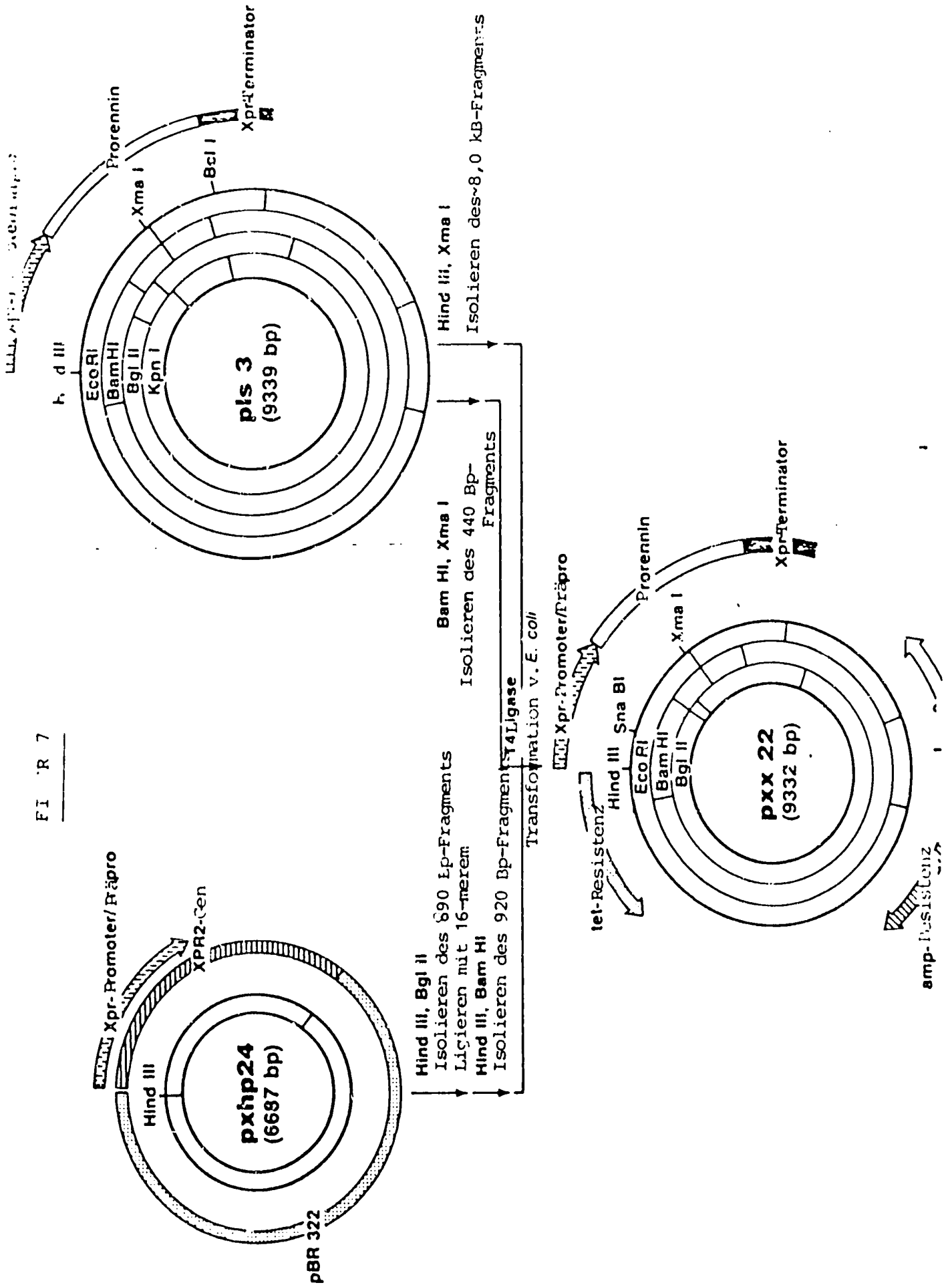
Bam HI, Xma I
 Isolieren des 440 Bp-Fragments

