



(12) Ausschließungspatent

Erteilt gemäß § 17 Absatz 1 Patentgesetz

(19) DD (11) 255 359 A5

4(51) C 12 Q 1/68

AMT FÜR ERFINDUNGS- UND PATENTWESEN

In der vom Anmelder eingereichten Fassung veröffentlicht

(21) AP C 12 Q / 300 376 0

(31) 655.841

675.298

(22) 19.06.85

(32) 28.09.84

30.11.84

(44) 30.03.88

(33) US

(71) siehe (73)

(72) Lin, Fu-Kuen, US

(73) Kirin-Amgen Inc., California 91320, US

(54) Verfahren zur Bestimmung eines spezifischen einzelsträngigen Polynucleotids unbekannter Sequenz

(55) Verfahren, Bestimmung, spezifisch, einzelsträngig, Polynucleotid, unbekannte Sequenz, heterogen, zellulär, viral, Probe

(57) Verfahren zur Bestimmung eines spezifischen einzelsträngigen Polynucleotids unbekannter Sequenz in einer heterogenen zellulären oder viralen Probe einschließlich multipler einzelsträngiger Polynucleotide, gekennzeichnet dadurch, daß

(a) ein Gemisch markierter einzelsträngiger Polynucleotidsonden mit gleichmäßig variierenden Sequenzen der Basen hergestellt wird, wobei jede der Sonden potentiell einer Sequenz von Basen komplementär ist, die zu dem zu bestimmenden Polynucleotid das vermeintlich nur einmal existierende Exemplar ist;

(b) die Probe auf ein festes Substrat fixiert wird;

(c) das Substrat mit der darauf fixierten Probe behandelt wird, um eine weitere Bindung von Polynucleotiden darauf abzuschwächen, ausgenommen durch Hybridisierung zu Polynucleotiden in der Probe;

(d) das behandelte Substrat mit der darauf fixierten Probe flüchtig mit dem genannten Gemisch markierter Sonden unter Bedingungen in Berührung gebracht wird, die nur die Hybridisierung zwischen vollständig komplementären Polynucleotiden erleichtern; und

(e) das spezifische Polynucleotid über die Beobachtung des Auftretens einer Hybridisierungsreaktion zwischen ihm und einer vollständig komplementären Sonde aus dem Gemisch der markierten Sonden bestimmt wird.

Patentanspruch:

Verfahren zur Bestimmung eines spezifischen einzelsträngigen Polynucleotids unbekannter Sequenz in einer heterogenen zellulären oder viralen Probe einschließlich multipler einzelsträngiger Polynucleotide, **gekennzeichnet dadurch**, daß :

- (a) ein Gemisch markierter einzelsträngiger Polynucleotidsonden mit gleichmäßig variierenden Sequenzen der Basen hergestellt wird, wobei jede der Sonden potentiell einer Sequenz von Basen komplementär ist, die zu dem zu bestimmenden Polynucleotid das vermeintlich nur einmal existierende Exemplar ist;
- (b) die Probe auf ein festes Substrat fixiert wird;
- (c) das Substrat mit der darauf fixierten Probe behandelt wird, um eine weitere Bindung von Polynucleotiden darauf abzuschwächen, ausgenommen durch Hybridisierung zu Polynucleotiden in der Probe;
- (d) das behandelte Substrat mit der darauf fixierten Probe flüchtig mit dem genannten Gemisch markierter Sonden unter Bedingungen in Berührung gebracht wird, die nur die Hybridisierung zwischen vollständig komplementären Polynucleotiden erleichtern; und
- (e) das spezifische Polynucleotid über die Beobachtung des Auftretens einer Hybridisierungsreaktion zwischen ihm und einer vollständig komplementären Sonde aus dem Gemisch der markierten Sonden bestimmt wird, was durch das Auftreten einer höheren Dichte des markierten Materials auf dem Substrat am Ort des spezifischen Polynucleotids im Vergleich zu einer Hintergrunddichte an markiertem Material, die sich aus dem nichtspezifischen Binden markierter Sonden an das Substrat ergibt, angezeigt wird, wobei die Verbesserung darin besteht, daß die Verwendung von 32 Mischsonden im Überschuß erfolgt und eines oder mehrere der folgenden Merkmale auftreten.
 - (1) Einsatz eines Papiers auf Nylonbasis als festes Substrat;
 - (2) Behandlung mit einer Protease in Stufe (c);
 - (3) Einsatz individuell markierter Sondenkonzentrationen von annähernd 0,025 Picomol; und
 - (4) Einsatz als einer der Hybridisierungsbedingungen in Stufe (d) strenge Temperaturen, die nahe bei 4°C von der niedrigsten berechneten Td irgendeiner der verwendeten Sonden liegen.

Hierzu 4 Seiten Zeichnungen

Anwendungsgebiet der Erfindung

Die Erfindung betrifft ein Verfahren zur Herstellung eines Polypeptids. Allgemein bezieht sich die Erfindung auf die Manipulation genetischer Materialien, insbesondere auf rekombinante Verfahren, die die Herstellung von Polypeptiden ermöglichen, die teilweise oder insgesamt die primäre strukturelle Anordnung aufweisen und/oder eine oder mehrere der biologischen Eigenschaften natürlich vorkommenden Erythropoietins.

Charakteristik der bekannten technischen Lösungen

A. Manipulation von genetischen Materialien

Genetische Materialien können weitgehend als jene chemischen Substanzen definiert werden, die für die Herstellung von Zellbestandteilen und Viren programmieren und diese Vorgänge lenken und die Reaktion von Zellen und Viren ausrichten. Eine langkettige polymere Substanz, die als Desoxyribonucleinsäure (DNA) bekannt ist, enthält das genetische Material aller lebenden Zellen und Viren, mit Ausnahme bestimmter Viren, die durch die Ribonucleinsäuren (RNA) programmiert werden.

Die wiederkehrenden Einheiten in DNA-Polymeren sind vier unterschiedliche Nucleotide, wobei jedes von ihnen entweder aus einem Purin (Adenin oder Guanin) oder einem Pyrimidin (Thymin oder Cytosin) besteht, gefunden an einen Desoxyribosezucker, an den eine Phosphatgruppe angefügt ist. Die Anfügung von Nucleotiden in linearer Form erfolgt über die Fusion 5'-Phosphates eines Nucleotids an die 3'-Hydroxy-Gruppe eines anderen. Funktionelle DNA tritt in Form stabiler doppelsträngiger Zusammenlagerungen einzelsträngiger Nucleotide auf (bekannt als Desoxyoligonucleotide), wobei sich diese Zusammenlagerungen durch Wasserstoffbindung zwischen Purin- und Pyrimidinbasen ergeben (z. B. „komplementäre“ Zusammenlagerungen, die entweder zwischen Adenin [A] und Thymin [T] oder Guanin [G] und Cytosin [C] bestehen). Durch Übereinkommen werden die Nucleotide durch die Namen ihrer konstitutionellen Purin- oder Pyrimidinbasen bezeichnet, und die komplementären Zusammenlagerungen der Nucleotide in der doppelsträngigen DNA (z. B. -A-T und G-C) werden als „Basenpaare“ bezeichnet. Ribonucleinsäure ist ein Polynucleotid, bestehend aus Adenin, Guanin, Cytosin und Uracil (U) anstelle von Thymin, gebunden an Ribose und eine Phosphatgruppe.

Zusammengefaßt wird die programmierende Funktion der DNA im allgemeinen durch einen Prozeß bewirkt, bei dem spezifische DNA-Nucleotidsequenzen (Gene) in relativ instabile Boten-RNA (mRNA)-Polymere „transkribiert“ werden. Die mRNA dient umgekehrt als eine Schablone für die Bildung struktureller, regulierender und katalytischer Proteine aus Aminosäuren. Zu diesem mRNA-„Translations“prozeß gehören die Operationen kleiner RNA-Stränge (tRNA), die einzelne Aminosäuren am mRNA-Strang entlang transportieren und ausrichten, um die Bildung von Polypeptiden in einwandfreier Aminosäurefolge zu ermöglichen. Die mRNA-„Botschaft“, abgeleitet aus der DNA und die Basis für die tRNA-Unterstützung bildend und die Orientierung jeder gegebenen der zwanzig Aminosäuren für die Polypeptid-„Expression“, liegt in Form von Dreifach-„Lodons“ vor — aufeinanderfolgende Gruppierungen der drei Nucleotidbasen. In diesem Sinne ist die Bildung eines Proteins die letzte Form der „Expression“ der programmierten genetischen Botschaft, die in der Nucleotidsequenz eines Gens vorliegt. „Promotor“-DNA-Sequenzen gehen üblicherweise einem Gen in einem DNA-Polymeren, „voraus“ und sehen eine Stelle für die Initiierung der Transskription in mRNA vor. „Regulator“-DNA-Sequenzen, auch üblicherweise oberhalb (upstream) des (z. B. vorausgehenden) Gens in ein einem gegebenen DNA-Polymeren, binden Proteine, die die Frequenz (oder Rate) der Transkriptionsinitiierung bestimmen.

Gemeinsam als „Promotor/Regulator“- oder „Steuerungs“-DNA-Sequenz bezeichnet, arbeiten diese Sequenzen, die einem ausgewählten Gen (oder eine Reihe von Genen) in einem funktionellen DNA-Polymeren vorausgehen, zusammen, um zu bestimmen, ob die Transskription (und eventuell Expression) eines Gens stattfindet. DNA-Sequenzen, die einem Gen in einem DNA-Polymeren „folgen“ und ein Signal für die Beendigung der Transkription in mRNA vorsehen, werden als Transkriptions„terminator“-Sequenzen bezeichnet.

Brennpunkt mikrobiologischer Verfahren war im letzten Jahrzehnt der Versuch zur Herstellung industriell und pharmazeutisch bedeutsamer Substanzen unter Verwendung von Organismen, die entweder anfänglich keine genetisch codierte Information hinsichtlich des gewünschten Produktes in ihrer DNA enthielten oder (im Falle menschlicher Zellen in Kultur) gewöhnlich kein chromosomales Gen mit entsprechenden Gehalten exprimieren. Vereinfacht dargestellt, wird ein Gen, das die Struktur eines gewünschten Polypeptidproduktes spezifiziert, entweder aus einem „Spender“ („Donor“-)organismus isoliert oder chemisch synthetisiert und dann stabil in einen anderen Organismus eingeführt, der vorzugsweise ein selbstreplizierender, vielzelliger Organismus wie Bakterien-, Hefe- oder Säugetierzellen in Kultur ist. Wenn dies einmal erfolgt ist, wird der vorhandene Mechanismus für die Genexpression in den „transformierten“ oder „transfektierten“ mikrobiellen Wirtszellen dazu benutzt, das gewünschte Produkt zu konstruieren, unter Verwendung der exogenen DNA als eine Schablone für die Transkription von mRNA, die dann in eine kontinuierliche Sequenz von Aminosäureresten translatiert wird.

Aus dem Stand der Technik gibt eine zahlreiche Patent- und Literaturpublikation zu „rekombinanter DNA“-Technik hinsichtlich Isolierung, Synthese, Reinigung und Amplifizierung genetischer Materialien zur Verwendung bei der Transformation ausgewählter Wirtsorganismen. So betrifft beispielsweise die US-PS 4237 224 (Cohen et al.) die Transformation einzelliger Wirtsorganismen mit „Hybrid“-viralen oder ringförmiger Plasmid-DNA, wozu ausgewählte exogene DNA-Sequenzen gehören. Zu den Verfahrensweisen des Cohen-Patents gehört die Herstellung eines Transformationsvektors durch enzymatisches Schneiden viralen oder ringförmiger Plasmid-DNA, um lineare DNA-Stränge zu bilden. Ausgewählte fremde („exogene“ oder „heterologe“) DNA-Stränge, zu denen üblicherweise Sequenzen, die für das gewünschte Produkt codieren, gehören, werden in linearer Form durch die Verwendung ähnlicher Enzyme hergestellt. Die lineare virale oder Plasmid-DNA wird mit der fremden DNA in Anwesenheit ligierender Enzyme inkubiert, die in der Lage sind, einen Wiederherstellungsprozeß zu bewirken, und es werden „Hybrid“-Vektoren gebildet, zu denen das ausgewählte exogene DNA-Segment gehört, „gespleißt“ in das virale oder kreisförmige DNA-Plasmid.

Die Transformation von kompatiblen einzelligen Wirtsorganismen mit dem Hybridvektor führt zur Bildung von vervielfachten Kopien der exogenen DNA in der Wirtszellenpopulation. In vielen Fällen ist das erwünschte Ergebnis einfach die Amplifizierung der Fremd-DNA, und das gewonnene „Produkt“ ist die DNA. Kürzlich wurde darüber berichtet, daß das Endziel der Transformation die Expression durch die Wirtszellen der exogenen DNA in Form einer Synthese im Großmaßstab von isolierbaren Mengen kommerziell bedeutender, dafür codierender Protein- oder Polypeptidfragmente durch die Fremd-DNA ist, siehe auch z. B. US-PS 4264 731 (Shine), 4273 875 (Manis), 4293 652 (Cohen) und EP-OS 093619, veröffentlicht am 9. November 1983.

Die Entwicklung spezifischer DNA-Sequenzen für das Spleißen in DNA-Vektoren erreicht man über eine Vielzahl von Techniken, abhängig im großen Maße vom Grad der „Fremdheit“ des „Donors“ zum vorgesehenen Wirt und der Größe des im Wirt zu exprimierenden Polypeptids. Zur Gefahr der Über-Vereinfachung kann festgestellt werden, daß drei alternative prinzipielle Methoden eingesetzt werden können: (1) die „Isolierung“ der doppelsträngigen DNA-Sequenz aus der genomischen DNA des Donors; (2) die chemische Herstellung einer DNA-Sequenz, die einen Code für ein Polypeptid von Interesse aufweist; und (3) die in vitro-Synthese einer doppelsträngigen DNA-Sequenz durch enzymatische „reverse Transskription“ von mRNA, die aus Donorzellen isoliert worden ist. Die letztgenannten Methoden, zu denen die Bildung eines DNA-„Komplements“ von mRNA gehört, werden allgemein als „cDNA“-Methoden bezeichnet.

Die Herstellung von DNA-Sequenzen wurde kürzlich als Methode der Wahl bezeichnet, wenn die gesamte Sequenz der Aminosäurereste des gewünschten Polypeptidproduktes bekannt ist. DNA-Herstellungsverfahren von der ebenfalls zugeeigneten, anhängigen US-Patentanmeldung Nr. 483481 von Alton et al. (eingereicht am 15. April 1983 und korrespondierend zu PCT US 83/00605, veröffentlicht am 24. November 1983 als WO 83/04053 sehen, beispielsweise weiterführende Mittel vor zur Erreichung solcher höchst wünschenswerter Ergebnisse wie: Vorsehen von alternierenden Codons, die kürzlich in Genen gefunden wurden, die in hohem Maße in den für die Expression ausgewählten Wirtsorganismen exprimiert werden (z. B. Vorsehen von Hefe- oder Ecoli-„Vorzugs“-codons! Vermeiden der Anwesenheit untranslatierter „Intron“-Sequenzen (üblicherweise in menschlichen genomischen DNA-Sequenzen und mRNA-Transskripten davon), die nicht leicht durch procaryotische Wirtszellen verarbeitet werden; Vermeiden der Expression unerwünschter „leader“-Polypeptidsequenzen, die üblicherweise dafür codieren, durch genomische DNA- und cDNA-Sequenzen, jedoch letztere nicht leicht von dem interessierenden Polypeptid durch bakterielle oder Hefe-Wirtszellen abgespalten werden; Vorsehen einer leichten Insertion der DNA in geeignete Expressionsvektoren in Zusammenlagerung mit gewünschten Promoter/Regulator- und Terminator-Sequenzen; und Vorsehen einer leichten Konstruktion von Genen, die für Polypeptidfragmente codieren und Analoge der gewünschten Polypeptide.

Wenn die gesamte Sequenz der Aminosäurereste der gewünschten Polypeptide nicht bekannt ist, ist die direkte Herstellung der DNA-Sequenzen nicht möglich, und die Isolierung von DNA-Sequenzen, die für das Polypeptid codieren, durch eine cDNA-Methode wird die Methode der Wahl, ungeachtet der potentiellen Rückgänge im Spielraum der Gruppe von Expressionsvektoren, die in der Lage sind, einen hohen Grad der mikrobiellen Expression, über die oben berichtet wurde, zu verwirklichen. Zu den Standardverfahren zur Isolierung von interessierenden cDNA-Sequenzen gehört die Bereitung von Plasmid-geborenen cDNA-„Bibliotheken“, abgeleitet von reverser Transskription von mRNA, reich an Donorzellen, ausgewählt als verantwortlich für einen hohen Grad an Expression von Genen (z. B. Bibliotheken von cDNA, abgeleitet von Hypophysenzellen, die relativ große Mengen an Wachstumshormonprodukten exprimieren).

Wo wesentliche Mengen der Aminosäuresequenzen des Polypeptids bekannt sind, können markierte, einzelsträngige DNA-Sondierungssequenzen eingesetzt werden, die eine vermutlich in der „Target“-cDNA vorhandene Sequenz duplizieren, in DNA/DNA-Hybridisierungsverfahren eingesetzt werden, die auf klonierten Kopien der cDNA durchgeführt werden, die in eine einzelsträngige Form abgebaut worden waren (siehe allgemein die Offenbarung und die Diskussion des Standes der Technik in der US-PS 4394443 von Weissman et al. und die kürzlich erfolgte Demonstration der Verwendung von langen Oligonucleotidhybridisierungssonden, die von Wallace et al. in *Nuc. Acids Res.* **6**, S. 3543–3557 [1979] und Reyes et al. in *P.N.A.S. (USA)*, **79**, S. 3270–3274 [1982] und Jaye et al. in *Nuc. Acids Res.* **11**, S. 2325–2335 [1983] berichtet wurden. Siehe auch US-PA 4358535 [Falkow et al.], die sich auf die DNA/DNA-Hybridisierungsverfahren zur Bewirkung der Diagnose bezieht; veröffentlichte Europäische Patentanmeldungen Nr. 0070685 und 0070687, die die leicht-emittierenden Markierungen an einzelsträngigen Polynucleotidsonden betreffen; Davis et al. „A Manual for Genetic Engineering, Advanced Bacterial Genetics“ [Handbuch der Gentechnik, Fortschritte bakterieller genetischer Methoden], Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y. [1980], Seiten 55–58 und 174–176, betreffend die Kolonie- und Plaque-Hybridisierungstechniken; sowie New England Nuclear [Boston, Mass.] Broschüren für „Gene Screen“ Hybridisierungs-Transfermembranmaterialien mit Anweisungen für den Transfer und die Hybridisierung von DNA und RNA, Katalog Nr. NEF-972).

