



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 600 24 659 T2** 2006.08.17

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 222 298 B1**

(21) Deutsches Aktenzeichen: **600 24 659.0**

(86) PCT-Aktenzeichen: **PCT/US00/28431**

(96) Europäisches Aktenzeichen: **00 970 900.7**

(87) PCT-Veröffentlichungs-Nr.: **WO 2001/027299**

(86) PCT-Anmeldetag: **12.10.2000**

(87) Veröffentlichungstag
der PCT-Anmeldung: **19.04.2001**

(97) Erstveröffentlichung durch das EPA: **17.07.2002**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **07.12.2005**

(47) Veröffentlichungstag im Patentblatt: **17.08.2006**

(51) Int Cl.⁸: **C12N 15/85** (2006.01)

C12N 5/10 (2006.01)

C12N 15/67 (2006.01)

C07K 14/715 (2006.01)

(30) Unionspriorität:

159177 P 13.10.1999 US

(73) Patentinhaber:

Immunex Corp., Thousand Oaks, Calif., US

(74) Vertreter:

**Huss und Kollegen, 82467
Garmisch-Partenkirchen**

(84) Benannte Vertragsstaaten:

**AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT,
LI, LU, MC, NL, PT, SE**

(72) Erfinder:

MCGREW, T., Jeffrey, Seattle, US

(54) Bezeichnung: **VEKTOREN UND VERFAHREN FÜR REKOMBINANTE PROTEINEXPRESSION**

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

TECHNISCHES GEBIET DER ERFINDUNG

[0001] Die vorliegende Erfindung betrifft die Expression von rekombinanten Proteinen in eukaryotischen Zellen.

ALLGEMEINER STAND DER TECHNIK

[0002] Die Entwicklung von Expressionssystemen zur Herstellung rekombinanter Proteine ist für das Bereitstellen einer Quelle eines gegebenen Proteins für Forschungs- oder therapeutische Zwecke wichtig. Expressionssysteme sind sowohl für prokaryotische Zellen, wie zum Beispiel *E. coli*, als auch für eukaryotische Zellen, wie zum Beispiel Hefe- (d.h. *Saccharomyces*-, *Pichia*- und *Kluyveromyces* spp.) und Säugetierzellen entwickelt worden. Die Expression in Säugetierzellen wird für die Herstellung von therapeutischen Proteinen oft bevorzugt, da die posttranslatorischen Modifikationen in solchen Expressionssystemen, jenen die auf endogenen Proteinen in einem Säugetier auftreten wahrscheinlich ähnlicher sind als die Art von posttranslatorischen Modifikationen, die in mikrobiellen Expressionssystemen auftreten.

[0003] Mehrere Vektoren stehen zur Expression in Säugetierwirten zur Verfügung, wobei jeder verschiedene Kombinationen von cis- und in einigen Fällen trans-regulatorischen Elementen enthalten, um hohe Konzentrationen an rekombinantem Protein in einem minimalen Zeitrahmen zu erreichen. Jedoch ist trotz der Verfügbarkeit zahlreicher solcher Vektoren das Expressionsniveau eines rekombinanten Proteins, das in Säugetiersystemen erreicht wird, oft niedriger als das mit einem mikrobiellen Expressionssystem erreichte. Weil weiters nur ein kleiner Prozentsatz von geklonten, transfizierten Säugetierzellen hohe Konzentrationen des interessierenden Proteins exprimieren, dauert es oft eine wesentlich längere Zeit zur Entwicklung von nutzbaren, stabil transfizierten Säugetierzelllinien als bei mikrobiellen Systemen.

[0004] Die Verwendung eines dicistronischen Expressionsvektors, wobei ein erster offener Leseraster Polypeptid von Interesse kodiert und ein zweiter offener Leseraster einen selektierbaren Marker kodiert, ist ein Verfahren, das zur Gewinnung von rekombinanten Proteinen verwendet worden ist. Ein bevorzugter Marker zur Verwendung in solchen Systemen ist Dihydrofolatreduktase (DHFR), das den Vorteil besitzt, ein amplifizierbares Gen zu sein, was Selektion nach Zellen erlaubt, die eine hohe Kopienzahl der eingefügten DNA aufweisen, indem sie in ansteigenden Konzentrationen an Methotrexat (MTX) kultiviert werden. Jedoch ist die Translation des selektierbaren Markergens bis zu 100fach weniger effizient als die Translation des interessierenden Gens, was die Effizienz des Selektionsverfahren verringert. Weiters besitzen dicistronische Expressionsvektoren die Neigung, eine Deletion oder eine Umordnung unter Amplifikationsbedingungen in unkontrollierter Weise zu erfahren, wodurch sich die Wahrscheinlichkeit erhöht, dass amplifizierte Zellen nicht mehr das interessierende Protein exprimieren.

[0005] Interne ribosomale-Startstellen (internal ribosome entry site: IRES) sind eine Art regulatorisches Element, die bei mehreren Viren und zellulären RNAs gefunden werden (besprochen in McBratney et al., *Current Opinion in Cell Biology* 5: 961, 1993). IRES erhöhen die Translationseffizienz des selektierbaren Markergens und sind daher zur Erhöhung des Selektions- und des Amplifikationsprozesses nützlich (Kaufmann, R. J., et al., *Nucleic Acids Res.* 19: 4485, 1991). Dennoch zeigen die verfügbaren Beweise an, dass dicistronische mRNAs sich mit einer niedrigeren Konzentration ansammeln als monocistronische mRNAs, möglicherweise wegen der verringerten mRNA-Stabilität der längeren Nachricht. Kaufmann et al., *Nucleic Acid Res.* 19: 4485, 1991, bezieht sich auf verbesserte dicistronische mRNA-Expressionsvektoren durch Nutzung einer mutmaßlichen IRES, die aus dem Encephalomyocarditis-Virus isoliert wurde.

[0006] Da die Menge an rekombinantem Protein, das von einer transfizierten Zelle erzeugt wird, im allgemeinen proportional zur Menge an mRNA ist, die für die Translation des Proteins zur Verfügung steht, kann die Verwendung von dicistronischen Expressionsvektoren in niedrigen Produktionsniveaus des gewünschten rekombinanten Proteins resultieren. Demgemäß besteht im Fachgebiet ein Bedarf, verbesserte Verfahren zu entwickeln, die die nützliche Einrichtung eines selektierbaren, amplifizierbaren Markers, wie zum Beispiel DHFR, beibehalten und gleichzeitig den Anteil an mRNAs erhöhen, die das gewünschte rekombinante Protein codieren. Weiters gibt es einen Bedarf an der Entwicklung von Verfahren, die die Selektion derjenigen Transfektanten erleichtern, die sich in transkriptionell aktive Stellen integrieren und die die Herstellung von zweckmäßigen Konzentrationen an rekombinantem Protein aus Säugetierzellen in einer relativ kurzen Zeitspanne ermöglichen.

ZUSAMMENFASSUNG DER ERFINDUNG

[0007] In einer Ausführungsform der Erfindung weist ein Expressionsvektor eine DNA, die ein erstes Protein kodiert, welches funktionsfähig mit einer DNA verbunden ist, die ein zweites Protein kodiert, auf, wobei eine DNA, die einen Polyadenylierungsstelle (polyA-Stelle) kodiert, zwischen die DNA, die das erste interessierende Protein kodiert, und die DNA, die das zweite Protein kodiert, eingefügt ist, so dass die DNA, die die Polyadenylierungsstelle kodiert, funktionsfähig mit der DNA verbunden ist, die das erste Protein kodiert. Ein bevorzugtes zweites Protein ist ein selektierbarer Marker, vorzugsweise Dihydrofolatreduktase (DHFR); andere amplifizierbare Marker sind ebenfalls zur Verwendung in den erfindungsgemäßen Expressionsvektoren geeignet.

[0008] Das Polyadenylierungssignal, das zur Bereitstellung der internen Polyadenylierungsstelle genutzt wird, ist vorzugsweise ein SV40-Polyadenylierungssignal, insbesondere das späte SV40-Polyadenylierungssignal und ganz besonders die Reste 80 bis 222 der SEQ ID Nr. 1. Die bevorzugten Polyadenylierungssignale werden in der Sequenzaufzählung dargestellt und weiter unten beschrieben. In einer anderen Ausführungsform dieser Erfindung ist das Polyadenylierungssignal induzierbar.

[0009] Der Expressionsvektor kann weiters eine IRES-Sequenz zwischen der DNA, die das erste Protein kodiert, und der DNA, die das zweite Protein kodiert, aufweisen, die funktionsfähig mit beiden verbunden ist und stromabwärts vom Polyadenylierungsort gelegen ist. Alternativ kann der Expressionsvektor mRNA-Spleißdonator- und Akzeptorstellen aufweisen, die im Wesentlichen so wie die von Lucas et al., siehe unten, beschrieben sind.

[0010] Ein weiterer Aspekt dieser Erfindung weist einen Expressionsvektor auf, in den eine DNA, die ein Protein kodiert, eingefügt ist. Solch ein Expressionsvektor weist eine Stelle auf, die in eine DNA, die ein rekombinantes, heterologes Protein kodiert, eingefügt werden kann (wird als Klonierungsstelle bezeichnet), so dass sie funktionsfähig an eine Polyadenylierungsstelle und eine DNA gebunden ist, die ein zweites Protein (wie zum Beispiel einen selektierbaren Marker) kodiert. Gegebenenfalls können auch andere regulatorische Elemente enthalten sein, zum Beispiel eine IRES-Sequenz stromabwärts von der Polyadenylierungsstelle oder mRNA-Spleißdonator- und Akzeptorstellen, die im wesentlichen so wie von Lucas et al., siehe unten, beschrieben sind, welche funktionsfähig mit der Polyadenylierungsstelle und der DNA, die das zweite Protein kodiert, verbunden sind. Ein expressionsverstärkendes Sequenzelement (EASE) kann ebenfalls stromaufwärts von der Klonierungsstelle, funktionsfähig damit verbunden, enthalten sein.

[0011] Wirtszellen können mit den Expressionsvektoren der Erfindung transfiziert werden. Dementsprechend stellt eine andere Ausführungsform dieser Erfindung eine transfizierte Wirtszelle bereit. Ebenfalls bereitgestellt werden Zelllinien, die aus den Wirtszellen kloniert wurden. Bevorzugte Wirtszellen sind Säugetierzellen. In einer stark bevorzugten Ausführungsform sind die Wirtszellen CHO-Zellen.

[0012] Die Erfindung stellt auch ein Verfahren zum Gewinnen eines rekombinanten Proteins bereit, welches die Transfizierung einer Wirtszelle mit einem Expressionsvektor der Erfindung, die Kultivierung der transfizierten Wirtszelle unter Bedingungen, die die Expression des Proteins fördern, und die Rückgewinnung des Proteins aufweist. In einer bevorzugten Anwendung dieser Erfindung werden transfizierte Wirtszellen durch zwei Selektionsschritte ausgewählt, wobei der erste Schritt der Selektion von Zellen, die das dominante Markergen exprimieren, und der zweite der Selektion hoher Expressionsniveaus und/oder der Amplifikation des Markergens sowie des interessierenden Gens dient. In einer stark bevorzugten Ausführungsform ist das Selektions- oder das Amplifikationsmittels Methotrexat, ein Inhibitor von DHFR, von dem gezeigt wurde, dass es eine Amplifikation von endogenen DHFR-Genen und transfizierten DHFR-Sequenzen bewirkt.

KURZBESCHREIBUNG DER ZEICHNUNGEN

[0013] [Fig. 1](#) ist eine grafische Darstellung von Konstrukten, die in Beispiel 1 hergestellt wurden. Das Konstrukt, in welches das frühe SV40-Polyadenylierungssignal aufgenommen wurde, wurde als SPA6 bezeichnet; jenes, in welches das späte Polyadenylierungssignal aufgenommen wurde, wurde als SPA4 bezeichnet. Ein Kontrollkonstrukt wurde als BGH bezeichnet.

