



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 600 33 583 T2** 2007.09.13

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 517 913 B1**

(21) Deutsches Aktenzeichen: **600 33 583.6**

(86) PCT-Aktenzeichen: **PCT/US00/27652**

(96) Europäisches Aktenzeichen: **00 970 619.3**

(87) PCT-Veröffentlichungs-Nr.: **WO 2001/025401**

(86) PCT-Anmeldetag: **06.10.2000**

(87) Veröffentlichungstag
der PCT-Anmeldung: **12.04.2001**

(97) Erstveröffentlichung durch das EPA: **30.03.2005**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **21.02.2007**

(47) Veröffentlichungstag im Patentblatt: **13.09.2007**

(51) Int Cl.⁸: **C12N 15/63** (2006.01)
C07K 19/00 (2006.01)

(30) Unionspriorität:
158585 P 07.10.1999 US

(73) Patentinhaber:
Corixa Corp. CSC, Wilmington, Del., US

(74) Vertreter:
HOFFMANN & EITLE, 81925 München

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

(72) Erfinder:
**SKEIKY, Yasir, Bellevue, WA 98006, US;
GUDERIAN, Jeffrey, Lynnwood, WA 98037, US**

(54) Bezeichnung: **EINE MYCOBACTERIUM TUBERCULOSIS KODIERENDE SEQUENZ ZUR EXPRESSION VON HE-TEROLOGEN PROTEINEN**

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

[0001] Die vorliegende Erfindung betrifft allgemein Nukleinsäure- und Aminosäuresequenzen eines Fusionspolypeptids, umfassend ein Mycobacterium tuberculosis-Polypeptid und ein heterologes Polypeptid von Interesse, Expressionsvektoren und Wirtszellen, umfassend solche Nukleinsäuren, und Verfahren zur Herstellung von Fusionspolypeptiden. Insbesondere betrifft die Erfindung Materialien und Verfahren unter Verwendung einer solchen M. tuberculosis-Sequenz als Fusionspartner, um die stabile Expression mit hoher Ausbeute von rekombinanten heterologen Polypeptiden sowohl eukaryotischen als auch prokaryotischen Ursprungs zu erleichtern.

Hintergrund der Erfindung

[0002] Der Beginn rekombinanter DNA-Technologie führte zur molekularen Klonierung einer großen Zahl von codierenden Sequenzen oder Genen diverser Zelltypen. Um die Funktion dieser Gene zu untersuchen oder die Produkte herzustellen, die durch solche Sequenzen codiert sind, werden diese Gene in Expressionsvektoren eingeführt und unter die Kontrolle geeigneter Regulationssequenzen gebracht. Die Übertragung dieses Expressionsvektors in eine eukaryotische oder prokaryotische Wirtszelle führt im allgemeinen zur Expression des codierten Produktes, das anschließend gereinigt werden kann. Die Großherstellung vieler Genprodukte ist in vielen Fällen besonders wichtig, wenn solche Produkte medizinischen oder industriellen Wert haben.

[0003] Trotz der Fortschritte der Genexpression, stellten jedoch bestimmte codierende Sequenzen ihre Produkte kaum in stabiler Form her. Zum Beispiel könnte die Expression von rekombinanten Proteinen in E. coli insbesondere für Proteine mit Transmembrandomänen oder umfangreichen hydrophoben Sequenzen problematisch sein. Darüber hinaus können rekombinante Proteine keine endterminalen Aminosäurereste mit der geeigneten Codonausrichtung enthalten. Folglich besteht die Notwendigkeit verbesserte Materialien und Verfahren zur Expression von rekombinanten Genen bereitzustellen.

[0004] WO 99/42076 und WO 99/51748 offenbaren das M. tuberculosis-Antigen Ra12 und Fusionsproteine, die dieses enthalten.