Zu den bedeutendsten Fortschritten der letzten Zeit bei Hybridisierungsverfahren für das Screening rekombinanter Klone gehört die Verwendung markierter gemischter synthetischer Oligonucleotidsonden, von denen jede potentiell das vollständige Gegenstück einer spezifischen DNA-Sequenz in der Hybridisierungsprobe einschließlich eines heterogenen Gemisches einzelsträngiger DNA's oder RNA's ist. Diese Verfahrensweisen wurden als besonders nützlich bei der Feststellung von cDNA-Klonen bestätigt, die aus solchen Quellen stammten, die extrem niedrige Mengen an mRNA-Sequenzen für die interessierenden Polypeptide zulassen. Zusammengefaßt, die Verwendung von scharfen Hybridisierungsbedingungen, die auf das Vermeiden nichtspezifischer Bindungen gerichtet ist, kann z. B. für die autoradiografische Sichtbarmachung eines spezifischen cDNA-Klons für den Fall der Hybridisierung der Target-DNA zu dieser einzelnen Sonde innerhalb des Gemisches, das deren vollständiges Gegenstück ist, gestatten. Siehe dazu allgemein Wallace et al., *Nucl. Acids Res.* **9**, S. 879–897 (1981); Suggs et al., *P.N.A.S. (USA)*, **78**, S. 6613–6617 (1981); Choo et al., *Nature*, **299**, S. 178–180 (1982); Kurachi et al., *P.N.A.S. (USA)*, **79**, S. 6461–6464 (1982); Ohkubo et al., *P.N.A.S. (USA)*, **80**, S. 2196–2200 (1983); und Kornblihtt et al., *P.N.A.S. (USA)*, **80**, S. 3218–3222 (1983). Allgemein wurden die gemischten Sondenverfahren von Wallace et al. (1981), s. o. von verschiedenen Forschern erweitert bis zu dem Punkt, wo zuverlässige Ergebnisse vermutlich erhalten wurden in einer cDNA-Klonisolierung unter Einsatz eines 32teiligen gemischten „Pools“ von 16-Basen-langen (16-mer) Oligonucleotidsonden von gleichmäßigen, verschiedenen DNA-Sequenzen zusammen mit einem einzelnen 11-mer, um eine Zwei-Stellen-„positive“-Bestätigung der Anwesenheit von interessierender cDNA zu bewirken. Siehe dazu Singer-Sam et al., *P.N.A.S. (USA)*, **80**, S. 802–806 (1983).

Die Verwendung von genomischen DNA-Isolaten ist die letzte übliche der drei obengenannten Methoden zur Entwicklung spezieller DNA-Sequenzen für den Einsatz in rekombinanten Verfahren. Dies ist insbesondere richtig auf dem Gebiet der rekombinanten Verfahren, die auf die Sicherung der mikrobiellen Expression von Säugetierpolypeptiden gerichtet ist und wird prinzipiell der Komplexität Säugetier-genomischer DNA gerecht. Während somit zuverlässige Verfahrensweisen für die Entwicklung phagegeborener Bibliotheken genomischer DNA des Menschen und anderer Säugetierspezies entstehen (siehe z. B. Lawn et al., *Cell* **15**, S. 1157–1174 [1978] hinsichtlich von Verfahren zur Erzeugung einer Human-Genbibliothek, bezeichnet als „Maniatis-Bibliothek“; Karn et al., *P.N.A.S. (USA)*, **77**, S. 5172–5176 [1980] hinsichtlich einer Human-Genbibliothek, basierend auf alternativem Restriktionsendonuclease-Fragmentierungsverfahren; und Blattner et al. *Science*, **196**, S. 161–169 [1977], wo die Konstruktion einer Rinder-Genbibliothek beschrieben wird), gibt es relativ wenig erfolgreiche Versuche, bei der Verwendung von Hybridisierungsverfahren zur Isolierung genomischer DNA beim Fehlen umfassender Vorkenntnisse zu Aminosäure- oder DNA-Sequenzen.

Als ein Beispiel berichten Fiddes et al., *J. Mol. and Appl. Genetics*, **1**, S. 3–18 [1981] über die erfolgreiche Isolierung eines Gens, das für die alpha-Untereinheit des menschlichen Hypophysenglycoproteinormons aus der Maniatis-Bibliothek codiert, durch Verwendung einer Sonde mit „voller Länge“, eines vollständigen 621-Basenpaar-Fragments einer vorher isolierten cDNA-Sequenz für die alpha-Untereinheit.

Als ein weiteres Beispiel berichten das et al., *P.N.A.S. (USA)*, **80**, S. 1531–1535 (1983) über die Isolierung von menschlichen genomischen Klonen für Human-HLA-DR unter Verwendung eines synthetischen Oligonucleotids mit 175 Basenpaaren. Schließlich berichten Anderson et al., *P.N.A.S. (USA)*, **80**, S. 6838–6842 (1983) über die Isolierung eines genomischen Klons für einen Rinderpankreatischen Trypsininhibitor (BPTI) unter Verwendung einer Einzelsonde mit 86 Basenpaaren in der Länge und konstruiert entsprechend der bekannten Aminosäuresequenz des BPTI. Die Autoren beobachteten eine Determinierung geringer Aussichten für die Isolierung von mRNA, die geeignet ist für die Synthese einer cDNA-Bibliothek, wegen scheinbar geringer Gehalte an mRNA in anfänglich „targetiertem“ Ohrspeicheldrüsen- und Lungengewebsmaterial und lenken dann die Erfolgsaussichten auf die Sondierung einer Gen-Bibliothek unter Verwendung eines Gemisches markierter Sonden mit der Feststellung:

- „Allgemeiner, es wurden Mischsequenz-Oligodesoxynucleotidsonden verwendet, um Proteingene unbekannter Sequenz aus cDNA-Bibliotheken zu isolieren. Derartige Sonden sind normalerweise Gemische von 8–32 Oligonucleotiden, 14–17 Nucleiden in der Länge, die jede mögliche Codonkombination für eine geringe Streckung (5–6 Reste) der Aminosäuresequenz repräsentieren. Unter strengen Hybridisierungsbedingungen, die nichtkorrekte Basenpaarsonden benachteiligen, sind diese Sonden in der Lage, den Platz spezifischer Gensequenzen in Klonbibliotheken nieder — bis mittelmäßiger Komplexität festzustellen. Dessen ungeachtet weisen Mischsonden wegen ihrer kurzen Länge und Heterogenität oft nicht die Spezifität auf, die für sondierende Sequenzen als Komplex als ein Säugetier-Genom erforderlich

ist. Dies macht ein derartiges Verfahren ungeeignet zur Isolierung von Säugetier-Proteingenen, wenn die entsprechenden mRNA's nicht erhältlich sind."

(Ende des Zitats).

Somit besteht weiterhin ein Bedürfnis, für verbesserte Methoden, die eine schnelle und wirksame Isolierung von cDNA-Klonen in Fällen bewirken, wo nur wenig von der Aminosäuresequenz des dafür codierenden Polypeptids bekannt ist und wo „angereichertes“ Gewebematerial von mRNA nicht leicht für den Einsatz bei der Konstruktion von cDNA-Bibliotheken erhältlich ist. Derartige verbesserte Methoden würden insbesondere dann nützlich sein, wenn sie zur Isolierung von Säugetier-genomischen Klonen einsetzbar wären, wo nur zerstreute Informationen erhältlich sind hinsichtlich der Aminosäuresequenzen der dafür codierenden Polypeptide durch die gesuchten Gene.

B. Erythropoietin als ein Polypeptid von Interesse

Erythropoese, die Produktion roter Blutzellen, läuft kontinuierlich während des menschlichen Lebens ab, um den Zellabbau auszugleichen. Die Erythrozytenbildung ist ein sehr genau gesteuerter physiologischer Mechanismus, die es ermöglicht, daß eine ausreichende Anzahl roter Blutzellen im Blut zur Versorgung des Gewebes mit Sauerstoff verfügbar ist, jedoch nicht so viele, daß die Zellen die Zirkulation behindern würden. Die Bildung von roten Blutzellen erfolgt im Rückenmark und wird von dem Hormon Erythropoietin gesteuert.

Erythropoietin, ein saures Glycoprotein von ungefähr 34000 Dalton Molekülmasse, kann in drei Formen auftreten: α , β und asialo. Die α - und β -Formen unterscheiden sich geringfügig in den Carbohydratkomponenten, haben aber die gleiche Wirksamkeit, biologische Aktivität und Molekülmasse. Die asialo-Form ist eine α - oder β -Form bei der das terminale Carbohydrat (Sialsäure) entfernt ist. Erythropoietin ist in sehr geringen Konzentrationen im Plasma vorhanden, wenn sich der Körper in einem gesunden Zustand befindet, wobei das Gewebe ausreichend Sauerstoff über die vorhandene Anzahl Erythrozyten hält. Diese normale geringe Konzentration reicht aus, den Ersatz der roten Blutzellen zu stimulieren, die normalerweise durch Alterung ausfallen.

Die Menge an Erythropoietin im Kreislauf steigt unter den Bedingungen von Gewebesauerstoffmangel (Hypoxie), wenn der Sauerstofftransport durch die Blutzellen im Blutkreislauf verringert ist. Hypoxie kann hervorgerufen werden durch Verlust großer Mengen an Blut infolge Hämorrhagie, Zerstörung roter Blutzellen durch Überdosis an Strahlung, Verminderung der Sauerstoffaufnahme in großen Höhen oder durch längere Bewußtlosigkeit oder durch verschiedene Formen der Anämie. Als Erwiderung auf den das Gewebe erleidenden hypoxischen Stress erhöht das Erythropoietin die Produktion roter Blutzellen durch Stimulierung der Umwandlung einfacher Precursor-Zellen im Rückenmark in Proerythroblasten, die später reifen, Hämoglobin synthetisieren und in den Blutkreislauf als rote Blutzellen freigesetzt werden. Wenn die Anzahl der roten Blutzellen im Kreislauf höher als für die normalen Sauerstoffanfordernisse des Gewebes benötigt ist, wird das Erythropoietin im Kreislauf verringert. Siehe dazu allgemein, Testa et al., *Exp. Hematol.*, **8** (Suppl. 8), 144–152 (1980); Tong et al., *J. Biol. Chem.*, **256** (24), 12666–12672 (1981); Goldwasser, *J. Cell. Physiol.*, **110** (Suppl. 1), 133–135 (1982); Finch, *Blood*, **60** (6), 1241–1246 (1982); Sytowski et al., *Expt. Hematol.*, **8** (Suppl. 8), 52–64 (1980); Naughton, *Ann. Clin. Lab. Sci.*, **13** (5), 432–438 (1983); Weiss et al., *Am. J. vet. Res.*, **44** (10), 1832–1835 (1983); Lappin et al., *Esp. Hematol.*, **11** (7), 661–666 (1983); Bacin et al., *Ann. N. Y. Acad. Sci.*, **414**, 66–72 (1983); Murphy et al., *Acta Haematologica Japonica*, **46** (7), 1380–1396 (1983); Dessypris et al., *Brit. J. Haematol.*, **56**, 295–306 (1984); und Emmanoul et al., *Am. J. Physiol.*, **247** (1 Pt2), F168–76 (1984).

Da Erythropoietin im Prozeß der roten Blutzellenbildung wesentlich ist, weist das Hormon eine potentiell nützliche Application sowohl bei der Diagnose als auch der Behandlung von Blutkrankheiten auf, die durch eine niedrige oder gestörte rote Blutzellenproduktion gekennzeichnet sind. Siehe dazu allgemein Pennathur-Das et al., *Blood*, **63** (5), 1168–71 (1984) und Haddy, *am. Jour. Ped. Hematol./Oncol.*, **4**, 191–196 (1982), die sich auf Erythropoietin für mögliche Therapien der Sichelzellen, anämie beziehen, und Eschbach et al., *J. Clin. Invest.*, **74** (2), S. 434–441 (1984), die ein therapeutisches Regime für urämische Schafe beschreiben, basierend auf in vivo-Reaktionen auf Erythropoietinreichen Plasmaintfusionen, und einer vorgeschlagenen Dosis von 10 Einheiten EPO/kg pro Tag für 15 bis 40 Tage als Korrektiv einer Anämie des mit chronischem Nierenversagen verbundenen Typs. Siehe auch Krane, *Henry Ford Hosp., Med.*, **31** (3) 177–181 (1983).

Kürzlich wurde eingeschätzt, daß die Erhältlichkeit von Erythropoietin in Mengen die Anämiebehandlung von 1600000 Menschen allein in den Vereinigten Staaten erlauben würde. Siehe z.B. Morrison „Bioprocessing in Space — an Overview“ Seite 557–571 in *The world Biotech Report 1984*, Bd. 2: USA (Online Publications, New York, N. Y. 1984). Kürzliche Studien haben eine Basis für die Planung der Wirksamkeit der Erythropoietintherapie in einer Vielzahl von Krankheitszuständen, Krankheiten und Zuständen hämatologischer Unregelmäßigkeit erbracht:

Vedovato, et al., *Acta. Haematol.*, **71**, 211–213 (1984) (beta-Thalassämie); Vichinsky, et al., *J. Pediatr.*, **105** (1), 15–21 (1984) (zystische Fibrose); Cotes, et al., *Brit. J. Obstet. Gynecol.*, **90** (4), 304–311 (1983) (Schwangerschaft, Menstruationsbeschwerden); Hage, et al., *Acta. Pediatr. Scand.*, **72**, 827–831 (1983) (frühe Anämie der Frühreife); Claus-Walker, et al., *Arch. Phys. Med. Rehabil.*, **65**, 370–374 (1984) (Rückenmarkverletzung) Dunn, et al., *Eur. J. Appl. Physiol.*, **52**, 178–182 (1984) (Raumflug); Miller, et al., *Brit. J. Haematol.*, **52**, 545–590 (akuter Blutverlust); Udupa, et al., *J. Lab. Clin. Med.*, **103** (4), 574–580 und 581–588 (1984); und Lipschitz, et al., *Blood*, **63** (3), 502–509 (1983) (Alterung); und Dainiak, et al., *Cancer*, **51** (6), 1101–1106 (1983) und Schwartz, et al., *Otolaryngol.*, **109**, 269–272 (1983) (verschiedene neoplastische Krankheitsstadien begleitet durch abnorme Erythropoiesie).

Frühere Versuche zur Gewinnung von Erythropoietin in guter Ausbeute aus Plasma oder Urin verliefen relativ erfolglos. Es sind komplizierte und umständliche Laborverfahren erforderlich, die im allgemeinen zur Gewinnung nur sehr geringer Mengen unreiner und instabiler Erythropoietin enthaltender Extrakte führten.

Anfängliche Versuche zur Isolierung von Erythropoietin aus Urin erbrachten instabile, biologisch inaktive Zubereitungen des Hormons. Die US-PS 3865801 beschreibt ein Verfahren zur Stabilisierung der biologischen Aktivität einer Rohsubstanz, die aus Urin gewonnenes Erythropoietin enthält. Das resultierende Rohpräparat, das Erythropoietin enthält, erhält scheinbar 90% der Erythropoietinaktivität zurück und ist stabil.

Ein anderes Reinigungsverfahren von Human-Erythropoietin aus Urin von Patienten mit aplastischer Anämie wird von Miyake et al., in *J. Biol. Chem.*, Bd. 252, Nr. 15 (10. August 1977) S. 5558–5564 beschrieben. Zu dem siebenstufigen Verfahren gehört Ionenaustauschchromatographie, Ethanol-fällung, Gelfiltration und Adsorptionschromatographie, wobei als Ausbeuten ein reines Erythropoietinpräparat mit einer Potenz von 70400 Einheiten/mg Protein mit 21% Ausbeute erhalten wird.

Die US-PS 4397840 (Takezawa et al.) beschreibt Verfahren zur Herstellung „eines Erythropoietinproduktes“ aus Urinproben gesunder Menschen mit schwachbasischen Ionenaustauschern und sagt aus, daß das erhaltene Produkt niedriger Molekülmasse „keine inhibitorischen Effekte gegen Erythropoietin aufweist“.

Die GB-PS 2085887 (Sugitomo et al.), veröffentlicht am 6. Mai 1982, beschreibt ein Verfahren zur Herstellung von Hybrid-Human-lymphoblastoiden Zellen und berichtet über Produktionsmengen im Bereich von 3 bis 420 Einheiten Erythropoietin pro ml Zellsuspension (verteilt in den Kulturen nach der Säugetierwirtsvermehrung, die bis zu 10^7 Zellen pro ml enthielt(en)). Bei den besten Ergebnissen, von denen behauptet wurde, daß man sie erreichte, konnte die Herstellungsrate von Erythropoietin auf 40 bis 4000 Einheiten/ 10^6 Zellen/48 Stunden in vitro-Kultur, die dem Zelltransfer aus in vivo-Vermehrungssystemen folgten, berechnet werden. (Siehe auch die analoge US-PS 4377513). Es wurden zahlreiche Vorschläge unterbreitet zur Isolierung von Erythropoietin aus Gewebe, einschließlich neoplastischen Zellen, die Ausbeuten waren jedoch sehr gering. Siehe z. B. Jelkman, et al., *Expt. Hematol.*, **11** (7), 581–588 (1983); Tambourin, et al., *P. N. A. S. (USA)*, **80**, 6269–6273 (1983); Katsucka, et al., *Gann*, **74**, 534–541 (1983); Hagiwara et al., *Blood*, **63** (4), 828–835 (1984); und Choppin, et al., *Blood* **64** (2), 341–347 (1984).

Zu anderen eingesetzten Isolierungsverfahren, um gereinigtes Erythropoietin zu erhalten, gehören immunologische Verfahrensweisen. Ein gegen Erythropoietin gerichteter polyklonaler, aus Serum herrührender Antikörper wurde durch Injizieren eines Tieres, vorzugsweise einer Ratte oder eines Kaninchens, mit Human-Erythropoietin entwickelt. Das injizierte Human-Erythropoietin wird durch das Immunsystem des Tieres als fremde Antigensubstanz erkannt und löst die Produktion von Antikörpern gegen das Antigen aus. Unterschiedliche Zellen, die auf die Stimulierung durch die Antigensubstanz reagieren, produzieren und setzen Antikörper in den Kreislauf frei, die sich von denen anderer reagierender Zellen geringfügig unterscheiden. Die Antikörperaktivität bleibt im Serum des Tieres erhalten, wenn das Blut extrahiert wird. Während ungereinigte Serum- oder Antikörperzubereitungen, gereinigt als eine Serum-Immunglobulin G-Fraktion, dann in Assays verwendet werden können, um Human-Erythropoietin zu bestimmen und mit ihm einen Komplex bilden zu können, weisen die Materialien einen wesentlichen Nachteil auf. Dieser Serum-Antikörper, zusammengesetzt aus all den unterschiedlichen, durch einzelne Zellen produzierten Antikörpern, ist von Natur aus polyklonal und wird mit Komponenten in Rohextrakten Komplexe bilden, anders als Erythropoietin allein.

