(19) **Bundesrepublik Deutschland**
Deutsches Patent- und Markenamt

(10) **DE 694 32 744 T2** 2004.03.25

(12) # Übersetzung der europäischen Patentschrift

(97) **EP 0 686 162 B1**
(21) Deutsches Aktenzeichen: **694 32 744.1**
(86) PCT-Aktenzeichen: **PCT/DK94/00054**
(96) Europäisches Aktenzeichen: **94 906 891.0**
(87) PCT-Veröffentlichungs-Nr.: **WO 94/018227**
(86) PCT-Anmeldetag: **04.02.1994**
(87) Veröffentlichungstag der PCT-Anmeldung: **18.08.1994**
(97) Erstveröffentlichung durch das EPA: **13.12.1995**
(97) Veröffentlichungstag der Patenterteilung beim EPA: **28.05.2003**
(47) Veröffentlichungstag im Patentblatt: **25.03.2004**

(51) Int Cl.[7]: **C07K 1/14**
C07K 14/47, C07K 14/61, C07D 471/04, C07K 14/81, C07D 413/14, C12N 9/68

(30) Unionspriorität:

13093	**04.02.1993**	**DK**
13993	**05.02.1993**	**DK**
PCT/GB93/02492	**03.12.1993**	**WO**

(84) Benannte Vertragsstaaten:
AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LI, LU, MC, NL, PT, SE

(73) Patentinhaber:
Borean Pharma A/S, Aarhus, DK

(74) Vertreter:
LEINWEBER & ZIMMERMANN, 80331 München

(72) Erfinder:
THOGERSEN, Christian, Hans, DK-8381 Mundelstrup, DK; HOLTET, Las, Thor, DK-8210 Arhus V, DK; ETZERODT, Michael, DK-8382 Hinnerup, DK

(54) Bezeichnung: **VERBESSERTES VERFAHREN ZUR RÜCKFALTUNG DER PROTEINE**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung

Gebiet der Erfindung

[0001] Diese Erfindung betrifft rekombinante DNA-Technologie und im Speziellen Protein-Engineering-Technologien zur Herstellung korrekt gefalteter Proteine durch heterologe oder homologe Expression von Genen oder Genfragmenten in einem Wirtsorganismus als rekombinante Proteinprodukte. Die Erfindung stellt neue allgemeine Prinzipen und Verfahren zur effizienten In-vitro-Neufaltung von fehlgefalteten und/oder unlöslichen Proteinen, einschließlich Disulfidbrücken enthaltenden Proteinen bereit. Diese Erfindung betrifft weiters die Neufaltung von ungefalteten oder fehlgefalteten Polypeptiden jeglicher Herkunft. Zwei Analoga des Rinderkoagulationsfaktors X_a, die für technologische Anwendungen im kleinen, mittleren und großen Maßstab, die die spezifische Spaltung chimärer Proteine an Stellen umfassen, die zur Spaltung durch Faktor X_a konstruiert sind, werden ebenfalls dargestellt. Die Erfindung kann reversible disulfidblockierende Reagenzien nützen, die als Hilfsverbindungen zur Neufaltung von Cystein enthaltenden Proteinen brauchbar sind. Ebenfalls dargestellt ist ein allgemeines Testverfahren, mit dem solche Disulfid-Austauschreagenzien für diesen speziellen Zweck evaluiert werden können.

Allgemeiner Hintergrund der Erfindung

[0002] Technologien zur Herstellung nahezu beliebiger Polypeptide durch Einführung eines natürlichen oder synthetischen DNA-Fragments, das für dieses spezielle Polypeptid kodieren, in einen geeigneten Wirt mittels DNA-Verfahren ist während der letzten fünfzehn Jahre intensiv entwickelt worden und sind derzeit essentielle Werkzeuge für die biochemische Forschung und für eine reihe industrieller Prozesse zur Produktion hochwertiger Proteinprodukte für biomedizinische oder andere industrielle Verwendungen.

[0003] Vier fundamentale Eigenschaften biologischer Systeme machen die heterologe Produktion von Proteinen möglich:

(i) Die funktionellen Eigenschaften eines Proteins sind durch ihre dreidimensionale Struktur vollständig definiert und wegen der molekularen Umgebung in der Struktur durch chemische Eigenschaften manifestiert, die von speziellen Teilen dieser Struktur entfaltet werden.

(ii) Die dreidimensionale Struktur eines Proteins ist ihrerseits durch die von der speziellen, aufeinander folgenden Anordnung von Aminosäureresten in der/den linearen Polypeptidkette(n) repräsentierten Sequenzinformation spezifiziert. Die in der Aminosäuresequenz eines Polypeptids verankerte Strukturinformation selbst ist ausreichend, um unter geeigneten Bedingungen den Faltungsprozess zu lenken, dessen Endprodukt das vollständig und korrekt gefaltete Protein ist.

(iii) Die lineare Sequenz von Aminosäureresten in der Polypeptidkette ist von der Nucleotidsequenz in der kodierenden Region des die Assemblierung der Polypeptidkette durch die Zellmaschinerie steuernden genetischen Materials festgelegt. Die Translation der Nucleinsäuresequenzinformation in eine die Aminosäuresequenz bestimmende Translationstabelle ist bekannt und ist unter den bekannten Organismen nahezu universell und ermöglicht daher Nucleinsäuresegmenten, die für ein beliebiges Polypeptidsegment kodieren, die Assemblierung des Polypeptidprodukts über nahezu alle Speziesbarrieren hinweg zu steuern.

(iv) Jeder Organismentyp baut auf sein eigenes charakteristisches Feld von genetischen Elementen, die innerhalb seiner eigenen Gene vorhanden ist, um mit der molekularen Maschinerie der Zelle in Wechselwirkung zu treten, die in Reaktion auf spezifische intrazelluläre und extrazelluläre Faktoren die Expression eines gegebenen Gens in Form von Transkription und Translation reguliert.

[0004] Um die Proteinsynthesemaschinerie einer Wirtszelle oder eines Organismus zu verwerten, um eine hohe Produktion eines gewünschten rekombinanten Proteinprodukts zu erzielen, ist es daher notwendig, das für das gewünschte Produkt kodierende DNA-Segment der Zelle an Kontrollsequenzen fusioniert zu präsentieren, die vom genetischen Kontrollsystem der Zelle erkannt werden.

[0005] Das unmittelbare Schicksal eines in einem Wirt exprimierten Polypeptids wird von der Natur des Polypeptids, der Natur des Wirts und möglichen Wirtsorganismus-Stresszuständen beeinflusst, die während der Produktion eines gegebenen Polypeptids hervorgerufen werden. Ein in mittelmäßigem Ausmaß und ähnlich oder gleich einem normalerweise in der Wirtszelle vorhandenen Protein exprimiertes Genprodukt wird häufig der normalen Prozessierung unterzogen und im entsprechenden Zellkompartiment akkumuliert oder sekretiert, was auch immer das natürliche Schicksal dieses endogenen Genprodukts ist. Im Gegensatz dazu aktivieret ein rekombinantes Genprodukt, das der Zelle fremd ist oder in großen Mengen produziert wird, häufig zelluläre Abwehrreaktionen ähnlich jenen, die durch Hitzeschock oder Exposition mit toxischen Aminosäureanaloga aktiviert werden, Stoffwechselwege, die von der Natur entworfen worden sind, um der Zelle zu helfen, „falsches" Polypeptidmaterial durch kontrollierte Proteolyse oder durch Segregation unerwünschten Proteinmaterials in Speicherteilchen („Einschlusskörper") loszuwerden. Das rekombinante Protein in diesen Speicherteilchen wird

häufig in einem fehlgefalteten und aggregierten Zustand abgelagert, wobei es in diesem Fall notwendig wird, das Produkt unter denaturierenden oder reduzierenden Bedingungen aufzulösen und dann das rekombinante Polypeptid mittels In-vitro-Verfahren zu falten, um ein brauchbares Proteinprodukt zu erhalten.

[0006] Die Expression eukaryotischer Gene in eukaryotischen Zellen ermöglicht häufig die direkte Isolierung des korrekt gefalteten und prozessierten Genprodukts aus Zellkulturflüssigkeiten oder aus Zellmaterial. Dieser Ansatz wird häufig angewendet, um relativ kleine Mengen eines Proteins für biochemische Untersuchungen zu erhalten und wird gegenwärtig auch industriell zur Produktion einer Reihe von biomedizinischen Produkten verwertet. Jedoch ist die eukaryotische Expressionstechnologie bezüglich technologischer Komplexität, Personal- und Materialkosten kostspielig. Darüber hinaus beträgt die Zeitdauer der Entwicklungsphase, die zur Etablierung eines Expressionssystems erforderlich ist, sogar für die Produktion im Labormaßstab zumindest mehrere Monate. Die Natur und das Ausmaß der posttranslationellen Modifizierung des rekombinanten Produkts unterscheiden sich häufig vom natürlichen Produkt, da solche Modifizierungen der indirekten genetischen Kontrolle in der Wirtszelle unterliegen. Sequenzsignale, die eine postsynthetische Modifikation auslösen, werden unter Eukaryoten häufig wechselseitig erkannt, jedoch ist die Verfügbarkeit des geeigneten Satzes von modifizierenden Enzymen von der Natur und dem Zustand der Wirtszelle vorgegeben.

[0007] Eine Vielzahl von Strategien sind zur Expression von Genprodukten in prokaryotischen Wirten entwickelt worden, die gegenüber eukaryotischen Wirten bezüglich Kapital-, Arbeitskräfte- und Materialbedarf vorteilhaft sind. Stämme der Eubakterien Escherichia coli werden als Wirtszellen häufig bevorzugt, da E. coli auch auf der molekularen Ebene genetisch viel besser als irgendein anderer Organismus charakterisiert ist.

[0008] Prokaryotische Wirtszellen besitzen nicht die enzymatische Maschinerie, die zur Durchführung der posttranslationellen Modifikation erforderlich ist und ein eukaryotisches Genprodukt wird daher notwendigerweise in seiner unmodifizierten Form zu prozessieren sein. Darüber hinaus muss das Produkt mit einer N-terminalen Extension, zumindest einem zusätzlichen, aus dem erforderlichen Translationsinitiationscodon hervorgehenden Methioninrest synthetisiert werden und umfasst häufiger auch ein N-terminales Segment, das jenem eines stark exprimierten Wirtsproteins entspricht. Allgemeine Verfahren zur Entfernung von solchen N-terminalen Extensionen durch sequenzspezifische Proteolyse an Linker-Segmenten, die an der Verbindung zwischen der N-terminalen Extension und dem gewünschten Polypeptidprodukt insertiert sind, sind beschrieben worden (Enterokinase-spaltbare Linker-Sequenz: EP-A-035.384, The Regents of the University of California; Faktor-X_a-spaltbare Linker-Sequenz: EP 161.937, Nagai und Thøgersen, Zessionar: Celltech Ltd.).

[0009] Im Laufe der Jahre waren beträchtliche Bemühungen auf die Entwicklung von Strategien zur heterologen Expression in Prokaryoten gerichtet, um rekombinante Genprodukte in löslicher Form oder Fusionsproteinkonstrukte zu erzeugen, die die Sekretion aus der Zelle in einer aktiven, möglicherweise N-terminal prozessierten Form ermöglichen, eine nur begrenzt erfolgreiche Bemühung, trotz neuerer Entwicklungen auf dem Chaperon-Gebiet. Typischerweise ist viel Zeit und Aufwand erforderlich, um ein Expressionssystem zu entwickeln und zu modifizieren, bevor auch nur eine kleine Menge lösliches und korrekt gefaltetes Fusionsprodukt isoliert werden kann. Häufiger wird das gesamte Polypeptidprodukt innerhalb der Wirtszelle in ungeeignet gefaltetem Zustand in „Einschlusskörpern" abgelagert. Dies trifft insbesondere zu, wenn exprimierende eukaryotische Proteine Disulfidbrücken enthalten.

[0010] Alle verfügbaren Verfahren zur In-vitro-Neufaltung von Proteinen beschreiben Prozesse, in denen das Protein in Lösung oder unspezifisch an Ionentauscherharze usw. gebunden einem Lösungsmittel exponiert ist, wobei die Zusammensetzung davon über die Zeit in einem einzigen Durchgang graduell von stark denaturierend (und möglicherweise denaturierend) bis nicht denaturierend geändert wird. Dies wird häufig durch Verdünnen einer konzentrierten, 6–8 M Guanidin-Hydrochlorid oder Harnstoff Proteinlösung enthaltenden Proteinlösung in ein beträchtliches Volumen von nicht denaturierendem Puffer oder durch Dialyse einer verdünnten Lösung des Proteins im denaturierenden Puffer gegen den nicht denaturierenden Puffer durchgeführt. Zahlreiche Varianten dieses grundlegenden Verfahrens sind beschrieben worden, einschließlich Zusatz spezifischer Liganden oder Cofaktoren des aktiven Proteins und Inkorporation von Polymersubstanzen, wie Polyethylenoxid (Polyethylenglykol), von denen angenommen wird, dass sie die gefaltete Struktur stabilisieren.

[0011] Obgleich effiziente Varianten des standardmäßigen In-vitro-Neufaltungsverfahrens für eine Reihe spezifischer Proteinprodukte, einschließlich für Proteine, die eine oder mehrere Disulfidbindungen enthalten, gefunden worden sind, sind die Neufaltungsausbeuten häufig schlecht und die Maßstabsvergrößerung unmöglich und kostspielig, und zwar wegen der niedrigen Löslichkeit der meisten unvollständig gefalteten Proteine, was die Verwendung übermäßiger Lösungsmittelmengen mit sich bringt.

[0012] Das gemeinsame Charakteristikum aller traditioneller In-vitro-Neufaltungsprotokolle ist, dass die durch plötzliche oder graduelle Verminderung des Denaturierungsmittels induzierte Neufaltung als einmaliger Arbeitsvorgang durchgeführt wird, wobei die Ausbeute davon dann als die bestmögliche für das fragliche Protein betrachtet wird.

[0013] Das allgemeine Gebiet der Proteinfaltung ist in einem neueren, von Thomas W. Creighton („Protein folding", T. E. Creighton (Hrsg.), Freeman (1992)) herausgegebenen Lehrbuch zusammengefasst worden und eine speziellere Übersicht praktischer Verfahren zur Proteinneufaltung wurde 1989 von Rainer Jaenicke und

Rainer Rudolph (S. 191–223 in „Protein Structure, a practical approach", T. E. Creighton (Hrsg.), IRL Press (1989)) publiziert worden. Unter zahlreichen weiteren Publikationen können Übersichtsartikel auf dem neuesten Stand der Technik, wie z. B. jene von Schein (C. H. Schein, Bio/Technology 8, 308–317 (1990)) oder Buchner und Rudolph (J. Buchner und R. Rudolph, Bio/ Technology 9, 157–162 (1991)) konsultiert werden.

[0014] Zusammenfassend besteht ein eindeutiger Bedarf für allgemein anwendbare Verfahren hoher Ausbeute für die Neufaltung von un- oder fehlgefalteten Proteinen verschiedenster Herkunft, wie z. B. aus prokaryotischen Expressionssystemen oder Peptidsynthese.

Zusammenfassung der Erfindung

[0015] Die Erfinder haben gefunden, dass Neufaltungsausbeuten stark erhöht werden können, indem in Betracht gezogen wurde, dass der Proteinfaltungsprozess ein kinetisch kontrollierter Vorgang ist, und dass Umwandlung zwischen gefalteten, ungefalteten und fehlgefalteten Konformeren untereinander der Hysterese und zeitabhängigen Phänomenen unterworfen ist und genutzt werden kann, um einen zyklischen Denaturierungs/Renaturierungsprozess zu entwerfen, in dem neugefaltetes Proteinprodukt in jedem Zyklus schrittweise auf Kosten ungefalteter und fehlgefalteter Konformere akkumuliert, um einen neuen Neufaltungsprozess von viel höherer Leistungsfähigkeit als der einfache traditionelle Ansatz zu entwickeln.

[0016] Mit dem Ausdruck „gefaltetes Protein" ist ein Polypeptid in einem Konformationszustand oder in Konformationszuständen gemeint, die jenem oder jenen entsprechen, die im Protein in seiner aktiven biologisch Form oder einzigartigen stabilen Zwischenprodukten auftreten, die in nachfolgenden Schritten umgewandelt werden können, um die biologisch aktive Spezies zu erzeugen. Die kovalente Struktur des gefalteten Proteins bezüglich der Vernetzung zwischen Paaren von Cysteinresten um Polypeptid ist mit jener des Proteins in seiner biologisch aktiven Form identisch.

[0017] Demgemäß bezieht sich der Ausdruck „ungefaltetes Protein" auf ein Polypeptid in Konformationszuständen, die weniger kompakt und wohldefiniert sind als jener oder jene, die dem Protein in seiner biologisch aktiven und daher gefalteten Form entsprechen. Die kovalente Struktur des ungefalteten Proteins bezüglich Vernetzung zwischen Paaren von Cysteinresten im Polypeptid kann, muss aber nicht identisch mit derjenigen des Proteins in seiner biologisch aktiven Form identisch sein. Nahe verwandt mit einem ungefalteten Protein ist ein „fehlgefaltetes Protein", das ein Polypeptid in einem Konformationszustand ist, der nahezu thermodynamisch stabil ist und manchmal stabiler als jene Zustände, die dem Protein in seinem gefalteten Zustand entsprechen, die jedoch, wenn überhaupt, nicht denselben Grad der biologischen Aktivität des gefalteten Proteins zeigt. Wie es auch für das ungefaltete Protein zutrifft, kann die kovalente Struktur bezüglich Vernetzung zwischen Paaren von Cysteinresten im Polypeptid dieselbe wie jene des gefalteten Proteins sein, muss es aber nicht sein.

[0018] Mit dem Ausdruck „neugefaltetes Protein" ist ein Polypeptid gemeint, das aus einem ungefalteten Zustand umgewandelt worden ist. Um seine biologisch aktive Konformation und kovalente Struktur bezüglich Vernetzung zwischen korrekten Paaren von Cysteinresten im Polypeptid zu erlangen.

[0019] Die neue, allgemein anwendbare Proteinneufaltungsstrategie ist auf Basis der folgenden allgemeinen Eigenschaften der Proteinstruktur entworfen worden.

(a) Die niedrige Löslichkeit ungefalteter Proteine, die nicht denaturierenden Lösungsmitteln exponiert sind, spiegelt eine hautsächliche Triebkraft wider, die das Polypeptid veranlasst, entweder die kompakte, korrekt neugefaltete Struktur zu bilden, oder sich falsch zu falten und Dead-End-Aggregate oder Präzipitate zu erzeugen, die nicht in der Lage sind, sich innerhalb einer angemessenen Zeit neu zu falten und die korrekt neugefaltete Struktur unter nicht denaturierenden Bedingungen zu erzeugen.

(b) Ein neu gebildetes Dead-End-Aggregat wird leichter „denaturiert", d. h. in eine ungefaltete Form umgewandelt, als das korrekt gefaltete Protein, weil die Struktur des Dead-End-Aggregats ungeordneter ist. Wahrscheinlich ist die Fehlfaltung ebenfalls im Allgemeinen ein kinetisch kontrollierter Vorgang.

(c) Ein ungefaltetes Protein ist häufig nicht (oder nur sehr langsam) in der Lage, sich bei Denaturierungsmittelkonzentrationen, die zur Denaturierung von Dead-End-Aggregaten innerhalb eines angemessen Zeitraums erforderlich sind, neu in die korrekt neugefaltete Form zu falten.

(d) Das verfügbare Beweismaterial, um (b) zu untermauern, umfasst eingehende Untersuchungen der Faltungs- und Entfaltungswege und Zwischenprodukte für mehrere Modellproteine. Ebenfalls erläuternd ist die für viele disulfidgebundene Proteine gemachte Beobachtung, dass die Stabilität von Disulfidbindungen gegen Reduktion bei Grenzkonzentrationen von Reduktions- und Denaturierungsmitteln häufig für jede Disulfidbrücke eines gegebenen Proteins signifikant verschieden ist und dass die Disulfidbrücken im gefalteten Protein im Allgemeinen viel weniger zur Reduktion oder zum Disulfidaustausch neigen als „nicht native" Disulfidbindungen in einem denaturierten Protein oder Proteinaggregat.

[0020] Die neue Strategie für ein Neufaltungsverfahren kann am leichtesten durch das folgende theoretische Beispiel veranschaulicht werden:

[0021] Man betrachte ein hypothetisches Protein – stabil gefaltet in einem nicht denaturierenden Puffer „A" und stabil ungefaltet im stark denaturierenden Puffer „B" (der beispielsweise ein 6 M Guanidin-HCl enthaltender Puffer ist) – Puffer A oder Puffer B exponiert und dann der Inkubation bei mittleren Konzentrationen der denaturierenden Mischungen der Puffer A und B unterzogen.

[0022] Konzentrationen zwischen beispielsweise 100–75% B bewirken die Umwandlung des gefalteten Proteins, sowie des aggregierten Dead-End-Proteins zur ungefalteten Form innerhalb kurzer Zeit.

[0023] Konzentrationen zwischen beispielsweise 75–50% N bewirken die Umwandlung des neu gebildeten Dead-End-Aggregats zur ungefalteten Form, wogegen fast das gesamte neugefaltete Protein in einer nativ-artigen Struktur verbleibt, die zumindest innerhalb eines Zeitraums von Stunden stabil ist, aus der es bei Entfernung des Denaturierungsmittels in die neugefaltete Form zurückschnappen kann.

[0024] Konzentrationen von über 10% B verhindern die rasche Bildung der neugefalteten Form aus ungefalteter Form.

[0025] Eine Lösungsmittelzusammensetzungsschritt von 100% B zu 0% B wandelt das ungefaltete Protein in Dead-End-Aggregat (75% Ausbeute) und neugefaltetes Protein (25%) um.

[0026] Man unterwerfe nun eine Probe dieses Proteins anfänglich in seiner ungefalteten Form in 100% B einer Zeitreihe von programmierten Denaturierungs-Renaturierungs-Zyklen, wie in **Fig.** 1 dargestellt ist, wobei jeder aus einer Renaturierungsphase (F_n)(< 10% B) und einer Denaturierungsphase (D_n) besteht. Am Ende der Renaturierungsphase von Zyklus (i) wird der Denaturierungsmittelgehalt auf ein Niveau k_i% weniger als die Denaturierungsmittelkonzentration des vorhergehenden Zyklus geändert. Nach einer kurzen Inkubation wird das Denaturierungsmittel wiederum entfernt und die nächste Renaturierungsphase F_{i+1} eingeleitet. Unter der Annahme, dass die Denaturierungsmittelkonzentration bei 100% B beginnt und k_i für jeden Zyklus mit 4% festgelegt ist, wird dieses Rezept eine abschwächende Reihe von „Denaturierungsschritten" erzeugen, die nach 25 Zyklen erlischt.

[0027] Über 25 Zyklen hinweg, wie oben dargestellt, würde die Akkumulierung von neugefaltetem Protein wie folgt fortschreiten:

[0028] In den Zyklen 1 bis 5 würde das gesamte Protein, gefaltetes sowie fehlgefaltetes, in jedem der Denaturierungsphasen D_n entfaltet werden.

[0029] Zyklen 7 bis einschließlich 12: Dead-End-Aggregate werden in jedem Schritt in ungefaltetes Protein umgewandelt, während als neugefaltetes Produkt gewinnbares Protein in den folgenden Mengen Zyklus für Zyklus akkumulieren wird: 25%, 44%, 58%, 68%, 76% und 82%.

[0030] Keine weiteren Umwandlungen finden während der Zyklen 13 bis 25 statt.

[0031] Der zyklische Neufaltungsprozess würde daher eine Neufaltungsgesamtausbeute von über 80% hervorbringen, wogegen traditionelle Einschritt-Renaturierung im besten Falle eine Ausbeute von 25% hervorbringen würde.

[0032] Es ist klar zu erkennen, dass eine große Zahl simplifizierender Annäherungen bezüglich der Alles-oder-Nichts-Abstufung jeder Eigenschaft der verschiedenen Konformationszustände des hypothetischen Proteins gemacht worden sind. Das grundlegende Funktionsprinzip bleibt aber trotzdem ähnlich, wenn ein komplizierterer Satz von Annahmen ins Modell aufgenommen wird.

[0033] Die Anordnung eines praktikablen Aufbaus zur Etablierung eines zyklischen Denaturierungs/Renaturierungs-Proteinneufaltungsverfahrens kann auf viele Weisen ins Auge gefasst werden.

[0034] Das Protein in Lösung kann z. B. in einem Ultrafiltrationsgerät zurückgehalten, in einem Dialyseschlauch zurückgehalten oder auf eine der Phasen in einem geeigneten Zweiphasensystem beschränkt werden, wobei alle die Kontrolle der Konzentration niedermolekularer chemischer Gelöststoffe in der Proteinlösung durch geeignete Geräte ermöglichen.

[0035] Alternativ dazu könnte das Protein an eine geeignete Oberfläche im Kontakt mit einer Flüssigphase adsorbiert werden, wobei deren chemische Zusammensetzung wie erforderlich gesteuert werden könnte. Eine geeignete Oberfläche wäre z. B. ein Filtrationsgerät, ein Hohlfasergerät oder ein gekörntes Chromatographiemedium. Die Adsorption des Proteins an die Oberfläche könnte durch unspezifische Wechselwirkungen, wie sie z. B. in WO 86/05809 (Thomas Edwin Creighton) beschrieben werden, durch neufaltungskompatible kovalente Bindungen zwischen Oberfläche und Protein oder über spezifische Konstruktionen von Affinitäts-Handles in einem rekombinanten Derivat des Proteins vermittelt werden, die eine spezifische und denaturierungsresistente Affinität für eine geeignet derivatisierte Oberfläche zeigen.

[0036] Die spezifische Implementierung des zyklischen Denaturierungs/Renaturierungs-Proteinneufaltungsverfahrens, die zur Untersuchung des Potentials des allgemeinen Verfahrens etabliert wurde, basiert auf einem Design spaltbarer Hybridproteine (EP 161.937, Nagai und Thøgersen, Zessionar: Calltech Ltd.), die ein Metallaffinitäts-Handle-Modul (EP 282.042 (Heinz Dobeli, Bernhard Eggimann, Reiner Gentz, Erich Hochuli; Hoffmann-La Roche) enthielten, das N-terminal zur konstruierten Faktor-X_a-Spaltstelle insertiert war. Rekombinante Proteine dieser allgemeinen Konstruktion, adsorbiert an Nickel komplexierende Agaroseperlen könnten dann dem vorliegendem zyklischen Neufaltungsverfahren in einer Chromatographiesäule („Neufaltungsreaktor") unterzogen werden, die mit einem Gemisch geeigneter denaturierender und nicht denaturierender Puffer

durchströmt wird, die von einer Reihe kalibrierter Pumpen zugeführt werden, wobei die Durchflussraten durch Computersteuerung zeitprogrammiert werden.

[0037] Ein allgemeines Schema der Festphasenneufaltung erfordert die Zyklisierung des immobilisierten Proteins, wie sie oben dargestellt wurde oder durch beliebige andere Mittel oder Implementierungen, zwischen denaturierenden und nicht denaturierenden Bedingungen auf progressive Weise, bei der die Konzentration des denaturierenden Mittels graduell von hohen Anfangswerten über eine Kolonne vieler Renaturierungs-Denaturierungs-Zyklen auf Null herabgesetzt wird. Unter Verwendung dieses Ansatzes ist es nicht notwendig, genau zu bestimmen, welche Denaturierungsmittel-Grenzkonzentration erforderlich ist, um im Verlauf der Zyklen eine Faltungsausbeuteanreicherung des speziell vorliegenden Proteins zu erhalten, da die progressive Zyklenkolonne (bis zu) drei Phasen durchlaufen wird, einer frühen Phase, in der am Ende des Zyklus (i) vorhandenes, gefaltetes Produkt beim Denaturierungsschritt des Zyklus (i + 1) vollständig denaturiert wird, einer produktiven Zwischenphase, während der neugefaltetes Produkt in ansteigenden Mengen akkumuliert, und einer Spätphase, während der die Konzentration des Denaturierungsmittels zu niedrig ist, um das neugefaltete Protein oder jegliche verbleibenden fehlgefalteten Strukturen zu stören. Das Unterziehen des Proteins einer progressiven Reihe von Denaturierungs-Renaturierungs-Schritten wird daher mehrere produktive Zyklen umfassen.

[0038] Für Disulfid enthaltende Proteine können die Denaturierungs-Renaturierungs-Zyklen durch Verwendung einer Ausrüstung gesteigert werden, die einer hoch entwickelter Chromatographieausrüstung mit On-line-Einrichtungen ähnlich ist, um Pufferzusammensetzungen des Faltungsreaktorabflusses zu beobachten. Informationen über die Abflusszusammensetzung bezüglich Reduktionsmittel- und Disulfid-Reshufflingreagens-Konzentrationsprofil würden eine produktive Zyklisierung erkennen lassen und könnten daher als Eingangswert für eine intelligente Prozessoreinheit verwendet werden, die ihrerseits das Fortschreiten der Denaturierungsmittelkonzentration in einer Rückkoppelungsschleife steuert, um sicherzustellen, dass der Großteil des Zyklisierungsaufwands innerhalb der produktiven Phase der Denaturierungs-Renaturierungs-Zyklenkolonne verwendet wird. Solche Auto-Optimierungen der Zyklusbedingungen wären möglich, da das analytische System verwendet werden kann, um das Ausmaß und die Richtung der Veränderungen des Redoxgleichgewichts im Pufferstrom zu messen, Messungen, welche die Titration der Thiolgruppen/Disulfid-Äquivalente in der immobilisierten Proteinprobe direkt widerspiegeln, und ist daher direkt in die mittlere Anzahl von Disulfidbindungen umrechenbar, die während verschiedener Phasen eines Zyklus zerstört oder gebildet werden.

[0039] Andere mögliche Eingangswerte für die intelligente Prozessorsteuerung des Fortschritts der Zyklisierung umfassen Messungen von Ligandenbindung, Substratumsetzung, Antikörperbindungsfähigkeit und in der Tat irgendein anderes lösliches Agens, das in entfernter Weise mit fehlgefaltetem oder gefaltetem Protein wechselwirkt und in der Beurteilungsphase der Faltungsmessung durch den Neufaltungsreaktor sickert und dann im Abfluss mittels geeigneter analytischer Geräte in-line gemessen werden kann.

[0040] Ein intelligentes Beobachtungs- und Steuersystem könnte darüber hinaus die verfügbare Information nützen, um verwendbare Teile des Reaktorabflusses zu Wiederverwendungs/Recycling-Subsystemen zu lenken, wodurch Ausgaben für Prozesse im Großmaßstab minimiert werden.

[0041] Nach Ausführung der Faltungsprozedur kann das endgültige Produkt in konzentrierter Form von der Affinitätsmatrix eluiert und prozessiert werden, um das reife authentische Protein durch Spaltung an der konstruierten Protease-Spaltstelle freizusetzen, und dann einer letzten Aufarbeitung unter Verwendung von standardmäßigen Proteinreinigungs- und Handhabungstechniken unterzogen werden, die auf dem Gebiet der Proteinchemie wohlbekannt sind.

Ausführliche Offenbarung der Erfindung

[0042] Folglich stellt die vorliegende Erfindung ein Verfahren zur Erzeugung einer prozessierten Anordnung von Polypeptidmolekülen bereit, wobei die in dieser prozessierten Anordnung repräsentierten Konformationszustände einen wesentlichen Anteil von Polypeptidmolekülen in einer bestimmten gefalteten Konformation enthalten, aus einer anfänglichen Anordnung von Polypeptidmolekülen, die dieselbe Aminosäuresequenz wie die prozessierte Anordnung von Polypeptidmolekülen aufweisen, wobei die in dieser anfänglichen Anordnung repräsentierten Konformationszustände einen wesentlichen Anteil von Polypeptidmolekülen in ungefalteten oder fehlgefalteten Konformationen enthalten, wobei das Verfahren das Unterziehen der anfänglichen Anordnung von Polypeptidmolekülen einer Reihe von zumindest drei aufeinander folgenden Zyklen umfasst, wobei jeder davon die folgende Abfolge umfasst:

1) Zumindest ein Denaturierungsschritt, der Bedingungen umfasst, die einen denaturierenden und/oder entfaltenden Einfluss auf die Polypeptidmoleküle der Anordnung ausübt, so dass ein Teil der Polypeptide in der Anordnung denaturiert und/oder entfaltet werden, gefolgt von

2) zumindest einem Renaturierungsschritt, der Bedingungen umfasst, die einen renaturierenden Einfluss auf die Polypeptidmoleküle mit Konformationen haben, die aus dem vorhergehenden Schritt resultieren, so dass ein Anteil der denaturierten und/oder entfalteten Polypeptide in der Anordnung renaturiert werden,

wobei die Reihe der zumindest drei aufeinander folgenden Zyklen so angepasst ist, dass die Bedingungen in

zumindest einem Denaturierungsschritt die Denaturierung und/oder Entfaltung eines kleineren Anteils der Polypeptide in der Anordnung bewirkt, als die Denaturierungsbedingungen in einem früheren Denaturierungsschritt der Reihe, und dass die prozessierte Anordnung von Polypeptidmolekülen einen höheren Anteil von Polypeptidmolekülen in der bestimmten gefalteten Konformation aufweist, als
a) die anfängliche Anordnung und
b) eine entsprechende anfängliche Anordnung, die nur einem der Zyklen unterzogen worden ist.

[0043] In der vorliegenden Patentbeschreibung und in den Ansprüchen wird der Begriff „Anordnung" in der Bedeutung verwendet, die er auf dem Gebiet der Erfindung angenommen hat, dass heißt, er bezeichnet eine Sammlung von Molekülen mit wesentlichen gemeinsamen Eigenschaften. Anfänglich („eine anfängliche Anordnung") haben sie zumindest ihre Aminosäuresequenz gemeinsam (und behalten selbstverständlich diese gemeinsame Eigenschaft bei). Wenn die Anordnung von Polypeptidmolekülen im Verfahren der Erfindung behandelt worden ist (um „eine prozessierte Anordnung" zu liefern), werden die in der Anordnung repräsentierten Konformationszustände einen beträchtlichen Anteil von Polypeptidmolekülen mit einer bestimmten Konformation enthalten. Wie aus der folgenden Diskussion verständlich wird, kann der beträchtliche Anteil von Polypeptidmolekülen mit einer bestimmten Konformation in der prozessierten Anordnung variieren, und zwar in Abhängigkeit von den Parametern der Behandlung durch das Verfahren der Erfindung, der Größe des Proteins in der bestimmten Konformation, der Länge und Identität der der Aminosäuresequenz der Moleküle usw. In den hierin berichteten Beispielen, bei denen die Prozessparameter noch nicht optimiert worden sind, variierte der Anteil von Polypeptidmolekülen mit einer bestimmten Konformation zwischen 15% und 100% der Anordnung, was in allen Fällen über den Werten liegt, die vor der vorliegende Erfindung erhalten werden konnten. In Beispiel 13 wird weiters nachgewiesen, dass die Reinigung der Polypeptidmoleküle, bevor sie dem Verfahren der Erfindung unterzogen werden, den Anteil der Polypeptidmoleküle mit einer Konformation erhöht.
[0044] „Denaturierungsschritt" bezieht sich auf die Exposition einer Anordnung von Polypeptidmolekülen während eines Zeitintervalls unter physikalischen und/oder chemischen Umständen, welche die Anordnung von Polypeptidmolekülen Bedingungen aussetzen, die durch eine stärkere Denaturierungskraft charakterisiert sind als jene charakterisierenden Bedingungen unmittelbar vor dem Denaturierungsschritt.
[0045] Demgemäß bezieht sich der Begriff „Renaturierungsschritt" aus die Exposition einer Anordnung von Polypeptidmolekülen während eines Zeitintervalls unter physikalischen und/oder chemischen Umständen, welche die Anordnung von Polypeptidmolekülen Bedingungen aussetzen, die durch eine weniger starke Denaturierungskraft charakterisiert sind als jene charakterisierenden Bedingungen unmittelbar vor dem Denaturierungsschritt.
[0046] Es versteht sich, dass der oben erwähnte „beträchtliche Anteil" größenordnungsmäßig von der Anordnung von Polypeptidmolekülen abhängt, die dem Verfahren der Erfindung unterzogen werden. Wenn die prozessierte Anordnung von Polypeptiden aus monomeren Proteinen relativ kurzer Länge und ohne intramolekulare Disulfidbrücken besteht, wird das Verfahren im Allgemeinen sehr hohe Ausbeuten liefern, wogegen komplizierte Moleküle (wie z. B. polymere Proteine mit einer komplizierten Disulfidbrückentopologie) niedrigere Ausbeuten liefern können, sogar wenn die Bedingungen des Verfahrens der Erfindung vollkommen optimiert sind.
[0047] Ein interessanter Aspekt der Erfindung betrifft ein oben beschriebenes Verfahren, worin die prozessierte Anordnung einen beträchtlichen Anteil von Polypeptidmolekülen in einem bestimmten Konformationszustand umfasst, wobei der beträchtliche Anteil zumindest 1% (w/w) der anfänglichen Anordnung von Polypeptidmolekülen ausmacht. Höhere Ausbeuten sind bevorzugt, wie z. B. zumindest 5%, zumindest 10%, zumindest 20% und zumindest 25% der anfänglichen Anordnung von Polypeptidmolekülen. Bevorzugter sind Ausbeuten von zumindest 30%, wie z. B. zumindest 40%, 50%, 60%, 70% und zumindest 80%. Insbesondere bevorzugt sind Ausbeuten von zumindest 85%, wie z. B. 90%, 95%, 97% und sogar 99%. Manchmal werden Ausbeuten nahe 100% beobachtet.
[0048] Wenn die Polypeptidmoleküle der Anordnung Cystein enthalten, wird die prozessierte Anordnung einen beträchtlichen Anteil von Polypeptidmolekülen in einer bestimmten einheitlichen Konformation umfassen, die zusätzlich im Wesentlichen identische Disulfidbrückentopologie aufweisen.
[0049] In den meisten Fällen werden die Moleküle, die dem Verfahren der Erfindung unterzogen worden sind, Moleküle sein, die eine Aminosäuresequenz aufweisen, die mit derjenigen eines authentischen Polypeptids identisch ist, oder werden Moleküle sein, die eine Aminosäuresequenz aufweisen, die jener eines authentischen Polypeptids entsprechen, das an ein oder zwei zusätzliche Polypeptidsegmente gebunden ist.
[0050] Mit dem Ausdruck „authentisches Protein oder Polypeptid" ist ein Polypeptid mit einer Primärstruktur gemeint, die einschließlich der N- und C-terminalen Strukturen mit derjenigen des entsprechenden natürlichen Proteins identisch ist. Der Begriff bezeichnet ferner ein Polypeptid, das eine bekannte Primärstruktur aufweist, die nicht notwendigerweise mit derjenigen eines natürlichen Proteins identisch ist, wobei dieses Polypeptid das beabsichtigte Endprodukt einer Proteinsynthese ist.
[0051] Mit dem Ausdruck „natürliches Protein" ist ein Protein gemeint, wie es in biologisch aktiver Form aus

einem Organismus isoliert wird, in dem es nicht als Folge genetischer Manipulation vorhanden ist.

[0052] Im Gegensatz dazu bezieht sich der Ausdruck „künstliches Protein oder Polypeptid", wie er in der vorliegenden Patentbeschreibung und in den Ansprüchen verwendet wird, auf ein Protein/Polypeptid, das aus keinen natürlichen Quellen verfügbar ist, d. h., es kann aus keiner natürlichen Quelle isoliert und gereinigt werden. Ein künstliches Protein/Polypeptid ist folglich das Resultat eines menschlichen Eingriffes und kann beispielsweise das Produkt rekombinanter DNA-Manipulation oder eine Form der In-vitro-Peptidsynthese sein. Gemäß den obigen Definitionen kann ein solches künstliches Protein ein authentisches Protein sein, nicht jedoch ein natürliches Protein.

[0053] Folglich betrifft die Erfindung auch ein Verfahren, worin natürliche Proteine sowie künstliche Protein den hierin beschriebenen Neufaltungsprozessen unterzogen werden.

[0054] Wie unten ausführlicher beschrieben wird, kann es aus verschiedenen Gründen vorteilhaft sein, dass das authentische Polypeptid an Polypeptidsegmente mit Hilfsfunktionen während der Zyklisierung und anderen vorhergehenden oder nachfolgenden Prozessierungen gebunden ist, z. B. als „Handles" zur Bindung des Polypeptids an einen Träger, als Löslichkeitsmodifikatoren, als Expressionsverstärker, die ihre vorteilhafte Funktion während der Translation von mRNA ausüben usw. Ein solches Hilfspolypeptidsegment wird vorzugsweise über eine spaltbare Verbindung an das authentische Polypeptid gebunden, und wo zwei solcher Hilfspolypeptidsegmente an das authentische Polypeptid gebunden sind, kann dies über ähnliche spaltbare Verbindungen erfolgen, die normalerweise gleichzeitig gespalten werden, oder über unähnliche spaltbare Verbindungen, die in beliebiger Zeitabfolge gespalten werden können.

[0055] Gemäß dem oben Erklärten wird der Umstand für eine neue charakteristische Haupteigenschaft der vorliegende Erfindung gehalten, dass das Zyklisieren (das wie oben erwähnt zumindest zwei aufeinander folgende Zyklen umfasst) zumindest ein Ereignis verursacht, worin ein Renaturierungsschritt einem Denaturierungsschritt nachfolgt, worin zumindest ein beträchtlicher Anteil der neugefalteten Polypeptide wiederum denaturiert wird.

[0056] In den meisten Fällen wird die Prozessierung zumindest 5 Zyklen und häufiger zumindest 9 Zyklen umfassen, wie z. B. zumindest 10 Zyklen und in einigen Fällen zumindest 25 Zyklen. Andererseits wird die Reihe von Zyklen normalerweise 2.000 Zyklen nicht überschreiten und wird häufig höchstens 1.000 Zyklen und häufiger höchstens 500 Zyklen umfassen. Die Anzahl verwendeter Zyklen wird teilweise von den Möglichkeiten abhängen, die durch die Ausrüstung verfügbar gemacht werden, mit der die Zyklisierung durchgeführt wird.

[0057] Wenn folglich die Zyklisierungsbehandlung mit den an eine Säule immobilisierten Polypeptidmolekülen, wie z. B. unten ausführlicher beschrieben durchgeführt wird, stellt die Geschwindigkeit, mit der die Flüssigphase im Kontakt mit der Säule ausgetauscht werden kann, eine der realistisch erreichbaren Grenzen dar. Andererseits ermöglicht eine Hochleistungsflüssigchromatographie- (HPLC-) Ausrüstung den sehr schnellen Austausch der Flüssigkeitsumgebung und macht Zyklenzahlen im Bereich von Hunderten bis Tausenden realistisch.

[0058] Andere Gesichtspunkte, die die erwünschte Anzahl von Zyklen bestimmen, sind z. B. inhärente kinetische Parameter, wie z. B. gegenseitige Umwandlung zwischen cis- und trans-Isomeren an Prolinresten, die dazu neigen, die Neuverteilung über die teilweise gefalteten Zustände zu verkomplizieren, und werden daher normalerweise eine reifliche Überlegung der Zeitvorgabe erfordern. Eine weitere zeitkritische Eigenschaft liegt in Kinetik der Disulfid-Reshuffling (vgl. untenstehende Diskussion der Disulfid-Reshufflingsysteme).

[0059] Unter reiflicher Überlegung des Obigen werden die Zyklisierungsreihen häufig höchstens 200 Zyklen, häufiger höchstens 100 Zyklen und noch häufiger höchstens 50 Zyklen umfassen.

[0060] Gemäß dem oben Dargelegten kann die Dauer jedes Denaturierungsschritts eine Dauer sein, die unter den speziell betreffenden Bedingungen zumindest eine Millisekunde und höchstens eine Stunde beträgt, und die Dauer jedes Renaturierungsschritts kann eine Dauer sein, die unter den speziell betreffenden Bedingungen zumindest 1 Sekunde und höchstens 12 Stunden beträgt.

[0061] In den meisten Ausführungsformen des Verfahrens werden die denaturierenden Bedingungen jedes einzelnen Denaturierungsschritts im Wesentlichen für einen Zeitraum konstant gehalten und es werden die Renaturierungsbedingungen jedes einzelnen Renaturierungsschritts im Wesentlichen für einen Zeitraum konstant gehalten, wobei die Zeiträume, währenddessen die Bedingungen im Wesentlichen konstant gehalten werden, durch Übergangszeiträume getrennt sind, währenddessen die Bedingungen geändert werden. Der Übergangszeitraum zwischen den Schritten, für die die Bedingungen im Wesentlichen konstant gehalten werden, können eine Dauer aufweisen, die über einen breiten Bereich variieren, wie z. B. zwischen 0,1 Sekunden und 12 Stunden, und werden normalerweise an die Dauer der Denaturierungs- und Renaturierungsschritte genau und richtig angepasst.

[0062] Dieses beachtend kann der Zeitraum, für den die denaturierenden Bedingungen eines Denaturierungsschritts im Wesentlichen konstant gehalten werden, z. B. eine Dauer von zumindest einer Millisekunde und höchstens einer Stunde, häufig höchstens 30 Minuten haben, und der Zeitraum, für den die renaturierenden Bedingungen eines Renaturierungsschritts im Wesentlichen konstant gehalten werden, hat eine Dauer von zumindest 1 Sekunde und höchstens 12 Stunden, und häufig höchstens 2 Stunden.

[0063] In der Praxis wird der Zeitraum, für den die denaturierenden Bedingungen des Denaturierungsschritts im Wesentlichen konstant gehalten werden, häufig eine Dauer zwischen 1 und 10 Minuten haben, und der Zeitraum, für den die renaturierenden Bedingungen des Renaturierungsschritts im Wesentlichen konstant gehalten werden, wird häufig eine Dauer zwischen 1 und 45 Minuten haben.

[0064] Es versteht sich aus dem obigen, dass die oben angegebenen Intervalle angepasst werden sollten, wobei die Änderung der Kinetik, die sich aus der Änderung der physikalischen Bedingungen ergibt, denen die Polypeptide unterzogen werden, berücksichtigt werden. Beispielsweise kann der Druck sehr hoch sein (bis zu 5.000 Bar), wenn bei der Durchführung des Verfahrens der Erfindung ein HPLC-System verwendet wird, und unter solchen Umständen können sehr schnelle Schritte erzielt werden und/oder notwendig sein. Weiters ist, wie aus den Beispielen ersichtlich, der Temperaturparameter von Bedeutung, da sich einige Proteine nur bei Temperaturen weit vom physiologischen Bereich entfernt richtig falten werden. Temperatur sowie Druck werden selbstverständlich eine Wirkung auf die Kinetik des Neufaltungsverfahrens der Erfindung haben, und daher sind die oben angegebenen Zeitintervalle der Renaturierungs- und Denaturierungsschritte realistische Grenzen für die vielen möglichen Ausführungsformen der Erfindung.

[0065] Für eine gegebene Anwendung des Verfahrens der Erfindung wird der Fachkundige in der Lage sein, die geeigneten Bedingungen z. B. auf Basis von Vorversuchen zu ermitteln.

[0066] Wie oben angedeutet stehen die Polypeptidmoleküle während den Denaturierungs- und Renaturierungsschritten normalerweise im Kontakt mit einer Flüssigphase, wobei die Flüssigphase normalerweise eine wässrige Phase ist. Dies bedeutet, dass alle im Verfahren verwendeten Reagenzien und Hilfssubstanzen in der Flüssigphase, normalerweise in einer wässrigen Phase gelöst sind. Jedoch kann die Flüssigphase bei Bedarf auch durch ein oder mehrere organische Lösungsmittel gebildet werden.

[0067] Im Zusammenhang mit der Renaturierung von Proteinen ist die Verwendung eines so genannten „Chaperons" oder „Chaperon-Komplexes" bekannt. Chaperons sind eine Gruppe von kürzlich beschriebenen Proteinen, die eine gemeinsame Eigenschaft bezüglich ihrer Fähigkeit zur Verstärkung der Neufaltung von ungefalteten und teilweise gefalteten Proteinen aufweisen. Häufig sind die Chaperons multimolekulare Komplexe. Viele dieser Chaperons sind Hitzeschockproteine, was bedeutet, dass sie in vivo als Faktoren dienen, die eine posttraumatische „Reparatur" an Proteinen durchführen, die durch das Trauma destabilisiert worden sind. Um diese Funktion zu erfüllen, neigen die Chaperons dazu, gegen traumatische Ereignisse stabiler als viele andere Proteine und Proteinkomplexe zu sein. Obwohl das Verfahren der Erfindung nicht von der Verwendung eines molekularen Chaperons oder eines molekularen Chaperon-Komplexes abhängt, ist es selbstverständlich möglich, ein geeignetes molekulares Chaperon oder einen geeigneten molekularen Chaperon-Komplex zumindest in einem Renaturierungsschritt zu verwenden, und es kann bevorzugt sein, dass im Wesentlichen in allen Zyklen ein molekulares Chaperon oder ein molekularer Chaperon-Komplex vorhanden ist.

[0068] Wie oben erwähnt sind die Polypeptide vorzugsweise im Wesentlichen auf eine Umgebung beschränkt, die ein Ändern oder Austauschen der Flüssigphase im Wesentlichen ohne Mitreißen der Polypeptidmoleküle ermöglicht. Dies kann auf vielfältige Weise erzielt werden. Beispielsweise können die Polypeptidmoleküle in einem Dialysegerät enthalten sein, oder sie können auf eine der Phasen eines geeigneten Zweiphasensystems beschränkt sein. Solch ein geeignetes, wässriges Zweiphasensystem kann z. B. ein Polymer enthalten, das aus der aus Polyethylenoxid (Polyethylenglykol), Polyvinylacetat, Dextran und Dextransulfat bestehenden Gruppe gewählt ist. In einem interessanten Aufbau enthält eine der Phasen Polyethylenoxid (Polyethylenglykol) und die andere Phase Dextran, wodurch die Polypeptidmoleküle auf die Dextran enthaltende Phase beschränkt werden.

[0069] Ein weiterer Weg zur Vermeidung des Mitreißens des Polypeptids ist die Bindung der Polypeptidmoleküle an eine festen oder halbfesten Träger, wie z. B. eine Filteroberfläche, Hohlfaser oder ein gekörntes Chromatographiemedium, z. B. ein Agarose- oder Polyacrylamidgel, eine Cellulosefasermatrix oder eine HPLC- oder FPLC- ("Fast Performance Liquid Chromatography"-) Matrix. Als weitere Maßnahme kann der Träger eine Substanz mit Molekülen solcher Größe sein, dass die Moleküle mit den daran gebundenen Polypeptidmolekülen, wenn sie in einer Flüssigphase gelöst oder dispergiert sind, mit Hilfe eines Filters zurückgehalten werden können, oder der Träger kann eine Substanz sein, die zur Micellenbildung fähig ist, was die Änderung oder den Austausch der Flüssigphase im Wesentlichen ohne Mitreißen der Micellen ermöglicht. In Fällen, wo die Micellen-bildenden Komponenten dazu neigen, als Monomere aus dem System zu entkommen, z. B. wo sie in gewissem Ausmaß in der Lage wären, ein zur Beschränkung des Systems verwendetes Ultrafilter zu passieren, könnte dies durch Ergänzung mit zusätzlichem Micellen-bildendem Monomer kompensiert werden.

[0070] Der Träger kann auch ein wasserlösliches Polymer mit Molekülen einer Größe sein, die im Wesentlichen nicht in der Lage sind, die Poren eines Filters oder anderen Mittels zur Beschränkung des Systems zu passieren.

[0071] Die Polypeptidmoleküle werden in geeigneter Weise nicht kovalent über eine Gruppierung an den Träger adsorbiert, die eine Affinität zu einer Komponente des Trägers aufweist. Eine solche Gruppierung kann z. B. eine an einen Aminosäureteil des Polypeptidmoleküls gebundene Biotingruppe oder ein Analogon davon sein, wobei an den Träger Avidin, Streptavidin oder Analoga davon gebunden sind, um ein System mit einer

starken Affinität zwischen den auf diese Weise modifizierten Polypeptidmolekülen und dem auf diese Weise modifizierten Träger zu etablieren. Es versteht sich, dass die Affinität zwischen dem modifizierten Polypeptid und dem modifizierten Träger ausreichend stabil sein sollte, so dass die Adsorption durch die denaturierenden Bedingungen im Wesentlichen unbeeinflusst bleibt; die Entfernung der Polypeptidmoleküle vom Träger nach der Zyklisierung sollte unter Anwendung spezifischer Spaltung durchgeführt werden, so wie sie z. B. im Folgenden erläutert wird.

[0072] Ein Beispiel eines geeigneten Aminosäurerests, an den eine Biotinylgruppe gebunden werden kann, ist Lysin.

[0073] Einer der interessanten Wege der Einführung einer Aminosäure, die eine Gruppierung mit Affinität zum Träger trägt, ist die CPY-Synthese. CPY (Carboxypeptidase Y) ist bekanntermaßen dazu fähig, eine beliebiges Aminosäureamid ungeachtet der Natur der Seitenkette dieser Aminosäure anzufügen.

[0074] In einer interessanten Ausführungsform ist die Affinität zum Träger aufweisende Gruppierung das Polypeptidsegment Seq.-ID Nr. 47, wobei in diesem Fall der Träger zweckdienlicherweise ein mit Ni^{++}-Ionen beladenes Nitriloessigsäurederivat (NZA) ist, beispielsweise eine NTA-Agarosematrix, die in einer Ni^{++} umfassenden Lösung gebadet worden ist.

[0075] Ein wichtiger Aspekt der Erfindung betrifft die Gegenwart geeigneter Möglichkeiten im Polypeptidmolekül, das Molekül für spätere Spaltung in zwei oder mehr Segmente vorzubereiten, worin eines der Segmente ein wie oben definiertes authentisches Polypeptid ist. Solche kombinierten Polypeptidmoleküle (Fusionspolypeptidmoleküle) können zu diesem Zweck ein Polypeptidsegment umfassen, das dazu fähig ist, die bevorzugte Spaltung durch ein Spaltungsmittel an einer spezifischen Peptidbindung zu dirigieren. Das betreffende Polypeptidsegment kann eines sein, das die Spaltung als Ergebnis der Konformation des Segments dirigiert, das als eine Erkennungsstelle für das Spaltungsmittel dient.

[0076] Das die Spaltung dirigierende Polypeptidsegment kann beispielsweise in der Lage sein, die bevorzugte Spaltung an einer spezifischen Peptidbindung durch ein Spaltungsmittel zu dirigieren, dass aus der aus Cyanogenbromid, Hydroxylamin, Iodosobenzoesäure und N-Bromsuccinimid gewählten Gruppe besteht.

[0077] Das die Spaltung dirigierende Polypeptidsegment kann eines sein, das in der Lage ist, die bevorzugte Spaltung an einer spezifischen Peptidbindung durch ein Spaltungsmittel zu dirigieren, das ein Enzym ist, und ein solches mögliches Enzym ist Rinder-Enterokinase oder ein Analogon und/oder Homolog davon.

[0078] In einem wichtigen Aspekt der Erfindung ist das Spaltungsmittel das Enzym Rinderkoagulationsfaktor X_a oder ein Analogon und/oder Homolog davon (solche Analoga werden weiter unten ausführlicher erörtert), und das Polypeptidsegment, das die bevorzugte Spaltung dirigiert, ist eine Sequenz, die im Wesentlichen vom Rinderkoagulationsfaktor X_a oder einem Analogon und/oder Homolog davon selektiv erkannt wird. Wichtige solcher Segmente sind Polypeptidsegmente, die eine aus der aus Seq.-ID Nr. 38, Seq.-ID Nr. 40, Seq.-ID Nr. 41 und Seq.-ID Nr. 42 bestehenden Gruppe gewählte Sequenz aufweisen.

[0079] Eine interessante Eigenschaft der Erfindung ist die Möglichkeit der Maskierung und Demaskierung von Polypeptidsegmenten bezüglich ihrer Fähigkeit, die Spaltung an einer spezifischen Peptidbindung zu dirigieren, wodurch hervorgerufen wird, dass verschiedene Segmente des Polypeptids in verschiedenen Stadien der Zyklen gespalten werden können.

[0080] Wenn die Polypeptidmoleküle ein Polypeptidsegment umfassen, das in vitro in ein derivatisiertes Polypeptidsegment umwandelbar ist, das in der Lage ist, die bevorzugte Spaltung durch ein Spaltungsmittel an einer spezifischen Peptidbindung zu dirigieren, wird folglich ein Maskierungs/Demaskierungs-Effekt wie erwähnt verfügbar. Eine besonders interessante Version dieser Strategie liegt vor, wo das in vitro umwandelbare Polypeptidsegment in ein derivatisiertes Polypeptidsegment umwandelbar ist, das im Wesentlichen vom Rinderkoagulationsfaktor X_a oder einem Analogon und/oder Homolog davon selektiv erkannt wird.