4 Ligase
 Transformation v. E. coli



emp-Resistenz

FIGURE 7



FIGUR 3

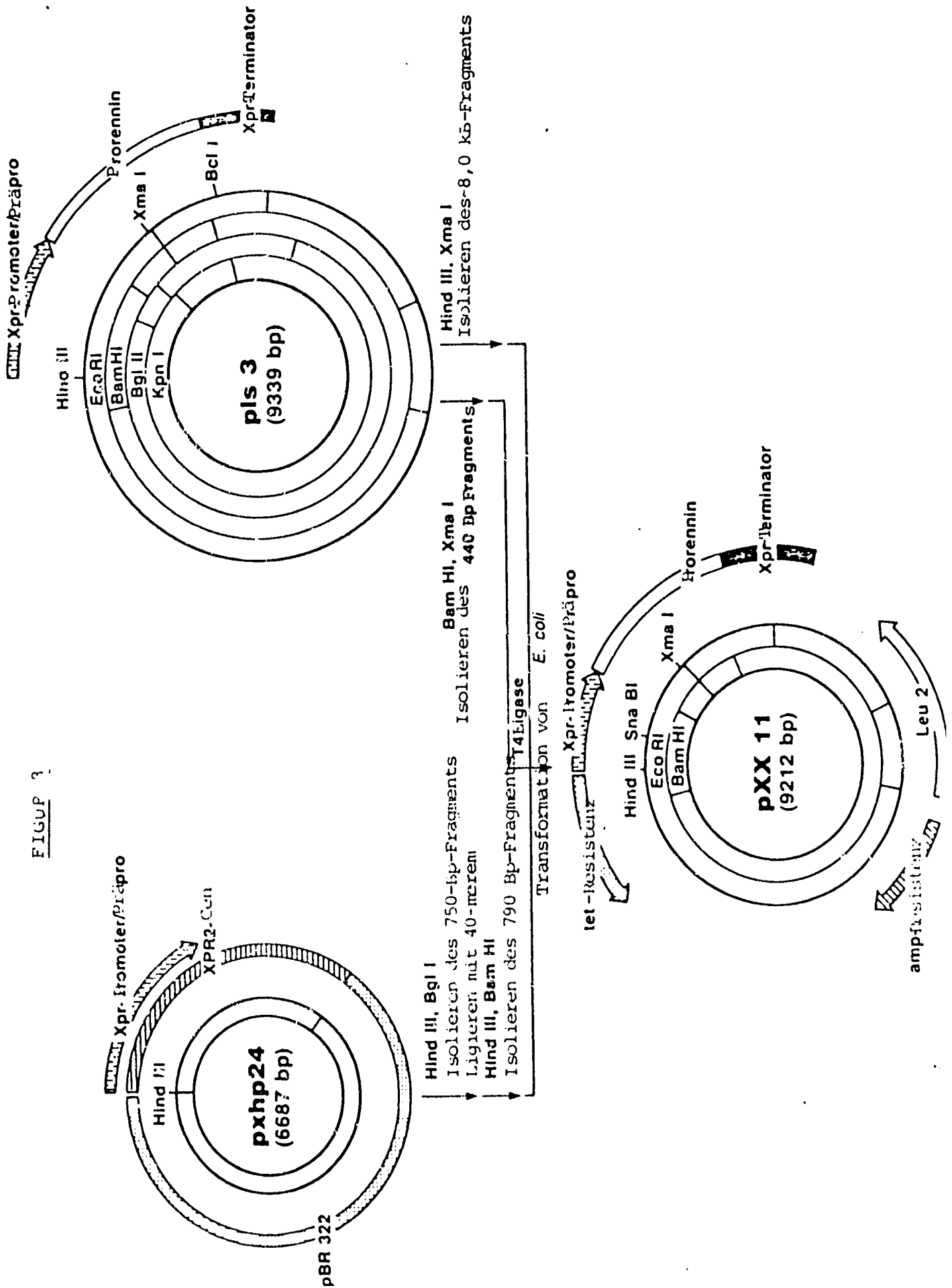


FIG. 9