Zusammenfassung der Erfindung

[0005] Die Erfindung stellt zum ersten Mal Verfahren unter Verwendung von rekombinanten Nukleinsäuremolekülen, die Fusionspolypeptide codieren, umfassend ein Ra12-Polypeptid und ein heterologes Polypeptid, Fusionspolypeptide, Expressionsvektoren und Wirtszellen, umfassend diese Nukleinsäuremoleküle, um eine stabile Expression von Fusionspolypeptiden von Interesse mit hoher Ausbeute bereit.

[0006] In einem Aspekt stellt die vorliegende Erfindung bereit, ein Verfahren zur stabilen Expression mit hoher Ausbeute eines rekombinanten heterologen Polypeptids, wobei in dem Verfahren ein für ein Fusionspolypeptid codierendes rekombinantes Nukleinsäuremolekül in einer Wirtszelle exprimiert wird, wobei das Fusionspolypeptid ein Ra12-Polypeptid und ein heterologes Polypeptid umfaßt, wobei das Ra12-Polypeptid von einer Ra12-Polypeptidnukleotidsequenz codiert wird, die unter stringenten Bedingungen an eine zu SEQ ID NO:3 komplementäre Sequenz hybridisiert, und wobei die Ra12-Polynukleotidsequenz 5' zu der für das heterologe Polypeptid codierenden heterologen Polynukleotidsequenz lokalisiert ist. In einer weiteren Ausführungsform umfassen die rekombinanten Nukleinsäuremoleküle außerdem eine Polynukleotidsequenz, die ein Linkerpeptid zwischen der Ra12-Polynukleotidsequenz und der heterologen Polynukleotidsequenz codiert, wobei das Linkerpeptid eine Spaltstelle umfassen kann. In wiederum einer anderen Ausführungsform codieren die rekombinanten Nukleinsäuremoleküle Fusionspolypeptide die außerdem ein Affinitäts-Taq umfassen. In wiederum einer anderen Ausführungsform codieren die rekombinanten Nukleinsäuremoleküle ein Fusionspolypeptid, umfassend ein DPPD, eine WT1, ein Mammaglobin oder ein H9-32A heterologes Polypeptid. In wiederum einer anderen Ausführungsform umfassen die rekombinanten Nukleinsäuremoleküle eine Ra12-Polynukleotidsequenz, umfassen mindestens ungefähr 30 Nukleotide, mindestens ungefähr 60 Nukleotide oder mindestens ungefähr 100 Nukleotide. In wiederum einer anderen Ausführungsform umfassen die rekombinanten Nukleinsäuremoleküle eine Ra12-Polynukleotidsequenz, die in SEQ ID NO:3 gezeigt ist. In wiederum einer anderen Ausführungsform umfassen die rekombinanten Nukleinsäuremoleküle eine Ra12-Polynukleotidsequenz, die in SEQ ID NO:4, SEQ ID NO:17 oder SEQ ID NO:18 gezeigtes Ra12-Polynukleotid codiert.

[0007] In einer bevorzugten Ausführungsform ist die Wirtszelle E. coli.

[0008] In einer Ausführungsform umfasst das Ra12-Polypeptid mindestens ungefähr 10 Aminosäuren, mindestens ungefähr 30 Aminosäuren oder mindestens ungefähr 100 Aminosäuren. In einer weiteren Ausführungsform weist das Ra12-Polynukleotid eine in SEQ ID NO:4, SEQ ID NO:17 oder SEQ ID NO:17 gezeigte Sequenz auf.

[0009] In einer Ausführungsform umfasst das Verfahren außerdem das Reinigen der Fusionspolypeptide nach deren Expression. In einer weiteren Ausführungsform umfasst das Verfahren außerdem das Spalten eines Fusionspolypeptids zwischen einem Ra12-Polypeptid und einem heterologen Polypeptid.

[0010] Diese und weitere erfindungsgemäßen Aspekte werden durch die folgende detaillierte Beschreibung und durch die beigefügten Abbildungen deutlich.