[0081] Es ist vorgesehen, dass Cystein- sowie Methioninreste in modifizierte Reste umgewandelt werden können, wobei die modifizierten Reste die Segmente mit Aminosäuresequenzen, die aus der aus Seq.-ID Nr. 43, Seq.-ID Nr. 44, Seq.-ID Nr. 45 und Seq.-ID Nr. 46 bestehenden Gruppe gewählt sind, zu in vitro umwandelbaren Segmenten machen, die vom Rinderkoagulationsfaktor X_a oder eine Analogon und/oder Homolog davon erkannt werden.

[0082] Gemäß der Erfindung ist eine der möglichen, den Methioninrest umfassenden Lösungen, dass ein Polypeptidsegment mit der Aminosäuresequenz Seq.-ID Nr. 44 in ein derivatisiertes Polypeptid umgewandelt wird, das im Wesentlichen selektiv vom Rinderkoagulationsfaktor X_a erkannt wird, idem der Cysteinrest mit N-(2-Mercaptoethyl)morpholyl-2-thiopyridyldisulfid oder Mercaptothioacetat-2-thiopyridyldisulfid zur Reaktion gebracht wird.

[0083] Eine mögliche, Methionin umfassende Strategie gemäß der Erfindung ist, dass ein Polypeptidsegment mit der Aminosäuresequenz Seq.-ID Nr. 45 oder Seq.-ID Nr. 46 in ein derivatisiertes Polypeptid umgewandelt wird, das im Wesentlichen selektiv vom Rinderkoagulationsfaktor X_a erkannt wird, indem die Thioethergruppe in der Methioninseitengruppe zu einem Sulfoxid- oder Sulfonderivat oxidiert wird.

[0084] Bevorzugte Ausführungsformen des Verfahrens gemäß der Erfindung sind jene, worin die die Spaltung dirigierenden Segmente mit den Aminosäuresequenzen Seq.-ID Nr. 38, Seq.-ID Nr. 40, Seq.-ID Nr. 41 oder

Seq.-ID Nr. 42 oder die maskierten, die Spaltung dirigierenden Segmente mit den Aminosäuresequenzen Seq.-ID Nr. 43, Seq.-ID Nr. 44, Seq.-ID Nr. 45 und Seq.-ID Nr. 46 N-terminal an das authentische Polypeptid gebunden sind, da dann keine weitere Prozessierung außer der selektiven Spaltung notwendig ist, um das authentische Polypeptid in Lösung zu erhalten. Andererseits wäre einer der mögliche Gründe für das Binden der die Spaltung dirigierenden Sequenz am C-terminalen Ende des authentischen Polypeptids, dass die korrekte Faltung der Polypeptidmoleküle von einem freien N-Terminus der Polypeptidmoleküle abhängt. In einem solchen Fall kann der nach der Spaltung verbleibende Teil der die Spaltung dirigierenden Sequenz durch geeignete Verwendung der Carboxypeptidasen A und B entfernt werden.

[0085] Die Änderung der Bedingungen während der Übergangszeiträume zwischen den Schritten kann gemäß der Erfindung durch Ändern der chemischen Zusammensetzung der Flüssigphase, mit der die Polypeptidmoleküle in Kontakt stehen, erzielt werden. Folglich kann die Denaturierung der Polypeptidmoleküle durch Kontaktieren der Polypeptidmoleküle mit einer Flüssigphase erzielt werden, in der zumindest eine denaturierende Verbindung gelöst ist, und die Renaturierung der Polypeptidmoleküle wird durch Kontaktieren der Polypeptidmoleküle mit einer Flüssigphase erzielt, die entweder zumindest eine gelöste denaturierende Verbindung in einer solchen Konzentration enthält, dass der Kontakt mit der Flüssigphase die Anordnung von Polypeptidmolekülen in ihren entsprechenden Konformationszuständen, die sich aus dem vorhergehenden Schritt ergeben, tendenziell renaturieren und nicht denaturieren wird, oder im Wesentlichen keine denaturierende Verbindung enthält.

[0086] Der Ausdruck „denaturierende Verbindung" bezieht sich auf eine Verbindung, die, wenn sie als einer der Gelöststoffe in einer Polypeptidmoleküle umfassenden Flüssigphase vorhanden ist, gefaltete Zustände der Polypeptidmoleküle destabilisieren kann, was die teilweise oder vollständige Entfaltung der Polypeptidketten bewirken kann. Die von einer denaturierenden Verbindung ausgeübte denaturierende Wirkung erhöht sich mit steigender Konzentration der denaturierenden Verbindung in der Lösung, kann jedoch aufgrund der Anwesenheit anderer Gelöststoffe in der Lösung oder durch Änderungen der physikalischen Parameter, z. B. Temperatur oder Druck weiter verstärkt oder abgeschwächt werden.

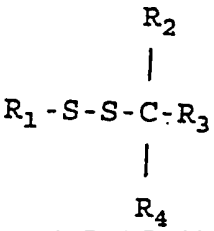
[0087] Als Beispiele von geeigneten, im Verfahren gemäß der Erfindung zu verwendenden denaturierenden Verbindungen können Harnstoff, Guanidin-HCl, Di-C_{1-6}-Alkylformamide, wie z. B. Dimethylformamid, und Di-C_{1-6}-Alkylsulfone genannt werden.

[0088] Die in zumindest einem der Denaturierungsschritte und/oder zumindest einem der Renaturierungsschritte verwendete Flüssigphase kann gemäß der Erfindung zumindest ein Disulfid-Reshufflingsystem enthalten.

[0089] „Disulfid-Reshufflingsysteme" sind Redoxsysteme, die Mischungen von reduzierenden und oxidierenden Mitteln enthalten, wobei deren Anwesenheit das Aufbrechen und die Herstellung von Disulfidbindungen in einem Polypeptid oder zwischen Polypeptiden erleichtert. Demgemäß sind „Disulfid-Reshufflingmittel" oder Disulfid-Reshufflingverbindungen" solche reduzierenden oder oxidierenden Mittel, welche das Aufbrechen oder die Herstellung von Disulfidbindungen in einem Polypeptid oder zwischen Polypeptiden erleichtern. In einem wichtigen Aspekt der Erfindung umfasst das Disulfid-Reshufflingsystem, das in der wässrigen Phase enthalten ist, die mit dem Protein in Kontakt steht, als ein Disulfid-Reshufflingsystem ein Gemisch eines Mercaptans und seine entsprechende Disulfidverbindung.

[0090] Als Beispiel könnten alle Cysteinreste in den Polypeptidmolekülen zu gemischten Disulfidprodukten von entweder Glutathion, Thiocholin, Mercaptoethanol oder Mercaptoessigsäure während zumindest einem der Denaturierungs/Renaturierungs-Zyklen umgewandelt worden sein. Ein solches umgewandeltes Polypeptid wird als „vollständig disulfidblockiertes Polypeptid oder Protein" bezeichnet, und dieser Ausdruck bezieht sich folglich auf ein Polypeptid oder ein Protein, in dem die Cysteinreste in ein gemischtes Disulfid umgewandelt worden sind, in dem jeder Cysteinrest an ein Mercaptan, z. B. Glutathion disulfidgebunden ist. Die Umwandlung der Cysteinreste zu gemischten Disulfidprodukten kann durch Reagieren einer vollständig denaturierten und vollständig reduzierten Anordnung von Polypeptidmolekülen mit einem Überschuss eines Reagens, das eine hochenergetische Mischung von Disulfidverbindungen ist, wie z. B. aliphatischaromatische Disulfidverbindungen, z. B. 2-Thiopyridylglutathionyldisulfid, oder durch ein beliebiges anderes geeignetes Verfahren erzielt werden.

[0091] Als Beispiele hochenergetischer gemischter Disulfide, das sind gemischte Disulfide mit einer relativ instabilen S-S-Bindung, können gemischte Disulfide mit der folgenden allgemeinen Formel erwähnt werden:



worin R_1 2-Pyridyl und R_2, R_3 und R_4 jeweils Wasserstoff oder eine gegebenenfalls substituierte, niedere aro-

matische oder aliphatische Kohlenwasserstoffgruppe ist. Beispiele solcher gemischten Disulfide sind Glutathionyl-2-thiopyridyldisulfid, 2-Thiocholyl-2-thiopyridyldisulfid, 2-Mercaptoethanol-2-thiopyridyldisulfid und Mercaptoacetat-2-thiopyridyldisulfid.

[0092] In interessanten Ausführungsformen enthält das Disulfid-Reshufflingsystem Glutathion, 2-Mercaptoethanol oder Thiocholin, wovon jedes ein Gemisch mit seinem entsprechenden symmetrischen Disulfid ist.

[0093] Die Eignung einer gegebenen Thiolmischung zur Verwendung als selektives Reduktionssystem und/oder Disulfid-Reshufflingsystem in einem zyklischen Neufaltungs/Reoxidationsverfahren für ein spezifisches Proteinprodukt kann direkt getestet werden, indem Anordnungen von Proben eines Gemischs von gefaltetem und fehlgefaltetem Protein mit einer Reihe von Thiolmischungen bei mehreren verschiedenen Konzentrationen von Denaturierungsmittel, die schwache, mittlere und starke denaturierende Wirkung auf das Protein ausüben, inkubiert werden. Nach der Inkubation wird die Disulfidtopologie in jeder Probe durch Reaktion mit einem Überschuss an thiolblockierendem Reagens (z. B. Iodacetamid) arretiert, bevor jeder Probensatz der SDS-PAGE und nicht reduzierenden Bedingungen unterzogen wird. Korrekt disulfidüberbrücktes Material und Material in unerwünschten kovalenten topologischen Zuständen werden als gesonderte Banden auftreten und werden daher die quantitative Bewertung des Faltungszustands des Proteins zur Zeit der Thioblockierung ermöglichen, da nur ein korrekt einzigartig disulfidgebundenes Topoisomer dem korrekt gefalteten Protein, das am Ende der Inkubation mit Thiol/Disulfid und denaturierenden Mitteln vorhanden ist, entsprechen wird. Dieser Satz von Experimenten erlaubt die Identifizierung einer Reihe von Denaturierungsmittelkonzentrationen, bei denen ein gegebenes Thiol/Disulfid-Reagens vorteilhaft als Disulfid-Reshuffling-Mittel verwendet werden kann, wie sich durch bevorzugte Reduktion und Reshuffling von falschen Disulfidbindungen und niedriger Tendenz zur Reduktion von Bindungen im vollständig gefalteten Protein zeigt. Dieses Verfahren der Reagenstestung kann als allgemeines Verfahren zur Auswahl vorteilhafter reduzierender Reagenzien und/oder Thiol/Disulfid-Reshufflingreagenzien verwendet werden. Beispiel 12 stellt die Anwendung dieses analytischen Verfahrens dar, um die Eignung für die selektive Reduktion fehlgefalteter Formen eines Modellproteins für 5 Thiolreagenzien dar und beweist dadurch die Durchführbarkeit des obigen Verfahrens.

[0094] Es versteht sich, dass das oben gekennzeichnete Verfahren zur Wahl geeigneter Disulfid-Reshufflingsysteme auch für die Wahl anderer Zusammensetzungen als Thiolgemische eingesetzt werden kann. Jedes Gemisch, das geeignete reduzierende/oxidierende Mittel enthält, kann gemäß dem oben gekennzeichneten Verfahren beurteilt werden, und die Zusammensetzung der Wahl im Verfahren der Erfindung wird diejenige sein, die die höchste Fähigkeit der bevorzugten Reduktion inkorrekt gebildeter Disulfidbrücken zeigt.

[0095] Folglich ist ein sehr wichtiger Aspekt der Erfindung ein Verfahren für die hierin beschriebene Proteinneufaltung, worin zumindest ein in Flüssigphase in zumindest einem Renaturierungs- und/der Denaturierungsschritt enthaltenes Disulfid-Reshufflingsystem eines ist, das zum Reduzieren und/oder Reshuffling inkorrekt gebildeter Disulfidbrücken unter Bedingungen bezüglich der Konzentration des denaturierenden Mittels fähig ist, bei der ungefaltete und/oder fehlgefaltete Proteine denaturiert werden und bei der im Wesentlichen keine Reduktion und/oder kein Reshuffling korrekt gebildeter Disulfidbrücken auftritt.

[0096] Eine interessante Ausführungsform der Erfindung ist ein wie oben beschriebenes Verfahren, worin ein Disulfid-Reshufflingsystem in zumindest einem Denaturierungs/Renaturierungsschritt verwendet wird und ein Verhältnis zwischen der relativen Menge von reduzierten/reshuffled, anfänglich inkorrekt gebildeten Disulfidbrücken und der relativen Menge von reduzierten/reshuffled, anfänglich korrekt gebildeten Disulfidbrücken von zumindest 1,05 liefert. Das Verhältnis wird vorzugsweise höher sein, wie z. B. 1,1, 1,5, 2,0, 3,0, 5,0, 10, 100, 1000, jedoch sind sogar höhere Verhältnisse realistisch und sind folglich gemäß der Erfindung insbesondere bevorzugt.

[0097] Mit den Ausdrücken „anfänglich inkorrekt/korrekt" bezüglich der Form der Disulfidbrücken ist die Brückentopologie gemeint, gerade bevor das Disulfid-Reshufflingsystem seine Wirkung ausübt.

[0098] Es versteht sich, dass das Verhältnis größer als 1 sein muss, um die Nettobildung korrekt gebildeter Disulfidbrücken in einer Proteinprobe zu ermöglichen. Normalerweise sollte das Verhältnis so hoch wie möglich sein, jedoch werden sogar Verhältnisse, die marginal über 1 liegen, die Nettobildung korrekt gebildeter Disulfidbrücken im Verfahren der Erfindung ermöglichen, wobei der wichtige Parameter zur Sicherstellung hoher Ausbeuten die Anzahl denaturierender/renaturierender Zyklen ist. Verhältnisse gerade über eins erfordern, dass viele Zyklen durchlaufen werden, bevor eine beträchtliche Ausbeute korrekt gebildeter Disulfidbrücken erzielt wird, wogegen hohe Verhältnisse nur eine beschränkte Anzahl von Zyklen erfordern.

[0099] In Fällen, wo nur ein Disulfid-Reshufflingsystem eingesetzt werden wird, kann ein solches Disulfid-Reshufflingsystem gemäß der Erfindung durch Folgendes gewählt werden:

1) Inkubieren von Proben des gefalteten und fehlgefalteten Proteins derselben Aminosäuresequenz wie das im Verfahren der Erfindung zu prozessierende Protein mit einer Reihe von Disulfid-Reshufflingsystemen bei verschiedenen Konzentrationen des gewählten Denaturierungsmittels,

2) Beurteilen der Fähigkeit von jedem der Disulfid-Reshufflingsysteme bei jeder der verschiedenen Konzentrationen des Denaturierungsmittels zum Reduzieren und/oder Reshuffling anfänglich inkorrekt gebildete Disulfidbrücken ohne wesentliches Reduzieren und/oder Reshuffling anfänglich richtig gebildeter Disulfidbrücken,

wie durch Berechnen des Verhältnisses zwischen der relativen Menge reduzierter/reshuffled, anfänglich inkorrekt gebildeter Disulfidbrücken und der relativen Menge reduzierter/reshuffled, anfänglich korrekt gebildeter Disulfidbrücken beurteilt wird, und

3) Wählen des Disulfid-Reshufflingsystems als das Disulfid-Reshufflingsystem X, das die Fähigkeit zur Reduktion anfänglich inkorrekt gebildeter Disulfidbrücken ohne wesentliches) Reduktion und/oder Reshuffling anfänglich korrekt gebildeter Disulfidbrücken im breitesten Konzentrationsbereich des gewählten Denaturierungsmittels zeigt.

[0100] Alternativ dazu kann mehr als ein Disulfid-Reshufflingsystem eingesetzt werden, beispielsweise in verschiedenen Zyklen im zyklischen Neufaltungsverfahren der Erfindung, jedoch auch gleichzeitig in denselben Zyklen. Dies wird z. B. der Fall sein, wenn es wahrscheinlich ist oder wenn z. B. mit dem oben dargestellten Verfahren festgestellt worden ist, dass die Gesamtausbeute des korrekt gefalteten Proteins mit der korrekten Disulfidbrückentopologie höher sein wird, wenn verschiedene Disulfid-Reshufflingsysteme im Verfahren der Erfindung verwendet werden.

[0101] Um das oben erwähnte Verhältnis zwischen der relativen Menge von reduzierten/reshuffled, anfänglich inkorrekt gebildeten Disulfidbrücken und der relativen Menge von reduzierten/reshuffled, anfänglich korrekt gebildeten Disulfidbrücken zu berechnen, kann das folgende Verfahren eingesetzt werden: Dem anfänglichen Gemisch von Reaktanten im Schritt 1) wird eine bekannte Menge radioaktiv markiertes, korrekt gefaltetes Protein zugegeben. Wenn die Mengen von korrekt und inkorrekt gefaltetem Protein im Schritt 2) abgeschätzt werden (zum Beispiel mittels nicht reduzierender SDS-PAGE), wird auch der Gehalt an Radioaktivität in der korrekt gefalteten Proteinfraktion ermittelt. Dadurch kann eine Abschätzung des nun inkorrekt gefalteten (jedoch anfänglich korrekt gefalteten) Proteins parallel mit der Bestimmung der Gesamtverteilung von korrekt/inkorrekt gefaltetem Protein vorgenommen werden. Das oben erwähnte Verhältnis kann folglich folgendermaßen berechnet werden:

$$R = \frac{C_2 - \frac{A_2}{A_1} \cdot C_1}{U_1 \cdot \frac{A_2}{A_1}}$$

worin C_1 und C_2 die anfänglichen bzw. endgültigen Mengen korrekt gefalteter Proteine sind, U_1 die Menge von anfänglich inkorrekt gefaltetem Protein und A_1 und A_2 die Radioaktivität in der anfänglichen korrekt gefalteten Proteinfraktion bzw. im endgültigen korrekt gefalteten Protein ist.

[0102] Zusätzlich zu den oben erwähnten Denaturierungsmitteln kann die Denaturierung auch durch Erniedrigung des pH-Wertes der Flüssigphase oder durch Erhöhung des pH-Werts der Flüssigphase erzielt werden.

[0103] Die Polarität der bei der Renaturierung verwendeten Flüssigphase kann gemäß der Erfindung durch Zugabe eines Salzes, eines Polymers und/oder einer Fluorwasserstoffverbindung, wie z. B. Trifluorethanol, modifiziert worden sein.

[0104] Gemäß der Erfindung kann die Denaturierung und Renaturierung der Polypeptidmoleküle auch durch direkte Änderungen der physikalischen Parameter erzielt werden, denen die Polypeptidmoleküle ausgesetzt sind, wie z. B. Temperatur oder Druck, oder es können diese Maßnahmen genützt werden, um die aus den anderen oben erwähnten Maßnahmen resultierende Denaturierung oder Renaturierung zu verstärken oder zu mäßigen.

[0105] Es versteht sich jedoch, dass eine höchst wichtige praktische Ausführungsform des Verfahrens durch Erzielen chemischer Veränderungen in der Flüssigphase durch Wechsel zwischen Denaturierungslösung A und Renaturierungslösung B durchgeführt wird. In diesem Fall wird die Konzentration einer oder mehrerer denaturierender Verbindungen in B häufig nach jedem Zyklus eingestellt, und als ein wichtiges Beispiel wird die Konzentration einer oder mehrerer denaturierender Verbindungen in B nach jedem Zyklus stufenweise vermindert, jedoch wird in einer weiteren wichtigen Ausführungsform die Konzentration einer oder mehrerer denaturierender Verbindungen im Medium B in jedem Zyklus konstant gehalten.

[0106] Diese Ausführungsform der Erfindung, worin die Konzentration der denaturierenden Verbindungen) in Medium B konstant gehalten wird, ist besonders interessant, wenn die produktivste Phase des zyklischen Prozesses (bezüglich des korrekt gefalteten Proteins) identifiziert worden ist und die Produktion von korrekt gefaltetem Protein im großen Maßstab erwünscht ist. Es versteht sich, dass die bevorzugte(n) Konzentrationen) der denaturierenden Verbindungen) des Mediums in dieser Ausführungsform diejenige(n) Konzentrationen) ist/sind, die etabliert worden ist/sind, um maximale Produktivität im zyklischen Prozess der Erfindung zu gewährleisten.

[0107] Die Polypeptidmoleküle der Anordnung, die dem Verfahren der Erfindung unterzogen wird, weisen normalerweise eine Länge von zumindest 25 Aminosäureresten auf, wie z. B. zumindest 30 Aminosäurereste oder zumindest 50 Aminosäurereste. Andererseits weisen die Polypeptidmoleküle der Anordnung normaler-

weise eine Länge von höchstens 5.000 Aminosäureresten auf, wie z. B. höchstens 2.000 Aminosäurereste oder höchstens 1.000 oder 800 Aminosäurereste.

[0108] Wie aus Beispiel 10 ersichtlich ist, hat das Verfahren der Erfindung die Produktion von korrekt gefalteten Diakörpermolekülen ermöglicht (Diakörper werden von Hollinger et al. (1993) beschrieben).

[0109] Ein wichtiger Aspekt der Erfindung betrifft daher ein Verfahren zur Herstellung korrekt gefalteter Diakörpermoleküle, worin eine anfängliche Anordnung von Polypeptidmolekülen, die ungefaltete und/oder fehlgefaltete Polypeptide mit Aminosäuresequenzen umfasst, die mit den Aminosäuresequenzen von Monomerfragmenten von Diakörpermolekülen identisch sind, einer Reihe von zumindest zwei aufeinander folgenden Zyklen unterzogen wird, wovon jeder eine Abfolge des Folgenden umfasst:

1) Zumindest ein Denaturierungsschritt, der Bedingungen umfasst, die einen denaturierenden und/oder entfaltenden Einfluss auf die Polypeptidmoleküle der Anordnung ausüben, um einen Anteil der Polypeptide in der Anordnung zu denaturieren und/oder zu entfalten, gefolgt von

2) zumindest einem Renaturierungsschritt, der Bedingungen umfasst, die einen renaturierenden Einfluss auf die Polypeptidmoleküle haben, die aus dem vorangegangenen Schritt resultierende Konformationen aufweisen, um einen Anteil der denaturierten und/ oder entfalteten Polypeptide der Anordnung zu renaturieren, wobei die Reihe von Zyklen so angepasst ist, dass ein beträchtlicher Anteil der anfänglichen Anordnung von Polypeptidmolekülen in einen Anteil korrekt gefalteter Diakörpermoleküle umgewandelt wird.

[0110] Ein solches Verfahren zur korrekten Faltung von Diakörpern kann in jedem der oben erwähnten Szenarien und Aspekte des Neufaltungsverfahrens der Erfindung ins Auge, das heißt, bezüglich der Wahl der physikalischen/chemischen Bedingungen sowie Zyklenabläufe. Jedoch ist ein wichtiger Aspekt des Verfahrens zur korrekten Faltung von Diakörpern ein Verfahren wie das oben gekennzeichnete, worin die Polypeptidmoleküle im Kontakt mit einer Flüssigphase stehen, die zumindest ein Disulfid-Reshufflingsystem in zumindest einem denaturierenden oder renaturierenden Schritt enthält. Das bevorzugte zu verwendende Denaturierungsmittel in einer solchen Flüssigphase ist Harnstoff und das bevorzugte Disulfid-Reshufflingsystem umfasst Glutathion als hautsächliches Reduktionsmittel.

[0111] Ein bestimmter Aspekt der Erfindung betrifft die Verwendung eines Polypeptids in den oben beschriebenen Verfahren, das ein Proenzym einer Serinprotease ist, die jedoch von jeglicher natürlich auftretenden Serinprotease verschieden ist und insbesondere eine Aminosäuresequenz aufweist, die von der des Rinderkoagulationsfaktors X (Protein Identification Resource (PIR), National Biomedical Research Foundation, Georgetown University, Medical Center, USA; Eintrag: P1;EXBO) verschieden ist und die proteolytisch aktiviert werden kann, um die aktive Serinprotease zu erzeugen, indem eine Lösung des Polypeptids in einem nicht denaturierenden Puffer mit einer Substanz inkubiert wird, die das Polypeptid spaltet, um einen neuen N-terminalen Rest freizusetzen, wobei die Substratspezifität der Serinprotease gleich oder besser als jene des Rinderblutkoagulationsfaktors X_a ist, wie sie durch die beiden Verhältnisse (k(I)/k(V) und k(III)/k(V) zwischen Spaltungsgeschwindigkeit für beide Substrate I und II:

I: Benzoyl-Val-Gly-Arg-p-Nitroanilid,

III: Tosyl-Gly-Pro-Arg-p-Nitroanilid,

gegen jene des Substrats

V: Benzoyl-Ile-Glu-Gly-Arg-p-Nitroanilid

bei 20°C, pH = 8 in einem Puffer, der aus 50 mM Tris, 100 mM NaCl, 1 mM $CaCl_2$ besteht, beurteilt werden kann, wobei es gleich oder niedriger als das entsprechende Verhältnis ist, das für Rinderkoagulationsfaktor X_a ermittelt wurde, der im Wesentlichen frei von kontaminierenden Proteasen ist.

[0112] Die Charakterisierung der oben identifizierten Polypeptide als Serinproteasen steht im Einklang mit der normalen Nomenklaturverwendung des Begriffs Serinprotease. Wie im Fach wohlbekannt ist, sind Serinproteasen Enzyme, von denen angenommen wird, dass sie ein katalytisches System aufweisen, das aus einem Aktivstellenserin besteht, das an einem Histidinrest ausgerichtet ist, und es wird angenommen, dass die Aktivierung des Enzyms aus dem entsprechenden Proenzymen auf der Freisetzung eines neuen N-terminalen Rests basiert, wovon die α-Aminogruppe in der Lage ist, sich innerhalb der Polypeptidstruktur neu zu positionieren, um eine Salzbrücke zu einem Asparaginsäurerest zu bilden, der einem Aktivstellenserinrest vorangeht, wodurch die für Serinproteasen charakteristische katalytische Stelle gebildet wird.

[0113] Die oben definierten „künstlichen" Serinproteasen äußerst wertvolle polypeptidspaltende Werkzeuge zur Verwendung im Verfahren der Erfindung und in anderen Verfahren, wo es entscheidend ist, ein Spaltungswerkzeug zu besitzen, das Protein selektiv spaltet, auch große, gefaltete Proteine. Analog zum Rinderkoagulationsfaktor X_a sind die oben definierten künstlichen Serinproteasen in aktivierter Form in der Lage, das die Spaltung dirigierende Polypeptidsegment Seq.-ID Nr. 38 selektiv zu erkennen, jedoch im Gegensatz zum Rinderkoagulationsfaktor X_a können sie mit solchen Aminosäuresequenzen aufgebaut werden, dass sie unter Anwendung rekombinanter DNA-Techniken leicht produziert werden können. Folglich sind die bevorzugten künstlichen Serinproteasen solche, die Aminosäuresequenzen aufweisen, die ihre Synthese durch rekombinante DNA-Techniken erlauben, insbesondere in einer prokaryotischen Zelle, wie z. B. E. coli. Wie aus der folgenden

Diskussion und den Beispielen hervorgeht, kann den künstlichen Serinproteasen, wenn sie in einem Prokaryoten produziert werden, durch Zyklisieren gemäß dem Verfahren der vorliegenden Erfindung eine enzymatisch aktive Konformation verliehen werden, in der die katalytisch aktiven Domänen in geeigneter Weise exponiert sind.

[0114] Der quantitative Test für die Selektivität der künstlichen Serinproteasen umfasst die Bestimmung der Spaltungsgeschwindigkeit k, die als Anfangssteigung einer Lichtabsorptionskurve bei 405 nm (Absorptionsmaximum des freien p-Nitroanilins) gegen die Zeit bei 20°C bestimmt wird.

[0115] Quantitativ ausgedrückt sollte die Selektivität der künstlichen Serinproteasen durch einen Wert (k(I)/k(V) von höchstens 0,06 und einen Wert k(III)/k(V) von höchstens 0,5 charakterisiert sein. Es wird bevorzugt, dass (k(I)/k(V) höchstens 0,05 und k(III)/k(V) höchstens 0,4 beträgt, und es wird stärker bevorzugt, dass (k(I)/k(V) höchstens 0,04 und k(III)/k(V) höchstens 0,15 beträgt.

[0116] Eine umfassendere Spezifitätscharakterisierung umfasst weitere Modellsubstrate: Folglich könnte die Substratspezifität als gleich oder besser derjenigen des Blutkoagulationsfaktors X_a durch jedes der Verhältnisse (k(I)/k(V), k(II)/k(V), k(III)/k(V) und k(IV)/k(V)) zwischen der Spaltungsgeschwindigkeit für jedes der Substrate I–IV:

I: Benzoyl-Val-Gly-Arg-p-Nitroanilid,
II: Tosyl-Gyl-Pro-Lys-p-Nitroanilid,
III: Tosyl-Gly-Pro-Arg-p-Nitroanilid,
IV: (d, 1)Val-Leu-Arg-p-Nitroanilid
gegen jene des Substrats
V: Benzoyl-Ile-Glu-Gly-Arg-p-Nitroanilid,
bei 20°C, pH 8 in einem Puffer bestehend aus 50 mM Tris, 100 mM NaCl, 1 mM $CaCl_2$ beurteilt werden, wobei es gleich oder niedriger als das entsprechende Verhältnis ist, das für Rinderkoagulationsfaktor X_a ermittelt wurde, der im Wesentlichen frei von kontaminierenden Proteasen ist.

[0117] Innerhalb dieser Charakterisierung sollte (k(I)/k(V) höchstens 0,06, k(III)/k(V) höchstens 0,0, k(III)/k(V) höchstens 0,5 und k(IV)/k(V)) höchstens 0,01 sein und es wird bevorzugt, dass (k(I)/k(V) höchstens 0,05, k(II)/k(V) höchstens 0,025, k(III)/k(V) höchstens 0,4 und k(IV)/k(V)) höchstens 0,008 ist, und es ist stärker bevorzugt, dass ((k(I)/k(V) höchstens 0,04, k(II)/k(V) höchstens 0,015, k(III)/k(V) höchstens 0,15 und k(IV)/k(V)) höchstens 0,005 ist.

[0118] Das wie oben definierte Serinprotease-artige Polypeptid wird normalerweise ein Molekulargewicht M_r von höchstens 70.000 und zumindest 15.000 aufweisen.

[0119] Ein solches Polypeptid besitzt die Aminosäuresequenz Seq.-ID Nr. 2 oder ist ein Analogon und/oder Homolog davon. Andere wichtige Ausführungsformen des Polypeptids der Erfindung besitzen eine Aminosäuresequenz, die eine Teilsequenz der Seq.-ID Nr. 2 oder das Analogon und/oder Homolog einer solchen Teilsequenz ist.

[0120] Mit der Verwendung des Ausdrucks „ein Analogon eines von der DNA-Sequenz kodierten Polypeptids" oder „ein Analogon eines Polypeptids mit der Aminosäuresequenz" ist jedes Polypeptid gemeint, das in den oben erwähnten Tests wie Rinderkoagulationsfaktor X_a arbeitet. Folglich sind auch Polypeptide aus anderen Quellen, wie z. B. aus anderen Säugetieren oder Wirbeltieren umfasst, die z. B. bis zu einem gewissen Grad bezüglich der Aminosäurezusammensetzung oder posttranslationellen Modifizierungen, z. B. Glykosylierung oder Phosphorylierung, im Vergleich zur im Beispiel beschriebenen künstlichen Serinprotease variieren.

[0121] Der Begriff „Analogon" wird folglich in vorliegenden Kontext verwendet, um ein Protein oder Polypeptid einer ähnlichen Aminosäurezusammensetzung oder Sequenz wie die charakteristische, von einer künstlichen, wie in Beispiel 5 beschriebenen Serinprotease hergeleiteten Seq.-ID Nr. 2 zu bezeichnen, was geringfügige, die Aminosäuresequenz verändernde Variationen erlaubt, z. B. Deletionen, ortsgerichtete Mutagenese, Insertionen zusätzlicher Aminosäuren oder Kombinationen davon, um künstliche Serinprotease-Analoga zu erzeugen.

[0122] Daher bezeichnet ein Analog in der vorliegenden Patentbeschreibung und in den Ansprüchen eine Variation des Polypeptids, in dem eine oder mehrere Aminosäuren deletiert oder ausgetauscht und/oder Aminosäuren eingeführt worden sein könnten, vorausgesetzt, dass die enzymatische Aktivität mit der oben definierten Spezifität erhalten bleibt, wie nach der obigen Beschreibung beurteilt werden kann.

[0123] Bezüglich der Homologie kann ein Analogon eines Polypeptids gemäß der Erfindung eine Sequenzhomologie auf der Polypeptidebene von zumindest 60% Identität im Vergleich zur Sequenz eines Fragments der Seq.-ID Nr. 2 aufweisen, was Deletionen und/ oder Insertionen von höchstens 50 Aminosäureresten erlaubt.

[0124] Solche Polypeptidsequenzen oder Analoga davon, die eine Homologie von zumindest 60% mit dem in Seq.-ID Nr. 2 gezeigten Polypeptid aufweisen, das von der DNA-Sequenz der Erfindung Seq.-ID Nr. 1 oder Analoga und/oder Homologe davon kodiert wird, stellen eine wichtige Ausführungsform dieser Erfindung dar.

[0125] Mit dem Begriff „Sequenzhomologie" ist die Identität bezüglich der Sequenz, entweder von Aminosäuren in Segmenten von zwei oder mehreren Aminosäuren in einer Aminosäuresequenz, oder von Nucleotiden

in Segmenten von zwei oder mehreren Nucleotiden in einer Nucleotidsequenz gemeint. Bezüglich des Polypeptids bedeuten die Begriffe daher eine Homologie zwischen den betreffenden Aminosäuren, zwischen denen die Homologie herzustellen ist, in Übereinstimmung bezüglich der Identität und Position der Aminosäuren des Polypeptids.

[0126] Der Begriff „homolog" wird folglich hierin verwendet, um den Grad der Identität zwischen der Aminosäuresequenz eines gegebenen Polypeptids und der in Seq.-ID Nr. 2 gezeigten Aminosäuresequenz zu erläutern. Die mit der in Seq.-ID Nr. 2 gezeigten Aminosäuresequenz zu vergleichenden Aminosäuresequenz kann aus einer Nucleotidsequenz, wie z. B. einer DNA- oder RNA-Sequenz, abgeleitet werden, die z. B. durch die im Folgenden definierte Hybridisierung oder durch herkömmliche Aminosäuresequenzierungsverfahren erhalten werden kann.

[0127] Eine weitere Ausführungsform betrifft ein Polypeptid mit einer Aminosäuresequenz, aus der eine fortlaufende Kette von 20 Aminosäuren mit einem Grad von zumindest 40% homolog zu einer Kette von Aminosäuren derselben Länge ist, die aus der in Seq.-ID Nr. 2 gezeigten Aminosäuresequenz gewählt ist.

[0128] Eines der Serinproteasepolypeptide besitzt die Aminosäuresequenz der Seq.-ID Nr. 2, Reste 82-484, oder ist ein Analogon und/oder Homolog davon. Ein weiteres Serinproteasepolypeptid gemäß der Erfindung besitzt die Aminosäuresequenz der Seq.-ID Nr. 2, Reste 166-484, oder ist ein Analogon und/oder Homolog davon.

[0129] Eine Reihe von Modifizierungen der hierin gezeigten Sequenzen ist besonders interessant: Die Insertion der die Spaltung dirigierenden Sequenzen Seq.-ID Nr. 38 oder 40–42 anstelle der Reste 230-233 in Seq.-ID Nr. 2, kombiniert mit dem Austausch des Cysteinrests 245, vorzugsweise durch Gly, Ser oder Arg in Seq.-ID Nr. 2. Eine weitere interessante Möglichkeit ist die Insertion von Seq.-ID Nr. 38 oder 40–42 anstelle der Reste 179-182 in Seq.-ID Nr. 2. Ganz allgemein wird in jeder der oben definierten künstlichen Serinproteasen der Ersatz der den Resten 230-233 in Seq.-ID Nr. 2 entsprechenden, spaltenden Sequenz durch eine der oben definierten, die Spaltung dirigierenden Sequenzen, äußert nützliche Spaltungsenzyme zur Verwendung im Verfahren gemäß der Erfindung hervorbringen, da diese selektiv und sehr effizient durch Enzyme mit der spezifischen enzymatischen Aktivität von Rinderkoagulationsfaktor X_a und folglich durch die oben definierten künstlichen Serinproteasen gespalten werden können, einschließlich durch Moleküle, die mit ihnen identisch sind. Die letztere Tatsache bedeutet, dass künstliche Serinproteasen, die durch eine solche Insertion der spezifisch die Spaltung dirigierenden Sequenzen modifiziert worden sind, höchst wirksam aktiviert werden können, so das die ersten gespaltenen und aktivierten Moleküle in der Lage sein werden, andere Moleküle zu spalten und folglich eine Kettenreaktion auszulösen.

[0130] Wie oben erwähnt, ist es eine höchst wichtige Eigenschaft, dass die künstlichen Serinproteasen durch rekombinante DNA-Techniken hergestellt werden können. Solche Techniken verwenden ein Nucleinsäurefragment, das zur Kodierung eines Polypeptids gemäß der obigen Definition fähig ist, insbesondere ein DNA-Fragment, das zur Kodierung eines wie oben definierten künstlichen Serinproteasepolypeptids fähig ist. Ein solches DNA-Fragment ist die in der DNA-Sequenz Seq.-ID Nr. 1 gezeigte Nucleotidsequenz oder ein Analogon davon, das eine Homologie mit irgendeiner der in Seq.-ID Nr. 1 gezeigten DNA-Sequenzen von zumindest 60% aufweist und/oder für ein Polypeptid kodiert, dessen Aminosäuresequenz zumindest zu 60% homolog mit den in Seq.-ID Nr. 2 gezeigten Aminosäuresequenzen ist.

[0131] Im Allgemeinen werden nur kodierende Regionen verwendet, wenn Nucleotidsequenzen verglichen werden, um ihre interne Homologie zu ermitteln.

[0132] Der Begriff „Analogon" bezeichnet bezüglich der DNA-Fragmente der Erfindung eine Nucleotidsequenz, die für ein Polypeptid kodiert, das mit dem Polypeptid identisch oder im Wesentlichen identisch ist, das von einem DNA-Fragment der Erfindung kodiert wird. Es ist wohlbekannt, dass dieselbe Aminosäure von verschiedenen Codons kodiert wird, wobei die Codonverwendung unter anderem mit der Präferenz des betreffenden Organismus in Beziehung stehen, der die Nucleotidsequenz exprimiert. Folglich können ein oder mehrere Nucleotide oder Codons des DNA-Fragments der Erfindung durch andere ersetzt werden, die, falls sie exprimiert werden, ein Polypeptid liefern, das identisch oder im Wesentlichen identisch mit dem vom betreffenden DNA-Fragment kodierten Polypeptid ist.

[0133] Darüber hinaus erlaubt der Begriff „Analogon" Variationen der Sequenz, wie z. B. Substitution, Insertion (einschließlich Introns), Addition und Neuordnung eines oder mehrerer Nucleotide, wobei die Variationen keine wesentliche Wirkung auf das vom DNA-Fragment kodierte Polypeptid haben, beispielsweise eine modifizierte Nucleotidsequenz, die sich von der in Seq.-ID Nr. 1 gezeigten DNA-Sequenz darin unterscheidet, dass zumindest ein Nucleotid substituiert, addiert, insertiert, deletiert und/oder neugeordnet worden ist.

[0134] Der Begriff „Substitution" bedeutet den Ersatz eines oder mehrerer Nucleotide in der vollständigen Nucleotidsequenz durch ein oder mehrere andere Nucleotide, „Addition" bedeutet die Addition eines oder mehrerer Nucleotide an eines der Enden der vollständigen Nucleotidsequenz, „Insertion" bedeutet die Einführung eines oder mehrerer Nucleotide innerhalb der vollständigen Nucleotidsequenz, „Deletion" bezeichnet, dass ein oder mehrere Nucleotide aus der vollständigen Nucleotidsequenz deletiert worden sind, ob an einem der Enden der Sequenz oder an jedem beliebigen Punkt im Inneren der Sequenz, und „Neuordnung" bedeutet, dass

zwei oder mehrere Nucleotidreste innerhalb der DNA bzw. des Polypeptids vertauscht worden sind. Das DNA-Fragment kann jedoch auch durch Mutagenese modifiziert werden, entweder vor oder nach seiner Insertion in den Organismus. Die DNA- oder Proteinsequenz der Erfindung kann solchermaßen modifiziert werden, dass sie überhaupt keine ihrer biophysikalischen, biochemischen oder biologischen Eigenschaften, oder einen Teil solcher Eigenschaften (eine und/ oder alle), oder alle dieser Eigenschaften (eine und/oder alle) verliert.

[0135] Ein Beispiel eines spezifischen Analogons ist eine DNA-Sequenz, welche die in Seq.-ID Nr. 1 gezeigte DNA-Sequenz umfasst und speziell zur Expression in E. coli angepasst ist. Diese DNA-Sequenz ist eine, die, wenn sie in E. coli gemeinsam mit geeigneten Regulationssequenzen insertiert wird, die Expression eines Polypeptids bewirkt, das im Wesentlichen die in Seq.-ID Nr. 2 gezeigte Aminosäuresequenz aufweist. Folglich umfasst diese DNA-Sequenz spezielle Codons, die von E. coli erkannt werden.

[0136] Die Begriffe „Fragment", „Sequenz", „Homolog" und „Analogon", wie sie in der vorliegenden Patentbeschreibung und den Ansprüchen bezüglich Fragmente, Sequenzen, Homologe und Analoga gemäß der Erfindung verwendet werden, sollten selbstverständlich dahingehend verstanden werden, dass sie diese Phänomene nicht in ihrer natürlichen Umgebung umfassen, sondern z. B. in isolierter, gereinigter, In-vitro- oder rekombinanter Form.

[0137] Ein solches Nucleinsäurefragment ist ein wie oben definiertes Nucleinsäurefragment, in dem zumindest 60% der kodierenden Triplets für dieselben Aminosäuren kodieren, wie ein Nucleinsäurefragment der Nucleinsäure, die für Rinderkoagulationsfaktor X kodiert, was Insertionen und/oder Deletionen von höchstens 150 Nucleotiden erlaubt. Ein Beispiel eines solchen Nucleinsäurefragments ist Seq.-ID Nr. 1, Nucleotide 76-1527, und Analoga und/oder Homologe davon. Ein weiteres Beispiel ist Seq.-ID Nr. 1, Nucleotide 319-1527, und Analoga und/oder Homologe davon. Noch eine weiteres Beispiel ist Seq.-ID Nr. 1, Nucleotide 571-1527, und Analoga und/oder Homologe davon.

[0138] Das oben beschriebene DNA-Fragment kann direkt aus genomischer DNA oder durch Isolieren von mRNA und ihr Umwandeln in die entsprechende DNA-Sequenz durch Verwendung reverser Transkriptase erhalten werden, wodurch eine cDNA produziert wird. Bei Verschaffung des DNA-Fragments aus genomischer DNA wird es direkt durch Screening auf genomische Sequenzen hergeleitet, wie dem Fachkundigen wohlbekannt ist. Dies kann durch Hybridisierung an eine DNA-Sonde erzielt werden, die auf Basis der Kenntnis der hierin beschriebenen Sequenzen, oder der durch Aminosäuresequenzierung einer gereinigten Serinprotease erhaltenen Sequenzinformation konstruiert wurde. Wenn die DNA von komplementärer DNA (cDNA) herrührt, kann sie durch Herstellen einer cDNA-Bibliothek mit mRNA aus Zellen erhalten werden, die eine künstliche Serinprotease enthalten. Die Hybridisierung kann mittels einer DNA-Sonde erzielt werden, die auf Basis der Kenntnis der cDNA-Sequenz oder der durch Aminosäuresequenzierung einer gereinigten künstlichen Serinprotease erhaltenen Sequenzinformation konstruiert wurde.

[0139] Das DNA-Fragment oder ein Analogon und/oder Homolog davon kann durch dessen Fusion mit einem Vektor und Insertieren des Komplexes in einen geeigneten Mikroorganismus oder in eine geeignete Säugetierzelllinie repliziert werden. Alternativ dazu kann das DNA-Fragment mittels chemischer Synthese hergestellt werden. Ferner können Polymerasekettenreaktions- (PCR-) Primer auf Basis der in Seq.-ID Nr. 1 gezeigten Nucleotidsequenz synthetisiert werden. Diese Primer können dann verwendet werden, um die gesamte oder einen Teil einer Sequenz zu amplifizieren, die für eine künstliche Serinprotease kodiert.

[0140] Geeignete Polypeptide zur Verwendung in der Erfindung können mittels rekombinanter DNA-Technologie hergestellt werden. Konkreter können die Polypeptide durch ein Verfahren hergestellt werden, das die Kultivierung oder Züchtung eines Organismus, der die in Seq.-ID Nr. 1 gezeigte DNA-Sequenz der Erfindung oder ein Analogon und/oder Homolog davon trägt, unter Bedingungen, die die Expression des besagten DNA-Fragments bewirken, und anschließend das Reinigen des exprimierten Polypeptids aus dem besagten Organismus umfasst.

[0141] Der Organismus, der zur Produktion des Polypeptids verwendet wird, kann ein höherer Organismus, z. B. ein Tier, oder ein niederer Organismus, z. B. ein Mikroorganismus sein. Ungeachtet der Art des verwendeten Organismus, sollte das DNA-Fragment der Erfindung (oben beschrieben) entweder direkt oder mithilfe eines geeigneten Vektors in den Organismus eingeführt werden. Alternativ dazu können die Polypeptide in den Säugetierzelllinien produziert werden, indem das DNA-Fragment der Erfindung oder ein Analogon und/oder Homolog davon entweder direkt oder mithilfe eines Expressionsvektors eingeführt wird.

[0142] Das DNA-Fragment kann auch in einen geeigneten stabilen Expressionsvektor kloniert und dann in eine geeignete Zelllinie platziert werden. Die die gewünschten Polypeptide exprimierenden Zellen werden dann unter denjenigen Bedingungen selektiert, die für den Vektor und die verwendete Zelllinie geeignet sind. Die selektierten Zellen werden dann weitergezüchtet und bilden eine wichtige und kontinuierliche Quelle der gewünschten Polypeptide.

[0143] Folglich wird hierin auch ein Expressionssystem veranschaulicht, dass ein wie oben definiertes Nucleinsäurefragment umfasst und für ein wie oben definiertes, künstliches Serinprotease-Polypeptidfragment kodiert, wobei das System eine 5'-flankierende Sequenz umfasst, die zur Vermittlung der Expression des besagten Nucleinsäurefragments fähig ist. Das Expressionssystem kann ein replizierbarer Expressionsvektor sein,

der das Nucleinsäurefragment trägt, wobei dieser Vektor fähig ist, sich in einem Wirtsorganismus oder einer Zelllinie zu replizieren; der Vektor kann z. B. ein Plasmid, Phage, Cosmid, Minichromosom oder Virus sein; der Vektor kann einer sein, der in das Wirtszellgenom integriert wird, wenn er in eine Wirtszelle eingeführt wird.

[0144] Weiters wird hierin ein Organismus erläutert, der das wie oben definierte Nucleinsäurefragment trägt und zu seiner Replikation fähig ist. Der Organismus kann ein Mikroorganismus sein, wie z. B. ein Bakterium, eine Hefe, ein Protozoon oder eine aus einem vielzelligen Organismus hergeleitete Zelle sein, wie z. B. ein Pilz, eine Insektenzelle, eine Pflanzenzellen, eine Säugetierzelle oder eine Zelllinie. Besonders interessante Wirtsorganismen sind Mikroorganismen, wie z. B. ein Bakterium der Gattung Escherichia, Bacillus oder Salmonella.

[0145] Eine oben definierte künstliche Serinprotease kann durch ein Verfahren hergestellt werden, das die folgenden Schritte umfasst:

1. Insertieren eines wie oben definierten Nucleinsäurefragments in einen Expressionsvektor,
1. Transformieren eines oben definierten Wirtsorganismus mit dem im Schritt a produzierten Vektor,
3. Kultivieren des in Schritt b produzierten Organismus, um das Polypeptid zu exprimieren,
4. Gewinnen des Polypeptids,
5. gegebenenfalls Unterziehen des Polypeptids der posttranslationellen Modifizierung,
6. wenn notwendig, Unterziehen des Polypeptids dem Denaturierungs/Renaturierungs-Zyklisierungsverfahren gemäß der vorliegende Erfindung und
7. gegebenenfalls Unterziehen des Polypeptid weiterer Modifizierung, um ein oben definiertes authentisches Polypeptid zu erhalten.

[0146] Weitere Modifizierungen der Polypeptide können beispielsweise erzielt werden, indem die Polypeptidmoleküle mit Carboxypeptidase A oder B behandelt werden, wodurch ausgewählte Aminosäurereste vom C-Terminus der Polypeptidmoleküle entfernt werden können. Dies ist unter Verhältnissen wünschenswert, bei denen die optimale Faltung der authentischen Polypeptidmoleküle nur dann erzielt wird, wenn der N-Terminus frei ist und das die Spaltung dirigierende Polypeptid (wie z. B. Seq.-ID Nr. 37) folglich C-terminal des authentischen Polypeptids platziert wird. Wie bekannt ist, spaltet Carboxypeptidase B nacheinander vom C-Terminus und spaltet nur basische Aminosäuren ab, wogegen Carboxypeptidase A nicht basische Aminosäuren abspaltet. Bei sorgfältiger Überlegung, welcher Rest C-terminal an das authentische Polypeptid angefügt wird, ist es möglich, zu gewährleisten, dass alle außer dem authentischen Polypeptid von der Carboxypeptidase abgespalten werden. Wenn der C-Terminus des authentischen Polypeptids eine basische Aminosäure ist, sollte man sicherstellen, dass der zu entfernende, C-terminal gebundene Rest nicht basisch ist und umgekehrt. Wenn einem die Sequenz der Aminosäurereste vom C-Terminus zum N-Terminus des authentischen Polypeptids bekannt ist, ist es möglich, zwischen Behandlungen mit den beiden Carboxypeptidasen abzuwechseln bis nur das nackte, authentische Polypeptid verbleibt. Eine praktische Ausführungsform wäre die Verwendung immobilisierter Carboxypeptidase.

[0147] Das produzierte Polypeptid kann durch ein Verfahren isoliert werden, das einen oder mehrere Schritte umfasst, wie z. B. Affinitätschromatographie unter Verwendung von immobilisiertem Polypeptid oder Antikörpern, die mit dem besagten Polypeptid reaktiv sind, und/oder durch andere chromatographische und elektrophoretische Verfahren.

[0148] Es versteht sich weiters, dass ein Polypeptid zur Verwendung in der Erfindung durch wohlbekannte Verfahren der Flüssig- oder Festphasen-Peptidsynthese unter Anwendung sukzessiver Kopplung der einzelnen Aminosäuren der Polypeptidsequenz hergestellt werden können. Alternativ dazu kann das Polypeptid durch Kopplung einzelner Aminosäuren synthetisiert werden, wobei Fragmente der Polypeptidsequenz gebildet werden, die später gekoppelt werden, um das gewünschte Polypeptid zu liefern. Diese Verfahren stellen daher einen weiteren interessanten Aspekt der Erfindung dar.

[0149] Die Erfindung betrifft ferner die Verwendung eines oben definierten künstlichen Serinproteasepolypeptids zur Spaltung von Polypeptiden an der Spaltstelle für Rinderkoagulationsfaktor X_a, wobei die Spaltstelle die Aminosäuresequenz aufweist, die aus der aus Seq.-ID Nr. 38, Seq.-ID Nr. 40, Seq.-ID Nr. 41 und Seq.-ID Nr. 42 bestehenden Gruppe gewählt ist, und betrifft die Verwendung eines wie oben definierten künstlichen Serinproteasepolypeptids zur Spaltung von Polypeptiden an der Spaltstelle für Rinderkoagulationsfaktor X_a, wobei die Spaltstelle eine modifizierte Version der Aminosäuresequenz aufweist, die aus der aus Seq.-ID Nr. 43, Seq.-ID Nr. 44, Seq.-ID Nr. 45 und Seq.-ID Nr. 46 bestehenden Gruppe gewählt ist, die wie weiter oben beschrieben in eine spaltbare Form umgewandelt worden ist.

Legenden zu den Figuren

[0150] **Fig.** 1: Schematische Darstellung eines Segments eines zyklischen Denaturierungs/Renaturierungs-Zeitprogramms.

[0151] Die Lösungsmittelzusammensetzung wird bezüglich eines binären Gemischs von einem nicht denatu-

rierenden „Puffer A" und einem denaturierenden „Puffer B" bezüglich des relativen Gehalts an Puffer B ausgedrückt. Drei aufeinander folgende Zyklen sind dargestellt, wobei jeder aus einer Renaturierungsphase „F" und einer Denaturierungsphase „D" besteht. Änderungen des Stärke der Denaturierungskraft des Lösungsmittelgemischs während der Denaturierungsphasen in aufeinander folgenden Zyklen sind mit „k" bezeichnet.

[0152] **Fig.** 2: Konstruktion des Expressionsplasmids pT_7H_6FX-hβ2m und pT_7H_6FX-mβ2m.

[0153] Die amplifizierten DNA-Fragmente, welche die Leseraster von Human- und Maus-β_2-Microglobulin von Aminosäurerest Ile_1 bis Met_{99}, enthalten, die am 5'-Ende an die für die FX_a-Spaltstelle kodierenden Nucleotidsequenzen (Seq.-ID Nr. 37) fusioniert sind, wurden unter Anwendung von Standardverfahren mit den Restriktionsendonucleasen Bam HI und Hind III (bezogen von Boehringer, Deutschland) geschnitten und mit T_4-DNA-Ligase (bezogen von Boehringer, Deutschland) in Bam HI- und Hind III-geschnittenes pT_7H6 ligiert.

[0154] **Fig.** 3: Aminosäuresequenzen von Human- und Maus-β_2-Microglobulin.

[0155] A: Die vorhergesagte Aminosäuresequenz des für Human-β_2-Microglobulin kodierenden Volllängen-Leserasters (Seq.-ID Nr. 49). Aminosäurerest Eins (Ile) im prozessierten reifen Protein ist gekennzeichnet. B. Vorhergesagte Aminosäuresequenz des für Maus-β_2-Microglobulin kodierenden Volllängen-Leserasters (Seq.-ID Nr. 50). Aminosäurerest Eins (Ile) im prozessierten reifen Protein ist gekennzeichnet.

[0156] **Fig.** 4: Konstruktion des Expressionsplasmids pT_7H_6FX-hGH. Das amplifizierte DNA-Fragment, das das an die für die FXa-Spaltstelle IEGR kodierende Nucleotidsequenz (Seq.-ID Nr. 38) fusionierte Leseraster des Human-Wachstumshormons von Aminosäurerest Phe_1 bis Phe_{191}, am 5'-Ende enthält, wurde mit den Restriktionsendonucleasen Bam HI und Hind III (bezogen von Boehringer, Deutschland) geschnitten und mit T_4-DNA-Ligase (bezogen von Boehringer, Deutschland) mittels Standardverfahren in mit Bam HI und Hind III geschnittenes pT_7H_6 ligiert.

[0157] **Fig.** 5: Aminosäuresequenz von Human-Wachstumshormon (Somatotropin).

[0158] Die vorhergesagten Aminosäuresequenz des für Human-Wachstumshormon kodierenden Volllängen-Leserasters (Seq.-ID Nr. 51). Der erste Aminosäurerest im prozessierten reifen Protein (Phe_1) ist gekennzeichnet.

[0159] **Fig.** 6: Konstruktion des Plasmids pT_7H_6FX-#1, #2 und #3, das den Aminosäurerest Nr. 20 (Ala) bis 109 (Arg), Aminosäurerest Nr. 20 (Ala) bis 190 (Ala) und Aminosäurerest Nr. 20 (Ala) bis 521 (Lys) von Human-α_2-Macroglobulin-Rezeptorprotein (α_2MR) (Seq.-ID Nr. 52) exprimiert.

[0160] Die vom Leseraster des α_2MR von #1: Aminosäurereste-Nr. 20 (Ala) bis 109 (Arg), #2: Aminosäurereste-Nr. 20 (Ala) bis 190 (Ala) und #3: Aminosäurereste-Nr. 20 (Ala) bis 521 (Lys) hergeleiteten, amplifizierten DNA-Fragmente, fusioniert am 5'-Ende an die für die FX_a-Spaltstelle IEGR kodierende Nucleotidsequenz (Seq.-ID Nr. 38), wurden mit den Restriktionsendonucleasen Bam HI und Hind III (bezogen vor Boehringer, Deutschland) geschnitten und mit T_4-DNA-Ligase (bezogen vor Boehringer, Deutschland) mittels Standardverfahren in mit Bam HI und Hind III geschnittenes pT_7H_6 ligiert.