Für den Hintergrund der vorliegenden Erfindung von Interesse sind auch kürzlich bekannt gewordene Fortschritte auf dem Gebiet hinsichtlich der Entwicklung kontinuierlicher Zellkulturen, die in der Lage sind, eine einzige Art von Antikörpern zu bilden, die spezifisch immunologisch reaktiv mit einer einzigen antigenen Bestimmungsgröße eines ausgewählten Antigens ist. Siehe allgemein Chisholm, *High Technology*, Bd. 3, Nr. 1, 57–63 (1983).

Es wurden Versuche unternommen, Zellfusions- und Hybridisierungstechniken anzuwenden, um „monoklonale“ Antikörper bei der Isolierung und quantitativen Bestimmung von Human-Erythropoietin absonderten, in Kurzform durch Lee-Huang, *Abstract No. 1463, Fed. Proc.*, **41**, 520 (1982). Als ein weiteres Beispiel erschien eine detaillierte Beschreibung der Herstellung und Verwendung monoklonaler Anti-Erythropoietin-Antikörper durch Weiss et al., *P. N. A. S. (USA)* **9**, 5465–5469 (1982). Siehe durch Sasaki, *Biomed. Biochim. Acta.*, **42/11** (12), 5202–5206 (1983); Yanagawa, et al., *Blood* **64** (2), 357–364 (1984), Yanagawa, et al., *J. Biol. Chem.*, **259** (5), 2707–2710 (1984); und US-PS 4465624.

Für den Hintergrund der Erfindung sind ebenso von Interesse Berichte über die immunologische Aktivität synthetischer Peptide die die in natürlich vorkommenden Proteinen, Glycoproteinen und Nucleoproteinen vorhandenen Aminosäuresequenzen im wesentlichen verdoppeln. Insbesondere Polypeptide mit relativ geringer Molekülmasse haben gezeigt, daß sie an Immunreaktionen teilnehmen, die in Dauer und Umfang den Immunreaktionen physiologisch bedeutender Proteine, wie virale Antikörper, Polypeptidhormone und ähnliche ähneln. Zu den Immunreaktionen derartiger Polypeptide gehört auch die Provozierung der Bildung spezifischer Antikörper in immunologisch aktiven Tieren. Siehe z. B. Lerner et al., *Cell*, **23**, 309–310 (1981); Ross, et al., *Nature*, **294**, 654–656 (1981); Walter, et al., *P. N. A. S. (USA)*, **77**, 5197–5200 (1980); Lerner, et al.: *P. N. A. S. (USA)*, **78**, 3403–3407 (1981); Walter et al., *P. N. A. S. (USA)*, **78**, 4882–4886 (1981); Wong, et al., *P. N. A. S. (USA)*, **78**, 7412–7416 (1981); Green, et al., *Cell*, **28**, 477–487 (1982); Nigg, et al., *P. N. A. S. (USA)*, **79**, 5322–5326 (1982); Baron, et al., *Cell*, **28**, 395–404 (1982); Dreesman, et al., *Nature*, **295**, 158–160 (1982); und Lerner, *Scientific American*, **248**, No. 2, 66–74 (1983). Siehe auch Kaiser, et al., *Science*, **223**, pp. 249–255 (1984) bezüglich biologischer und immunologischer Aktivitäten synthetischer Peptide, die die sekundären Strukturen der Peptidhormone annähernd teilen, nicht jedoch deren primäre strukturelle Anordnung. Die oben genannten Studien beziehen sich natürlich auf Aminosäuresequenzen von Proteinen, die sich von Erythropoietin unterscheiden, einer Substanz, für die keine wesentliche Aminosäuresequenz bisher beschrieben worden ist. In der anhängigen US-Patentanmeldung Nr. 463724, an der Miteigentum besteht, eingereicht am 4. Februar 1983 durch J. Egrie, veröffentlicht am 22. August 1984 als EP-OS 0 116446 wird eine Maus-Maus-Hybridoma-Zelllinie (ATCC HB 8209) beschrieben, die einen hoch spezifischen, monoklonalen, Anti-Erythropoietin-Antikörper produziert, der auch spezifisch immunoreaktiv mit einem Polypeptid ist, das aus der folgenden Sequenz von Aminosäuren besteht:

$\text{NH}_2\text{-Ala-Pro-Pro-Arg-Leu-Ile-Cys-Asp-Ser-Arg-Val-Leu-Glu-Arg-Tyr-Leu-Leu-Glu-Ala-Lys-COOH}$.

Die Polypeptidsequenz ist eine, die den ersten zwanzig Aminosäureresten reifen Human-Erythropoietins zuzuschreiben ist, die nach dem Verfahren von Miyake et al., *J. Biol. Chem.*, **252**, 5558–5564 (1977) isoliert wurden, und für die eine Aminosäureanalyse mittels Gasphasensequenzierung (Applied Biosystems, Inc.) durchgeführt wurde nach der Verfahrensweise von Hewick, M. et al., *J. Biol. Chem.*, **256**, 7990–7997 (1981). Siehe auch Sue et al., *Proc. Nat. Acad. Sci. (USA)*, **80**, S. 3651–3655 (1983) betreffend die Entwicklung polyklonaler Antikörper gegen ein synthetisches 26-mer, basierend auf einer unterschiedlichen Aminosäuresequenz und Sytowski et al., *J. Immunol. Methods*, **69**, S. 181–186 (1984).

Während polyklonale und monoklonale Antikörper, wie oben beschrieben, zu sehr nützlichen Produkten für den Einsatz in Immunoassays zur Bestimmung und mengenmäßigen Erfassung von Erythropoietin führen und bei der Affinitätsreinigung von Erythropoietin verwendbar ist, erscheint es unwahrscheinlich, daß diese Materialien leicht zur großtechnischen Isolierung größerer Mengen Erythropoietin aus Säugetiermaterial eingesetzt werden können, die ausreichend sind für weitere Analysen, klinische Untersuchungen und eine potentiell weitreichende therapeutische Verwendung der Substanz bei der Behandlung von beispielsweise chronischen Nierenerkrankungen, wobei krankes Gewebe nicht in der Lage ist, die Produktion von Erythropoietin aufrechtzuerhalten. Es besteht daher in der Wissenschaft die Auffassung, daß die besten Aussichten zur vollständigen Kennzeichnung von Säugetier-Erythropoietin und Bereitstellung großer Mengen davon zur potentiellen diagnostischen und klinischen Verwendung darin bestehen, erfolgreich die rekombinanten Verfahren einzusetzen, um eine großtechnische mikrobielle Synthese der Verbindung zu ermöglichen.

Während es so erscheint, daß bei der versuchten Isolierung von DNA-Sequenzen, die für Human- und andere Säugetierarten-Erythropoietin kodieren, wesentliche Fortschritte gemacht wurden, scheint dies doch nicht so erfolgreich gewesen zu sein. Dies liegt prinzipiell an dem Mangel an Gewebematerial, insbesondere menschlichem Gewebematerial, das mRNA-angereichert ist, und das gestatten würde, eine cDNA-Bibliothek aufzubauen, aus der eine für Erythropoietin codierende DNA-Sequenz durch Standardverfahren isoliert werden könnte. Darüber hinaus ist über die fortlaufende Sequenz der Aminosäurereste des Erythropoietins so wenig bekannt, daß es nicht möglich ist, beispielsweise lange Polynucleotidsonden zu konstruieren, die in der Lage sind, zuverlässig bei dem DNA/DNA-Hybridisierungsscreening von cDNA und insbesondere genomischen DNA-Bibliotheken eingesetzt zu werden.

Als Erläuterung dafür sei genannt, daß die 20-Aminosäuresequenz, die zur Erzeugung des oben genannten monoklonalen Antikörpers verwendet wurde, produziert durch ATCC Nr. HB 8209, keine Konstruktion einer zweifelsfreien 60-Basenpaar-Oligonucleotidsonde in der von Anderson et al., ebenda, beschriebenen Weise. Es wird eingeschätzt, daß das Humangen für Erythropoietin als ein „Einzelkopie-Gen“ (single copy gen) innerhalb des menschlichen Genoms erscheint und in jedem Fall das für Human-Erythropoietin codierende genetische Material wahrscheinlich aus weniger als 0,00005% der gesamten menschlichen genomischen DNA besteht, die in einer Gen-Bibliothek vorhanden sein würde.

Bis heute sind die erfolgreichsten der in der Literatur bekannt gewordenen Versuche mit Rekombinanten verwandten Verfahren zur Bereitstellung von DNA-Sequenzen für die Verwendung bei der mikrobiellen Expression isolierbarer Mengen von Säugetier-Erythropoietin weit vor dem Ziel stehengeblieben. Als Beispiel berichtet Farber et al., Exp. Hematol. 11, Suppl. 14, Abstract 101 (1983) über die Extraktion von mRNA aus Nierengewebe von Phenylhydrazin-behandelten Pavianen und die Injizierung von mRNA in *Xenopus laevis*-Eizellen mit dem eher flüchtigen Ergebnis der in vitro Produktion eines Gemisches von „Translationsproduktion“, das unter anderem biologische Eigenschaften von Erythropoietin entfaltete.

In der jüngsten Zeit berichteten Farber et al., Blood, 62, Nr. 5, Suppl. Nr. 1, Abstract 392 auf Seite 122a (1983) über die in vitro-Translation von Humannieren-mRNA durch Frosch-Eizellen. Es wurde geschätzt, daß das resultierende Translationsproduktgemisch in der Größenordnung von 220 (Millieinheiten) (mU) eines Translationsproduktes mit der Aktivität von Erythropoietin pro Mikrogramm injizierte mRNA enthielt. Während sich solche Gehalte bei der in vitro-Translation exogener mRNA als ziemlich niedrig bestätigten (verglichen mit den schon früher bekannt gewordenen Gehalten der Pavian-mRNA-Translation in das gesuchte Produkt), wurde geglaubt, daß die Ergebnisse die menschliche Niere als Ort der Erythropoietinexpression bestätigen und die Konstruktion einer angereicherten Humannieren-cDNA-Bibliothek gestatten würden, aus der das gewünschte Gen isoliert werden könnte. (siehe auch Farber, Clin. Res., 31 [4], 769A [1983]). Seit Einreichen der US-Patentanmeldungen Nr. 561 024 und 582 185 ist ein einziger Bericht zur Klonierung und Expression erschienen, in dem behauptet wird, Human-Erythropoietin-cDNA in *E. coli* erhalten zu haben. Dabei wurden eine Anzahl von cDNA-Klonen in *E. coli*-Plasmide inseriert, und es wurden β -Lactamase-Fusionsprodukte festgestellt, die sich immunreaktiv mit einem monoklonalen Antikörper zu einem unspezifisierten „Epitop“ von Human-Erythropoietin verhielten. Siehe Lee-Huang, Proc. Nat. Acad. Sci. (USA), 81, Ss. 2708-2712 (1984).

Ziel der Erfassung

Durch die Erfindung wird ein verbessertes Verfahren zur Bestimmung eines spezifischen einzelsträngigen Polynucleotids unbekannter Sequenz zur Verfügung gestellt.

Darlegung des Wesens der Erfindung

Der Erfindung liegt die Aufgabe zugrunde, ein Verfahren zur Bestimmung eines spezifischen einzelsträngigen Polynucleotids unbekannter Sequenz zur Verfügung zu stellen.

Gegenstand der Erfindung ist ein Verfahren zur Bestimmung eines spezifischen einzelsträngigen Polynucleotids unbekannter Sequenz in einer heterogenen zellulären oder viralen Probe einschließlich multipler einzelsträngiger Polynucleotide, gekennzeichnet dadurch, daß

- (a) ein Gemisch markierter einzelsträngiger Polynucleotidsonden mit gleichmäßig variierenden Sequenzen der Basen hergestellt wird, wobei jede der Sonden potentiell einer Sequenz von Basen komplementär ist, die zu dem zu bestimmenden Polynucleotid das vermeintlich nur einmal existierende Exemplar ist;
- (b) die Probe auf ein festes Substrat fixiert wird;
- (c) das Substrat mit der darauf fixierten Probe behandelt wird, um eine weitere Bindung von Polynucleotiden darauf abzuschwächen, ausgenommen durch Hybridisierung zu Polynucleotiden in der Probe;
- (d) das behandelte Substrat mit der darauf fixierten Probe flüchtig mit dem genannten Gemisch markierter Sonden unter Bedingungen in Berührung gebracht wird, die nur die Hybridisierung zwischen vollständig komplementären Polynucleotiden erleichtern; und
- (e) das spezifische Polynucleotid über die Beobachtung des Auftretens einer Hybridisierungsreaktion zwischen ihm und einer vollständig komplementären Sonde aus dem Gemisch der markierten Sonden bestimmt wird.

Die vorliegende Erfindung erläutern klonierte DNA-Sequenzen vom Affen oder Menschen und auf geeignete Weise daraus abgeleitete Polypeptidsequenzen, die die primäre Strukturordnung von Erythropoietinen mit ihrem entsprechenden Ursprung vom Affen oder Menschen darstellen.

Unter Darstellung der Sequenz der Aminosäurereste von Erythropoietin sieht die vorliegende Erfindung die totale und/oder partielle Herstellung von DNA-Sequenzen vor, die für Erythropoietin codieren und die solch vorteilhafte Charakteristika wie Einbeziehung von für die Expression „bevorzugten“ Codons durch Nicht-Säugetierwirte, mit der Maßgabe, daß Stellen für das Schneiden durch Restriktionsendonuclease-Enzyme vorgesehen sind und der weiteren Maßgabe von zusätzlichen Anfangs-, End- oder Zwischen-DNA-Sequenzen, die die Konstruktion leicht exprimierbarer Vektoren erleichtern. Dementsprechend betrifft die vorliegende Erfindung die Herstellung (und Entwicklung durch Stellen-spezifische Mutagenese von cDNA und genomischer DNA) von DNA-Sequenzen, die für die mikrobielle Expression von Polypeptidanalogen oder Derivaten von Erythropoietin

codieren, die sich von natürlich vorkommenden Formen durch Identität oder Ort der Anordnung eines oder mehrerer Aminosäurereste unterscheiden (d. i. Deletierungsanaloga, die weniger als alle der für EPO spezifizierenden Reste enthalten und/oder Substitutionsanaloga, worin einer oder mehrere der spezifizierten Reste durch andere Reste ersetzt ist und/oder Additionsanaloga, worin einer oder mehrere Aminosäurereste an einen terminalen oder medialen Teil des Polypeptids angefügt ist); und die an einigen oder allen Eigenschaften natürlich vorkommender Formen teilhaben.

Die neuen DNA-Sequenzen der Erfindung enthalten alle für eine sichere Expression in procaryotischen oder eucaryotischen Wirtszellen von Polypeptidprodukten erforderlichen Sequenzen, die zumindest einen Teil der primären Strukturanordnung aufweisen sowie eine oder mehrere der biologischen Eigenschaften von Erythropoietin, von denen umfaßt werden:

- (a) die in den Tabellen V und VI aufgeführten DNA-Sequenzen oder deren Komplementärstränge;
- (b) DNA-Sequenzen, die (unter Hybridisierungsbedingungen, wie sie hier erläutert werden oder strengeren Bedingungen) zu DNA-Sequenzen, wie sie unter (a) definiert sind, hybridisieren, oder zu Fragmenten davon; und
- (c) DNA-Sequenzen, die allerdings unter Degenerierung des genetischen Codes zu unter (a) und (b) definierten DNA-Sequenzen hybridisieren würden. Speziell vom Teil (b) umfaßt sind genomische DNA-Sequenzen, die allelische Variatenformen von Affen- oder Human-Erythropoietin verschlüsseln und/oder andere Säugetierarten von Erythropoietin decodieren. Speziell vom Teil (c) umfaßt sind hergestellte DNA-Sequenzen, die EPO, EPO-Fragmente und EPO-Analoga verschlüsseln, deren DNA-Sequenzen Condons enthalten können, die die Translation von Boten-RNA in Nicht-Wirbeltierwirten erleichtern.

Weiterhin der vorliegenden Erfindung zuzuordnen ist die Klasse von Polypeptiden, die für Teil des DNA-Komplements zum oberen Strang der Human-genomischen DNA-Sequenz von Tabelle VI codieren, d. i. „komplementär umgekehrte Proteine“, wie sie von Tramontano et al., *Nucleic Acids Research*, **12**, S. 5049-5059 (1984).

Die erfindungsgemäßen Polypeptidprodukte können „markiert“ werden durch kovalente Bindung an eine bestimmbare Markersubstanz (z. B. radiomarkiert mit ¹²⁵I), um zu Reagentien zu gelangen, die bei der Ermittlung und mengenmäßigen Bestimmung von Erythropoietin in festem Gewebe und flüssigen Proben wie Blut oder Urin von Nutzen sind. Die erfindungsgemäßen DNA-Produkte können auch mit bestimmbaren Markern (wie Radiomarkern und Nicht-Isotopenmarkern wie Biotin) markiert und bei DNA-Hybridisierungsverfahren eingesetzt werden, um die Erythropoietin-Genposition und/oder die Genposition irgendeiner verwandten Genfamilie aus der Human-, Affen- und anderen Säugetierspezies-Chromosomenkarte zu orten. Sie können auch zum Herausfinden von Erythropoietin-Genunstimmigkeiten an Hand des DNA-Gehalts verwendet werden und als Genmarker zur Identifizierung benachbarter Gene und deren Unstimmigkeiten.

Wie hier anschließend im Detail beschrieben wird, bringt die vorliegende Erfindung signifikante Verbesserungen bei den Bestimmungsverfahren eines speziellen einzelsträngigen Polynucleotids unbekannter Sequenz in einer heterogenen zellulären oder viralen Probe mit sich, die multipler einzelsträngiger Polynucleotide enthält, wobei

- (a) ein Gemisch markierter, einzelsträngiger Nucleotidsonden mit gleichmäßig variierenden Basensequenzen hergestellt wird, wobei jede dieser Sonden potentiell spezifisch komplementär zu einer Sequenz von Basen ist, die in dem zu bestimmenden Polynucleotid vermutlich nur einmal existierend ist,
- (b) die Probe auf einem festen Substrat fixiert ist,
- (c) das Substrat mit der darauf fixierenden Probe behandelt wird, um die weitere Bindung von Polynucleotiden daran zu verhindern, ausgenommen mittels Hybridisierung an Polynucleotide in der Probe,
- (d) das behandelte Substrat mit der darauf fixierten Probe flüchtig mit dem genannten Gemisch markierter Sonden unter solchen Bedingungen in Kontakt gebracht wird, die die Hybridisierung nur zwischen vollständig komplementären Polynucleotiden erleichtern, und
- (e) das spezifische Polynucleotid durch Beobachten des Ablaufens einer Hybridisierungsreaktion zwischen diesem und einer vollständig komplementären Sonde innerhalb des Gemisches markierter Sonden bestimmt wird, was durch Vorhandensein einer höheren Dichte von markiertem Material auf dem Substrat am Ort des spezifischen Polynucleotids bewiesen wird, im Vergleich zu einer Hintergrund-Dichte am markierten Material, die sich aus der nichtspezifischen Bindung markierter Sonden an das Substrat ergibt.