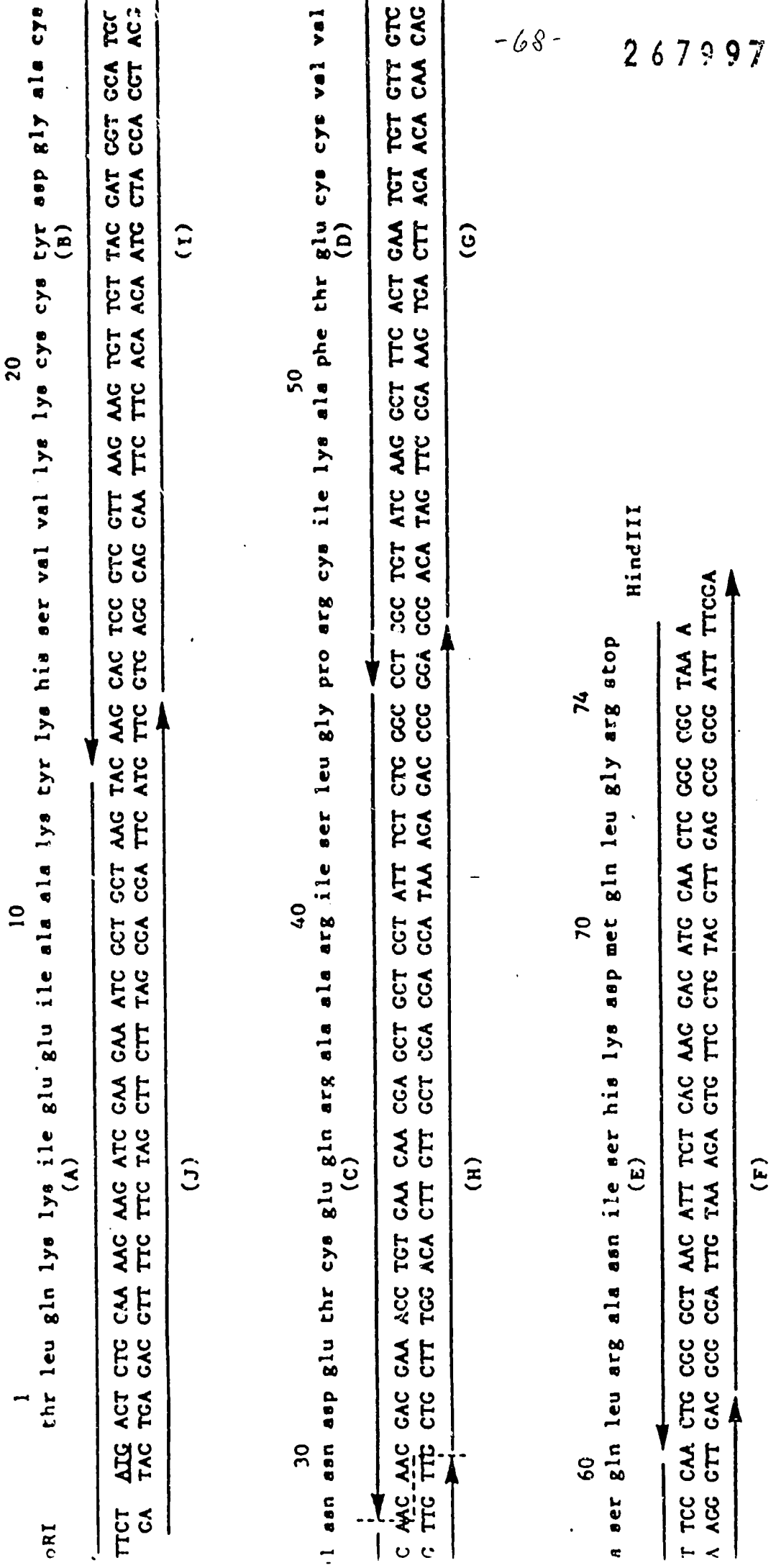
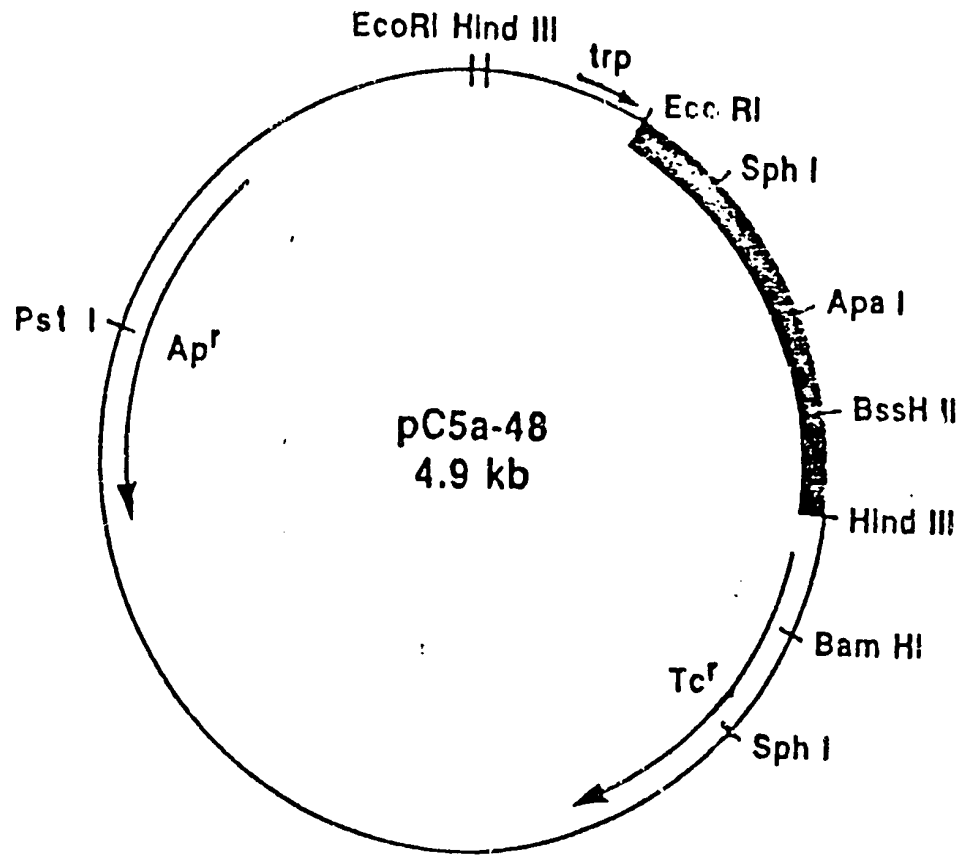