[0161] **Fig.** 7: Konstruktion des Plasmids $pLcIIMLCH_6FX$-#4, #5 und #6, das den Aminosäurerest Nr. 903 (Gly) bis 1265 (Asp), Aminosäurerest Nr. 849 (Val) bis 1184 (Gln) und Aminosäurerest Nr. 1184 (Gln) bis 1582 (Lys) von Human-α_2-Macroglobulin-Rezeptorprotein (α_2MR) (Seq.-ID Nr. 52) exprimiert.

[0162] Die vom Leseraster des α_2MR von #4: Aminosäurereste-Nr. 803 (Gly) bis 1265 (Asp), #5: Aminosäurereste-Nr. 849 (Val) bis 1184 (Gln) und #6: Aminosäurereste-Nr. 1184 (Gln) bis 1582 (Lys) hergeleiteten, amplifizierten DNA-Fragmente, fusioniert am 5'-Ende an die für die FX_a-Spaltstelle IEGR kodierende Nucleotidsequenz (Seq.-ID Nr. 38), wurden mit den Restriktionsendonucleasen Bam HI oder Bcl und Hind III (bezogen vor Boehringer, Deutschland) geschnitten und mit T_4-DNA-Ligase (bezogen vor Boehringer, Deutschland) mittels Standardverfahren in mit Bam HI und Hind III geschnittenes $pLcIIMLCH_6FX$ ligiert.

[0163] **Fig.** 8: Konstruktion des Plasmids $pLcIIMLCH_6FX$-#7, #8 und #9, das den Aminosäurerest Nr. 803 (Gly) bis 1582 (Lys), Aminosäurerest Nr. 2519 (Ala) bis 2941 (Ile) und Aminosäurerest Nr. 3331 (Val) bis 3778 (Ile) von Human-α_2-Macroglobulin-Rezeptorprotein (α_2MR) (Seq.-ID Nr. 52) exprimiert.

[0164] Die vom Leseraster des α_2MR von #7: Aminosäurereste-Nr. 803 (Gly) bis 1582 (Lys), #8: Aminosäurereste-Nr. 2519 (Ala) bis 2941 (Ile) und #9: Aminosäurereste-Nr. 3331 (Val) bis 3778 (Ile) hergeleiteten, amplifizierten DNA-Fragmente, fusioniert am 5'-Ende an die für die FX_a-Spaltstelle IEGR kodierende Nucleotidsequenz (Seq.-ID Nr. 38), wurden mit den Restriktionsendonucleasen Bam HI und Hind III (bezogen vor Boehringer, Deutschland) geschnitten und mit T_4-DNA-Ligase (bezogen vor Boehringer, Deutschland) mittels Standardverfahren in mit Bam HI und Hind III geschnittenes $pLcIIMLCH_6FX$ ligiert.

[0165] **Fig.** 9a und 9b: Aminosäuresequenz von Human-α_2-Macroglobulin-Rezeptorprotein (α_2MR)(Seq.-ID Nr. 52).

[0166] Die vorhergesagte Aminosäuresequenz des für α_2MR kodierenden Volllängen-Leserasters. In den rekombinanten Proteinen als N- oder C-terminale Reste vorhandene Aminosäuren sind anhand ihrer Nummern über der α_2MR-Sequenz gekennzeichnet.

[0167] **Fig.** 10: Konstruktion des Expressionsplasmids $pLcIIMLCH_6FX$-FXΔγ.

[0168] Das amplifizierte DNA-Fragment, welches das am 5'-Ende an die für die FX_a-Spaltstelle IEGR kodierende Nucleotidsequenz (Seq.-ID Nr. 38) fusionierte Leseraster des Rinderblutkoagulationsfaktors X von

Aminsäurerest Ser_{82} bis Trp_{484} (FXΔγ) enthält, wurde mit den Restriktionsendonucleasen Bam HI und Hind III (bezogen vor Boehringer, Deutschland) geschnitten und mit T_4-DNA-Ligase (bezogen vor Boehringer, Deutschland) mittels Standardverfahren in mit Bam HI und Hind III geschnittenes pLcIIMLCH$_6$FX ligiert.

[0169] **Fig.** 11: Aminosäuresequenz des Rinderblutkoagulationsfaktors X (FX).

[0170] Die vorhergesagte Aminosäuresequenz des für Rinder-FX kodierenden Volllängen-Leserasters (Seq.-ID Nr. 53). Der N-terminale Aminosäurerest Ser_{82} und der C-terminale Trp_{484}-Rest im FXΔγ-Konstrukt sind gekennzeichnet.

[0171] **Fig.** 12: Konstruktion des Expressionsplasmids pLcIIMLCH$_6$FX-K1.

[0172] Das das Leseraster von Human-Plasminogen-Kringle 1 (K1) von Aminosäurerest Ser_{82} bis Glu_{162} (Nummerierung wie in „Glu"-Plasminogen) enthaltende, am 5'-Ende an die für die FXa-Spaltstelle IEGR (Seq.-ID Nr. 38) kodierende Nucleotidsequenz fusionierte, amplifizierte DNA-Fragment wurde mit den Restriktionsendonucleasen Bam HI und Hind III (bezogen vor Boehringer, Deutschland) geschnitten und mit T_4-DNA-Ligase (bezogen vor Boehringer, Deutschland) mittels Standardverfahren in mit Bam HI und Hind III geschnittenes pLcIIMLCH$_6$FX ligiert.

[0173] **Fig.** 13: Konstruktion des Expressionsplasmids pLcIIH$_6$FX-K4.

[0174] Das das Leseraster von Human-Plasminogen-Kringle 4 (K4) von Aminosäurerest Val_{354} bis Ala_{439} (Nummerierung wie in „Glu"-Plasminogen) enthaltende, am 5'-Ende an die für die FX_a-Spaltstelle IEGR (Seq.-ID Nr. 38) kodierende Nucleotidsequenz fusionierte, amplifizierte DNA-Fragment wurde mit den Restriktionsendonucleasen Bam HI und Hind III (bezogen vor Boehringer, Deutschland) geschnitten und mit T_4-DNA-Ligase (bezogen von Boehringer, Deutschland) mittels Standardverfahren in mit Bam HI und Hind III geschnittenes pLcIIH$_6$FX ligiert.

[0175] **Fig.** 14: Aminosäuresequenz von Human-„Glu"-Plasminogen (Seq.-ID Nr. 54). Die N- und C-terminalen Aminosäurereste in den K1- und K5-Konstrukten sind durch ihre Nummern in der Sequenz gekennzeichnet.

[0176] **Fig.** 15: SDS-PAGE-Analyse der Produktion und In-vitro-Faltung von rekombinantem β_2-Microglobulin.

[0177] Spur 1: Proteinrohextrakt vor Aufgabe auf die Ni^{2+}NTA-Agarosesäule (reduzierte Probe).

[0178] Spur 2: Säulendurchfluss während der Aufgabe des Proteinrohextrakts auf die Ni^{2+}NTA-Agarosesäule (reduzierte Probe).

[0179] Spur 3: Human-β_2-Microglobulin, eluiert aus der Ni^{2+}NTA-Agarosesäule nach dem zyklischen Faltungsverfahren durch den nicht denaturierenden Elutionspuffer (reduzierte Probe).

[0180] Spur 4: Proteinmarker (Pharmacia, Schweden): Vom oberen Gel-Ende; 94 kDa, 67 kDa, 43 kDa, 30 kDa 20,1 kDa und 14,4 kDa (reduzierte Probe).

[0181] Spur 5: Wie Spur 3 (nicht reduzierte Probe).

[0182] Spur 6: Rekombinantes Human-β_2-Microglobulin nach FX_a-Spaltung und Endreinigung (nicht reduzierte Probe).

[0183] **Fig.** 16: SDS-PAGE-Analyse der In-vitro-Faltung von rekombinantem Human-Wachstumshormon, hGH (Somatotropin).

[0184] Spur 1: Proteinmarker (Pharmacia, Schweden): Vom oberen Gel-Ende; 94 kDa, 67 kDa, 43 kDa, 30 kDa 20,1 kDa und 14,4 kDa (reduzierte Probe).

[0185] Spur 2: Human-hGH, eluiert aus der Ni^{2+}NTA-Agarosesäule nach dem zyklischen Faltungsverfahren durch den nicht denaturierenden Elutionspuffer (nicht reduzierte Probe).

[0186] Spur 3: Human-hGH, eluiert aus der Ni^{2+}NTA-Agarosesäule nach dem zyklischen Faltungsverfahren durch den denaturierenden Elutionspuffer B aus dem Faltungsverfahren (nicht reduzierte Probe).

[0187] Spuren 4–18: Fraktionen, die während der Trennung des monomeren hGH-Fusionsproteins von Dimer- und Multimer-Fusionsproteinen nach dem zyklischen Faltungsverfahren durch Ionenaustauschchromatographie an Q-Sepharose (Pharmacia, Schweden) gesammelt wurden. Das monomere Protein wurde in einem Peak eluiert, der vom Peak, der die Dimer- und Multimer-Proteine enthielt, gut getrennt war (nicht reduzierte Proben).

[0188] **Fig.** 17: SDS-PAGE-Analyse der In-vitro-Faltung von rekombinantem Kringle 1 und 4 aus Human-Plasminogen und rekombinantem Fusionsprotein #4, hergeleitet vom Humanα_2-Macroglobulin-Rezeptorprotein (α_2MR).

[0189] Spur 1: Proteinmarker (Pharmacia, Schweden): Vom oberen Gel-Ende; 94 kDa, 67 kDa, 43 kDa, 30 kDa 20,1 kDa und 14,4 kDa (reduzierte Probe).

[0190] Spur 2: K1-Fusionsprotein-Rohextrakt vor Aufgabe auf Ni^{2+}NTA-Agarosesäule (reduzierte Probe).

[0191] Spur 3: K1-Fusionsprotein, eluiert aus der Ni^{2+}NTA-Agarosesäule nach dem zyklischen Faltungsverfahren durch den nicht denaturierenden Elutionspuffer (reduzierte Probe).

[0192] Spur 4: Wie Spur 3 (nicht reduzierte Probe).

[0193] Spur 5: Durchfluss aus der Lysin-Agarose-Säule während der Auftragung des K1-Fusionsproteins (nicht reduzierte Probe).

[0194] Spur 6: K1-Fusionsprotein, eluiert aus der Lysin-Agarose-Säule (nicht reduzierte Probe).
[0195] Spur 7: K4-Fusionsprotein, eluiert aus der Ni^{2+}NTA-Agarosesäule nach dem zyklischen Faltungsverfahren durch den nicht denaturierenden Elutionspuffer (reduzierte Probe).
[0196] Spur 8: Wie Spur 7 (nicht reduzierte Probe).
[0197] Spur 9: α_2MR#4-Fusionsprotein, eluiert aus der Ni^{2+}NTA-Agarosesäule nach dem zyklischen Faltungsverfahren durch den nicht denaturierenden Elutionspuffer (reduzierte Probe).
[0198] Spur 10: Wie Spur 9 (nicht reduzierte Probe).
[0199] **Fig.** 18: Konstruktion des Expressionsplasmids pT_7H_7FX-α_2MRBDv.
[0200] Das das Leseraster von α_2-Macroglobulin von Aminosäurerest Val_{1299} bis Ala_{1451} enthaltende, am 5'-Ende an die für die FX_a-Spaltstelle IEGR (Seq.-ID Nr. 38) kodierende Nucleotidsequenz fusionierte, amplifizierte DNA-Fragment wurde mit den Restriktionsendonucleasen Bam HI und Hind III (bezogen vor Boehringer, Deutschland) geschnitten und mit T_4-DNA-Ligase (bezogen vor Boehringer, Deutschland) mittels Standardverfahren in mit Bam HI und Hind III geschnittenes pT_7H_6 ligiert.
[0201] **Fig.** 19: Aminosäuresequenz der Rezeptorbindungsdomäne von Human-α_2-Macroglobulin (von Rest Val_{1299} bis Ala_{1451}) (Seq.-ID Nr. 55).
[0202] **Fig.** 20: Konstruktion des Expressionsplasmids pT_7H_6FX-TETN.
[0203] Das das Leseraster von reifem monomeren Human-Tetranectin von Aminosäurerest Glu_1 bis Val_{181} enthaltende, am 5'-Ende an die für die FX_a-Spaltstelle IEGR (Seq.-ID Nr. 38) kodierende Nucleotidsequenz fusionierte, amplifizierte DNA-Fragment wurde mit den Restriktionsendonucleasen Bam HI und Hind III (bezogen vor Boehringer, Deutschland) geschnitten und mit T_4-DNA-Ligase (bezogen vor Boehringer, Deutschland) mittels Standardverfahren in mit Bam HI und Hind III geschnittenes pT_7H_6 ligiert.
[0204] **Fig.** 21: Aminosäuresequenz von monomerem Human-Tetranectin.
[0205] Die vorhergesagte Aminosäuresequenz des für Human-Tetranectin kodierenden Volllängenleserasters (Seq.-ID Nr. 56). Der erste Aminosäurerest im prozessierten, reifen Protein (Glu_1) ist gekennzeichnet.
[0206] **Fig.** 22: Konstruktion des Expressionsplasmids pT_7H_6FX-DB32.
[0207] Das das Leseraster des künstlichen Diakörpers DB32 von Aminosäurerest Gln_1 bis Asn_{246} enthaltende, am 5'-Ende an die für die FX_a-Spaltstelle IEGR (Seq.-ID Nr. 38) kodierende Nucleotidsequenz fusionierte, amplifizierte DNA-Fragment wurde mit den Restriktionsendonucleasen Bam HI und Hind III (bezogen vor Boehringer, Deutschland) geschnitten und mit T_4-DNA-Ligase (bezogen vor Boehringer, Deutschland) mittels Standardverfahren in mit Bam HI und Hind III geschnittenes pT_7H_6 ligiert.
[0208] **Fig.** 23: Aminosäuresequenz des künstlichen Diakörpers DB32 (Seq.-ID Nr. 57).
[0209] **Fig.** 24: Das Expressionsplasmid pT_7H_6FX-PS.4.
[0210] Die Konstruktion von pT_7H_6FX-PS.4, das Human-Psoriasin von Aminosäurerest Ser_2 bis Gln_{101} exprimiert, ist kürzlich beschreiben worden (Hoffmann (1994)).
[0211] **Fig.** 25: Aminosäuresequenz von Human-Psoriasin.
[0212] Die vorhergesagte Aminosäuresequenz des für Human-Psoriasin kodierenden Volllängen-Leserasters (Seq.-ID Nr. 58).
[0213] **Fig.** 26: SDS-PAGE-Analyse der Reinigung und FX_a-Spaltung des rekombinanten Mab 32-Diakörpers.

a: Verschiedene Stadien der Reinigung.

[0214] Spuren 1 und 2: Rohprodukt der Faltung.
[0215] Spur 3: Endgültiges gereinigtes Mab 32-Diakörper-Fusionsproteinprodukt.
[0216] Spur 4: Überstand des Faltungsrohprodukts nach 50facher Konzentrierung und Zentrifugation.
[0217] Spur 5: Pellet des Faltungsrohprodukts nach 50facher Konzentrierung und Zentrifugation.

b: FX_a-Spaltung des Mab 32-Diakörper-Fusionsproteins.

[0218] Spuren 1 und 5: Endgültiges gereinigtes Mab32-Diakörper-Fusionsprotein.
[0219] Spur 2: Molverhältnis 1 : 5 FX_a : Mab 32-Diakörper-Fusionsprotein bei 37°C für 20 Stunden.
[0220] Spur 3: Molverhältnis 1 : 2 FX_a : Mab 32-Diakörper-Fusionsprotein bei 37°C für 20 Stunden.
[0221] Spur 4: Molverhältnis 1 : 1 FX_a : Mab 32-Diakörper-Fusionsprotein bei 37°C für 20 Stunden.
[0222] **Fig.** 27: Eignung von Glutathion als Reduktionsmittel bei der zyklischen Neufaltung von Human-β_2-Microglobulin-Fusionsprotein.
[0223] Spur 1: Reduzierte Probe des Tests Nr. 1.
[0224] Spur 2: Nicht reduzierte Probe des Tests Nr. 1.
[0225] Spur 3: Nicht reduzierte Probe des Tests Nr. 2.
[0226] Spur 4: Nicht reduzierte Probe des Tests Nr. 3.
[0227] Spur 5: Nicht reduzierte Probe des Tests Nr. 4.
[0228] Spur 6: Nicht reduzierte Probe des Tests Nr. 5.

[0229] Spur 7: Nicht reduzierte Probe des Tests Nr. 6.
[0230] Spur 8: Nicht reduzierte Probe des Tests Nr. 7.
[0231] Spur 9: Nicht reduzierte Probe des Tests Nr. 8.
[0232] Spur 10: Nicht reduzierte Probe des Tests Nr. 9.
[0233] Spur 11: Nicht reduzierte Probe des Tests Nr. 10.
[0234] Spur 12: Nicht reduzierte Probe des Tests Nr. 11.
[0235] **Fig.** 28: Eignung von L-Cystein-Ethylester als Reduktionsmittel bei der zyklischen Neufaltung von Human-β_2-Microglobulin-Fusionsprotein.
[0236] Spur 1: Reduzierte Probe des Tests Nr. 1.
[0237] Spur 2: Nicht reduzierte Probe des Tests Nr. 1.
[0238] Spur 3: Nicht reduzierte Probe des Tests Nr. 2.
[0239] Spur 4: Nicht reduzierte Probe des Tests Nr. 3.
[0240] Spur 5: Nicht reduzierte Probe des Tests Nr. 4.
[0241] Spur 6: Nicht reduzierte Probe des Tests Nr. 5.
[0242] Spur 7: Nicht reduzierte Probe des Tests Nr. 6.
[0243] Spur 8: Nicht reduzierte Probe des Tests Nr. 7.
[0244] Spur 9: Nicht reduzierte Probe des Tests Nr. 8.
[0245] Spur 10: Nicht reduzierte Probe des Tests Nr. 9.
[0246] **Fig.** 29: Eignung von 2-Mercaptoethanol als Reduktionsmittel bei der zyklischen Neufaltung von Human-β_2-Microglobulin-Fusionsprotein.
[0247] Spur 1: Reduzierte Probe des Tests Nr. 1.
[0248] Spur 2: Nicht reduzierte Probe des Tests Nr. 1.
[0249] Spur 3: Nicht reduzierte Probe des Tests Nr. 2.
[0250] Spur 4: Nicht reduzierte Probe des Tests Nr. 3.
[0251] Spur 5: Nicht reduzierte Probe des Tests Nr. 4.
[0252] Spur 6: Nicht reduzierte Probe des Tests Nr. 5.
[0253] Spur 7: Nicht reduzierte Probe des Tests Nr. 6.
[0254] Spur 8: Nicht reduzierte Probe des Tests Nr. 7.
[0255] Spur 9: Nicht reduzierte Probe des Tests Nr. 8.
[0256] Spur 10: Nicht reduzierte Probe des Tests Nr. 9.
[0257] **Fig.** 30: Eignung von Mercaptobernsteinsäure als Reduktionsmittel bei der zyklischen Neufaltung von Human-β_2-Microglobulin-Fusionsprotein.
[0258] Spur 1: Nicht reduzierte Probe des Tests Nr. 1.
[0259] Spur 2: Nicht reduzierte Probe des Tests Nr. 2.
[0260] Spur 3: Nicht reduzierte Probe des Tests Nr. 3.
[0261] Spur 4: Nicht reduzierte Probe des Tests Nr. 4.
[0262] Spur 5: Nicht reduzierte Probe des Tests Nr. 5.
[0263] Spur 6: Nicht reduzierte Probe des Tests Nr. 6.
[0264] Spur 7: Nicht reduzierte Probe des Tests Nr. 7.
[0265] Spur 8: Nicht reduzierte Probe des Tests Nr. 8.
[0266] Spur 9: Nicht reduzierte Probe des Tests Nr. 9.
[0267] **Fig.** 31: Eignung von N-Acetyl-L-Cystein als Reduktionsmittel bei der zyklischen Neufaltung von Human-β_2-Microglobulin-Fusionsprotein.
[0268] Spur 1: Reduzierte Probe des Tests Nr. 1.
[0269] Spur 2: Nicht reduzierte Probe des Tests Nr. 1.
[0270] Spur 3: Nicht reduzierte Probe des Tests Nr. 2.
[0271] Spur 4: Nicht reduzierte Probe des Tests Nr. 3.
[0272] Spur 5: Nicht reduzierte Probe des Tests Nr. 4.
[0273] Spur 6: Nicht reduzierte Probe des Tests Nr. 5.
[0274] Spur 7: Nicht reduzierte Probe des Tests Nr. 6.
[0275] Spur 8: Nicht reduzierte Probe des Tests Nr. 7.
[0276] Spur 9: Nicht reduzierte Probe des Tests Nr. 8.
[0277] Spur 10: Nicht reduzierte Probe des Tests Nr. 9.
[0278] **Fig.** 32: SDS-PAGE-Analyse der zyklischen Neufaltung des Human-β_2-Microglobulin-Fusionsproteins.
[0279] Spur 1: Proteinrohextrakt vor Aufgabe auf die Ni^{2+}NTA-Agarosesäule (reduzierte Probe).
[0280] Spur 2: 8 µl-Probe der löslichen Fraktion von neugefaltetem hβ_2m, wie beschrieben in Beispiel 1.
[0281] Spur 3: 4 µl-Probe der löslichen Fraktion von neugefaltetem hβ_2m, wie beschrieben in Beispiel 1.
[0282] Spur 4: 2 µl-Probe der löslichen Fraktion von neugefaltetem hβ_2m, wie beschrieben in Beispiel 1.
[0283] Spur 5: 8 µl-Probe der unlöslichen Fraktion von neugefaltetem hβ_2m, wie beschrieben in Beispiel 1.
[0284] Spuren 6 und 7: hβ_2m-Endprodukt nach Reinigung durch Ionenaustauschchromatographie.

[0285] Spuren 8 und 9: Neugefaltetes hβ_2m nach optimiertem Neufaltungsprotokoll, wie beschrieben in Beispiel 13.
[0286] **Fig.** 33: SDS-PAGE-Analyse der Neufaltung von Human-β_2-Microglobulin-Fusionsprotein durch Pufferschritt und linearen Gradienten.
[0287] Spur 1: Probe der löslichen Fraktion von neugefaltetem hβ_2m, gefaltet durch das in Beispiel 13 beschriebene Pufferschrittprotokoll.
[0288] Spur 2 und 3: Probe der unlöslichen Fraktion von neugefaltetem hβ_2m, gefaltet durch das in Beispiel 13 beschriebene Pufferschrittprotokoll.
[0289] Spur 4: Proteinmarker (Pharmacia, Schweden): Vom oberen Gel-Ende; 94 kDa, 67 kDa, 43 kDa, 30 kDa 20,1 kDa und 14,4 kDa (reduzierte Probe).
[0290] Spur 5: Probe löslicher Fraktion von neugefaltetem hβ_2m, gefaltet durch das in Beispiel 13 beschriebene lineare Gradientenprotokoll.
[0291] Spuren 6 und 7: Probe unlöslicher Fraktion von neugefaltetem hβ_2m, gefaltet durch das in Beispiel 13 beschriebene lineare Gradientenprotokoll.
[0292] **Fig.** 34: Das allgemeine Schema der Konstruktion der in den Beispielen beschriebenen Fusionsproteine.
[0293] In das N-terminale Ende des Fusionsproteins wird gegebenenfalls ein „Booster-Segment" insertiert, das die Expressionsstärke des Fusionsproteins in der Zelle verstärkt, die die für das Fusionsprotein kodierende DNA exprimiert. C-terminal davon bezeichnet „6H" die 6 Histidinylreste, die eine Ionenkomplexierungsstelle darstellen, die als „Affinitäts-Handle" während der Reinigung und Neufaltung der Fusionsproteine dient. Das „FX" am C-Terminus der 6 Histidinyl-Stelle ist die FX_a-Spaltstelle. Schließlich repräsentiert der mit „Protein" bezeichnete Teil des Fusionsproteins das Protein, das gemäß dem Verfahren der Erfindung neu gefaltet werden wird.

Beispiele

[0294] Die in diesem Abschnitt angeführten Beispiele 1 bis 11, die der Erläuterung des „zyklischen Faltungsverfahrens" dienen, beschreiben alle den Prozess der Faltung eines rekombinanten, spaltbaren, in E. coli produzierten Hybridproteins (Fusionsproteins), das durch ein einziges, allgemeines Verfahren aus einem Proteinrohextrakt gereinigt und ohne weitere Reinigung der Faltung unterzogen wurde.
[0295] Die herzustellende, für das rekombinante Protein kodierende Nucleotidsequenz ist am 5'-Ende an eine Nucleotidsequenz fusioniert, die für eine Aminosäuresequenz, kodiert, die eine FX_a-Spaltstelle (FX) definiert, die ihrerseits N-terminal an ein sechs Histidinylreste (Seq.-ID Nr. 47) enthaltendes Segment gebunden ist. Das Binden der FX_a-Spaltstelle wird normalerweise während einer Polymerasekettenreaktion erzielt, worin der 5'-terminale Primer Nucleotide umfasst, die für diese Sequenz kodiert. Das Binden der sechs Histidinylreste wird normalerweise durch Einsetzen eines Vektors erreicht, der ein für die Seq.-ID Nr. 47 kodierendes Nucleotidfragment umfasst. Die sechs Histidinylreste bilden eine Metallionen komplexierende Stelle, die als Affinitäts-Handle während der Reinigung des Fusionsproteins und anschließend als Kontaktpunkt zur Festphasenmatrix während des zyklischen Faltungsprozesses genützt wird. Gelegentlich werden „Booster-Segmente" (z. B. ein vom N-Terminus des λcII-Proteins hergeleitetes Segment, an das sich in manchen Fällen ein von der Myosin-Leichtkette hergeleitetes Segment anschließt) N-terminal des Affinitäts-Handles insertiert, um die Expressionsstärke des Fusionsproteins in E. coli zu verbessern.
[0296] Die Fusionsproteine sind alle nach demselben allgemeinen Schema konstruiert (vgl. **Fig.** 34). Die Anwesenheit von Booster-Segmenten, Affinitäts-Handle und FX_a-Spaltstelle könnte die Neufaltung des rekombinanten Proteins von Interesse komplizieren. Darüber hinaus wird der zyklische Faltungsprozess unmittelbar nach der Affinitätsreinigung des Fusionsproteins begonnen. Dies bedeutet, dass Fusionsmaterial, das teilweise vom E. coli-Wirt abgebaut worden ist, an der Affinitätsmatrix zusätzlich zum Volllängen-Fusionsprotein an der Säule zurückgehalten wird. Dieses abgebaute Fusionsprotein kann die Neufaltung des Volllängen-Fusionsproteins sehr wohl ernstlich stören, wodurch die scheinbare Effizienz des Prozesses herabgesetzt wird. Die in Beispielen 1 bis 11 dargelegten Faltungseffizienzergebnisse können daher nicht direkt mit der Effizienz des Prozesses der Neufaltung eines gereinigten Fusionsproteins verglichen werden.
[0297] Die Beispiele 1 bis 11 beschreiben das Neufaltungsverfahren für 21 verschiedene Proteine, Proteindomänen oder Domänen-Cluster im Größenbereich von 82 Aminosäuren (K1, Beispiel 6) bis 780 Aminosäuren (α_2MR#7, Beispiel 4), und die Anzahl der Disulfidbrücken in den Proteinen liegen im Bereich von Null (α_2MRAP, Beispiel 3) bis 33 (α_2MR#4, Beispiel 4) und 36 (α_2MR#7, Beispiel 4).
[0298] Die Effizienz der Neufaltung der Proteine liegt im Bereich von 15 bis 95% und die Ausbeute von aktivem Protein liegt in der Größenordnung von 10–100 mg für die Neufaltung an einer 40 ml-Ni + NTA-Agarosesäule (NTA bezeichnet eine substituierte Nitrilotriessigsäure).
[0299] Die folgenden Tabellen 1–5 stellen die in den Beispielen verwendeten Gradientenprofile dar. „Zeit" ist in Minuten angegeben und „Fluss" in ml/min.

Tabelle 1

Schritt	Zeit	Fluss	%A	%B
1	0	2	100	0
2	45	2	100	0
3	46	2	0	100
4	52	2	0	100
5	60	2	100	0
6	105	2	100	0
7	106	2	4	96
8	113	2	4	96
9	120	2	100	0
10	165	2	100	0
11	166	2	8	92
12	172	2	8	92
13	180	2	100	0
14	225	2	100	0
15	226	2	12	88
16	232	2	12	88
17	240	2	100	0
18	285	2	100	0
19	286	2	16	84
20	292	2	16	84
21	300	2	100	0
22	345	2	100	0
23	346	2	20	80
24	352	2	20	80
25	360	2	100	0
26	405	2	100	0
27	406	2	24	76
28	412	2	24	76
29	420	2	100	0
30	465	2	100	0
31	466	2	28	72
32	472	2	28	72
33	480	2	100	0
34	525	2	100	0
35	526	2	32	68
36	532	2	32	68
37	540	2	100	0
38	585	2	100	0
39	586	2	36	64
40	592	2	36	64
41	600	2	100	0
42	645	2	100	0
43	646	2	40	60
44	652	2	40	60
45	660	2	100	0
46	705	2	100	0
47	706	2	44	56
48	713	2	44	56
49	720	2	100	0
50	765	2	100	0
51	766	2	48	52
52	772	2	48	52
53	780	2	100	0
54	825	2	100	0
55	826	2	52	48
56	832	2	52	48
57	840	2	100	0
58	885	2	100	0
59	886	2	56	44
60	892	2	56	44
61	900	2	100	0
62	945	2	100	0
63	946	2	60	40
64	952	2	60	40
65	960	2	100	0
66	1005	2	100	0
67	1006	2	62	38
68	1012	2	62	38
69	1020	2	100	0
70	1065	2	100	0
71	1066	2	64	36
72	1072	2	64	36
73	1080	2	100	0
74	1125	2	100	0
75	1126	2	66	34
76	1132	2	66	34
77	1140	2	100	0
78	1185	2	100	0
79	1186	2	68	32
80	1192	2	68	32
81	1200	2	100	0
82	1245	2	100	0
83	1246	2	70	30
84	1252	2	70	30
85	1260	2	100	0
86	1305	2	100	0
87	1306	2	72	28
88	1312	2	72	28
89	1319	2	100	0
90	1364	2	100	0
91	1365	2	74	26
92	1371	2	74	26
93	1378	2	100	0
94	1423	2	100	0
95	1424	2	76	24
96	1430	2	76	24
97	1437	2	100	0
98	1482	2	100	0
99	1483	2	78	22
100	1489	2	78	22
101	1496	2	100	0
102	1541	2	100	0
103	1542	2	80	20
104	1548	2	80	20
105	1555	2	100	0
106	1556	2	82	18
107	1562	2	82	18
108	1569	2	100	0
109	1614	2	100	0
110	1615	2	84	16
111	1621	2	84	16
112	1628	2	100	0
113	1673	2	100	0
114	1674	2	88	12
115	1732	2	88	12
116	1733	2	100	0
117	1778	2	100	0

Tabelle 2

Schritt	Zeit	Fluss	%A	%B
1	0	2	100	0
2	45	2	100	0
3	46	2	0	100
4	52	2	0	100
5	60	2	100	0
6	105	2	100	0
7	106	2	8	92
8	113	2	8	92
9	120	2	100	0
10	165	2	100	0
11	166	2	20	80
12	172	2	20	80
13	180	2	100	0
14	225	2	100	0
15	226	2	28	72
16	232	2	28	72
17	240	2	100	0
18	285	2	100	0
19	286	2	34	66
20	292	2	34	66
21	300	2	100	0
22	345	2	100	0
23	346	2	42	58
24	352	2	42	58
25	360	2	100	0
26	405	2	100	0
27	406	2	50	50
28	412	2	50	50
29	420	2	100	0
30	465	2	100	0
31	466	2	54	46
32	472	2	54	46
33	480	2	100	0
34	525	2	100	0
35	526	2	58	42
36	532	2	58	42
37	540	2	100	0
38	585	2	100	0
39	586	2	62	38
40	592	2	62	38
41	600	2	100	0
42	645	2	100	0
43	646	2	66	34
44	652	2	66	34
45	660	2	100	0
46	705	2	100	0
47	706	2	70	30
48	713	2	70	30

Schritt	Zeit	Fluss	%A	%B
49	720	2	100	0
50	765	2	100	0
51	766	2	74	26
52	772	2	74	26
53	780	2	100	0
54	825	2	100	0
55	826	2	76	24
56	832	2	76	24
57	840	2	100	0
58	885	2	100	0
59	886	2	78	22
60	892	2	78	22
61	900	2	100	0
62	945	2	100	0
63	946	2	80	20
64	952	2	80	20
65	960	2	100	0
66	1005	2	100	0
67	1006	2	82	18
68	1012	2	82	18
69	1020	2	100	0
70	1065	2	100	0
71	1066	2	84	16
72	1072	2	84	16
73	1080	2	100	0
74	1125	2	100	0
75	1126	2	86	14
76	1132	2	86	14
77	1140	2	100	0
78	1185	2	100	0
79	1186	2	88	12
80	1192	2	88	12
81	1200	2	100	0
82	1245	2	100	0
83	1246	2	90	10
84	1252	2	90	10
85	1260	2	100	0
86	1305	2	100	0
87	1306	2	95	5
88	1312	2	95	5
89	1319	2	100	0
90	1364	2	100	0

Tabelle 3

Schritt	Zeit	Fluss	%A	%B
1	0,0	1,0	0,0	100,0
2	10,0	1,0	0,0	100,0
3	40,0	1,0	100,0	0,0
4	70,0	1,0	100,0	0,0
5	70,5	1,0	10,0	90,0
6	80,0	1,0	10,0	90,0
7	110,0	1,0	100,0	0,0
8	140,0	1,0	100,0	0,0
9	140,5	1,0	20,0	80,0
10	150,0	1,0	20,0	80,0
11	180,0	1,0	100,0	0,0
12	210,0	1,0	100,0	0,0
13	210,5	1,0	30,0	70,0
14	220,0	1,0	30,0	70,0
15	250,0	1,0	100,0	0,0
16	280,0	1,0	100,0	0,0
17	280,5	1,0	40,0	60,0
18	290,0	1,0	40,0	60,0
19	320,0	1,0	100,0	0,0
20	350,0	1,0	100,0	0,0
21	350,5	1,0	50,0	50,0
22	360,0	1,0	50,0	50,0
23	390,0	1,0	100,0	0,0
24	420,0	1,0	100,0	0,0

Schritt	Zeit	Fluss	%A	%B
25	420,5	1,0	60,0	40,0
26	430,0	1,0	60,0	40,0
27	460,0	1,0	100,0	0,0
28	490,0	1,0	100,0	0,0
29	490,5	1,0	70,0	30,0
30	500,0	1,0	70,0	30,0
31	530,0	1,0	100,0	0,0
32	560,0	1,0	100,0	0,0
33	560,5	1,0	80,0	20,0
34	570,0	1,0	80,0	20,0
35	600,0	1,0	100,0	0,0
36	630,0	1,0	100,0	0,0
37	630,5	1,0	85,0	15,0
38	640,0	1,0	85,0	15,0
39	670,0	1,0	100,0	0,0
40	700,0	1,0	100,0	0,0
41	700,5	1,0	88,0	12,0
42	710,0	1,0	88,0	12,0
43	740,0	1,0	100,0	0,0
44	770,0	1,0	100,0	0,0
45	770,5	1,0	90,0	10,0
46	780,0	1,0	90,0	10,0
47	810,0	1,0	100,0	0,0
48	850,0	1,0	100,0	0,0

Tabelle 4

Schritt	Zeit	Fluss	%A	%B
1	0	2	100	0
2	45	2	100	0
3	46	2	0	100
4	52	2	0	100
5	60	2	100	0
6	105	2	100	0
7	106	2	4	96
8	113	2	4	96
9	120	2	100	0
10	165	2	100	0
11	166	2	8	92
12	172	2	8	92
13	180	2	100	0
14	225	2	100	0
15	226	2	12	88
16	232	2	12	88
17	240	2	100	0
18	285	2	100	0
19	286	2	16	84
20	292	2	16	84
21	300	2	100	0
22	345	2	100	0
23	346	2	20	80
24	352	2	20	80
25	360	2	100	0
26	405	2	100	0
27	406	2	24	76
28	412	2	24	76
29	420	2	100	0
30	465	2	100	0
31	466	2	28	72
32	472	2	28	72
33	480	2	100	0
34	525	2	100	0
35	526	2	32	68
36	532	2	32	68
37	540	2	100	0
38	585	2	100	0
39	586	2	36	64
40	592	2	36	64
41	600	2	100	0
42	645	2	100	0
43	646	2	40	60
44	652	2	40	60
45	660	2	100	0
46	705	2	100	0
47	706	2	44	56
48	713	2	44	56

Schritt	Zeit	Fluss	%A	%B
49	720	2	100	0
50	765	2	100	0
51	766	2	48	52
52	772	2	48	52
53	780	2	100	0
54	825	2	100	0
55	826	2	52	48
56	832	2	52	48
57	840	2	100	0
58	885	2	100	0
59	886	2	56	44
60	892	2	56	44
61	900	2	100	0
62	945	2	100	0
63	946	2	60	40
64	952	2	60	40
65	960	2	100	0
66	1005	2	100	0
67	1006	2	64	36
68	1012	2	64	36
69	1020	2	100	0
70	1065	2	100	0
71	1066	2	68	32
72	1072	2	68	32
73	1080	2	100	0
74	1125	2	100	0
75	1126	2	70	30
76	1132	2	70	30
77	1140	2	100	0
78	1185	2	100	0
79	1186	2	72	28
80	1192	2	72	28
81	1200	2	100	0
82	1245	2	100	0
83	1246	2	75	25
84	1252	2	75	25
85	1260	2	100	0
86	1305	2	100	0
87	1306	2	80	20
88	1312	2	80	20
89	1319	2	100	0
90	1364	2	100	0
91	1365	2	85	15
92	1371	2	85	15
93	1378	2	100	0
94	1423	2	100	0

Tabelle 5

Schritt	Zeit	Fluss	%A	%B
1	0	2	100	0
2	45	2	100	0
3	46	2	0	100
4	52	2	0	100
5	60	2	100	0
6	105	2	100	0
7	106	2	13	87
8	113	2	13	87
9	120	2	100	0
10	165	2	100	0
11	166	2	25	75
12	172	2	25	75
13	180	2	100	0
14	225	2	100	0
15	226	2	29	71
16	232	2	29	71
17	240	2	100	0
18	285	2	100	0
19	286	2	34	66
20	292	2	34	66
21	300	2	100	0
22	345	2	100	0
23	346	2	38	62
24	352	2	38	62
25	360	2	100	0
26	405	2	100	0
27	406	2	40	60
28	412	2	40	60
29	420	2	100	0
30	465	2	100	0
31	466	2	42	58
32	472	2	42	58
33	480	2	100	0
34	525	2	100	0
35	526	2	44	56
36	532	2	44	56
37	540	2	100	0
38	585	2	100	0
39	586	2	46	54
40	592	2	46	54
41	600	2	100	0
42	645	2	100	0
43	646	2	48	52
44	652	2	48	52
45	660	2	100	0
46	705	2	100	0
47	706	2	50	50
48	713	2	50	50

Schritt	Zeit	Fluss	%A	%B
49	720	2	100	0
50	765	2	100	0
51	766	2	52	48
52	772	2	52	48
53	780	2	100	0
54	825	2	100	0
55	826	2	54	46
56	832	2	54	46
57	840	2	100	0
58	885	2	100	0
59	886	2	56	44
60	892	2	56	44
61	900	2	100	0
62	945	2	100	0
63	946	2	58	42
64	952	2	58	42
65	960	2	100	0
66	1005	2	100	0
67	1006	2	60	40
68	1012	2	60	40
69	1020	2	100	0
70	1065	2	100	0
71	1066	2	62	38
72	1072	2	62	38
73	1080	2	100	0
74	1125	2	100	0
75	1126	2	66	34
76	1132	2	66	34
77	1140	2	100	0
78	1185	2	100	0
79	1186	2	70	30
80	1192	2	70	30
81	1200	2	100	0
82	1245	2	100	0
83	1246	2	74	26
84	1252	2	74	26
85	1260	2	100	0
86	1305	2	100	0
87	1306	2	78	22
88	1312	2	78	22
89	1319	2	100	0
90	1364	2	100	0
91	1365	2	82	18
92	1371	2	82	18
93	1378	2	100	0
94	1423	2	100	0

Beispiel 1

Produktion und Faltung von Human- und Maus-β_2-Microglobulin

[0300] Dieses Beispiel beschreibt die Produktion von Human-β_2-Microglobulin sowie Maus-β_2-Microglobulin als FX_a-spaltbare Fusionsproteine in E. coli und die Reinigung des rekombinanten Human- und Maus-β_2-Microglobulins nach FX_a-Spaltung.

[0301] Plasmidklone, die die für Human- und Maus-β_2-Microglobulinprotein kodierenden Volllängen-cDNAs enthalten (großzügigerweise von Dr. David N. Garboczi Dr. Søren Buus bereitgestellt), wurden als Template in einer Polymerasekettenreaktion (PCR) (Saiki et al. (1988)) verwendet, die zur Produktion von cDNA-Fragmenten entworfen war, die den reifen Human- (entsprechend den Aminosäureresten Ile_1 bis Met_{99}) und reifen Maus-(entsprechend den Aminosäureresten Ile_1 bis Met_{99}) β_2-Microglobulinproteinen entsprechen, und zwar durch Verwendung der Primer Seq.-ID Nr. 3 und Seq.-ID Nr. 4 (für Human-β_2-Microglobulin) und Seq.-ID Nr. 5 und Seq.-ID Nr. 6 (für Maus-β_2-Microglobulin). Die amplifizierten, kodierenden Leseraster wurden an ihren 5'-Enden über die PCR-Reaktion an in Seq.-ID Nr. 3 und 5 umfassten Nucleotidsequenzen gebunden, die für die Aminosäuresequenz Seq.-ID Nr. 37 kodieren, die eine Spaltstelle für Rinderrestriktionsprotease FX_a darstellen (Nagai und Thrøgersen (1987)). Die amplifizierten DNA-Fragmente wurden in den E. coli-Expressions-

vektor pT_7H_6 subkloniert (Christensen et al. (1991)). Die Konstruktion der resultierenden Plasmide pT_7H_6FX-hβ_2m (Human-β_2-Microglobulin exprimierend) und pT_7H_6FX-mβ_2m (Maus-β_2-Microglobulin exprimierend) ist in **Fig.** 2 umrissen und in **Fig.** 3 sind die Aminosäuresequenzen der exprimierten Proteine gezeigt (in Seq.-ID Nr. 49 (Human) und Seq.-ID Nr. 50 (Maus) sind die von den Volllängen-Leserastern kodierten Aminosäuresequenzen gezeigt).

[0302] Human- und Maus-β_2-Microglobulin wurden durch Züchten und Exprimieren der Plasmide pT_7H_6FX-hβ_2m und -mβ_2m in E. coli BL21-Zellen im mittleren Maßstab (2 × 1 Liter) wie von Studier und Moffat, J. Mol. Biol. 189, 113–130 (1986), beschrieben produziert. Exponentiell wachsende Kulturen bei 37°C wurden bei einer OD_{600} von 0,8 mit Bakteriophagen λCE6 in einer Multiplizität von ungefähr 5 infiziert. Die Kulturen wurden bei 37°C für weitere drei Stunden gezüchtet, bevor die Zellen durch Zentrifugation geerntet wurden. Die Zellen wurden durch osmotischen Schock und Ultraschallbehandlung lysiert und das Gesamtzellprotein in Phenol extrahiert (mit Trisma-Base auf pH 8 gestellt). Protein wurde aus dem Phenol durch Zugabe von 2,5 Volumina Ethanol und Zentrifugation präzipitiert. Das Proteinpellet wurde in einem Puffer gelöst, der 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8, und 0,1 M Dithioerythrol enthielt. Nach Gelfiltration an Sephadex G-25 (Pharmacia, LKB, Schweden) in 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, 10 mM 2-Mercaptoethanol und 3 mM Methionin wurde das rohe Proteinpräparat auf Ni^{2+}-aktivierte NTA-Agarosesäulen zur Reinigung (Hochuli et al. (1988)) der Fusionsproteine, MGSHHHHHHGSIEGR-Human bzw. Maus-β_2-Microglobulin (worin MGSHHHHHHGSIEGR Seq.-ID Nr. 48 ist) aufgegeben und anschließend dem zyklischen Faltungsverfahren unterworfen.

[0303] Alle für die Flüssigchromatographie hergestellten Puffer wurden vor Zugabe von Reduktionsmittel und/oder Verwendung im Vakuum entgast.

[0304] Ni^{2+}-aktivierte NTA-Agarosematrix (Ni^{2+}NTA-Agarose) ist im Handel von der Diagen GmbH, Deutschland, erhältlich. Im Verlauf dieser Arbeit ist jedoch gefunden worden, dass dieses handelsübliche Produkt nicht die erwartete Leistung erbrachte. Die Beobachtungen der Erfinder waren, dass die handelsübliche Ni^{2+}NTA-Agarosematix leicht blockiert wurde, wenn der denaturierte und reduzierte Gesamtproteinextrakt aufgegeben wurde, dass die Kapazität für Fusionsprotein niedriger als erwartet war und dass die Matrix nur wenige Male erfolgreich regeneriert werden konnte.

[0305] Um die Leistungsfähigkeit der Ni^{2+}NTA-Agarose zu verbessern, wurde beschlossen, eine Carbodiimid-Kopplung des N-(5-Amino-1-carboxypentyl)iminodiessigsäure-Metallliganden (Syntheseweg wie von Döbeli und Hochuli beschrieben (EPO 0253 303)) an eine stabilere Agarosematrix (d. h. Sepharose CL-6B, Pharmacia, Schweden) durchzuführen: 8 g N-(5-Amino-1-carboxypentyl)iminodiessigsäure aus dem Syntheseverfahren in 50 ml wurden durch Zugabe von 29 g Na_2CO_3 (10 H_2O) auf pH 10 gestellt und einer gerührten Suspension aktivierter Sepharose CL-6B in 1 M Na_2CO_3 zugegeben. Die Reaktion erfolgte über Nacht.

[0306] Die Sepharose CL-6B (anfänglich 100 ml Suspension) wurde nach Entfernung des Wassers durch Aceton mit 7 g 1,1'-Carbonyldiimidazol unter Rühren für 15 bis 30 Minuten aktiviert. Nach Aktivierung wurde die Sepharose CL6B mit Aceton, gefolgt von Wasser und 1 M Na_2CO_3 gewaschen. Die NTA-Agarosematrix wurde in eine Säule gefüllt und mit Ni^{2+} „geladen", indem 5 Volumina einer 10%-igen $NiSO_4$-Lösung langsam durch die Säule gepumpt wurden. Die Menge von Ni^{2+} an der NTA-Agarosematrix, die mit diesem Verfahren hergestellt wurde, ist mit 14 µmol pro ml Matrix ermittelt worden. Die Ni^{2+}NTA-Agarosematrix wurde in eine standardmäßige Säule für Flüssigchromatographie (Innendurchmesser 2,6 cm) auf ein Volumen von 40 ml gepackt. Nach dem Beladen wurde die Ni^{2+}NTA-Agarosesäule mit zwei Säulenvolumina Wasser, einem Säulenvolumen 1 M Tris-HCl, pH 8, und zwei Säulenvolumina Beladungspuffer gewaschen, bevor der Proteinrohextrakt aufgegeben wurde.

[0307] Beim Aufgeben der Proteinrohextrakte auf die Ni^{2+}NTA-Agarosesäule wurden die Fusionsproteine MGSHHHHHHGSIEGR-hβ_2m bzw. MGSNHHHHHGSIEGR-mβ_2m (worin MGSHHHHHHGSIEGR Seq.-ID Nr. 48 ist) vom Hauptteil der E. coli- und λ-Phagen-Proteine durch Waschen mit einem Säulenvolumen des Beladungspuffers, gefolgt von 6 M Guanidiniumchlorid, 50 mM Tris-HCl, 10 mM 2-Mercaptoethanol und 3 mM Methionin gereinigt, bis die optische Dichte (OD) bei 280 nm der Säuleneluate stabil waren.

[0308] Die Fusionsproteine wurden an der Ni^{2+}NTA-Agarosesäule unter Anwendung des Gradientenmanagerprofils wie in Tabelle 1 beschrieben und 0,5 M NaCl, 50 mM Tris-HCl, pH 8, und 1,2 mM/0,4 mM reduziertes/oxidiertes Glutathion als Puffer A und 8 M Harnstoff, 0,5 M NaCl, 50 mM Tris-HCl, pH 8, 3 mM Methionin und 6 mM reduziertes Glutathion als Puffer B neu gefaltet. Die reduzierte/oxidierte Glutathionlösung wurde als eine 20fache Stammlösung durch Zugabe von 9,9 M H_2O_2 zu einer gerührten Lösung von 0,2 M reduziertem Glutathion vor Zugabe zu Puffer A frisch hergestellt.

[0309] Nach Beendigung des zyklischen Faltungsverfahrens wurden die hβ_2m- und mβ_2m-Fusionsproteine mit einem 0,5 M NaCl, 50 mM Tris-HCl, 20 mM EDTA, pH 8, enthaltenden Puffer aus den Ni^{2+}NTA-Agarosesäulen eluiert.

[0310] Fusionsprotein, das an den Ni^{2+}NTA-Agarosesäulen aggregiert und präzipitiert war, wurde in Puffer B eluiert. Ungefähr 75% des Fusionsproteinmaterials wurde vom nicht denaturierenden Elutionspuffer eluiert (siehe **Fig.** 16, Spuren 2 und 3).

[0311] Wie durch nicht reduzierende SDS-PAGE-Analyse beurteilt wurde, schienen 70% des löslichen hβ_2m-Fusionsproteinmaterials (entsprechend 40 mg hβ_2m-Fusionsprotein) monomer zu sein (siehe **Fig.** 15, Spuren 5 und 3), wogegen 25% des mβ_2m-Fusionsproteins monomer zu sein schienen (entsprechend 20 mg mβ_2m-Fusionsprotein). Die Gesamteffizienzen des Faltungsverfahrens sind daher ungefähr 50% für das hβ_2m-Fusionsprotein und weniger als 20% für das mβ_2m-Fusionsprotein.

[0312] Monomere hβ_2m- und mβ_2m-Fusionsproteine wurden von Dimer und Multimeren höherer Ordnung durch Ionenaustauschchromatographie an S-Sepharose (Pharmacia, Schweden) gereinigt: Die durch den nicht reduzierenden Puffer eluierten Fusionsproteine (ungefähr 70% des Fusionsproteinmaterials) wurden in einen Puffer, der 5 mM NaCl und 5 mM Tris-HCl, pH 8, enthielt, an Sephadex G-25 gelfiltriert und 1 : 1 mit Wasser verdünnt, bevor sie auf S-Sepharose-Ionentauschersäulen geladen wurden. Die Fusionsproteine wurden über 5 Säulenvolumina mit einem linearen Gradienten von 2,5 mM NaCl, 2,5 mM Tris-HCl, pH 8, bis 100 mM NaCl, 25 mM Tris-HCl, pH 8, eluiert. Die monomeren hβ_2m- sowie mβ_2m-Fusionsproteine eluierten ganz am Anfang des Gradienten, wogegen Dimere und Multimere höherer Ordnung später eluierten. Die monomeren Fusionsproteine enthaltende Fraktionen wurden mit Wasser verdünnt und neuerlich auf die S-Sepharose-Säulen aufgegeben und in einem Schritt in 1 M NaCl, 50 mM Tris-HCl, pH 8 eluiert.

[0313] Die monomeren Fusionsproteine wurden mit der Restriktionsprotease FX_a über Nacht bei Raumtemperatur in einem Gewichtsverhältnis von ungefähr 200 zu Eins gespalten.

[0314] Nach Spaltung wurden die rekombinanten hβ_2m- und mβ_2m-Proteine durch Ionenaustauschchromatographie an Q-Sepharose-Säulen (Pharmacia, Schweden) vom N-terminalen Fusionsschwanz gereinigt, vom abgespaltenen Fusionsprotein und FX_a befreit: Bei Gelfiltration an Sephadex G-25 in 5 mM NaCl, 5 mM Tris-HCl, pH 8, und 1 : 1-Verdünnung mit Wasser wurden rekombinantes hβ_2m und mβ_2m in einem linearen Gradienten (über 5 Säulenvolumina) von 2,5 mM NaCl, 2,5 mM Tris-HCl, pH 8, bis 100 mM NaCl, 25 mM Tris-HCl, pH 8, eluiert. Die gespaltenen rekombinanten Proteine enthaltenden Fraktionen wurde mit Wasser verdünnt und wiederum auf die Q-Sepharose-Säulen aufgegeben und in einem Schritt in 1 M NaCl, 50 mM Tris-HCl, pH 8, eluiert. Rekombinante hβ_2m- und mβ_2m-Proteine wurden in frisch hergestelltes NH_4HCO_3 gelfiltriert und zweimal lyophilisiert.

[0315] SDS-PAGE-Analyse der Produktion von rekombinantem Human-β_2-Microglobulin ist in **Fig.** 15 dargestellt.

[0316] Die Ausbeute des völlig prozessierten rekombinanten Human-β_2-Microglobulins, das mit diesem Verfahren produziert wurde, betrug 30 mg.

[0317] Die Ausbeute des völlig prozessierten rekombinanten Maus-β_2-Microglobulins, das mit diesem Verfahren produziert wurde, betrug 10 mg.

[0318] Der Vergleich von rekombinantem Human- mit gereinigtem natürlichen Human-β_2-Microglobulin wurde freundlicherweise von Dr. Søren Buus in zwei verschiedenen Tests durchgeführt:

1. Es wurde gefunden, das rekombinantes Human-β_2-Microglobulin und natürliches Human-β_2-Microglobulin sowohl mit einem monoklonalen, als auch mit einem monospezifischen Antikörper mit gleicher Affinität reagierten.
2. Es erwies sich, dass rekombinantes Human-β_2-Microglobulin und natürliches Human-β_2-Microglobulin in einem Bindungshemmexperiment mit radiomarkierten Liganden natürliche affinitätsgereinigte Schwerketten-Klasse I-K^d-Moleküle mit identischer Affinität binden.

[0319] Es erwies sich, dass Maus-β_2-Microglobulin natürliche Klasse I-Schwerkettenmoleküle mit einer Affinität binden, die 5 Mal niedriger ist, als die des Human-β_2-Microglobulins. Dieses Ergebnis ist in gutem Einklang mit früheren Ergebnissen aus der Literatur unter Verwendung natürlichen Materials.

Beispiel 2

Produktion und Faltung von Human-Wachstumshormon (Somatotropin)

[0320] Dieses Beispiel beschreibt die Produktion von Human-Wachstumshormon (hGH) als FX_a-spaltbares Fusionsprotein in E. coli und die Reinigung des rekombinanten hGH nach FX_a-Spaltung.

[0321] Ein Plasmidklon, der die für hGH kodierende cDNA enthielt (großzügigerweise von Dr. Henrik Dalbrøge bereitgestellt (Dalbøge et al. (1987)), wurde als Templat in einer Polymerasekettenreaktion (PCR) (Saiki et al. (1988)), unter Verwendung der Primer Seq.-ID Nr. 7 und Seq.-ID Nr. 8 verwendet, die zur Produktion eines dem reifen hGH entsprechenden cDNA-Fragments (entsprechend den Aminosäureresten Phe_1 bis Phe_{191}) entworfen waren. Das amplifizierte, kodierende Leseraster war am 5'-Ende über die PCR-Reaktion an eine in Seq.-ID Nr. 7 umfasste Nucleotidsequenz gebunden, die für die Aminosäuresequenz Seq.-ID Nr. 37 kodiert, die eine Spaltstelle für Rinderrestriktionsprotease FX_a darstellen (Nagai und Thøgersen (1987)). Das amplifizierte DNA-Fragment wurde in den E. coli-Expressionsvektor pT_7H_6 subkloniert (Christensen et al. (1991)). Die Konstruktion des resultierenden Plasmids pT_7H_6FX-hGH (Human-Wachstumshormon exprimierend) ist in

Fig. 4 umrissen und in **Fig.** 5 ist die Aminosäuresequenz des exprimierten Proteins gezeigt (in Seq.-ID Nr. 51 ist die vom Volllängen-Leseraster kodierte Aminosäuresequenz gezeigt).