Kurze Beschreibung der Abbildungen

[0011] **Fig. 1** stellt eine Nukleotidsequenz (SEQ ID NO:1) und eine Aminosäuresequenz (SEQ ID NO:2) von MTB32A dar.

[0012] **Fig. 2** stellt eine Nukleotidsequenz (SEQ ID NO:3) und eine Aminosäuresequenz (SEQ ID NO:4) von Ra12 dar.

[0013] **Fig. 3** stellt eine rekombinante Nukleinsäuresequenz, umfassend eine Nukleotidsequenz (SEQ ID NO:5) und eine Aminosäuresequenz (SEQ ID NO:6) des Ra12-DPPD-Fusionspolypeptids dar.

[0014] **Fig. 4** stellt eine rekombinante Nukleinsäuresequenz, umfassend eine Nukleotidsequenz (SEQ ID NO:7) und eine Aminosäuresequenz (SEQ ID NO:8) des Ra12-WT1-Fusionspolypeptids dar.

[0015] **Fig. 5** stellt eine rekombinante Nukleinsäuresequenz, umfassend eine Nukleotidsequenz (SEQ ID NO:9) und eine Aminosäuresequenz (SEQ ID NO:10) des Ra12-Mammaglobin-Fusionspolypeptids dar.

[0016] **Fig. 6** stellt eine rekombinante Nukleotidsequenz, umfassend eine Nukleotidsequenz (SEQ ID NO:11) und eine Aminosäuresequenz (SEQ ID NO:12) des Ra12-H9-32A-Fusionspolypeptids dar.

[0017] **Fig. 7** stellt das Ra12 (kurze)-Polypeptid (SEQ ID NO:17), das die Aminosäuren 1-30 der SEQ ID NO:3 aufweist, dar.

[0018] **Fig. 8** stellt das Ra12 (lange)-Polypeptid (SEQ ID NO:18), das die Aminosäuren 1-128 der SEQ ID NO:4 aufweist, dar.

[0019] **Fig. 9** stellt ein Konstrukt eines Ra12 (kurzen)-Polypeptids, dar, das an ein humanes Mammaglobin-Gen gebunden ist.

Kurze Beschreibung der Erfindung

[0020] Wie oben erwähnt, stellt die vorliegende Erfindung zum ersten Mal Verfahren bereit unter Verwendung von Ra12-Sequenzen als Fusionspartner, um die stabile Expression mit hoher Ausbeute von rekombinanten heterologen Polypeptiden sowohl eukaryotischen als auch prokaryotischen Ursprungs zu erleichtern.

[0021] MTB32A ist eine Serinprotease mit einem Molekulargewicht von 32 kDa, die durch ein Gen in virulenten und avirulenten Stämmen von *M. tuberculosis* codiert wird. Die vollständige Nukleotidsequenz (SEQ ID NO:1) und Aminosäuresequenz (SEQ ID NO:2) von MTB32A sind in **Fig. 1** offenbart. Siehe außerdem Skeiky et al., *Infection and Immun.* (1999) 67:3998-4007. Dieses Protein wird normalerweise in den Überstand von Bakterienkulturen abgesondert. Der offene Leserahmen der codierenden Sequenz enthält endterminale hydrophobe Sekretionssignale. Er stimuliert periphere mononukleare Blutzellen von gesunden gereinigten Protein-derivat (PPd)-positiven Spendern, um Interferon zu vermehren und abzusondern. Folglich ist MTB32A ein Kandidat-Antigen zur Verwendung bei einer Impfstoffentwicklung gegen Tuberkulose.