FIG. 10



FIGUR 11

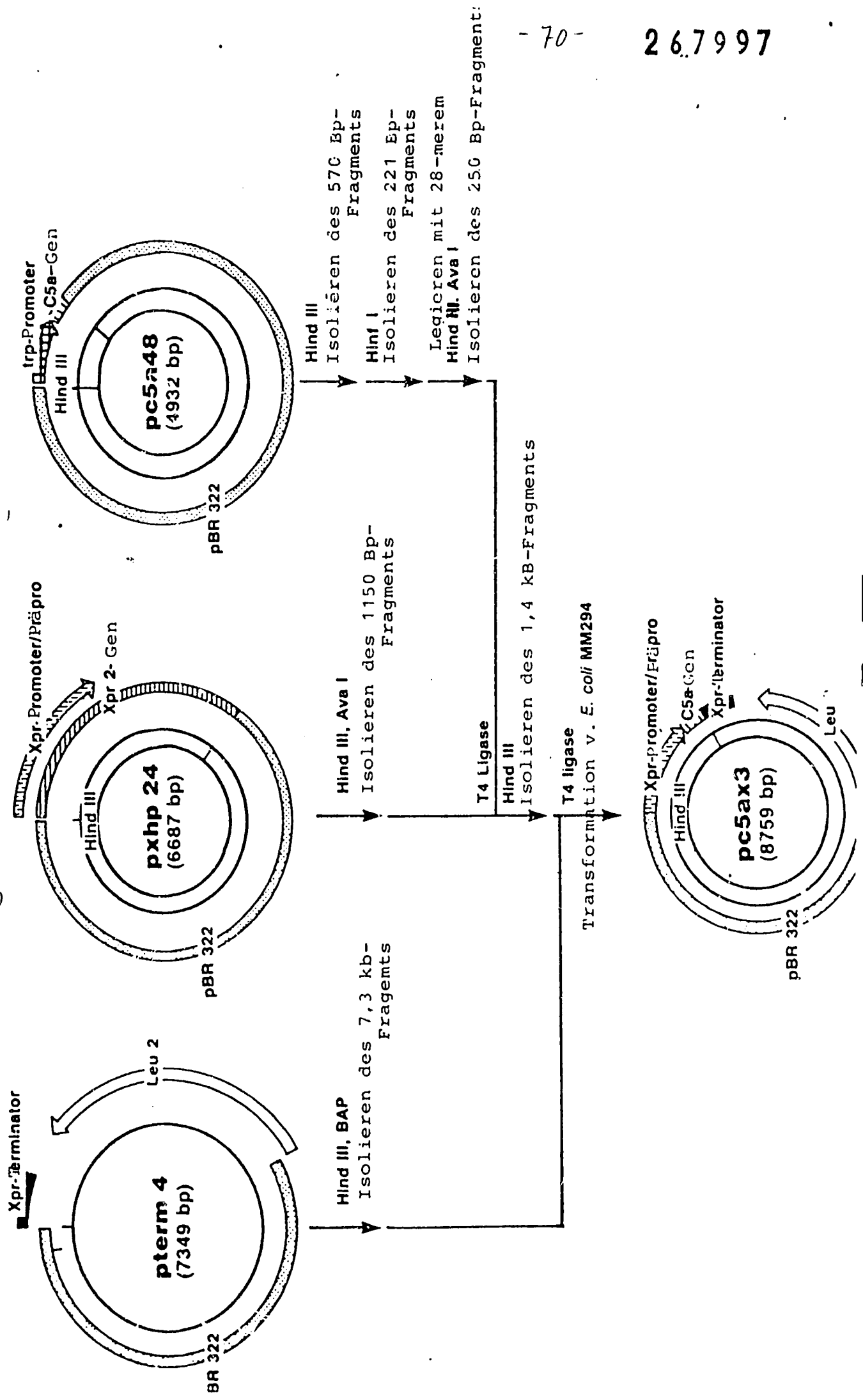


Figure 12

GTCGACACCATATCATATAAAACTAACAATGCATTGCTTATTACGAAGACTACCCGTTGCTATCTCCAC
 ACCGTTATCTCCACGGTCCAAAGGCTGCTCAATGTGCTGCATACGTAACTGGGGTGCAACCTTGAGCA
 CATAGTACTTTTCCGAAAACCGGCGATAATTAAGTGTGCACTCCAACCTTTTCACACTGAGCGTAAAATG
 TGGAGAAGAAATCSGCACTAAAAAGTCAGGTAGACTGGAAAATGCGCCATGAAATGAATATCTCTTGCT
 ACAGTAATGCCCAGCATCGAGGGGTATTGTGTACCAACACTATAGTGGCAGCTGAAGCGCTCGTGATT
 GTAGTATGAGTCTTTATTGGTGATGGGAAGAGTTCACTCAATATTCTCGTTACTGCCAAAACACCACGG
 TAATCGGCCAGACACCATGGATGTAGATCACCAAGCCTGTGAATGTTATTTCGAGCTAAAATGCACATGG
 TTGGTGAAAGGAGTAGTTGCTGTGCAATTCCGTCGTCGCCTGAGTCATCATTTATTTACCAGTTGGCCA
 CAAACCCCTGACGATCTCGTATGTCCCCTCCGACATACTCCCGGCCGGCTGGGGTACGTTTCGATAGCGC
 .ATCGGCATCGACAAGGTTTGGGTCCCTAGCCGATACCGCACTACCTGAGTCACAATCTTCGGAGGTTT
 AGTCTTCCACATAGCACGGGCAAAGTGCGTATATATACAAGAGCGTTTGCCAGCCACAGATTTTCACT