[0322] Rekombinantes Human-Wachstumshormon wurde durch Züchten und Exprimieren des Plasmids pT_7H_6FX-hGH in E. coli BL21-Zellen im mittleren Maßstab (2 × 1 Liter) wie von Studier und Moffat, J. Mol. Biol. 189, 113–130 (1986), beschrieben produziert. Exponentiell wachsende Kulturen bei 37°C wurden bei einer OD_{600} von 0,8 mit Bakteriophagen λCE6 in einer Multiplizität von ungefähr 5 infiziert. Die Kulturen wurden bei 37°C für weitere drei Stunden gezüchtet, bevor die Zellen durch Zentrifugation geerntet wurden. Die Zellen wurden durch osmotischen Schock und Ultraschallbehandlung lysiert und das Gesamtzellprotein in Phenol extrahiert (mit Trisma-Base auf pH 8 gestellt). Protein wurde aus dem Phenol durch Zugabe von 2,5 Volumina Ethanol und Zentrifugation präzipitiert. Das Proteinpellet wurde in einem Puffer gelöst, der 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8, und 50 mM Dithioerythrol enthielt. Nach Gelfiltration an Sephadex G-25 (Pharmacia, LKB, Schweden) in 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, 5 mM 2-Mercaptoethanol und 1 mM Methionin wurde das rohe Proteinpräparat auf eine Ni^{2+}-aktivierte NTA-Agarosesäule zur Reinigung (Hochuli et al. (1988)) des Fusionsproteins, MGSHHHHHHGSIEGR-hGH (worin MGSHHHHHHGSI-EGR Seq.-ID Nr. 48 ist) aufgegeben und anschließend dem zyklischen Faltungsverfahren unterworfen.

[0323] Die Herstellung und „Ladung" der Ni^{2+}NTA-Agarosesäule ist in Beispiel 1 beschrieben.

[0324] Alle für die Flüssigchromatographie hergestellten Puffer wurden vor Zugabe von Reduktionsmittel und/oder Verwendung im Vakuum entgast.

[0325] Beim Aufgeben des Proteinrohextrakts auf die Ni^{2+}NTA-Agarosesäule wurde das Fusionsprotein MGSHHHHHHGSIEGR-hGH (worin MGSHHHHHHGSIEGR Seq.-ID Nr. 48 ist) vom Hauptteil der E. coli- und λ-Phagen-Proteine durch Waschen mit einem Säulenvolumen des Beladungspuffers, gefolgt von 6 M Guanidiniumchlorid, 50 mM Tris-HCl, 5 mM 2-Mercaptoethanol und 1 mM Methionin gereinigt, bis die optische Dichte (OD) bei 280 nm der Säuleneluate stabil waren.

[0326] Das Fusionsprotein wurde an der Ni^{2+}NTA-Agarosesäule unter Anwendung des Gradientenmanagerprofils wie in Tabelle 2 beschrieben und 0,5 M NaCl, 50 mM Tris-HCl, pH 8, und 1,0 mM/0,1 mM reduziertes/oxidiertes Glutathion als Puffer A und 8 M Harnstoff, 0,5 M NaCl, 50 mM Tris-HCl, pH 8, 1 mM Methionin und 5 mM reduziertes Glutathion als Puffer B neu gefaltet. Die reduzierte/oxidierte Glutathionlösung wurde als 200fache Stammlösung durch Zugabe von 9,9 M H_2O_2 zu einer gerührten Lösung von 0,2 M reduziertem Glutathion vor Zugabe zu Puffer A frisch hergestellt.

[0327] Nach Beendigung des zyklischen Faltungsverfahrens wurde das hGH-Fusionsprotein mit einem 0,5 M NaCl, 50 mM Tris-HCl, 20 mM EDTA, pH 8, enthaltenden Puffer aus den Ni^{2+}NTA-Agarosesäulen eluiert. Fusionsprotein, das an der Ni^{2+}NTA-Agarosesäule aggregiert und präzipitiert war, wurde in Puffer B eluiert.

[0328] Ungefähr 80% des Fusionsproteinmaterials wurde vom nicht denaturierenden Elutionspuffer eluiert (siehe **Fig.** 16, Spuren 2 und 3). Wie durch nicht reduzierende SDS-PAGE-Analyse beurteilt wurde, schienen 90% des löslichen Fusionsproteinmaterials (entsprechend 70 mg Fusionsprotein) monomer zu sein (siehe **Fig.** 16, Spur 2), was Gesamteffizienzen des Faltungsverfahrens von ungefähr 80% ergab.

[0329] Monomeres hGH-Fusionsproteins wurde von Dimer und Multimeren höherer Ordnung durch Ionenaustauschchromatographie an Q-Sepharose (Pharmacia, Schweden) gereinigt: Nach Gelfiltration in einen 25 mM NaCl und 25 mM Tris-HCl, pH 8, enthaltenden Puffer an Sephadex G-25 wurde das vom nicht denaturierenden Puffer eluierte Fusionsproteinmaterial auf eine Q-Sepharose-Ionentauschersäule geladen. Das Fusionsprotein wurde über 5 Säulenvolumina mit einem linearen Gradienten von 25 mM NaCl, 25 mM Tris-HCl, pH 8, bis 200 mM NaCl, 50 mM Tris-HCl, pH 8, eluiert. Das monomere hGH-Fusionsprotein eluierte am Anfang des Gradienten, wogegen Dimere und Multimere höherer Ordnung später eluierten. Fraktionen, die das reine monomere Fusionsprotein enthielten, wurde $NiSO_4$ und Iminodiessigsäure (IDA, eingestellt auf pH 8 mit NaOH) auf 1 mM zugegeben und mit der Restriktionsprotease FX_a für 5 Stunden bei 37°C in einem Gewichtsverhältnis von ungefähr 100 zu Eins gespalten. FX_a wird nach der Spaltung durch Zugabe von Benzamidin-Hydrochlorid auf 1 mM inhibiert.

[0330] Nach der Spaltung wurde das rekombinante hGH-Protein nach Gelfiltration an Sephadex G-25 in 8 M Harnstoff, 50 mM Tris-HCl, pH 8, zur Entfernung von Ni^{2+}IDA und Benzamidin vom ungespaltenen Fusionsprotein und dem freigesetzten Fusionsschwanz isoliert, indem es über eine kleine Ni^{2+}NTA-Agarosesäule, gefolgt von einer in-line nachgeschalteten kleinen Nd^{3+}NTA-Agarosesäule und anschließend einer nicht Ni^{2+}-aktivierten NTA-Agarosesäule geschickt wurde, um die vollständige Entfernung von FX_a und Ni^{2+} bzw. Nd^{2+} zu gewährleisten. Das rekombinante hGH wurde aus einer kleinen Fraktion des rekombinanten Abbauprodukts durch Ionenaustauschchromatographie an Q-Sepharose gereinigt: hGH wurde in einem linearen Gradienten (über 5 Säulenvolumina) von 8 M Harnstoff, 50 mM Tris-HCl, pH 8, bis 8 M Harnstoff, 250 mM NaCl, 25 mM Tris-HCl, pH 8, eluiert. Die das gespaltene, gereinigte rekombinante Protein enthaltenden Fraktionen wurden in frisch hergestelltem 20 mM NH_4HCO_3 gelfiltriert und zweimal lyophilisiert.

[0331] Die SDS-PAGE-Analyse der Produktion und Faltung von rekombinantem Human-Wachstumshormon ist in **Fig.** 16 dargestellt.

[0332] Die Ausbeute des durch dieses Verfahren vollständig prozessierten rekombinanten Human-Wachs-

tumshormons betrug 10 mg.

[0333] Das mit diesem Verfahren produzierte rekombinante Human-Wachstumshormon migrierte sowohl bei reduzierender, als auch nicht reduzierender SDS-PAGE gemeinsam mit biologisch aktivem rekombinanten Wachstumshormon, das großzügigerweise von Novo-Nordisk A/S bereitgestellt worden ist.

Beispiel 3

Produktion und Faltung von Human-α_2MRAP

[0334] Das zur Expression des Human-α_2-Macroglobulinrezeptor-assoziierten Proteins (α_2MRAP) in E. coli BL21-Zellen verwendete Plasmid pT7H6FX-α_2MRAP und die zur Produktion des Fusionsproteins verwendeten Bedingungen sind von den Erfindern früher in Nykjaer et al., J. Biol. Chem. 267, 14543–14546 (1992), beschrieben worden. Die Primer Seq.-ID Nr. 9 und Seq.-ID Nr. 10 wurden in der zur Vervielfachung der für α_2MRAP kodierenden DNA eingesetzten PCR verwendet.

[0335] Aus der Phenolphase der Proteinextraktion von Zellen aus 2 Litern einer Kultur MGS-HHHHHH-SIEGR-α_2MRAP- (worin MGSHHHHHHGSIEGR Seq.-ID Nr. 48 ist) exprimierender E. coli BL21-Zellen präzipitierter Proteinrohextrakt wurde in einem 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8 und 50 mM Dithioerythrol enthaltendem Puffer gelöst. Nach Gelfiltration an Sephadex G-25 (Pharmacia, Schweden) in 8 M Harnstoff, 0,5 M NaCl, 50 mM Tris-HCl, pH 8, und 1 mM Methionin wurde des Proteinrohpräparat auf eine Ni^{2+}-aktivierte NTA-Agarosematrix (Ni^{2+}NTA-Agarose) zur Reinigung (Hochuli et al. (1988)) des Fusionsproteins MGSHHHHHHSIEGR-α_2MRAP (worin MGSHHH-HHHGSIEGR Seq.-ID Nr. 48 ist) aufgegeben und anschließend dem zyklischen Faltungsprozess unterzogen.

[0336] Alle zur Flüssigchromatographie hergestellten Puffer wurden vor der Zugabe von Reduktionsmittel und/oder Verwendung im Vakuum entgast.

[0337] Die Herstellung und „Ladung" der Ni^{2+}NTA-Agarosesäule ist unter Beispiel 1 beschrieben.

[0338] Beim Aufgeben des Proteinrohextrakts auf die Ni^{2+}NTA-Agarosesäule wurde das Fusionsprotein MGSHHHHHHGSIEGR-α_2MRAP (worin MGSHHHHHHGSIEGR Seq.-ID Nr. 48 ist) vom Hauptteil der E. coli- und λ-Phagen-Proteine durch Waschen mit einem Säulenvolumen des Beladungspuffers, gefolgt von 6 M Guanidiniumchlorid, 50 mM Tris-HCl und 1 mM Methionin gereinigt, bis die optische Dichte (OD) des Säuleneluats bei 280 nm stabil war.

[0339] Das Fusionsprotein wurde an der Ni^{2+}NTA-Agarosesäule unter Anwendung des Gradientenmanagerprofils wie in Tabelle 3 beschrieben und 0,5 M NaCl, 50 mM Tris-HCl, 2 mM $CaCl_2$ und 1 mM 2-Mercaptoethanol als Puffer A und 6 M Harnstoff, 50 mM Tris-HCl, pH 8, 2 mM $CaCl_2$ und 1 mM 2-mercaptoethanol als Puffer B neu gefaltet.

[0340] Nach Beendigung des zyklischen Faltungsverfahrens wurde das α_2MRAP-Fusionsprotein mit einem 0,5 M NaCl, 50 mM Tris-HCl, 20 mM EDTA, pH 8, enthaltenden Puffer aus der Ni^{2+}NTA-Agarosesäule eluiert.

[0341] Es fand sich praktisch kein an der Ni^{2+}NTA-Agarosesäule aggregiertes oder präzipitiertes Fusionsprotein. Die geschätzte Ausbeute von α_2MRAP-Fusionsprotein betrug 60 mg, und die Effizienz des Faltungsverfahrens war nahezu 95%.

[0342] Das Fusionsprotein MGSHHHHHHGSIEGR-α_2MRAP (worin MGSHHHHHHGSIEGR Seq.-ID Nr. 48 ist) wurde mit der Rinderrestriktionsprotease FX_a über Nacht bei Raumtemperatur in einem Gewichtsverhältnis von 200 : 1 im Elutionspuffer gespalten. Nach Gelfiltration an Sephadex G-25 in 100 mM NaCl, 25 mM Tris-HCl, pH 8, wurde die Proteinlösung über eine Ni^{2+}NTA-Agarosesäule geschickt, wodurch ungespaltenes Fusionsprotein und der von gespaltenen Fusionsproteinen stammende, freigesetzte N-terminale Fusionsschwanz entfernt wurde. Schließlich wurde die Proteinlösung 1 : 4 mit Wasser verdünnt und das α_2MRAP-Protein vom FX_a durch Ionenaustauschchromatographie an Q-Sepharose (Pharmacia, Schweden) gereinigt. Die Q-Sepharosesäule wurde mit einem linearen Gradienten über 6 Säulenvolumina von 25 mM NaCl, 25 mM Tris-HCl, pH 8, bis 250 mM NaCl, 25 mM Tris-HCl, pH 8, eluiert. Das α_2MRAP-Protein eluierte ganz am Anfang des Gradienten, wogegen FX_a später eluierte.

[0343] Die Ausbeute des durch dieses Verfahren produzierten und neugefalteten α_2MRAP-Proteins betrug 40 mg.

[0344] Die Ligandenbindungseigenschaften (d. h. Bindung an den α_2-Macroglobulinrezeptor und Störung der Bindung von Human-Urokinase-Plasminogenaktivator-Plasminogenaktivator-Inhibitor-Typ-I-Komplex an den α_2-M-Rezeptor) hat sich nach Dr. Nykjaer als identisch mit den Ligandenbindungseigenschaften des gereinigten natürlichen Proteins erwiesen.

Beispiel 4

Produktion und Neufaltung von Domänen und Domänen-Clustern aus dem α_2-M-Rezeptor

[0345] Das Human-α_2-Macroglobulinrezeptor/Low-Density-Lipoproteinrezeptor-bezogene Protein (α_2MR) ist ein Endozytose-Membranrezeptor von 600 kDa. α_2MR wird als einkettiges Vorläuferprotein von 4524 Aminosäuren synthetisiert. Der Vorläufer wird in eine 85 kDa-Transmembran-β-Kette und eine 500 kDa-α-Kette prozessiert, die nicht kovalent an die extrazelluläre Domäne der β-Kette gebunden ist. Das α_2MR bindet bekanntermaßen Ca^{2+} in strukturabhängiger Weise (d. h., das reduzierte Proteine bindet Ca^{2+} nicht) und wird in dem Sinne als multifunktionell angenommen, als dass α_2MRAP Liganden unterschiedlicher Klassen bindet.
[0346] Die gesamte Aminsäuresequenz der α-Kette kann durch Cluster dreier Formen von Wiederholungen dargestellt werden, die sich auch in anderen membrangebundenen Rezeptoren in verschiedenen Plasmaproteinen finden:
A. Diese Wiederholungsform überspannt ungefähr 40 Aminosäurereste und ist durch das aufeinander folgende Auftreten der sechs in der Wiederholung enthaltenen Cysteinylreste charakterisiert. Einige Autoren haben diese Wiederholung Domäne der Komplement-Form benannt.
B: Diese Wiederholungsform überspannt ebenfalls ungefähr 40 Aminosäurereste und ist durch das aufeinander folgende Auftreten der sechs in der Wiederholung enthaltenen Cysteinylreste charakterisiert. In der Literatur ist diese Wiederholung Domäne der EGF-Form benannt worden.
C: Diese Wiederholungsform überspannt ungefähr 55 Aminosäurereste und ist durch die Anwesenheit der Konsensussequenz Seq.-ID Nr. 39 charakterisiert.

[0347] Dieses Beispiel beschreibt die Produktion einer Reihe von Domänen und Domänen-Clustern, die vom α_2-MRProtein als FX_a-spaltbare Fusionsproteine hergeleitet sind, in E. coli und die Reinigung, In-vitroFaltung und die FX_a-Spaltung und Prozessierung dieser rekombinanten Proteine.
[0348] Ein die für Human-α_2-MR-Protein kodierende Volllängen-cDNA enthaltender Plasmidklon (großzügigerweise von Dr. Joachim Herz bereitgestellt; Herz et al., EMBO J. 7, 4119–4127 (1988)) wurde als Templat in einer Serie von Polymerasekettenreaktionen (PCR) verwendet, die zur Produktion von cDNA-Fragmenten entworfen waren, die einer Reihe von Polypeptiden entsprechen, welche die vom α_2-MR-Protein hergeleiteten Domänen und Domänen-Cluster repräsentieren:
#1: Enthält zwei Domänen der A-Form, entsprechend den Aminosäureresten 20 bis 109 im α_2-MR-Protein. Die Primer Seq.-ID Nr. 11 und Seq.-ID Nr. 12 wurden in der PCR verwendet.
#2: enthält zwei Domänen der A-Form, gefolgt von zwei Domänen der B-Form, entsprechend den Aminosäureresten 20 bis 190 im α_2-MR-Protein. Die Primer Seq.-ID Nr. 11 und Seq.-ID Nr. 13 wurden in der PCR verwendet.
#3: Identisch mit #2, gefolgt von einer YWTD-Wiederholungen enthaltenden Region, entsprechend den Aminosäureresten 20 bis 521. Die Primer Seq.-ID Nr. 11 und Seq.-ID Nr. 14 wurden in der PCR verwendet.
#4: Enthält eine Domäne der B-Form, gefolgt von 8 Domänen der A-Form und schließlich zwei Domänen der B-Form, entsprechend den Aminosäureresten 803 bis 1265 im α_2-MR-Protein. Die Primer Seq.-ID Nr. 15 und Seq.-ID Nr. 16 wurden in der PCR verwendet.
#5: Enthält nur die 8 Domänen der A-Form, die auch in #4 vorhanden sind, entsprechend den Aminosäureresten 849-1184 im α_2-MR-Protein. Die Primer Seq.-ID Nr. 17 und Seq.-ID Nr. 18 wurden in der PCR verwendet.
#6: Enthält die beiden C-terminalen Domänen der B-Form aus #4, gefolgt von 8 YWTD-Wiederholungen und einer Domäne der B-Form, entsprechend den Aminosäureresten 1184-1582 im α_2-MR-Protein. Die Primer Seq.-ID Nr. 19 und Seq.-ID Nr. 20 wurden in der PCR verwendet.
#7: Enthält die gesamte in den Konstrukten #4 bis #6 enthaltene Region, entsprechend den Aminosäureresten 803 bis 1582 im α_2-MR-Protein. Die Primer Seq.-ID Nr. 15 und Seq.-ID Nr. 20 wurden in der PCR verwendet.
#8: Enthält 10 Domänen der A-Form, entsprechend den Aminosäureresten 2520 bis 2941 im α_2-MR-Protein. Die Primer Seq.-ID Nr. 21 und Seq.-ID Nr. 22 wurden in der PCR verwendet.
#9: Enthält 11 Domänen der A-Form, entsprechend den Aminosäureresten 3331-3778 im α_2-MR-Protein. Die Primer Seq.-ID Nr. 23 und Seq.-ID Nr. 24 wurden in der PCR verwendet.

[0349] Die für Domänen und Domänen-Cluster kodierenden, amplifizierten Nucleotidsequenzen wurden an ihren 5'-Enden über die PCR-Reaktion an Nucleotidsequenzen (in Seq.-ID Nr. 11, 15, 17, 19, 21 und 23 umfasst) gebunden, die für Aminosäuresequenz Seq.-ID Nr. 37, die eine Spaltstelle für die Rinderrestriktionsprotease FX_a darstellt (Nagai und Thrøgersen, Methods in Enzymology 152, 461–481 (1987)) kodieren. Die amplifizierten DNA-Fragmente wurden entweder in den E. coli-Expressionsvektor pT_7H_6 (Christensen et al., FEBS Letters 295, 181–184 (1991)) oder in das Expressionsplasmid $pLcIIMLCH_6$, welches eine modifizierte Form von pLcIIMLC ist (Nagai et al., Nature 332, 284–286 (1988)), durch Insertion eines Oligonucleotids subkloniert, das für sechs Histidinylreste C-terminal des Myosin-Leichtkettenfragments kodiert. Die Konstruktion der resul-

tierenden Plasmide pT_7H_6FX-#1 bis #3 und $pLcIIMLCH_6FX$-#4 bis #9 ist in **Fig.** 6–8 dargestellt und in **Fig.** 9 ist die Aminosäuresequenz des exprimierten Proteins gezeigt (in Seq.-ID Nr. 52 ist die vom Volllängen-Leseraster kodierte Aminosäuresequenz gezeigt).

[0350] Die in der pT_7H_6FX-Reihe subklonierten Domänen und Domänen-Cluster wurden in E. coli BL21-Zellen im mittleren Maßstab (2 Liter) wie von Studier und Moffat,). Mol. Biol. 189, 113–130 (1986), beschrieben gezüchtet und exprimiert. Exponentiell wachsende Kulturen bei 37°C wurden bei einer OD_{600} von 0,8 mit Bakteriophagen λCE6 in einer Multiplizität von ungefähr 5 infiziert. Die Kulturen wurden bei 37°C für weitere drei Stunden gezüchtet, bevor die Zellen durch Zentrifugation geerntet wurden. Die Zellen wurden durch osmotischen Schock lysiert und ultraschallbehandelt und das Gesamtzellprotein in Phenol (mit Trisma-Base auf pH 8 gestellt) extrahiert.

[0351] Die in der $pLcIIMLCH_6$-Reihe subklonierten Domänen-Cluster wurden in E. coli QY13-Zellen wie in Nagai und Thøgersen, Methods in Enzymology 152, 461–481 (1987), beschrieben gezüchtet und exprimiert. Exponentiell wachsende Kulturen bei 30°C wurden bei einer OD_{600} von 1,0 für 15 Minuten übertragen. Dieser Hitzschock induziert die Synthese der Fusionsproteine. Die Kulturen werden bei 37°C für weitere drei bis vier Stunden inkubiert, bevor die Zellen durch Zentrifugation geerntet werden. Die Zellen wurden durch osmotischen Schock und Ultraschallbehandlung lysiert und das Gesamtzellprotein in Phenol (mit Trisma-Base auf pH 8 gestellt) extrahiert.

[0352] Rohprotein wurde aus der Phenolphase durch Zugabe von 2,5 Volumina Ethanol und Zentrifugation präzipitiert. Das Proteinpellet wurde in einem Puffer gelöst, der 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8, und 0,1 M Dithioerythrol enthielt. Nach Gelfiltration an Sephadex G-25 (Pharmacia, Schweden) in 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, 10 mM 2-Mercaptoethanol und 2 mM Methionin wurden die Proteinrohpräparate auf Ni^{2+}-aktivierte NTA-Agarosesäulen zur Reinigung (Hochuli et al. (1988)) der Fusionsproteine aufgegeben und anschließend dem zyklischen Faltungsverfahren unterzogen.

[0353] Alle für die Flüssigchromatographie hergestellten Puffer wurden vor der Zugabe von Reduktionsmittel und/oder Verwendung unter Vakuum entgast.

[0354] Die Herstellung und „Ladung" der Ni^{2+}NTA-Agarosesäule ist unter Beispiel 1 beschrieben.

[0355] Nach Aufgeben der Proteinrohextrakte auf die Ni^{2+}NTA-Agarosesäule wurden die Fusionsproteine vom Großteil der E. coli- und λ-Phagenproteine durch Waschen mit einem Säulenvolumen des Beladungspuffers, gefolgt von 6 M Guanidiniumchlorid, 50 mM Tris-HCl, 10 mM 2-Mercaptoethanol und 2 mM Methionin gereinigt, bis die optische Dichte (OD) bei 280 nm stabil war.

[0356] Jedes der Fusionsproteine wurde an der Ni^{2+}-NTA-Agarosesäule unter Anwendung des Gradientenmanagerprofils wie in Tabelle 4 beschrieben und 0,5 M NaCl, 50 mM Tris-HCl, pH 8, 2 mM $CaCl_2$, 0,33 mM Methionin und 2,0 mM/0,2 mM reduziertes Glutathion als Puffer B neu gefaltet. Die reduzierte/oxidierte Glutathionlösung wurde als 100fache Stammlösung durch Zugabe von 9,9 M H_2O_2 zu einer gerührten Lösung von 0,2 M reduziertem Glutathion vor Zugabe zu Puffer A frisch hergestellt.

[0357] Nach Beendigung des zyklischen Neufaltungsverfahrens wurden die Fusionsproteine, die die vom α_2-MR-Protein hergeleiteten Domänen und Domänen-Cluster repräsentieren, von der Ni^{2+}NTA-Agarosesäule mit einem 0,5 M NaCl, 50 mM Tris-HCl, 5 mM EDTA, pH 8, enthaltendem Puffer eluiert. Fusionsproteine, die an der Ni^{2+}NTA-Agarosesäule aggregiert und präzipitiert waren, wurden in Puffer B eluiert.

[0358] Ungefähr 75% des aus den Plasmiden pT_7H_6FX-#1 und #2 exprimierten Fusionsproteinmaterials, das die N-terminalen zwei und vier cysteinreichen Domänen des α_2-MR-Proteins repräsentieren, wurden aus der Ni^{2+}NTA-Agarosesäule durch den nicht denaturierenden Puffer eluiert. Der Großteil dieses Fusionsproteinmaterials erschien monomer, wie durch nicht reduzierende SDS-PAGE-Analyse beurteilt wurde. Die Ausbeute des monomeren Fusionsproteins #1 und #2 wurde auf 50 mg geschätzt.

[0359] Ungefähr 50% des aus allen anderen Expressionsplasmiden, welche die vom α_2-MR-Protein hergeleiteten Domänen-Cluster repräsentieren, exprimierten Fusionsproteinmaterials wurde aus der Ni^{2+}NTA-Agarosesäule durch den nicht denaturierenden Puffer eluiert. Zwischen 30% (Fusionsproteine #5 und #7) und 65% (Fusionsprotein #4) dieser Fusionsproteine schienen monomer zu sein, wie mittels nicht reduzierender SDS-PAGE-Analyse beurteilt wurde (siehe **Fig.** 17, Spuren 9 und 10).

[0360] Jedes vom nicht denaturierenden Elutionspuffer eluierte Fusionsprotein wurde mit der Restriktionsprotease FX_a über Nacht bei Raumtemperatur in einem Gewichtsverhältnis von 100 zu Eins gespalten.

[0361] Nach Gelfiltration an Sephadex G-25 in 100 mM NaCl, 25 mM Tris-HCl, pH 8, wurde die Proteinlösung durch eine Ni^{2+}NTA-Agarosesäule geschickt, wodurch ungespaltenes Fusionsprotein entfernt und der aus den gespaltenen Fusionsproteinen stammende N-terminale Fusionsschwanz freigesetzt wurden. FX_a wurde aus der Lösung entfernt, indem die rekombinanten Proteinlösungen durch eine kleine Säule von SBTI-Agarose (Sojabohnen-Trypsininhibitor, immobilisiert an Sepharose CL-6B (Pharmacia, Schweden)) geschickt wurden.

[0362] Die SDS-PAGE-Analyse des neugefalteten, löslichen Fusionsproteinprodukts #4 ist in **Fig.** 17, Spuren 9 und 10 dargestellt und zeigt die reduzierten bzw. nicht reduzierten Proben. Der für die nicht reduzierte Probe beobachtete Mobilitätsanstieg spiegelt die Kompaktheit des Polypeptids aufgrund der Anwesenheit von 33 Disulfidbrücken wider.

[0363] Jedes der rekombinanten Proteine erwies sich in strukturabhängiger Weise als Ca^{2+}-bindend.
[0364] Von Dr. Søren Moestrup ist gefunden worden, dass ein monoklonaler, vom natürlichen Human-α_2MR hergeleiteter Antikörper, A2MRα-5, die von den Konstrukten #4, #6 und #7 exprimierten rekombinanten Proteine band, wogegen ein monospezifischer, ebenfalls vom natürlichen α_2-MR hergeleiteter Antikörper, A2MRα-3, das vom Konstrukt #8 exprimierte rekombinante Protein band. Die Bindungsspezifität beider Antikörper ist strukturabhängig (d. h., die Antikörper reagieren weder mit reduziertem α_2-MR, noch mit dem reduzierten rekombinanten Protein).

Beispiel 5

Produktion und Faltung von Rinderkoagulationsfaktor X_a (FX_a)

[0365] Dieses Beispiel beschreibt die Produktion eines der vom Rinder-FX_a als FX_a spaltbares Fusionsprotein hergeleiteten Fragmente und die Reinigung, In-vitro-Faltung und Prozessierung des rekombinanten Proteins.
[0366] Die für Rinder-FX kodierende cDNA wurde durch spezifische Amplifikation der für Rinder-FX von Aminosäurerest Ser_{82} bis Trp_{484} (Seq.-ID Nr. 2, Reste 82-484) (FXΔγ, Aminosäurenummerierung bezieht sich auf das Volllängen-Leseraster) kodierenden Nucleotidsequenzen in einer Polymerasekettenreaktion (PCR) kloniert, und zwar unter Verwendung von Erststrang-Oligo-dt-geprimter, aus Gesamtrinderleber-DNA als Templat synthetisierter cDNA. Die in der PCR verwendeten Primer waren Seq.-ID Nr. 35 und Seq.-ID Nr. 26. RNA-Extraktion und cDNA-Synthese wurden mittels Standardverfahren durchgeführt.
[0367] Das für FXΔγ kodierende, amplifizierte Leseraster war am 5'-Ende über die PCR-Reaktion an Nucleotidsequenzen gebunden, die für die Aminosäuresequenz Seq.-ID Nr. 37 kodieren, die eine Spaltstelle für die Rinderrestriktionsprotease FX_a darstellt (Nagai und Thøgersen, Methods in Enzymology 152, 461–481 (1987)). Die amplifizierten DNA-Fragmente wurden in den E. coli-Expressionsvektor $pLcIIMLCH_6$, der eine modifizierte Form von pLcIIMLC ist (Nagai et al., Nature 332, 284–286 (1988)), durch Insertion eines Oligonucleotids kloniert, das für sechs Histidinylreste C-terminal des Myosin-Leichtkettenfragments kodiert. Die Konstruktion des resultierenden Plasmids $pLcIIMLCH_6$FX-FXΔγ ist in **Fig.** 10 dargestellt und in **Fig.** 1 1 ist die Aminosäuresequenz des exprimierten Proteins gezeigt (in Seq.-ID Nr. 53 ist die vom Volllängen-Leseraster kodierte Aminosäuresequenz gezeigt).
[0368] Das $pLcIIMLCH_6$FX-FXΔγ-Plasmid wurde in E. coll QY13-Zellen wie in Nagai und Thøgersen beschrieben (Methods in Enzymology 152, 461–481 (1987)) gezüchtet und exprimiert. Exponentiell wachsende Kulturen bei 30°C wurden bei einer OD_{600} von 1,0 bei 42°C für 15 Minuten inkubiert. Dieser Hitzeschock induziert die Synthese der Fusionsproteine. Die Kulturen werden bei 37°C für drei bis vier Stunden weiter inkubiert, bevor die Zellen durch Zentrifugation geerntet werden. Die Zellen wurden durch osmotischen Schock und Ultraschallbehandlung lysiert und das Gesamtzellprotein in Phenol extrahiert (mit Trisma-Base auf pH 8 eingestellt).
[0369] Rohprotein wurde aus der Phenolphase durch Zugabe von 2,5 Volumina Ethanol und Zentrifugation präzipitiert. Das Proteinpellet wurde in einem Puffer gelöst, der 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8, und 0,1 M Dithioerythrol enthielt. Nach Gelfiltration an Sephadex G-25 (Pharmacia, Schweden) in 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, 10 mM 2-Mercaptoethanol wurde das Proteinrohpräparat auf eine Ni^{2+}-aktivierte NTA-Agarosesäule zur Reinigung (Hochuli et al. (1988)) des FXΔγ-Fusionsproteins aufgegeben und anschließend dem zyklischen Faltungsverfahren unterzogen.
[0370] Alle für die Flüssigchromatographie hergestellten Puffer wurden vor der Zugabe von Reduktionsmittel und/oder Verwendung unter Vakuum entgast.
[0371] Die Herstellung und „Ladung" der Ni^{2+}NTA-Agarosesäule ist unter Beispiel i beschrieben.
[0372] Nach Aufgeben der Proteinrohextrakte auf die Ni^{2+}NTA-Agarosesäule wurden die Fusionsproteine vom Großteil der E. coli- und λ-Phagenproteine durch Waschen mit einem Säulenvolumen des Beladungspuffers, gefolgt von 6 M Guanidiniumchlorid, 50 mM Tris-HCl und 10 mM 2-Mercaptoethanol gereinigt, bis die optische Dichte (OD) bei 280 nm stabil war.
[0373] Das Fusionsprotein wurde an der Ni^{2+}-NTA-Agarosesäule unter Anwendung des Gradientenmanagerprofils wie in Tabelle 5 beschrieben und 0,5 M NaCl, 50 mM Tris-HCl, pH 8, 2 mM $CaCl_2$, und 3,0 mM reduziertes Glutathion als Puffer B neu gefaltet. Die reduzierte/oxidierte Glutathionlösung wurde als 100fache Stammlösung durch Zugabe von 9,9 M H_2O_2 zu einer gerührten Lösung von 0,2 M reduziertem Glutathion vor Zugabe zu Puffer A frisch hergestellt.
[0374] Nach Beendigung des zyklischen Neufaltungsverfahrens wurde das FXΔγ-Fusionsprotein von der Ni^{2+}NTA-Agarosesäule mit dem nicht denaturierenden Puffer eluiert. Die Menge des FXΔγ-Fusionsproteins wurde auf 15 mg geschätzt. Nur ungefähr ein Drittel dieses Fusionsproteinmaterials schien monomer zu sein, wie durch nicht reduzierende SDS-PAGE-Analyse beurteilt wurde, was einer Gesamteffizienz des Faltungsverfahrens von ungefähr 10% entspricht.
[0375] FXΔγ-Fusionsprotein in nicht denaturierendem Puffer wurde aktiviert, indem die Lösung des rekombi-

nanten Proteins durch eine kleine Säule von Trypsin-Agarose (an Sepharose CL-6B (Pharmacia, Schweden) immobilisiertes Trypsin) geschickt wurde.

[0376] Das aktivierte rekombinante FXΔγ-Fusionsprotein wurde mittels Standardverfahren mit chromogenen Substraten auf das proteolytische Aktivitäts- und Substratspezifitätsprofil hin getestet. Aktivitäts- und Substratspezifitätsprofil waren von den für natürliches Rinder-FX_a erhaltenen nicht unterscheidbar.

Beispiel 6

Produktion und Faltung der Kringle-Domänen 1 und 4 aus Human-Plasminogen

[0377] Dieses Beispiel beschreibt die Produktion der Lysin bindenden Kringle-Domänen 1 und 4 aus Human-Plasminogen (K1 bzw. K4) als FX_a-spaltbare Fusionsproteine in E. coli und die Reinigung und In-vitro-Spaltung der K1- und K4-Fusionsproteine.

[0378] Ein die für Human-Plasminogen kodierende Volllängen-cDNA enthaltender, in den allgemeinen Klonierungsvektor pUC18 (großzügigerweise von Dr. Earl Davie, Seattle, USA) klonierter Plasmidklon wurde als Templat in einer zur Produktion von K1 (entsprechend den Aminosäureresten Ser_{81} bis Glu_{162} im so genannten Glu-Plasminogen) und K4 (entsprechend den Aminosäureresten Val_{354} bis Ala_{439} im so genannten Glu-Plasminogen) entsprechenden cDNA-Fragmenten entworfenen Polymerasekettenreaktion (PCR) verwendet. Die Primer Seq.-ID Nr. 27 und Seq.-ID Nr. 28 wurden in der K1 produzierenden PCR und die Primer Seq.-ID Nr. 29 und Seq.-ID Nr. 30 in der K4 produzierenden PCR verwendet.

[0379] Die für K1 und K4 kodierenden, amplifizierten Leseraster waren an ihren 5'-Enden über die PCR-Reaktion an Nucleotidsequenzen, die in Seq.-ID Nr. 27 und Seq.-ID Nr. 28 umfasst sind und für die Aminosäuresequenz Seq.-ID Nr. 37 kodieren, die eine Spaltstelle für die Rinderrestriktionsprotease FX_a darstellen (Nagai und Thøgersen, Methods in Enzymology 152, 461–481 (1987)). Das amplifizierte K1-DNA-Fragment wurde in den E. coli-Expressionsvektor $pLcIIMLCH_6$, der eine modifizierte Form von pLcIIMLC ist (Nagai et al., Nature 332, 284–286 (1988)), durch Insertion eines Oligonucleotids kloniert, das für sechs Histidinylreste C-terminal des Myosin-Leichtkettenfragments kodiert. Die Konstruktion des resultierenden Plasmids $pLcIIMLCH_6FX$-K1 ist in **Fig.** 12 dargestellt. Das amplifizierte K4-DNA-Fragment wurde in den E. coli-Expressionsvektor $pLcIIH_6$, der eine modifizierte Form von pLcII ist (Nagai und Thøgersen, Methods in Enzymology 152, 461–481 (1987)), durch Insertion eines Oligonucleotids kloniert, das für sechs Histidinylreste C-terminal des cII-Fragments kodiert. Die Konstruktion des resultierenden Plasmids $pLcIIH_6FX$-K4 ist in **Fig.** 13 dargestellt und in **Fig.** 14 ist die Aminosäuresequenz von Human-„Glu"-Plasminogen gezeigt (Seq.-ID Nr. 54).

[0380] Beide Plasmide, $pLcIIMLCH_6$K1 und $pLcIIH_6FX$-K4 wurden in E. coli QY13-Zellen wie in Nagai und Thøgersen, Methods in Enzymology 152, 461–481 (1987), beschrieben gezüchtet und exprimiert. Exponentiell wachsende Kulturen bei 30°C wurden bei einer OD_{600} von 1,0 bei 42°C für 15 Minuten inkubiert. Dieser Hitzeschock induziert die Synthese der Fusionsproteine. Die Kulturen werden bei 37°C für drei bis vier Stunden weiter inkubiert, bevor die Zellen durch Zentrifugation geerntet werden. Die Zellen wurden durch osmotischen Schock und Ultraschallbehandlung lysiert und das Gesamtzellprotein in Phenol extrahiert (mit Trisma-Base auf pH 8 eingestellt).

[0381] Rohprotein wurde aus der Phenolphase durch Zugabe von 2,5 Volumina Ethanol und Zentrifugation präzipitiert. Das Proteinpellet wurde in einem Puffer gelöst, der 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8, und 0,1 M Dithioerythrol enthielt. Nach Gelfiltration an Sephadex G-25 (Pharmacia, Schweden) in 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, 10 mM 2-Mercaptoethanol und 2 mM Methionin wurde das Proteinrohpräparat auf eine Ni^{2+}-aktivierte NTA-Agarosesäule zur Reinigung (Hochuli et al. (1988)) der K1- und K4- Fusionsproteins aufgegeben und anschließend dem zyklischen Faltungsverfahren unterzogen.

[0382] Alle für die Flüssigchromatographie hergestellten Puffer wurden vor der Zugabe von Reduktionsmittel und/oder Verwendung unter Vakuum entgast.

[0383] Die Herstellung und „Ladung" der Ni^{2+}NTA-Agarosesäule ist unter Beispiel 1 beschrieben.

[0384] Nach Aufgeben der Proteinrohextrakte auf die Ni^{2+}NTA-Agarosesäule wurden die Fusionsproteine vom Großteil der E. coli- und λ-Phagenproteine durch Waschen mit einem Säulenvolumen des Beladungspuffers, gefolgt von 6 M Guanidiniumchlorid, 50 mM Tris-HCl, 10 mM 2-Mercaptoethanol und 2 mM Methionin gereinigt, bis die optische Dichte (OD) des Säuleneluats bei 280 nm stabil war.

[0385] Das Fusionsprotein wurde an der Ni^{2+}-NTA-Agarosesäule unter Anwendung des Gradientenmanagerprofils wie in Tabelle 4 beschrieben und 0,5 M NaCl, 50 mM Tris-HCl, pH 8, 10 mM 6-Aminohexansäure (Ɛ-Aminocapronsäure, ?-ACA), 0,33 mM Methionin und 2,0 mM/0,2 mM reduziertes/oxidiertes Glutathion als Puffer A und 4 M Harnstoff, 0,5 M NaCl, 50 mM Tris-HCl, pH 8, 10 mM Ɛ-ACA, 2 mM Methionin und 3 mM reduziertes Glutathion als Puffer B neu gefaltet. Die reduzierte/oxidierte Glutathionlösung wurde als 100fache Stammlösung durch Zugabe von 9,9 M H_2O_2 zu einer gerührten Lösung von 0,2 M reduziertem Glutathion vor Zugabe zu Puffer A frisch hergestellt.

[0386] Nach Beendigung des zyklischen Faltungsverfahrens wurden beide der K1- und K4-Fusionsproteine

aus der Ni^{2+}NTA-Agarosesäule mit einem Puffer eluiert, der 0,5 M NaCl, 50 mM Tris-HCl, 5 mM EDTA, pH 8, enthielt. Die Fusionsproteine, die an der Ni^{2+}NTA-Agarosesäule aggregiert und präzipitiert waren, wurden in Puffer B eluiert.

[0387] Praktisch das gesamte K1- und K4-Fusionsproteinmaterial wurde durch den nicht reduzierenden Puffer aus den Ni^{2+}NTA-Agarosesäulen eluiert. Die geschätzte Ausbeute von K1-Fusionsprotein und K4-Fusionsprotein betrug ungefähr 60 mg. Praktisch das gesamte K1-Fusionsprotein sowie K4-Fusionsprotein schien monomer zu sein, wie mittels nicht reduzierender SDS-PAGE-Analyse beurteilt wurde, was einer Effizienz des obigen Faltungsverfahrens von 90% entspricht.

[0388] Die SDS-PAGE-Analyse der Produktion der rekombinanten Kringles 1 und 4 ist in **Fig.** 17 dargestellt.

[0389] Das K1-Fusionsprotein und K4-Fusionsprotein wurden mittels Affinitätschromatographie an Lysin-Sepharose CL6B (Pharmacia, Schweden) weiter gereinigt. Die Fusionsproteine wurden mit einem 0,5 M NaCl, 50 mM Tris-HCl, pH 8, 10 mM ?-ACA enthaltendem Puffer aus den Affinitätssäulen eluiert.

[0390] Die Bindung an Lysin-Agarose ist normalerweise als Hinweis der korrekten Faltung von Lysin bindenden Kringle-Domänen anerkannt.

[0391] Die dreidimensionale Struktur der rekombinanten K1- und K4-Proteindomänen, die mit diesem zyklischen Faltungsverfahren produziert und die durch Freisetzung aus dem terminalen Fusionsschwanz vollständig prozessiert und anschließend mittels Ionenaustauschchromatographie gereinigt worden sind, ist mittels Röntgenstrahlenbeugung (durchgeführt von Dr. Robert Huber) und zweidimensionale NMR-Analyse (durchgeführt von stud. scient. Peter Reinholdt und Dr. Flemming Poulsen) bestätigt worden.

Beispiel 7

Produktion rekombinanter Fragmente aus Human-α_2-Macroglobulin und Hühner-Ovostatin in E. coli und Neufaltung

[0392] Dieses Beispiel beschreibt die Produktion der rezeptorbindenden Domäne von Human-α_2-Macroglobulin (α_2-MRBDv) als FX_a-spaltbares Fusionsprotein in E. coli und die Reinigung des rekombinanten α_2-MRBDv nach FX_a-Spaltung.

[0393] Das für das α_2-Macroglobulin-Leseraster von Aminosäurerest Val_{1299} bis Ala_{1451} (α_2-MRDv) kodierende 462 bp-DNA-Fragment wurde in einer Polymerasekettenreaktion (PCR). im wesentlichen nach dem Protokoll von Saiki et al. (1988) amplifiziert. pA2M (großzügigerweise von Dr. T. Kristensen bereitgestellt), das die Volllängen-cDNA von α_2-Macroglobulin enthält, wurde als Templat und die Oligonucleotide Seq.-ID Nr. 31 und Seq.-ID Nr. 32 als Primer verwendet. Das amplifizierte, kodierende Leseraster war am 5'-Ende über die PCR-Reaktion an eine in Seq.-ID Nr. 7 umfasste Nucleotidsequenz gebunden, die für die Aminosäuresequenz Seq.-ID Nr. 37 kodiert, die eine Spaltstelle für Rinderrestriktionsprotease FX_a darstellen (Nagai und Thøgersen (1987)). Das amplifizierte DNA-Fragment wurde in den E. coli-Expressionsvektor pT_7H_6 subkloniert (Christensen et al. (1991)). Die Konstruktion des resultierenden Plasmids pT_7H_6FX-α_2MDRv (Human-α_2-MRDv exprimierend) ist in **Fig.** 18 dargestellt und die Aminosäuresequenz des exprimierten Proteins ist in **Fig.** 19 gezeigt (Seq.-ID Nr. 55).

[0394] Rekombinantes Human-α_2-MRDv wurde durch Züchten und Exprimieren des Plasmids pT_7H_6FX-α_2-MRDv in E. coli BL21-Zellen im mittleren Maßstab (2 × 1 Liter) wie von Studier und Moffat, J. Mol. Biol. 189, 113–130 (1986), beschrieben produziert. Exponentiell wachsende Kulturen bei 37°C wurden bei einer OD_{600} von 0,8 mit Bakteriophagen λCE6 in einer Multiplizität von ungefähr 5 infiziert. Die Kulturen wurden bei 37°C für weitere drei Stunden gezüchtet, bevor die Zellen durch Zentrifugation geerntet wurden. Die Zellen wurden durch osmotischen Schock und Ultraschallbehandlung lysiert und das Gesamtzellprotein in Phenol extrahiert (mit Trisma-Base auf pH 8 gestellt). Protein wurde aus dem Phenol durch Zugabe von 2,5 Volumina Ethanol und Zentrifugation präzipitiert. Das Proteinpellet wurde in einem Puffer gelöst, der 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8, und 50 mM Dithioerythrol enthielt. Nach Gelfiltration an Sephadex G-25 (Pharmacia, LKB, Schweden) in 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, 10 mM 2-Mercaptoethanol wurde das rohe Proteinpräparat auf eine Ni^{2+}-aktivierte NTA-Agarosesäule zur Reinigung (Hochuli et al. (1988)) des Fusionsproteins MGS-HHHHHHGSIEGR-α_2-MRDv (worin MGSHHHHHHGSIEGR Seq.-ID Nr. 48 ist) aufgegeben und anschließend dem zyklischen Faltungsverfahren unterworfen.

[0395] Die Herstellung und „Ladung" der Ni^{2+}NTA-Agarosesäule ist in Beispiel 1 beschrieben.

[0396] Alle für die Flüssigchromatographie hergestellten Puffer wurden vor Zugabe von Reduktionsmittel und/oder Verwendung im Vakuum entgast.

[0397] Beim Aufgeben des Proteinrohextrakts auf die Ni^{2+}NTA-Agarosesäule wurde das Fusionsprotein MGSHHHHHHGSIEGR-α_2-MRDv (worin MGSHHHHHHGSIEGR Seq.-ID Nr. 48 ist) vom Hauptteil der E. coli- und λ-Phagen-Proteine durch Waschen mit einem Säulenvolumen des Beladungspuffers, gefolgt von 6 M Guanidiniumchlorid, 50 mM Tris-HCl und 10 mM 2-Mercaptoethanol gereinigt, bis die optische Dichte (OD) bei 280 nm des Eluats stabil war.

[0398] Das Fusionsprotein wurde an der Ni^{2+}NTA-Agarosesäule unter Anwendung des Gradientenmanagerprofils wie in Tabelle 4 beschrieben und 0,5 M NaCl, 50 mM Tris-HCl, pH 8, und 2,0 mM/0,2 mM reduziertes/oxidiertes Glutathion als Puffer A und 8 M Harnstoff, 0,5 M NaCl, 50 mM Tris-HCl, pH 8 und 5 mM reduziertes Glutathion als Puffer B neu gefaltet. Die reduzierte/oxidierte Glutathionlösung wurde als eine 200fache Stammlösung durch Zugabe von 9,9 M H_2O zu einer gerührten Lösung von 0,2 M reduziertem Glutathion vor Zugabe zu Puffer A frisch hergestellt.

[0399] Nach Beendigung des zyklischen Faltungsverfahrens wurde das und 1 mM α_2-MRDv-Fusionsprotein mit einem 0,5 M NaCl, 50 mM Tris-HCl, 20 mM EDTA, pH 8, enthaltenden Puffer aus den Ni^{2+}NTA-Agarosesäulen eluiert. Fusionsprotein, das an der Ni^{2+} NTA-Agarosesäule aggregiert und präzipitiert war, wurde in Puffer B eluiert.

[0400] Ungefähr 50% des Fusionsproteinmaterials wurde in dem wässrigen Elutionspuffer eluiert. Die Hälfte dieses Fusionsproteinmaterials schien monomer zu sein, wie mittels nicht reduzierender SDS-PAGE-Analyse beurteilt wurde.

[0401] Rekombinantes α_2-MRDv-Protein wurde vom N-terminalen Fusionsschwanz durch Spaltung mit der Restriktionsprotease FX_a bei Raumtemperatur in einem Gewichtsverhältnis von ungefähr 50 zu Eins für vier Stunden freigesetzt. Nach der Spaltung wurde das α_2-MRDv-Protein vom ungespaltenen Fusionsprotein, dem freigesetzten Fusionsschwanz und FX_a durch Gelfiltration an Sephadex G-25 in 10 mM NaCl, 50 mM Tris-HCl, pH 8, gefolgt von Ionenaustauschchromatographie an Q-Sepharose isoliert: α_2MRDv wurde in einem linearen Gradienten (über 10 Säulenvolumina) von 10 mM NaCl, 10 mM Tris-HCl, pH 8, bis 500 mM NaCl, 10 mM Tris-HCl, pH 8, eluiert. Das α_2-MRDv-Protein eluierte bei 150 mM NaCl.

[0402] Die rekombinante α_2-MRDv-Domäne bindet an den α_2-M-Rezeptor mit ähnlicher Affinität für den Rezeptor, wie sie das vollständigen α_2-Macroglobulin-Molekül aufweist (in Bezug auf das geschätzte K_D bei der Ein-Ligand/Ein-Rezeptor-Bindung (Moestrup und Gliemann 1991)). Die Bindungsanalyse wurde von Dr. Søren K. Moestrup und stud. scient. Kåre Lehmann durchgeführt).

Beispiel 8

Produktion von rekombinanten, vom Forellen-Virus VHS-Hüll-Glykoprotein G hergeleiteten Fragmenten in E. coli und Neufaltung

[0403] Die Expression und In-vitro-Neufaltung von rekombinanten, vom Forellen-Virus VHS hergeleiteten Fragmenten in E. coli als FX_a-spaltbare Fusionsproteine wird unter Anwendung allgemeiner Strategien und Verfahren durchgeführt, die analog zu jenen sind, die in der allgemeinen Beschreibung des „zyklischen Neufaltungsverfahrens" dargestellt und in Beispielen 1 bis einschließlich 6 angeführt sind.

Beispiel 9

Produktion von rekombinantem Human-Tetranectin und von rekombinanten von Human-Tetranectin hergeleiteten Fragmenten in E. coli und Neufaltung

[0404] Tetranectin ist ein tetrameres Proteine, das aus vier identischen und nicht kovalent verbundenen Einzelkettenuntereinheiten von 181 Aminosäureresten (17 kDa) besteht. Jede Untereinheit enthält drei Disulfidbrücken und bindet Ca^{2+}. Tetranectin findet sich in Plasma und assoziiert mit der extrazellulären Matrix. Tetranectin bindet spezifisch an Plasminogen-Kringle 4. Diese Bindung kann spezifisch mit Lysin oder ω-Aminosäuren titriert werden.

[0405] Die für das der reifen Tetranectin-Einzelkettenuntereinheit entsprechenden Leseraster kodierende cDNA wurde durch spezifische Amplifikation kloniert, und zwar in einer Polymerasekettenreaktion (PCR) (Saiki et al. (1988)) der Nucleotidsequenzen vom Aminosäurerest Glu_1 bis Val_{181} unter Verwendung von Erststrang-Oligo-dT-geprimter cDNA, die aus Human-Plazenta-Gesamt-RNA als Templat synthetisiert wurde. Die in der PCR verwendeten Primer waren Seq.-ID Nr. 33 und Seq.-ID Nr. 34. RNA-Extraktion und cDNA-Synthese wurden mittels Standardverfahren durchgeführt.

[0406] Das amplifizierte, für die Monomeruntereinheit von Tetranectin kodierende Leseraster war am 5'-Ende über die PCR-Reaktion an ein Nucleotidsequenzen gebunden, die für die Aminosäuresequenz Seq.-ID Nr. 37 kodieren, die eine Spaltstelle für Rinderrestriktionsprotease FX_a darstellen (Nagai und Thøgersen (1987)). Ein Glycinrest wurde aufgrund der speziellen Konstruktion des 5'-PCR-Primers (Seq.-ID Nr. 33) zwischen dem C-terminalen Argininrest der FX_a-Spaltstelle (Seq.-ID Nr. 37) und dem Tetranectin-Glu_1-Rest insertiert. Das amplifizierte DNA-Fragment wurde in den E. coli-Expressionsvektor pT_7H_6 subkloniert (Christensen et al. (1991)). Die Konstruktion des resultierenden Plasmids pT_7H_6FX-TETN (das Tetranectin-Monomer exprimierend) ist in **Fig.** 20 dargestellt und die Aminosäuresequenz des exprimierten Proteins ist in **Fig.** 21 gezeigt (in Seq.-ID Nr. 56 ist die vom Volllängen-Leseraster kodierte Aminosäuresequenz gezeigt).

[0407] Um das Tetranectin-Monomer zu produzieren, wurde das Plasmid pT_7H_6FX-TETN im mittleren Maßstab (4 × 1 Liter; 2 × TY-Medium, 5 mM $MgSO_4$ und 100 µg Ampicillin) in E. coli-BL21-Zellen wie von Studier und Moffat, J. Mol. Biol. 189, 113–130 (1986), beschrieben gezüchtet. Exponentiell wachsende Kulturen bei 37°C wurden bei einer OD_{600} von 0,8 mit Bakteriophagen λCE6 in einer Multiplizität von ungefähr 5 infiziert. Die Kulturen wurden bei 37°C für weitere drei Stunden gezüchtet und die Zellen durch Zentrifugation geerntet. Die Zellen wurden in 150 ml 0,5 M NaCl, 10 mM Tris-HCl, pH 8, und 1 mM EDTA, pH 8, resuspendiert. Phenol (100 ml, auf pH 8 gestellt) wurde zugesetzt und das Gemisch ultraschallbehandelt, um das Gesamtprotein zu extrahieren. Protein wurde aus der Phenolphase durch 2,5 Volumina Ethanol und Zentrifugation präzipitiert.

[0408] Das Proteinpellet wurde in einem Puffer gelöst, der 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8, und 0,1 M Dithioerythrol enthielt. Nach Gelfiltration an Sephadex G-25 (Pharmacia, LKB, Schweden) in 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, und 10 mM 2-Mercaptoethanol wurde das rohe Proteinpräparat auf eine Ni^{2+}-aktivierte NTA-Agarosesäule (Ni^{2+}NTA-Agarosesäule, 75 ml, vorgewaschen mit 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, und 10 mM 2-Mercaptoethanol) zur Reinigung (Hochuli et al. (1988)) des Fusionsproteins MGSHHHHHHGSIEGR-TETN (worin MGSHHH-HHHGSIEGR Seq.-ID Nr. 48 ist) aufgegeben.

[0409] Die Herstellung und „Ladung" der Ni^{2+}NTA-Agarosesäule ist in Beispiel 1 beschrieben.

[0410] Alle für die Flüssigchromatographie hergestellten Puffer wurden vor Zugabe von Reduktionsmittel und/oder Verwendung im Vakuum entgast.

[0411] Die Säule wurde mit 200 ml Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, und 10 mM 2-Mercaptoethanol (Puffer I) und 100 ml 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8, und 10 mM 2-Mercaptorethanol (Puffer II) gewaschen. Das MGSHHHHHHGSIEGR-TETN-Fusionsprotein wurde mit 10 mM EDTA, pH 8, enthaltendem Puffer II eluiert und das Eluat wurde an Sephadex G-25 mit Puffer I als Elutionsmittel gelfiltriert.

[0412] Das eluierte Protein wurde dann neu gefaltet. Das Fusionsprotein MGSHHHHHHGSI-EGR-TETN (worin MGSHHHHHHGSIEGR Seq.-ID Nr. 48 ist) wurde mit 100 ml Ni^{2+} NTA-Agarose gemischt. Das das gebundene Protein enthaltende Harz wurde in eine Säule mit 5 cm Durchmesser gepackt und mit Puffer I, das mit $CaCl_2$ auf 2 mM ergänzt war, gewaschen. Das Fusionsprotein wurde an der Ni^{2+}NTA-Agarosesäule bei 11–12°C unter Anwendung des Gradientenmanagerprofils wie in Tabelle 4 beschrieben und 0,5 M NaCl, 50 mM Tris-HCl, pH 8, 2 mM $CaCl_2$ und 2,0 mM/0,2 mM reduziertem/oxidiertem Glutathion als Puffer A und 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, 2 mM $CaCl_2$ und 3 mM reduziertem Glutathion als Puffer B neu gefaltet. Die reduzierte/oxidierte Glutathionlösung wurde als eine 200fache Stammlösung durch Zugabe von 9,9 M H_2O zu einer gerührten Lösung von 0,2 M reduziertem Glutathion vor Zugabe zu Puffer A frisch hergestellt.