Die Verfahren sind insbesondere effektiv bei Situationen, die die Verwendung von 64, 128, 256, 512, 1024 oder mehr gemischten Polynucleotidsonden mit einer Länge von 17 bis 20 Basen bei DNA/DNA- oder RNA/RNA- oder DNA/RNA-Hybridisierungen erfordern.

Wie weiter unten ausgeführt wird, haben die oben genannten verbesserten Verfahren die Erkennung von cDNA-Klonen gestattet, die für Erythropoietin mit Ursprung von Affenspezies codieren, innerhalb einer Bibliothek, die aus anämischer Affennierenzell-mRNA hergestellt wurde. Insbesondere wurde ein Gemisch von 128 gleichmäßig variierten 20-mer-Sonden basierend auf einer Aminosäuresequenzinformation, die von sequenzierenden Fraktionen von Human-Erythropoietin abgeleitet wurde, in Koloniehybridisierungsverfahren eingesetzt, um sieben „positive“ Erythropoietin-cDNA-Klone innerhalb einer Gesamtmenge von 200 000 Kolonien zu identifizieren. Noch bemerkenswerter ist, daß der praktische Einsatz der verbesserten erfindungsgemäßen Verfahren die schnelle Isolierung von drei positiven Klonen gestattet, und zwar über ein Screening von 1 500 000 Phagenplaques, die eine Human-Genbibliothek darstellen. Dies wurde erreicht durch Verwendung des oben genannten Gemisches von 128 20-mer-Sonden zusammen mit einem zweiten Satz von 128 17-mer-Sonden, basierend auf einer Aminosäureanalyse einer unterschiedlichen fortlaufenden Sequenz von Human-Erythropoietin.

Die oben genannten erläuternden Verfahren bilden den ersten bekannten Fall der Verwendung multipler gemischter Oligonucleotidsonden bei DNA/DNA-Hybridisierungsverfahren, die auf die Isolierung von Säugetiergenomklonen gerichtet sind, und den ersten bekannten Fall der Verwendung eines Gemisches von mehr als 32 Oligonucleotidsonden bei der Isolierung von cDNA-Klonen. Für den Fachmann ergeben sich zahlreiche Aspekte und Vorteile der Erfindung durch die folgende detaillierte Beschreibung, in der auch praktische Seiten der Erfindung in ihren gegenwärtig bevorzugten Ausführungsformen erläutert werden.

Erfindungsgemäß wurden DNA-Sequenzen isoliert und bestimmt, die einen Teil oder die gesamte Polypeptid-Sequenz von Human- und Affenspezies-Erythropoietin (anschließend teilweise als „EPO“ bezeichnet) verschlüsseln. Weiterhin wurde die vom Affen oder Menschen stammende DNA zum Subjekt eucaryotischer und procaryotischer Expression gemacht, wobei isolierbare Mengen an Polypeptiden entstehen, die biologische (z. B. immunologische) Eigenschaften von natürlich vorkommenden EPO zeigen sowie sowohl in vivo- als auch in vitro-biologische Aktivitäten von EPO.

Die DNA aus Affenspezies wurde aus einer cDNA-Bibliothek isoliert, die mit einer mRNA konstruiert worden war, die aus Affennierengewebe in einem auf chemischem Wege ausgelösten anämischen Zustand gewonnen wurde, und deren Serum nach immunologischer Bestimmung hohe Anteile EPO enthielt, verglichen mit normalem Affenserum. Die Isolierung der gewünschten cDNA-Klone, die DNA verschlüsselnde EPO enthielten, erreichte man unter Einsatz von DNA/DNA-Koloniehybridisierung durch Verwendung eines Pools von 128 gemischten, radiomarkierten 20-mer-Oligonucleotidsonden unter Einbeziehung des Schnellscreenings von 200000 Kolonien. Das Muster der Oligonucleotidsonden wurde auf Aminosäuresequenzinformation stationiert, die durch enzymatische Fragmentierungen und Sequenzieren einer kleinen Human-EPO-Probe gewonnen wurden.

Die DNA von Menschen wurde aus einer humangenomischen DNA-Bibliothek isoliert. Die Isolierung von Klonen, die EPO-verschlüsselnde DNA enthielten, wurde durch DNA/DNA-Plauehybridisierung erreicht unter Verwendung des oben angeführten Pools von 128 gemischten 20-mer-Oligonucleotidsonden und einem zweiten Pool von 128 radiomarkierten 17-mer-Sonden, deren Sequenzen auf Aminosäuresequenzinformationen stationiert sind, die man aus einem unterschiedlichen enzymatischen Human-EPO-Fragment erhielt.

Positive Kolonien und Plaques wurden mittels Dideoxy-Sequenzieren klonaler DNA unter Verwendung eines Teilsatzes von 16 Sequenzen innerhalb des Pools von 20-mer-Sonden auf Richtigkeit nachgeprüft, und ausgewählte Klone wurden einer Nucleotid-Sequenzanalyse unterworfen, die die Herleitung primärer Strukturübereinstimmung der damit verschlüsselten EPO-Polypeptide zuließ. Die hergeleiteten Polypeptidsequenzen zeigten einen hohen Grad an Homologie zueinander und zu einer durch Aminosäureanalyse von Human-EPO-Fragmenten erzeugten Partialsequenz.

Ein selektierter positiver Affen-cDNA-Klon und ein selektierter positiver Human-Genklon wurden jeweils in einen „Zubringer“- („shuttle“) DNA-Vektor inseriert, der in *E. coli* amplifiziert und dazu verwendet wurde, Säugetierzellen in Kultur zu transfektieren. Kultiviertes Wachstum transfektierter Wirtszellen führte zu Präparationen des Überstehenden des Kulturmediums, die mehr als 3000 mE EPO pro ml Kulturflüssigkeit enthielten.

Ausführungsbeispiel

Die folgenden Beispiele dienen der Erläuterung der Erfindung und sind speziell auf Verfahren gerichtet, die vorher zur Erkennung von EPO verschlüsselnden Affen-cDNA-Klonen und humangenomischen Klonen durchgeführt wurden, auf Verfahren, die in einer solchen Erkennung führen und auf das Sequenzieren.

Insbesondere Beispiel 1 ist auf das Aminosäuresequenzieren von Human-EPO-Fragmenten gerichtet und auf die Konstruktion von Gemischen radiomarkierter Sonden, die auf den Ergebnissen dieses Sequenzierens basieren. Beispiel 2 ist allgemein auf Verfahren gerichtet, zu denen die Erkennung von positiven Affen-cDNA-Klonen gehört und erbringt eine Information zur Behandlung des Tieres und zur vorläufigen Radioimmunoassay-(RIA)-analyse tierischer Seren. Beispiel 3 ist gerichtet auf die Herstellung der cDNA-Bibliothek, das Kolonie-Hybridisierungsscreening und die Überprüfung der positiven Klone auf Richtigkeit, das DNA-Sequenzieren eines positiven cDNA-Klons und die Erzeugung einer primären Strukturübereinstimmung(Aminosäuresequenz)-Information des Affen-EPO-Polypeptids. Beispiel 4 ist auf Verfahren gerichtet, die die Erkennung positiver humangenomischer Klone betreffen und zu einer Information über die Herkunft der Gen-Bibliothek, über Plauehybridisierungsverfahren und die Überprüfung auf Richtigkeit positiver Klone führen. Beispiel 5 ist auf das DNA-Sequenzieren eines positiven genomischen Klons und die Erzeugung einer Aminosäuresequenzinformation des Human-EPO-Polypeptids gerichtet, einschließlich eines Vergleichs dieser mit der Affen-EPO-Sequenzinformation.

Beispiel 1

A. Human-EPO-Fragment-Aminosäuresequenzieren

Human-EPO wurde aus Urin isoliert und einem tryptischen Abbau unterworfen, was zur Entwicklung und Isolierung von 17 getrennten Fragmenten in Mengen von annähernd 100 bis 150 Picomolen führte.

Den Fragmenten wurden willkürlich Nummern zugeordnet, und sie wurden auf Aminosäuresequenzen durch Mikrosequenzanalyse unter Verwendung eines Gasphasensequenzers (Applied Biosystems) analysiert, wobei man die in Tabelle I dargestellte Sequenzinformation erhielt. Es wurden darin Einzelbuchstabencodes verwendet, und „X“ bezeichnet einen Rest, der nicht genauer weiter bestimmt wurde.

Tabelle I

Fragment No.	Ergebnis Sequenzanalyse
T 4 a	A-P-P-R
T 4 b	G-K-L-K
T 9	A-L-G-A-Q-K
T 13	V-L-E-R
T 16	A-V-S-G-L-R
T 18	L-F-R
T 21	K-L-F-R
T 25	Y-L-L-E-A-K
T 26 a	L-I-C-D-S-R
T 26 b	L-Y-T-G-E-A-C-R
T 27	T-I-T-A-D-T-F-R
T 28	E-A-I-S-P-P-D-A-A-M-A-A-P-L-R
T 30	E-A-E-X-I-T-T-G-X-A-E-H-X-S-L-N-E-X-I-T-V-P
T 31	V-Y-S-N-F-L-R
T 33	S-L-T-T-L-L-R
T 35	V-N-F-Y-A-W-K
T 38	G-Q-A-L-L-V-X-S-S-Q-P-W-E-P-L-Q-L-H-V-D-K

B. Muster und Konstruktion von Oligonucleotidsendengemischen

Die in Tabelle I aufgeführten Aminosäuresequenzen wurden im Zusammenhang mit der Degenerierung des genetischen Codes dahingehend nochmals überprüft, ob die Mischsondenverfahren auf DNA/DNA-Hybridisierungsverfahren in cDNA- und/oder genomischen DNA-Bibliotheken angewandt werden können. Diese Analyse zeigte, daß innerhalb des Fragmentes Nr. T 35 eine Serie von 7 Aminosäureresten (Val-Asn-Phe-Tyr-Ala-Trp-Lys) existierte, die einzig und allein dadurch gekennzeichnet werden kann, daß sie durch eine von 128 möglichen DNA-Sequenzen, die 20 Basenpaare überspannt, verschlüsselt wird. Ein erster Satz von 128 20-mer-Oligonucleotiden wurde somit durch Standard-Phosphoramiditverfahren (siehe z. B. Beaucage et al., Tetrahedron Letters, **22**, S. 1859–1862 (1981) auf festen Träger entsprechend der in Tabelle II aufgeführten Sequenz synthetisiert.

Tabelle II

Rest 3'	—	Val	Asn	Phe	Tyr	Ala	Trp	Lys
		CAA	TTG	AAG	ATG	CGA	ACC	TT—5'
		T	A	A	A	T		
		G				G		
		C				C		

Die weitere Analyse zeigte, daß innerhalb des Fragments T 38 eine Serie von 6 Aminosäureresten (Gln-Pro-Trp-Glu-Pro-Leu) auf der Basis existierte, von der aus man einen Pool von 128 Misch-Oligonucleotid-17-mer-Sonden herstellen konnte, wie unten in Tabelle III aufgeführt sind.

Tabelle III

Rest 3'	Gln	Pro	Trp	Glu	Pro	Leu
	GTT	GGA	ACC	CTT	GGA	GA—5'
	C	T		C	T	kA
		G			G	
		C			C	

Die Oligonucleotidsonden wurden am 5'-Ende mit Gamma-³²-p-ATP, 7500–8000 ci/mmol (ICN) unter Verwendung von T₄-Polynucleotidkinase (NEN) markiert.

Beispiel 2

A. Affenbehandlungsverfahren und RIA-Analyse

Weibliche Cynomolgus-Affen *Macaca fascicularis* (2,5–3 kg), eineinhalb bis zwei Jahre alt) wurden subkutan mit einer pH 7-Lösung von Phenylhydrazinhydrochlorid bei einer Dosis von 12,5 mg/kg am ersten, dritten und fünften Tag behandelt. Vor jeder Injektion wurde der Hämatokritwert beobachtet. Am siebenten Tag oder zu dem Zeitpunkt, wo der Hämatokritspiegel unter 25% des Anfangswertes fiel, wurden Serum und Nieren nach Verabreichung von 25 mg/kg Dosen Ketaminhydrochlorid entnommen. Die entnommenen Materialien wurden unmittelbar danach in flüssigem Stickstoff eingefroren und bei –70°C gelagert.

B. RIA für EPO

Es wurden Radioimmunoassays zur quantitativen Bestimmung von EPO in Proben nach den folgenden Verfahrensweisen durchgeführt:

Ein Erythropoietin-Standard oder eine unbekannte Probe wurde zusammen mit Antiserum für 2 Stunden bei 37°C inkubiert. Nach der zweistündigen Inkubierung wurden die Proberöhrchen in Eis gekühlt, ¹²⁵I-markiertes Erythropoietin hinzugegeben und die Röhrchen bei 0°C für wenigstens 15 weitere Stunden inkubiert. Jedes Assayröhrchen enthielt 500 µl Inkubationsgemisch, bestehend aus 50 µl verdünnten Immunseren, 10 000 cpm ¹²⁵I-Erythropoietin, 5 µl Trasylol und 0 bis 250 µl entweder des EPO-Standards oder der unbekannten Probe, wobei mit 0,1% BSA enthaltendem PBS das restliche Volumen erbracht wurde. Das verwendete Antiserum war das zweite Testblut eines mit 1%igem Reinpräparat von Humanurin-Erythropoietin immunisierten Kaninchens. Die Endverdünnung des Antiserums bei der Assay wurde so eingestellt, daß das Antikörper-gebundene ¹²⁵I-EPO 10–20% der Eingangs-Gesamtcounts überschritt. Allgemein entsprach dies einer Endverdünnung des Antiserums von 1:50 000 bis 1:100 000.

Das Antikörper-gebundene ¹²⁵I-Erythropoietin wurde durch Zusatz von 150 µl Staph A gefällt. Nach einer 40minütigen Inkubation wurden die Proben zentrifugiert und die Pellets zweimal mit 0,75 ml 10 mM Tris-HCl von pH 8,2, die 0,15 M NaCl, 2 mM EDTA und 0,05% Triton x-100 enthielt, gewaschen. Die gewaschenen Pellets wurden in einem Gammazähler gezählt, um den Prozentsatz der ¹²⁵I-Erythropoietinbindung zu bestimmen. Die durch Prä-Immunseren gebundenen Counts wurden von allen Gesamtwerten abgezogen zwecks Korrektur der nicht-spezifischen Fällung. Der Erythropoietingehalt der unbekannten Proben wurde durch Vergleich mit der Standardkurve ermittelt.

Das oben genannte Verfahren wurde auf im Teil A (s. c.) erhaltenes Affenserum angewandt und ebenso auf das unbehandelte Affenserum. Normale Serumspiegel wurden geprüft und zeigten annähernd 36 mE/ml, während behandeltes Affenserum von 1000 bis 1700 mE/ml enthielt.

Beispiel 3

A. Konstruktion der Affen-cDNA-Bibliothek

Aus normalen und anämischen Affennieren wurde Boten-RNA isoliert mit Hilfe des Guanidin-Thiocyanatverfahrens von Chirgwin et al., *Biochemistry*, **18**, S. 5294 (1979), und die Poly (A)⁺mRNA wurde durch zwei Durchläufe der Oligo(dT)-Zellulose-Säulenchromatografie gereinigt, beschrieben bei „Maniatis“ et al., „Molecular Cloning, A Laboratory Manual“, S. 197–198 Cold Springs Harbor Laboratory, Cold Springs, Harbor, N.Y., 1982). Die cDNA-Bibliothek wurde gemäß einer Modifikation der allgemeinen Verfahrensweisen von Okayama et al., *Mol. and Cell. Biol.*, **2**, S. 161–170 (1982) konstruiert. Die Schlüsselmerkmale der gegenwärtig bevorzugten Verfahren sind die folgenden:

- (1) pUC8 wurde als alleiniger Vektor verwendet, geschnitten mit **Pst** I und dann mit einem Schwanzstück von Oligo dT von 60 bis 80 Basen in der Länge versehen;
- (2) **Hinc** II-Verdauung wurde eingesetzt, um das Oligo dT-Schwanzstück vom einen Ende des Vektors zu entfernen;
- (3) die erste Strangsynthese und das Versehen von Oligo dG mit einem Schwanzstück erfolgte nach dem veröffentlichten Verfahren
- (4) es wurde mit einer **Bam** HI-Verdauung gearbeitet, um den Oligo dG-Schwanz von einem Ende des Vektors zu entfernen; und
- (5) Ersatz des RNA-Stranges durch DNA in Anwesenheit von zwei Linkern (GATCTAAAGACCGTCCCCCCCC und ACGGTCTTTA) in einem dreifachen molaren Überschuß über den mit den Oligo dG-Schwanzstück versehenen Vektor.

B. Koloniehybridisierungsverfahren für das Screenen der Affen-cDNA-Bibliothek

Transformierte *E. coli* wurde bei einer Dichte von 9000 Kolonien pro 10 × 10 cm Platte auf Nährmediumplatten abgebracht, die 50 Mikrogramm/ml Ampicillin enthielten. Gene-Screen-Filter (New England Nuclear Catalog Nr. NEF-972) wurden auf einer BHI-CAM-Platte vorbenetzt (Bacto Hirn-Herz-Infusion 37 g/l, Casaminsäuren 2 g/l und Agar 15 g/l, mit 500 Mikrogramm/ml Chloramphenicol) und dazu verwendet, die Kolonien von der Platte abzuheben. Die Kolonien wurden im gleichen Medium 12 Stunden wachsen gelassen, um die Plasmidkopieanzahl zu amplifizieren. Die amplifizierten Kolonien (Kolonieseite oben) wurden durch reihenweises Auflegen der Filter über 2 Stücke Whatman 3 MM-Papier behandelt, gesättigt mit jeweils den folgenden Lösungen:

- (1) 50 mM Glucose — 25 mM Tris-HCl (pH 8,0) — 10 mM EDTA (pH 8,0) für fünf Minuten;
- (2) 0,5 M NaOH für zehn Minuten; und
- (3) 1,0 M Tris-HCl (pH 7,5) für drei Minuten

Die Filter wurden anschließend im Vakuum bei 80°C für zwei Stunden luftgetrocknet.