1. ----->
 met glu pro glu thr lys lys
 CCACACACCACATCACACATAACAACCACACACATCCACA ATG GAA CCC GAA ACT AAG AAG

10
 thr lys thr asp ser lys lys ile val leu leu gly gly asp phe
 ACC AAG ACT GAC TCC AAG AAG ATT GTT CTT CTC GGC GGC GAC TTC

20
 cys gly pro glu val ile ala glu ala val lys val leu lys ser
 TGT GGC CCC GAG GTG ATT GCC GAG GCC GTC AAG GTG CTC AAG TCT

30
 al ala glu ala ser gly thr glu phe val phe glu asp arg leu
 GTT GCT GAG GCC TCC GGC ACC GAG TTT GTG TTT GAG GAC CGA CTC

40
 ile gly gly ala ala ile glu lys glu gly glu pro ile thr asp
 ATT GGA GGA GCT GCC ATT GAG AAG GAG GGC GAG CCC ATC ACC GAC

50
 ala thr leu asp ile cys arg lys ala asp ser ile met leu gly
 GCT ACT CTC GAC ATC TGC CGA AAG GCT GAC TCT ATT ATG CTC GGT

60
 ala val gly gly ala ala asn thr val trp thr thr pro asp gly
 GCT GTC GGA GGC GCT GCC AAC ACC GTA TGG ACC ACT CCC GAC GGA

70
 arg thr asp val arg pro glu gln gly leu leu lys leu arg lys
 CGA ACC GAC GTG CGA CCC GAG CAG GGT CTC CTC AAG CTG CGA AAG

80
 asp leu asn leu tyr ala asn leu arg pro cys gln leu leu ser
 GAC CTG AAC CTG TAC GCC AAC CTG CGA CCC TGC CAG CTG CTG TCG

90
 asp leu asn leu tyr ala asn leu arg pro cys gln leu leu ser
 GAC CTG AAC CTG TAC GCC AAC CTG CGA CCC TGC CAG CTG CTG TCG

100
 asp leu asn leu tyr ala asn leu arg pro cys gln leu leu ser
 GAC CTG AAC CTG TAC GCC AAC CTG CGA CCC TGC CAG CTG CTG TCG

110
 asp leu asn leu tyr ala asn leu arg pro cys gln leu leu ser
 GAC CTG AAC CTG TAC GCC AAC CTG CGA CCC TGC CAG CTG CTG TCG

120
 asp leu asn leu tyr ala asn leu arg pro cys gln leu leu ser
 GAC CTG AAC CTG TAC GCC AAC CTG CGA CCC TGC CAG CTG CTG TCG

130 140
pro lys leu ala asp leu ser pro ile arg asn val glu gly thr
CCC AAG CTC GCC GAT CTC TCC CCC ATC CGA AAC GTT GAG GGC ACC

150
asp phe ile ile val arg glu leu val gly gly ile tyr phe gly
GAC TTC ATC ATT GTC CGA GAG CTC GTC GGA GGT ATC TAC TTT GGA

160 170
glu arg lys glu asp asp gly ser gly val ala ser asp thr glu
GAG CGA AAG GAG GAT GAC GGA TCT GGC GTC GCT TCC GAC ACC GAG

180
thr tyr ser val pro glu val glu arg ile ala arg met ala ala
ACC TAC TCC GTT CCT GAG GTT GAG CGA ATT GCC CGA ATG GCC GCC

190 200
phe leu ala leu gln his asn pro pro leu pro val trp ser leu
TTC CTG GCC CTT CAG CAC AAC CCC CCT CTT CCC GTG TGG TCT CTT

210
asp lys ala asn val leu ala ser ser arg leu trp arg lys thr
GAC AAG GCC AAC GTG CTG GCC TCC TCT CGA CTT TGG CGA AAG ACT

220 230
val thr arg val leu lys asp glu phe pro gln leu glu leu asn
GTC ACT CGA GTC CTC AAG GAC GAA TTC CCC CAG CTC GAG CTC AAC