AUSFÜHRLICHE BESCHREIBUNG DER ERFINDUNG

[0014] Expressionsvektoren, die den Nutzen eines selektierbaren, amplifizierbaren Markers, wie zum Beispiel DHFR, behalten, während sie gleichzeitig den Anteil von mRNAs, die ein gewünschtes rekombinantes Protein codieren, erhöhen, werden hierin bereitgestellt. Die Expressionsvektoren der Erfindung weisen ein Polyadeny-

lierungssignal auf, das zwischen eine erste kodierende Sequenz und eine zweite oder nachfolgende kodierende Sequenz (die als interne Polyadenylierungsstelle bezeichnet wird) eingefügt ist. In den Vektoren der Erfindung können Transkripte, die am Promotor ihren Ursprung haben, nach der ersten kodierenden Sequenz (monocistronische Mitteilung) oder nach der zweiten oder nachfolgenden kodierenden Sequenz (multicistronische Mitteilung) polyadenyliert werden. In einer Ausführungsform der Erfindung kodiert die erste kodierende Sequenz ein interessierendes Protein, und die zweite (oder nachfolgende) kodierende Sequenz kodiert einen selektierbaren Marker. In einer anderen Ausführungsform folgt eine zweite Polyadenylierungsstelle der zweiten oder nachfolgenden kodierenden Sequenz und ist damit funktionsfähig verbunden. In dieser Ausführungsform wird die interne Polyadenylierungsstelle daher zur ersten Polyadenylierungsstelle.

[0015] Weil viele Transkripte nur das interessierende Gen und nicht den selektierbaren Marker codieren, produzieren die Vektoren der Erfindung weniger selektierbares Markerprotein, und nur diejenigen Transfektanten, die sich in stärker transfectionsaktive Stellen integrieren, überstehen das Selektionsverfahren. Dementsprechend erleichtert die Verwendung der Vektoren der Erfindung das Isolieren der transfizierten Pools und Klone, die hohe Konzentrationen an rekombinantem Protein exprimieren, wobei niedrigere Konzentrationen eines Selektionsmittels, als bei Fehlen des internen Polyadenylierungsortssignals möglich ist, verwendet werden.

[0016] Ein weiterer Nutzen der Verwendung der Expressionsvektoren der Erfindung ist, dass monocistronische Mitteilungen stabiler oder effizienter bearbeitet werden können als dicistronische Mitteilungen, was möglicherweise zur erhöhten Ansammlung der Mitteilung führt, die das interessierende Protein kodiert, und daher zu einem höheren Niveau der Proteinproduktion. Die Verwendung der internen Polyadenylierungsstelle der Erfindung erleichtert daher die Erzeugung brauchbarer Konzentrationen an rekombinantem Protein durch transfizierte Zellen in einer relativ kurzen Zeitspanne.

[0017] Die Vektoren und Verfahren der Erfindung sind auch in der Entwicklung von multicistronischen Vektoren nützlich. Die multicistronischen Expressionsvektoren ermöglichen die koordinierte Expression von zwei oder mehr Genen (siehe zum Beispiel Fussenegger et al., Biotechnol. Prog. 13: 733, 1997). Das Einfügen einer Polyadenylierungsstelle nach dem ersten Cistron würde zu einem hohen Expressionsniveau des ersten Cistrons und zu einem niedrigeren Expressionsniveau aller folgenden Cistrons führen. Potenzielle Anwendungen dieser Technologie bestünden darin, die Expression von großen Mengen eines therapeutischen Proteins (oder anderer, gewünschter rekombinanter Proteine) und geringerer Mengen anderer Proteine, wie zum Beispiel selektierbare Marker, Transkriptionsfaktoren, Enzyme, die an der Proteinfaltung beteiligt sind, und anderer Proteine, die den Zellstoffwechsel und die Expression regulieren, zu erleichtern.

[0018] In einer anderen Ausführungsform wird die Polyadenylierungsstelle nach dem zweiten oder dritten (oder nachfolgenden) Cistron eingefügt. Dies würde eine hohe Expression der ersten beiden (oder drei oder mehr) Cistrons ermöglichen, gefolgt von geringerer Expression des Cistrons, das der Polyadenylierungsstelle folgt. Diese Ausführungsform findet zum Beispiel in der rekombinanten Antikörper-Synthese Anwendung, wo die schweren und leichten Ketten unabhängig voneinander auf hohem Niveau synthetisiert werden. Ein tricistronischer Vektor wird mit den schweren und leichten Ketten konstruiert, die von den ersten zwei Cistrons kodiert werden. Die Polyadenylierungsstelle wird nach dem zweiten Cistron eingefügt, was eine der Expression der ersten zwei Cistrons auf hohem Niveau ermöglicht. Der selektierbare Marker wird vom dritten Cistron (d.h., nach der Polyadenylierungsstelle) und auf niedrigeren Niveau exprimiert.

Expression von rekombinanten Proteinen

[0019] Wie hierin verwendet, soll der Begriff 'Expressionsvektor' einen Vektor beschreiben, der verschiedene regulatorische, im Detail unten beschriebene Elemente aufweist, die für die Expression von rekombinanten, heterologen Proteinen in Zellen notwendig sind. Der Expressionsvektor kann Signale enthalten, die für die Erhaltung in prokaryotischen oder eukaryotischen Zellen geeignet sind, und/oder der Expressionsvektor kann in ein Chromosom integriert werden.

[0020] Rekombinante Expressionsvektoren können eine kodierende Sequenz enthalten, die ein interessierendes Protein (oder ein Fragment davon), Ribozyme, ribosomale mRNAs, Antisense-RNAs und dergleichen kodieren. Die kodierende Sequenz kodiert vorzugsweise ein Protein oder Peptid. Die kodierende Sequenz kann synthetisch, eine aus cDNA abgeleitete Nucleinsäuresequenz oder ein Nucleinsäurefragment sein, das durch die Polymerasekettenreaktion (PCR) isoliert wurde. Die kodierende Sequenz wird funktionsfähig mit geeigneten transkriptionalen oder translatorischen regulatorischen Elementen verbunden, die aus Säugetier-, Virus- oder Insektengen abgeleitet wurden. Solche regulatorischen Elemente schließen einen transkriptionalen Promotor, eine Sequenz, die geeignete ribosomale mRNA-Bindungsstellen kodiert, und Sequenzen, die die

Termination von Transkription und Translation kontrollieren (d.h. ein Polyadenylierungssignal) ein, wie im Detail unten beschrieben wird.

[0021] Expressionsvektoren können auch nicht transkribierte Elemente, wie zum Beispiel einen geeigneten Promotor und/oder Enhancer, der mit dem zu exprimierenden Gen verbunden ist, andere 5'- oder 3'-flankierende nicht transkribierte Sequenzen, 5'- oder 3'-nichttranslatierte Sequenzen, wie zum Beispiel Ribosomenbindungsstellen, eine Polyadenylierungsstelle, Spleiß-Donator- und Akzeptorstellen, und transkriptionale Terminationssequenzen, aufweisen. Ein Replikationsursprung, der die Fähigkeit zum Replizieren in einem Wirt verleiht, und ein selektierbares Gen zur Erleichterung der Erkennung von Transfektanten können ebenfalls enthalten sein.

[0022] DNA-Regionen sind funktionsfähig verbunden, wenn sie funktionell zueinander in Beziehung stehen. Zum Beispiel ist DNA für ein Signalpeptid (sekretorische Vorsequenz) funktionsfähig mit DNA für ein Polypeptid verbunden, wenn sie als Präkursor exprimiert wird, der an der Sekretion des Polypeptids beteiligt ist; daher bedeutet 'funktionsfähig verbunden' im Fall der DNA, die sekretorische Vorsequenzen kodiert, sich berührend und in einem Leseraster. Ein Promotor ist funktionsfähig mit einer kodierenden Sequenz verbunden, wenn er die Transkription der Sequenz kontrolliert; und eine Ribosomenbindungsstelle ist funktionsfähig mit einer kodierenden Sequenz verbunden, wenn sie so angeordnet ist, dass sie die Translation ermöglicht.

[0023] Dicistronische Expressionsvektoren, die für die Expression mehrerer Transkripte verwendet werden, sind bereits beschrieben worden (Kim S. K. und Wold B. J., Cell 42: 129, 1985; Kaufman et al., 1991, siehe oben). Dicistronische Expressionsvektoren weisen zwei Cistrons oder offene Leseraster, wobei sie in der Lage sind, zwei Proteine, zum Beispiel ein rekombinantes interessierendes und einen selektierbaren Marker, zu codieren. Ein Beispiel für einen solchen dicistronischen Expressionsvektor ist pCAVDHFR, ein Derivat von pCD302 (Mosley et al., Cell 1989), der die kodierende Sequenz für Maus-DHFR enthält (Subramani et al., Mol. Cell. Biol. 1: 854, 1981). Ein weiteres Beispiel für einen solchen dicistronischen Expressionsvektor ist der pC-DE-Vektor, ein Derivat von pCAVDHFR, der die interne ribosomale Startstellen (IRES) des Maus-Encephalomyocarditis-Virus enthält (Nukleotide 260 bis 824; Jang und Wimmer, Genes and Dev. 4: 1560, 1990), welcher zwischen die dreiteiligen Adenovirus-Vorsequenz und die DHFR-cDNA kodierenden Sequenz kloniert wird. Andere Arten von Expressionsvektoren sind ebenfalls in Kombination mit der Erfindung nützlich, zum Beispiel die in den US-Patenten 4,634,665 (Axel et al.) und 4,656,134 (Ringold et al.) beschriebenen.

[0024] Die transkriptionalen und translatorischen Kontrollsequenzen in Expressionsvektoren, die in transfizierten Zellen zu verwenden sind, können aus viralen Quellen bereitgestellt werden. Zum Beispiel werden häufig verwendete Promotoren und Enhancer aus Polyomavirus, Adenovirus 2, Simian-Virus 40 (SV40) und menschlichem Zytomegalievirus abgeleitet. Virale genomische Promotoren, Kontroll- und/oder Signalsequenzen können dazu verwendet werden, die Expression voranzutreiben, vorausgesetzt, solche Kontrollsequenzen sind mit der gewählten Wirtszelle kompatibel. Beispiele für solche Vektoren können so konstruiert werden, wie von Okayama und Berg (Mol. Cell. Biol. 3: 280, 1983) offenbart wurde. Nichtvirale zelluläre Promotoren können ebenfalls verwendet werden (d.h. das Beta-Globin und die EF-1 alpha-Promotoren), abhängig vom Zelltyp, in dem das rekombinante Protein exprimiert werden soll.

[0025] DNA-Sequenzen, die aus dem SV40-Virusgenom abgeleitet werden, zum Beispiel SV40-Ursprung, früher und später Promotor, Enhancer, Spleiß- und Polyadenylierungsstellen, können dazu verwendet werden, die anderen genetischen Elemente bereitzustellen, die für die Expression einer heterologen DNA-Sequenz benötigt werden. Die frühen und späten Promotoren sind besonders nützlich, weil sie beide leicht aus dem Virus als Fragment zu gewinnen sind, das auch den SV40-Replikationsursprung enthält (Fiers et al., Nature 273: 113, 1978). Kleinere oder größere SV40-Fragmente können ebenfalls verwendet werden, vorausgesetzt, die etwa 250 bp große Sequenz, die sich von der HindIII-Stelle zur BglII-Stelle, der sich im viralen Replikationsursprung befindet, hin erstreckt, ist enthalten.

[0026] In dicistronischen Expressionsvektoren wird oft eine Polyadenylierungsstelle, die stromabwärts vom zweiten Cistron (normalerweise eine DNA, die einen selektierbaren Marker kodiert) eingefügt ist und mit diesem funktionsfähig verbunden ist, dazu verwendet, die Transkription und Translation zu regulieren. Viele solcher Polyadenylierungssignale sind bekannt (siehe zum Beispiel Tabelle 1 unten). Die vorliegende Erfindung verwendet ein internes Polyadenylierungssignal, z.B. eines, das zwischen die beiden Cistrons eines dicistronischen Expressionsvektors eingefügt wird, zusätzlich zu einem Polyadenylierungssignal oder einem anderen regulatorischen Element stromabwärts vom zweiten Cistron.

[0027] Sowohl die frühen als auch die späten Polyadenylierungssignale von SV40 sind in der vorliegenden

Erfindung anwendbar. Diese Sequenzen werden innerhalb des 237 Basenpaar großen Fragmentes zwischen der BamHI-Stelle bei Nukleotid 2533 und der BclI-Stelle bei Nukleotid 2770 des SV40-Genoms kodiert (Carswell und Alwine, Mol. Cell. Biol., 9: 4248, 1989). Carswell und Alwine kamen zum Schluss, dass von den zwei SV40-Polyadenylierungssignalen das späte Signal wirksamer war, höchstwahrscheinlich, weil es sowohl stromabwärts als auch stromaufwärts gelegene Sequenzelemente aufweist, die die wirksame Spaltung und Polyadenylierung erleichtern.