Durch Behandlung mit einer Lösung, die 50 Mikrogramm/ml des Proteaseenzym in Puffer K [0,1 M Tris-HCl (pH 8,0) — 0,15 NaCl — 10 mM EDTA (pH 8,2) — 0,2% SDS] wurden die Filter danach dem Protease K-Abbau unterworfen. Jedem Filter wurden speziell 5 ml der Lösung hinzugesetzt, und man ließ den Abbau bei 55°C über 30 Minuten vor sich gehen. Danach wurde die Lösung entfernt.

Die Filter wurden dann mit 4 ml eine Prä-Hybridisierungspuffers (5 × SSPE — 0,5% SDS — 100 Mikrogramm/ml *SS E. coli*-DNA — 5 × BFP) behandelt. Die Prä-Hybridisierungsbehandlung erfolgte bei 55°C, im allgemeinen über 4 Stunden oder länger.

Danach wurde der Prä-Hybridisierungspuffer entfernt.

Das Hybridisierungsverfahren wurde auf folgende Weise durchgeführt. Zu jedem Filter wurden 3 ml Hybridisierungspuffer (5 × SSPE — 0,5% SDS — 100 Mikrogramm/ml Hefe-tRNA), der 0,025 Picomole jeder der 128 Sondensequenzen von Tabelle II enthielt (das Gesamtgemisch wurde als EPV-Gemisch bezeichnet), gegeben, und die Filter wurden 20 Stunden bei 48°C gehalten. Diese Temperatur lag um 2°C niedriger als die niedrigste der berechneten Dissoziationstemperaturen (Td), die für eine beliebige der Sonden berechnet worden war.

Nach der Hybridisierung wurden die Filter dreimal für zehn Minuten auf einem Schüttler mit 6 × SSC — 0,1% SDS bei Raumtemperatur gewaschen sowie zwei — bis dreimal mit 6 × SSC — 1% SDS bei der Hybridisierungstemperatur (48°C) gewaschen. Die Autoradiografie der Filter zeigte sieben positive Klone unter 200000 gescreenten Kolonien.

Die Initialsequenzanalyse eines der vermeintlichen Affen-cDNA-Klone (als Klon 83 bezeichnet) erfolgte aus Gründen der Überprüfung auf Richtigkeit durch eine Modifikation der Verfahrensweise von Wallace et al., *Gene*, **16**, S. 21–26 (1981). Dabei wurde die Plasmid-DNA des Affen-cDNA-Klons 83 linearisiert durch Abbau mit **Eco** RI und durch Erhitzen im siedenden Wasserbad denaturiert. Die Nucleotidsequenz wurde mittels des Dideoxy-Verfahrens von Sanger et al., *P. N. A. S. (USA)*, **74**, S. 5463–5467 (1977) bestimmt. Ein Teilsatz des EPV-Gemisches der Sonden, bestehend aus 16 Sequenzen, wurde als Primer für die Sequenzierungsreaktionen verwendet.

C. Sequenzieren von Affen-EPO-cDNA

Die Nucleotid-Sequenzanalyse des Klons 83 wurde mittels des Verfahrens von Messing, *Methods in Enzymology*, **101**, S. 20–78 (1983) durchgeführt. Aus Tabelle IV ist eine vorläufige Restriktionsmappenanalyse des annähernd 1600 Basenpaar-**Eco** RI/**Hind** III-klonierten Fragments des Klons 83 zu entnehmen. Die annähernden Ortungen der Restriktionsendonuclease-enzym-Erkennungsstellen erfolgten mit Hilfe der Anzahl von Basen 3' zur **Eco** RI-Stelle am 5'-Ende des Fragments. Das Nucleotidsequenzieren erfolgte durch Sequenzen der individuellen Restriktionsfragmente mit der Absicht, überlappende Fragmente anzupassen. Beispielsweise bestimmte die Analyse der Nucleotide in einem Restriktionsfragment die Überlappung der Sequenzinformation C 113 (**Sau** bei ~ 111/**Sma** I bei ~ 324), und das reverse Ordersequenzieren eines Fragments C 73 (**Alu** I bei ~ 424/**Bst** II bei ~ 203).

Tabelle IV

Restriktionsenzym-Erkennungsstelle	Ungefähre(r) Ort(e)
Eco RI	1
Sau 3A	111
Sma I	180
Bst II	203
Sma I	324
Kpn I	371

Fortsetzung der Tabelle IV

Restriktionsenzym-Erkennungsstelle	Ungefähre(r) Ort(e)
RsaI	372
AluI	424
PstI	426
AluI	430
HpaI	466
AluI	546
PstI	601
PvuII	604
AluI	605
AluI	782
AluI	788
RsaI	792
PstI	807
AluI	841
AluI	927
NcoI	946
Sau3A	1014
AluI	1072
AluI	1115
AluI	1223
PstI	1301
RsaI	1343
AluI	1384
HindIII	1449
AluI	1450
HindIII	1585

Das Sequenzieren von annähernd 1342 Basenpaaren (innerhalb der von der **Sau 3A**-Stelle 3' bis zur **Eco RI**-Stelle und der **Hind III**-Stelle überspannten Region) und die Analyse der Ableseraster gestattete die Aufstellung der DNA- und Aminosäuresequenzinformationen, die in der Tabelle V enthalten sind. In der Tabelle wurde der vermutliche Anfangsaminosäurerest des Aminoterminals der reifen EPO (auf Richtigkeit überprüft durch Korrelation der vorher genannten Sequenzanalyse von zwanzig Aminoterminalesten) mit der Ziffer +1 bezeichnet. Die Anwesenheit eines Methioninspezifizierenden ATG-Codons (bezeichnet als -27) „oberhalb“ (upstream) des anfänglichen Aminoterminalesten als des ersten Restes, der für die Aminosäuresequenz des reifen Proteins bezeichnet ist, zeigt die Wahrscheinlichkeit an, daß EPO anfänglich im Cytoplasma in einer Precursorform exprimiert wurde mit einer 27 Aminosäuren-„leader“-Region, die vor dem Eintritt des reifen EPO in den Kreislauf herausgeschnitten wird. Potentielle Glykosylierungsstellen innerhalb des Polypeptids sind durch Sternchen gekennzeichnet. Die geschätzte Molekülmasse der translatierten Region betrug 21 117 Daltons und die Molekülmasse der 165 Reste des Polypeptids, das reife Affen-EPO bildete, wurde mit 18236 Daltons bestimmt.

Tabelle V
Translation von Affen-EPO cDNA
Sau 3A

GATCCGCGCCCCCTGGACAGCGCCCTCTCCTCCAGGCCCGTGGGGCTGGCCCTGCCCCGCTGAAC TTCCCGGGATGAGGACT
CCCGGTGTGGTACCGCGCGCCTAGGTCGCTGAG

						-27 Met	Gly	Val	His	Glu	Cys	Pro	-20 Ala	Trp	
GGACCCCGGCCAGGCGCGGAGATG						GGG	GTG	CAC	GAA	TGT	CCT	GCC	TGG		
								-10							
Leu	Trp	Leu	Leu	Leu	Ser	Leu	Val	Ser	Leu	Pro	Leu	Gly	Leu	Pro	
CTG	TGG	CTT	CTC	CTG	TCT	CTC	GTG	TCG	CTC	CCT	CTG	GGC	CTC	CCA	
		-1	+1									10			
Val	Pro	Gly	Ala	Pro	Pro	Arg	Leu	Ile	Cys	Asp	Ser	Arg	Val	Leu	
GTC	CCG	GGC	GCC	CCA	CCA	CGC	CTC	ATC	TGT	GAC	AGC	CGA	GTC	CTG	
							20								
											x				
Glu	Arg	Tyr	Leu	Leu	Glu	Ala	Lys	Glu	Ala	Glu	Asn	Val	Thr	Met	
GAG	AGG	TAC	CTC	TTG	GAG	GCC	AAG	GAG	GCC	GAG	AAT	GTC	ACG	ATG	
		30										40			
										x					
Gly	Cys	Ser	Glu	Ser	Cys	Ser	Leu	Asn	Glu	Asn	Ile	Thr	Val	Pro	
GGC	TGT	TCC	GAA	AGC	TGC	AGC	TTG	AAT	GAG	AAT	ATC	ACC	GTC	CCA	
							50								
Asp	Thr	Lys	Val	Asn	Phe	Tyr	Ala	Trp	Lys	Arg	Met	Glu	Val	Gly	
GAC	ACC	AAA	GTT	AAC	TTC	TAT	GCC	TGG	AAG	AGG	ATG	GAG	GTC	GGG	
		60										70			
Gln	Gln	Ala	Val	Glu	Val	Trp	Gln	Gly	Leu	Ala	Leu	Leu	Ser	Glu	

Tabelle V (Forts.)

CAG	CAG	GCT	GTA	GAA	GTC	TGG	CAG 80	GGC	CTG	GCC	CTG	CTC	TCA	GAA
Ala	Val	Leu	Arg	Gly	Gln	Ala	Val	Leu	Ala	x Asn	Ser	Ser	Gln	Pro
GCT	GTC	CTG 90	CGG	GGC	CAG	GCC	GTG	TTG	GCC	AAC	TCT	TCC 100	CAG	CCT
Phe	Glu	Pro	Leu	Gln	Leu	His	Met	Asp	Lys	Ala	Ile	Ser	Gly	Leu
TTC	GAG	CCC	CTG	CAG	CTG	CAC	ATG 110	GAT	AAA	GCC	ATC	AGT	GGC	CTT
Arg	Ser	Ile	Thr	Thr	Leu	Leu	Arg	Ala	Leu	Gly	Ala	Gln	Glu	Ala
CGC	AGC	ATC 120	ACC	ACT	CTG	CTT	CGG	GCG	CTG	GGA	GCC	CAG 130	GAA	GCC
Ile	Ser	Leu	Pro	Asp	Ala	Ala	Ser	Ala	Ala	Pro	Leu	Arg	Thr	Ile
ATC	TCC	CTC	CCA	GAT	GCG	GCC	TCG 140	GCT	GCT	CCA	CTC	CGA	ACC	ATC
Thr	Ala	Asp	Thr	Phe	Cys	Lys	Leu	Phe	Arg	Val	Tyr	Ser	Asn	Phe
ACT	GCT	GAC 150	ACT	TTC	TGC	AAA	CTC	TTC	CGA	GTC	TAC	TCC 160	AAT	TTC
Leu	Arg	Gly	Lys	Leu	Lys	Leu	Tyr	Thr	Gly	Glu	Ala	Cys	Arg	Arg
CTC	CGG	GGA 165	AAG	CTG	AAG	CTG	TAC	ACG	GGG	GAG	GCC	TGC	AGG	AGA

Gly Asp Arg OP
 GGG GAC AGA TGA CCAGGTGCGTCCAGCTGGGCGACATCCACCACCTCCCTCACCAACACTGCCTGTGCCACACCCCTCCC
 CACCACTCCCCGAACCCCATCGAGGGGCTCTCAGCTAAGCGCCAGCCTGTCCCATGGACACTCCAGTGCCAGCAATGACATCTCAG
 GGCCAGAGGAACTGTCCAGAGCACAACCTCTGAGATCTAAGGATGTGCGAGGCCAACTTGAGGGCCAGAGCAGGAAGCATTCA
 GAGAGCAGCTTTAAACTCAGGAGCAGAGACAATGCAGGGAAAACACCTGAGCTCACTCGGCCACCTGCAAAAGATGCAGGA
 CACGCTTTGGAGGCAATTTACCTGTTTTGCACCTACCATCAGGGACAGGATGACTGGAGAAGCTTAGGTGGCAAGCTGTGACTT
 CTCAGGCCTCAGGGGCACTCCCTTGGTGGCAAGAGCCCCCTTGACACTGAGAGAATATTTTGCAATCTGCAGCAGGAAAAATTA
 CGGACAGGTTTTGGAGGTTGGAGGGTACTTGACAGGTGTGTGGGGAAGCAGGGCGGTAGGGGTGGAGCTGGGATGCGAGT
 GAGAACCGTGAAGACAGGATGGGGGCTGGCCTCTGTTCTCGTGGGGTCCAAGCTT

HindIII

Die Polypeptidsequenz von Tabelle V kann auf analytischem Wege leicht auf das Vorhandensein hydrophiler Bereiche und/oder sekundärer übereinstimmender Charakteristika überprüft werden die potentiell hoch immunogene Bereiche anzeigen, z. B. durch die Methoden von Hopp et al., *P.N.A.S. (USA)*, **78**, S. 3824–3828 (1981) und Kyte et al., *J. Mol. Biol.* **157**, S. 105–132 (1982) und/oder Chou et al., *Biochem.*, **13**, S. 222–245 (1974) und *Advances in Enzymology*, **47**, S. 45–47 (1978). Die Computergestützte Analyse nach der Methode von Hopp et al. ist über ein Programm möglich, das als PEP Reference Section 6.7 bezeichnet wird und über Intelligenetics, Inc. 124 University Avenue, Palo Alto, Californien, erhältlich ist.

Beispiel 4**A. Human-Genbibliothek**

Eine Ch4A Phagen-geborene menschliche fötale Leber-Genbibliothek, hergestellt nach den Verfahren von Lawn et al., *Cell*, **18**, S. 533–543 (1979), wurde erhalten und zur Verwendung in einer Plaque-Hybridisierungsassay bereitgestellt.

B. Plaque-Hybridisierungsverfahren zum Screening einer Human-Genbibliothek

Phagenteilchen wurden lysiert und die DNA's auf Filtern fixiert (50 000 Plaques pro Filter) nach den Verfahrensweisen von Woo, *Methods in Enzymology*, **68**, S. 389–395 (1979), ausgenommen für die Verwendung von Gene-Screene Plus-Filter (New England Nuclear Catalog Nr. NEF-976) und NZYAM-Platten (NaCl, 5g; MgCl₂·6H₂O, 2g; NZ-Amin A, 10g; Hefeextrakt, 5g; Casaminoäuren, 2g; Maltose, 2g; und Agar, 15g/l).

Die luftgetrockneten Filter wurden für eine Stunde bei 80°C wärmebehandelt und im Anschluß daran mit Proteinase K wie in Beispiel 3, Teil B, abgebaut. Die Prä-Hybridisierung wurde mit 1 M NaCl — 1% SDS-Puffer bei 55°C über 4 Stunden oder mehr durchgeführt, wonach der Puffer entfernt wurde. Hybridisierungs- und Nachhybridisierungswaschungen wurden wie in Beispiel 3, Teil B beschrieben, durchgeführt. Sowohl das Gemisch von als EPV bezeichneten 128 20-mer-Sonden als auch das Gemisch von 128 17-mer-Sonden aus Tabelle III (bezeichnet als EPQ-Gemisch) wurde eingesetzt. Die Hybridisierung wurde bei 48°C unter Verwendung des EPV-Sondengemisches durchgeführt. Die EPQ-Sondergemisch-Hybridisierung wurde bei 46°C durchgeführt — 4 Grad unterhalb der niedrigsten berechneten T_d für Bestandteile des Gemisches. Die Entfernung der Hybridisierungssonde für die Rehybridisierung erfolgte durch Kochen mit 1 × SSC-0,1% SDS für zwei Minuten. Die Autoradiografie der Filter zeigte drei positive Klone (mit beiden Sondengemischen reaktionsfähig) von 1 500 000 gescreenten Phagenplaques. Die Überprüfung auf Richtigkeit der positiven Klone als EPO-verschlüsselnd erhielt man über DNA-Sequenzieren und elektronenmikrografische Sichtbarmachung der Heteroduplex-Bildung mit der Affen-cdNA von Beispiel 3. Dieses Verfahren gab auch über multiple Introns in der genomischen DNA-Sequenz Aufschluß.

Beispiel 5

Es wurde eine Nucleotid-Sequenzanalyse eines der positiven Klone (bezeichnet als λhE1) durchgeführt. Die erhaltenen Ergebnisse sind in Tabelle VI aufgeführt.

Tabelle VI

AAGCTTCTGGGCTTCCAGACCCAGCTACTTTGCGGAACCTCAGCAACCCAGGCATCTCTGAGTCTCCGCCCAAGACCGGGATGCC
 CCCCAGGGGAGGTGTCCGGGAGCCAGCCTTTCCAGATAGCAGCTCCGCCAGTCCCAAGGGTGCACAACCGGTGCACT
 CCCCTCCCGCAGCCAGGGCCCGGAGGAGCCCCCATGACCCACACGCACGTCTGCAGCAGCCCCGCTCAGCCCCGGCGAGC
 CTCAACCCAGGCGTCTGCTGCTGCTGACCCCGGGTGGCCCTACCCCTGGCGACCCCTCAGCAGACAGCCTCTCCCCACC
 CCCACCCGCGCAGCAGACATGCAGATAACAGCCCCGACCCCGGCCAGAGCCGXAGAGTCCCTGGGCCACCCCGGCCGCTCG
 CCTGCCGCTGCGCCGACCCGCTGTCTCCCGAGCCGACCGGGGCCACCGCGCCXGCTCTGCTCCGACACCGCGCCCTT
 GGACAGCCGCCCTCTCTCTAGGCCCGTGGGGCTGGCCCTGCACCGCGAGCTTCCCGGGATGAGGX

CCCAGTGACCGCGCGCCCCAAGTCGCTGAGGGACCCCGGCCAAGCGCGGAG ATG GGG GTG CAC G
 GTGAGTACTCGCGGGCTGGGCGCTCCCGCGGGCGGGTCTGTTTGTAGCGGGGATTTAGCGCCCGGCTATTGGCCAAGAGG
 TGGCTGGGTTCAAGGACCGCGCACTTGTCAAGGACCCCGAAGGGGGAGGGGGTGGGGCAGCTCCAGCTGCCGCGGGGA
 CTTGGGGGAGTTCTTGGGGATGGCAAAAACCTGGCCTGTTGAGGGGCACAGTTTGGGGTTGGGGAGGAGTTTGGGGTTCTG
 CTGTGCAGTTGTGCTGTTGTCACTGTCTCG I.S. TTGCACACGCACAGATCAATAAGCCAGAGGCAGCACCTGAGTGCTTGCATG
 GTTGGGACAGGAAGGACGAGCTGGGGCAGAGACGTGGGGATGAAGGAAGCTGTCTTCCACAGCCACCTTCTCCCCCCC
 GCCTGACTCT