[0413] Nach Beendigung des zyklischen Faltungsverfahrens wurde das Tetranectin-Fusionsprotein von der Ni^{2+}NTA-Agarosesäule mit einem 0,5 M NaCl, 50 mM Tris-HCl, 25 mM EDTA, pH 8, enthaltendem Puffer eluiert. Das Tetranectin-Fusionsprotein wurde mit FX_a bei 4°C über Nacht in einem Molverhältnis von 1 : 300 gespalten. Nach der FX_a-Spaltung wurde die Proteinprobe mittels Ultrafiltration an einer YM10-Membran (Amicon) konzentriert. Rekombinantes Tetranectin wurde nach 10facher Verdünnung der Proteinprobe mit 2 mM $CaCl_2$ mittels Ionenaustauschchromatographie an Q-Sepharose (Pharmacia, Schweden) in einem linearen Gradienten von 10 mM Tris-HCl, pH 8, 2 mM $CaCl_2$ bis 10 mM Tris, pH 8, 2 mM $CaCl_2$ und 0,5 M NaCl isoliert.

[0414] Mit diesem Verfahren produziertes Tetranectin wurde von Dr. Inge Clemmensen Rigshospitalet, Copenhagen, analysiert. Dr. Clemmensen fand, dass sich das rekombinante Tetranectin bezüglich der Bindung an Plasminogen-Kringle 4 und Expression antigener Stellen gleich wie natürlich isoliertes Human-Tetranectin verhielt.

[0415] Vorversuche zum Vergleich der Effizienz der Neufaltung unter Anwendung des „zyklischen Neufaltungsverfahrens" auf an die Ni^{2+}NTA-Agarosesäule gebundenes Tetranectin-Fusionsprotein mit in einem Dialyseschlauch enthaltenem rekombinantem Tetranectin weisen auf eine signifikant verbesserte Ausbeute von löslichem Monomer aus der Neufaltungsstrategie in Lösung hin. Wird jedoch eines der Produkte der zyklischen Verfahren dem Disulfid-Reshuffling in Lösung in Gegenwart von 5 mM $CaCl_2$ unterzogen, wird praktisch das gesamte Polypeptidmaterial zum korrekt gefalteten Tetranectin-Tetramer umgesetzt.

[0416] In einem Dialyseschlauch enthaltenes, denaturiertes und rekombinantes authentisches Tetranectin wurde über 15 zyklische Behandlungen mit Puffer B (6 M Harnstoff, 100 mM NaCl, 50 mM Tris-HCl, pH 8, 2 mM/0,2 mM reduziertes/oxidiertes Glutathion, 2 mM $CaCl_2$ und 0,5 mM Methionin) und Puffer A (100 mM NaCl, 50 mM Tris-HCl, pH 8, 2 mM/0,2 mM reduziertes/oxidiertes Glutathion, 2 mM $CaCl_2$ und 0,5 mM Methionin) neu gefaltet.

Beispiel 10

Produktion und Faltung eines in intrazellulär in E. coli exprimierten Diakörpers: Gegen Tumornekrosefaktor gerichteter Mab 32-Diakörper

[0417] Diakörper (beschrieben in Hollinger et al. (1993)) sind künstliche zweiwertige und bispezifische Anti-

körperfragmente.

[0418] Dieses Beispiel beschreibt die Produktion eines gegen Tumornekrosefaktor Alpha (TNF-α) gerichteten Diakörpers, hergeleitet vom monoklonalen Maus-Antikörper Mab 32 (Rathjen et al. (1991, 1992); Australische Patentanmeldung 7.576; EP-A-486.526).

[0419] Ein Phagemidklon, pCANTAB5-myc-Mab32-5, der im Diakörper-Format kodiertes Mab32 enthält (PCT/GB93/02492), wurde großzügigerweise von Dr. G. Winter, Cambridge Antibody Technology (CAT) Ltd., Cambridge, UK, bereitgestellt. pCANTAB5-myc-Mab32-5-DNA wurde als Templat in einer Polymerasekettenreaktion (PCR) (Saiki et al. (1988) unter Verwendung der zur Produktion eines dem vollständigen künstlichen Diakörper entsprechenden cDNA-Fragments entworfenen Primer Seq.-ID Nr. 35 und Seq.-ID Nr. 36 verwendet. Das amplifizierte, kodierende Leseraster war am 5'-Ende über die PCR-Reaktion an eine in Seq.-ID Nr. 35 umfasste Nucleotidsequenz gebunden, die für die Aminosäuresequenz Seq.-ID Nr. 37 kodiert, die eine Spaltstelle für die Rinderrestriktionsprotease FX_a (Nagai und Thøgersen (1987)) darstellt. Das amplifizierte DNA-Fragment wurde in den E. coli-Expressionsvektor pT_7H_6 (Christensen et al. (1991)) subkloniert. Die Konstruktion des resultierenden Plasmids pT_7H_6FX-DB32 (den Mab32-Diakörper exprimierend) ist in **Fig.** 22 dargestellt und die Aminosäuresequenz des exprimierten Proteins ist in **Fig.** 23 gezeigt (in Seq.-ID Nr. 57 ist die vom Volllängen-Leseraster kodierte Aminosäuresequenz gezeigt).

[0420] Um das Diakörperfragment herzustellen, wurde das Plasmid pT_7H_6FX-DB32 im mittleren Maßstab (4 × 1 Liter; 2 × TY-Medium, 5 mM $MgSO_4$ und 100 µg Ampicillin) in E. coli BL21-Zellen wie von Studier und Moffat, J. Mol. Biol. 189, 113–130 (1986), beschrieben gezüchtet. Exponentiell wachsende Kulturen bei 37°C wurde bei einer OD_{600} von 0,8 mit Bakteriophagen λCE6 in einer Multiplizität von ungefähr 5 infiziert. Vierzig Minuten nach der Infektion wurde Rifampicin zugesetzt (0,2 g in 2 ml Methanol je Liter Medium). Die Kulturen wurden bei 37°C für weitere drei Stunden gezüchtet und die Zellen durch Zentrifugation geerntet. Die Zellen wurden in 150 ml 0,5 M NaCl, 10 mM Tris-HCl, pH 8, und 1 mM EDTA, pH 8, resuspendiert. Es wurde Phenol (100 ml, auf pH 8 gestellt) zugesetzt und das Gemisch ultraschallbehandelt, um das Gesamtprotein zu extrahieren. Protein wurde aus der Phenolphase mit 2,5 Volumina Ethanol und Zentrifugation präzipitiert.

[0421] Das Proteinpellet wurde in einem 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8, und 0,1 M Dithioerythrol enthaltendem Puffer gelöst. Nach Gelfiltration an Sephadex G-25 (Pharmacia, LKB, Schweden) in 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, und 10 mM 2-Mercaptoethanol wurde das Proteinrohpräparat auf eine Ni^{2+}-aktivierte NTA-Agarosesäule (Ni^{2+}NTA-Agarose, 75 ml, vorgewaschen mit 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, und 10 mM 2-Mercaptoethanol) zur Reinigung (Hochuli et al. (1988)) des Fusionsproteins MGSHHHHHHGSIEGR-DB32 (worin MGSHHHHHHGSI-EGR Seq.-ID Nr. 48 ist) aufgegeben.

[0422] Die Herstellung und „Ladung" der Ni^{2+}NTA-Agarosesäule ist in Beispiel 1 beschrieben.

[0423] Alle für die Flüssigchromatographie hergestellten Puffer wurden vor der Zugabe von Reduktionsmittel und/oder Verwendung unter Vakuum entgast.

[0424] Die Säule wurde mit 200 ml 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, und 10 mM 2-Mercaptoethanol (Puffer I) und 100 ml 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8, und 10 mM 2-Mercaptoethanol (Puffer II) gewaschen. Das MGSHHHHHH-GSIEGR-DB32-Fusionsprotein wurde mit 10 mM EDTA, pH 8, enthaltendem Puffer II eluiert und das Eluat an Sephadex G25 mit Puffer I als Elutionsmittel gelfiltriert.

[0425] Das Protein wurde eluiert und dann neu gefaltet. Das Fusionsprotein MGSHHHHHH-GSIEGR-DB32 (worin MGSHHHHHHGSIEGR Seq.-ID Nr. 48 ist) wurde mit 100 ml Ni^{2+}NTA-Agarose gemischt. Das das gebundene Protein enthaltende Harz wurde in eine Säule mit 5 cm Durchmesser gepackt und mit Puffer I gewaschen. Das Fusionsprotein wurde an der Ni^{2+}NTA-Agarosesäule bei 11–12°C unter Anwendung des Gradientenmanagerprofils wie in Tabelle 4 beschrieben und 0,5 M NaCl, 50 mM Tris-HCl, pH 8, und 2,0 mM/0,2 mM reduziertem/oxidiertem Glutathion als Puffer A und 8 M Harnstoff, 1 M NaCl, 50 mM Tris-HCl, pH 8, und 3 mM reduziertem Glutathion als Puffer B neu gefaltet. Die reduzierte/oxidierte Glutathionlösung wurde als eine 200fache Stammlösung durch Zugabe von 9,9 M H_2O_2 zu einer gerührten Lösung von 0,2 M reduziertem Glutathion vor Zugabe zu Puffer A frisch hergestellt.

[0426] Nach Beendigung des zyklischen Faltungsverfahrens wurde das DB32-Fusionsprotein von der Ni^{2+}NTA-Agarosesäule mit einem 0,5 M NaCl, 50 mM Tris-HCl, 25 mM EDTA, pH 8, enthaltendem Puffer eluiert und auf 5 mM GSH, 0,5 mM GSSG eingestellt und für 12 bis 15 Stunden bei 20°C inkubiert. Das Fusionsprotein wurde dann mittels Ultrafiltration unter Verwendung von YM10-Membranen 50fach konzentriert und mittels Zentrifugation geklärt.

[0427] Das DB32-Fusionsproteindimer wurde mittels Gelfiltration mit einer Superose 12-Säule (Pharmacia, Schweden) mit PBS als Elutionsmittel gereinigt.

[0428] Die Gesamtausbeute von korrekt gefaltetem DB32-Fusionsprotein aus diesem Verfahren betrug 4 mg pro Liter.

[0429] Eine Analyse mittels nicht reduzierender SDS-PAGE aus verschiedenen Reinigungsstadien ist in **Fig.** 26 gezeigt.

[0430] Das N-terminale MGSHHHHHHGSIEGR- (Seq.-ID Nr. 48) Fusionspeptid wurde vom DB32-Protein durch Spaltung mit der Restriktionsprotease FX_a (Molverhältnis 1 : 5 Fx_a : DB32-Fusionsprotein) bei 37°C für

20 Stunden abgespalten. Dies zeigt sich als Auftreten einer niedermolekularen Bande gerade unterhalb des ungespaltenen Fusionsproteins in **Fig.** 26.

[0431] Das neu gefaltete DB32-Protein wurde von Cambridge Antibody Technology Ltd. (CAT) analysiert. Es wurde gefunden, das DB32 spezifisch an TNF-α bindet und mit dem Mab32-Gesamtantikörper um die Bindung an TNF-α konkurriert. Darüber hinaus konkurrierte Schaf-Anti-301-Antiserum, das durch Immunisieren von Schafen mit einem Peptid hergestellt wurde, das für die ersten 18 Aminosäuren, die zumindest einen Teil des vom Maus-Mab32 erkannten Epitops umfassen, von Human-TGF-α kodiert, sowohl mit DB32, als auch mit Mab32 um die Bindung an TGF-α

Beispiel 13

Produktion und Neufaltung von Human-Psoriasin in E. coli

[0432] Psoriasin ist ein Ca^{2+}-bindendes Protein von 100 Aminosäureresten (11,5 kDa) mit einer einzigen Domäne. Psoriasin enthält eine einzige Disulfidbrücke. Das Protein, von dem angenommen wird, das der S100-Proteinfamilie angehört, ist in psoriatischer Haut im höchsten Maße up-reguliert und erfährt in primären Human-Keratinozyten eine abnormale Differenzierung.

[0433] Das Plasmid pT_7H_6FX-PS.4 (großzügigerweise von Dr. P. Madsen, Institute of Medical Biochemistry, University of Aarhus, Dänemark, bereitgestellt) ist früher von Hoffmann et al. (1994) beschrieben worden. Die für das Psoriasin-Protein kodierende Nucleotidsequenz von Ser_2 bis Gln_{101} ist am 5'-Ende an die für die Aminosäuresequenz MGS-HHHHHHGSIEGR (Seq.-ID Nr. 48) kodierende Nucleotidsequenz gebunden. Eine Karte des pT_7H_6FX-PS.4 ist in **Fig.** 24 dargestellt und die Aminosäuresequenz von Psoriasin ist in **Fig.** 25 aufgeführt (in Seq.-ID Nr. 58 ist die vom Volllängen-Leseraster kodierte Aminosäuresequenz gezeigt).

[0434] Rekombinantes Human-Psoriasin wurde aus dem Plasmid pT_7H_6FX-PS.4 in E. coli BL21-Zellen gezüchtet und exprimiert und das Gesamtzellprotein wie beschrieben (Hoffmann et al. (1994) extrahiert. In Ethanol präzipitiertes Gesamtprotein wurde in einem Puffer gelöst, der 6 M Guanidiniumchlorid, 50 mM Tris-HCl, pH 8, und 50 mM Dithioerythrit enthielt. Nach Gelafiltration an Sephadex G-25 (Pharmacia, LKB, Schweden) in 8 M Harnstoff, 0,5 M NaCl, 50 mM Tris-HCl, pH 8, und 5 mM 2-Mercaptoethanol wurde das Rohprotein auf eine Ni^{2+}-aktivierte NTA-Agarosesäule (Ni^{2+}NTA-Agarose) zur Reinigung (Hochuli et al. (1988)) des Fusionsproteins MGSHHHHHHGSIEGR-Psoriasin (worin MGSHHHHHHGSIEGR die Seq.-ID Nr. 48 ist) aufgegeben und anschließend dem zyklischen Faltungsverfahren unterzogen.

[0435] Die Herstellung und „Ladung" der Ni^{2+}NTA-Agarosesäule ist in Beispiel 1 beschrieben.

[0436] Alle für die Flüssigchromatographie hergestellten Puffer wurden vor der Zugabe von Reduktionsmittel und/oder Verwendung unter Vakuum entgast.

[0437] Nach Aufgeben des Proteinrohextrakts auf die Ni^{2+}NTA-Agarosesäule wurde das Fusionsprotein MGSHHHHHHGSIEGR-Psoriasin (worin MGSHHHHHHGSIEGR Seq.-ID Nr. 48 ist) vom Hauptteil der E. coli- und λ-Phagen-Proteine durch Waschen mit einem Säulenvolumen des Beladungspuffers, gefolgt von 6 M Guanidiniumchlorid, 50 mM Tris-HCl und 5 mM 2-Mercaptoethanol gereinigt, bis die optische Dichte (OD) bei 280 nm des Eluats stabil war.

[0438] Das Fusionsprotein wurde an der Ni^{2+}NTA-Agarosesäule unter Anwendung des Gradientenmanagerprofils wie in Tabelle 4 beschrieben und 0,5 M NaCl, 50 mM Tris-HCl, pH 8, 2 mM $CaCl_2$ und 1,0 mM/0,1 mM reduziertes/oxidiertes Glutathion als Puffer A und 8 M Harnstoff, 0,5 M NaCl, 50 mM Tris-HCl, pH 8, 2 mM $CaCl_2$ und 5 mM reduziertes Glutathion als Puffer B neu gefaltet. Die reduzierte/oxidierte Glutathionlösung wurde als eine 200fache Stammlösung durch Zugabe von 9,9 M H_2O_2 zu einer gerührten Lösung von 0,2 M reduziertem Glutathion vor Zugabe zu Puffer A frisch hergestellt.

[0439] Nach Beendigung des zyklischen Faltungsverfahrens wurde das Psoriasin-Fusionsprotein mit einem 0,5 M NaCl, 50 mM Tris-HCl, 10 mM EDTA, pH 8, enthaltenden Puffer aus der Ni^{2+}NTA-Agarosesäule eluiert. Fusionsprotein, das an der Ni^{2+}NTA-Agarosesäule aggregiert und präzipitiert war, wurde in Puffer B eluiert.

[0440] Ungefähr 95% des Fusionsproteinmaterials wurde vom nicht denaturierenden Elutionspuffer eluiert. Wie mittels nicht reduzierender SDS-PAGE-Analyse beurteilt wurde, schienen 75% des löslichen Fusionsproteinmaterials monomer zu sein, was eine Gesamteffizienz des Faltungsverfahrens von ungefähr 70% ergibt. Die Effizienz des früher beschriebenen Faltungsverfahrens zur Produktion von rekombinantem Human-Psoriasin (Hoffman et al. (1994)) wurde auf weniger als 25% geschätzt.

[0441] Das Psoriasin-Fusionsprotein wurde mit FX_a in einem Molverhältnis von 100 : 1 für 48 Stunden bei Raumtemperatur gespalten. Nach Gelfiltration in einem 20 mM Na-Acetat, pH 5, und 20 mM NaCl enthaltendem Puffer an Sephadex G-25 wurde die Proteinprobe auf eine S-Sepharose-Ionentauschersäule (Pharmacia) aufgegeben. Monomeres rekombinantes Psoriasin wurde über 5 Säulenvolumina mit einem linearen Gradienten von 20 mM Na-Acetat, pH 5, 20 mM NaCl bis 0,5 M NaCl eluiert. Monomeres Psoriasin eluierte bei 150 mM NaCl. Psoriasin-Dimere und Multimere höherer Ordnung gemeinsam mit ungespaltenem Fusionsprotein eluierten später im Gradienten. Das gespaltene, gereinigte rekombinante Protein enthaltende Fraktionen wur-

den an Sephadex G25 in einen Puffer gelfiltriert, der 150 mM NaCl, 10 mM Tris-HCl, pH 7,4, enthielt und bei 4°C gelagert.

Beispiel 12

Evaluierungsverfahren zur Eignungstestung von Thiolverbindungen zur Verwendung als Reduktionsmittel in der zyklischen Neufaltung und Bestimmung der optimalen Konzentrationen von Denaturierungsmitteln und Disulfid-Shuffling-Mitteln zur Optimierung der zyklischen Neufaltungsverfahren

[0442] Um die Ausbeute an korrekt gefaltetem Protein aus der zyklischen Neufaltung zu verbessern, sollte die Anzahl produktiver Zyklen maximiert werden (siehe „Zusammenfassung der Erfindung"). Produktiver Zyklen sind durch Schritte der Denaturierung gekennzeichnet, wo fehlgefaltetes Protein auf dem Weg zu Dead-End-Aggregat-Konformationszuständen in ungefaltete Konformationszustände gerettet wird, während der Großteil des bereits korrekt gefalteten Proteins in Konformationszuständen verbleibt, die dazu fähig sind, während des Neufaltungsschritts des Zyklus in den neugefalteten Zustand zurückzuschnappen.

[0443] Eine Reihe von Disulfidbrücken enthaltenden Proteinen, wie z. B. β_2-Microglobulin falten sich bekanntermaßen mit hoher Effizienz (< 95%) neu, wenn sie hohen Konzentrationen von Denaturierungsmitteln ausgesetzt werden, solange ihre Disulfidbrücken intakt bleiben.

[0444] Dieses Beispiel beschreibt, wie die Eignung einer Thiolverbindung zur Verwendung in der zyklischen Neufaltung auf Basis ihrer Fähigkeit evaluiert wird, korrekt von inkorrekten Disulfidbrücken zu unterscheiden und wie die Konzentrationen des Denaturierungsmittels und/oder Reduktionsmittels, das in den Denaturierungsschritten zu verwenden ist, optimiert wird, um die Anzahl produktiver Zyklen zu maximieren. Als Modellsystem wählten die Erfinder ein Gemisch von mono-, di- und multimeren Formen des gereinigten rekombinanten β_2-Microglobulins. Das spezielle Ziel der Erfinder war es, die Stabilität verschiedener topologischer Formen des Human-β_2-Microglobulins gegen Reduktion durch fünf verschiedene Reduktionsmittel bei verschiedenen Denaturierungsmittelkonzentrationen zu analysieren.

[0445] Human-β_2-Microglobulin (wie in Beispiel 13 beschrieben produziert) in 6 M Guanidiniumhydrochlorid, 50 mM Tris-HCl und 10 mM 2-Mercaptoethanol, pH 8, wurde in nicht denaturierenden Puffer (50 mM Tris-HCl, 0,5 M NaCl, pH 8) gelfiltriert. Nur ein Teil des Proteins in der Probe war im nicht denaturierenden Puffer löslich. Nach 48 Stunden Luftexposition erschien die Proteinlösung trüb. Nicht reduzierende SDS-PAGE-Analyse zeigte, das der Großteil des Proteins zu multimeren Formen oxidiert worden und nur ein Teil oxidiert und monomer war (**Fig.** 27, Spur 1).

[0446] Die Proteinlösung wurde in eine Anzahl von Röhrchen aliquotiert und es wurden unterschiedliche Mengen Harnstoff zugesetzt, während die Konzentration von Protein und Salz konstant gehalten wurde.

[0447] Reduktionsmittel, entweder Glutathion, Cysteinethylester, N-Acetyl-L-cystein, Mercaptobernsteinsäure oder 2-Mercaptoethanol wurde der Anordnung von Proteinproben mit unterschiedlichen Harnstoffkonzentrationen zugesetzt. Jedes der Reduktionsmittel wurde auf eine Endkonzentration von 4 mM zugesetzt. Die Proteinproben wurden bei Raumtemperatur für 10 Minuten inkubiert und dann wurden die Thiolguppen durch Zugabe von Iodessigsäure auf eine Endkonzentration von 12 mM blockiert. Schließlich wurden die Proteinproben mittels nicht reduzierender SDS-PAGE analysiert (**Fig.** 27-32). Die Zusammensetzungen der in der nicht reduzierenden SDS-PAGE verwendeten Testproben, sowie die Ergebnisse sind unten in den folgenden Tabellen angeführt; in den Zeilen, welche die Fähigkeit des gewählten Reduktionsmittels angeben, Disulfidbrücken zu reduzieren, zeigt die Markierung „++" gute Fähigkeit, „++" mittlere Fähigkeit, „+" schwache Fähigkeit an, wogegen keine Markierung anzeigt, das keine messbare Wirkung beobachtet werden konnte.

Zusammensetzung von Proben, die in der SDS-PAGE von Fig. 27 verwendet wurden

Test-Nr.	1	2	3	4	5	6	7	8	9	10	11
µl Proteinlösung	36	36	36	36	36	36	36	36	36	36	36
µl Puffer A	160	160	140	120	100	80	70	60	50	40	20
µl Puffer B	0	0	20	40	60	80	90	100	110	120	140
µl GSH	0	4	4	4	4	4	4	4	4	4	4
M Harnstoff	0	0	1	2	3	4	4.5	5	5.5	6	7
Fähigkeit, falsche Disulfidbrücken zu reduzieren			+	+	++	++	+++	+++	+++	+++	+++
Fähigkeit, korrekte Disulfidbrücken zu reduzieren										+	+++

Puffer A : 50 mM Tris.HCl pH 8, 0.5 M NaCl
Puffer B: 10 M Harnstoff, 50 mM Tris.HCl pH 8, 0.5 M NaCl
GSH: 0.2 M Gluthation
Proteinlösung: 2 mg/ml hβ_2m, 50 mM Tris.HCl pH 8, 0.5 M NaCl

Zusammensetzung von Proben, die in der SDS-PAGE von Fig. 28 verwendet wurden

Test-Nr.	1	2	3	4	5	6	7	8	9
µl Proteinlösung	36	36	36	36	36	36	36	36	36
µl Puffer A	160	160	140	120	100	80	60	40	20
µl Puffer B	0	0	20	40	60	80	100	120	140
µl CE	0	4	4	4	4	4	4	4	4
M Harnstoff	0	0	1	2	3	4	5	6	7
Fähigkeit, falsche Disulfidbrücken zu reduzieren		++	++	++	+++	+++	+++	+++	+++
Fähigkeit, korrekte Disulfidbrücken zu reduzieren							++	+++	+++

Puffer A : 50 mM Tris.HCl pH 8, 0.5 M NaCl
Puffer B: 10 M Harnstoff, 50 mM Tris.HCl pH 8, 0.5 M NaCl
CE: 0.2 M L-Cysteinethylester
Proteinlösung: 2 mg/ml hβ_2m, 50 mM Tris.HCl pH 8, 0.5 M NaCl

Zusammensetzung von Proben, die in der SDS-PAGE von Fig. 29 verwendet wurden

Test-Nr.	1	2	3	4	5	6	7	8	9
µl Proteinlösung	36	36	36	36	36	36	36	36	36
µl Puffer A	160	160	140	120	100	80	60	40	20
µl Puffer B	0	0	20	40	60	80	100	120	140
µl ME	0	4	4	4	4	4	4	4	4
M Harnstoff	0	0	1	2	3	4	5	6	7
Fähigkeit, falsche Disulfidbrücken zu reduzieren		++	++	++	+++	+++	+++	+++	+++
Fähigkeit, korrekte Disulfidbrücken zu reduzieren						+	++	+++	+++

Puffer A : 50 mM Tris.HCl pH 8, 0.5 M NaCl
Puffer B: 10 M urea, 50 mM Tris.HCl pH 8, 0.5 M NaCl
ME: 0.2 M 2-Mercaptoethanol
Proteinlösung 2 mg/ml hβ_2m, 50 mM Tris.HCl pH 8, 0.5 M NaCl

Zusammensetzung von Proben, die in der SDS-PAGE von Fig. 30 verwendet wurden

Test-Nr.	1	2	3	4	5	6	7	8	9
µl Proteinlösung	36	36	36	36	36	36	36	36	36
µl Puffer A	160	160	140	120	100	80	60	40	20
µl Puffer B	0	0	20	40	60	80	100	120	140
µl MSA	0	4	4	4	4	4	4	4	4
M Harnstoff	0	0	1	2	3	4	5	6	7
Fähigkeit, falsche Disulfidbrücken zu reduzieren		++	++	++	++	++	+++	+++	+++
Fähigkeit, korrekte Disulfidbrücken zu reduzieren							++	+++	+++

Puffer A : 50 mM Tris.HCl pH 8, 0.5 M NaCl
Puffer B: 10 M Harnstoff, 50 mM Tris.HCl pH 8, 0.5 M NaCl
MSA: 0.2 M Mercaptobernsteinsäure
Proteinlösung: 2 mg/ml hβ_2m, 50 mM Tris.HCl pH 8, 0.5 M NaCl

Zusammensetzung von Proben, die in der SDS-PAGE von Fig. 31 verwendet wurden

Test-Nr.	1	2	3	4	5	6	7	8	9
µl Proteinlösung	36	36	36	36	36	36	36	36	36
µl Puffer A	160	160	140	120	100	80	60	40	20
µl Puffer B	0	0	20	40	60	80	100	120	140
µl AC	0	4	4	4	4	4	4	4	4
M Harnstoff	0	0	1	2	3	4	5	6	7
Fähigkeit, falsche Disulfidbrücken zu reduzieren		+	++	++	+++	+++	+++	+++	+++
Fähigkeit, korrekte Disulfidbrücken zu reduzieren					+	++	+++	+++	+++

Puffer A : 50 mM Tris.HCl pH 8, 0.5 M NaCl
Puffer B: 10 M Harnstoff, 50 mM Tris.HCl pH 8, 0.5 M NaCl
AC: 0.2 M N-Acetyl-L-cystein
Proteinlösung: 2 mg/ml hβ_2m, 50 mM Tris.HCl pH 8, 0.5 M NaCl

[0448] Die verschiedenen topologischen Formen von β_2-m könnten mittels nicht reduzierender SDS-PAGE-Gelelektrophorese getrennt werden. Die am schnellsten wandernde Bande stellt die oxidierte monomere Form dar. Dieser Bande folgt unmittelbar das reduzierte β_2-m mit einer etwas niedrigeren Wanderungsgeschwindigkeit, wogegen die multimeren Formen des Proteins im Gel viel langsamer wandern.

[0449] In dieser Analyse sondieren die Erfinder die Fähigkeit jedes der fünf getesteten Reduktionsmittel, die Disulfidbrücken multimerer Formen von β_2-Microglobulin zu reduzieren, ohne die korrekt gebildete Disulfidbrücke der monomeren, oxidierten Form signifikant zu reduzieren.

[0450] Die Ergebnisse aus den Analysen (**Fig.** 27–32) sind, zusammenfassend, die folgenden: N-Acetyl-L-Cystein und Mercaptobernsteinsäure sind unter die angewendeten Bedingungen im Wesentlichen nicht in der Lage, zwischen korrekten und inkorrekten Disulfidbrücken zu unterscheiden. Glutathion, Cysteinethylester und 2-Mercaptorethanol sind alle in der Lage – innerhalb von 10 Minuten und innerhalb einzelner charakteristischer Bereiche von Harnstoffkonzentrationen – Disulfidbrücken multimerer Formen signifikant zu reduzieren, während der Großteil des oxidierten, monomeren β_2-m in der oxidierten Form verbleibt. Glutathion hat im Vergleich zu Cysteinethylester und 2-Mercaptoethanol eindeutig die Fähigkeit, inkorrekte Disulfidbrücken bei höheren Harnstoffkonzentrationen selektiv zu reduzieren, und daher währe Glutathion unter der Auswahl von getesteten Thiolen das Reduktionsmittel der Wahl für die zyklische Neufaltung von Human-β_2-Microglobulin. Als Konsequenz dieser Experimente wurde die Harnstoffkonzentration im reduzierenden Puffer B für das in Beispiel 13 verwendete Neufaltungsverfahren von 8 M (Beispiel 1) auf 6 M herabgesetzt, was eine Verbesserung der Neufaltungsausbeute von Human-β_2-Microglobulin von 53% auf 87% bewirkte.

Beispiel 13

Neufaltung von gereinigtem Human-β_2-Microglobulin:

Vergleichende Analyse von drei Neufaltungsverfahren

[0451] Der folgende Satz von Experimenten wurde durchgeführt, um vergleichbare quantitative Daten zu erhalten, um die Bedeutung der Zyklisierung für die Neufaltungsausbeute gegenüber simplen Neufaltungsverfahren zu beurteilen, die einen schrittweisen oder graduellen Übergang von stark denaturierenden und reduzierenden Bedingungen zu nicht denaturierenden und nicht reduzierenden Bedingungen umfassen.

[0452] Gereinigtes, neu gefaltetes rekombinantes, wie in Beispiel 1 beschrieben erhaltenes Human-β_2-Microglobulin-Fusionsprotein wurde reduziert und denaturiert, um Anfangsmaterialien frei von Verunreinigungen, wie z. B. proteolytischen Abbauprodukten oder geringfügigen Anteilen von durch irreversible Oxidation oder andere chemische Derivatisierung beschädigtem Fusionsprotein zu erhalten.

[0453] In einem ersten Schritt wurde das in Beispiel 12 beschriebene Optimierungsverfahren angewendet,

um die Bedingungen für die in Beispiel 1 beschriebene zyklische Neufaltung zu modifizieren, um die Anzahl produktiver Zyklen zu erhöhen. Das optimierte Neufaltungsprotokoll war identisch mit jenem in Beispiel 1 beschriebenen, wie auch die Puffer und andere experimentelle Parameter mit der Ausnahme, dass der Puffer B in den vorliegenden Experimenten aus 6 M Harnstoff, 50 mM Tris-HCl, pH 8, 0,5 M NaCl, 4 mM Glutathion bestand.

[0454] Drei Chargen von reinem Fusionsprotein wurden neu gefaltet, während es an Ni^{2+}-beladener NTA-Agarose wie in Beispiel 1 beschrieben gebunden war, und zwar unter Verwendung der vorliegenden Puffer B-Zusammensetzung. Eine Charge wurde wie in Beispiel 1 beschrieben der Puffer-Zyklisierung unterzogen, und für Chargen zwei und drei wurde die Zyklisierung durch einen monotonen linearen Puffergradienten (100% B bis 0% B über 24 Stunden) bzw. einen Stufengradienten (100% B bis 0% B in einem Schritt, gefolgt von 0% B-Puffer für 24 Stunden) ersetzt. In jedem Neufaltungsexperiment wurde das gesamte Polypeptidmaterial wie in Beispiel 1 beschrieben als lösliche Fraktion, die unter nicht denaturierenden Bedingungen eluierbar war, und als eine verbleibende unlösliche Fraktion gewonnen, die nur unter denaturierenden und reduzierenden Bedingungen eluierbar war. Die Ausbeuten von korrekt gefaltetem Fusionsprotein wurden durch quantitative densitometrische Analyse (optischer Scanner HW und GS-370 densitometrische Analyse-SW-Packet von Hoeffer Scientific, CA, USA) von Coomassie-gefärbten SDS-PAGE-Gelen gemessen, an denen geeignet verdünnte, abgemessene Aliquote löslicher und unlöslicher Fraktionen unter reduzierenden oder nicht reduzierenden Bedingungen getrennt worden sind, wie es erforderlich war, um die Trennung von korrekten Disulfid-überbrücktem Monomer von löslichen Polymeren in löslichen Fraktionen zu erlauben. Wo es zur Erlangung verlässlicher densitometrischer Daten sowohl für intensive, als auch für schwache Banden in einer Gel-Spur erforderlich war, wurden mehrere Probenverdünnungen gescannt und analysiert, um neu skalierte Datensätze zu erhalten.

Details und Ergebnisse der Versuche

Gereinigtes, denaturiertes und reduziertes Fusionsprotein:

[0455] Eine Charge Human-β_2-Microglobulin-Fusionsprotein wurde wie in Beispiel 1 neu gefaltet. 96% des Fusionsproteins wurde in der löslichen Fraktion gewonnen (**Fig.** 32, Spuren 2–5). 56% dieser löslichen Fraktion befand sich in monomerer und Disulfid-überbrückter Form. Daher betrug die Neufaltungsgesamteffizienz 53%. Monomeres Fusionsprotein wurde von Multimeren mittels Ionenaustauschchromatographie an S-Sepharose (Pharmacia, Schweden) gereinigt: Eine nach der Neufaltung erhaltene lösliche Fraktion wurde an Sephadex G-25 (Pharmacia, Schweden) in einen 5 mM NaCl und 5 mM Tris-HCl, pH 8, enthaltenden Puffer gelfiltriert, mit Wasser auf das Doppelte verdünnt und dann auf die S-Sepharose-Säule aufgegeben, die dann mit einem Gradienten (5 Säulenvolumina von 2,5 mM Tris-HCl, pH 8, 2,5 mM NaCl bis 25 mM Tris-HCl, pH 8, 100 mM NaCl) eluiert wurde. Das monomere, korrekt gefaltete, auf > 95% Reinheit gereinigte Fusionsprotein (**Fig.** 32, Spuren 6 und 7) wurde dann auf 6 M Guanidiniumhydrochlorid und 0,1 M DTE gebracht, in einen 8 M Harnstoff, 50 mM Tris-HCl, pH 8, 1 M NaCl und 10 mM 2-Mercaptorethanol enthaltenden Puffer gelfiltriert und dann in Aliquote geteilt, die als Anfangsmaterial für die unten beschriebenen Neufaltungsexperimente verwendet wurde.

Zyklische Neufaltung des gereinigten Fusionsproteins:

[0456] Ein Aliquot des denaturierten, reduzierten Fusionsproteins wurde auf eine Ni^{2+}-beladene NTA-Säule aufgegeben, die dann mit einem Säulenvolumen eines 6 M Guanidiniumhydrochlorid, 50 mM Tris-HCl, pH 8, und 10 mM 2-Mercaptoethanol enthaltenden Puffers gewaschen wurde.

[0457] Das Fusionsprotein wurde dann der Puffer-Zyklisierung nach dem in Tabelle 1 gezeigten Schema unter Verwendung von Puffer A: 50 mM Tris-HCl, pH 8, 0,5 M NaCl, und 3,2 mM/0,4 mM reduziertes/oxidiertes Glutathion, und Puffer B: 50 mM Tris-HCl, pH 8, 0,5 M NaCl, 6 M Harnstoff und 4 mM reduziertes Glutathion unterzogen. Nach Beendigung des Puffer-Zyklisierung wurde das Fusionsprotein durch Elution der Säule mit einem 50 mM Tris-HCl, pH 8, 0,5 M NaCl und 20 mM EDTA enthaltendem Puffer quantitativ in löslicher Form gewonnen. 87% wurde in der korrekten monomeren, Disulfid-überbrückten Form erhalten (**Fig.** 32, Spuren 8 und 9).

Neufaltung des gereinigten Fusionsproteins mittels linearem Gradienten:

[0458] Ein Aliquot von denaturiertem, reduziertem Fusionsprotein wurde auf eine Ni^{2+}-beladene NTA-Säule aufgegeben, die dann mit einem Säulenvolumen eines 6 M Guanidiniumhydrochlorid, 50 mM Tris-HCl, pH 8 und 10 mM 2-Mercaptoethanol enthaltendem Puffer, gefolgt von einem Säulenvolumen eines 50 mM Tris-HCl, pH 8, 0,5 M NaCl, 6 M Harnstoff und 4 mM reduziertes Glutathion enthaltendem Puffer gewaschen wurde.

[0459] Ein 24-stündiger linearer Gradient von 100% B bis 100% A wurden dann bei 2 ml/min angewendet, wobei Puffer A: 50 mM Tris-HCl, pH 8, 0,5 M NaCl und 3,2 mM/0,4 mM reduziertes/oxidiertes Glutathion, und Puffer B: 50 mM Tris-HCl, pH 8, 0,5 M NaCl, 6 M Harnstoff und 4 mM reduziertes Glutathion verwendet wurden. Nach Beendigung des Gradienten wurde die lösliche Fraktion des Fusionsproteins in einem 50 mm Tris-HCl, pH 8, 0,5 M NaCl und 20 mM EDTA enthaltendem Puffer eluiert. Die verbleibende unlösliche Fraktion wurde von der Säule in einen 50 mM Tris-HCl, pH 8, 1 M NaCl, 8 M Harnstoff, 10 mM 2-Mercaptoethanol und 20 mM EDTA enthaltenden Puffer extrahiert.

[0460] 48% des Fusionsproteins wurde in der löslichen Fraktion gewonnen und 60% der löslichen Fraktion wurde in der korrekt Disulfid-überbrückten Form gewonnen. Die Gesamteffizienz der erhaltenen Faltung betrug daher 29% (**Fig.** 33, Spuren 5–7).

Neufaltung des gereinigten Fusionsproteins mittels Pufferschritt:

[0461] Ein Aliquot von denaturiertem, reduziertem Fusionsprotein wurde auf eine Ni^{2+}-beladene NTA-Säule aufgegeben, die dann mit einem Säulenvolumen eines 6 M Guanidiniumhydrochlorid, 50 mM Tris-HCl, pH 8 und 10 mM 2-Mercaptoethanol enthaltendem Puffer gewaschen wurde.

[0462] 50 mM Tris-HCl, pH 8, 0,5 M NaCl und 3,2 mM/0,4 mM reduziertes/oxidiertes Glutathion enthaltender Puffer wurde dann bei 2 ml/min für 24 Stunden auf die Säule aufgegeben, bevor die lösliche Fraktion des Fusionsproteins in einem 50 mM Tris-HCl, pH 8, 0,5 M NaCl und 20 mM EDTA enthaltendem Puffer gewonnen wurde. Die verbleibende unlösliche Fraktion wurde von der Säule in einem 50 mM Tris-HCl, pH 8, 1 M NaCl, 8 M Harnstoff, 10 mM 2-Mercaptoethanol und 20 mM EDTA enthaltendem Puffer extrahiert.

[0463] 34% des Fusionsproteins wurde in der löslichen Fraktion gewonnen und 28% der löslichen Fraktion wurde in der korrekten, monomeren Disulfid-überbrückten Form gewonnen. Die Gesamteffizienz der erhaltenen Faltung betrug daher 9,5% (**Fig.** 33, Spuren 1– 3).

Schlussfolgerungen:

[0464] Zusammenfassend kann unter Verwendung von Human-β_2-Microglobulin als Modellprotein gefolgert werden, dass a) die unkomplizierte Pufferoptimierung und verbesserte Reinigung des Fusionsproteins vor der zyklischen Neufaltung die Neufaltungsausbeute signifikant erhöhte (von 53% auf 87%) und b) progressive Denaturierungs/Renaturierungs-Zyklisierung der einstufigen Neufaltung unter ansonsten vergleichbaren experimentellen Bedingungen um einen sehr großen Faktor (87% gegenüber 29% oder 9,5% Ausbeute) überlegen ist.

Literatur

[0465] Christensen, J. H., Hansen, P. K., Lillelund, O., and Thøgersen, H. C. (1991). Sequence-specific binding of the N-terminal three-finger fragment of Xenopus transcription factor IIIA to the internal control region of a 5S RNA gene. FEBS Letters, 295: 181–184.

Dalbøge, H., Dahl, H. -H., M., Pedersen, J., Hansen, J., W., and T., Kristensen (1987). A Novel Enzymatic Method for Production of Authentic hGH From an Eschericia coli Produced hGh-Precursor. Bio/Technology, 5: 161–164.

Datar, R., V., Cartwright, T., and C. -G. Rosen (1993). Process Economics of Animal Cell and Bacterial Fermentations: A Case Study Analysis of Tissue Plasminogen Activator. Bio/Technology, 11: 349–357.

Herz, J., Hanmann, U., Rogne, S., Myklebost, O., Gausepohl, H., and Stanley, R. K. (1988), Surface location and high affinity for calcium of a 500 kd liver membrane protein closely related to the LDL-receptor suggest a physiological role as lipoprotein receptor. EMBO J., 7: 4119–4127.

Hoffmann, H. J., Olsen, E., Etzerodt, M., Madsen, P., Thrøgersen, H. C., Kruse, T., and Celis J. E. (1994). Psoriasin Binds Calcium and Is Differentially Regulated With Respect to Other Members of the S100 Protein Family. J. Dermatol. Invest. In press.

Hochuli, E., W. Bannwarth, H. Dōbeli, R. Gentz, and D. Stüber. 1988. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. Bio/Technology, 6: 1321–1325.

Holliger., P., Prospero, T., and G. Winter (1993). "Diabodies": Small bivalent and bispecific antibody fragments. Proc. Natl. Acad. Sci. USA. 90: 6444–6448.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.

Nagai, K., and H. C. Thøgersen. 1987. Synthesis and Sequence-Specific Proteolysis of Hybrid Proteins Produced in Escherichia coli. Methods in Enzymology, 152: 461–481.

Nagai, K., Nakaseko, Y., Nasmyth, K., and Rhodes, D. (1988). Zinc-finger motifs expressed in E. coli and folded in vitro direct specific binding to DNA. Nature, 332: 284–286.

Nykjær A., Petersen C. M., Møller B., Jensen P. H., Moestrup S. K., Holtet T. L., Etzerodt M., Thøgersen H. C., Munch M., Andreasen P. A., and Gliemann J. (1992). Purified α_2-Macroglobulin Receptor/LDL Receptor-related Protein Binds Urokinase-Plasminogen Activator Inhibitor Type-1 Complex. J. Biol. Chem. 267: 14543–14546.
Rathjen, D. et al. (1991), Mol. Immunol. 28, p29.
Rathjen, D. et al. (1992), Brit. J. Cancer 65, 852–856.
Saiki, R. K., Gelfant, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487–491.
Studier, F. W. and Moffat, B. A. 1986. Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-level Expression of Cloned Genes. J. Mol. Biol., 189: 113–130.
The Regents of the University of California. Enterokinasecleavable linker sequence. EP 035384.

SEQUENZPROTOKOLL

(1) ALLGEMEINE INFORMATIONEN:

(i) ANMELDER:

(A) NAME: Denzyme ApS

(B) STRASSE: Gustav Wieds Vej 10

(C) STADT: Aarhus C

(E) BUNDESSTAAT: Dänemark

(E) LAND: Dänemark

(F) PLZ: 8000

(ii) TITEL DER ERFFINDUNG: Verbessertes Verfahren zum Falten von Proteinen

(iii) ANZAHL DER SEQUENZEN: 47

(iv) COMPUTERLESBARE FORM:

(A) ART DES MEDIUMS: Diskette

(B) COMPUTER: IBM PC-kompatibel

(C) BETRIEBSSYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release Nr. 1.0, Version Nr. 1.25 (EPO)

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 1:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 1554 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: doppelt

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: cDNA

(iii) HYPOTHETISCH: JA

(iii) ANTI-SENSE: NEIN

(vi) URSPRÜNGLICHE QUELLE:

(A) ORGANISMUS: Bos taurus

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: CDS

(B) POSITION: 76..1551

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID Nr. 1:

```
AGCCTGGGCG AGCGGACCTT GCCCTGGAGG CCTGTTGCGG CAGGGACTCA CGGCTGTCCT    60

CGGAAGGGCC CCACC ATG GCG GGC CTG CTG CAT CTC GTT CTG CTC AGC ACC    111
                 Met Ala Gly Leu Leu His Leu Val Leu Leu Ser Thr
                   1               5                  10

GCC CTG GGC GGC CTC CTG CGG CCG GCG GGG AGC GTG TTC CTG CCC CGG    159
Ala Leu Gly Gly Leu Leu Arg Pro Ala Gly Ser Val Phe Leu Pro Arg
          15                  20                  25

GAC CAG GCC CAC CGT GTC CTG CAG AGA GCC CGC AGG GCC AAC TCA TTC    207
Asp Gln Ala His Arg Val Leu Gln Arg Ala Arg Arg Ala Asn Ser Phe
     30                  35                  40

TTG GAG GAG GTG AAG CAG GGA AAC CTG GAG CGA GAG TGC CTG GAG GAG    255
Leu Glu Glu Val Lys Gln Gly Asn Leu Glu Arg Glu Cys Leu Glu Glu
 45                  50                  55                  60

GCC TGC TCA CTA GAG GAG GCC CGC GAG GTC TTC GAG GAC GCA GAG CAG    303
Ala Cys Ser Leu Glu Glu Ala Arg Glu Val Phe Glu Asp Ala Glu Gln
                 65                  70                  75

ACG GAT GAA TTC TGG AGT AAA TAC AAA GAT GGA GAC CAG TGT GAA GGC    351
Thr Asp Glu Phe Trp Ser Lys Tyr Lys Asp Gly Asp Gln Cys Glu Gly
           80                  85                  90

CAC CCG TGC CTG AAT CAG GGC CAC TGT AAA GAC GGC ATC GGA GAC TAC    399
His Pro Cys Leu Asn Gln Gly His Cys Lys Asp Gly Ile Gly Asp Tyr
         95                  100                 105
```

```
ACC TGC ACC TGT GCG GAA GGG TTT GAA GGC AAA AAC TGC GAG TTC TCC     447
Thr Cys Thr Cys Ala Glu Gly Phe Glu Gly Lys Asn Cys Glu Phe Ser
    110                 115                 120

ACG CGT GAG ATC TGC AGC CTG GAC AAT GGA GGC TGC GAC CAG TTC TGC     495
Thr Arg Glu Ile Cys Ser Leu Asp Asn Gly Gly Cys Asp Gln Phe Cys
125                 130                 135                 140

AGG GAG GAG CGC AGC GAG GTG CGG TGC TCC TGC GCG CAC GGC TAC GTG     543
Arg Glu Glu Arg Ser Glu Val Arg Cys Ser Cys Ala His Gly Tyr Val
                145                 150                 155

CTG GGC GAC GAC AGC AAG TCC TGC GTG TCC ACA GAG CGC TTC CCC TGT     591
Leu Gly Asp Asp Ser Lys Ser Cys Val Ser Thr Glu Arg Phe Pro Cys
            160                 165                 170

GGG AAG TTC ACG CAG GGA CGC AGC CGG CGG TGG GCC ATC CAC ACC AGC     639
Gly Lys Phe Thr Gln Gly Arg Ser Arg Arg Trp Ala Ile His Thr Ser
        175                 180                 185

GAG GAC GCG CTT GAC GCC AGC GAG CTG GAG CAC TAC GAC CCT GCA GAC     687
Glu Asp Ala Leu Asp Ala Ser Glu Leu Glu His Tyr Asp Pro Ala Asp
    190                 195                 200

CTG AGC CCC ACA GAG AGC TCC TTG GAC CTG CTG GGC CTC AAC AGG ACC     735
Leu Ser Pro Thr Glu Ser Ser Leu Asp Leu Leu Gly Leu Asn Arg Thr
205                 210                 215                 220

GAG CCC AGC GCC GGG GAG GAC GGC AGC CAG GTG GTC CGG ATA GTG GGC     783
Glu Pro Ser Ala Gly Glu Asp Gly Ser Gln Val Val Arg Ile Val Gly
                225                 230                 235

GGC AGG GAC TGC GCG GAG GGC GAG TGC CCA TGG CAG GCT CTG CTG GTC     831
Gly Arg Asp Cys Ala Glu Gly Glu Cys Pro Trp Gln Ala Leu Leu Val
            240                 245                 250

AAC GAA GAG AAC GAG GGA TTC TGC GGG GGC ACC ATC CTG AAC GAG TTC     879
Asn Glu Glu Asn Glu Gly Phe Cys Gly Gly Thr Ile Leu Asn Glu Phe
        255                 260                 265

TAC GTC CTC ACG GCT GCC CAC TGC CTG CAC CAG GCC AAG AGG TTC ACG     927
Tyr Val Leu Thr Ala Ala His Cys Leu His Gln Ala Lys Arg Phe Thr
    270                 275                 280

GTG AGG GTC GGC GAC CGG AAC ACA GAG CAG GAG GAG GGC AAC GAG ATG     975
Val Arg Val Gly Asp Arg Asn Thr Glu Gln Glu Glu Gly Asn Glu Met
285                 290                 295                 300
```

```
GCA CAC GAG GTG GAG ATG ACT GTG AAG CAC AGC CGC TTT GTC AAG GAG   1023
Ala His Glu Val Glu Met Thr Val Lys His Ser Arg Phe Val Lys Glu
                305                 310                 315

ACC TAC GAC TTC GAC ATC GCG GTG CTG AGG CTC AAG ACG CCC ATC CGG   1071
Thr Tyr Asp Phe Asp Ile Ala Val Leu Arg Leu Lys Thr Pro Ile Arg
            320                 325                 330

TTC CGC CGG AAC GTG GCG CCC GCC TGC CTG CCC GAG AAG GAC TGG GCG   1119
Phe Arg Arg Asn Val Ala Pro Ala Cys Leu Pro Glu Lys Asp Trp Ala
        335                 340                 345

GAG GCC ACG CTG ATG ACC CAG AAG ACG GGC ATC GTC AGC GGC TTC GGG   1167
Glu Ala Thr Leu Met Thr Gln Lys Thr Gly Ile Val Ser Gly Phe Gly
    350                 355                 360

CGC ACG CAC GAG AAG GGC CGC CTG TCG TCC ACG CTC AAG ATG CTG GAG   1215
Arg Thr His Glu Lys Gly Arg Leu Ser Ser Thr Leu Lys Met Leu Glu
365                 370                 375                 380

GTG CCC TAC GTG GAC CGC AGC ACC TGT AAG CTG TCC AGC AGC TTC ACC   1263
Val Pro Tyr Val Asp Arg Ser Thr Cys Lys Leu Ser Ser Ser Phe Thr
                385                 390                 395

ATT ACG CCC AAC ATG TTC TGC GCC GGC TAC GAC ACC CAG CCC GAG GAC   1311
Ile Thr Pro Asn Met Phe Cys Ala Gly Tyr Asp Thr Gln Pro Glu Asp
            400                 405                 410

GCC TGC CAG GGC GAC AGT GGC GGC CCC CAC GTC ACC CGC TTC AAG GAC   1359
Ala Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Arg Phe Lys Asp
        415                 420                 425

ACC TAC TTC GTC ACA GGC ATC GTC AGC TGG GGA GAA GGG TGC GCG CGC   1407
Thr Tyr Phe Val Thr Gly Ile Val Ser Trp Gly Glu Gly Cys Ala Arg
    430                 435                 440

AAG GGC AAG TTC GGC GTC TAC ACC AAG GTC TCC AAC TTC CTC AAG TGG   1455
Lys Gly Lys Phe Gly Val Tyr Thr Lys Val Ser Asn Phe Leu Lys Trp
445                 450                 455                 460

ATC GAC AAG ATC ATG AAG GCC AGG GCA GGG GCC GCG GGC AGC CGC GGC   1503
Ile Asp Lys Ile Met Lys Ala Arg Ala Gly Ala Ala Gly Ser Arg Gly
                465                 470                 475

CAC AGT GAA GCC CCT GCC ACC TGG ACG GTC CCG CCG CCC CTC CCC CTC   1551
His Ser Glu Ala Pro Ala Thr Trp Thr Val Pro Pro Pro Leu Pro Leu
            480                 485                 490

TAA                                                               1554
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 2:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 492 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 2:

```
Met Ala Gly Leu Leu His Leu Val Leu Leu Ser Thr Ala Leu Gly Gly
 1               5                  10                  15

Leu Leu Arg Pro Ala Gly Ser Val Phe Leu Pro Arg Asp Gln Ala His
            20                  25                  30

Arg Val Leu Gln Arg Ala Arg Arg Ala Asn Ser Phe Leu Glu Glu Val
        35                  40                  45

Lys Gln Gly Asn Leu Glu Arg Glu Cys Leu Glu Glu Ala Cys Ser Leu
    50                  55                  60

Glu Glu Ala Arg Glu Val Phe Glu Asp Ala Glu Gln Thr Asp Glu Phe
65                  70                  75                  80

Trp Ser Lys Tyr Lys Asp Gly Asp Gln Cys Glu Gly His Pro Cys Leu
                85                  90                  95

Asn Gln Gly His Cys Lys Asp Gly Ile Gly Asp Tyr Thr Cys Thr Cys
            100                 105                 110

Ala Glu Gly Phe Glu Gly Lys Asn Cys Glu Phe Ser Thr Arg Glu Ile
        115                 120                 125

Cys Ser Leu Asp Asn Gly Gly Cys Asp Gln Phe Cys Arg Glu Glu Arg
    130                 135                 140

Ser Glu Val Arg Cys Ser Cys Ala His Gly Tyr Val Leu Gly Asp Asp
145                 150                 155                 160

Ser Lys Ser Cys Val Ser Thr Glu Arg Phe Pro Cys Gly Lys Phe Thr
                165                 170                 175
```

```
Gln Gly Arg Ser Arg Arg Trp Ala Ile His Thr Ser Glu Asp Ala Leu
            180                 185                 190

Asp Ala Ser Glu Leu Glu His Tyr Asp Pro Ala Asp Leu Ser Pro Thr
        195                 200                 205

Glu Ser Ser Leu Asp Leu Leu Gly Leu Asn Arg Thr Glu Pro Ser Ala
    210                 215                 220

Gly Glu Asp Gly Ser Gln Val Val Arg Ile Val Gly Gly Arg Asp Cys
225                 230                 235                 240

Ala Glu Gly Glu Cys Pro Trp Gln Ala Leu Leu Val Asn Glu Glu Asn
                245                 250                 255

   Glu Gly Phe Cys Gly Gly Thr Ile Leu Asn Glu Phe Tyr Val Leu Thr
            260                 265                 270

   Ala Ala His Cys Leu His Gln Ala Lys Arg Phe Thr Val Arg Val Gly
        275                 280                 285

   Asp Arg Asn Thr Glu Gln Glu Glu Gly Asn Glu Met Ala His Glu Val
    290                 295                 300

   Glu Met Thr Val Lys His Ser Arg Phe Val Lys Glu Thr Tyr Asp Phe
305                 310                 315                 320

   Asp Ile Ala Val Leu Arg Leu Lys Thr Pro Ile Arg Phe Arg Arg Asn
                325                 330                 335

   Val Ala Pro Ala Cys Leu Pro Glu Lys Asp Trp Ala Glu Ala Thr Leu
            340                 345                 350

   Met Thr Gln Lys Thr Gly Ile Val Ser Gly Phe Gly Arg Thr His Glu
        355                 360                 365

   Lys Gly Arg Leu Ser Ser Thr Leu Lys Met Leu Glu Val Pro Tyr Val
    370                 375                 380

   Asp Arg Ser Thr Cys Lys Leu Ser Ser Ser Phe Thr Ile Thr Pro Asn
385                 390                 395                 400

   Met Phe Cys Ala Gly Tyr Asp Thr Gln Pro Glu Asp Ala Cys Gln Gly
                405                 410                 415