[0022] Es wurde überraschenderweise durch die vorliegenden Erfinder festgestellt, dass ein 14 KD-C-terminales Fragment der MTB32A-codierenden Sequenz selbst in hohen Mengen exprimiert wird und während des Reinigungsverfahrens ein lösliches Protein bleibt. Dieses 14 KC-C-terminale MTB32A-Fragment wird hiernach als Ra12 bezeichnet (mit den Aminosäureresten 192 bis 323 von MTB32A). Die Nukleinsäure- und Aminosäure-

resequenzen von nativem Ra12 sind gezeigt, z.B. in den [Fig. 2-6](#). Wie hiernach im Detail beschrieben wird, bezieht sich der Ausdruck "Ra12-Polypeptid" oder "Ra12-Polynukleotid" hier auf die nativen Ra12-Sequenzen (z.B. SEQ ID NO:3 oder SEQ ID NO:4), deren Varianten oder Fragmente davon (z.B. SEQ ID NO:17 oder SEQ ID NO:18). Die vorliegende Erfindung nutzt die Eigenschaften der Ra12-Polypeptide und stellt Verfahren bereit zur stabilen Expression mit hoher Ausbeute von Fusionspolypeptiden, umfassend ein Ra12-Polypeptid und ein heterologes Polypeptid von Interesse. Die erfindungsgemäßen Verfahren eignen sich insbesondere zur Expression von bestimmten heterologen Polypeptiden (z.B. DPPD), die von anderen herkömmlichen Expressionsverfahren nicht in beachtlicher Menge hergestellt werden.

Rekombinante Fusionsnukleinsäuren

[0023] Rekombinante Nukleinsäuren, die ein Fusionspolypeptid, umfassend ein Ra12-Polypeptid und ein heterologes Polypeptid von Interesse codieren, können leicht durch herkömmliche Gentechnikverfahren erzeugt werden. Rekombinante Nukleinsäuren werden hergestellt, so dass sich eine Ra12-Polynukleotidsequenz 5' zu der ausgewählten heterologen Polynukleotidsequenz befindet.

[0024] Erfindungsgemäß kann jedes geeignete heterologe Polypeptid von Interesse als Fusionspartner der Ra12-Nukleinsäuren ausgewählt werden, um ein Fusionspolypeptid zu erzeugen. Eine "heterologe Sequenz" oder eine "heterologe Nukleinsäure", wie sie hier verwendet wird, ist eine, die von einer Quelle abstammt, die für die ausgewählte Wirtszelle fremd ist oder wenn sie von der gleichen Quelle abstammt, ist gegenüber ihrer ursprünglichen Form modifiziert. Folglich umfasst eine heterologe Nukleinsäure in einer prokaryotischen Wirtszelle eine heterologe Nukleinsäure, die endogen zu der bestimmten Wirtszelle ist, die modifiziert worden ist. Die Modifizierung der heterologen Sequenz kann zum Beispiel auftreten durch Behandeln der DNA mit einem Restriktionsenzym, um ein DNA-Fragment zu erzeugen, das an einen Promotor operativ gebunden werden kann. Techniken der ortsspezifischen Mutagenese sind ebenfalls zum Modifizieren einer heterologen Sequenz nützlich.

[0025] Eine heterologe Nukleinsäure sowohl von eukaryotischen als auch prokaryotischen Ursprung kann als Fusionspartner ausgewählt werden. Diese Nukleinsäuren umfassen, sind aber nicht darauf beschränkt, Nukleinsäuren, die pathogene Antigene, bakterielle Antigen, virale Antigene, Krebsantigene, Tumorantigene und Tumorunterdrücker codieren. Beispielhafte heterologe Nukleinsäuren von Interesse umfassen DPPD, WT1, Mammaglobin, H9-32A-Nukleinsäuren und anderer Mycobacterium tuberculosis-Nukleinsäuren (siehe z.B. Cole et al., Nature (1999), 393:537-544; <http://www.sanger.ac.uk> und <http://www.pasteur.fr/mycdb/> für die vollständigen Genomsequenzen von M. tuberculosis siehe außerdem WO 98/53075 und WO 98/53076, von denen beide am 26. November 1998 veröffentlicht wurden, für Nukleinsäuresequenzen, die M. tuberculosis-Proteine codieren). Jede der hier offenbarten Nukleinsäuren kann allein oder in Kombination als eine heterologe Nukleinsäure verwendet werden, die als Fusionspartner ausgewählt werden kann.