CAGCCTGGCTATCTGTTCTAG
 -10
 Leu Ser -Leu Pro Leu Gly Leu Pro Val Leu Gly Ala Pro Pro Arg Leu Ile Cys
 CTG TCG CTC CCT CTG GGC CTC CCA GTC CTG GGC GCC CCA CCA CGC CTC ATC TGT
 10
 Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile
 GAC AGC CGA GTC CTG GAG AGG TAC CTC TTG GAG GCC AAG GAG GCC GAG AAT ATC
 26
 Thr
 ACG GTGAGACCCCTTCCCCAGCACATTCCACAGAACTCACGCTCAGGGCTTCAGGGAACCTCTCCAGATCCAGGAACCTGG
 CACTTGGTTTGGGGTGGAGTTGGGAAGCTAGACACTGCCCCCTACATAAGAATAAGTCTGGTGGCCCCAAACCATACTGAAA
 CTAGGCAAGGAGCAAAGCCAGCAGATCCTACGCCTGTGGGCCAGGG

CCAGAGCCTTCAGGGACCCCTTGACTCCCCGGGCTGTGTGCATTTCAG
 40
 His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr
 CAC TGC AGC TTG AAT GAG AAT ATC ACT GTC CCA GAC ACC AAA GTT AAT TTC TAT
 50
 Ala Trp Lys Arg Met Glu
 GCC TGG AAG AGG ATG GAG GTGAGTTCCTTTTTTTTTTTTTCTTTCTTTGGAGAATCTCATTTGCGAGCCTGAT
 TTTGGATGAAAGGGGAGAAATGATCGGGGGAAGGTAATAATGGAGCAGCAGAGATGAGGCTGCCTGGGCGCAGAGGGCTCACGTC
 TATAATCCCAGGCTGAGATGGCCGAGATGGGAGAATTGCTTGAGCCCTGGAGTTTCAGACCAACCTAGGCGAGCATAGTGAGATC
 CCCCATCTCACAACATTTAAAAAATTAGTCAGGTGAAGTGGTGATGGTGGTAGTCCAGATATTTGGAAGGCTGAGGCGGG
 AGGATCGCTTGAGCCAGGAATTTGAGGCTGCAGTGAGCTGTGATCACACCACTGCACTCCAGCCTCAGTGACAGAGTGAGGCC
 CTGTCTCAAAAAGAAAAGAAAAGAAAATAATGAGGGCTGTATGGAATACATTCAATTATTCACTCACTCACTCACTCATT
 CATTCACT
 AGGGTGACATGGGTGAGCTCGACTCCAGAGTCCACTCCCTGTAG

56 60 70
 Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala
 GTC GGG CAG CAG GCC GTA GAA GTC TGG CAG GGC CTG GCC CTG CTG TCG GAA GCT
 80
 Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu
 GTC CTG CGG GGC CAG GCC CTG TTG GTC AAC TCT TCC CAG CCG TGG GAG CCC CTG
 100
 Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu
 CAG CTG CAT GTG GAT AAA GCC GTC AGT GGC CTT CGC AGC CTC ACC ACT CTG CTT
 110 115
 Arg Ala Leu Gly Ala Gln
 CGG GCT CTG GGA GCC CAG GTGAGTAGGAGCGGACACTTCTGCTTGCCCTTTCTGTAAGAAGGGGAGAAGGGTC
 TTGCTAAGGAGTACAGGAAGTGTCCGTATTCTTCCCTTTCTGTGGCACTGCAGCGACCTCCT
 116 120
 Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
 AAG GAA GCC ATC TCC CCT CCA GAT GCG GCC TCA GCT GCT
 GTTTTCTCCTTGGCAG

[illegible]

In Tabelle VI bezeichnet die anfängliche fortlaufende DNA-Sequenz einen Kopfstrang von 620 Basen, in dem offensichtlich eine untranslatierte Sequenz unmittelbar einem translatierten Teil des Human-EPO-Gens vorausgeht. Insbesondere erscheint es so, daß die Sequenz das 5' Ende des Gens umfaßt, das zu einer translatierten DNA-Region überleitet, die für die ersten vier Aminosäuren (–27 bis –24) einer Führungssequenz („Pre-Sequenz“) codiert. Vier Basenpaare in der Sequenz vor der die den Anfang der Führungssequenz verschlüsselt, wurden nicht genau bestimmt und sind daher als „X“ bezeichnet worden. Es folgt dann ein Intron von etwa 639 Basenpaaren (439 Basenpaare davon wurden sequenziert und die restlichen 200 Basenpaare davon wurden als „I.S.“ bezeichnet) und unmittelbar davor ein Codon für Glutamin, das als Rest –23 des translatierten Polypeptids bezeichnet wurde. Die unmittelbar folgende Exonsequenz wird als codierend für Aminosäurereste durch einen Alaninrest angesehen (bezeichnet als der +1-Rest der Aminosäuresequenz von reifem Human-EPO) zu dem Codon, das Threonin an Position +26 spezifiziert, wonach ein zweites Intron folgt, das aus 256 Basen besteht, wie sie spezifisch bezeichnet sind. Diesem Intron folgt eine Exonsequenz für die Aminosäurereste 27 bis 55 und danach beginnt ein 612 Basenpaare enthaltendes drittes Intron.

Das anschließende Exon codiert für die Reste 56 bis 115 das Human-EPO und dann beginnt ein viertes Intron von 134 Basen, wie aufgeführt. Dem vierten Intron folgt ein Exon, das für die Reste Nr. 116 bis 166 codiert und ein Stopcodon (TGA). Schließlich wird eine Sequenz von 568 Basenpaaren identifiziert, worin eine untranslatierte 3'-Region des Human-EPO-Gens erscheint, zwei Basenpaare, von denen („X“) bisher noch nicht genau sequenziert worden ist.

Tabelle VI dient somit zur Ermittlung der Strukturübereinstimmung (Aminosäuresequenz) von reifem Human-EPO, das 166 spezifische Aminosäurereste enthält (geschätzte Molmasse = 18399). Außerdem erscheint in der Tabelle die DNA-Sequenz, die für eine Führungssequenz mit 27 Resten zusammen mit 5'- und 3'-DNA-Sequenzen codiert, die für Promoter/Operatorfunktionen des Humangen-Operons signifikant sein können.

Stellen für die potentielle Glycosylierung des reifen Human-EPO-Polypeptids sind in der Tabelle durch Sternchen bezeichnet. Es sollte beachtet werden, daß die spezifische Aminosäuresequenz von Tabelle VI wahrscheinlich jene einer natürlich auftretenden allelischen Form von Human-Erythropoietin bildet. Diese Position wird durch die Ergebnisse der fortlaufenden Bemühungen beim Sequenzieren von Urinisolation von Human-Erythropoietin gestützt, das dazu führte, daß man eine bedeutende Anzahl von Erythropoietinmolekülen darin herausfand, die ein Methionin am Rest 126 gegenüber einem Serin aufwiesen, wie in der Tabelle gezeigt.

Die folgende Tabelle VII erläutert den Grad der Polypeptidsequenz-Homologie zwischen Human- und affen-EPO. In der oberen fortlaufenden Reihe der Tabelle wurden Einzelbuchstabenbezeichnungen verwendet, um die hergeleiteten, translatierten Polypeptidsequenzen von Human-EPO darzustellen, beginnend mit dem Rest -27, und die untere fortlaufende Reihe zeigt die hergeleitete Polypeptidsequenz von Affen-EPO, beginnend mit dem als Nummer -27 bezeichneten Rest. Sternchen zeigen die Sequenzhomologien an. Es sollte festgehalten werden, daß die hergeleiteten Human- und Affen-EPO-Sequenzen einen „zusätzlichen“ Lysin(K)-rest an der (Human-)Position 116 aufweisen. Ein Vergleich mit Tabelle VI zeigt, daß dieser Rest am Rande einer vermutlichen mRNA-Spleißverbindung in der genomischen Sequenz auftritt. Das Vorhandensein des Lysinrestes in der Human-Polypeptidsequenz wurde darüber hinaus durch Sequenzieren eines cDNA-Humansequenzklons überprüft, hergestellt aus mRNA, die aus COS-1 Zellen isoliert worden waren, die mit der humangenomischen DNA von Beispiel 7 (siehe unten) transformiert wurden.

Tabelle VII
Vergleich von Human- und Affen-EPO-Polypeptiden

		-20	-10	+1	10	20	30	40	
Human		MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITTCGAEHCSLNENITVPDTK							
		XXXXXXXXXXXXXXXXXX	XXXXXXX	XXXXXXXXXXXXXXXXXXXXXXXXXXXX		XXXXXX	X	XXXXXXXXXXXXXXXXXX	
Affen		MGVHECPAWLWLLLSLVSLPLGLPVPGAPPRLICDSRVLERYLLEAKEAENVMTGCSESCSLNENITVPDTK							
		50	60	70	80	90	100	110	
Human		VNIFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLRLALGAQKE							
		XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		X	XXXXX	XXXXXX	XXX	XXXXX XXXXXXXXXXXXX	

Tabelle VII (Fortsetzung)

Affen	VNFWYAWKRMEVGGQAVEVWQGLALLSEAVLRGQAVLANSSQPFEPQLHMDKAISGLRSITTLRLALGAQ-E
	120 130 140 150 160
Human	AISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKCLKLYTGEACRTGDR
	xxx xxxxxxxxxxxxxxxxxxxx xxxxxxxxxxxxxxxxxxxxxxxxxxxx xxx
Affen	AISLPDAASAAPLRTITADTFCKLFRVYSNFLRGKCLKLYTGEACRRGDR

Ein Vergleich der Aktivität der rekombinanten Affen- und Human-EPO-Materialien gegen einen Standard einer natürlich auftretenden Human-EPO und die Ergebnisse sind in grafischer Form in Fig. 1 aufgeführt. Kurz zusammengefaßt offenbaren die Ergebnisse erwartetermaßen, daß das rekombinante Affen-EPO signifikant mit Anti-Human-EPO-Antikörper konkurrierte, obgleich es nicht in der Lage war, die Bindung unter den Testbedingungen vollständig zu inhibieren. Der höchste Inhibierungsprozensatz für rekombinante Human-EPO lag allerdings nahe etwa bei dem des Human-EPO-Standards. Die parallele Lage der Dosisreaktionskurven deutet auf immunologische Identität der gemeinsamen Sequenzen (Epitope) hin.

Beispiel 11

Das vorliegende Beispiel betrifft die gesamte Herstellung durch Zusammenstellung der Nucleotidbasen von zwei strukturellen Genen, die die Human-EPO-Sequenz von Tabelle VI verschlüsseln und schließt entsprechend „bevorzugte“ Codons für die Expression in E. coli- und Hefe(s. *cevevisiae*)zellen ein.

Beschrieben wird auch die Konstruktion von Genen, die Analoge von Human-EPO verschlüsseln. Kurz zusammengefaßt entsprach der Ablauf dem in der bereits vorher genannten Literaturstelle von Alton et al., (WO 83/04053). Die Gene wurden für eine anfängliche Zusammenstellung der Komponenten-Oligonucleotide in mehrfachen Verdoppelungen gestaltet, die der Reihe nach in drei getrennte Abschnitte (section) vereinigt wurden. Diese Sectionen wurden für eine fertige Amplifizierung gestaltet, und nach Entfernung aus dem Amplifizierungssystem konnten sie sequentiell oder durch eine mehrfache Fragmentligation in einem geeigneten Expressionsvektor vereinigt werden.

Die folgenden Tabellen VIII bis XIV erläutern die Gestaltung und Vereinigung eines hergestellten Gens, das ein Human-EPO-Translationsprodukt verschlüsselt, dem eine beliebige Führungs- oder Presequenz fehlt, zu dem jedoch ein anfänglicher Methioninrest an Position -1 gehört. Darüber hinaus gehören zu dem Gen in seinem wesentlichen Teil E. coli-Präferenzcodons, und die Konstruktion wurde daher als das „ECEPO“-Gen bezeichnet.

Tabelle VIII

ECEPO SECTION 1 OLIGONUCLEOTIDES

1. AATTCTAGAAACCATGAGGGTAATAAAATA
2. CCATTATTTTATTACCCTCATGGTTTCTAG
3. ATGGCTCCGCCGCGTCTGATCTGCGAC
4. CTCGAGTCGCAGATCAGACGCGGCGGAG
5. TCGAGAGTTCTGGAACGTTACCTGCTG
6. CTTCCAGCAGGTAACGTTCCAGAACT
7. GAAGCTAAAGAAGCTGAAAACATC
8. GTGGTGATGTTTTCAGCTTCTTTAG
9. ACCACTGGTTGTGCTGAACACTGTTT
10. CAAAGAACAGTGTTTCAGCACAAACCA
11. TTTGAACGAAAACATTACGGTACCG
12. GATCCGGTACCGTAATGTTTTCGTT

Tabelle IX
ECEPO SECTION 1

XbaI

EcoRI

AAATTCAG AAACCATGAG¹ GGTAATAAAA TAATGGCTCC³ GCCGCTCTG
GATC TTTGGTACTC² CCATTATTTT ATTACGAGG⁴ CGGCGCAGAC

ATCTGCCACT⁵ CGAGAGTTCT GGAACGTTAC CTGCTGGAAG⁶ CTAAAGAAGC
TAGACGCTGA GCTCTCAAGA CCTTGCAATG GACGACCTTC⁷ GATTTCTTCG

TGAAAACATC⁸ ACCACTGGTT⁹ GTGCTGAACA CTGTTCTTTG¹⁰ AACGAAAACA¹¹
ACTTTTGTAG TGGTGACCAA CACGACTTGT GACAAGAAAC¹² TTGCTTTTGT

KpnI BamHI

TTACGGTACC G
AATGCCATGG CCTAG

12

Tabelle X
ECEPO SECTION 2 OLIGONUCLEOTIDES

1. AATTCGGTACCAGACACCAAGGT
2. GTTAACCTTGGTGTCTGGTACCG
3. TAACTTCTACGCTTGGAACGTAT
4. TTCCATACGTTTCCAAGCGTAGAA
5. GGAAGTTGGTCAACAAGCAGTTGAAGT
6. CCAAACCTTCAACTGCTTGTTGACCAAC
7. TTGGCAGGGTCTGGCACTGCTGAGCG
8. GCCTCGCTCAGCAGTGCCAGACCCTG
9. AGGCTGTAAGTGGTGGCCAGGCA
10. GCAGTGCCTGGCCACGCAGTACA
11. CTGCTGGTAACTCCTCTCAGCCGT
12. TTCCACGGCTGAGAGGAGTTTACCA
13. GGGAACCGCTGCAGCTGCATGTTGAC
14. GCTTTGTCAACATGCAGCTGCAGCGG
15. AAAGCAGTATCTGGCCTGAGATCTG
16. GATCCAGATCTCAGGCCAGATACT

Tabelle XI
ECEPO SECTION 2

EcoRI KpnI

A ATTCGGTACC AGACACCAAG GTTAACTTCT ACGCTTGGAA ACGTATGGAA
GCCATGG TCTGTGGTTC CAATTGAAGA TCGGAACCTT TGCATACCTT

GTTGGTCAAC AAGCAGTTGA AGTCTGGCAG GGTCTGGCAC TGCTGAGCCTA
CAACCAGTTG TTCGTCAACT TCAAACCGTC CCAGACCGTG ACGACTCGCT

GGCTGTACTG CGTGGCCAGG CACTGCTGGT AAATCCTCT CAGCCGTGGG
CCGACATGAC GCACCGGTCC GTGACGACCA TTTGAGGAGA GTCGGCACCC

AACCGCTGCA GCTGCATGTT GACAAAGCAG TATCTGGCCT GAGATCTG
TTGGCGACGT CGACGTACAA CTGTTTCGTC ATAGACCGGA CTCTAGACCTAC

Tabelle XII
ECEPO SECTION 3

1. GATCCAGATCTCTGACTACTCTGC
2. ACGCAGCAGAGTAGTCAGAGATCTG
3. TCGTGCTCTGGGTGCACAGAAAGAGG
4. GATAGCCTCTTTCTGTGCACCCAGAGC
5. CTATCTCTCCGCCGGATGCTGCATCT
6. CAGCAGATGCAGCATCCGGCGGAGA
7. GCTGCACCGCTGCGTACCATCACTG
8. ATCAGCAGTGATGGTACGCAGCGGTG
9. CTGATACCTTCCGCAAACTGTTTCG
10. ATACACGAAACAGTTTGCGGAAGGT
11. TGTATACTCTAACTTCTGCGTGGTA
12. CAGTTTACCACGCAGGAAGTTAGAGT
13. AACTGAACTGTATACTGGCGAAGC
14. GGCATGCTTCGCCAGTATACAGTTT
15. ATGCCGTACTGGTGACCGCTAATAG
16. TCGACTATTAGCGGTCACCAGTAC

Tabelle XIII
ECEPO SECTION 3

BamHI BglII

GA TCCAGATCTCTG

GTCTAGAGAC

ACTACT¹CTGC | TGGTGTCTCT ³GGGTGCACAG AAAGAGGCTA ⁵TCTCTCCGCC
TGATGAGACG ²ACGCACGAGA CCCACGTGTC ⁴TTTCTCCGAT AGAGAGGCCG

GGATGCTGCA TCT⁷GCTGCAC CGCTGCGTAC CATCACTGCT ⁹GATACCTTCC
CCTACGACGT ⁶AGACGACGTG GCGACGCATG ⁸GTAGTGACGA CTATGGAAGG

GCAAACTGTT TCG¹¹GTATAC TCTAACTTCC TGCGTGGTAA ¹³ACTGAAACTG
CGTTTGACAA ¹⁰AGCACATATG AGATTGAAGG ¹²ACGCACCATT TGACTTTGAC

TATACTGGCG ¹⁵AAGCATGCCG TACTGGTGAC ^{Sal}I CGCTAATAG
ATATGACCGC ¹⁴TTCGTACGGC ¹⁶ATGACCACTG GCGATTATC AGCT

Tabelle XIV
ECEPO GENE

XbaI			-1 1 MetAla	
CTAG	AAACCATGAG TTTGGTACTC	GGTAATAAAA CCATTATTTT	TAATGGCTCC ATTACCGAGG	GCCGCGTCTG CGGCGCAGAC
ATCTGCGACT TAGACGCTGA	CGAGAGTTCT GCTCTCAAGA	GGAACGTTAC CCTTGCAATG	CTGCTGGAAG GACGACCTTC	CTAAAGAAGC GATTTCTTCG
TGAAAACATC ACTTTTGTAG	ACCACTGGTT TGGTGACCAA	GTGCTGAACA CACGACTTGT	CTGTTCTTTG GACAAGAAAC	AACGAAAACA TTGCTTTTGT
TTACGGTACC AATGCCATGG	AGACACCAAG TCTGTGGTTC	GTAACTTCT CAATTGAAGA	ACGCTTGGAA TGCGAACCTT	ACGTATGGAA TGCATACCTT
GTTGGTCAAC CAACCACTTG	AAGCAGTTGA TTCGTCAACT	AGTTTGCCAG TCAAACCGTC	GGTCTGGCAC CCAGACCGTG	TGCTGAGCGA ACGACTCGCT
GGCTGTAAGT CCGACATGAC	CGTGGCCAGG GCACCGGTCC	CACTGCTGGT GTGACGACCA	AAACTCCTCT TTTGAGGAGA	CAGCCGTGGG GTCGGCACCC
AACCGCTGCA TTGGCGACGT	GCTGCATGTT CGACGTACAA	GACAAAGCAG CTGTTTCGTC	TATCTGGCCT ATAGACCGGA	GAGATCTCTG CTCTAGAGAC
ACTACTCTGC TGATGAGACG	TGCGTGTCT ACGCACGAGA	GGGTGCACAG CCCACGTGTC	AAAGAGGCTA TTTCTCCGAT	TCTCTCCGCC AGAGAGGCCG
GCATGCTGCA CCTACGACGT	TCTGCTGCAC AGACGACGTG	CGCTGCGTAC GCGACGCATG	CATCACTGCT GTAGTGACGA	GATACCTTCC CTATGGAAGG