   Asp Ser Gly Gly Pro His Val Thr Arg Phe Lys Asp Thr Tyr Phe Val
            420                 425                 430

   Thr Gly Ile Val Ser Trp Gly Glu Gly Cys Ala Arg Lys Gly Lys Phe
        435                 440                 445

   Gly Val Tyr Thr Lys Val Ser Asn Phe Leu Lys Trp Ile Asp Lys Ile
    450                 455                 460

   Met Lys Ala Arg Ala Gly Ala Ala Gly Ser Arg Gly His Ser Glu Ala
465                 470                 475                 480

   Pro Ala Thr Trp Thr Val Pro Pro Pro Leu Pro Leu
                485                 490
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 3:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 42 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 3:

CGTCCTGGAT CCATCGAGGG TAGAATCCAG CGTACTCCAA AG 42

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 4:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 22 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 4:

GCGAAGCTTG ATCACATGTC TCG 23

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 5:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 44 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 5:

CGTCCTGGAT CCATCGAGGG TAGAATCCAG AAAACCCCTC AAAT 44

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 6:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 23 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 6:

GCGAAGCTTA CATGTCTCGA TC 22

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 7:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 40 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 7:

CCTGGATCCA TCGAGGGTAG GTTCCCAACC ATTCCCTTAT 40

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 8:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 26 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 8:

CCGAAGCTTA GAAGCCACAG CTGCCC 26

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 9:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 39 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 9:

CGTCCTGGAT CCATCGAGGG TAGGTACTCG CGGGAGAAG 39

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 10:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 26 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 10:

CGACCGAAGC TTCAGAGTTC GTTGTG 26

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 11:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 42 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 11:

CGTCCTGGAT CCATCGAGGG TAGGGCTATC GACGCCCCTA AG 42

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 12:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 30 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 12:

CGACCGAAGC TTATCGGCAG TGGGGCCCCT 30

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 13:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 29 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 13:

CGACCGAAGC TTAGGCCTTG CAGGAGCGG 29

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 14:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 32 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 14:

CGACCGAAGC TTACTTCTTG CATGACTTCC CG 32

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 15:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 42 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 15:

CGTCCTGGAT CCATCGAGGG TAGGGGCACC AACAAATGCC GG 42

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 16:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 29 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 16:

CGACCGAAGC TTAGTCCAGG CTGCGGCAG 29

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 17:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 41 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 17:

CGTCCTGGAT CCATCGAGGG TAGGGTGCCT CCACCCCAGT G 41

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 18:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 29 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 18:

CGACCGAAGC TTACTGGTCG CAGAGCTCG 29

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 19:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 46 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 19:

CCTTGATCAA TCGAGGGTAG GGGTGGTCAG TGCTCTCTGA ATAACG 46

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 20:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 29 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 20:

CGCAAGCTTA CTTAAACTCA TAGCAGGTG 29

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 21:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 44 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 21:

CGTCCTGGAT CCATCGAGGG TAGGGCGGTG AATTCCTCTT GCCG 44

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 22:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 30 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 22:

CGACCGAAGC TTAGATGTGG CAGCCACGCT 30

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 23:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 42 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 23:

CGTCCTGGAT CCATCGAGGG TAGGGTGTCC AACTGCACGG CT 42

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 24:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 30 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 24:

CGACCGAAGC TTAGATGCTG CAGTCCTCCT 31

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 25:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 47 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 25:

CGTCCTGGAT CCATCGAGGG TAGGAGTAAA TACAAAGATG GAGACCA 47

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 26:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 30 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 26:

CGACCGAAGC TTACCAGGTG GCAGGGGCTT 30

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 27:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 46 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 27:

CTGCCTGGAT CCATCGAGGG TAGGAAAGTG TATCTCTCAT CAGAGTGCAA GACTGGGAAT GG 62

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 28:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 33 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 28:

CGACCGAAGC TTATTCACAC TCAAGAATGT CGC 33

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 29:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 41 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 29:

CTGCCTGGAT CCATCGAGGG TAGGGTCCAG GACTGCTACC AT 42

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 30:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 31 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 30:

CGACCGAAGC TTACGCTTCT GTTCCTGAGC A 31

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 31:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 40 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 31:

CCTGGATCCA TCGAGGGTAG GGTCTACCTC CAGACATCCT 40

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 32:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 26 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 32:

CCGAAGCTTC AAGCATTTCC AAGATC 26

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 33:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 39 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 33:

CCTGGATCCA TCGAGGGTAG GGGCGAGCCA CCAACCCAG 39

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 34:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 25 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 34:

CCGAAGCTTA CACGATCCCG AACTG 25

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 35:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 38 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 35:

CCGAGATCTA TCGAGGGTAG GCAGGTCAAA CTGCAGCA 38

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 36:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 29 Basenpaare

(B) TYP: Nukleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (synthetisch)

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 36:

GCCAAGCTTA ATTCAGATCC TCTTCTGAG 29

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 37:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 6 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 37:

```
Gly Ser Ile Glu Gly Arg
1               5
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 38:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 4 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 38:

```
Ile Glu Gly Arg
1
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 39:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 4 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 39:

```
Tyr Trp Thr Asp
1
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 40:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 4 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 40:

```
Ile Gln Gly Arg
1
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 41:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 4 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 41:

```
Ala Glu Gly Arg
1
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 42:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 4 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 42:

```
Ala Gln Gly Arg
1
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 43:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 4 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 43:

```
Ile Cys Gly Arg
1
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 44:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 4 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 44:

```
Ala Cys Gly Arg
1
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 45:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 4 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 45:

```
Ile Met Gly Arg
1
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 46:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 4 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 46:

```
Ala Met Gly Arg
1
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 47:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 6 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 47:

```
His His His His His His
1               5
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 48:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 15 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 48:

```
Met Gly Ser His His His His His His Gly Ser Ile Glu Gly Arg
1               5                   10                  15
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 49:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 119 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 49:

```
Met Ser Arg Ser Val Ala Leu Ala Val Leu Ala Leu Leu Ser Leu Ser
1               5                  10                  15

Gly Leu Glu Ala Ile Gln Arg Thr Pro Lys Ile Gln Val Tyr Ser Arg
            20                  25                  30

His Pro Ala Glu Asn Gly Lys Ser Asn Phe Leu Asn Cys Tyr Val Ser
        35                  40                  45

Gly Phe His Pro Ser Asp Ile Glu Val Asp Leu Leu Lys Asn Gly Glu
    50                  55                  60

Arg Ile Glu Lys Val Glu His Ser Asp Leu Ser Phe Ser Lys Asp Trp
65                  70                  75                  80

Ser Phe Tyr Leu Leu Tyr Tyr Thr Glu Phe Thr Pro Thr Glu Lys Asp
                85                  90                  95

Glu Tyr Ala Cys Arg Val Asn His Val Thr Leu Ser Gln Pro Lys Ile
            100                 105                 110

Val Lys Trp Asp Arg Asp Met
        115
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 50:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 119 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 50:

```
Met Ala Arg Ser Val Thr Leu Val Phe Leu Val Leu Val Ser Leu Thr
1               5                  10                  15

Gly Leu Tyr Ala Ile Gln Lys Thr Pro Gln Ile Gln Val Tyr Ser Arg
            20                  25                  30

His Pro Pro Glu Asn Gly Lys Pro Asn Ile Leu Asn Cys Tyr Val Thr
        35                  40                  45

Gln Phe His Pro Pro His Ile Glu Ile Gln Met Leu Lys Asn Gly Lys
    50                  55                  60

Lys Ile Pro Lys Val Glu Met Ser Asp Met Ser Phe Ser Lys Asp Trp
65                  70                  75                  80

Ser Phe Tyr Ile Leu Ala His Thr Glu Phe Thr Pro Thr Glu Thr Asp
                85                  90                  95

Thr Tyr Ala Cys Arg Val Lys His Asp Ser Met Ala Glu Pro Lys Thr
            100                 105                 110

Val Tyr Trp Asp Arg Asp Met
        115
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 51:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 217 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 51:

```
Met Ala Thr Gly Ser Arg Thr Ser Leu Leu Leu Ala Phe Gly Leu Leu
1               5                  10                  15

Cys Leu Pro Trp Leu Gln Glu Gly Ser Ala Phe Pro Thr Ile Pro Leu
            20                  25                  30

Ser Arg Leu Phe Asp Asn Ala Ser Leu Arg Ala His Arg Leu His Gln
        35                  40                  45

Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys
    50                  55                  60

Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe
65                  70                  75                  80

Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys
                85                  90                  95

Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu Leu Leu Ile Gln Ser Trp
            100                 105                 110

Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val
        115                 120                 125

Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu
    130                 135                 140

Glu Gly Ile Gln Thr Leu Met Gly Arg Leu Glu Asp Gly Ser Pro Arg
145                 150                 155                 160

Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser
                165                 170                 175
```

```
His Asn Asp Asp Ala Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe
            180                 185                 190

Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys
        195                 200                 205

Arg Ser Val Glu Gly Ser Cys Gly Phe
    210                 215
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 52:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 4544 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 52:

```
Met Leu Thr Pro Pro Leu Leu Leu Leu Pro Leu Leu Ser Ala Leu
1               5                  10                  15

Val Ala Ala Ala Ile Asp Ala Pro Lys Thr Cys Ser Pro Lys Gln Phe
            20                  25                  30

Ala Cys Arg Asp Gln Ile Thr Cys Ile Ser Lys Gly Trp Arg Cys Asp
        35                  40                  45

Gly Glu Arg Asp Cys Pro Asp Gly Ser Asp Glu Ala Pro Glu Ile Cys
    50                  55                  60

Pro Gln Ser Lys Ala Gln Arg Cys Gln Pro Asn Glu His Asn Cys Leu
65                  70                  75                  80

Gly Thr Glu Leu Cys Val Pro Met Ser Arg Leu Cys Asn Gly Val Gln
                85                  90                  95

Asp Cys Met Asp Gly Ser Asp Glu Gly Pro His Cys Arg Glu Leu Gln
            100                 105                 110

Gly Asn Cys Ser Arg Leu Gly Cys Gln His His Cys Val Pro Thr Leu
        115                 120                 125

Asp Gly Pro Thr Cys Tyr Cys Asn Ser Ser Phe Gln Leu Gln Ala Asp
    130                 135                 140

Gly Lys Thr Cys Lys Asp Phe Asp Glu Cys Ser Val Tyr Gly Thr Cys
145                 150                 155                 160

Ser Gln Leu Cys Thr Asn Thr Asp Gly Ser Phe Ile Cys Gly Cys Val
                165                 170                 175

Glu Gly Tyr Leu Leu Gln Pro Asp Asn Arg Ser Cys Lys Ala Lys Asn
            180                 185                 190
```

```
Glu Pro Val Asp Arg Pro Pro Val Leu Leu Ile Ala Asn Ser Gln Asn
        195                 200                 205

Ile Leu Ala Thr Tyr Leu Ser Gly Ala Gln Val Ser Thr Ile Thr Pro
    210                 215                 220

Thr Ser Thr Arg Gln Thr Thr Ala Met Asp Phe Ser Tyr Ala Asn Glu
225                 230                 235                 240

Thr Val Cys Trp Val His Val Gly Asp Ser Ala Ala Gln Thr Gln Leu
            245                 250                 255

Lys Cys Ala Arg Met Pro Gly Leu Lys Gly Phe Val Asp Glu His Thr
            260                 265                 270

Ile Asn Ile Ser Leu Ser Leu His His Val Glu Gln Met Ala Ile Asp
        275                 280                 285

Trp Leu Thr Gly Asn Phe Tyr Phe Val Asp Asp Ile Asp Asp Arg Ile
    290                 295                 300

Phe Val Cys Asn Arg Asn Gly Asp Thr Cys Val Thr Leu Leu Asp Leu
305                 310                 315                 320

Glu Leu Tyr Asn Pro Lys Gly Ile Ala Leu Asp Pro Ala Met Gly Lys
                325                 330                 335

Val Phe Phe Thr Asp Tyr Gly Gln Ile Pro Lys Val Glu Arg Cys Asp
            340                 345                 350

Met Asp Gly Gln Asn Arg Thr Lys Leu Val Asp Ser Lys Ile Val Phe
        355                 360                 365

Pro His Gly Ile Thr Leu Asp Leu Val Ser Arg Leu Val Tyr Trp Ala
    370                 375                 380

Asp Ala Tyr Leu Asp Tyr Ile Glu Val Val Asp Tyr Glu Gly Lys Gly
385                 390                 395                 400

Arg Gln Thr Ile Ile Gln Gly Ile Leu Ile Glu His Leu Tyr Gly Leu
                405                 410                 415

Thr Val Phe Glu Asn Tyr Leu Tyr Ala Thr Asn Ser Asp Asn Ala Asn
            420                 425                 430

Ala Gln Gln Lys Thr Ser Val Ile Arg Val Asn Arg Phe Asn Ser Thr
        435                 440                 445

Glu Tyr Gln Val Val Thr Arg Val Asp Lys Gly Gly Ala Leu His Ile
    450                 455                 460

Tyr His Gln Arg Arg Gln Pro Arg Val Arg Ser His Ala Cys Glu Asn
465                 470                 475                 480

Asp Gln Tyr Gly Lys Pro Gly Gly Cys Ser Asp Ile Cys Leu Leu Ala
                485                 490                 495
```

```
Asn Ser His Lys Ala Arg Thr Cys Arg Cys Arg Ser Gly Phe Ser Leu
            500                 505                 510

Gly Ser Asp Gly Lys Ser Cys Lys Lys Pro Glu His Glu Leu Phe Leu
        515                 520                 525

Val Tyr Gly Lys Gly Arg Pro Gly Ile Ile Arg Gly Met Asp Met Gly
    530                 535                 540

Ala Lys Val Pro Asp Glu His Met Ile Pro Ile Glu Asn Leu Met Asn
545                 550                 555                 560

Pro Arg Ala Leu Asp Phe His Ala Glu Thr Gly Phe Ile Tyr Phe Ala
                565                 570                 575

Asp Thr Thr Ser Tyr Leu Ile Gly Arg Gln Lys Ile Asp Gly Thr Glu
            580                 585                 590

Arg Glu Thr Ile Leu Lys Asp Gly Ile His Asn Val Glu Gly Val Ala
        595                 600                 605

Val Asp Trp Met Gly Asp Asn Leu Tyr Trp Thr Asp Asp Gly Pro Lys
    610                 615                 620

Lys Thr Ile Ser Val Ala Arg Leu Glu Lys Ala Ala Gln Thr Arg Lys
625                 630                 635                 640

Thr Leu Ile Glu Gly Lys Met Thr His Pro Arg Ala Ile Val Val Asp
                645                 650                 655

Pro Leu Asn Gly Trp Met Tyr Trp Thr Asp Trp Glu Glu Asp Pro Lys
            660                 665                 670

Asp Ser Arg Arg Gly Arg Leu Glu Arg Ala Trp Met Asp Gly Ser His
        675                 680                 685

Arg Asp Ile Phe Val Thr Ser Lys Thr Val Leu Trp Pro Asn Gly Leu
    690                 695                 700

Ser Leu Asp Ile Pro Ala Gly Arg Leu Tyr Trp Val Asp Ala Phe Tyr
705                 710                 715                 720

Asp Arg Ile Glu Thr Ile Leu Leu Asn Gly Thr Asp Arg Lys Ile Val
                725                 730                 735

Tyr Glu Gly Pro Glu Leu Asn His Ala Phe Gly Leu Cys His His Gly
            740                 745                 750

Asn Tyr Leu Phe Trp Thr Glu Tyr Arg Ser Gly Ser Val Tyr Arg Leu
        755                 760                 765

Glu Arg Gly Val Gly Gly Ala Pro Pro Thr Val Thr Leu Leu Arg Ser
    770                 775                 780

Glu Arg Pro Pro Ile Phe Glu Ile Arg Met Tyr Asp Ala Gln Gln Gln
785                 790                 795                 800
```

Gln Val Gly Thr Asn Lys Cys Arg Val Asn Asn Gly Gly Cys Ser Ser
805 810 815

Leu Cys Leu Ala Thr Pro Gly Ser Arg Gln Cys Ala Cys Ala Glu Asp
820 825 830

Gln Val Leu Asp Ala Asp Gly Val Thr Cys Leu Ala Asn Pro Ser Tyr
835 840 845

Val Pro Pro Pro Gln Cys Gln Pro Gly Glu Phe Ala Cys Ala Asn Ser
850 855 860

Arg Cys Ile Gln Glu Arg Trp Lys Cys Asp Gly Asp Asn Asp Cys Leu
865 870 875 880

Asp Asn Ser Asp Glu Ala Pro Ala Leu Cys His Gln His Thr Cys Pro
885 890 895

Ser Asp Arg Phe Lys Cys Glu Asn Asn Arg Cys Ile Pro Asn Arg Trp
900 905 910

Leu Cys Asp Gly Asp Asn Asp Cys Gly Asn Ser Glu Asp Glu Ser Asn
915 920 925

Ala Thr Cys Ser Ala Arg Thr Cys Pro Pro Asn Gln Phe Ser Cys Ala
930 935 940

Ser Gly Arg Cys Ile Pro Ile Ser Trp Thr Cys Asp Leu Asp Asp Asp
945 950 955 960

Cys Gly Asp Arg Ser Asp Glu Ser Ala Ser Cys Ala Tyr Pro Thr Cys
965 970 975

Phe Pro Leu Thr Gln Phe Thr Cys Asn Asn Gly Arg Cys Ile Asn Ile
980 985 990

Asn Trp Arg Cys Asp Asn Asp Asn Asp Cys Gly Asp Asn Ser Asp Glu
995 1000 1005

Ala Gly Cys Ser His Ser Cys Ser Ser Thr Gln Phe Lys Cys Asn Ser
1010 1015 1020

Gly Arg Cys Ile Pro Glu His Trp Thr Cys Asp Gly Asp Asn Asp Cys
1025 1030 1035 1040

Gly Asp Tyr Ser Asp Glu Thr His Ala Asn Cys Thr Asn Gln Ala Thr
1045 1050 1055

Arg Pro Pro Gly Gly Cys His Thr Asp Glu Phe Gln Cys Arg Leu Asp
1060 1065 1070

Gly Leu Cys Ile Pro Leu Arg Trp Arg Cys Asp Gly Asp Thr Asp Cys
1075 1080 1085

Met Asp Ser Ser Asp Glu Lys Ser Cys Glu Gly Val Thr His Val Cys
1090 1095 1100

Asp Pro Ser Val Lys Phe Gly Cys Lys Asp Ser Ala Arg Cys Ile Ser
1105 1110 1115 1120

Lys Ala Trp Val Cys Asp Gly Asp Asn Asp Cys Glu Asp Asn Ser Asp
1125 1130 1135

Glu Glu Asn Cys Glu Ser Leu Ala Cys Arg Pro Pro Ser His Pro Cys
1140 1145 1150

Ala Asn Asn Thr Ser Val Cys Leu Pro Pro Asp Lys Leu Cys Asp Gly
1155 1160 1165

Asn Asp Asp Cys Gly Asp Gly Ser Asp Glu Gly Glu Leu Cys Asp Gln
1170 1175 1180

Cys Ser Leu Asn Asn Gly Gly Cys Ser His Asn Cys Ser Val Ala Pro
1185 1190 1195 1200

Gly Glu Gly Ile Val Cys Ser Cys Pro Leu Gly Met Glu Leu Gly Pro
1205 1210 1215

Asp Asn His Thr Cys Gln Ile Gln Ser Tyr Cys Ala Lys His Leu Lys
1220 1225 1230

Cys Ser Gln Lys Cys Asp Gln Asn Lys Phe Ser Val Lys Cys Ser Cys
1235 1240 1245

Tyr Glu Gly Trp Val Leu Glu Pro Asp Gly Glu Ser Cys Arg Ser Leu
1250 1255 1260

Asp Pro Phe Lys Pro Phe Ile Ile Phe Ser Asn Arg His Glu Ile Arg
1265 1270 1275 1280

Arg Ile Asp Leu His Lys Gly Asp Tyr Ser Val Leu Val Pro Gly Leu
1285 1290 1295

Arg Asn Thr Ile Ala Leu Asp Phe His Leu Ser Gln Ser Ala Leu Tyr
1300 1305 1310

Trp Thr Asp Val Val Glu Asp Lys Ile Tyr Arg Gly Lys Leu Leu Asp
1315 1320 1325

Asn Gly Ala Leu Thr Ser Phe Glu Val Val Ile Gln Tyr Gly Leu Ala
1330 1335 1340

Thr Pro Glu Gly Leu Ala Val Asp Trp Ile Ala Gly Asn Ile Tyr Trp
1345 1350 1355 1360

Val Glu Ser Asn Leu Asp Gln Ile Glu Val Ala Lys Leu Asp Gly Thr
1365 1370 1375

Leu Arg Thr Thr Leu Leu Ala Gly Asp Ile Glu His Pro Arg Ala Ile
1380 1385 1390

Ala Leu Asp Pro Arg Asp Gly Ile Leu Phe Trp Thr Asp Trp Asp Ala
1395 1400 1405

Ser Leu Pro Arg Ile Glu Ala Ala Ser Met Ser Gly Ala Gly Arg Arg
1410 1415 1420

Thr Val His Arg Glu Thr Gly Ser Gly Gly Trp Pro Asn Gly Leu Thr
1425 1430 1435 1440

Val Asp Tyr Leu Glu Lys Arg Ile Leu Trp Ile Asp Ala Arg Ser Asp
1445 1450 1455

Ala Ile Tyr Ser Ala Arg Tyr Asp Gly Ser Gly His Met Glu Val Leu
1460 1465 1470

Arg Gly His Glu Phe Leu Ser His Pro Phe Ala Val Thr Leu Tyr Gly
1475 1480 1485

Gly Glu Val Tyr Trp Thr Asp Trp Arg Thr Asn Thr Leu Ala Lys Ala
1490 1495 1500

Asn Lys Trp Thr Gly His Asn Val Thr Val Val Gln Arg Thr Asn Thr
1505 1510 1515 1520

Gln Pro Phe Asp Leu Gln Val Tyr His Pro Ser Arg Gln Pro Met Ala
1525 1530 1535

Pro Asn Pro Cys Glu Ala Asn Gly Gly Gln Gly Pro Cys Ser His Leu
1540 1545 1550

Cys Leu Ile Asn Tyr Asn Arg Thr Val Ser Cys Ala Cys Pro His Leu
1555 1560 1565

Met Lys Leu His Lys Asp Asn Thr Thr Cys Tyr Glu Phe Lys Lys Phe
1570 1575 1580

Leu Leu Tyr Ala Arg Gln Met Glu Ile Arg Gly Val Asp Leu Asp Ala
1585 1590 1595 1600

Pro Tyr Tyr Asn Tyr Ile Ile Ser Phe Thr Val Pro Asp Ile Asp Asn
1605 1610 1615

Val Thr Val Leu Asp Tyr Asp Ala Arg Glu Gln Arg Val Tyr Trp Ser
1620 1625 1630

Asp Val Arg Thr Gln Ala Ile Lys Arg Ala Phe Ile Asn Gly Thr Gly
1635 1640 1645

Val Glu Thr Val Val Ser Ala Asp Leu Pro Asn Ala His Gly Leu Ala
1650 1655 1660

Val Asp Trp Val Ser Arg Asn Leu Phe Trp Thr Ser Tyr Asp Thr Asn
1665 1670 1675 1680

Lys Lys Gln Ile Asn Val Ala Arg Leu Asp Gly Ser Phe Lys Asn Ala
1685 1690 1695

Val Val Gln Gly Leu Glu Gln Pro His Gly Leu Val Val His Pro Leu
1700 1705 1710

Arg Gly Lys Leu Tyr Trp Thr Asp Gly Asp Asn Ile Ser Met Ala Asn
1715 1720 1725

Met Asp Gly Ser Asn Arg Thr Leu Leu Phe Ser Gly Gln Lys Gly Pro
1730 1735 1740

Val Gly Leu Ala Ile Asp Phe Pro Glu Ser Lys Leu Tyr Trp Ile Ser
1745 1750 1755 1760

Ser Gly Asn His Thr Ile Asn Arg Cys Asn Leu Asp Gly Ser Gly Leu
1765 1770 1775

Glu Val Ile Asp Ala Met Arg Ser Gln Leu Gly Lys Ala Thr Ala Leu
1780 1785 1790

Ala Ile Met Gly Asp Lys Leu Trp Trp Ala Asp Gln Val Ser Glu Lys
1795 1800 1805

Met Gly Thr Cys Ser Lys Ala Asp Gly Ser Gly Ser Val Val Leu Arg
1810 1815 1820

Asn Ser Thr Thr Leu Val Met His Met Lys Val Tyr Asp Glu Ser Ile
1825 1830 1835 1840

Gln Leu Asp His Lys Gly Thr Asn Pro Cys Ser Val Asn Asn Gly Asp
1845 1850 1855

Cys Ser Gln Leu Cys Leu Pro Thr Ser Glu Thr Thr Arg Ser Cys Met
1860 1865 1870

Cys Thr Ala Gly Tyr Ser Leu Arg Ser Gly Gln Gln Ala Cys Glu Gly
1875 1880 1885

Val Gly Ser Phe Leu Leu Tyr Ser Val His Glu Gly Ile Arg Gly Ile
1890 1895 1900

Pro Leu Asp Pro Asn Asp Lys Ser Asp Ala Leu Val Pro Val Ser Gly
1905 1910 1915 1920

Thr Ser Leu Ala Val Gly Ile Asp Phe His Ala Glu Asn Asp Thr Ile
1925 1930 1935

Tyr Trp Val Asp Met Gly Leu Ser Thr Ile Ser Arg Ala Lys Arg Asp
1940 1945 1950

Gln Thr Trp Arg Glu Asp Val Val Thr Asn Gly Ile Gly Arg Val Glu
1955 1960 1965

Gly Ile Ala Val Asp Trp Ile Ala Gly Asn Ile Tyr Trp Thr Asp Gln
1970 1975 1980

Gly Phe Asp Val Ile Glu Val Ala Arg Leu Asn Gly Ser Phe Arg Tyr
1985 1990 1995 2000

Val Val Ile Ser Gln Gly Leu Asp Lys Pro Arg Ala Ile Thr Val His
2005 2010 2015

Pro Glu Lys Gly Tyr Leu Phe Trp Thr Glu Trp Gly Gln Tyr Pro Arg
2020 2025 2030

Ile Glu Arg Ser Arg Leu Asp Gly Thr Glu Arg Val Val Leu Val Asn
2035 2040 2045

Val Ser Ile Ser Trp Pro Asn Gly Ile Ser Val Asp Tyr Gln Asp Gly
2050 2055 2060

Lys Leu Tyr Trp Cys Asp Ala Arg Thr Asp Lys Ile Glu Arg Ile Asp
2065 2070 2075 2080

Leu Glu Thr Gly Glu Asn Arg Glu Val Val Leu Ser Ser Asn Asn Met
2085 2090 2095

Asp Met Phe Ser Val Ser Val Phe Glu Asp Phe Ile Tyr Trp Ser Asp
2100 2105 2110

Arg Thr His Ala Asn Gly Ser Ile Lys Arg Gly Ser Lys Asp Asn Ala
2115 2120 2125

Thr Asp Ser Val Pro Leu Arg Thr Gly Ile Gly Val Gln Leu Lys Asp
2130 2135 2140

Ile Lys Val Phe Asn Arg Asp Arg Gln Lys Gly Thr Asn Val Cys Ala
2145 2150 2155 2160

Val Ala Asn Gly Gly Cys Gln Gln Leu Cys Leu Tyr Arg Gly Arg Gly
2165 2170 2175

Gln Arg Ala Cys Ala Cys Ala His Gly Met Leu Ala Glu Asp Gly Ala
2180 2185 2190

Ser Cys Arg Glu Tyr Ala Gly Tyr Leu Leu Tyr Ser Glu Arg Thr Ile
2195 2200 2205

Leu Lys Ser Ile His Leu Ser Asp Glu Arg Asn Leu Asn Ala Pro Val
2210 2215 2220

Gln Pro Phe Glu Asp Pro Glu His Met Lys Asn Val Ile Ala Leu Ala
2225 2230 2235 2240

Phe Asp Tyr Arg Ala Gly Thr Ser Pro Gly Thr Pro Asn Arg Ile Phe
2245 2250 2255

Phe Ser Asp Ile His Phe Gly Asn Ile Gln Gln Ile Asn Asp Asp Gly
2260 2265 2270

Ser Arg Arg Ile Thr Ile Val Glu Asn Val Gly Ser Val Glu Gly Leu
2275 2280 2285

Ala Tyr His Arg Gly Trp Asp Thr Leu Tyr Trp Thr Ser Tyr Thr Thr
2290 2295 2300

Ser Thr Ile Thr Arg His Thr Val Asp Gln Thr Arg Pro Gly Ala Phe
2305 2310 2315 2320

Glu Arg Glu Thr Val Ile Thr Met Ser Gly Asp Asp His Pro Arg Ala
2325 2330 2335

Phe Val Leu Asp Glu Cys Gln Asn Leu Met Phe Trp Thr Asn Trp Asn
2340 2345 2350

Glu Gln His Pro Ser Ile Met Arg Ala Ala Leu Ser Gly Ala Asn Val
2355 2360 2365

Leu Thr Leu Ile Glu Lys Asp Ile Arg Thr Pro Asn Gly Leu Ala Ile
2370 2375 2380

Asp His Arg Ala Glu Lys Leu Tyr Phe Ser Asp Ala Thr Leu Asp Lys
2385 2390 2395 2400

Ile Glu Arg Cys Glu Tyr Asp Gly Ser His Arg Tyr Val Ile Leu Lys
2405 2410 2415

Ser Glu Pro Val His Pro Phe Gly Leu Ala Val Tyr Gly Glu His Ile
2420 2425 2430

Phe Trp Thr Asp Trp Val Arg Arg Ala Val Gln Arg Ala Asn Lys His
2435 2440 2445

Val Gly Ser Asn Met Lys Leu Leu Arg Val Asp Ile Pro Gln Gln Pro
2450 2455 2460

Met Gly Ile Ile Ala Val Ala Asn Asp Thr Asn Ser Cys Glu Leu Ser
2465 2470 2475 2480

Pro Cys Arg Ile Asn Asn Gly Gly Cys Gln Asp Leu Cys Leu Leu Thr
2485 2490 2495

His Gln Gly His Val Asn Cys Ser Cys Arg Gly Gly Arg Ile Leu Gln
2500 2505 2510

Asp Asp Leu Thr Cys Arg Ala Val Asn Ser Ser Cys Arg Ala Gln Asp
2515 2520 2525

Glu Phe Glu Cys Ala Asn Gly Glu Cys Ile Asn Phe Ser Leu Thr Cys
2530 2535 2540

Asp Gly Val Pro His Cys Lys Asp Lys Ser Asp Glu Lys Pro Ser Tyr
2545 2550 2555 2560

Cys Asn Ser Arg Arg Cys Lys Lys Thr Phe Arg Gln Cys Ser Asn Gly
2565 2570 2575

Arg Cys Val Ser Asn Met Leu Trp Cys Asn Gly Ala Asp Asp Cys Gly
2580 2585 2590

Asp Gly Ser Asp Glu Ile Pro Cys Asn Lys Thr Ala Cys Gly Val Gly
2595 2600 2605

Glu Phe Arg Cys Arg Asp Gly Thr Cys Ile Gly Asn Ser Ser Arg Cys
2610 2615 2620

Asn Gln Phe Val Asp Cys Glu Asp Ala Ser Asp Glu Met Asn Cys Ser
2625 2630 2635 2640

Ala Thr Asp Cys Ser Ser Tyr Phe Arg Leu Gly Val Lys Gly Val Leu
2645 2650 2655

Phe Gln Pro Cys Glu Arg Thr Ser Leu Cys Tyr Ala Pro Ser Trp Val
2660 2665 2670

Cys Asp Gly Ala Asn Asp Cys Gly Asp Tyr Ser Asp Glu Arg Asp Cys
2675 2680 2685

Pro Gly Val Lys Arg Pro Arg Cys Pro Leu Asn Tyr Phe Ala Cys Pro
2690 2695 2700

Ser Gly Arg Cys Ile Pro Met Ser Trp Thr Cys Asp Lys Glu Asp Asp
2705 2710 2715 2720

Cys Glu His Gly Glu Asp Glu Thr His Cys Asn Lys Phe Cys Ser Glu
2725 2730 2735

Ala Gln Phe Glu Cys Gln Asn His Arg Cys Ile Ser Lys Gln Trp Leu
2740 2745 2750

Cys Asp Gly Ser Asp Asp Cys Gly Asp Gly Ser Asp Glu Ala Ala His
2755 2760 2765

Cys Glu Gly Lys Thr Cys Gly Pro Ser Ser Phe Ser Cys Pro Gly Thr
2770 2775 2780

His Val Cys Val Pro Glu Arg Trp Leu Cys Asp Gly Asp Lys Asp Cys
2785 2790 2795 2800

Ala Asp Gly Ala Asp Glu Ser Ile Ala Ala Gly Cys Leu Tyr Asn Ser
2805 2810 2815

Thr Cys Asp Asp Arg Glu Phe Met Cys Gln Asn Arg Gln Cys Ile Pro
2820 2825 2830

Lys His Phe Val Cys Asp His Asp Arg Asp Cys Ala Asp Gly Ser Asp
2835 2840 2845

Glu Ser Pro Glu Cys Glu Tyr Pro Thr Cys Gly Pro Ser Glu Phe Arg
2850 2855 2860

Cys Ala Asn Gly Arg Cys Leu Ser Ser Arg Gln Trp Glu Cys Asp Gly
2865 2870 2875 2880

Glu Asn Asp Cys His Asp Gln Ser Asp Glu Ala Pro Lys Asn Pro His
2885 2890 2895

Cys Thr Ser Pro Glu His Lys Cys Asn Ala Ser Ser Gln Phe Leu Cys
2900 2905 2910

Ser Ser Gly Arg Cys Val Ala Glu Ala Leu Leu Cys Asn Gly Gln Asp
2915 2920 2925

Asp Cys Gly Asp Ser Ser Asp Glu Arg Gly Cys His Ile Asn Glu Cys
2930 2935 2940

Leu Ser Arg Lys Leu Ser Gly Cys Ser Gln Asp Cys Glu Asp Leu Lys
2945 2950 2955 2960

Ile Gly Phe Lys Cys Arg Cys Arg Pro Gly Phe Arg Leu Lys Asp Asp
2965 2970 2975

Gly Arg Thr Cys Ala Asp Val Asp Glu Cys Ser Thr Thr Phe Pro Cys
2980 2985 2990

Ser Gln Arg Cys Ile Asn Thr His Gly Ser Tyr Lys Cys Leu Cys Val
2995 3000 3005

Glu Gly Tyr Ala Pro Arg Gly Gly Asp Pro His Ser Cys Lys Ala Val
3010 3015 3020

Thr Asp Glu Glu Pro Phe Leu Ile Phe Ala Asn Arg Tyr Tyr Leu Arg
3025 3030 3035 3040

Lys Leu Asn Leu Asp Gly Ser Asn Tyr Thr Leu Leu Lys Gln Gly Leu
3045 3050 3055

Asn Asn Ala Val Ala Leu Asp Phe Asp Tyr Arg Glu Gln Met Ile Tyr
3060 3065 3070

Trp Thr Asp Val Thr Thr Gln Gly Ser Met Ile Arg Arg Met His Leu
3075 3080 3085

Asn Gly Ser Asn Val Gln Val Leu His Arg Thr Gly Leu Ser Asn Pro
3090 3095 3100

Asp Gly Leu Ala Val Asp Trp Val Gly Gly Asn Leu Tyr Trp Cys Asp
3105 3110 3115 3120

Lys Gly Arg Asp Thr Ile Glu Val Ser Lys Leu Asn Gly Ala Tyr Arg
3125 3130 3135

Thr Val Leu Val Ser Ser Gly Leu Arg Glu Pro Arg Ala Leu Val Val
3140 3145 3150

Asp Val Gln Asn Gly Tyr Leu Tyr Trp Thr Asp Trp Gly Asp His Ser
3155 3160 3165

Leu Ile Gly Arg Ile Gly Met Asp Gly Ser Ser Arg Ser Val Ile Val
3170 3175 3180

Asp Thr Lys Ile Thr Trp Pro Asn Gly Leu Thr Leu Asp Tyr Val Thr
3185 3190 3195 3200

Glu Arg Ile Tyr Trp Ala Asp Ala Arg Glu Asp Tyr Ile Glu Phe Ala
3205 3210 3215

Ser Leu Asp Gly Ser Asn Arg His Val Val Leu Ser Gln Asp Ile Pro
3220 3225 3230

His Ile Phe Ala Leu Thr Leu Phe Glu Asp Tyr Val Tyr Trp Thr Asp
3235 3240 3245

Trp Glu Thr Lys Ser Ile Asn Arg Ala His Lys Thr Thr Gly Thr Asn
3250 3255 3260

Lys Thr Leu Leu Ile Ser Thr Leu His Arg Pro Met Asp Leu His Val
3265 3270 3275 3280

Phe His Ala Leu Arg Gln Pro Asp Val Pro Asn His Pro Cys Lys Val
3285 3290 3295

Asn Asn Gly Gly Cys Ser Asn Leu Cys Leu Leu Ser Pro Gly Gly Gly
3300 3305 3310

His Lys Cys Ala Cys Pro Thr Asn Phe Tyr Leu Gly Ser Asp Gly Arg
3315 3320 3325

Thr Cys Val Ser Asn Cys Thr Ala Ser Gln Phe Val Cys Lys Asn Asp
3330 3335 3340

Lys Cys Ile Pro Phe Trp Trp Lys Cys Asp Thr Glu Asp Asp Cys Gly
3345 3350 3355 3360

Asp His Ser Asp Glu Pro Pro Asp Cys Pro Glu Phe Lys Cys Arg Pro
3365 3370 3375

Gly Gln Phe Gln Cys Ser Thr Gly Ile Cys Thr Asn Pro Ala Phe Ile
3380 3385 3390

Cys Asp Gly Asp Asn Asp Cys Gln Asp Asn Ser Asp Glu Ala Asn Cys
3395 3400 3405

Asp Ile His Val Cys Leu Pro Ser Gln Phe Lys Cys Thr Asn Thr Asn
3410 3415 3420

Arg Cys Ile Pro Gly Ile Phe Arg Cys Asn Gly Gln Asp Asn Cys Gly
3425 3430 3435 3440

Asp Gly Glu Asp Glu Arg Asp Cys Pro Glu Val Thr Cys Ala Pro Asn
3445 3450 3455

Gln Phe Gln Cys Ser Ile Thr Lys Arg Cys Ile Pro Arg Val Trp Val
3460 3465 3470

Cys Asp Arg Asp Asn Asp Cys Val Asp Gly Ser Asp Glu Pro Ala Asn
3475 3480 3485

Cys Thr Gln Met Thr Cys Gly Val Asp Glu Phe Arg Cys Lys Asp Ser
3490 3495 3500

Gly Arg Cys Ile Pro Ala Arg Trp Lys Cys Asp Gly Glu Asp Asp Cys
3505 3510 3515 3520

Gly Asp Gly Ser Asp Glu Pro Lys Glu Glu Cys Asp Glu Arg Thr Cys
3525 3530 3535

Glu Pro Tyr Gln Phe Arg Cys Lys Asn Asn Arg Cys Val Pro Gly Arg
3540 3545 3550

Trp Gln Cys Asp Tyr Asp Asn Asp Cys Gly Asp Asn Ser Asp Glu Glu
3555 3560 3565

Ser Cys Thr Pro Arg Pro Cys Ser Glu Ser Glu Phe Ser Cys Ala Asn
3570 3575 3580

Gly Arg Cys Ile Ala Gly Arg Trp Lys Cys Asp Gly Asp His Asp Cys
3585 3590 3595 3600

Ala Asp Gly Ser Asp Glu Lys Asp Cys Thr Pro Arg Cys Asp Met Asp
3605 3610 3615

Gln Phe Gln Cys Lys Ser Gly His Cys Ile Pro Leu Arg Trp Arg Cys
3620 3625 3630

Asp Ala Asp Ala Asp Cys Met Asp Gly Ser Asp Glu Glu Ala Cys Gly
3635 3640 3645

Thr Gly Val Arg Thr Cys Pro Leu Asp Glu Phe Gln Cys Asn Asn Thr
3650 3655 3660

Leu Cys Lys Pro Leu Ala Trp Lys Cys Asp Gly Glu Asp Asp Cys Gly
3665 3670 3675 3680

Asp Asn Ser Asp Glu Asn Pro Glu Glu Cys Ala Arg Phe Val Cys Pro
3685 3690 3695

Pro Asn Arg Pro Phe Arg Cys Lys Asn Asp Arg Val Cys Leu Trp Ile
3700 3705 3710

Gly Arg Gln Cys Asp Gly Thr Asp Asn Cys Gly Asp Gly Thr Asp Glu
3715 3720 3725

Glu Asp Cys Glu Pro Pro Thr Ala His Thr Thr His Cys Lys Asp Lys
3730 3735 3740

Lys Glu Phe Leu Cys Arg Asn Gln Arg Cys Leu Ser Ser Ser Leu Arg
3745 3750 3755 3760

Cys Asn Met Phe Asp Asp Cys Gly Asp Gly Ser Asp Glu Glu Asp Cys
3765 3770 3775

Ser Ile Asp Pro Lys Leu Thr Ser Cys Ala Thr Asn Ala Ser Ile Cys
3780 3785 3790

Gly Asp Glu Ala Arg Cys Val Arg Thr Glu Lys Ala Ala Tyr Cys Ala
3795 3800 3805

Cys Arg Ser Gly Phe His Thr Val Pro Gly Gln Pro Gly Cys Gln Asp
3810 3815 3820

Ile Asn Glu Cys Leu Arg Phe Gly Thr Cys Ser Gln Leu Cys Asn Asn
3825 3830 3835 3840

Thr Lys Gly Gly His Leu Cys Ser Cys Ala Arg Asn Phe Met Lys Thr
3845 3850 3855

His Asn Thr Cys Lys Ala Glu Gly Ser Glu Tyr Gln Val Leu Tyr Ile
3860 3865 3870

Ala Asp Asp Asn Glu Ile Arg Ser Leu Phe Pro Gly His Pro His Ser
3875 3880 3885

Ala Tyr Glu Gln Ala Phe Gln Gly Asp Glu Ser Val Arg Ile Asp Ala
3890 3895 3900

Met Asp Val His Val Lys Ala Gly Arg Val Tyr Trp Thr Asn Trp His
3905 3910 3915 3920

Thr Gly Thr Ile Ser Tyr Arg Ser Leu Pro Pro Ala Ala Pro Pro Thr
3925 3930 3935

Thr Ser Asn Arg His Arg Arg Gln Ile Asp Arg Gly Val Thr His Leu
3940 3945 3950

Asn Ile Ser Gly Leu Lys Met Pro Arg Gly Ile Ala Ile Asp Trp Val
3955 3960 3965

Ala Gly Asn Val Tyr Trp Thr Asp Ser Gly Arg Asp Val Ile Glu Val
3970 3975 3980

Ala Gln Met Lys Gly Glu Asn Arg Lys Thr Leu Ile Ser Gly Met Ile
3985 3990 3995 4000

Asp Glu Pro His Ala Ile Val Val Asp Pro Leu Arg Gly Thr Met Tyr
4005 4010 4015

Trp Ser Asp Trp Gly Asn His Pro Lys Ile Glu Thr Ala Ala Met Asp
4020 4025 4030

Gly Thr Leu Arg Glu Thr Leu Val Gln Asp Asn Ile Gln Trp Pro Thr
4035 4040 4045

Gly Leu Ala Val Asp Tyr His Asn Glu Arg Leu Tyr Trp Ala Asp Ala
4050 4055 4060

Lys Leu Ser Val Ile Gly Ser Ile Arg Leu Asn Gly Thr Asp Pro Ile
4065 4070 4075 4080

Val Ala Ala Asp Ser Lys Arg Gly Leu Ser His Pro Phe Ser Ile Asp
4085 4090 4095

Val Phe Glu Asp Tyr Ile Tyr Gly Val Thr Tyr Ile Asn Asn Arg Val
4100 4105 4110

Phe Lys Ile His Lys Phe Gly His Ser Pro Leu Val Asn Leu Thr Gly
4115 4120 4125

Gly Leu Ser His Ala Ser Asp Val Val Leu Tyr His Gln His Lys Gln
4130 4135 4140

Pro Glu Val Thr Asn Pro Cys Asp Arg Lys Lys Cys Glu Trp Leu Cys
4145 4150 4155 4160

Leu Leu Ser Pro Ser Gly Pro Val Cys Thr Cys Pro Asn Gly Lys Arg
4165 4170 4175

Leu Asp Asn Gly Thr Cys Val Pro Val Pro Ser Pro Thr Pro Pro Pro
4180 4185 4190

Asp Ala Pro Arg Pro Gly Thr Cys Asn Leu Gln Cys Phe Asn Gly Gly
4195 4200 4205

Ser Cys Phe Leu Asn Ala Arg Arg Gln Pro Lys Cys Arg Cys Gln Pro
4210 4215 4220

Arg Tyr Thr Gly Asp Lys Cys Glu Leu Asp Gln Cys Trp Glu His Cys
4225 4230 4235 4240

Arg Asn Gly Gly Thr Cys Ala Ala Ser Pro Ser Gly Met Pro Thr Cys
4245 4250 4255

Arg Cys Pro Thr Gly Phe Thr Gly Pro Lys Cys Thr Gln Gln Val Cys
4260 4265 4270

Ala Gly Tyr Cys Ala Asn Asn Ser Thr Cys Thr Val Asn Gln Gly Asn
4275 4280 4285

Gln Pro Gln Cys Arg Cys Leu Pro Gly Phe Leu Gly Asp Arg Cys Gln
4290 4295 4300

Tyr Arg Gln Cys Ser Gly Tyr Cys Glu Asn Phe Gly Thr Cys Gln Met
4305 4310 4315 4320

Ala Ala Asp Gly Ser Arg Gln Cys Arg Cys Thr Ala Tyr Phe Glu Gly
4325 4330 4335

Ser Arg Cys Glu Val Asn Lys Cys Ser Arg Cys Leu Glu Gly Ala Cys
4340 4345 4350

Val Val Asn Lys Gln Ser Gly Asp Val Thr Cys Asn Cys Thr Asp Gly
4355 4360 4365

Arg Val Ala Pro Ser Cys Leu Thr Cys Val Gly His Cys Ser Asn Gly
4370 4375 4380

Gly Ser Cys Thr Met Asn Ser Lys Met Met Pro Glu Cys Gln Cys Pro
4385 4390 4395 4400

Pro His Met Thr Gly Pro Arg Cys Glu Glu His Val Phe Ser Gln Gln
4405 4410 4415

Gln Pro Gly His Ile Ala Ser Ile Leu Ile Pro Leu Leu Leu Leu Leu
4420 4425 4430

Leu Leu Val Leu Val Ala Gly Val Val Phe Trp Tyr Lys Arg Arg Val
4435 4440 4445

Gln Gly Ala Lys Gly Phe Gln His Gln Arg Met Thr Asn Gly Ala Met
4450 4455 4460

Asn Val Glu Ile Gly Asn Pro Thr Tyr Lys Met Tyr Glu Gly Gly Glu
4465 4470 4475 4480

Pro Asp Asp Val Gly Gly Leu Leu Asp Ala Asp Phe Ala Leu Asp Pro
4485 4490 4495

Asp Lys Pro Thr Asn Phe Thr Asn Pro Val Tyr Ala Thr Leu Tyr Met
4500 4505 4510

Gly Gly His Gly Ser Arg His Ser Leu Ala Ser Thr Asp Glu Lys Arg
4515 4520 4525

Glu Leu Leu Gly Arg Gly Pro Glu Asp Glu Ile Gly Asp Pro Leu Ala
4530 4535 4540

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 53:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 487 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 53:

```
Met Ala Gly Leu Leu His Leu Val Leu Leu Ser Thr Ala Leu Gly Gly
1               5                  10                  15

Leu Leu Arg Pro Ala Gly Ser Val Phe Leu Pro Arg Asp Gln Ala His
            20                  25                  30

Arg Val Leu Gln Arg Ala Arg Arg Ala Asn Ser Phe Leu Glu Glu Val
        35                  40                  45

Lys Gln Gly Asn Leu Glu Arg Glu Cys Leu Glu Glu Ala Cys Ser Leu
    50                  55                  60

Glu Glu Ala Arg Glu Val Phe Glu Asp Ala Glu Gln Thr Asp Glu Phe
65                  70                  75                  80

Trp Ser Lys Tyr Lys Asp Gly Asp Gln Cys Glu Gly His Pro Cys Leu
                85                  90                  95

Asn Gln Gly His Cys Lys Asp Gly Ile Gly Asp Tyr Thr Cys Thr Cys
            100                 105                 110

Ala Glu Gly Phe Glu Gly Lys Asn Cys Glu Phe Ser Thr Arg Glu Ile
        115                 120                 125

Cys Ser Leu Asp Asn Gly Gly Cys Asp Gln Phe Cys Arg Glu Glu Arg
    130                 135                 140
```

```
Ser Glu Val Arg Cys Ser Cys Ala His Gly Tyr Val Leu Gly Asp Asp
145                 150                 155                 160

Ser Lys Ser Cys Val Ser Thr Glu Arg Phe Pro Cys Gly Lys Phe Thr
                165                 170                 175

Gln Gly Arg Ser Arg Arg Trp Ala Ile His Thr Ser Glu Asp Ala Leu
            180                 185                 190

Asp Ala Ser Glu Leu Glu His Tyr Asp Pro Ala Asp Leu Ser Pro Thr
        195                 200                 205

Glu Ser Ser Leu Asp Leu Leu Gly Leu Asn Arg Thr Glu Pro Ser Ala
    210                 215                 220

Gly Glu Asp Gly Ser Gln Val Val Arg Ile Val Gly Gly Arg Asp Cys
225                 230                 235                 240

Ala Glu Gly Glu Cys Pro Trp Gln Ala Leu Leu Val Asn Glu Glu Asn
                245                 250                 255

Glu Gly Phe Cys Gly Gly Thr Ile Leu Asn Glu Phe Tyr Val Leu Thr
            260                 265                 270

Ala Ala His Cys Leu His Gln Ala Lys Arg Phe Thr Val Arg Val Gly
        275                 280                 285

Asp Arg Asn Thr Glu Gln Glu Glu Gly Asn Glu Met Ala His Glu Val
    290                 295                 300

Glu Met Thr Val Lys His Ser Arg Phe Val Lys Glu Thr Tyr Asp Phe
305                 310                 315                 320

Asp Ile Ala Val Leu Arg Leu Lys Thr Pro Ile Arg Phe Arg Arg Asn
                325                 330                 335

Val Ala Pro Ala Cys Leu Pro Glu Lys Asp Trp Ala Glu Ala Thr Leu
            340                 345                 350

Met Thr Gln Lys Thr Gly Ile Val Ser Gly Phe Gly Arg Thr His Glu
        355                 360                 365

Lys Gly Arg Leu Ser Ser Thr Leu Lys Met Leu Glu Val Pro Tyr Val
    370                 375                 380

Asp Arg Ser Thr Cys Lys Leu Ser Ser Ser Phe Thr Ile Thr Pro Asn
385                 390                 395                 400

Met Phe Cys Ala Gly Tyr Asp Thr Gln Pro Glu Asp Ala Cys Gln Gly
                405                 410                 415

Asp Ser Gly Gly Pro His Val Thr Arg Phe Lys Asp Thr Tyr Phe Val
            420                 425                 430

Thr Gly Ile Val Ser Trp Gly Glu Gly Cys Ala Arg Lys Gly Lys Phe
        435                 440                 445
```

```
Gly Val Tyr Thr Lys Val Ser Asn Phe Leu Lys Trp Ile Asp Lys Ile
    450                 455                 460

Met Lys Ala Arg Ala Gly Ala Ala Gly Ser Arg Gly His Ser Glu Ala
465                 470                 475                 480

Pro Ala Thr Trp Thr Val Pro
                485
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 54:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 790 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 54:

```
Glu Pro Leu Asp Asp Tyr Val Asn Thr Gln Gly Ala Ser Leu Phe Ser
1               5                   10                  15

Val Thr Lys Lys Gln Leu Gly Ala Gly Ser Ile Glu Glu Cys Ala Ala
            20                  25                  30

Lys Cys Glu Glu Asp Glu Glu Phe Thr Cys Arg Ala Phe Gln Tyr His
        35                  40                  45

Ser Lys Glu Gln Gln Cys Val Ile Met Ala Glu Asn Arg Lys Ser Ser
    50                  55                  60

Ile Ile Arg Met Arg Asp Val Val Leu Phe Glu Lys Lys Val Tyr Leu
65                  70                  75                  80
```

Ser Glu Cys Lys Thr Gly Asn Gly Lys Asn Tyr Arg Gly Thr Met Ser
85 90 95

Lys Thr Lys Asn Gly Ile Thr Cys Gln Lys Trp Ser Ser Thr Ser Pro
100 105 110

His Arg Pro Arg Phe Ser Pro Ala Thr His Pro Ser Glu Gly Leu Glu
115 120 125

Glu Asn Tyr Cys Arg Asn Pro Asp Asn Asp Pro Gln Gly Pro Trp Cys
130 135 140

Tyr Thr Thr Asp Pro Glu Lys Arg Tyr Asp Tyr Cys Asp Ile Leu Glu
145 150 155 160

Cys Glu Glu Glu Cys Met His Cys Ser Gly Glu Asn Tyr Asp Gly Lys
165 170 175

Ile Ser Lys Thr Met Ser Gly Leu Glu Cys Gln Ala Trp Asp Ser Gln
180 185 190

Ser Pro His Ala His Gly Tyr Ile Pro Ser Lys Phe Pro Asn Lys Asn
195 200 205

Leu Lys Lys Asn Tyr Cys Arg Asn Pro Asp Arg Glu Leu Arg Pro Trp
210 215 220

Cys Phe Thr Thr Asp Pro Asn Lys Arg Trp Glu Leu Cys Asp Ile Pro
225 230 235 240

Arg Cys Thr Thr Pro Pro Pro Ser Ser Gly Pro Thr Tyr Gln Cys Leu
245 250 255

Lys Gly Thr Gly Glu Asn Tyr Arg Gly Asn Val Ala Val Thr Val Ser
260 265 270

Gly His Thr Cys Gln His Trp Ser Ala Gln Thr Pro His Thr His Asn
275 280 285

Arg Thr Pro Glu Asn Phe Pro Cys Lys Asn Leu Asp Glu Asn Tyr Cys
290 295 300

Arg Asn Pro Asp Gly Lys Arg Ala Pro Trp Cys His Thr Thr Asn Ser
305 310 315 320

Gln Val Arg Trp Glu Tyr Cys Lys Ile Pro Ser Cys Asp Ser Ser Pro
325 330 335

Val Ser Thr Glu Glu Leu Ala Pro Thr Ala Pro Pro Glu Leu Thr Pro
340 345 350

Val Val Gln Asp Cys Tyr His Gly Asp Gly Gln Ser Tyr Arg Gly Thr
355 360 365

Ser Ser Thr Thr Thr Thr Gly Lys Lys Cys Gln Ser Trp Ser Ser Met
370 375 380

Thr Pro His Arg His Gln Lys Thr Pro Glu Asn Tyr Pro Asn Ala Gly
385 390 395 400

Leu Thr Met Asn Tyr Cys Arg Asn Pro Asp Ala Asp Lys Gly Pro Trp
405 410 415

Cys Phe Thr Thr Asp Pro Ser Val Arg Trp Glu Tyr Cys Asn Leu Lys
420 425 430

Lys Cys Ser Gly Thr Glu Ala Ser Val Val Ala Pro Pro Pro Val Val
435 440 445

Leu Leu Pro Asn Val Glu Thr Pro Ser Glu Glu Asp Cys Met Phe Gly
450 455 460

Asn Gly Lys Gly Tyr Arg Gly Lys Arg Ala Thr Thr Val Thr Gly Thr
465 470 475 480

Pro Cys Gln Asp Trp Ala Ala Gln Glu Pro His Arg His Ser Ile Phe
485 490 495

Thr Pro Glu Thr Asn Pro Arg Ala Gly Leu Glu Lys Asn Tyr Cys Arg
500 505 510

Asn Pro Asp Gly Asp Val Gly Gly Pro Trp Cys Tyr Thr Thr Asn Pro
515 520 525

Arg Lys Leu Tyr Asp Tyr Cys Asp Val Pro Gln Cys Ala Ala Pro Ser
530 535 540

Phe Asp Cys Gly Lys Pro Gln Val Glu Pro Lys Lys Cys Pro Gly Arg
545 550 555 560

Val Val Gly Gly Cys Val Ala His Pro His Ser Trp Pro Trp Gln Val
565 570 575

Ser Leu Arg Thr Arg Phe Gly Met His Phe Cys Gly Gly Thr Leu Ile
580 585 590

Ser Pro Glu Trp Val Leu Thr Ala Ala His Cys Leu Glu Lys Ser Pro
595 600 605

Arg Pro Ser Ser Tyr Lys Val Ile Leu Gly Ala His Gln Glu Val Asn
610 615 620

Leu Glu Pro His Val Gln Glu Ile Glu Val Ser Arg Leu Phe Leu Glu
625 630 635 640

```
Pro Thr Arg Lys Asp Ile Ala Leu Leu Lys Leu Ser Ser Pro Ala Val
645                 650                 655

Ile Thr Asp Lys Val Ile Pro Ala Cys Leu Pro Ser Pro Asn Tyr Val
660                 665                 670

Val Ala Asp Arg Thr Glu Cys Phe Ile Thr Gly Trp Gly Glu Thr Gln
675                 680                 685

Gly Thr Phe Gly Ala Gly Leu Leu Lys Glu Ala Gln Leu Pro Val Ile
690                 695                 700

Glu Asn Lys Val Cys Asn Arg Tyr Glu Phe Leu Asn Gly Arg Val Gln
705                 710                 715                 720

Ser Thr Glu Leu Cys Ala Gly His Leu Ala Gly Gly Thr Asp Ser Cys
725                 730                 735

Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Phe Glu Lys Asp Lys Tyr
740                 745                 750

Ile Leu Gln Gly Val Thr Ser Trp Gly Leu Gly Cys Ala Arg Pro Asn
755                 760                 765

Lys Pro Gly Val Tyr Val Arg Val Ser Arg Phe Val Thr Trp Ile Glu
770                 775                 780

Gly Val Met Arg Asn Asn
785                 790
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 55:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 153 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 55:

```
Val Tyr Leu Gln Thr Ser Leu Lys Tyr Asn Ile Leu Pro Glu Lys Glu
1               5                   10                  15

Glu Phe Pro Phe Ala Leu Gly Val Gln Thr Leu Pro Gln Thr Cys Asp
            20                  25                  30

Glu Pro Lys Ala His Thr Ser Phe Gln Ile Ser Leu Ser Val Ser Tyr
        35                  40                  45

Thr Gly Ser Arg Ser Ala Ser Asn Met Ala Ile Val Asp Val Lys Met
    50                  55                  60

Val Ser Gly Phe Ile Pro Leu Lys Pro Thr Val Lys Met Leu Glu Arg
65                  70                  75                  80
```

```
Ser Asn His Val Ser Arg Thr Glu Val Ser Ser Asn His Val Leu Ile
                85                  90                  95

Tyr Leu Asp Lys Val Ser Asn Gln Thr Leu Ser Leu Phe Phe Thr Val
            100                 105                 110

Leu Gln Asp Val Pro Val Arg Asp Leu Lys Pro Ala Ile Val Lys Val
        115                 120                 125

Tyr Asp Tyr Tyr Glu Thr Asp Glu Phe Ala Ile Ala Glu Tyr Asn Ala
    130                 135                 140

Pro Cys Ser Lys Asp Leu Gly Asn Ala
145                 150
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 56:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 202 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 56:

```
  Met Glu Leu Trp Gly Ala Tyr Leu Leu Leu Cys Leu Phe Ser Leu Leu
  1               5                  10                  15

Thr Gln Val Thr Thr Glu Pro Pro Thr Gln Lys Pro Lys Lys Ile Val
            20                  25                  30

Asn Ala Lys Lys Asp Val Val Asn Thr Lys Met Phe Glu Glu Leu Lys
        35                  40                  45

Ser Arg Leu Asp Thr Leu Ala Gln Glu Val Ala Leu Leu Lys Glu Gln
    50                  55                  60

Gln Ala Leu Gln Thr Val Cys Leu Lys Gly Thr Lys Val His Met Lys
65                  70                  75                  80

Cys Phe Leu Ala Phe Thr Gln Thr Lys Thr Phe His Glu Ala Ser Glu
                85                  90                  95

Asp Cys Ile Ser Arg Gly Gly Thr Leu Ser Thr Pro Gln Thr Gly Ser
            100                 105                 110
```

```
Glu Asn Asp Ala Leu Tyr Glu Tyr Leu Arg Gln Ser Val Gly Asn Glu
        115                 120                 125

Ala Glu Ile Trp Leu Gly Leu Asn Asp Met Ala Ala Glu Gly Thr Trp
    130                 135                 140

Val Asp Met Thr Gly Ala Arg Ile Ala Tyr Lys Asn Trp Glu Thr Glu
145                 150                 155                 160

Ile Thr Ala Gln Pro Asp Gly Gly Lys Thr Glu Asn Cys Ala Val Leu
                165                 170                 175

Ser Gly Ala Ala Asn Gly Lys Trp Phe Asp Lys Arg Cys Arg Asp Gln
            180                 185                 190

Leu Pro Tyr Ile Cys Gln Phe Gly Ile Val
        195                 200
```

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 57:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 246 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 57:

```
Gln Val Lys Leu Gln Gln Ser Gly Ala Glu Leu Val Lys Pro Gly Ala
1               5                   10                  15

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Ala Ser Tyr
            20                  25                  30

Trp Ile Asn Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile
        35                  40                  45
```

Gly His Ile Tyr Pro Val Arg Ser Ile Thr Lys Tyr Asn Glu Lys Phe
50 55 60

Lys Ser Lys Ala Thr Leu Thr Leu Asp Thr Ser Ser Ser Thr Ala Tyr
65 70 75 80

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys
85 90 95

Ser Arg Gly Asp Gly Ser Asp Tyr Tyr Ala Met Asp Tyr Trp Gly Gln
100 105 110

Gly Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Asp Ile Glu
115 120 125

Leu Thr Gln Ser Pro Ala Ile Leu Ser Ala Ser Pro Gly Gly Lys Val
130 135 140

Thr Met Thr Cys Arg Ala Ser Ser Ser Val Ser Tyr Met His Trp Tyr
145 150 155 160

Gln Gln Lys Pro Gly Ser Ser Pro Lys Pro Trp Ile Tyr Ala Thr Ser
165 170 175

Asn Leu Ala Ser Gly Val Pro Thr Arg Phe Ser Gly Thr Gly Ser Gly
180 185 190

Thr Ser Tyr Ser Leu Thr Ile Ser Arg Val Glu Ala Glu Asp Ala Ala
195 200 205

Thr Tyr Tyr Cys Gln Gln Trp Ser Arg Asn Pro Phe Thr Phe Gly Ser
210 215 220

Gly Thr Lys Leu Glu Ile Lys Arg Ala Ala Ala Glu Gln Lys Leu Ile
225 230 235 240

Ser Glu Glu Asp Leu Asn
245

(2) INFORMATIONEN FÜR SEQ.-ID Nr. 58:

(i) SEQUENZMERKMALE:

(A) LÄNGE: 101 Aminosäuren

(B) TYP: Aminosäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: Seq.-ID Nr. 58:

```
Met Ser Asn Thr Gln Ala Glu Arg Ser Ile Ile Gly Met Ile Asp Met
1               5                  10                  15

Phe His Lys Tyr Thr Arg Arg Asp Asp Lys Ile Asp Lys Pro Ser Leu
            20                  25                  30

Leu Thr Met Met Lys Glu Asn Phe Pro Asn Phe Leu Ser Ala Cys Asp
        35                  40                  45

Lys Lys Gly Thr Asn Tyr Leu Ala Asp Val Phe Glu Lys Lys Asp Lys
    50                  55                  60

Asn Glu Asp Lys Lys Ile Asp Phe Ser Glu Phe Leu Ser Leu Leu Gly
65                  70                  75                  80

Asp Ile Ala Thr Asp Tyr His Lys Gln Ser His Gly Ala Ala Pro Cys
                85                  90                  95

Ser Gly Gly Ser Gln
            100
```

Patentansprüche

1. Verfahren zur Erzeugung einer bearbeiteten Anordnung von Polypeptidmolekülen, wobei in der bearbeiteten Anordnung die vorhandenen Konformationszustände einen beträchtlichen Anteil an Polypeptidmolekülen in einer bestimmten gefalteten Konformation enthalten, aus einer Ausgangsanordnung aus Polypeptidmolekülen, welche dieselbe Aminosäuresequenz wie die bearbeitete Anordnung von Polypeptidmolekülen aufweisen, wobei die Ausgangsanordnung der vorhandenen Konformationszustände einen beträchtlichen Anteil an Polypeptidmolekülen in ungefalteten oder fehlgefalteten Konformationen enthält, wobei das Verfahren das Vornehmen einer Reihe aus zumindest drei aufeinanderfolgenden Zyklen mit der Ausgangsanordnung von Polypeptidmolekülen umfasst, die jeweils eine Abfolge aus Folgenden umfassen:
1) zumindest einem Denaturierungsschritt, der Bedingungen umfasst, die einen denaturierenden und/oder auffaltenden Einfluss auf die Polypeptidmoleküle der Anordnung ausüben, um einen Teil der Polypeptide in der Anordnung zu denaturieren und/ oder aufzufalten, gefolgt von
2) zumindest einem Renaturierungsschritt, der Bedingungen umfasst, die einen renaturierenden Einfluss auf die Polypeptidmoleküle mit Konformationen, die aus dem vorhergehenden Schritt resultieren, ausübt, so dass ein Teil der denaturierten und/oder aufgefalteten Polypeptide in der Anordnung renaturiert wird,
wobei die Abfolge aus den zumindest drei aufeinanderfolgenden Zyklen so angepasst ist, dass die Bedingun-

gen in zumindest einem Denaturierungsschritt zur Denaturierung und/oder Auffaltung eines kleineren Anteils der Polypeptide in der Anordnung führt als die Denaturierungsbedingungen in einem früheren Denaturierungsschritt in der Abfolge und dass die bearbeitete Anordnung von Polypeptidmolekülen einen höheren Anteil der Polypeptidmoleküle in der bestimmten gefalteten Konformation aufweist als
a) die Ausgangsanordnung und
b) eine entsprechende Ausgangsanordnung, die nur einem der Zyklen unterzogen worden ist.

2. Verfahren nach Anspruch 1, worin der beträchtliche Anteil an Polypeptidmolekülen in einer bestimmten gefalteten Konformation in der bearbeiteten Anordnung zumindest 5 Gew.-% der Ausgangsanordnung aus Polypeptidmolekülen ausmacht.

3. Verfahren nach Anspruch 1 oder 2, worin die Polypeptidmoleküle der bearbeiteten Anordnung Cystein-hältige Moleküle umfassen und die bearbeitete Anordnung einen beträchtlichen Anteil an Polypeptidmolekülen in einer bestimmten gefalteten Konformation umfasst, die zusätzlich identische Disulfidbrücken-Topologie aufweisen.

4. Verfahren nach einem der Ansprüche 1 bis 3, worin die Polypeptidmoleküle Moleküle sind, die eine Aminosäuresequenz aufweisen, die mit jener eines authentischen Polypeptids identisch ist, oder Moleküle sind, die eine Aminosäuresequenz umfassen, die jener eines authentischen Polypeptids entspricht, das an ein oder zwei zusätzliche Polypeptidsegmente gebunden ist.

5. Verfahren nach Anspruch 4, worin die Aminosäuresequenz, die jener eines authentischen Polypeptids entspricht, über eine spaltbare Bindung oder ähnliche oder unähnliche spaltbare Bindungen an das bzw. die zusätzliche Polypeptidsegment oder -segmente gebunden ist.

6. Verfahren nach einem der Ansprüche 1 bis 5, worin die Abfolge zumindest 5 Zyklen, beispielsweise zumindest 8, zumindest 10 und zumindest 25 Zyklen, und maximal 2.000 Zyklen, beispielsweise maximal 1.000, maximal 500 und maximal 200 Zyklen, maximal 100 und maximal 50 Zyklen, umfasst.

7. Verfahren nach einem der vorangegangenen Ansprüche, worin die Dauer jedes Denaturierungsschritts zumindest 1 ms und maximal 1 h beträgt und die Dauer jedes Renaturierungsschritts zumindest 1 s und maximal 12 h beträgt.

8. Verfahren nach Anspruch 7, worin die Denaturierungsbedingungen jedes einzelnen Denaturierungsschritts für einen Zeitraum konstant gehalten werden und die Renaturierungsbedingungen jedes einzelnen Renaturierungsschritts für einen Zeitraum konstant gehalten werden, wobei die Zeiträume, während deren die Bedingungen konstant gehalten werden, durch Übergangszeiträume getrennt sind, während deren die Bedingungen verändert werden.

9. Verfahren nach Anspruch 8, worin der Übergangszeitraum zwischen Schritten, für die die Bedingungen konstant gehalten werden, zwischen 0,1 s und 12 h dauert.

10. Verfahren nach Anspruch 9, worin der Zeitraum, für den die Denaturierungsbedingungen des Denaturierungsschritts konstant gehalten werden, zwischen 1 und 10 min dauert und der Zeitraum, für den die Renaturierungsbedingungen des Renaturierungsschritts konstant gehalten werden, zwischen 1 und 45 min dauert.

11. Verfahren nach einem der vorangegangenen Ansprüche, worin die Polypeptidmoleküle während der Denaturierungs- und Renaturierungsschritte mit einer flüssigen Phase in Kontakt stehen, wobei die flüssige Phase eine wässrige Phase oder eine organische Phase ist.

12. Verfahren nach Anspruch 11, worin die Polypeptidmoleküle auf eine Umgebung begrenzt sind, die eine Veränderung oder einen Austausch der flüssigen Phase ermöglicht, ohne dass die Polypeptidmoleküle mitgerissen werden.

13. Verfahren nach Anspruch 12, worin die Polypeptide auf eine Dialysevorrichtung oder ein flüssiges Zweiphasen-System begrenzt sind.

14. Verfahren nach Anspruch 12, worin die Polypeptidmoleküle an einen festen oder halbfesten Träger, wie z. B. eine Filteroberfläche, eine Hohlfaser oder ein kugelförmiges Chromatographiemedium, z. B. ein Agarose- oder Polyacrylamidgel, eine faserförmige Zellulosematrix, eine HPLC- oder FPLC-Matrix, eine Substanz mit

Molekülen einer solchen Größe, dass die Moleküle mit den daran gebundenen Polypeptidmolekülen, wenn sie in einer flüssigen Phase gelöst oder dispergiert sind, durch ein Filter zurückgehalten werden können, eine Substanz, die fähig ist, Micellen zu bilden oder an der Bildung von Micellen teilzunehmen, was ermöglicht, die flüssige Phase zu wechseln oder auszutauschen, ohne die Micellen mitzureißen, oder ein wasserlösliches Polymer gebunden werden.

15. Verfahren nach Anspruch 14, worin die Polypeptidmoleküle über eine Gruppierung, die Affinität für eine Komponente des Trägers aufweist, wie z. B. eine Biotingruppe oder ein Analog davon, das an eine Aminosäuregruppierung des Polypeptids gebunden ist, nichtkovalent an den Träger adsorbiert werden, wobei an den Träger Avidin, Streptavidin oder Analoge davon gebunden sind.

16. Verfahren nach Anspruch 15, worin die Gruppierung eine Aminosäuresequenz aufweist, die mit Seq.-ID Nr. 47 identisch ist, wobei der Träger ein Nitrilotriessigsäure-Derivat (NTA) aufweist, das mit Ni^{++}-Ionen beladen ist.

17. Verfahren nach einem der vorangegangenen Ansprüche, worin die Polypeptidmoleküle ein Polypeptidsegment umfassen, das fähig ist, bevorzugte Spaltung durch ein Spaltungsmittel zu einer spezifischen Peptidbindung zu dirigieren.

18. Verfahren nach Anspruch 17, worin das die Spaltung dirigierende Polypeptidsegment eines ist, das bevorzugte Spaltung zu einer spezifischen Peptidbindung durch ein Spaltungsmittel lenken kann, das aus der aus Bromcyan, Hydroxylamin, Iodbenzoesäure, N-Bromsuccinimid und Enzymen, wie z. B. Rinderkoagulationsfaktor X_a oder Analogen und/oder Homologen davon und Rinder-Enterokinase oder Analogen und/oder Homologen davon bestehenden Gruppe ausgewählt ist.

19. Verfahren nach Anspruch 17 oder 18, worin das Polypeptidsegment, das bevorzugte Spaltung dirigiert, eine Sequenz ist, die selektiv vom Rinderkoagulationsfaktor X_a oder einem Analog und/oder Homolog davon erkannt wird, wie z. B. ein Polypeptidsegment, das eine Aminosäuresequenz aufweist, die aus der aus Seq.-ID Nr. 38, Seq.-ID Nr. 40, Seq.-ID Nr. 41 oder Seq.-ID Nr. 42 bestehenden Gruppe ausgewählt ist.

20. Verfahren nach einem der vorangegangenen Ansprüche, worin die Polypeptidmoleküle ein Polypeptidsegment umfassen, das in vitro in ein derivatisiertes Polypeptidsegment umgewandelt werden kann, das in der Lage ist, bevorzugte Spaltung durch ein Spaltungsmittel zu einer spezifische Peptidbindung zu dirigieren.

21. Verfahren nach Anspruch 20, worin das in vitro umwandelbare Polypeptidsegment in ein derivatisiertes Polypeptidsegment umgewandelt werden kann, das selektiv von Rinderkoagulationsfaktor X_a oder einem Analog und/oder Homolog davon erkannt wird.

22. Verfahren nach Anspruch 21, worin das in vitro umwandelbare Polypeptidsegment eine Aminosäuresequenz aufweist, die aus der aus Seq.-ID Nr. 43, Seq.-ID Nr. 44, Seq.-ID Nr. 45 und Seq.-ID Nr. 46 bestehenden Gruppe ausgewählt ist.

23. Verfahren nach Anspruch 22, worin die Polypeptidmoleküle ein Polypeptidsegment umfassen, mit entweder
der Aminosäuresequenz Seq.-ID Nr. 43 oder Seq.-ID Nr. 44, die durch Umsetzen des Cysteinrests mit N-(2-Mercaptoethyl)morpholyl-2-thiopyridyldisulfid oder Mercaptothioacetat-2-thiopyridiyldisulfid in ein derivatisiertes Polypeptid umgewandelt wird, das selektiv von Rinderkoagulationsfaktor X_a oder einem Analog und/oder Homolog davon erkannt wird, oder
mit der Aminosäuresequenz Seq.-ID Nr. 45 oder Seq.-ID Nr. 46, die durch Oxidation der Thioethergruppierung in der Methionin-Seitengruppe zu einem Sulfoxid- oder Sulfonderivat in ein derivatisiertes Polypeptid umgewandelt wird, das selektiv von Rinderkoagulationsfaktor X_a erkannt wird.

24. Verfahren nach Anspruch 19, 22 oder 23, worin das Polypeptidsegment, das aus der aus Seq.-ID Nr. 38, Seq.-ID Nr. 40, Seq.-ID Nr. 41 oder Seq.-ID Nr. 42 bestehenden Gruppe ausgewählt ist oder das aus der aus Seq.-ID Nr: 43, Seq.-ID Nr. 44, Seq.-ID Nr. 45 und Seq.-ID Nr. 46 bestehenden Gruppe ausgewählt ist, N-terminal an das authentische Polypeptid gebunden ist.

25. Verfahren nach einem der Ansprüche 8 bis 24, worin die Veränderung der Bedingungen während des Übergangszeitraums erreicht wird, indem die chemische Zusammensetzung der flüssigen Phase verändert wird, mit der die Polypeptidmoleküle in Kontakt stehen.

26. Verfahren nach Anspruch 25, worin die Denaturierung der Polypeptidmoleküle erreicht wird, indem die Polypeptidmoleküle mit einer flüssigen Phase in Kontakt gebracht werden, in der zumindest eine denaturierende Verbindung gelöst ist, und worin Renaturierung der Polypeptidmoleküle erreicht wird, indem die Polypeptidmoleküle mit einer flüssigen Phase in Kontakt gebracht werden, die entweder zumindest eine gelöste denaturierende Verbindung in einer solchen Konzentration enthält, dass der Kontakt mit der flüssigen Phase die Anordnung von Polypeptidmolekülen in ihren jeweiligen Konformationszuständen, die aus dem vorherigen Schritt resultieren, tendenziell eher renaturiert als denaturiert, oder keine denaturierende Verbindung enthält.

27. Verfahren nach Anspruch 26, worin die Denaturierung der Polypeptidmoleküle durch Verringerung oder Erhöhung des pH-Werts der flüssigen Phase erreicht wird.

28. Verfahren nach Anspruch 26 oder 27, worin die denaturierende Verbindung aus Harnstoff, Guanidin-HCl und Di-C_{1-6}-alkylformamid, wie z. B. Dimethylformamid, und Di-C_{1-6}-alkylsulfon, ausgewählt ist.

29. Verfahren nach einem der Ansprüche 11 bis 28, worin die in zumindest einem der Denaturierungsschritte und/oder in zumindest einem der Renaturierungsschritte verwendete flüssige Phase zumindest ein Disulfid-Reshufflingsystem enthält.

30. Verfahren nach Anspruch 29, worin es sich bei dem zumindest einen Disulfid-Reshufflingsystem um eines handelt, das fähig ist, inkorrekt gebildete Disulfidbrücken in ungefalteten und/oder fehlgefalteten Polypeptidmolekülen unter Bedingungen der Konzentration des Denaturierungsmittels, bei der ungefaltete und/oder fehlgefaltete Proteine denaturiert werden, mit höherer Effizienz Reduktion und/oder Reshuffling zu unterziehen, als korrekt gebildete Disulfidbrücken in ungefalteten und/oder fehlgefalteten Polypeptidmolekülen Reduktion und/oder Reshuffling unterzogen werden.

31. Verfahren nach Anspruch 30, worin die Gegenwart des Disulfid-Reshufflingsystems in zumindest einem Schritt zu einem Verhältnis zwischen der relativen Menge an Reduktion/Reshuffling unterzogenen anfänglich inkorrekt gebildeten Disulfidbrücken und der relativen Menge an Reduktion/Reshuffling unterzogenen anfänglich korrekt gebildeten Disulfidbrücken von zumindest 1,05 führt.

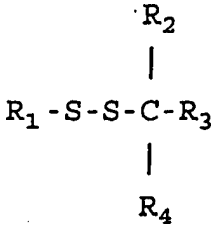
32. Verfahren nach einem der Ansprüche 29 bis 31, worin das Disulfid-Reshufflingsystem Glutathion, 2-Mercaptoethanol oder Thiocholin enthält, jedes davon im Gemisch mit seinem entsprechenden symmetrischen Disulfid.

33. Verfahren nach einem der Ansprüche 11 bis 32, worin alle Cysteinreste in den Polypeptidmolekülen während zumindest eines der Denaturierungs/Renaturierungs-Zyklen in gemischte Disulfidprodukte von entweder Glutathion, Thiocholin, Mercaptoethanol oder Mercaptoessigsäure umgewandelt worden sind.

34. Verfahren nach Anspruch 33, worin die Umwandlung der Cysteinreste in gemischte Disulfidprodukte erzielt wird, indem die vollständig denaturierte und vollständig reduzierte Anordnung von Polypeptidmolekülen mit einem Überschuss eines Reagens umgesetzt wird, das eine hochenergetische gemischte Disulfidverbindung ist.

35. Verfahren nach Anspruch 34, worin die hochenergetische gemischte Disulfidverbindung aliphatisch-aromatisch ist.

36. Verfahren nach Anspruch 34 oder 35, worin die hochenergetische gemischte Disulfidverbindung die allgemeine Formel:



hat, worin R_1 2-Pyridyl ist, R_2, R_3 und R_4 Wasserstoff oder eine gegebenenfalls substituierte, niedere aromatische oder aliphatische Kohlenwasserstoffgruppe sind.

37. Verfahren nach einem der Ansprüche 34 bis 36, worin die hochenergetische gemischte Disulfidverbin-

dung aus der aus Glutathionyl-2-thiopyridyldisulfid, 2-Thiocholyl-2-thiopyridiyldisulfid, 2-Mercaptoethanol-2-thiopyridyldisulfid und Mercaptoacetat-2-thiopyridyldisulfid bestehenden Gruppe ausgewählt ist.

38. Verfahren nach einem der Ansprüche 11 bis 37, worin die Polarität der bei der Renaturierung der Polypeptidmoleküle verwendeten flüssigen Phase durch Zusatz eines Salzes, eines Polymers und/oder einer Hydrofluorverbindung, wie z. B. Trifluorethanol, modifiziert worden ist.

39. Verfahren nach einem der Ansprüche 1 bis 24 oder 29 bis 38, worin die Denaturierung und Renaturierung der Polypeptidmoleküle durch direkte Änderungen der physikalischen Parameter, denen die Polypeptidmoleküle ausgesetzt sind, wie z. B. Temperatur oder Druck, oder durch Änderungen der physikalischen Parameter, die die Denaturierungs- und Renaturierungsbedingungen verstärken oder mäßigen, wie z. B. Temperatur oder Druck, erreicht wird.

40. Verfahren nach Anspruch 25, worin die chemischen Veränderungen in der flüssigen Phase durch einen Wechsel zwischen einer Denaturierungslösung B und einer Renaturierungslösung A erreicht werden.

41. Verfahren nach Anspruch 40, worin die Konzentration einer oder mehrerer denaturierenden Verbindungen in B nach jedem Zyklus eingestellt wird.

42. Verfahren nach Anspruch 41, worin die Konzentration einer oder mehrerer denaturierenden Verbindungen in B nach jedem Zyklus verringert wird.

43. Verfahren nach Anspruch 40, worin die Konzentration einer oder mehrerer denaturierenden Verbindungen in Medium B in jedem Zyklus konstant gehalten wird.

44. Verfahren nach einem der vorangegangenen Ansprüche, worin die Polypeptidmoleküle der Anordnung eine Länge von zumindest 25 Aminosäureresten und maximal 5.000 Aminosäureresten aufweisen.

45. Verfahren nach einem der vorangegangenen Ansprüche, worin die Polypeptide der Ausgangsanordnung künstliche Polypeptide sind, die in prokaryotischen Zellen durch DNA-Rekombinationstechniken produziert wurden.

46. Verfahren nach Anspruch 45, worin die Anfangsanordnung von Polypeptidmolekülen ungefaltete oder fehlgefaltete Diakörpermoleküle (künstliche bispezifische und zweiwertige Antikörperfragmente) oder Monomerfragmente von Diakörpermolekülen sind.

47. Verfahren zur Herstellung von korrekt gefalteten Diakörpermolekülen, worin eine Anfangsanordnung von Polypeptidmolekülen, die ungefaltete und/oder fehlgefaltete Polypeptide mit Aminosäuresequenzen umfassen, die mit Monomerfragmenten von Diakörpermolekülen identisch sind, einer Abfolge von zumindest zwei aufeinander folgenden Zyklen unterzogen wird, die jeweils eine Abfolge von Folgendem umfassen:
1) zumindest einem Denaturierungsschritt, der Bedingungen umfasst, die einen denaturierenden und/oder auffaltenden Einfluss auf die Polypeptidmoleküle der Anordnung ausüben, um einen Teil der Polypeptide in der Anordnung zu denaturieren, gefolgt von
2) zumindest einem Renaturierungsschritt, der Bedingungen umfasst, die einen renaturierenden Einfluss auf die Polypeptidmoleküle mit Konformationen, die aus dem vorhergehenden Schritt resultieren, ausübt, so dass ein Teil der denaturierten und/oder aufgefalteten Polypeptide in der Anordnung renaturiert wird,
wobei die Abfolge von Zyklen so beschaffen ist, dass ein beträchtlicher Anteil der Anfangsanordnung von Polypeptidmolekülen in einen Anteil aus korrekt gefalteten Diakörpermolekülen umgewandelt wird.

48. Verfahren nach Anspruch 47, worin die Polypeptidmoleküle in zumindest einem Denaturierungs/Renaturierungs-Zyklus mit einer flüssigen Phase in Kontakt stehen, die zumindest ein Disulfid-Reshufflingsystem enthält.

Es folgen 34 Blatt Zeichnungen

Anhängende Zeichnungen

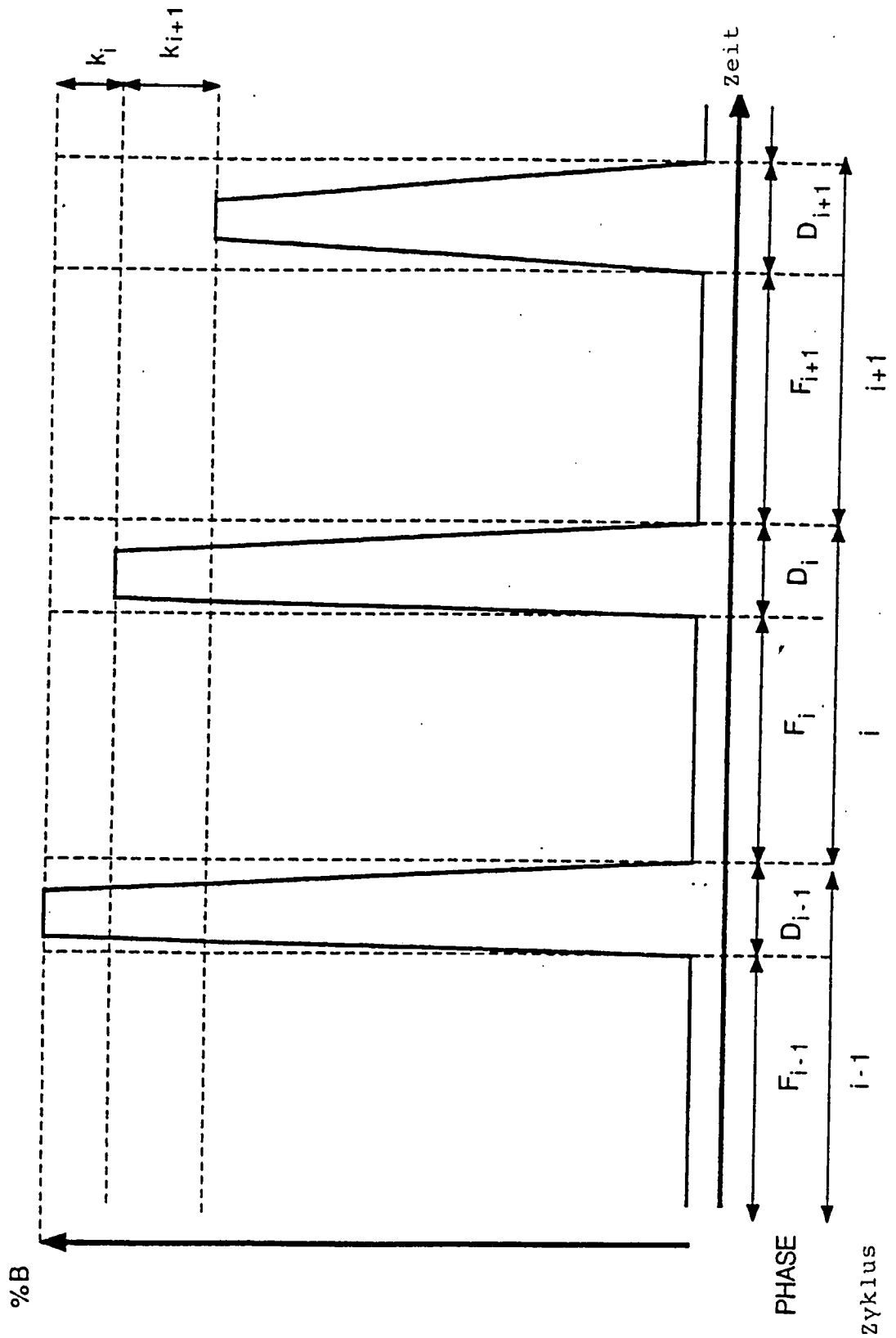


Fig. 1

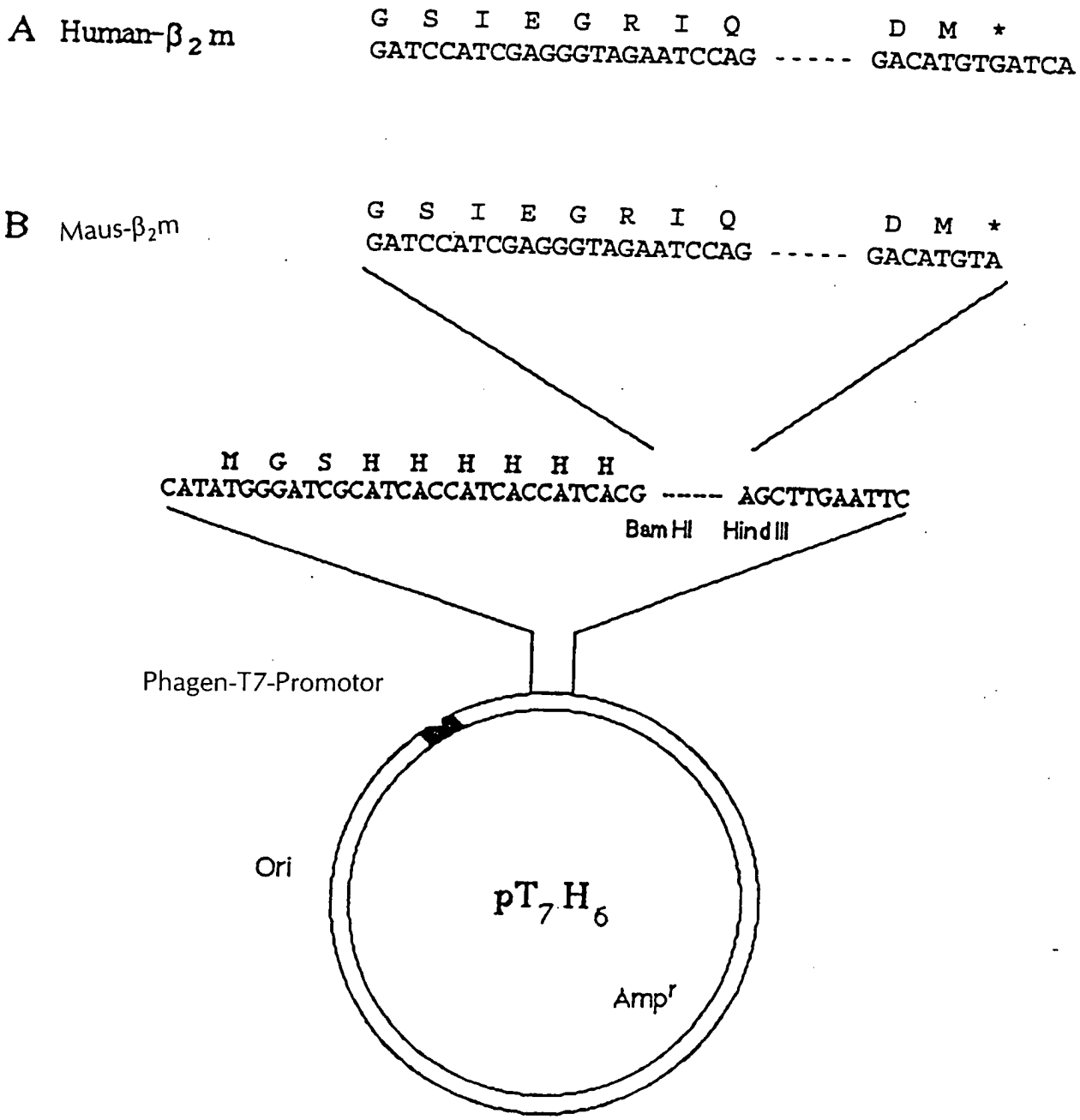

A Human-β2m
G S I E G R I Q D M *
GATCCATCGAGGGTAGAATCCAG ----- GACATGTGATCA
B Maus-β2m
G S I E G R I Q D M *
GATCCATCGAGGGTAGAATCCAG ----- GACATGTA
M G S H H H H H H
CATATGGGATCGCATCACCATCACCATCACG ----- AGCTTGAATTC
BamHI HindIII
Phagen-T7-Promotor
Ori
pT7H6
Ampr

Fig. 2

A:

Human-β_2-Microglobulin

```
-20                -10                -1 1                10
 M S R S V A L A V L A L L S L S G L E A I Q R T P K I Q V Y

                   20                 30                  40
 S R H P A E N G K S N F L N C Y V S G F H P S D I E V D L L

                   50                 60                  70
 K N G E R I E K V E H S D L S F S K D W S F Y L L Y Y T E F

                   80                 90
 T P T E K D E Y A C R V N H V T L S Q P K I V K W D R D M
```

B:

Maus-β_2-Microglobulin

```
-20                -10                -1 1                10
 M A R S V T L V F L V L V S L T G L Y A I Q K T P Q I Q V Y

                   20                 30                  40
 S R H P P E N G K P N I L N C Y V T Q F H P P H I E I Q M L

                   50                 60                  70
 K N G K K I P K V E M S D M S F S K D W S F Y I L A H T E F

                   80                 90
 T P T E T D T Y A C R V K H D S M A E P K T V Y W D R D M
```

Fig. 3

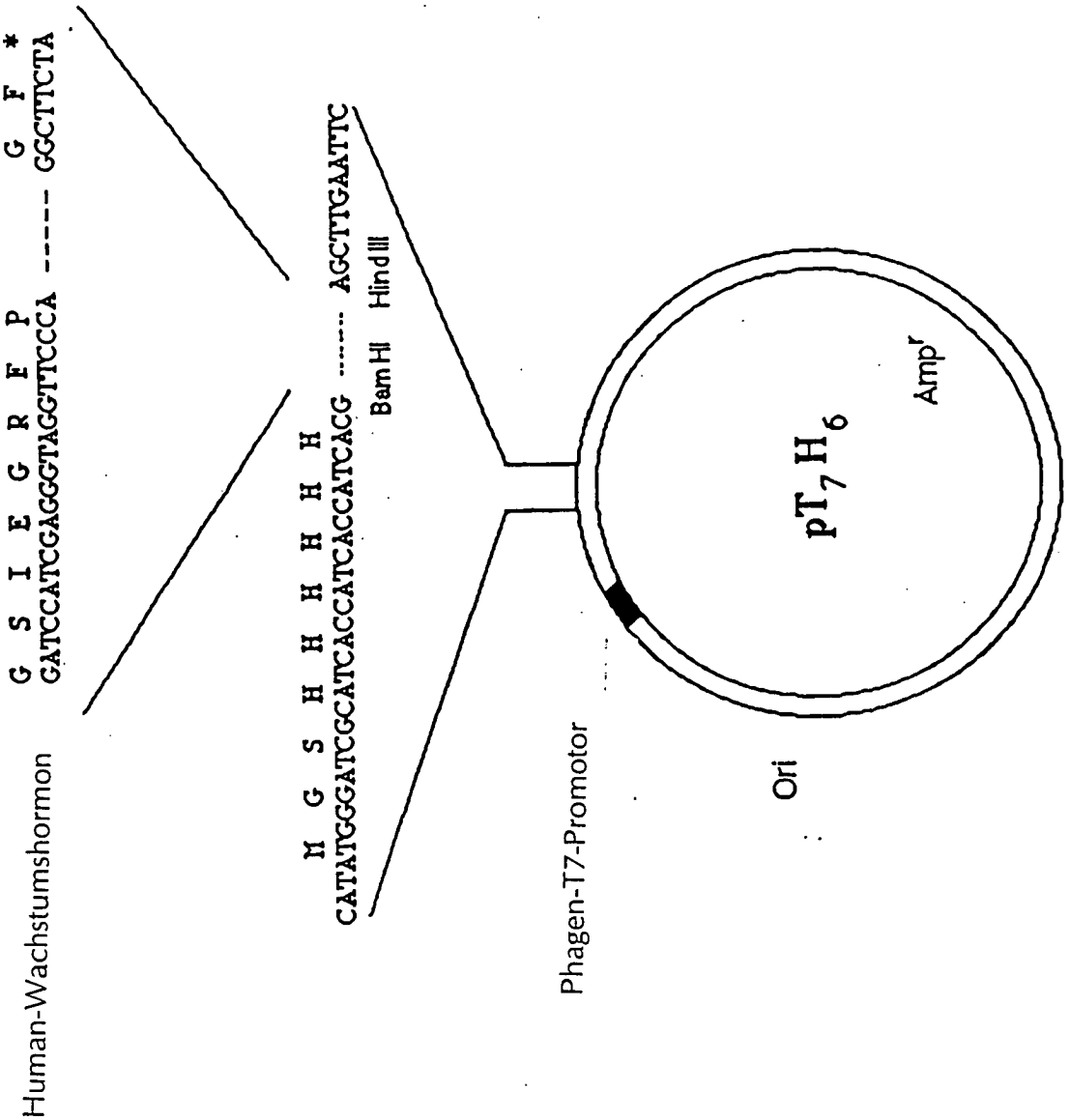

Human-Wachstumshormon
G S I E G R F P G F *
GATCCATCGAGGGTAGGGTTCCCA ----- GGCTTCTA
M G S H H H H H H
CATATGGGATCGCATCACCATCACCATCACG ----- AGCTTGAATTC
BamHI HindIII
Phagen-T7-Promotor
Ori
pT7H6
Ampr

Fig. 4

Human-Wachstumshormon (Somatotropin)

```
-26                                                  -1 1     4
 M A T G S R T S L L L A F G L L C L P W L Q E G S A F P T I

          10                  20                  30
 P L S R L F D N A S L R A H R L H Q L A F D T Y Q E F E E A

          40                  50                  60
 Y I P K E Q K Y S F L Q N P Q T S L C F S E S I P T P S N R

          70                  80                  90
 E E T Q Q K S N L E L L R I S L L L I Q S W L E P V Q F L R

         100                 110                 120
 S V F A N S L V Y G A S D S N V Y D L L K D L E E G I Q T L

         130                 140                 150
 M G R L E D G S P R T G Q I F K Q T Y S K F D T N S H N D D

         160                 170                 180
 A L L K N Y G L L Y C F R K D M D K V E T F L R I V Q C R S

         190
 V E G S C G F
```

Fig. 5

α_2 MR:

#1
G S I E G R A I C R *
GATCCATCGAGGGTAGGGCTATC ----- TGCCGATA

#2
G S I E G R A I K A *
GATCCATCGAGGGTAGGGCTATC ----- AAGGCCTA

#3
G S I E G R A I K K *
GATCCATCGAGGGTAGGGCTATC ----- AAGAAGTA

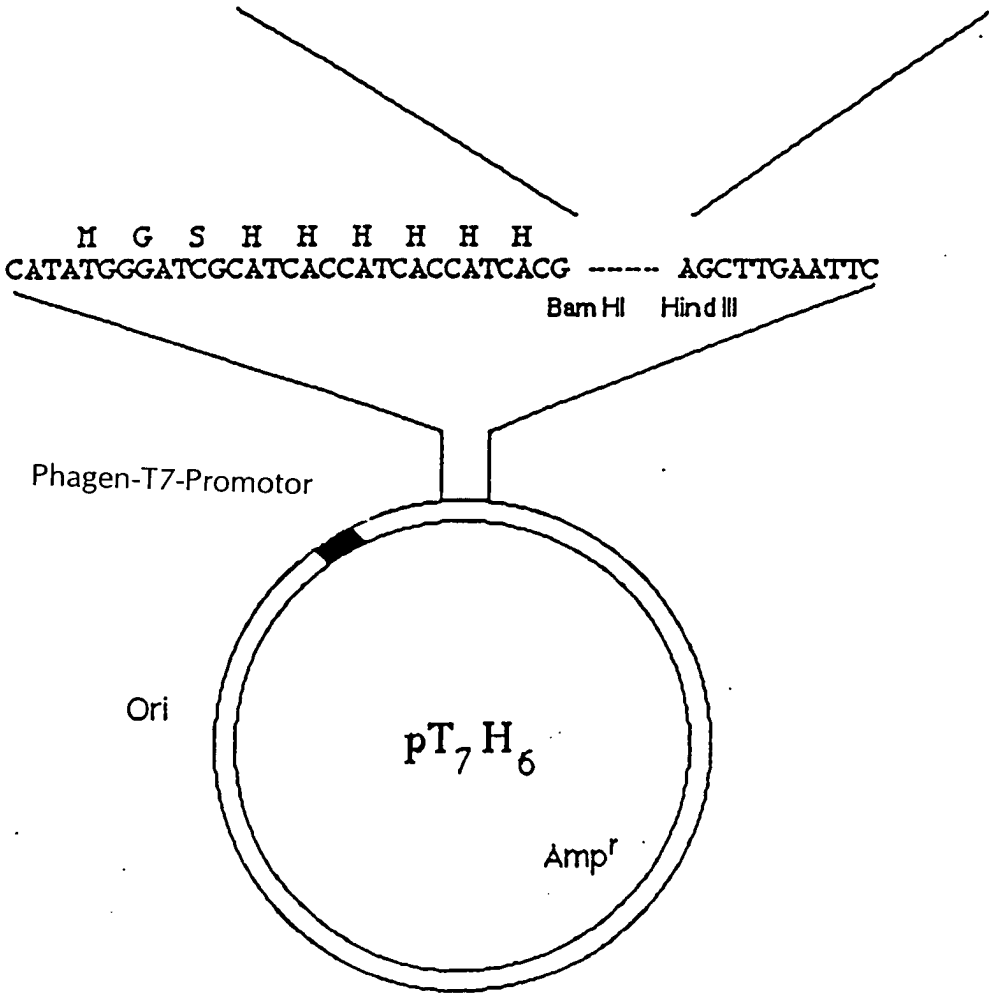


Fig. 6

α_2MR:

#4
G S I E G R G T L D *
GATCCATCGAGGGTAGGGGCACC ----- CTGGACTA

#5
G S I E G R V P D Q *
GATCCATCGAGGGTAGGGTGCCT ----- GACCAGTA

#6
G S I E G R G G Q C F K *
GATCAATCGAGGGTAGGGGTGGTCAGTGC---- TTTAAGTA

G K G S H H H H H H
GGGAAGGGATCGCATCACCATCACCATCACG ----- AGCTTGGCGTA
BamHI HindIII

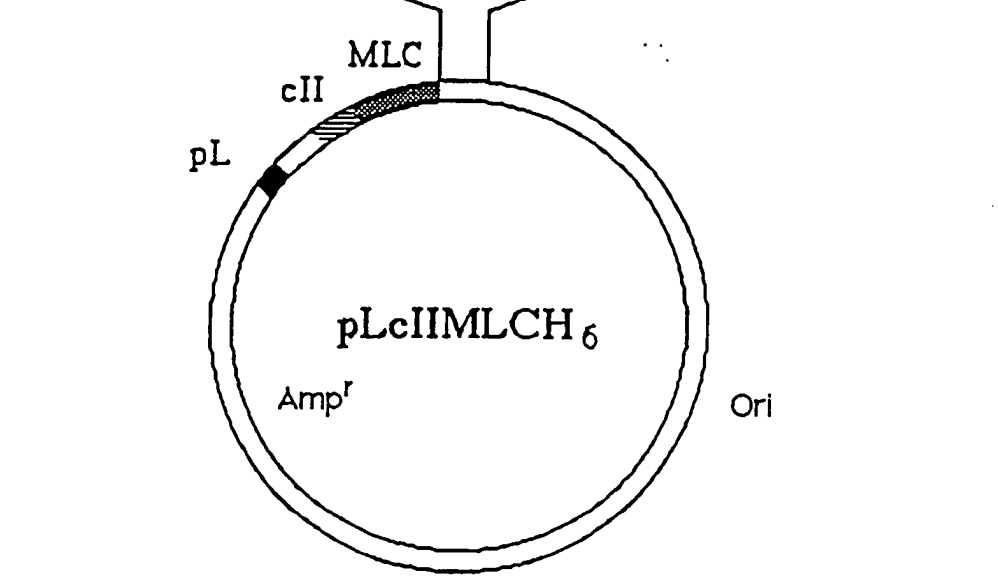


Fig. 7

α_2MR:

#7
G S I E G R G T F K *
GATCCATCGAGGGTAGGGGCACC ----- TTTAAGTA

#8
G S I E G R A V H I *
GATCCATCGAGGGTAGGGCGGTG ----- CACATCTA

#9
G S I E G R V S S I *
GATCCATCGAGGGTAGGGTGTCC ----- AGCATCTA

G K G S H H H H H H
GGGAAGGGATCGCATCACCATCACCATCACG ----- AGCTTGGCGTA
BamHI HindIII

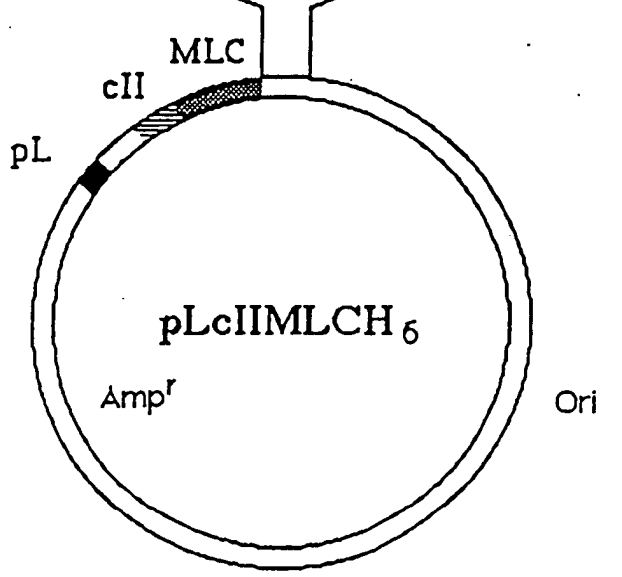


Fig. 8

α_2-Macroglobulinrezeptor

```
                          20
   1 MLTPPLLLLLPLLSALVAAAIDAPKTCSPKQFACRDQITCISKGWRCDGERDCPDGSDEA
                                                              109
  61 PEICPQSKAQRCQPNEHNCLGTELCVPMSRLCNGVQDCMDGSDEGPHCRELQGNCSRLGC
 121 QHHCVPTLDGPTCYCNSSFQLQADGKTCKDFDECSVYGTCSQLCTNTDGSFICGCVEGYL
               190
 181 LQPDNRSCKAKNEPVDRPPVLLIANSQNILATYLSGAQVSTITPTSTRQTTAMDFSYANE
 241 TVCWVHVGDSAAQTQLKCARMPGLKGFVDEHTINISLSLHHVEQMAIDWLTGNFYFVDDI
 301 DDRIFVCNRNGDTCVTLLDLELYNPKGIALDPAMGKVFFTDYGQIPKVERCDMDGQNRTK
 361 LVDSKIVFPHGITLDLVSRLVYWADAYLDYIEVVDYEGKGRQTIIQGILIEHLYGLTVFE
 421 NYLYATNSDNANAQQKTSVIRVNRFNSTEYQVVTRVDKGGALHIYHQRRQPRVRSHACEN
                                                         521
 481 DQYGKPGGCSDICLLANSHKARTCRCRSGFSLGSDGKSCKKPEHELFLVYGKGRPGIIRG
 541 MDMGAKVPDEHMIPIENLMNPRALDFHAETGFIYFADTTSYLIGRQKIDGTERETILKDG
 601 IHNVEGVAVDWMGDNLYWTDDGPKKTISVARLEKAAQTRKTLIEGKMTHPRAIVVDPLNG
 661 WMYWTDWEEDPKDSRRGRLERAWMDGSHRDIFVTSKTVLWPNGLSLDIPAGRLYWVDAFY
 721 DRIETILLNGTDRKIVYEGPELNHAFGLCHHGNYLFWTEYRSGSVYRLERGVGGAPPTVT
                               803
 781 LLRSERPPIFEIRMYDAQQQQVGTNKCRVNNGGCSSLCLATPGSRQCACAEDQVLDADGV
 841 TCLANPSYVPPPQCQPGEFACANSRCIQERWKCDGDNDCLDNSDEAPALCHQHTCPSDRF
 901 KCENNRCIPNRWLCDGDNDCGNSEDESNATCSARTCPPNQFSCASGRCIPISWTCDLDDD
 961 CGDRSDESASCAYPTCFPLTQFTCNNGRCININWRCDNDNDCGDNSDEAGCSHSCSSTQF
1021 KCNSGRCIPEHWTCDGDNDCGDYSDETHANCTNQATRPPGGCHTDEFQCRLDGLCIPLRW
1081 RCDGDTDCMDSSDEKSCEGVTHVCDPSVKFGCKDSARCISKAWVCDGDNDCEDNSDEENC
                                                    1184
1141 ESLACRPPSHPCANNTSVCLPPDKLCDGNDDCGDGSDEGELCDQCSLNNGGCSHNCSVAP
1201 GEGIVCSCPLGMELGPDNHTCQIQSYCAKHLKCSQKCDQNKFSVKCSCYEGWVLEPDGES
      1265
1261 CRSLDPFKPFIIFSNRHEIRRIDLHKGDYSVLVPGLRNTIALDFHLSQSALYWTDVVEDK
1321 IYRGKLLDNGALTSFEVVIQYGLATPEGLAVDWIAGNIYWVESNLDQIEVAKLDGTLRTT
1381 LLAGDIEHPRAIALDPRDGILFWTDWDASLPRIEAASMSGAGRRTVHRETGSGGWPNGLT
1441 VDYLEKRILWIDARSDAIYSARYDGSGHMEVLRGHEFLSHPFAVTLYGGEVYWTDWRTNT
1501 LAKANKWTGHNVTVVQRTNTQPFDLQVYHPSRQPMAPNPCEANGGQGPCSHLCLINYNRT
                          1582
1561 VSCACPHLMKLHKDNTTCYEFKKFLLYARQMEIRGVDLDAPYYNYIISFTVPDIDNVTVL
1621 DYDAREQRVYWSDVRTQAIKRAFINGTGVETVVSADLPNAHGLAVDWVSRNLFWTSYDTN
1681 KKQINVARLDGSFKNAVVQGLEQPHGLVVHPLRGKLYWTDGDNISMANMDGSNRTLLFSG
1741 QKGPVGLAIDFPESKLYWISSGNHTINRCNLDGSGLEVIDAMRSQLGKATALAIMGDKLW
1801 WADQVSEKMGTCSKADGSGSVVLRNSTTLVMHMKVYDESIQLDHKGTNPCSVNNGDCSQL
1861 CLPTSETTRSCMCTAGYSLRSGQQACEGVGSFLLYSVHEGIRGIPLDPNDKSDALVPVSG
1921 TSLAVGIDFHAENDTIYWVDMGLSTISRAKRDQTWREDVVTNGIGRVEGIAVDWIAGNIY
1981 WTDQGFDVIEVARLNGSFRYVVISQGLDKPKAITVHPEKGYLFWTEWGQYPRIERSRLDG
2041 TERVVLVNVSISWPNGISVDYQDGKLYWCDARTDKIERIDLETGENREVVLSSNNMDMFS
2101 VSVFEDFIYWSDRTHANGSIKRGSKDNATDSVPLRTGIGVQLKDIKVFNRDRQKGTNVCA
2161 VANGGCQQLCLYRGRGQRACACAHGMLAEDGASCREYAGYLLYSERTILKSIHLSDERNL
2221 NAPVQPFEDPEHMKNVIALAFDYRAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRITIVE
2281 NVGSVEGLAYHRGWDTLYWTSYTTSTITRHTVDQTRPGAFERETVITMSGDDHPRAFVLD
2341 ECQNLMFWTNWNEQHPSIMRAALSGANVLTLIEKDIRTPNGLAIDHRAEKLYFSDATLDK
2401 IERCEYDGSHRYVILKSEPVHPFGLAVYGEHIFWTDWVRRAVQRANKHVGSNMKLLRVDI
                                                         2520
2461 PQQPMGIIAVANDTNSCELSPCRINNGGCQDLCLLTHQGHVNCSCRGGRILQDDLTCRAV
2521 NSSCRAQDEFECANGECINFSLTCDGVPHCKDKSDEKPSYCNSRRCKKTFRQCSNGRCVS
2581 NMLWCNGADDCGDGSDEIPCNKTACGVGEFRCRDGTCIGNSSRCNQFVDCEDASDEMNCS
2641 ATDCSSYFRLGVKGVLFQPCERTSLCYAPSWVCDGANDCGDYSDERDCPGVKRPRCPLNY
2701 FACPSGRCIPMSWTCDKEDDCEHGEDETHCNKFCSEAQFECQNHRCISKQWLCDGSDDCG
```