[0028] Viele Polyadenylierungssignale sind im Fachgebiet bekannt und sind auch in der vorliegenden Erfindung anwendbar. Beispiele sind u.a. die in Tabelle 1 unten angeführten.

Tabelle 1. Polyadenylierungssignale

Spätes SV40-polyA und Deletionsmutanten davon	Schek, N, Cook, C. and J. C. Alwine (1992): Definition of the upstream efficiency element of the simian virus 40 late polyadenylation signal by using in vitro analysis. Mol. Cell Biol. 12: 5386–5393
HIV-1-poly A	Klasens, B. I. F., Das, A. T. and B. Berkhout (1998): Inhibition of polyadenylation by stable RNA secondary structure. Nucleic Acids Res. 26: 1870–1876
β-Globin-poly A)	Gil, A., and N. J. Proudfoot (1987): Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit β-globin mRNA formation. Cell 49: 399–406
HSV TK poly A	Cole, C. N. and T. P. Stacy (1985): Identification of sequence in the herpes simplex virus thymidine kinase gene required for efficient processing and polyadenylation. Mol. Cell. Biol. 5: 2104–2113
Polyomavirus-poly A	Batt, D. B., and G. G. Carmichael (1995): Characterization of the polyomavirus late polyadenylation signal. Mol. Cell. Biol. 15: 4783–4790
Rinderwachstumshormon	Gimmi, E. R., Reff, M. E. and I. C. Deckman (1989): Alterations in pre-mRNA topology of the bovine growth hormone polyadenylation region decrease polyA site efficiency. Nucleic Acids Res. 17: 6983–6998

[0029] Zusätzliche Polyadenylierungsstellen können mit Verfahren, die im Fachgebiet bekannt sind, identifiziert oder konstruiert werden. Eine minimale Polyadenylierungsstelle besteht aus AAUAAA und einer zweiten Erkennungssequenz, im allgemeinen einer G/U-reichen Sequenz, die etwa 30 Nukleotide stromabwärts zu finden ist. In der Sequenzauflistung werden Sequenzen als DNA dargestellt und nicht als RNA, um die Herstellung geeigneter DNAs zum Einbau in Expressionsvektoren zu erleichtern. Bei Darstellung als DNA besteht die Polyadenylierungsstelle aus AATAAA, mit zum Beispiel einer G/T-reichen Region stromabwärts (siehe zum Beispiel die Nukleotide 123 bis 128 bzw. 151 bis 187 der SEQ ID Nr. 1).

[0030] Beide Sequenzen müssen vorhanden sein, um eine effiziente Polyadenylierungsstelle zu bilden. Der Zweck dieser Stellen ist es, spezifische RNA bindende Proteine die an die RNA binden zu rekrutieren. Das AAUAAA bindet den Spaltungspolyadenylierungsspezifitätsfaktor (CPSF; Murthy, K. G., und Manley, J. I. (1995), Genes Dev 9: 2672–2683), und die zweite Stelle, häufig eine G/U-Sequenz, bindet sich an den spaltungsstimulierenden Faktor (CstF; Takagaki, Y., und Manley, J. L. (1997) Mol Cell Biol. 17: 3907–3914). CstF besteht aus mehreren Proteinen, das für die RNA-Bindung verantwortliche Protein ist jedoch CstF-64, ein Mitglied der Ribonucleoprotein-Domänenfamilie von Proteinen (Takagaki et al. (1992) Proc. Natl. Acad. Sci. USA 89: 1403–1407).

[0031] Es ist gezeigt worden, dass die Konzentration von CstF-64-Protein wichtig bei der regulierenden Verwendung verschiedener Polyadenylierungsstellen in B-Zellen ist (Takagaki, Y., und Manley, J. L. (1998) Mol Cell 2: 761–771). Dementsprechend kann eine induzierbare Polyadenylierungsstelle auf der Grundlage dieser natürlich vorkommenden Regulierung des Polyadenylierungsgebrauchs in B-Zellen konstruiert werden, indem die Wechselwirkung von CstF-64 mit einer mRNA nach Wahl kontrolliert wird, um die Polyadenylierung zu induzieren. Zum Beispiel kann das CstF-64 mit der RNA-bindenden Domäne des MS2-Phagen-Hüllproteins fusioniert sein, welches eine spezifische RNA-Sequenz bindet (ACAUGAGGAUUACCCAUGU; SEQ ID Nr. 4), die sich von dem G/U-reichen Element unterscheidet (Lowary und Uhlenbeck (1987) Nucleic Acids Res. 15:

10483–10493). Die Ziel-mRNA enthält normalerweise eine AAUAAA-Sequenz und eine MS2-Hüllprotein-RNA-Erkennungssequenz. Durch Regulieren der Konzentration des MS2-CstF-64-Fusionsproteins auf transkriptionalem Wege unter Verwendung von induzierbaren Standardexpressionssystemen (zum Beispiel eines Ecdysone-induzierbaren Säugetieryexpressionssystems, das von No et al. (1996) Proc. Natl. Acad. Sci. USA 93: 3346–3351, beschrieben wurde), könnte die Verwendung der induzierbaren poly A-Stelle kontrolliert werden.

[0032] Die Polyadenylierung kann auch durch Entwicklung von temperaturempfindlichen MS2-RNA-bindenden Domänenmutanten reguliert werden. MS2-RNA-bindende Domänenmutanten können unter Einsatz der Zufallsmutagenese erzeugt und auf Temperaturempfindlichkeit durchsucht werden. Bei Verwendung als Fusionspartner mit CstF-64, wie oben beschrieben, wäre das temperaturempfindliche MS2-Hüllprotein inaktiv und würde RNA bei 37°C nicht binden; daher würde die interne poly A-Stelle bei dieser Temperatur nicht funktionieren. Bei verringerter Temperatur jedoch, zum Beispiel bei 32°C, wäre das MS2-Hüllprotein aktiv, würde das RNA-Sequenzziel erkennen und die Mitteilung würde polyadenyliert. Die Temperaturregulierung wäre besonders für die rekombinante Proteinexpression nützlich, da die Verringerung der Temperatur von Expressionskulturen normalerweise dazu verwendet wird, die Proteinexpression zu erhöhen.

[0033] Ein weiteres Verfahren, das in Verbindung mit den Vektoren der Erfindung verwendet werden kann, wird von Lucas et al. (Nucleic Acids Res. 24: 1774, 1996) beschrieben. In dem Bemühen, die Produktion eines gewünschten Proteins zu erhöhen, nutzten Lucas et al. mRNA-Spleiß-Donator- und Akzeptorstellen zur Entwicklung von stabilen Klonen, die sowohl selektierbaren Marker als auch rekombinante Proteine erzeugten. Laut diesen Forschern führten die Vektoren, die sie herstellten, zur Transkription eines hohen Anteils von mRNA, die das gewünschte Protein kodiert, und zu einem festen, relativ niedrigen Konzentration des Selektionsmarkers, der die Selektion von stabilen Transfektanten ermöglichte.

Wirtszellen

[0034] Transfizierte Wirtszellen sind Zellen, die mit heterologer DNA transfiziert wurden (was manchmal als "transformiert" bezeichnet wird). Viele Verfahren zum Transfizieren von Zellen sind bekannt; bei einem Ansatz werden Zellen mit Expressionsvektoren transfiziert, die unter Verwendung von rekombinanten DNA-Techniken konstruiert werden und die Sequenzen enthalten, welche rekombinante Proteine codieren. Exprimierte Proteine werden vorzugsweise in den Kulturüberstand sekretiert, können aber mit der Zellmembran assoziiert sein, abhängig vom speziellen Polypeptid, das exprimiert wird. Säugetier-Wirtszellen werden für die vorliegende Erfindung bevorzugt. Verschiedene Säugetierzellkultursysteme können zum Exprimieren von rekombinantem Protein eingesetzt werden. Beispiele für geeignete Säugetierwirtszelllinien sind u.a. die COS-7-Linien von Affenfarnierenzellen, die von Gluzman (Cell 23: 175, 1981) beschrieben wurden, CV-1/EBNA (ATCC CRL 10478), L-Zellen, C127, 3T3, Ovarialzellen des Chinesischen Hamsters (CHO), HeLa- und BHK-Zelllinien.

[0035] Eine häufig verwendete Zelllinie sind DHFR-CHO-Zellen, die auxotroph für Glycin, Thymidin und Hypoxanthin sind und in den DHFR+-Phänotyp mittels DHFR-cDNA als amplifizierbarem dominantem Marker transformiert werden können. Eine solche DHFR-CHO-Zelllinie, DXB11, wurde von Urlaub und Chasin (Proc. Natl. Acad. Sci. USA 77: 2416, 1980) beschrieben. Ein weiteres Beispiel einer DHFR-CHO-Zelllinie ist DG44 (siehe zum Beispiel Kaufman, R. J., Meth. Enzymology 185: 537 (1988)). Andere Zelllinien, die für spezifische Selektions- oder Amplifikationsprogramme entwickelt wurden, sind ebenfalls in der Erfindung nützlich.

[0036] Zahlreiche andere eukaryotische Zellen sind ebenfalls in der vorliegenden Erfindung nützlich, einschließlich Zellen von anderen Wirbeltieren, und Insektenzellen. Die Fachleute auf diesem Gebiet sind in der Lage, geeignete Vektoren, regulatorische Elemente, Transfektions- und Kulturprogramme entsprechend den Bedürfnissen ihres bevorzugten Kultursystems auszuwählen.

Herstellung von transfizierten Säugetierzellen

[0037] Mehrere Transfektionsprotokolle sind im Fachgebiet bekannt und werden in Kaufman, R. J. oben, besprochen. Das gewählte Transfektionsprotokoll hängt von der Art der Wirtszelle und der Natur des interessierenden Proteins ab und kann auf der Grundlage von Routineexperimentierarbeiten gewählt werden. Die grundlegenden Anforderungen an ein solches Protokoll sind erstens, DNA, die das interessierende Protein kodiert, in eine geeignete Wirtszelle einzuführen, und dann die Wirtszellen, die die heterologe DNA in einer stabilen, exprimierbaren Weise aufgenommen haben, zu identifizieren und zu isolieren.

[0038] Ein häufig verwendetes Verfahren zur Einführung von heterologer DNA ist die Kalziumphosphatpräzi-

pitation, wie zum Beispiel wie von Wigler et al. (Proc. Natl. Acad. Sci. USA 77: 3567, 1980) beschrieben wurde. DNA, die mit diesem Verfahren in eine Wirtszelle eingeführt wurde, erfährt häufig eine Umordnung, was dieses Verfahren nützlich für die Co-Transfektion unabhängiger Gene macht.

[0039] Die polyethyleninduzierte Fusion von Bakterienprotoplasten mit Säugetierzellen (Schaffner et al., Proc. Natl. Acad. Sci. USA 77: 2163, 1980) ist ein weiteres nützliches Verfahren zur Einführung von heterologer DNA. Protoplastenfusionsprotokolle ergeben häufig mehrere Kopien der Plasmid-DNA, die in das Säugetier-Wirtszellgenom integriert ist. Dieses Verfahren erfordert, dass der Selektions- und Amplifikationsmarker sich auf demselben Plasmid wie das interessierende Gen befindet.

[0040] Die Elektroporation kann ebenfalls dazu verwendet werden, DNA direkt in das Zytoplasma einer Wirtszelle einzuführen, wie von Potter et al. (Proc. Natl. Acad. Sci. USA 81: 7161, 1988) oder Shigekawa und Dower (BioTechniques 6: 742, 1988) beschrieben wird. Anders als die Protoplastenfusion erfordert die Elektroporation nicht, dass sich der Selektionsmarker und das interessierende Gen auf demselben Plasmid befinden.

[0041] Vor kurzem wurden mehrere Reagenzien beschrieben, die für die Einführung von heterologer DNA in eine Säugetierzelle nützlich sind. Dazu gehören das Lipofection® Reagenz und das Lipofectamine™-Reagenz (Gibco BRL, Gaithersburg, MD). Beide Reagenzien sind handelsübliche Reagenzien, die zur Bildung von Lipid-Nucleinsäure-Komplexen (oder Liposomen) verwendet werden, welche bei Anwendung auf kultivierte Zellen die Aufnahme der Nucleinsäure in die Zellen erleichtern.