GCAAACTGTT TCGTGTATAC TCTAACTTCC TCGTGGTAA ACTGAACTG
CGTTTGACAA AGCACATATG AGATTGAAGG ACGCACCATT TGACTTTGAC

Sal I

TATACTGGCG AAGCATGCCG TACTGGTGAC CGCTAATAG
ATATGACCGC TTCGTACGGC ATGACCACTG GCGATTATCA GCT

Im Detail sind aus Tabelle VIII Oligonucleotide zu entnehmen, die dazu eingesetzt wurden, die Section 1 des ECEPO-Gens zu erzeugen, das die Aminoterminalreste menschlichen Polypeptids verschlüsselt. Oligonucleotide wurden zu Doppelungen vereinigt (1 und 2, 3 und 4 usw.), und die Doppelungen wurden dann ligiert, um zu der ECEPO Section 1 wie in Tabelle IX zu gelangen. Bemerkt werden muß, daß zu der vereinigten Section terminale einzelsträngige **EcoRI**- und **BamHI**-Enden gehören, daß „stromabwärts“ des einzelsträngigen **EcoRI**-Endes eine **XbaI**-Restriktionsenzym-Erkennungsstelle ist; und daß „stromaufwärts“ von einzelsträngigem **BamHI**-Ende eine **KpnI**-Erkennungsstelle ist. Section 1 konnte leicht unter Verwendung des M13 Phagenvektors amplifiziert werden, der zur Überprüfung der Sequenz der Section eingesetzt wurde. Einige Schwierigkeiten traten auf beim Isolieren der Section als **XbaI/KpnI**-Fragment aus RF DNA, in *E. coli* erzeugt, wahrscheinlich, wegen der Methylierung der **KpnI**-Erkennungsstellenbasen innerhalb des Wirtes. Es wurde daher einzelsträngige Phagen-DNA isoliert und in die doppelsträngige Form in vitro übersetzt mittels Primerverlängerung (extension), und das gewünschte doppelsträngige Fragment wurde danach leicht isoliert.

Die ECEPO-Gensectionen 2 und 3 (Tabellen XI und XIII) wurden in ähnlicher Weise aus den Oligonucleotiden der Tabellen X und XII konstruiert. Jede Section wurde in dem M13 Vektor amplifiziert, der für die Sequenzüberprüfung auf Richtigkeit eingesetzt wurde, und wurde aus Phagen DNA isoliert.

Wie aus Tabelle XI ersichtlich, wurde die ECEPO-Section 2 mit einzelsträngigen **EcoRI**- und **BamHI**-Enden konstruiert und konnte als ein **KpnI/BglII**-Fragment isoliert werden. In ähnlicher Weise wurde die ECEPO-Section 3 mit einzelsträngigen **BamHI**- und **SalI**-Enden hergestellt und konnte aus Phagen RF DNA als ein **BglII-SalI**-Fragment isoliert werden. Die auf diese Weise hergestellten drei Sectionen konnten leicht zu einer fortlaufenden DNA-Sequenz vereinigt werden (Tabelle XIV), die das gesamte menschliche EPO-Polypeptid mit einem Aminoterminal-Methionincodon (ATG) für den *E. coli*-Translationsbeginn verschlüsselt. Weiterhin ist zu vermerken, daß sich „stromaufwärts“ vom anfänglichen ATG eine Serie von Basenpaaren befindet, die die Ribosomenbindungsstellensequenz des hoch exprimierten OMP-f-Gens von *E. coli* dupliziert.

Um das ECEPO zu tragen, kann ein beliebiger geeigneter Expressionsvektor eingesetzt werden. Der besondere Vektor, der für die Expression des ECEPO-Gens als das „temperaturempfindliche“ Plasmid pCFM536 — ein Derivat des Plasmids pCFM414 (ATCC 40076) — gewählt wurde, ist in der noch anhängigen US-Patentanmeldung Nr. 636727 eingereicht am 6. August 1984 von Charles F. Morris beschrieben.

Insbesondere wurde pCFM536 mit **XbaI** und **HindIII** abgebaut; das große Fragment wurde isoliert und in einer zweiteiligen Ligierung mit dem ECEPO-Gen eingesetzt. Sectionen 1 (**XbaI/KpnI**), 2 (**KpnI/BglII**) und 3 (**BglII/SalI**) wurden vorher in korrekter Weise in M13 vereinigt, und das EPO-Gen wurde daraus isoliert als ein einzelnes **XbaI/HindIII**-Fragment.

Zu diesem Fragment gehörte ein Teil des Polylinkers aus dem M13 und mp9 Phagen, der die **SalI** bis **HindIII**-Stellen darin überspannte. Die Steuerung der Expression im erhaltenen Expressionsplasmid p536 erfolgte mittels eines Lambda- P_L -Promotors, der selbst von dem C_{11857} -Repressor-Gen (wie vom *E. coli* Stamm K12 Htrp vorgesehen) gesteuert wurde. Das hergestellte obige ECEPO-Gen konnte auf verschiedene Weise modifiziert werden, um Erythropoietin-Analoga zu verschlüsseln, wie [Asn², des Pro² bis Ile⁶] hEPO und [His⁷] hEPO, wie unten beschrieben.

A. [Asn², des Pro² bis Ile⁶] hEPO

Das Plasmid 53b, das das ECEPO-hergestellte Gen von Tabelle XIV als ein **XbaI**- bis **HindIII**-Insert trug, wurde mit **HindIII** und **XhoI** abgebaut. Letztere Endonuclease schneidet das ECEPO-Gen an einer einzigen 6-Basenpaare-Erkennungsstelle, die die letzte Base des Asp⁸ verschlüsselnden Codons bis zur zweiten Base des Arg¹⁰-Codons überspannt.

Eine **XbaI/XhoI**-„Linker“-sequenz wurde hergestellt, die die folgenden Sequenzen enthielt:

XbaI		+1	2	7	8	9
Met	Ala	Asn	Cys	Asp	XhoI	
5'-CTAG ATG	GCT	AAT	TGC	GAC-3'		
3'-TAC	CGA	TTA	ACG	CTG AGCT-5'		

Der **XbaI/XhoI**-Linker und das **XhoI/HindIII**-ECEPO-Gensequenzfragment wurde in das große Fragment insertiert, das man aus dem **XbaI**- und **HindIII**-Abbau des Plasmids pCFM526 — einem Derivat des Plasmids pCFM414 (ARCC 40076) — erhielt, wie in der noch anhängigen US-Patentanmeldung Nr. 636727, eingereicht am 6. August 1984 von Charles F. Morris beschrieben, um eine Plasmid-geborene DNA-Sequenz zu erzeugen, die die *E. coli*-Expression der Met⁻¹-Form des gewünschten Analogens verschlüsselt.

B. [His⁷] hEPO

Das Plasmid 536 wurde mit **HindIII** und **XhoI** wie im Teil A oben abgebaut. Ein **XbaI/XhoI**-Linker wurde hergestellt, der die folgende Sequenz hatte:

XbaI		+1	2	3	4	5	6	7	8	9	XhoI
Met	Ala	Pro	Pro	Arg	Leu	Ile	His	Asp			
5'-CTAG ATG	GCT	CCG	CCA	CGT	CTG	ATC	CAT	CAC-3'			
3'-TAC	CGA	GGC	GGT	GCA	GAC	TAG	GTA	CTG	AGCT-5'		

Der Linker und das **XhoI/HindIII**-ECEPO-Sequenzfragment wurde dann in pCFM526 insertiert, um eine Plasmid-gebotene DNA-Sequenz zu erzeugen, die die *E. coli*-Expression der Met⁻¹-Form des gewünschten Analogens verschlüsselt.

Die Konstruktion eines hergestellten Gens („SCEPO“) einschließlich der Hefepreferenzcodons ist in den folgenden Tabellen XV bis XXI beschrieben. Wie beim ECEPO-Gen gehört zu der gesamten Konstruktion die Bildung von drei Sätzen Oligonucleotide (Tabellen XV, XVII und XIX) die in Doppelungen formiert wurden und zu Sectionen vereinigt wurden (Tabellen XVI, XVIII und XX). Es ist festzuhalten, daß die Synthese teilweise erleichtert wurde durch Verwendung von einigen suboptimalen Codons, sowohl bei der SCEPO- als auch ECEPO-Konstruktion, d.h. die Oligonucleotide 7–12 der Section 1 beider Gene waren identisch, ebenso die Oligonucleotide 1–6 der Section 2 in jedem Gen.

Tabelle XV
SCEPO SECTION 1 OLIGONUCLEOTIDE

1. AATTCAAGCTTGGATAAAAGAGCT
2. GTGGAGCTCTTTTATCCAAGCTTG
3. CCACCAAGATTGATCTGTGACTC
4. TCTCGAGTCACAGATCAATCTTG
5. GAGAGTTTTGGAAAGATACTTGTG
6. CTTCCAACAAGTATCTTTCCAAAC
7. GAAGCTAAAGAAGCTGAAACATC
8. GTGGTGATGTTTTAGCTTCTTTAG
9. ACCACTGGTTGTGCTGAACACTGTT
10. CAAAGAACAGTGTTTCAACACAACCA
11. TTTGAACGAAACATTACGGTACCG
12. GATCCGGTACCGTAATGTTTTCGTT

Tabelle XVI
SCEPO SECTION 1

EcoRI Hind III 1
AATTCA AGCTTGGATA
GT TCGAACCTAT
2

3
AAAGAGCTCC ACCAAGATTG ATCTGTGACT dGAGAGTTTT
TTTCTCGAGG TGCTTCTAAC TAGACACTGA GCTCTCAAAA
4

5 7
GGAAAGATAC TTGTTGGAAG CTAAAGAAGC TGAAAACATC ACCACTGGTT
CCTTTCTATG AACAACTTC GATTTCTTCG ACTTTTGTAG TGGTGACCAA
6 8

9 11 KpnI BamHI
GTGCTGAACA CTGTTGTTT AACGAAAACA TTACGGTACC G
CACGACTTGT GACAAGAAAC TTGCTTTTGT AATGCCATGG CCTAG
12

Tabelle XVII
SCEPO SECTION 2 OLIGONUCLEOTIDE

1. AATTCGGTACCAGACACCAAGGT
2. GTTAACCTTGGTGTCTGGTACCG
3. TAATTCTACGCTTGGAAACGTAT
4. TTCCATACGTTTCCAAGCGTAGAA
5. GGAAGTTGGTCAACAAGCAGTTGAAGT
6. CCAAACCTCAACTGCTTGTGACCAAC
7. TTGGCAAGGTTTGGCCTTGTTATCTG
8. GCTTCAGATAACAAGGCCAACCTTG
9. AAGCTGTTTTGAGAGGTCAAGCCT
10. AACAAAGGCTTGACCTCTCAAAACA
11. TGTTGGTTAACTCTTCTCAACCATGGG
12. TGGTTCCCATGGTTGAGAAGAGTTAACC
13. AACCATTGCAATTGCACGTCGAT
14. CTTTATCGACGTGCAATTGCAA
15. AAAGCCGTCTCTGGTTTGAGATCTG
16. GATCCAGATCTCAAACCAGAGACGG

Tabelle XVIII
SCEPO SECTION 2

EcoRI KpnI
A ATTCGGTACC AGACACCAAG
GCCATGG TCTGTGGTTC
2

GTTAACTTCT 3 ACGCTTGGAA ACGTATGGAA 5 GTTGGTCAAC AAGCTGTTGA
CAATTGAAGA TCGGAACCTT TGCATACCTT CAACCAGTTG 6 TCGACAACCT
4

AGTTTGGCAA 7 GGTTTGGCCT TGTTATCTGA 9 AGCTGTTTTG AGAGGTCAAG
TCAAACCGTT CCAAACCGGA ACAATAGACT TCGACAAAAC 10 TCTCCAGTTC
10

CCTTTGTTGGT 11 TAACTCTTCT CAACCATGGG 13 AACCATTGCA ATTGCACGTC
GGAACACACCA ATTGAGAAGA GTTGGTACCC TTGGTAACGT TAACGTGCAG 14
14

GATAAAGCCG 15 TCTCTGGTTT BglII BamHI GAGATCTG
CTATTTCGGC AGAGACCAAA CTCTAGACCTA G
16

Tabelle XIX
SCEPO SECTION 3 OLIGONUCLEOTIDES

1. GATCCAGATCTTTGACTACTTTGTT
2. TCTCAACAAAGTAGTCAAAGATCTG
3. GAGAGCTTTGGGTGCTCAAAAGGAAG
4. ATGGCTTCCTTTGAGCACCCAAAGC
5. CCATTTCCTCCACAGACGCTGCTT
6. GCAGAAGCAGCGTCTGGTGGGGAA
7. CTGCCGCTCCATTGAGAACCATC
8. CAGTGATGGTTCTCAATGGAGCG
9. ACTGCTGATACCTTCAGAAAGTT
10. GAATAACTTTCTGAAGGTATCAG
11. ATTCAGAGTTTACTCCAATTCT
12. CTCAAGAAGTTGGAGTAACTCT
13. TGAGAGGTAAATTGAAGTTGTACAC
14. ACCGGTGTACAACCTCAATTTACCT
15. CGGTGAAGCCTGTAGAACTGGT
16. CTGTCACCAGTTCTACAGGCTTC
17. GACAGATAAGCCCGACTGATAA
18. GTTGTTATCAGTCGGGCTTAT
19. CAACAGTGTAGATGTAACAAAG
20. TCGACTTTGTTACATCTACACT

Tabelle XX
SCEPO SECTION 3

BamHI BglIII 1
GATC CAGATCTTTG ACTACTTTGT TGAGAGCTTT
GTCTAGAAAC TGATGAAACA ACTCTCGAAA
2

3 5
GGGTGCTCAA AAGGAACCA TTTCCCCACC AGACGCTGCT TCTGCCGCTC
CCCACGAGTT TTCCTTCGGT AAAGGGGTGG TCTGCGACGA AGACGGCGAG
4 6

7 9 11
CATTGAGAAC CATCACTGCT GATACCTTCA GAAAGTTATT CAGAGTTTAC
GTAACCTCTG GTAGTGACGA CTATGGAAGT CTTTCAATAA GTCTCAAATG
8 10 12

13 15
TCCAACTTCT TGAGAGGTAA ATTGAAGTTG TACACCGGTG AAGCCTGTAG
AGGTTGAAGA ACTCTCCATT TAACTTCAAC ATGTGGCCAC TTCGGACATC
14 16

17 19
AACTGGTGAC AGATAAGCCC GACTGATAAC AACAGTTGTAG
TTGACCCTG TCTATTCCGG CTGACTATTG TTCTCACATC
18

SalI
ATGTAACAAA G
TACATTGTTT CAGCT
20

Tabelle XXI
SCEPO GEN

-1+1
HindIII ArgAla
AGCTTGATA AAAGAGCTCC ACCAAGATTG ATCTGTGACT CGAGAGTTTT
ACCTAT TTTCTCGAGG TGGTTCTAAC TAGACACTGA GCTCTCAA

GGAAAGATAC TTGTTGGAAG CTAAAGAAGC TGAAAACATC ACCACTGGTT
CCTTTCTATG AACAACCTTC GATTTCTTCG ACTTTTGTAG TGGTGACCAA

GTGCTGAACA CTGTTCTTTG AACGAAAACA TTACGGTACC AGACACCAAG
CACGACTTGT GACAAGAAAC TTGCTTTTGT AATGCCATGG TCTGTGGTTC

GTAACTTCT ACGCTTGGAA ACGTATGGAA GTTGGTCAAC AAGCTGTTGA
CAATTGAAGA TGCGAACCTT TGCATACCTT CAACCAAGTTG TTCGACAACT

AGTTTGGCAA GGTTCGGCT TGTTATCTGA AGCTGTTTTG AGAGGTCAAG
TCAAACCGTT CCAAACCGGA ACAATAGACT TCGACAAAAC TCTCCAGTTC

CCTTGTTGGT TAACTCTTCT CAACCATGGG AACCATTGCA ATTGCACGTC
GGAACAACCA ATTGAGAAGA GTTGGTACCC TTGGTAACGT TAACGTGCAG

GATAAAGCCG TCTCTGGTTT GAGATCTTTG ACTACTTTGT TGAGAGCTTT
CTATTTCCGGC AGAGACCAAA CTCTAGAAAC TGATGAAACA ACTCTCGAAA

GGGTGCTCAA AAGGAAGCCA TTTCCCCACC AGACGCTGCT TCTGCCGCTC
CCCACGAGTT TTCCTTCGGT AAAGGGGTGG TCTGCGACGA AGACGGCGAG

CATTGAGAAC CATCACTGCT GATACCTTCA GAAAGTTATT CAGAGTTTAC
GTAACCTCTG GTAGTGACGA CTATGGAAGT CTTTCAATAA GTCTCAAATG

Tabelle XXI (Fortsetzung)

	-1 +1			
HindIII	ArgAla			
TCCAATTCT	TGAGAGGTAA	ATTGAAGTTG	TACACCGGTG	AAGCCTGTAG
AGGTTGAAGA	ACTCTCCATT	TAACITCAAC	ATGTGGCCAC	TTCGGACATC
AACTGGTGAC	AGATAAGCCC	GAAGTATAAC	AACAGTGTAG	
TTGACCACTG	TCTATTCGGG	CTGACTATTG	TTGTACATC	
	Sall			
ATGTAACAAA	G			
TACATTGTTT	CAGCT			

Die vereinigten SCEPO-Sectionen wurden in M 13 sequenziert, und die Sectionen 1, 2 und 3 waren aus dem Phagen als **HindIII/KpnI**-, **KpnI/BglII**- und **BglII/Sall**-Fragmente isolierbar.