Fig. 9a

2761 DGSDEAAHCEGKTCGPSSFSCPGTHVCVPERWLCDGDKDCADGADESIAAGCLYNSTCDD
2821 REFMCQNRQCIPKHFVCDHDRDCADGSDESPECEYPTCGPSEFRCANGRCLSSRQWECDG
2881 ENDCHDQSDEAPKNPHCTSPEHKCNASSQFLCSSGRCVAEALLCNGQDDCGDSSDERGCH
2941
2941 INECLSRKLSGCSQDCEDLKIGFKCRCRPGFRLKDDGRTCADVDECSTTFPCSQRCINTH
3001 GSYKCLCVEGYAPRGGDPHSCKAVTDEEPFLIFANRYYLRKLNLDGSNYTLLKQGLNNAV
3061 ALDFDYREQMIYWTDVTTQGSMIRRMHLNGSNVQVLHRTGLSNPDGLAVDWVGGNLYWCD
3121 KGRDTIEVSKLNGAYRTVLVSSGLREPRALVVDVQNGYLYWTDWGDHSLIGRIGMDGSSR
3181 SVIVDTKITWPNGLTLDYVTERIYWADAREDYIEFASLDGSNRHVVLSQDIPHIFALTLF
3241 EDYVYWTDWETKSINRAHKTTGTNKTLLISTLHRPMDLHVFHALRQPDVPNHPCKVNNGG
3331
3301 CSNLCLLSPGGGHKCACPTNFYLGSDGRTCVSNCTASQFVCKNDKCIPFWWKCDTEDDCG
3361 DHSDEPPDCPEFKCRPGQFQCSTGICTNPAFICDGDNDCQDNSDEANCDIHVCLPSQFKC
3421 TNTNRCIPGIFRCNGQDNCGDGEDERDCPEVTCAPNQFQCSITKRCIPRVWVCDRDNDCV
3481 DGSDEPANCTQMTCGVDEFRCKDSGRCIPARWKCDGEDDCGDGSDEPKEECDERTCEPYQ
3541 FRCKNNRCVPGRWQCDYDNDCGDNSDEESCTPRPCSESEFSCANGRCIAGRWKCDGDHDC
3601 ADGSDEKDCTPRCDMDQFQCKSGHCIPLRWRCDADADCMDGSDEEACGTGVRTCPLDEFQ
3661 CNNTLCKPLAWKCDGEDDCGDNSDENPEECARFVCPPNRPFRCKNDRVCLWIGRQCDGTD
3778
3721 NCGDGTDEEDCEPPTAHTTHCKDKKEFLCRNQRCLSSSLRCNMFDDCGDGSDEEDCSIDP
3781 KLTSCATNASICGDEARCVRTEKAAYCACRSGFHTVPGQPGCQDINECLRFGTCSQLCNN
3841 TKGGHLCSCARNFMKTHNTCKAEGSEYQVLYIADDNEIRSLFPGHPHSAYEQAFQGDESV
3901 RIDAMDVHVKAGRVYWTNWHTGTISYRSLPPAAPPTTSNRHRRQIDRGVTHLNISGLKMP
3961 RGIAIDWVAGNVYWTDSGRDVIEVAQMKGENRKTLISGMIDEPHAIVVDPLRGTMYWSDW
4021 GNHPKIETAAMDGTLRETLVQDNIQWPTGLAVDYHNERLYWADAKLSVIGSIRLNGTDPI
4081 VAADSKRGLSHPFSIDVFEDYIYGVTYINNRVFKIHKFGHSPLVNLTGGLSHASDVVLYH
4141 QHKQPEVTNPCDRKKCEWLCLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPPPDAPRPGTC
4201 NLQCFNGGSCFLNARRQPKCRCQPRYTGDKCELDQCWEHCRNGGTCAASPSGMPTCRCPT
4261 GFTGPKCTQQVCAGYCANNSTCTVNQGNQPQCRCLPGFLGDRCQYRQCSGYCENFGTCQM
4321 AADGSRQCRCTAYFEGSRCEVNKCSRCLEGACVVNKQSGDVTCNCTDGRVAPSCLTCVGH
4381 CSNGGSCTMNSKMMPECQCPPHMTGPRCEEHVFSQQQPGHIASILIPLLLLLLLVLVAGV
4441 VFWYKRRVQGAKGFQHQRMTNGAMNVEIGNPTYKMYEGGEPDDVGGLLDADFALDPDKPT
4501 NFTNPVYATLYMGGHGSRHSLASTDEKRELLGRGPEDEIGDPLA

Fig. 9b

DE 694 32 744 T2 2004.03.25

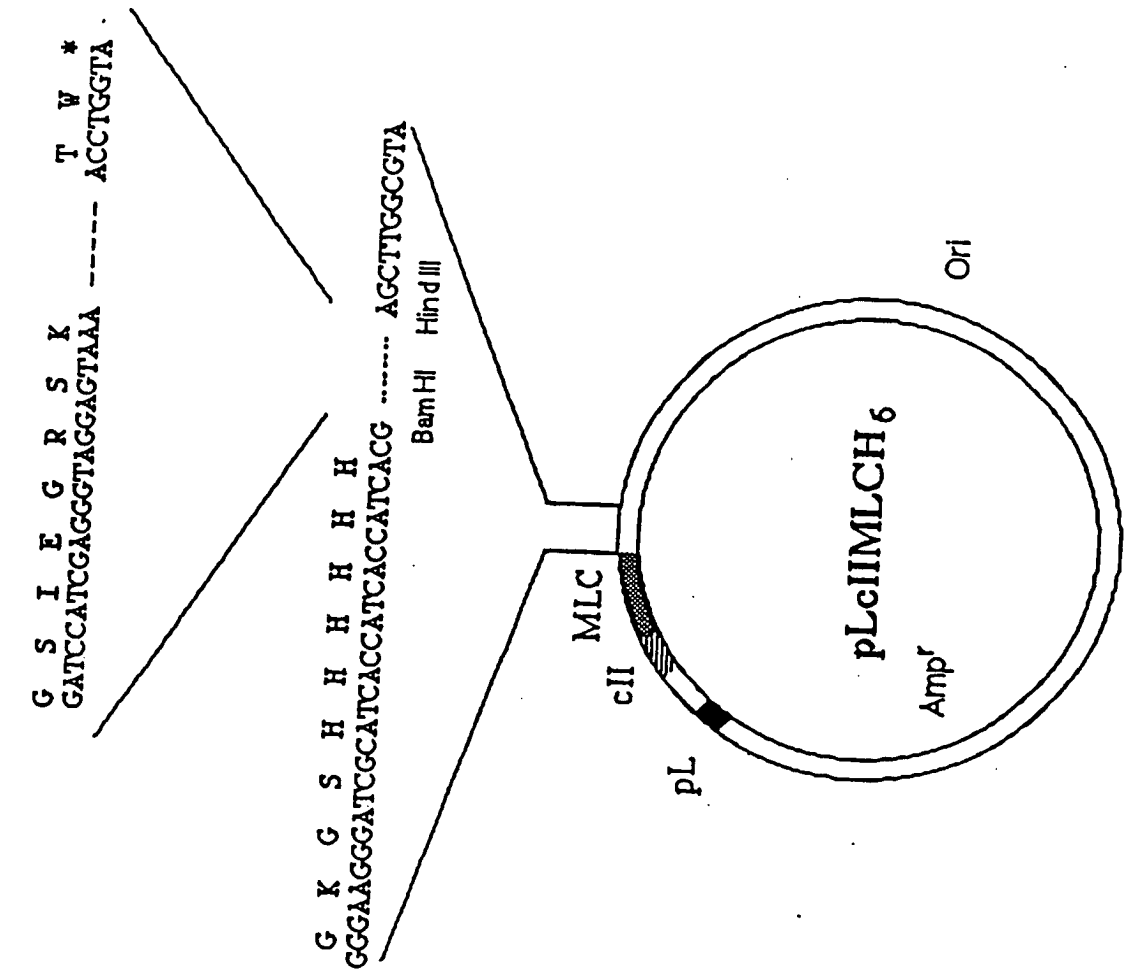


Fig. 10

Rinder-FX

Rinder-FX

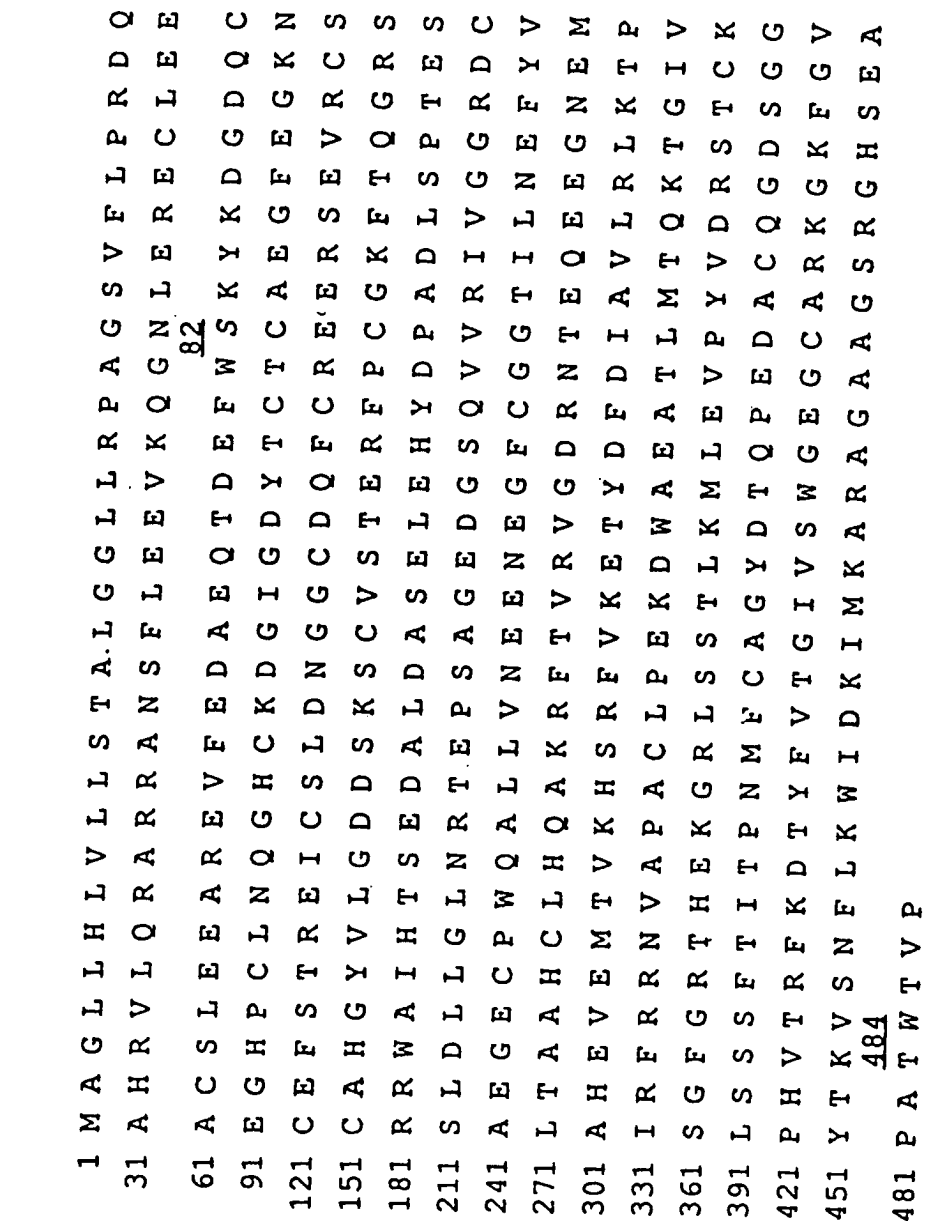


Fig. 11

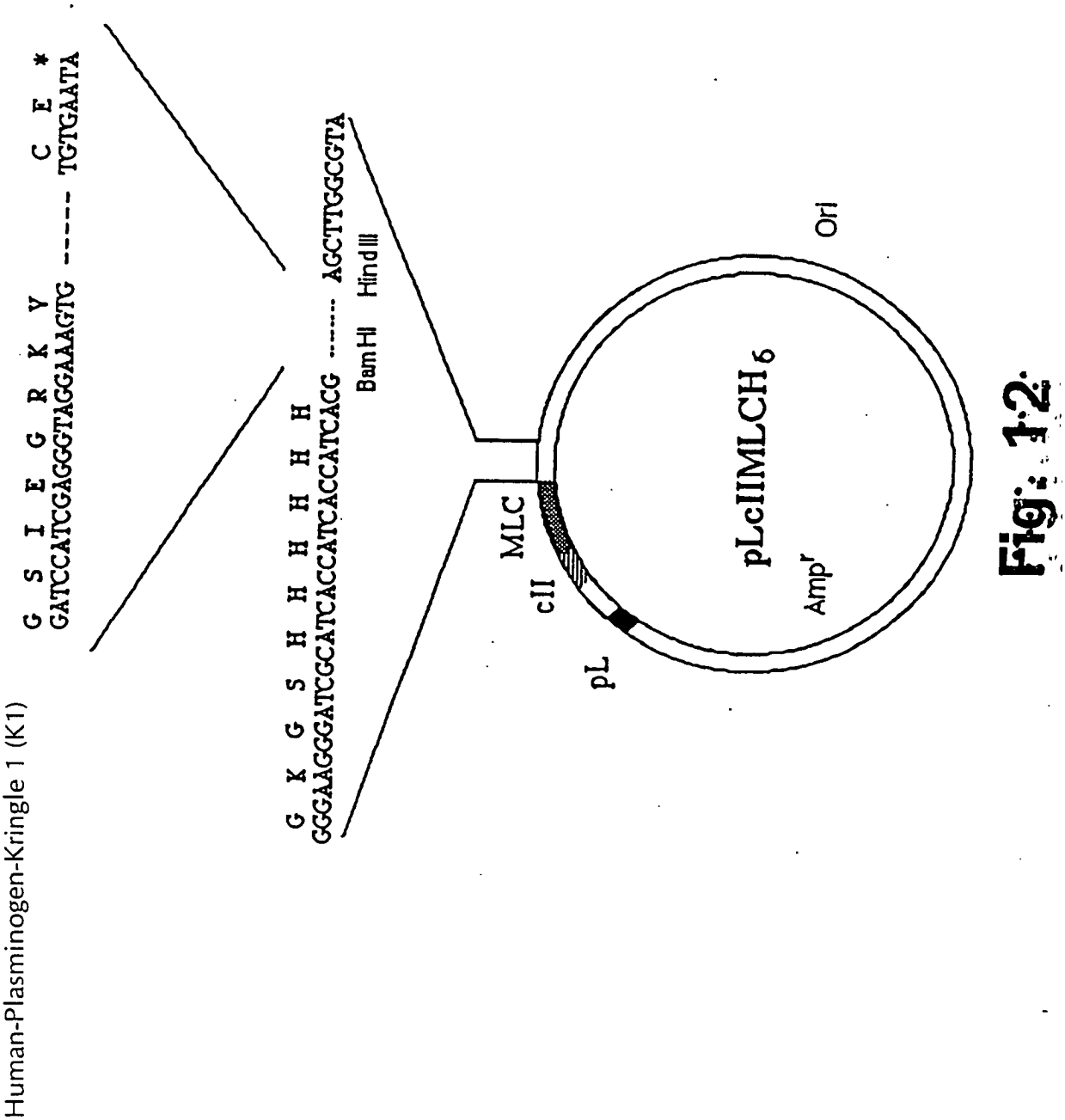

Human-Plasminogen-Kringle 1 (K1)
G S I E G R K Y C E *
GATCCATCGAGGGTAGGAAAGTG ----- TGTGAATA
G K G S H H H H H H
GGGAAGGGATCGCATCACCATCACCATCACG ----- AGCTTGGCGTA
BamHI HindIII
MLC
cII
pL
pLcIIMLCH6
Amp^r
Ori

Fig. 12

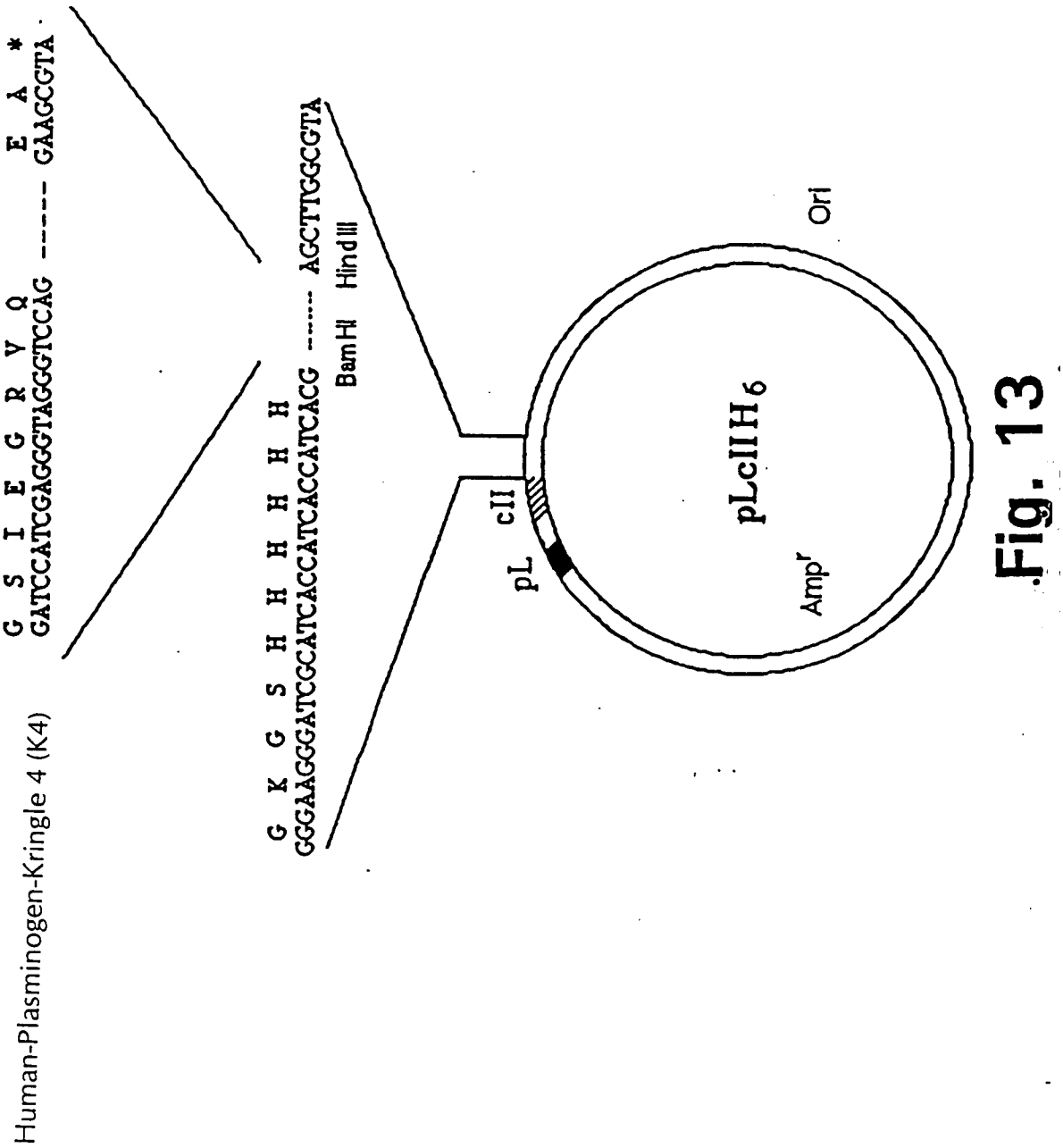

Human-Plasminogen-Kringle 4 (K4)
G S I E G R V Q E A *
GATCCATCGAGGGTAGGGTCCAG ----- GAAGCGTA
G K G S H H H H H H
GGGAAGGGATCGCATCACCATCACCATCACG ----- AGCTTGGCGTA
BamHI HindIII
pL
cII
pLcIIH6
Ori
Ampr

Fig. 13

Glu - Plasminogen.

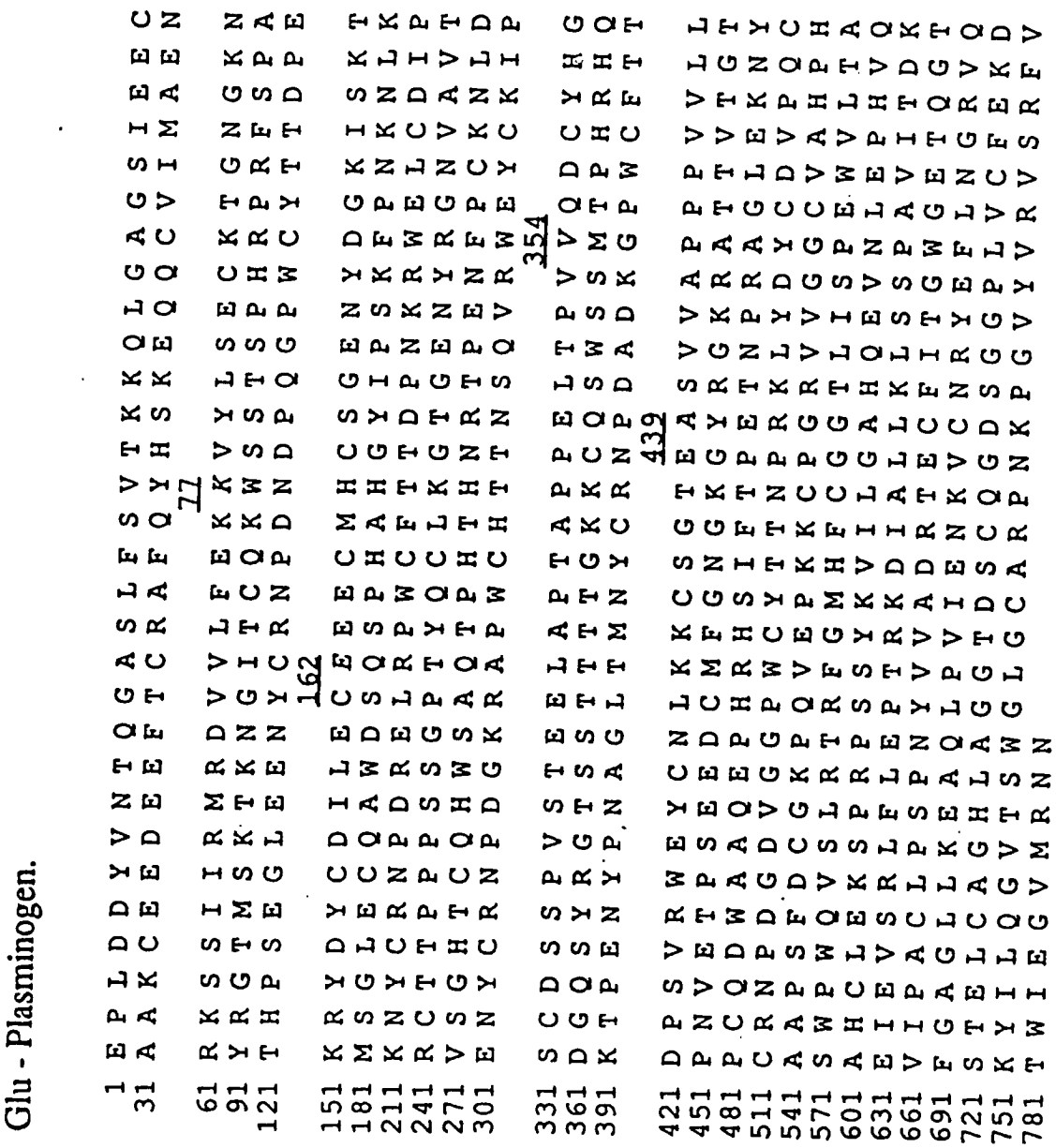


Fig. 14.

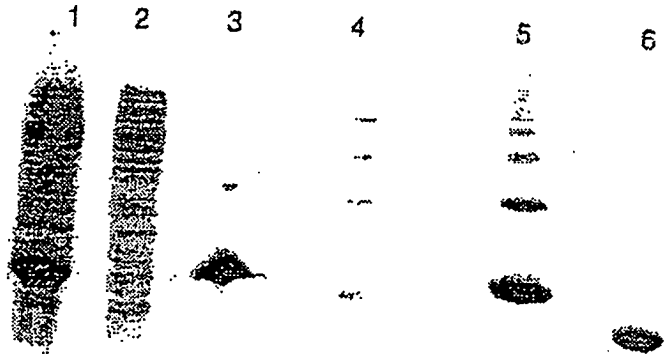


Fig. 15

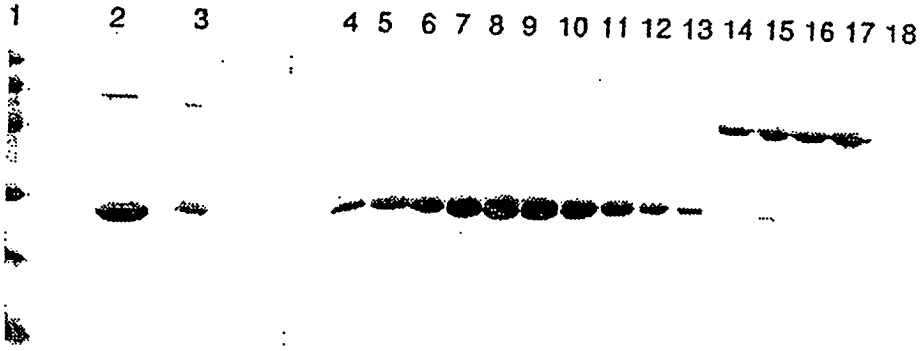


Fig. 16

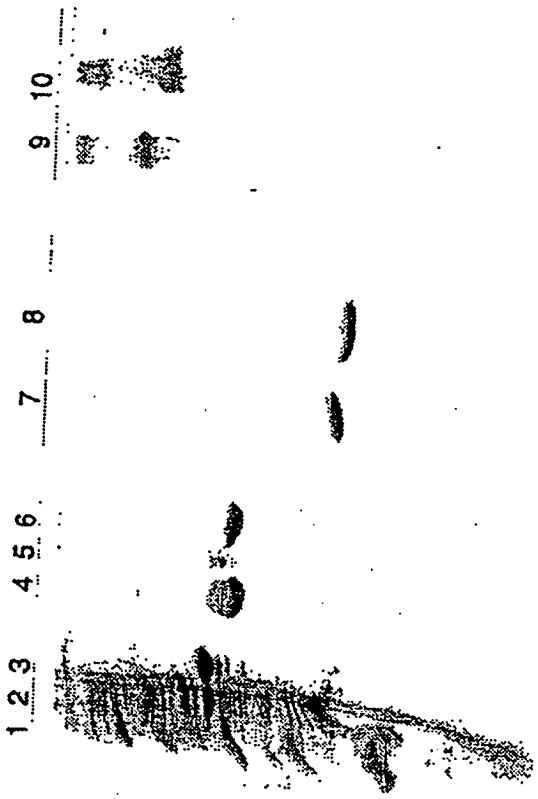


Fig. 17

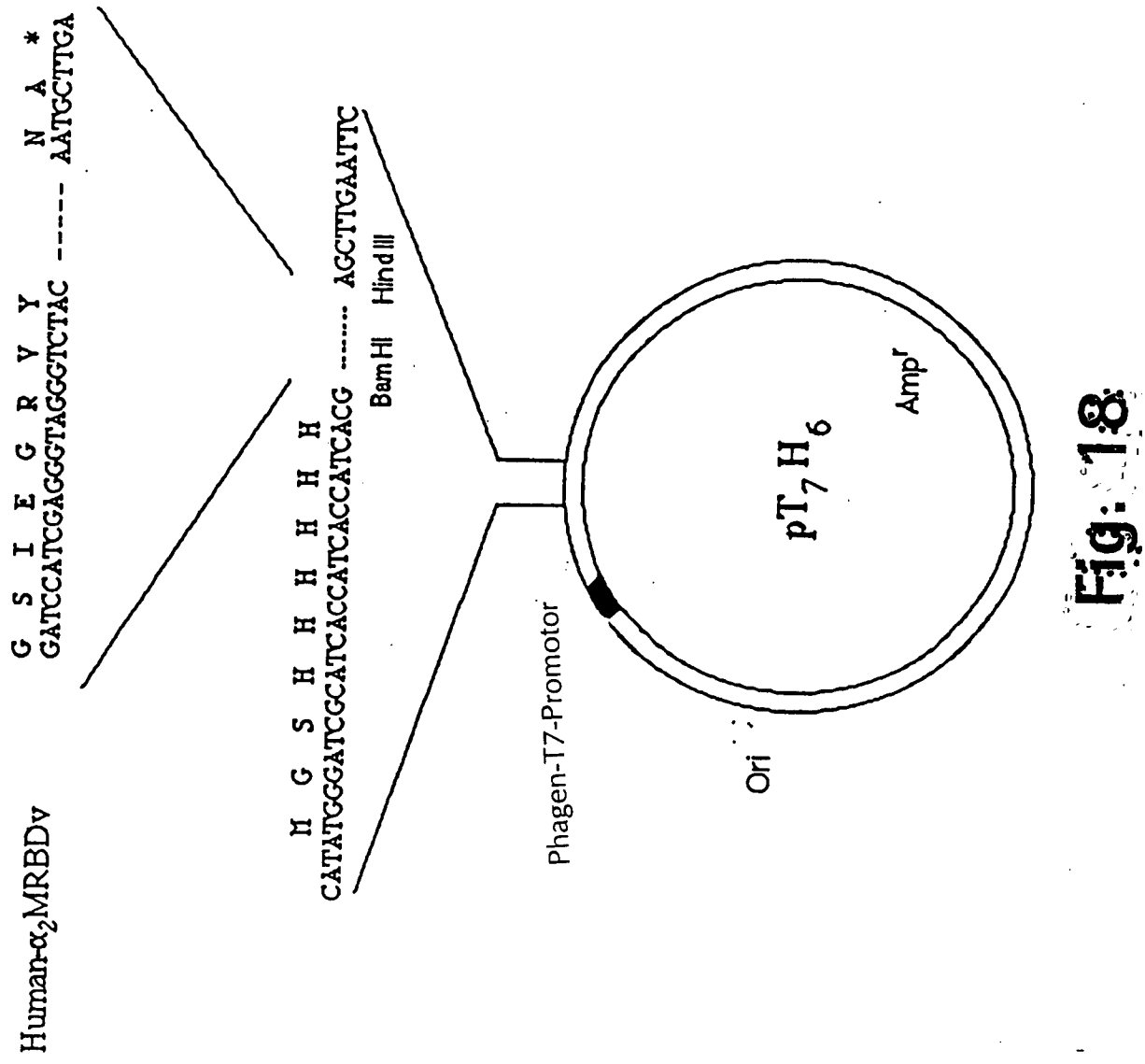

Human-α2MRBDv
G S I E G R V Y N A *
GATCCATCGAGGGTAGGGTCTAC ----- AATGCTTGA
H G S H H H H H H
CATATGGGATCGCATCACCATCACCATCACG ------ AGCTTGAATTC
BamHI HindIII
Phagen-T7-Promotor
Ori
pT7H6
Ampr

Fig. 18

Human-α_2MRBDv.

```
                        10                  20                  30
(1299)V Y L Q T S L K Y N I L P E K E E F P F A L G V Q T L P Q T
                        40                  50                  60
      C D E P K A H T S F Q I S L S V S Y T G S R S A S N M A I V
                        70                  80                  90
      D V K M V S G F I P L K P T V K M L E R S N H V S R T E V S
                       100                 110                 120
      S N H V L I Y L D K V S N Q T L S L F F T V L Q D V P V R D
                       130                 140                 150
      L K P A I V K V Y D Y Y E T D E F A I A E Y N A P C S K D L
                   153
      G N A(1451)
```

Fig. 19

DE 694 32 744 T2 2004.03.25

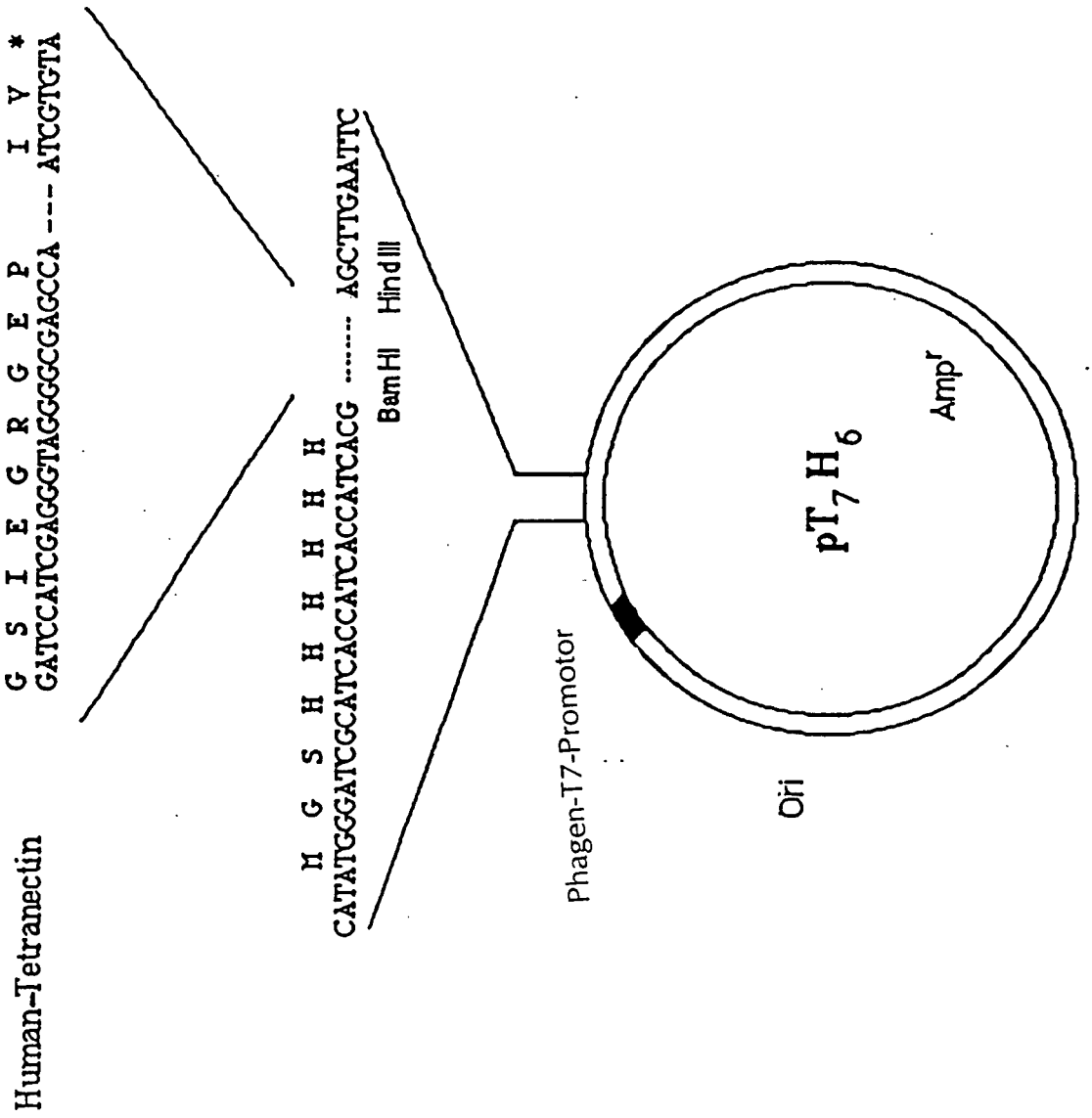

Human-Tetranectin
G S I E G R G E P I V *
GATCCATCGAGGGTAGGGGCGAGCCA --- ATCGTGTA
M G S H H H H H H
CATATGGGATCGCATCACCATCACCATCACG ..... AGCTTGAATTC
BamHI HindIII
Phagen-T7-Promotor
Ori
pT7H6
Ampr

Fig. 20

Human-Tetranectin.

```
-21                                          -1 1                  9
  M E L W G A Y L L L C L F S L L T Q V T T E P P T Q K P K K

 10                  20                  30
  I V N A K K D V V N T K M F E E L K S R L D T L A Q E V A L

 40                  50                  60
  L K E Q Q A L Q T V C L K G T K V H M K C F L A F T Q T K T

 70                  80                  90
  F H E A S E D C I S R G G T L S T P Q T G S E N D A L Y E Y

100                 110                 120
  L R Q S V G N E A E I W L G L N D M A A E G T W V D M T G A

130                 140                 150
  R I A Y K N W E T E I T A Q P D G G K T E N C A V L S G A A

160                 170                 180
  N G K W F D K R C R D Q L P Y I C Q F G I V
```

Fig. 21

DE 694 32 744 T2 2004.03.25

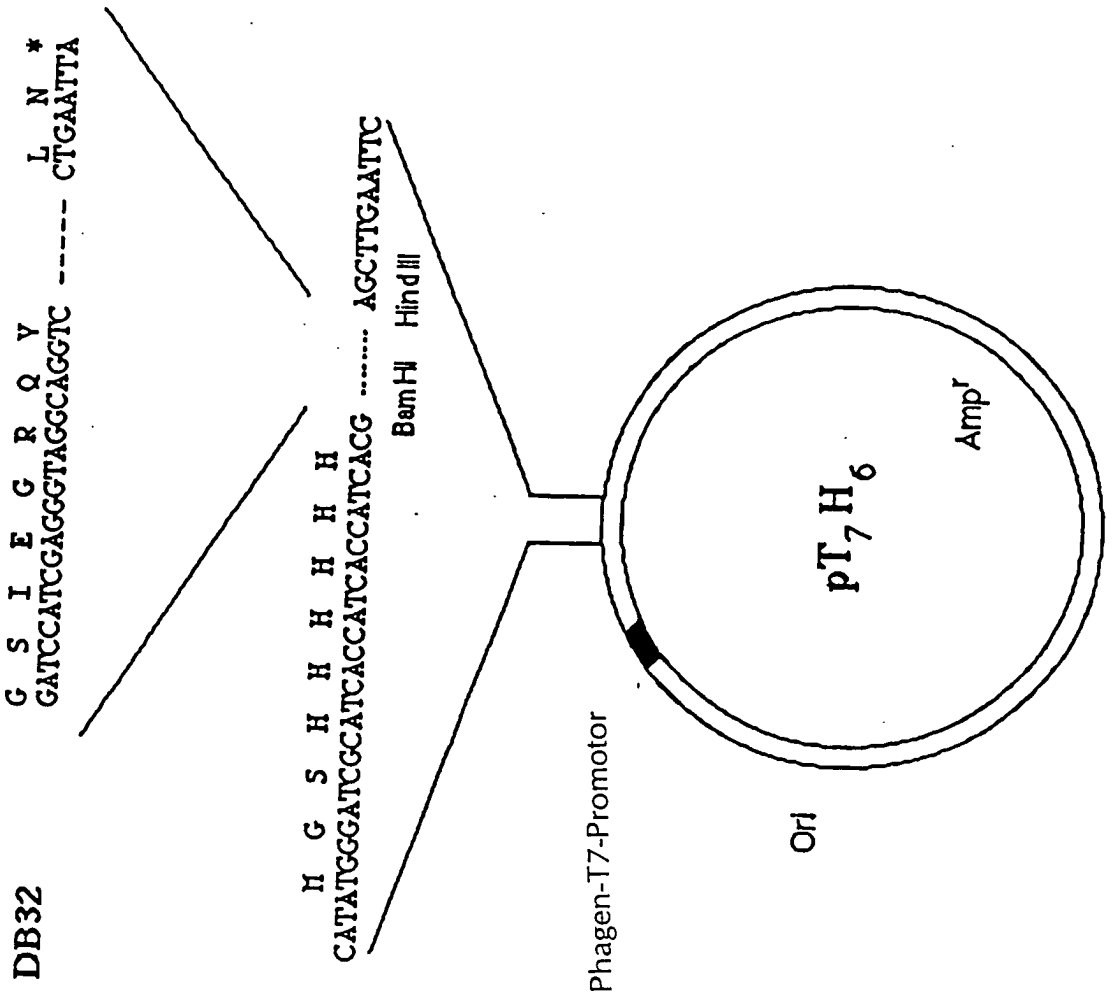

DB32
G S I E G R Q V
GATCCATCGAGGGTAGGCAGGTC ----- L N *
CTGAATTA
M G S H H H H H H
CATATGGATCGCATCACCATCACCATCACG ..... AGCTTGAATTC
BamHI HindIII
Phagen-T7-Promotor
Ori
pT7H6
Ampr

Fig.:22

DB32.

```
  1 Q V K L Q Q S G A E L V K P G A S V K M S C K A S G Y T F A
 31 S Y W I N W V K Q R P G Q G L E W I G H I Y P V R S I T K Y
 61 N E K F K S K A T L T L D T S S S T A Y M Q L S S L T S E D
 91 S A V Y Y C S R G D G S D Y Y A M D Y W G Q G T T V T V S S
121 G G G G S D I E L T Q S P A I L S A S P G G K V T M T C R A
151 S S S V S Y M H W Y Q Q K P G S S P K P W I Y A T S N L A S
181 G V P T R F S G S G T S Y S L T I S R V E A E D A A T Y Y
211 Y C Q Q W S R N P F T F G S G T K L E I K R A A A E Q K L I
241 S E E D L N
```

Fig. 23

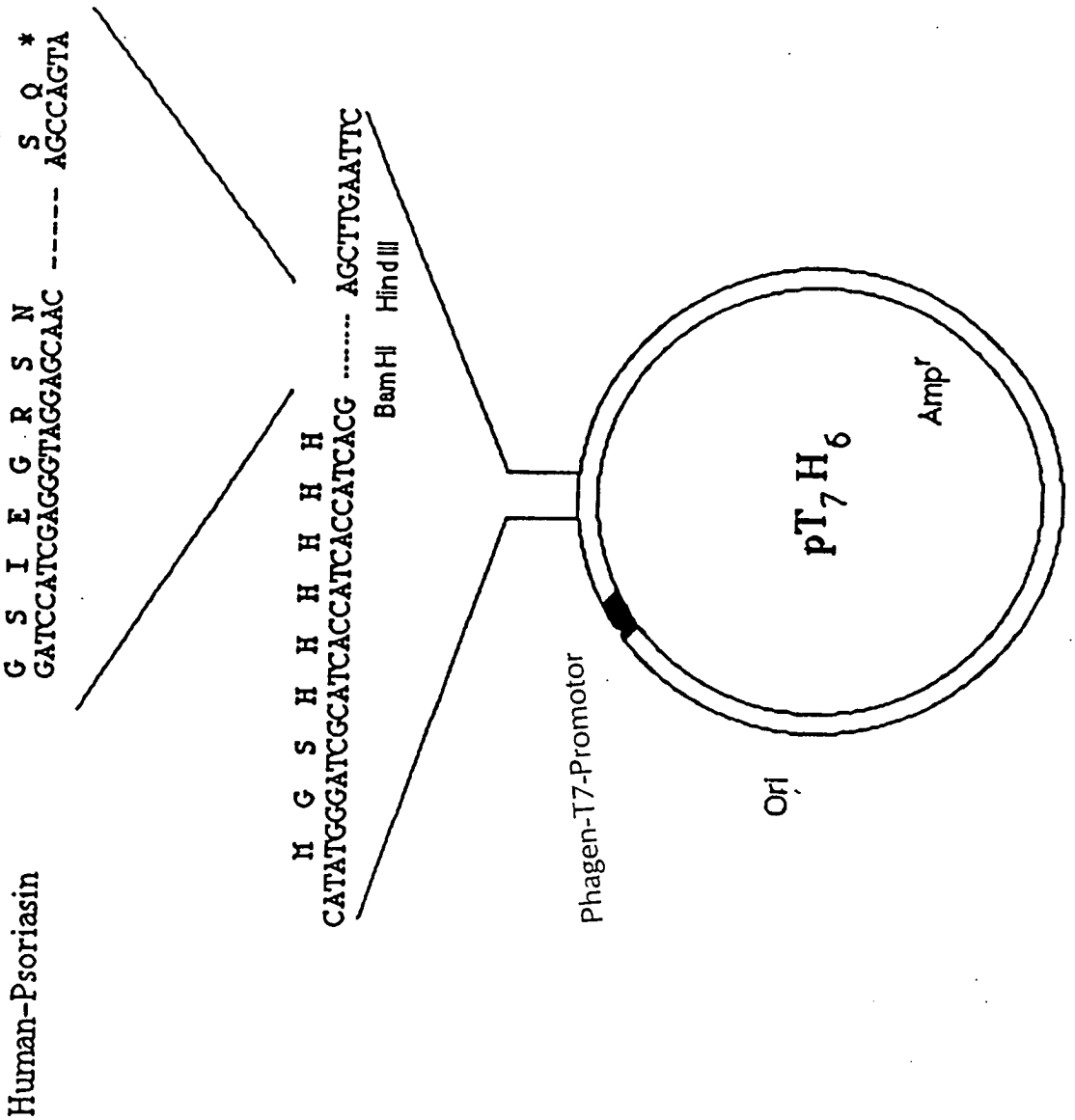

Human-Psoriasin
G S I E G R S N
GATCCATCGAGGGTAGGAGCAAC ----- S Q *
AGCCAGTA
M G S H H H H H H
CATATGGGATCGCATCACCATCACCATCACG ..... AGCTTGAATTC
BamHI HindIII
Phagen-T7-Promotor
Ori
pT7H6
Ampr

Fig. 24

Human-Psoriasin.

```
1                   10                  20                  30
M S N T Q A E R S I I G M I D M F H K Y T R R D D K I D K P

                    40                  50                  60
S L L T M M K E N F P N F L S A C D K K G T N Y L A D V F E

                    70                  80                  90
K K D K N E D K K I D F S E F L S L L G D I A T D Y H K Q S

                   100
H G A A P C S G G S Q
```

Fig. 25

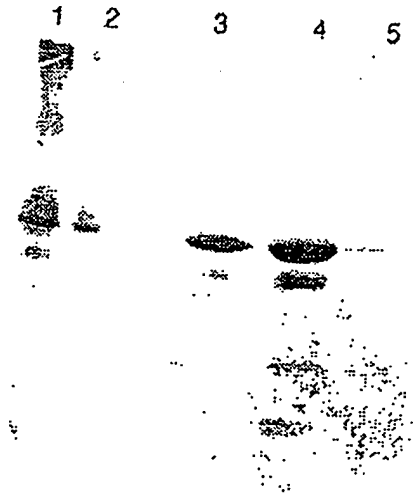


Fig. 26a

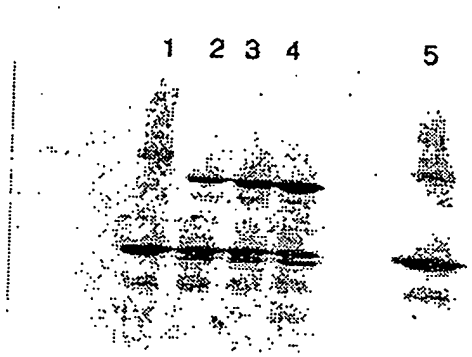


Fig. 26b

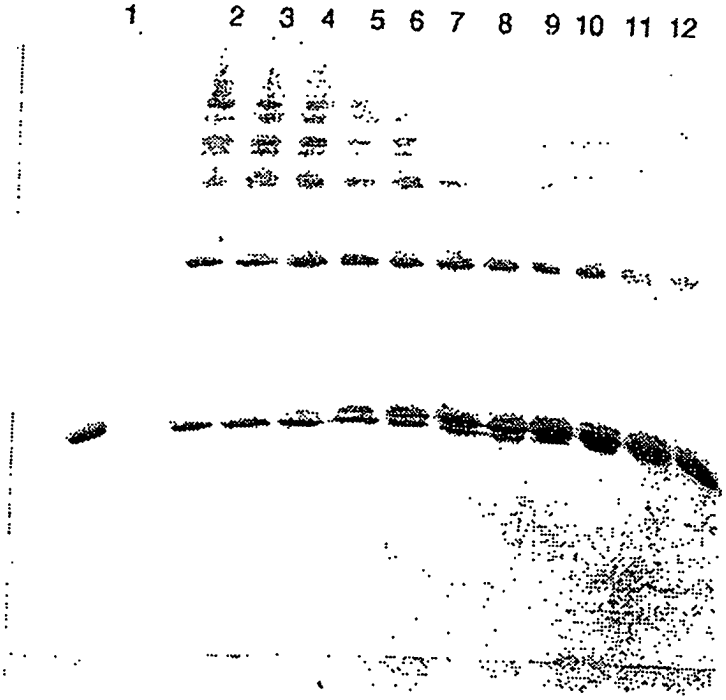


Fig. 27

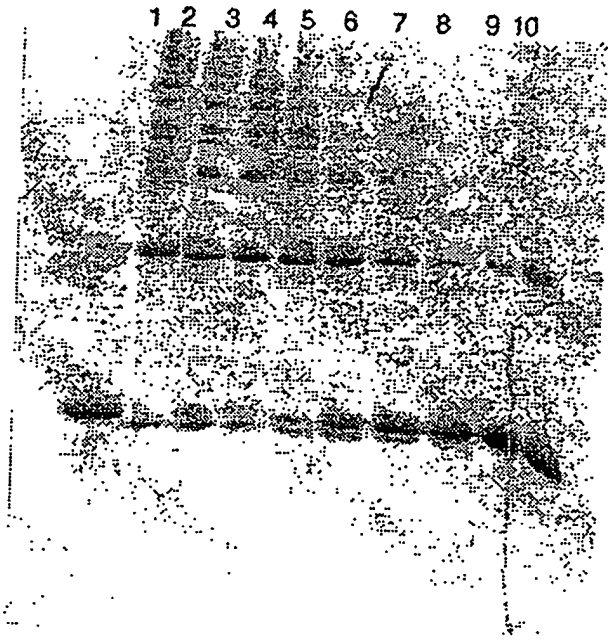


Fig. 28



Fig. 29

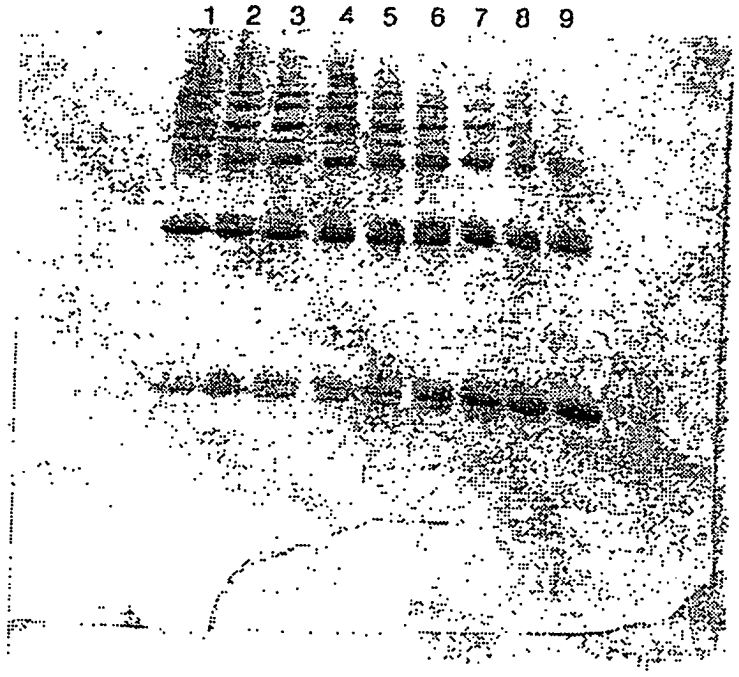


Fig. 30

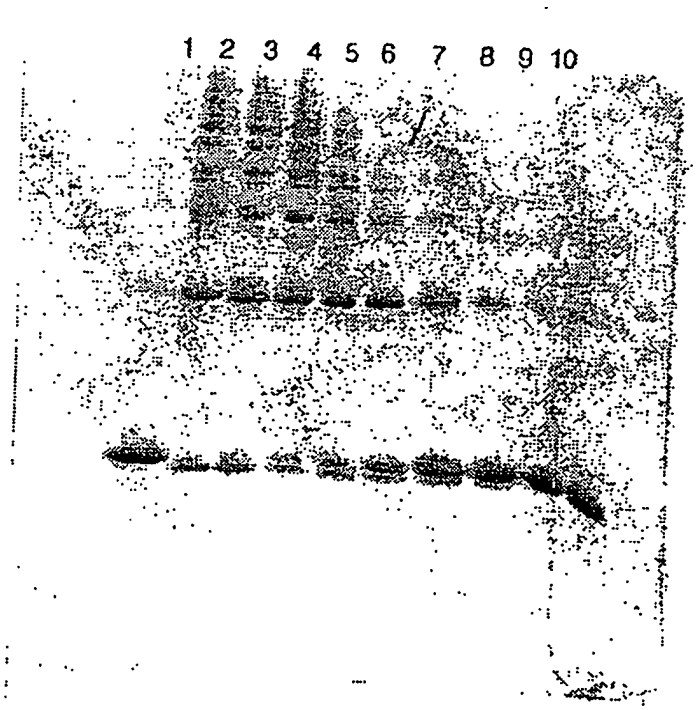


Fig. 31

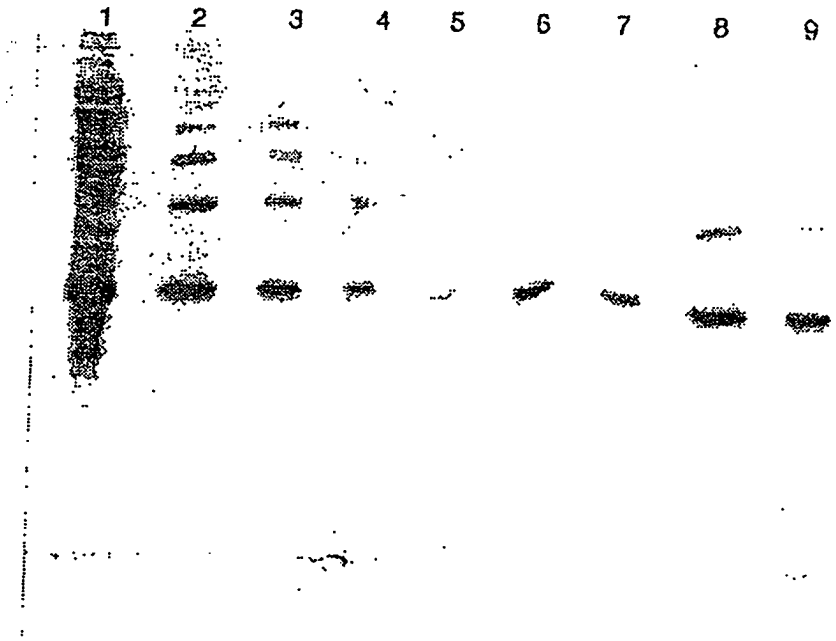


Fig. 32

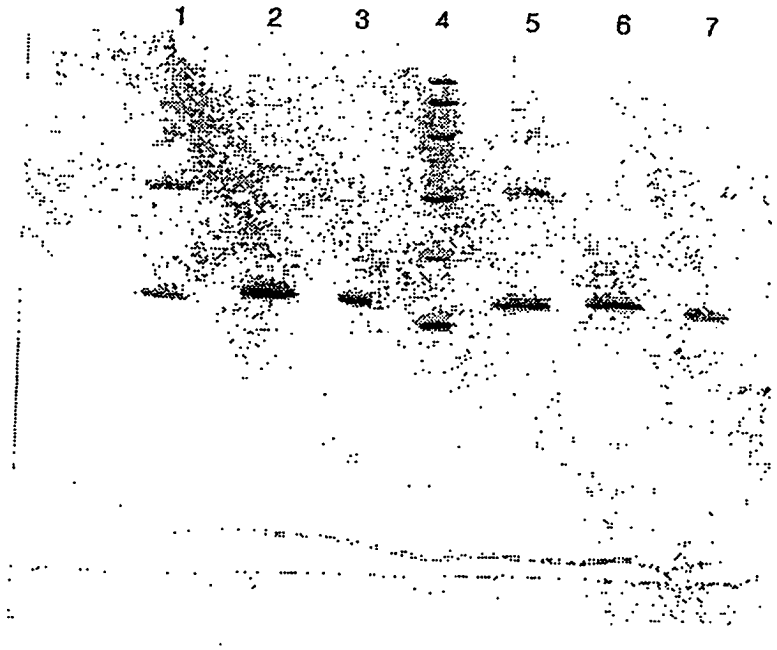


Fig. 33

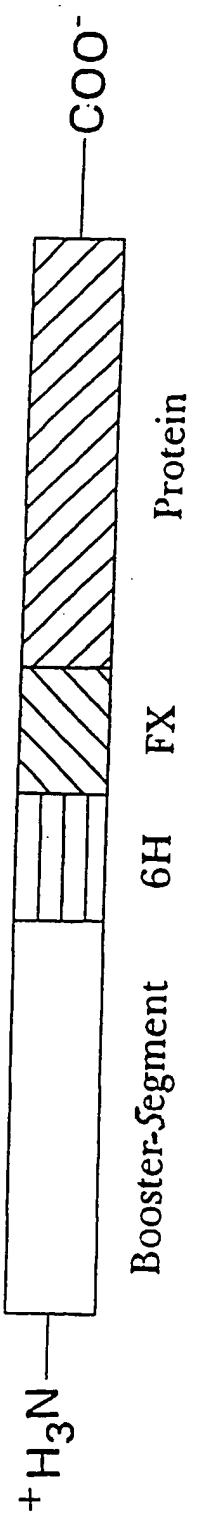


Fig: 34