[0042] Die Transfektion von Zellen mit heterologer DNA und die Selektion auf Zellen, die die heterologe DNA aufgenommen haben und den selektierbaren Marker exprimieren, führt zu einem Pool an transfizierten Zellen. Einzelne Zellen in diesen Pools werden sich in der Menge der aufgenommenen DNA und in der chromosomalen Lage der transfizierten DNA unterscheiden. Nach wiederholtem Durchgang verlieren die Pools häufig die Fähigkeit, das heterologe Protein zu exprimieren. Um stabile Zelllinien zu erzeugen, können einzelne Zellen aus den Pools isoliert und kultiviert werden (ein Verfahren, das als Klonieren bezeichnet wird), ein arbeits- und zeitaufwendiges Verfahren. In einigen Fällen können jedoch die Pools selbst stabil sein (d.h. die Produktion des heterologen rekombinanten Proteins bleibt stabil). Die Fähigkeit, solche stabilen Pools an Zellen zu selektieren und zu kultivieren, wäre wünschenswert, da sie die schnelle Herstellung von relativ großen Mengen von rekombinantem Protein aus Säugetierzellen ermöglichen würden.

[0043] Ein Verfahren zum Amplifizieren des interessierenden Gens ist ebenfalls wünschenswert für die Expression des rekombinanten Proteins und beinhaltet normalerweise die Verwendung eines Selektionsmarkers (was in Kaufman, R. J., oben, besprochen wird). Die Widerstandsfähigkeit gegenüber zytotoxischen Medikamenten ist das Kennzeichen, das am häufigsten als Selektionsmarker verwendet wird und kann das Ergebnis entweder eines dominanten Merkmals (d.h. es kann unabhängig vom Wirtszelltyp verwendet werden) oder eines rezessiven Merkmals (d.h. nützlich in besonderen Wirtszelltypen, denen die Aktivität fehlt, auf die gerade selektiert wird) sein. Mehrere amplifizierbare Marker sind für die Verwendung in den Expressionsvektoren der Erfindung geeignet (wie zum Beispiel in Maniatis, Molecular Biology: A Laboratory Manual, Cold Spring Harbor Laboratory, NY, 1989, S. 16.9–16.14, beschrieben ist).

[0044] Nützliche selektierbare Marker für die Genamplifikation in wirkstoff-resistenten Säugetierzellen werden in Tabelle 1 von Kaufman, R. J., oben, gezeigt und beinhalten DHFR-MTX-Resistenz, P-Glycoprotein- und multiple Wirkstoff-Resistenz Medikamentenbeständigkeit (MDR) – verschiedene lipophile zytotoxische Mittel (d.h. Adriamycin, Colchicin, Vincristin) und Adenosindeaminase (ADA)-Xyl-A oder Adenosin und 2'-Desoxycoformycin. Spezifische Beispiele von Genen, die selektierbare Marker kodieren, sind diejenigen, die die Antimetabolitenresistenz kodieren, wie zum Beispiel das DHFR-Protein, welches die Resistenz gegenüber Methotrexat vermittelt (Wigler et al., 1980, Proc. Natl. Acad. Sci. USA 77: 3567; O'Hare et al., 1981, Proc. Natl. Acad. Sci. USA 78: 1527); das GPT-Protein, welches die Resistenz gegenüber Mycophenolsäure vermittelt (Mulligan und Berg, 1981, Proc. Natl. Acad. Sci. USA 78: 2072); der Neomycin-Resistenzmarker, der die Resistenz gegenüber dem Aminoglycosid G-418 vermittelt (Colberre-Garapin et al., 1981, J. Mol. Biol. 150: 1); das Hygro-Protein, welches die Resistenz gegenüber Hygromycin vermittelt (Santerre et al., 1984, Gene 30: 147) und der Zeocin™-Resistenzmarker (der kommerziell von Invitrogen bezogen werden kann). Weiters können die Gene für die Herpes simplex-Virus-Thymidinkinase (Wigler et al., 1977, Cell 11: 223), Hypoxanthin-Guanin-Phosphoribosyltransferase (Szybalska und Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48: 2026) und Adeninphosphoribosyltransferase (Lowy et al., 1980, Cell 22: 817) in tk-, hgprr- bzw. aprt-Zellen eingesetzt werden.

[0045] Andere dominante selektierbare Marker sind u.a. mikrobiell abgeleitete Antibiotikaresistenzgene, zum Beispiel Neomycin-, Kanamycin- oder Hygromycin-Resistenz. Diese Selektionsmarker haben sich jedoch nicht

als amplifizierbar herausgestellt (Kaufman, R. J., oben). Es gibt mehrere geeignete Selektionssysteme für Säugetierwirte (Maniatis, oben, S. 16.9–16.15). Cotransfektionsprotokolle, die zwei dominante selektierbare Marker verwenden, sind ebenfalls beschrieben worden (Okayama und Berg, Mol Cell Biol 5: 1136, 1985).

[0046] Eine besonders nützliches Selektions- und Amplifikationsprogramm nutzt die DHFR-MTX-Resistenz. MTX ist ein Inhibitor von DHFR, von dem gezeigt wurde, dass es die Amplifikation von endogenen DHFR-Genen (Alt, F. W., et al., J. Biol Chem 253: 1357, 1978) und transfizierten DHFR-Sequenzen verursacht (Wigler, M., et al., Proc. Natl. Acad. Sci. USA 77: 3567, 1980). Zellen werden mit DNA transfiziert, die das interessierende Gen und DNA umfasst, welche DHFR in einer dicistronischen Expressionseinheit kodiert, aufweist (Kaufman et al., 1991, oben, und Kaufman, R. J., et al., EMBO J 6: 187, 1987). Transfizierte Zellen werden in Medien gezüchtet, die zunehmend höheren Mengen von MTX enthalten, was zu einer größeren Expression des DHFR-Gens sowie des interessierenden Gens führt.

[0047] Nützliche regulatorische Elemente, die vorher beschrieben wurden, können ebenfalls in die Plasmide oder Expressionsvektoren, die zum Transfizieren von Säugetierzellen verwendet werden, aufgenommen werden. Das gewählte Transfektionsprotokoll und die Elemente, die zur Verwendung darin ausgewählt werden, hängen von der Art der verwendeten Wirtszelle ab. Fachleute auf diesem Gebiet kennen zahlreiche verschiedene Protokolle und Wirtszellen und können ein geeignetes System für die Expression eines gewünschten Proteins auf der Grundlage der Anforderungen an ihr(e) gewählte(s/n) Zellkultursystem(e) auswählen.

Verwendungen der Erfindung

[0048] Die Vektoren und Verfahren der Erfindung werden für die Expression einer großen Vielzahl von rekombinanten Polypeptiden angewendet. Beispiele für solche Polypeptide sind u.a. Cytokine und Wachstumsfaktoren, wie zum Beispiel die Interleukine 1 bis 18, die Interferone, RANTES, Lymphotoxin- β , Fas-Ligand, flt-3-Ligand, der Ligand für den Rezeptoraktivator von NF-kappa B (RANKL), der TNF-bezogene apoptoseinduzierende Ligand (TRAIL), CD40-Ligand, Ox40-Ligand, 4-1BB-Ligand (und andere Mitglieder der TNF-Familie), vom Thymusstroma abgeleitetes Lymphopoietin, Granulozytenkolonie-stimulierender Faktor, Granulozytenmakrophagenkolonie-stimulierender Faktor, Mastzellwachstumsfaktor, Stammzellenwachstumsfaktor, epidermaler Wachstumsfaktor, Wachstumshormon, Tumornekrosefaktor, Leukämiehemmungsfaktor, Oncostatin-M und hämopoietische Faktoren, wie zum Beispiel Erythropoietin und Thrombopoietin.

[0049] Ebenfalls enthalten sind neurotrophe Faktoren, wie zum Beispiel der vom Hirn abgeleitete neurotrophe Faktor, der ziliare neurotrophe Faktor, der aus glialen Zelllinien abgeleitete neurotrophe Faktor und verschiedene Liganden für die Zelloberflächenmoleküle Elk und Hek (wie zum Beispiel die Liganden für eph-verwandte Kinasen oder LERKS). Beschreibungen von Proteinen, die gemäß den Verfahren der Erfindung exprimiert werden können, sind zum Beispiel in Human Cytokines: Handbook for Basic and Clinical Research, Vol. II (Aggawal und Gutterman, eds., Blackwell Sciences, Cambridge, MA, 1998); Growth Factors: A Practical Approach (McKay und Leigh, eds., Oxford University Press Inc., New York, 1993) und The Cytokine Handbook (AW Thompson, ed., Academic Press, San Diego, CA, 1991) zu finden.

[0050] Rezeptoren für jedes der vorher genannten Proteine können auch unter Verwendung der Vektoren und Verfahren der Erfindung exprimiert werden, einschließlich beider Formen des Tumornekrosefaktors (die als p55 und p75 bezeichnet werden), der Interleukin-1-Rezeptoren (Typ 1 und 2), des Interleukin-4-Rezeptors, des Interleukin-15-Rezeptors, des Interleukin-17-Rezeptors, des Interleukin-18-Rezeptors, des Granulozytenmakrophagenkolonie-stimulierenden Faktor Rezeptors, des Granulozytenkolonie-stimulierenden Faktor Rezeptors, der Rezeptoren für Oncostatin-M und des Leukämiehemmungsfaktors, des Rezeptoraktivators von NF-kappa B (RANK), der Rezeptoren für TRAIL und Rezeptoren, die Todesdomänen aufweisen, wie zum Beispiel Fas oder der apoptoseinduzierende Rezeptor (AIR).

[0051] Andere Proteine, die mittels der Vektoren und Verfahren der Erfindung exprimiert werden können, schließen Gruppen von Differenzierungsantigenen ein (die als CD-Proteine bezeichnet werden), zum Beispiel die in Leukocyte Typing VI (Proceedings of the VIth International Workshop and Conference; Kishimoto, Kikutani et al., eds., Kobe, Japan, 1996) offenbart, oder CD-Moleküle, die in nachfolgenden Workshops offenbart wurden. Beispiele für solche Moleküle sind u.a. CD27, CD30, CD39, CD40; sowie Liganden dafür (CD27-Ligand, CD30-Ligand und CD40-Ligand). Mehrere davon sind Mitglieder der TNF-Rezeptor-Familie, die auch 4-1BB und OX40 einschließen; die Liganden sind oft Mitglieder der TNF-Familie (wie zum Beispiel der 4-1BB-Ligand und der OX40-Ligand); dementsprechend können die Mitglieder der TNF- und TNFR-Familien ebenfalls unter Verwendung der vorliegenden Erfindung exprimiert werden.

[0052] Proteine, die enzymatisch aktiv sind, können ebenfalls gemäß der vorliegenden Erfindung exprimiert werden. Beispiele schließen die Mitglieder der Metallproteinase-Disintegrin-Familie, verschiedene Kinasen (einschließlich Streptokinase und Gewebsplasminogenaktivator sowie die todesassoziierte Kinase, die wiederholte Ankyrin-Sequenzen enthält, und IKR 1 und 2), TNF- α Converting-Enzym und zahlreiche andere Enzyme ein. Liganden für enzymatisch aktive Proteine können ebenfalls durch Anwenden der vorliegenden Erfindung exprimiert werden.

[0053] Die Vektoren und Verfahren der Erfindung sind auch zur Expression anderer Arten von rekombinanten Proteinen nützlich, einschließlich von Immunglobulinmolekülen oder Teilen davon, und von chimären Antikörpern (d.h. ein Antikörper, der eine menschliche konstante Region, gekoppelt an eine murine antigenbindende Region besitzt) oder Fragmenten davon. Es sind zahlreiche Verfahren bekannt, durch welche DNA, die Immunglobulinmoleküle kodiert, manipuliert werden kann, um DNA zu ergeben, die in der Lage ist, rekombinante Proteine, wie zum Beispiel einkettige Antikörper, Antikörper mit verstärkter Affinität oder andere Polypeptide auf Antikörperbasis zu kodieren (siehe zum Beispiel Larrick et al., *Biotechnology* 7: 934–938, 1989; Reichmann et al., *Nature* 332: 323–327, 1988; Roberts et al., *Nature* 328: 731–734, 1987; Verhoeyen et al., *Science* 239: 1534–1536, 1988; Chaudhary et al., *Nature* 339: 394–397, 1989).

[0054] Verschiedene Fusionsproteine können auch mittels der Vektoren und Verfahren der Erfindung exprimiert werden. Beispiele für solche Fusionsproteine schließen Proteine, die als Fusion mit einem Teil eines Immunglobulinmoleküls exprimiert werden, Proteine, die als Fusionsproteine mit einem Zipper-Rest exprimiert werden, und neuartige polyfunktionale Proteine, wie zum Beispiel Fusionsproteine eines Zytokins und eines Wachstumsfaktors (d.h. GM-CSF und IL-3, MGF und IL-3) ein. WO 93/08207 und WO 96/40918 beschreiben die Herstellung verschiedener löslicher oligomerer Formen eines Moleküls, das als CD40L bezeichnet wird, welches ein Immunglobulininfusionsprotein bzw. ein Zipper-Fusionsprotein enthält; die hierin diskutierten Verfahren sind ohne weiteres auf andere Proteine anwendbar.