240
his gln leu ile asp ser ala ala met ile leu ile lys gln pro
CAC CAG CTG ATC GAC TCG GCC GCC ATG ATC CTC ATC AAG CAG CCC

250 260
ser lys met asn gly ile ile ile thr thr asn met phe gly asp
TCC AAG ATG AAT GGT ATC ATC ATC ACC ACC AAC ATG TTT GGC GAT

270
ile ile ser asp glu ala ser val ile pro gly ser leu gly leu
ATC ATC TCC GAC GAG GCC TCC GTC ATC CCC GGT TCT CTC GGT CTG

280 290
leu pro ser ala ser leu ala ser leu pro asp thr asn glu ala
CTG CCC TCC GCC TCT CTG GCT TCT CTG CCC GAC ACC AAC GAG GCG

300
phe gly leu tyr glu pro cys his gly ser ala pro asp leu gly
TTC GGT CTG TAC GAG CCC TGT CAC GGA TCT GCC CCC GAT CTC GGC

310 320
lys gln lys val asn pro ile ala thr ile leu ser ala ala met
AAG CAG AAG GTC AAC CCC ATT GCC ACC ATT CTG TCT GCC GCC ATG

330
met leu lys phe ser leu asn met lys pro ala gly asp ala val
ATG CTC AAG TTC TCT CTT AAC ATG AAG CCC GCC GGT GAC GCT GTT

340 350
 glu ala ala val lys glu ser val glu ala gly ile thr thr ala
 GAG GCT GCC GTC AAG GAG TCC GTC GAG GCT GGT ATC ACT ACC GCC

360
 asp ile gly gly ser ser ser thr ser glu val gly asp leu leu
 GAT ATC GGA GGC TCT TCC TCC ACC TCC GAG GTC GGA GAC TTG TTG

370 380
 pro thr arg ser arg ser cys ser arg arg ser lys ser phe leu
 CA ACA AGG TCA AGG AGC TGC TCA AGA AGG AGT AAG TCG TTT CTA

390
 arg arg ile asp gly arg ser lys leu thr arg leu arg val gly
 CGA CGC ATT GAT GGA AGG AGC AAA CTG ACG CGC CTG CGG GTT GGT

400 405
 leu pro ala gly ser ala ser val OC
 CTA CCG GCA GGG TCC GCT AGT GTA TAA GACTCTATAAAAAGGGCCCTGCCCTGCTAATGAA

ATGATGATTTATAATTTACCGGTGTAGCAACCTTGACTAGAGAAAGCAGATTGGGTGTGTTTGTAGTGGA
 GGACAGTGGTACGTTTTGGAAACAGTCTTCTTGAAAGTGTCTTGTCTACAGTATATTCACTCATAACCTC
 AATAGCCAAGGGTGTAGTCGGTTTATTAAAGGAAGGGAGTTGTGGCTGATGTGGATAGATATCTTTAAGC
 TGGCGACTGCACCCAACGAGTGTGGTGGTAGCTTGTTACTGTATATTCGGTAAGATATATTTTGTGGGGT
 TAGTGGTGTGGTAGGTTAGTGCTTGGTATATGAGTTGTAGGCATGACAATTTGGAAAGGGGTGGAC
 TTTGGGAATATTGTGGGATTTCAATACCTTAGTTTGTACAGGGTAATTGTTACAAATGATACAAAGA
 ACTGTATTTCTTTTCATTTGTTTTAATTGGTTGTATATCAAGTCCGTTAGACGAGCTCAGTGCCATGGCTTTT
 GGCACTGTATTTCATTTTAGAGGTACACTACATTCAGTGAGGTATGGTAAGGTTGAGGGCATAATGAAG
 GCACCTTGACTGACAGTCACAGACCTCTCACCGAGAATTTTATGAGATATACTCGGGTTCATTTTAGGC
 TCCGATTCGATTCAAATTACTGTGCAAATCGGTTGAGCATCCGTTGATTTCCGAACAGATCTCGGCA
 GTCTCTCGGATGTAGAATTAGGTTTCCTTGAGGCGAAGATCGGTTTGTGTGACATGAATT