Das gegenwärtig bevorzugteste Expressionssystem für SCEPO-Genprodukte ist ein Sekretionssystem, basierend auf einer *S. cerevisiae*- α -Faktor-sekretion, wie in der anhängigen US-Patentanmeldung Nr. 487 753, eingereicht am 22. April 1983 von Grant A. Bitter, veröffentlicht am 31. Oktober 1984 als Europäische Patentanmeldung 0 123 294, beschrieben. Kurz zusammengefaßt gehören zu dem System Konstruktionen, worin DNA, die die Führungssequenz des Hefe- α -Faktor-Genproduktes verschlüsselt, unmittelbar 5' zur codierenden Region des zu exprimierenden exogenen Gens angeordnet ist. Als ein Ergebnis enthält das translatierte Genprodukt eine Führungs- oder Signalsequenz, die „wegbehandelt“ (processed off) wird durch ein endogenes Hefezym im Verlauf der Sekretion des verbliebenen Produktes. Da bei der Konstruktion mit dem α -Faktor-Translationsbeginn (ATG)-codon gearbeitet wurde, war es nicht erforderlich, ein solches Codon an der -1-Position des SCEPO-Gens einzubauen. Wie aus Tabelle XXI entnommen werden kann, geht der Alanin (+1) verschlüsselnden Sequenz eine Linkersequenz voraus, die eine direkte Insertion in ein Plasmid gestattet, enthaltend die DNA für die ersten 80 Reste der α -Faktorführungssequenz, die dem α -Faktor-Promoter folgt. Zu der speziell bevorzugten Konstruktion für die SCEPO-Genexpression gehört eine vierteilige Ligierung unter Einschluß des oben genannten SCEPO-Section-Fragmentes und des großen Fragmentes des **HindIII/Sall**-Abbaus des Plasmids p α C3. Aus dem erhaltenen p α C3/SCEPO wurden der α -Faktorpromotor und die Führungssequenz und das SCEPO-Gen isoliert durch Abbau mit **Bam**HI und ligiert in das **Bam**HI-abgebaute Plasmid pYE, um zu dem Expressionsplasmid pYE/SCEPO zu gelangen.

Es sollte bei Betrachtung der Beispiele deutlich geworden sein, daß zahlreiche außerordentlich wertvolle Produkte und Verfahrensweisen durch die vorliegende Erfindung in ihren vielen Aspekten hervorgebracht wurden.

Nach einem weiteren Aspekt der vorliegenden Erfindung sind die hier beschriebenen klonierten DNA-Sequenzen, die Human- und Affen-EPO-Polypeptide verschlüsseln, bemerkenswert wertvoll für die darin enthaltenen Informationen hinsichtlich der Aminosäuresequenz von Säugetier-Erythropoietin, die bis heute nicht zu erhalten waren, abgesehen von Teilen aus analytischen Verfahren von Isolaten natürlich auftretender Produkte. Die DNA-Sequenzen sind auch bemerkenswert wertvoll als Produkte, die dazu eingesetzt werden können, die mikrobielle Synthese von Erythropoietin im großen Maße zu beschreiben mittels einer Vielzahl rekombinanter Techniken. Andererseits sind die durch die vorliegende Erfindung vorgesehenen DNA-Sequenzen nützlich für die Erzeugung neuer und nutzbringender viraler und ringförmiger Plasmid-DNA-Vektoren, neuer und nutzbringender transformierter und transfektierter mikrobieller procaryotischer und eucaryotischer Wirtszellen (einschließlich bakterieller und Hefezellen sowie in Kultur gewachsener Säugetierzellen sowie neuer und nutzbringender Verfahren des kultivierten Wachstums solcher mikrobieller Wirtszellen, die zur Expression von EPO- und EPO-Produkten in der Lage sind.

Die erfindungsgemäßen DNA-Sequenzen sind auch bemerkenswert geeignete Materialien für den Einsatz als markierte Sonden bei der Isolierung von EPO und verwandter Protein-verschlüsselnder cDNA- und genomischer DNA-Sequenzen von anderen Säugetieren als den hier speziell erläuterten Menschen und Affenarten. Der Gehalt, bis zu dem die erfindungsgemäßen DNA-Sequenzen bei dem unterschiedlichen alternativen Verfahren zur Prokaryotensynthese eingesetzt werden (z. B. in Insektenzellen) oder bei der genetischen Therapie von Menschen oder anderen Säugetieren kann noch nicht berechnet werden. Es ist zu erwarten, daß die erfindungsgemäßen DNA-Sequenzen bei der Entwicklung transgenerischer Säugetierarten nutzbringend sein können, die als eucaryotische „Wirt“ zur Herstellung von Erythropoietin und Erythropoietinprodukten in Mengen dienen können. Siehe allgemein Polmiter et al., Science, **222** (4625), 809-814 (1983).

Unter diesem Gesichtspunkt ist klar nicht beabsichtigt, daß die erläuternden Beispiele den Schutzzumfang der vorliegenden Erfindung einschränken, und für den Fachmann sind zahlreiche Modifikationen und Variationen denkbar. Als ein Beispiel:

Während zu den DNA-Sequenzen, die durch die erläuternden Beispiele gehören, da diese Anmeldung Aminosäuresequenzinformationen vorsieht, die zur Herstellung der DNA-Sequenz wesentlich sind, umfaßt die Erfindung auch solche hergestellten DNA-Sequenzen, wie sie infolge des Wissens um die EPO-Aminosäuresequenzen konstruiert werden können. Diese können für EPO codieren (wie im Beispiel 12), ebenso wie für EPO-Fragmente und kEPO-Polypeptidanaloge (d. i. „EPO-Produkte“), die eine oder mehrere der biologischen Eigenschaften von natürlich auftretendem EPO aufweisen, andere jedoch nicht (oder die andere in unterschiedlichem Grad besitzen).

Die durch die vorliegende Erfindung offengelegten DNA-Sequenzen umfassen somit alle DNA-Sequenzen, die für eine sichere Expression eines Polypeptidproduktes in einer procaryotischen oder eucaryotischen Wirtszelle geeignet sind, wobei dieses Produkt wenigstens einen Teil der primären strukturellen Anordnung und eine oder mehrere der biologischen Eigenschaften von Erythropoietin aufweist und unter folgenden ausgewählt ist:

- den in den Tabellen V und VI aufgeführten DNA-Sequenzen;
- DNA-Sequenzen, die zu den unter (a) aufgeführten oder deren Fragmenten hybridisieren; und
- DNA-Sequenzen, die, allerdings unter Degenerierung des genetischen Codes, zu den unter (a) und (b) definierten DNA-Sequenzen hybridisieren. In diesem Zusammenhang ist es zum Beispiel bemerkenswert, daß zu erwarten ist, daß existierende allelische Affen- und Human-EPO-Gensequenzen und die Gensequenzen anderer Säugetierarten zu Sequenzen der Tabellen V und VI oder zu Fragmenten davon hybridisieren. Darüber hinaus würden auch — allerdings unter Degenerierung des genetischen

Codes — die SCEPO — und ECEPO-Gene und die hergestellten oder mutierten cDNA- oder genomischen DNA-Sequenzen, die verschiedene EPO-Fragmente und -analoge verschlüsseln, ebenfalls zu den oben genannten DNA-Sequenzen hybridisieren. Derartige Hybridisierungen könnten leicht unter den hier beschriebenen Hybridisierungsbedingungen durchgeführt werden im Hinblick auf die anfängliche Isolierung der Affen- und Human-EPO verschlüsselnden DNA oder schärferen Bedingungen, um die Hintergrundhybridisierung gewünschtenfalls zu verringern.

Erfindungsgemäße verbesserte Hybridisierungsmethodologien, die erläuternd auf die obigen DNA/DNA-Hybridisierungsscreenings angewandt wurden, sind gleichfalls auf RNA/RNA- und RNA/DNA-Screening anwendbar. Mischsondentechniken, wie sie hier erläutert wurden, bestehen allgemein aus einer Anzahl von Verbesserungen bei Hybridisierungsmethoden, die schnellere und zuverlässigere Polynucleotidisolierungen gestatten. Zu diesen vielen einzelnen Verfahrensverbesserungen gehören:

verbesserte Kolonieübertragungs- und -erhaltungsverfahren; Verwendung von Filtern auf Nylonbasis wie Gene-Screen und Gene-Screen-Plus, um mit demselben Filtern ein Re-Sondieren zu gestatten und die Filter wiederholt verwenden zu können, Anwendung neuer Proteasebehandlungen [z. B. Taub et al., *Anal. Biochem.*, **126**, S. 222–230 (1982)];

Verwendung von sehr geringen Einzelkonzentrationen (im Bereich von 0,025 picomol) einer größeren Anzahl von Mischsonden (z. B. die über 32 liegenden Zahlen); und Durchführung von Hybridisierungs- und Nachhybridisierungsschritten unter strengen Temperaturen, die in der Nähe (d. i. innerhalb 4°C und vorzugsweise innerhalb 2°C entfernt von) der niedrigsten berechneten Dissoziations-temperatur einer beliebigen der eingesetzten Mischsonden liegt. Die Verbesserungen führen gemeinsam zu Ergebnissen, die von vornherein nicht zu erwarten gewesen sind. Verstärkt wird dies durch die Tatsache unterstrichen, daß Mischsondenverfahren mit der vierfachen Zahl der Sonden, über die je bisher als erfolgreich eingesetzt in gleichen cDNA-Screens von Boten-RNA-spezies relativ geringer Menge berichtet wurde, erfolgreich für die Isolierung eines einzelnen Gens in einer Genbibliothek unter Screenen von 1 500 000 Phagenplaques angewandt wurden. Dies wurde erreicht im wesentlichen entgegen der von Anderson et al., siehe oben, publizierten Auffassung, daß Mischsondenscreeningverfahren ... für die Isolierung von Säugetierproteingenen untunlich sind, wenn die entsprechenden RNA's nicht erhältlich sind.

Die in der vorangegangenen Beschreibung, in den folgenden Patentansprüchen und/oder in den dazugehörigen Zeichnungen offenbarten Merkmale können sowohl getrennt voneinander als auch in Kombination Grundlage für die Realisierung der Erfindung in ihren diversen Formen sein.

FIG. 1

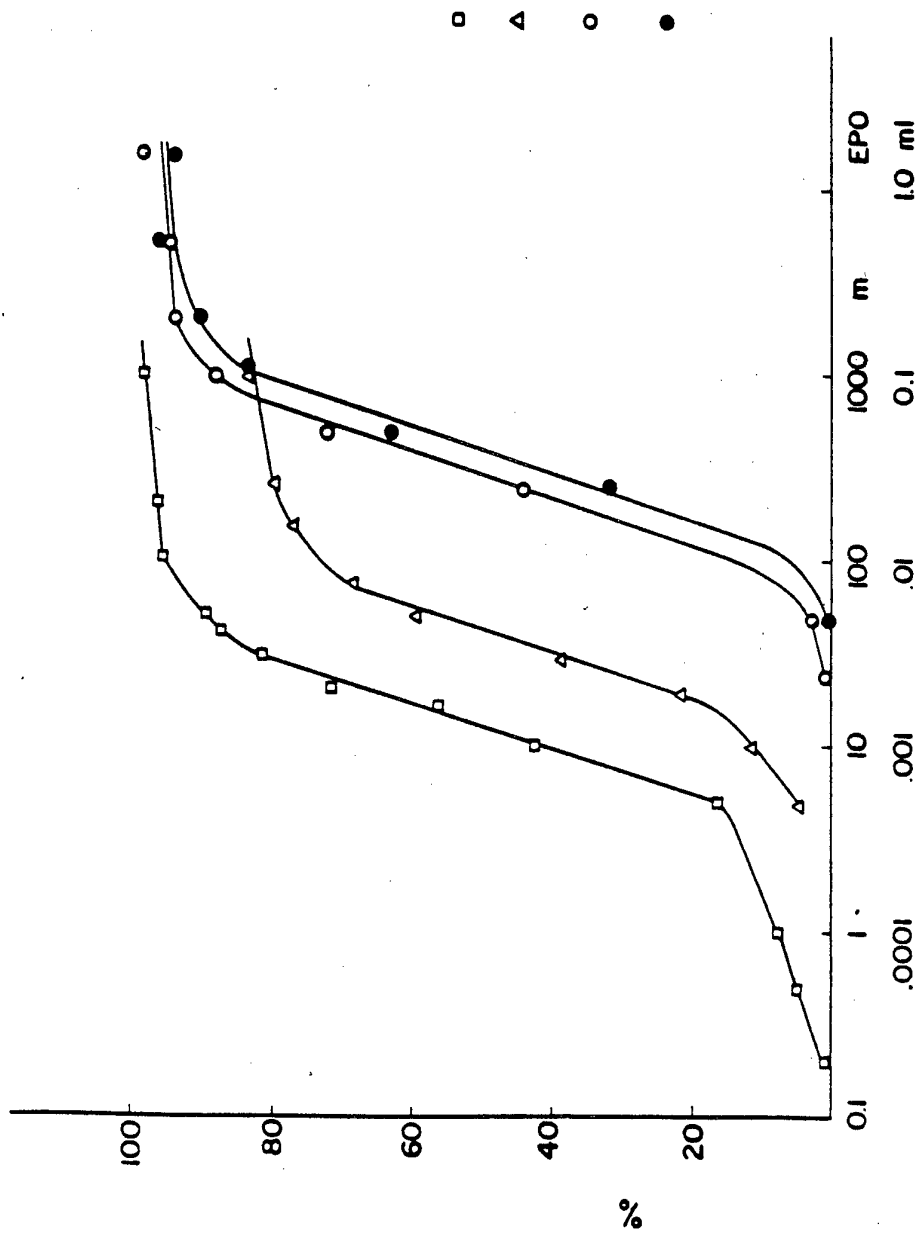


FIG. 2

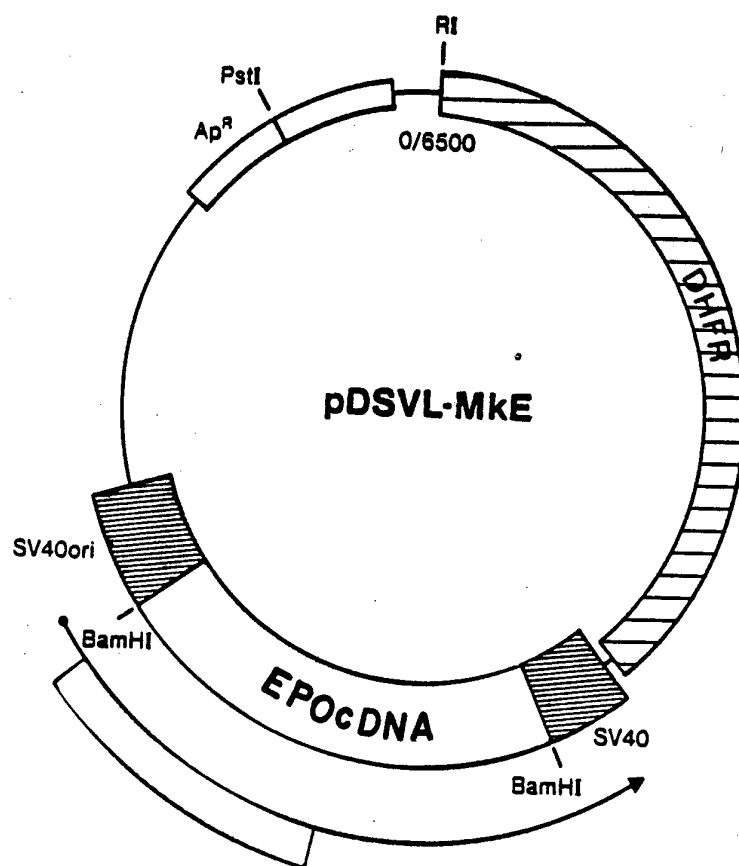


FIG.3

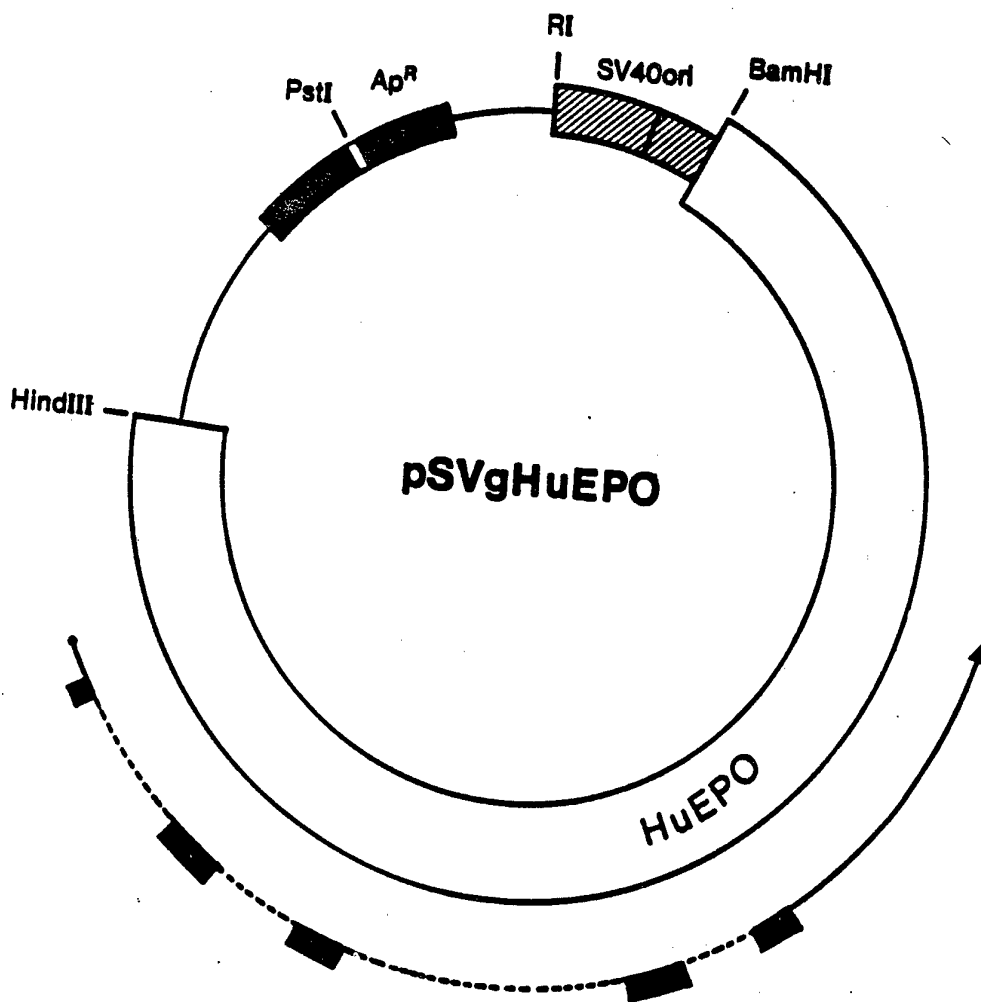


FIG. 4