[0055] Weitere Beispiele sind: DNAs, die auf einem oder mehreren exprimierten Sequenztags (EST) aus einer Bibliothek von ESTs beruhen, können hergestellt, in die Vektoren der Erfindung eingefügt und exprimiert werden, um ein rekombinantes Polypeptid zu erhalten. Weiters können DNAs, die durch die Verwendung von ESTs (d.h. durch PCR oder die Anwendung anderer Klonierungsverfahren) isoliert wurden, ebenfalls durch Anwendung der vorliegenden Erfindung exprimiert werden. Informationen über die oben genannten Polypeptide sowie viele andere können aus einer Vielzahl von öffentlichen Quellen gewonnen werden, einschließlich elektronischen Datenbanken, wie zum Beispiel GenBank. Eine besonders nützliche Stelle ist die Web-Seite des National Center for Biotechnology Information/National Library of Medicine/National Institutes of Health (www.ncbi.nlm.nih.gov). Die Fachleute auf diesem Gebiet können Informationen erhalten, die zum Exprimieren eines gewünschten Polypeptids benötigt werden, und die hierin beschriebenen Verfahren durch routinemäßige Experimentierarbeit anwenden.

[0056] Für Zwecke dieser Anwendung schließt die Definition eines interessierenden Proteins jedoch Gene aus, die Proteine codieren, welche normalerweise als selektierbare Marker in Zellkulturen verwendet werden, wie zum Beispiel als auxotrophe, Antimetabolit- und/oder antibiotische Marker. Trotzdem schließt die Erfindung die Verwendung eines selektierbaren Markers als Hilfe bei der Selektion von Zellen und/oder bei der Amplifikation von Klonen ein, die genetisch werden, um ein interessierendes Gen zu exprimieren. Das selektierbare Markergen ist vorzugsweise angrenzend an das interessierende Gen angeordnet, sodass die Selektion und/oder Amplifikation des Markergens das benachbarte Gen selektiert oder amplifiziert.

[0057] Die folgenden Beispiele sollen zur Erläuterung spezieller Ausführungsformen dieser Erfindung dienen. Fachleute auf diesem Gebiet werden ohne weiteres erkennen, dass zusätzliche Ausführungsformen von der Erfindung eingeschlossen werden.

BEISPIELE

Beispiel 1

[0058] Dieses Beispiel beschreibt die Herstellung mehrerer Expressionsvektoren zur Expression einer löslichen Form eines Rezeptors für humanes Interleukin-4, der als sIL-4R bezeichnet wird. Humane IL-4R-cDNA und Protein werden in den US-Patenten 5,840,869, ausgestellt am 24. November 1998; 5,599,905, ausgestellt am 4. Februar 1997, und 5,856,296, ausgestellt am 5. Januar 1999, offenbart. SEQ ID Nrn: 5 und 6 stellen die Nucleotid- bzw. Aminosäuresequenz von humanem IL-4R dar. Die Aminosäuren –25 bis –1 weisen ein mutmaßliches Leit-Peptid auf; es wurde festgestellt, dass die Spaltung zwischen den Aminosäuren –1 und 1 und

zwischen den Aminosäuren –3 und –2 stattfindet. Die Aminosäuren 208 bis 231 bilden eine Transmembranregion. DNA, die sIL-4R von der Aminosäure –25 bis Aminosäure 207 kodiert, wurde in den Expressionsvektoren verwendet.

[0059] Der ursprüngliche Expressionsvektor, pCAVDHFR, ist ein Derivat von pCD302 (Mosley et al., Cell 89: 335–348, 1989), das die kodierende Sequenz für Mäuse-DHFR enthält (Subramani et al., Mol. Cell. Biol. 1: 854, 1981). Der pCDE-Vektor ist ein Derivat von pCAVDHFR, der die murine-Encephalomyocarditis-Virus-IRES (Nucleotide 260 bis 824; Jang und Wimmer, Genes and Dev. 4: 1560, 1990) enthält, die zwischen die dreiteilige Adenovirus-Vorsequenz und die DHFR-cDNA-Kodierendessequenz kloniert ist. Ein expressionsverstärkendes Sequenzelement (EASE) wurde stromaufwärts von der CMV-Vorsequenz eingeschlossen. Das EASE wird in US-Patent 6,027,915, ausgestellt am 22. Februar 2000, und in USSN 09/435,377, eingereicht am 5. November 1999, das nun gewährt ist, beschrieben.

[0060] Um die Polyadenylierung der dicistronischen Mitteilung zu ermöglichen, wurde das Polyadenylierungssignal des Rinderwachstumshormons am 3' des DHFR-Gens platziert. Das Plasmid pBGH ist ein dicistronischer Standardvektor und dient als Kontrolle. Die alternativen Polyadenylierungsvektoren der vorliegenden Erfindung wurden durch Einfügen verschiedener Polyadenylierungsstellen zwischen dem IL-4R und der IRES konstruiert. Die Plasmide pSPA4, pSPA6 und pMLPA wurden durch Einfügen der späten SV40-polyA-Stelle, der frühen SV40-polyA-Stelle bzw. einer Deletionsmutante der späten SV40-polyA-Stelle konstruiert. Die Deletionsmutante der späten SV40-polyA-Stelle wurde unter Verwendung von PCR zum Isolieren eines Fragmentes des späten SV40-polyA, und zwar der Nucleotide 80 bis 222 der SEQ ID Nr. 1, konstruiert. Ein Diagramm der verschiedenen Konstrukte wird in [Fig. 1](#) gezeigt; die Nucleotidsequenzen der verschiedenen polyA-Stellen werden in der Sequenzauflistung gezeigt (SV40 spät: SEQ ID Nr. 1; BGH: SEQ ID Nr. 2; SV40 früh: SEQ ID Nr. 3).

[0061] Die Plasmide wurden in Standardtransfektionen verwendet, um transfizierte Zellen, die IL-4R exprimieren, herzustellen. Ovarialzellen des Chinesischen Hamster (CHO-Zellen), denen die Dihydrofolatreduktase (DHFR) fehlt, DXB11-Zellen (Chasin und Urlaub, oben) wurden auf ein serumfreies Medium auf der Basis von DMEM: F12, das mit 2 mM L-Glutamin, 90 mM Thymidin, 90 mM Hypoxanthin, 120 mM Glycin, 5% Hy-Soja-Pepton und 100 mg/l insulinähnlichem Wachstumsfaktor 1 (Rasmussen et al., Cytotechnology 28: 31–42, 1998) ergänzt wurde, umgestellt. Für die DHFR-Selektion und Methotrexat-Amplifikationen wurden die Zellen im selben Medium kultiviert, dem Thymidin, Hypoxanthin und Glycin fehlte. Für die Methotrexat-Selektion wird Methotrexat (MTX; Lederle Laboratories, Pearl River, NY) dem Selektionsmedium in geeigneten Konzentrationen hinzugefügt. Wenn die Neomycin-Selektion eingesetzt wird, werden 400 µg/ml G418 (Gibco, Grand Island, NY) dem Medium hinzugefügt. Die Zellen werden unter Verwendung der Kalziumphosphattransfektion (Wigler et al., oben) oder der Lipofectamine™-Transfektion, wie vom Lieferant (Gibco BRL, Gaithersburg, MD) empfohlen wird, transfiziert. Das Lipofectamine™-Reagenz ist ein kommerziell erhältliches Reagenz, das zur Bildung von Lipid-Nucleinsäure-Komplexen (oder Liposomen) verwendet wird, das bei Anwendung auf kultivierte Zellen die Aufnahme der Nucleinsäure in die Zellen erleichtert.

Beispiel 2

[0062] Dieses Beispiel beschreibt ein semiquantitatives Polymerase-Kettenreaktions-(PCR)-Verfahren, das zur Bestätigung verwendet wurde, dass die IL-4R- und DHFR-Mitteilungen, die von den oben beschriebenen Plasmiden kodiert wurden, auf den relativen Niveaus der verschiedenen mRNAs erzeugt wurden und Informationen bereitstellen. Zellen wurden, wie beschrieben transfiziert und kultiviert, und mRNA wurde mittels eines RNeasy Gesamt-RNA-Isolierungssatzes (Quiagen, Chatsworth, CA) gewonnen und mit RNase-freier DNase behandelt, um die DNA-Kontamination zu verringern. Oligo-dT-Primer wurden verwendet, um den ersten Strang cDNA herzustellen; ein Kontroll-Primer für Aktin wurde aufgenommen, um die Quantifizierung zu erleichtern.

[0063] Der erste Strang wurde amplifiziert und die Menge an eingegebener RNA mittels einer GeneAmp 5700 von PE Biosystems (Forster City, CA) bestimmt. Dreißig PCR-Zyklen wurden ausgeführt, und eine Quantifizierung der PCR-Produkte wurde in Echtzeit unter Verwendung des doppelsträngigen DNA-Bindestoffes SYBR Green I (PE Biosystems, Forster City, CA) erreicht. Es wurde eine Standardkurve erstellt, unter Verwendung von bekannten Mengen an Aktin-cDNA, IL-4R-cDNA und DHFR-cDNA. Die Menge an cDNA in jeder Probe wurde mit der Menge an Aktin-DNA normalisiert. Die relativen Mengen an IL-4R und DHFR in jeder Probe sind in Tabelle 2 gezeigt.

Tabelle 2

Konstrukt	IL-4R/Aktin	DHFR/Aktin
pBGH	4,2	7,4
pMPLA	32,5	2,0

[0064] Diese Daten demonstrieren, dass die Zellen, die mit dem alternativen Polyadenylierungsvektor transfiziert wurden, etwa 8 mal so viel IL-4R-spezifische Mitteilungen wie die Kontrolle enthalten, und die Menge an DHFR ist im Verhältnis zur Kontrolle um das 3,5fache reduziert. Dieses Verfahren kann dazu verwendet werden, um zusätzliche Polyadenylierungssignale zum Gebrauch in den Expressionsvektoren der Erfindung zu evaluieren.

Beispiel 3

[0065] Dieses Beispiel beschreibt eine Enzymgebundene-Immunosorbent-Untersuchung (ELISA), der zur Überwachung der Erzeugung von rekombinanten Proteinen verwendet werden kann. Der ELISA ist im Fachgebiet gut bekannt; Anpassungen der Techniken, die in Engval et al., Immunochem. 8: 871, 1971, und im US-Patent 4,703,004 offenbart wurden, sind zur Überwachung der Herstellung von verschiedenen rekombinanten Proteinen eingesetzt worden. In dieser Untersuchung wird ein erster Antikörper, der spezifisch für ein interessierendes Protein ist (normalerweise ein monoklonaler Antikörper) auf einem Substrat (am häufigsten eine 96-Kammer-Mikrotiterplatte) immobilisiert, dann wird eine Probe, die das Protein enthält, hinzugefügt und inkubiert. Eine Reihe von Verdünnungen mit einer bekannten Konzentration des Proteins wird ebenfalls hinzugefügt und inkubiert, um eine Standardkurve zu erhalten. Nach einem Waschschrift zum Entfernen von ungebundenen Proteinen und anderen Materialien wird ein zweiter Antikörper zum Protein hinzugefügt. Der zweite Antikörper ist gegen ein anderes Epitop des Proteins gerichtet und kann entweder ein monoklonaler Antikörper oder ein polyklonaler Antikörper sein.

[0066] Ein konjugiertes Reagenz, das einen an ein Enzym, wie z.B. Meerrettich-Peroxidase (HRP) konjugierten Antikörper aufweist, der an den zweiten Antikörper bindet, wird hinzugefügt, entweder nach einem zweiten Waschschrift zum Entfernen von ungebundenem Protein oder zur gleichen Zeit wie der zweite Antikörper. Nach einer geeigneten Inkubationszeit wird ungebundenes konjugiertes Reagenz durch Waschen entfernt, und eine Entwicklungslösung, die das Substrat für das Enzymkonjugat enthält, wird der Platte hinzugefügt, wodurch sich Farbe entwickelt. Die Messwerte für die optische Dichte bei der richtigen Wellenlänge ergeben Zahlenwerte für jede Kammer. Die Werte für die Probe werden mit den Werten der Standardkurve verglichen, was die Quantifizierung der Werte für das gewünschte Protein ermöglicht.

[0067] Zur Quantifizierung von sIL-4R wurde ein ELISA mit zwei monoklonalen Antikörpern (MAb) entwickelt, die gegen unterschiedliche Epitope von IL-4R gerichtet sind. Der erste MAb (der als M10 bezeichnet wird) wurde über Nacht auf den Platten adsorbiert, und der peroxidase (HRP)-konjugierte zweite Antikörper (der als HRP-M8 bezeichnet wird) wurde nach dem Waschschrift hinzugefügt.

Beispiel 4

[0068] Dieses Beispiel beschreibt die Transfektion von CHO-Zellen mit verschiedenen Konstrukten und vergleicht die Herstellung von sIL-4R durch Pools von transfizierten Zellen. Die verschiedenen sIL-4R-Expressionsplasmide wurden in CHO-Zellen mittels LipofectamineTM transfiziert. Die Zellen wurden zuerst auf den DHFR+-Phänotyp selektiert, dann in einen Pool gebracht und bei verschiedenen MTX-Konzentrationen selektiert. Zellpools wurden zwei bis drei Tage lang gezüchtet, dann wurde Überstandsflüssigkeit abgeschöpft und mittels ELISA wie in Beispiel 3 beschrieben ist, analysiert, und die spezifische Produktivität (definiert als µg erzeugtes Protein pro Tag für 10⁶ Zellen) wurde bestimmt. Die Ergebnisse eines repräsentativen Experimentes werden in Tabelle 3 unten gezeigt.

Tabelle 3

Konstrukt	Zellen/ml x 10 ⁶	% lebensfähig	ELISA (µg Protein)	Spez. Produktivität µg/10 ⁶ Zellen/Tag
BGH, Kontrolle	1,98	94	0,3	0,12
SPA4, SV40 spät	1,57	89	2,3	1,11
SPA6, SV40 früh	2,42	91	3,2	1,10
MLPA	1,81	92	2,5	1,08
PY, Polyomavirus	2,4	93	0,4	0,14

[0069] Diese Ergebnisse zeigten, dass die Einfügung von internen polyA-Stellen zwischen eine DNA, die ein gewünschtes rekombinantes Protein kodiert, und eine DNA, die einen selektierbaren Marker kodiert, die Expression des gewünschten rekombinanten Proteins aus Pools von transfizierten Zellen verstärken kann.

Beispiel 5

[0070] Dieses Beispiel illustriert die Herstellung von sIL-4R durch Pools von transfizierten CHO-Zellen mit der Zeit. Ein hohes Expressionsniveau war über viele Passagen stabil. Vier unabhängige Transfektionen mit dem MLPA-Plasmid wurden im wesentlichen wie vorher beschrieben durchgeführt und durchliefen mehr als 20 Generationen. Die Expression von jeder Kultur wurde einzeln überwacht, und von den spezifischen Produktivitätsergebnissen wurde der Durchschnitt gebildet; die Durchschnitte werden in Tabelle 4 gezeigt.

Tabelle 4

Durchgangsnummer	Spezifische Produktivität µg/10 ⁶ Zellen/Tag	Standardabweichung
5	1,22	0,41
10	1,26	0,39
15	1,18	0,55
18	1,09	0,49
20	1,02	0,62

[0071] Wie an den Daten in Tabelle 4 zu erkennen ist, blieb die Expression über 20 Durchgänge stabil. Zellen von Durchlauf 20 von zwei der Pools wurden dann in 5 nM Methotrexat amplifiziert und auf IL-4R-Expression überwacht; die Ergebnisse werden in Tabelle 5 gezeigt. Amplifizierte Pools zeigten eine erhöhte Expression im Vergleich zu den nicht amplifizierten Pools.

Tabelle 5

Durchgangsnummer	Spezifische Produktivität Pool Nr. 1	Spezifische Produktivität Pool Nr. 2
27	1,95	1,37
29	2,10	1,58
33	2,08	1,42

Beispiel 6

[0072] Dieses Beispiel erläutert die Wirkung von internen polyA-Stellen auf Klone von Zellen, die aus transfizierten Pools abgeleitet wurden. BGH-, SPA4- und SPA6-Zellen wurden durch Beschränken der Verdünnung in Gegenwart von MTX kloniert. Mehrere Kolonien wurden ausgewählt und auf spezifische Produktivität von sIL-4R durchsucht, wie für die Pools beschrieben ist. Die Ergebnisse werden in Tabelle 6 gezeigt.

Tabelle 6

Konstrukt	Klon-Nr.	MTX-Konzentration	Zellen/ml	% lebensfähig	ELISA (µg Protein)	Spez. Produktiv. (µg/10 ⁶ Zellen/Tag)
BGH	3	200	1,84	69	14,6	4,25
SPA4	2	50	2,26	89	16,2	3,99
SPA6	10	100	1,82	94	1,6	0,47
SPA6	11	100	1,46	83	7	2,44
SPA6	13	100	0,96	67	9,3	4,40
SPA6	16	100	1,02	73	0	0

[0073] Diese Ergebnisse zeigen, dass die Klone, die aus den Pools ausgesucht wurden, welche mit Expressionsvektoren, die eine interne polyA-Stelle aufweisen, transfiziert waren, hohe Konzentrationen des gewünschten rekombinanten Proteins exprimieren können. Für die Zwecke der Erzeugung großer Mengen an rekombinantem Protein zur Verwendung als Pharmazeutikum werden Klone oft in Methotrexat erneut amplifiziert. Um die Wirkung einer internen polyA-Stelle auf den erneuten Amplifizierungsprozess zu bewerten, wurde Klon 2 aus dem SPA4-Pool erneut durch Kultivieren der Zellen über mehrere Durchgänge in steigenden Konzentrationen von Methotrexat amplifiziert. Sobald die Zellen aus der Methotrexat-Amplifizierung mit Viabilitäten von etwa 90% gewonnen wurden, wurde die spezifische Produktivität durch Kultivieren der Zellen über zwei bis drei Tage, Abschöpfen der Überstandsflüssigkeit und Testen der Überstandsflüssigkeit auf IL-4R durch ELISA bestimmt; die Ergebnisse werden in Tabelle 7 gezeigt.

Tabelle 7

Konstrukt	Methotrexat-konzentration	Zellen/ml	% lebensfähig	ELISA ($\mu\text{g Protein}$)	Spez. Produktivität $\mu\text{g}/10^6$ Zellen/Tag
SPA4-2	50 nM	2,32	94	18,7	4,58
SPA4-2	100 nM	1,66	91	21,1	6,83
SPA4-2	150 nM	1,95	90	27,7	7,86
SPA4-2	200 nM	1,93	91	28,8	8,24

[0074] Diese Ergebnisse zeigen, dass Klone von Zellen, die mit Expressionsvektoren transfiziert wurden, welche eine interne polyA-Stelle aufweisen, erneut amplifiziert werden können; es kann erwartet werden, dass sie eine höhere spezifische Produktivität zeigen.

Beispiel 7

[0075] Dieses Beispiel beschreibt die Herstellung von mehreren Expressionsvektoren für die Expression von rekombinanten Proteinen. Ein Expressionsvektor, der ein Markerprotein kodiert (sekretierte alkalische Phosphatase oder SEAP; Bergen et al., Gene 66: 1, 1988), wird im Wesentlichen so wie beschrieben hergestellt, wobei die MLPA-polyA-Stelle intern verwendet wird; eine andere polyA-Stelle als BGH kann als terminale polyA-Stelle verwendet werden. Mehrere Änderungen werden an der IRES-Sequenz innerhalb der Expressionsvektoren vorgenommen. Wie in Davies und Kaufman (J. Virology 66: 1924, 1992) diskutiert ist, kann die Effizienz der Translation eines zweiten Gens durch Ändern der IRES-Sequenz bei der oder in der Nähe der Verbindung der IRES mit dem zweiten Gen, in diesem Fall DHFR, manipuliert werden. Tabelle 8 stellt die Nucleotidsequenz dar, die der IRES hinzugefügt wird; wobei die erste Base, die in der Tabelle angegeben ist, direkt nach dem Nucleotid 566 der EMCV-IRES (SEQ ID NO: 7) liegt. Die translatorischen Startstellen (ATG) sind unterstrichen; das 3'-ATG ist das erste ATG von muDHFR.

Tabelle 8

Konstrukt	DNA-Sequenz an der IRES-DHFR-Verbindungsstelle
IX-312	ATTGCTCGAGATCCGTGCCATCATG (SEQ ID Nr: 8)
IXED-1	ATGATAATATG (SEQ ID Nr: 9)
IXED-3	ATGATAATATGGCCACAACCATG (SEQ ID Nr: 10)

[0076] Das Anhängen der Nucleotidsequenzen an die IRES moduliert die Expression von DHFR ausreichend, um den Prozentsatz von Zellen zu erhöhen, die ohne wesentliche Verringerung der Konzentrationen des gewünschten rekombinanten Proteins transfiziert werden. Die Vektoren (einschließlich der Kontrollvektoren) werden bei Standardtransfektionen verwendet, um transfizierte Zellen herzustellen, die SEAP, wie im Wesentlichen hierin beschrieben ist, exprimieren. Die Expressionsniveaus des Markerproteins, SEAP, werden mit Hilfe einer quantitativen Untersuchung, wie zum Beispiel dem von CLONTECH Laboratories (Palo Alto, CA, USA; Yang et al., Biotechniques 2: 1110, 1997) erhältlichen, bestimmt.

Sequenzliste

<110> Immunex Corporation

<120> Vektor und Verfahren zur Expression rekombinanter Proteine

<130> 2902-WO

<140> --ist zu übertragen--

<141> 2000-10-12

<150> 60/159,177

<151> 1999-10-13

<160> 10

<170> PatentIn Version 3.0

<210> 1

<211>222

<212> DNA

<213> SV40

<400> 1

```

atccagacat gataagatac attgatgagt ttggacaaac cacaactaga atgcagtgaa      60
aaaaatgctt tatttctgaa atttctgatg ctattgcttt atttctaacc attataagct      120
gcaataaaca agttcaacaa caattgcatt cattttatgt ttcagggttca gggggagggtg      180
tgggagggttc tttaaagcaa gtaaaacctc tacaatgtg gt                          222

```

<210> 2

<211> 285

<212> DNA

<213> Rind

<400> 2

```

aattgtctag agctcgctga tcagcctcga ctgtgccttc tagttgccag ccattctgtg      60
tttgcccttc ccccgctgct tccttgacct tggaagggtgc cactccact gtcctttcct      120
aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg      180
gggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctgggggatg      240
cggtagggctc tatggcttct gaggcggaaa gaaccagctg gggca                      285

```

<210> 3

<211> 222

<212> DNA

<213> SV40

<400> 3

accacatttg tagagggtttt acttgcttta aaaaacctcc cacacctccc cctgaacctg 60

aaacataaaa tgaatgcaat tgttggtgaa cttggttatt gcagcttata atgggtacaa 120

ataaagcaat agcatcacia atttcacaaa taaagcattt ttttcaactgc attctagttag 180

tggtttgtcc aaactcatca atgtatctta tcatgtctgg at 222

<210> 4

<211> 19

<212> RNA

<213> RNA Erkennungssequenz

<400> 4

acaugaggau uacccaugu 19

<210> 5

<211> 2478

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(2478)

<220>

<221> mat.-Peptid

<222> (76)..()

<220>

<221> sig.-Peptid

<222> (1)..(75)

<400> 5

atg ggg tgg ctt tgc tct ggg ctc ctg ttc cct gtg agc tgc ctg gtc	48
Met Gly Trp Leu Cys Ser Gly Leu Leu Phe Pro Val Ser Cys Leu Val	
-25 -20 -15 -10	
ctg ctg cag gtg gca agc tct ggg aac atg aag gtc ttg cag gag ccc	96
Leu Leu Gln Val Ala Ser Ser Gly Asn Met Lys Val Leu Gln Glu Pro	
-5 -1 1 5	
acc tgc gtc tcc gac tac atg agc atc tct act tgc gag tgg aag atg	144
Thr Cys Val Ser Asp Tyr Met Ser Ile Ser Thr Cys Glu Trp Lys Met	
10 15 20	
aat ggt ccc acc aat tgc agc acc gag ctc cgc ctg ttg tac cag ctg	192
Asn Gly Pro Thr Asn Cys Ser Thr Glu Leu Arg Leu Leu Tyr Gln Leu	
25 30 35	
gtt ttt ctg ctc tcc gaa gcc cac acg tgt atc cct gag aac aac gga	240
Val Phe Leu Leu Ser Glu Ala His Thr Cys Ile Pro Glu Asn Asn Gly	
40 45 50 55	
ggc gcg ggg tgc gtg tgc cac ctg ctc atg gat gac gtg gtc agt gcg	288
Gly Ala Gly Cys Val Cys His Leu Leu Met Asp Asp Val Val Ser Ala	
60 65 70	
gat aac tat aca ctg gac ctg tgg gct ggg cag cag ctg ctg tgg aag	336
Asp Asn Tyr Thr Leu Asp Leu Trp Ala Gly Gln Gln Leu Leu Trp Lys	
75 80 85	

ggc tcc ttc aag ccc agc gag cat gtg aaa ccc agg gcc cca gga aac Gly Ser Phe Lys Pro Ser Glu His Val Lys Pro Arg Ala Pro Gly Asn 90 95 100	384
ctg aca gtt cac acc aat gtc tcc gac act ctg ctg ctg acc tgg agc Leu Thr Val His Thr Asn Val Ser Asp Thr Leu Leu Leu Thr Trp Ser 105 110 115	432
aac ccg tat ccc cct gac aat tac ctg tat aat cat ctc acc tat gca Asn Pro Tyr Pro Pro Asp Asn Tyr Leu Tyr Asn His Leu Thr Tyr Ala 120 125 130 135	480
gtc aac att tgg agt gaa aac gac ccg gca gat ttc aga atc tat aac Val Asn Ile Trp Ser Glu Asn Asp Pro Ala Asp Phe Arg Ile Tyr Asn 140 145 150	528
gtg acc tac cta gaa ccc tcc ctc cgc atc gca gcc agc acc ctg aag Val Thr Tyr Leu Glu Pro Ser Leu Arg Ile Ala Ala Ser Thr Leu Lys 155 160 165	576
tct ggg att tcc tac agg gca cgg gtg agg gcc tgg gct cag tgc tat Ser Gly Ile Ser Tyr Arg Ala Arg Val Arg Ala Trp Ala Gln Cys Tyr 170 175 180	624
aac acc acc tgg agt gag tgg agc ccc agc acc aag tgg cac aac tcc Asn Thr Thr Trp Ser Glu Trp Ser Pro Ser Thr Lys Trp His Asn Ser 185 190 195	672
tac agg gag ccc ttc gag cag cac ctc ctg ctg gcc gtc agc gtt tcc Tyr Arg Glu Pro Phe Glu Gln His Leu Leu Leu Gly Val Ser Val Ser 200 205 210 215	720
tgc att gtc atc ctg gcc gtc tgc ctg ttg tgc tat gtc agc atc acc Cys Ile Val Ile Leu Ala Val Cys Leu Leu Cys Tyr Val Ser Ile Thr 220 225 230	768
aag att aag aaa gaa tgg tgg gat cag att ccc aac cca gcc cgc agc Lys Ile Lys Lys Glu Trp Trp Asp Gln Ile Pro Asn Pro Ala Arg Ser 235 240 245	816
cgc ctc gtg gct ata ata atc cag gat gct cag ggg tca cag tgg gag Arg Leu Val Ala Ile Ile Ile Gln Asp Ala Gln Gly Ser Gln Trp Glu 250 255 260	864
aag cgg tcc cga gcc cag gaa cca gcc aag tgc cca cac tgg aag aat Lys Arg Ser Arg Gly Gln Glu Pro Ala Lys Cys Pro His Trp Lys Asn 265 270 275	912
tgt ctt acc aag ctc ttg ccc tgt ttt ctg gag cac aac atg aaa agg Cys Leu Thr Lys Leu Leu Pro Cys Phe Leu Glu His Asn Met Lys Arg 280 285 290 295	960
gat gaa gat cct cac aag gct gcc aaa gag atg cct ttc cag gcc tct Asp Glu Asp Pro His Lys Ala Ala Lys Glu Met Pro Phe Gln Gly Ser 300 305 310	1008
gga aaa tca gca tgg tgc cca gtg gag atc agc aag aca gtc ctc tgg Gly Lys Ser Ala Trp Cys Pro Val Glu Ile Ser Lys Thr Val Leu Trp 315 320 325	1056

cca gag agc atc agc gtg gtg cga tgt gtg gag ttg ttt gag gcc ccg	1104
Pro Glu Ser Ile Ser Val Val Arg Cys Val Glu Leu Phe Glu Ala Pro	
330 335 340	
gtg gag tgt gag gag gag gag gag gta gag gaa gaa aaa ggg agc ttc	1152
Val Glu Cys Glu Glu Glu Glu Glu Val Glu Glu Glu Lys Gly Ser Phe	
345 350 355	
tgt gca tgc cct gag agc agc agg gat gac ttc cag gag gga agg gag	1200
Cys Ala Ser Pro Glu Ser Ser Arg Asp Asp Phe Gln Glu Gly Arg Glu	
360 365 370 375	
ggc att gtg gcc cgg cta aca gag agc ctg ttc ctg gac ctg ctc gga	1248
Gly Ile Val Ala Arg Leu Thr Glu Ser Leu Phe Leu Asp Leu Leu Gly	
380 385 390	
gag gag aat ggg ggc ttt tgc cag cag gac atg ggg gag tca tgc ctt	1296
Glu Glu Asn Gly Gly Phe Cys Gln Gln Asp Met Gly Glu Ser Cys Leu	
395 400 405	
ctt cca cct tgc gga agt acg agt gct cac atg ccc tgg gat gag ttc	1344
Leu Pro Pro Ser Gly Ser Thr Ser Ala His Met Pro Trp Asp Glu Phe	
410 415 420	
cca agt gca ggg ccc aag gag gca cct ccc tgg ggc aag gag cag cct	1392
Pro Ser Ala Gly Pro Lys Glu Ala Pro Pro Trp Gly Lys Glu Gln Pro	
425 430 435	
ctc cac ctg gag cca agt cct cct gcc agc ccg acc cag agt cca gac	1440
Leu His Leu Glu Pro Ser Pro Pro Ala Ser Pro Thr Gln Ser Pro Asp	
440 445 450 455	
aac ctg act tgc aca gag acg ccc ctc gtc atc gca ggc aac cct gct	1488
Asn Leu Thr Cys Thr Glu Thr Pro Leu Val Ile Ala Gly Asn Pro Ala	
460 465 470	
tac cgc agc ttc agc aac tcc ctg agc cag tca ccg tgt ccc aga gag	1536
Tyr Arg Ser Phe Ser Asn Ser Leu Ser Gln Ser Pro Cys Pro Arg Glu	
475 480 485	
ctg ggt cca gac cca ctg ctg gcc aga cac ctg gag gaa gta gaa ccc	1584
Leu Gly Pro Asp Pro Leu Leu Ala Arg His Leu Glu Val Glu Pro	
490 495 500	
gag atg ccc tgt gtc ccc cag ctc tct gag cca acc act gtg ccc caa	1632
Glu Met Pro Cys Val Pro Gln Leu Ser Glu Pro Thr Thr Val Pro Gln	
505 510 515	
cct gag cca gaa acc tgg gag cag atc ctc cgc cga aat gtc ctc cag	1680
Pro Glu Pro Glu Thr Trp Glu Gln Ile Leu Arg Arg Asn Val Leu Gln	
520 525 530 535	
cat ggg gca gct gca gcc ccc gtc tgc gcc ccc acc agt ggc tat cag	1728
His Gly Ala Ala Ala Ala Pro Val Ser Ala Pro Thr Ser Gly Tyr Gln	
540 545 550	
gag ttt gta cat gcg gtg gag cag ggt ggc acc cag gcc agt gcg gtg	1776
Glu Phe Val His Ala Val Glu Gln Gly Gly Thr Gln Ala Ser Ala Val	
555 560 565	

gtg ggc ttg ggt ccc cca gga gag gct ggt tac aag gcc ttc tca agc Val Gly Leu Gly Pro Pro Gly Glu Ala Gly Tyr Lys Ala Phe Ser Ser 570 575 580	1824
ctg ctt gcc agc agt gct gtg tcc cca gag aaa tgt ggg ttt ggg gct Leu Leu Ala Ser Ser Ala Val Ser Pro Glu Lys Cys Gly Phe Gly Ala 585 590 595	1872
agc agt ggg gaa gag ggg tat aag cct ttc caa gac ctc att cct ggc Ser Ser Gly Glu Glu Gly Tyr Lys Pro Phe Gln Asp Leu Ile Pro Gly 600 605 610 615	1920
tgc cct ggg gac cct gcc cca gtc cct gtc ccc ttg ttc acc ttt gga Cys Pro Gly Asp Pro Ala Pro Val Pro Val Pro Leu Phe Thr Phe Gly 620 625 630	1968
ctg gac agg gag cca cct cgc agt ccg cag agc tca cat ctc cca agc Leu Asp Arg Glu Pro Pro Arg Ser Pro Gln Ser Ser His Leu Pro Ser 635 640 645	2016
agc tcc cca gag cac ctg ggt ctg gag ccg ggg gaa aag gta gag gac Ser Ser Pro Glu His Leu Gly Leu Glu Pro Gly Glu Lys Val Glu Asp 650 655 660	2064
atg cca aag ccc cca ctt ccc cag gag cag gcc aca gac ccc ctt gtg Met Pro Lys Pro Pro Leu Pro Gln Glu Gln Ala Thr Asp Pro Leu Val 665 670 675	2112
gac agc ctg ggc agt ggc att gtc tac tca gcc ctt acc tgc cac ctg Asp Ser Leu Gly Ser Gly Ile Val Tyr Ser Ala Leu Thr Cys His Leu 680 685 690 695	2160
tgc ggc cac ctg aaa cag tgt cat ggc cag gag gat ggt ggc cag acc Cys Gly His Leu Lys Gln Cys His Gly Gln Glu Asp Gly Gly Gln Thr 700 705 710	2208
cct gtc atg gcc agt cct tgc tgt ggc tgc tgc tgt gga gac agg tcc Pro Val Met Ala Ser Pro Cys Cys Gly Cys Cys Cys Gly Asp Arg Ser 715 720 725	2256
tcg ccc cct aca acc ccc ctg agg gcc cca gac ccc tct cca ggt ggg Ser Pro Pro Thr Thr Pro Leu Arg Ala Pro Asp Pro Ser Pro Gly Gly 730 735 740	2304
gtt cca ctg gag gcc agt ctg tgt ccg gcc tcc ctg gca ccc tcg ggc Val Pro Leu Glu Ala Ser Leu Cys Pro Ala Ser Leu Ala Pro Ser Gly 745 750 755	2352
atc tca gag aag agt aaa tcc tca tca tcc ttc cat cct gcc cct ggc Ile Ser Glu Lys Ser Lys Ser Ser Ser Ser Phe His Pro Ala Pro Gly 760 765 770 775	2400
aat gct cag agc tca agc cag acc ccc aaa atc gtg aac ttt gtc tcc Asn Ala Gln Ser Ser Ser Gln Thr Pro Lys Ile Val Asn Phe Val Ser 780 785 790	2448
gtg gga ccc aca tac atg agg gtc tct tat Val Gly Pro Thr Tyr Met Arg Val Ser Tyr 795 800	2478

<210> 6

<211> 826

<212> PRT

<213> Homo sapiens

<400> 6

Met Gly Trp Leu Cys Ser Gly Leu Leu Phe Pro Val Ser Cys Leu Val
 -25 -20 -15 -10

Leu Leu Gln Val Ala Ser Ser Gly Asn Met Lys Val Leu Gln Glu Pro
 -5 -1 1 5

Thr Cys Val Ser Asp Tyr Met Ser Ile Ser Thr Cys Glu Trp Lys Met
 10 15 20

Asn Gly Pro Thr Asn Cys Ser Thr Glu Leu Arg Leu Leu Tyr Gln Leu
 25 30 35

Val Phe Leu Leu Ser Glu Ala His Thr Cys Ile Pro Glu Asn Asn Gly
 40 45 50 55

Gly Ala Gly Cys Val Cys His Leu Leu Met Asp Asp Val Val Ser Ala
 60 65 70

Asp Asn Tyr Thr Leu Asp Leu Trp Ala Gly Gln Gln Leu Leu Trp Lys
 75 80 85

Gly Ser Phe Lys Pro Ser Glu His Val Lys Pro Arg Ala Pro Gly Asn
 90 95 100

Leu Thr Val His Thr Asn Val Ser Asp Thr Leu Leu Leu Thr Trp Ser
 105 110 115

Asn Pro Tyr Pro Pro Asp Asn Tyr Leu Tyr Asn His Leu Thr Tyr Ala
 120 125 130 135

Val Asn Ile Trp Ser Glu Asn Asp Pro Ala Asp Phe Arg Ile Tyr Asn
 140 145 150

Val Thr Tyr Leu Glu Pro Ser Leu Arg Ile Ala Ala Ser Thr Leu Lys
 155 160 165

Ser Gly Ile Ser Tyr Arg Ala Arg Val Arg Ala Trp Ala Gln Cys Tyr
 170 175 180

Asn Thr Thr Trp Ser Glu Trp Ser Pro Ser Thr Lys Trp His Asn Ser
 185 190 195

Tyr Arg Glu Pro Phe Glu Gln His Leu Leu Leu Gly Val Ser Val Ser
 200 205 210 215
 Cys Ile Val Ile Leu Ala Val Cys Leu Leu Cys Tyr Val Ser Ile Thr
 220 225 230
 Lys Ile Lys Lys Glu Trp Trp Asp Gln Ile Pro Asn Pro Ala Arg Ser
 235 240 245
 Arg Leu Val Ala Ile Ile Ile Gln Asp Ala Gln Gly Ser Gln Trp Glu
 250 255 260
 Lys Arg Ser Arg Gly Gln Glu Pro Ala Lys Cys Pro His Trp Lys Asn
 265 270 275
 Cys Leu Thr Lys Leu Leu Pro Cys Phe Leu Glu His Asn Met Lys Arg
 280 285 290 295
 Asp Glu Asp Pro His Lys Ala Ala Lys Glu Met Pro Phe Gln Gly Ser
 300 305 310
 Gly Lys Ser Ala Trp Cys Pro Val Glu Ile Ser Lys Thr Val Leu Trp
 315 320 325
 Pro Glu Ser Ile Ser Val Val Arg Cys Val Glu Leu Phe Glu Ala Pro
 330 335 340
 Val Glu Cys Glu Glu Glu Glu Glu Val Glu Glu Glu Lys Gly Ser Phe
 345 350 355
 Cys Ala Ser Pro Glu Ser Ser Arg Asp Asp Phe Gln Glu Gly Arg Glu
 360 365 370 375
 Gly Ile Val Ala Arg Leu Thr Glu Ser Leu Phe Leu Asp Leu Leu Gly
 380 385 390
 Glu Glu Asn Gly Gly Phe Cys Gln Gln Asp Met Gly Glu Ser Cys Leu
 395 400 405
 Leu Pro Pro Ser Gly Ser Thr Ser Ala His Met Pro Trp Asp Glu Phe
 410 415 420
 Pro Ser Ala Gly Pro Lys Glu Ala Pro Pro Trp Gly Lys Glu Gln Pro
 425 430 435

Leu His Leu Glu Pro Ser Pro Pro Ala Ser Pro Thr Gln Ser Pro Asp
 440 445 450 455

Asn Leu Thr Cys Thr Glu Thr Pro Leu Val Ile Ala Gly Asn Pro Ala
 460 465 470

Tyr Arg Ser Phe Ser Asn Ser Leu Ser Gln Ser Pro Cys Pro Arg Glu
 475 480 485

Leu Gly Pro Asp Pro Leu Leu Ala Arg His Leu Glu Glu Val Glu Pro
 490 495 500

Glu Met Pro Cys Val Pro Gln Leu Ser Glu Pro Thr Thr Val Pro Gln
 505 510 515

Pro Glu Pro Glu Thr Trp Glu Gln Ile Leu Arg Arg Asn Val Leu Gln
 520 525 530 535

His Gly Ala Ala Ala Ala Pro Val Ser Ala Pro Thr Ser Gly Tyr Gln
 540 545 550

Glu Phe Val His Ala Val Glu Gln Gly Gly Thr Gln Ala Ser Ala Val
 555 560 565

Val Gly Leu Gly Pro Pro Gly Glu Ala Gly Tyr Lys Ala Phe Ser Ser
 570 575 580

Leu Leu Ala Ser Ser Ala Val Ser Pro Glu Lys Cys Gly Phe Gly Ala
 585 590 595

Ser Ser Gly Glu Glu Gly Tyr Lys Pro Phe Gln Asp Leu Ile Pro Gly
 600 605 610 615

Cys Pro Gly Asp Pro Ala Pro Val Pro Val Pro Leu Phe Thr Phe Gly
 620 625 630

Leu Asp Arg Glu Pro Pro Arg Ser Pro Gln Ser Ser His Leu Pro Ser
 635 640 645

Ser Ser Pro Glu His Leu Gly Leu Glu Pro Gly Glu Lys Val Glu Asp
 650 655 660

Met Pro Lys Pro Pro Leu Pro Gln Glu Gln Ala Thr Asp Pro Leu Val
 665 670 675

Asp Ser Leu Gly Ser Gly Ile Val Tyr Ser Ala Leu Thr Cys His Leu
 680 685 690 695

Cys Gly His Leu Lys Gln Cys His Gly Gln Glu Asp Gly Gly Gln Thr
 700 705 710

Pro Val Met Ala Ser Pro Cys Cys Gly Cys Cys Cys Gly Asp Arg Ser
 715 720 725

Ser Pro Pro Thr Thr Pro Leu Arg Ala Pro Asp Pro Ser Pro Gly Gly
 730 735 740

Val Pro Leu Glu Ala Ser Leu Cys Pro Ala Ser Leu Ala Pro Ser Gly
 745 750 755

Ile Ser Glu Lys Ser Lys Ser Ser Ser Ser Phe His Pro Ala Pro Gly
 760 765 770 775

Asn Ala Gln Ser Ser Ser Gln Thr Pro Lys Ile Val Asn Phe Val Ser
 780 785 790

Val Gly Pro Thr Tyr Met Arg Val Ser Tyr
 795 800

<210> 7

<211> 566

<212> DNA

<213> EMCV

<400> 7

```

ccccctctccc tccccccccc ctaacgttac tggccgaagc cgcttggaat aaggccggtg      60
tgcgtttgtc tatatgttat tttccaccat attgccgtct tttggcaatg tgagggcccg      120
gaaacctggc cctgtcttct tgacgagcat tcctaggggt ctttcccctc tcgccaaagg      180
aatgcaaggt ctgttgaatg tcgtgaagga agcagttcct ctggaagctt cttgaagaca      240
aacaacgtct gtagcgaccc ttgacaggca gcggaacccc ccacctggcg acaggtgcct      300
ctgcggccaa aagccacgtg tataagatac acctgcaaag gcggcacaac ccagtgcca      360
cgttgtgagt tggatagttg tggaaagagt caaatggctc tcctcaagcg tattcaacaa      420
ggggctgaag gatgcccaga aggtacccca ttgtatggga tctgatctgg ggcctcgggtg      480
cacatgcttt acatgtgttt agtcgaggtt aaaaaacgtc taggcccccc gaaccacggg      540
gacgtggttt tcctttgaaa aacacg                                     566

```

<210> 8

<211> 25

<212> DNA

<213> EMCV

<400> 8

```
attgctcgag atccgtgcc tcatg                                     25
```

<210> 9

<211> 11

<212> DNA

<213> EMCV

<400> 9

```
atgataatat g                                     11
```

<210> 10

<211> 23

<212> DNA

<213> EMCV

<400> 10

```
atgataatat ggccacaacc atg                                     23
```

Patentansprüche

1. Expressionsvektor, der in folgender Reihenfolge aufweist: eine Promotersequenz, eine erste codierende Sequenz, eine Polyadenylierungsstelle und eine zweite codierende Sequenz, wobei keine Promotersequenz zwischen der Polyadenylierungsstelle und der codierenden Sequenz aufscheint.

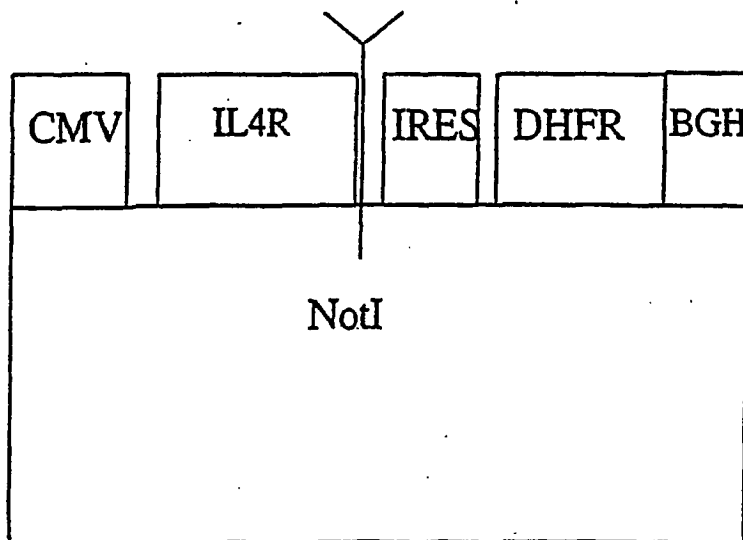
2. Expressionsvektor nach Anspruch 1, wobei die zweite codierende Sequenz einen wählbaren Marker codiert.
3. Expressionsvektor nach Anspruch 2, dadurch gekennzeichnet, dass der wählbare Marker Dihydrofolatreduktase ist.
4. Expressionsvektor nach einem der Ansprüche 1–3, wobei die Polyadenylierungsstelle eine Sequenz aufweist, die aus der Gruppe bestehend aus der SV40 Spät-Polyadenylierungsstelle (SEQ ID Nr: 1), SV40 Früh-Polyadenylierungsstelle (SEQ ID Nr: 3) und Resten 80 bis 222 von SEQ ID Nr: 1 gewählt wird.
5. Expressionsvektor nach einem der Ansprüche 1–4, wobei eine interne Ribosomen-Rezeptorstelle (IRES) zwischen der Polyadenylierungsstelle und der zweiten codierenden Sequenz derart eingefügt ist, dass der IRES wirkungsmäßig mit der zweiten codierenden Sequenz verbunden ist.
6. Expressionsvektor nach einem der Ansprüche 1–5, der weiters eine zweite Polyadenylierungsstelle nach der zweiten codierenden Sequenz aufweist und wirkungsmäßig mit derselben verbunden ist.
7. Wirtszelle, die den Expressionsvektor nach Anspruch 6 enthält.
8. Wirtszelle nach Anspruch 7, die eine Säugetier-Wirtszelle aufweist.
9. Säugetier-Wirtszelle nach Anspruch 8, wobei die erste codierende Sequenz ein rekombinantes Protein codiert.
10. Verfahren zur Gewinnung eines rekombinanten Proteins, das die Kultivierung der Säugetier-Wirtszelle gemäß Anspruch 9 unter Bedingungen, die die Expression des rekombinanten Proteins fördern, und die Gewinnung des rekombinanten Proteins aufweist.
11. Expressionsvektor, der in folgender Reihenfolge aufweist: eine Promotersequenz, eine erste codierende Sequenz, eine Polyadenylierungsstelle, die aus den Nucleotiden 80 bis 222 von SEQ ID Nr: 1 besteht, und eine zweite codierende Sequenz, wobei keine Promotersequenz zwischen der Polyadenylierungsstelle und der zweiten codierenden Sequenz auftritt.
12. Wirtszelle, die den Expressionsvektor nach Anspruch 11 enthält.
13. Wirtszelle nach Anspruch 12, die eine Säugetier-Wirtszelle aufweist.
14. Säugetier-Wirtszelle nach Anspruch 13, wobei die erste codierende Sequenz ein rekombinantes Protein codiert.
15. Verfahren zur Gewinnung eines rekombinanten Proteins, das die Kultivierung der Säugetier-Wirtszelle nach Anspruch 14 unter Bedingungen, die die Expression des rekombinanten Proteins fördern, und die Gewinnung des rekombinanten Proteins aufweist.
16. Expressionsvektor nach Anspruch 10, der weiters eine zweite Polyadenylierungsstelle nach der zweiten codierenden Sequenz aufweist und wirkungsmäßig mit derselben verbunden ist.
17. Säugetier-Wirtszelle, die einen Expressionsvektor nach Anspruch 16 enthält.

Es folgt ein Blatt Zeichnungen

Anhängende Zeichnungen

 SV 40-Früh-Polyadenylierungsstelle

 SV 40pA-Spät-Polyadenylierungsstelle



Figur 1