[0026] Zusätzlich kann jedes geeignete Ra12-Polynukleotid (z.B. natives Ra12-Polynukleotid mit der SEQ ID NO:3, Varianten oder Fragmente davon) bei der Herstellung von erfindungsgemäßen rekombinanten Fusionsnukleinsäuren verwendet werden. Bevorzugte Ra12-Polynukleotide umfassen mindestens ungefähr 15 konsekutive Nukleotide, mindestens ungefähr 30 Nukleotide, mindestens etwa 60 Nukleotide, mindestens ungefähr 100 Nukleotide, mindestens ungefähr 200 Nukleotide, oder mindestens ungefähr 300 Nukleotide. Polynukleotide können einzel- oder doppelsträngig sein und können DNA (Genom, cDNA oder synthetische) oder RNA-Moleküle sein.

[0027] In einer Ausführungsform ist die Ra12-Polynukleotidsequenz die in SEQ ID NO:3 gezeigt. In einer weiteren Ausführungsform codiert die Ra12-Polynukleotidsequenz ein Ra12-Polynukleotid, das in SEQ ID NO:4 gezeigt ist. In einigen weiteren Ausführungsformen umfasst die Ra12-Polynukleotidsequenz einen Teil der SEQ ID NO:3 oder codiert einen Teil der SEQ ID NO:4. Zum Beispiel kann ein Ra12-Polynukleotid, umfassend 90 Nukleotide (z.B. Nukleotide 1-90 der SEQ ID NO:3) oder ein Ra12-Polynukleotid, umfassend 384 Nukleotide (z.B. Nukleotide 1-384 der SEQ ID NO:3) als Fusionspartner verwendet werden. Siehe nachfolgende Beispiele 2 und 3.

[0028] Polynukleotide können eine native Sequenz (d.h. eine endogene Sequenz, die ein Ra12-Polynukleotid der SEQ ID NO:3 oder einen Teil davon codiert), oder eine Variante einer solchen Sequenz umfassen. Polynukleotidvarianten können ein oder mehrere Substitutionen, Additionen und/oder Insertionen enthalten, so dass die biologische Aktivität des codierten Fusionspolypeptids im Verhältnis zu einem Fusionspolypeptid, umfassend ein natives Ra12-Polynukleotid nicht verringert ist. Varianten entfalten bevorzugt mindestens ungefähr 70% Identität, bevorzugter mindestens ungefähr 80% Identität und am meisten bevorzugt mindestens unge-

fähr 90% Identität zu einer Polynukleotidsequenz, die ein natives Ra12-Polypeptid (SEQ ID NO:4) oder einen Teil davon codiert. Optional besteht Identität über einen Bereich, der mindestens ungefähr 25 bis 50 Aminosäuren oder Nukleotide lang ist oder optional über einen Bereich der 75-100 Aminosäuren oder Nukleotide lang ist.

[0029] Zwei Polynukleotid- oder Polypeptidsequenzen werden als "identisch" bezeichnet wenn die Sequenz der Nukleotide oder Aminosäuren in den beiden Sequenzen die gleiche ist wenn sie wie nachfolgend beschrieben auf maximale Übereinstimmung abgeglichen ist. Vergleiche zwischen zwei Sequenzen werden üblicherweise durch Vergleichen der Sequenzen in einem Vergleichsfenster durchgeführt, um lokale Bereiche der Sequenzähnlichkeit zu identifizieren und zu vergleichen. Ein "Vergleichsfenster" wie es hier verwendet wird, betrifft ein Segment von mindestens ungefähr 20 zusammenhängenden Positionen, üblicherweise 30 bis ungefähr 75, z.B. 40 bis ungefähr 50, wobei eine Sequenz mit einer Bezugssequenz mit der gleichen Anzahl von zusammenhängenden Positionen verglichen wird, nachdem die beiden Sequenzen optimal angeordnet wurden.

[0030] Eine optimale Anordnung der Sequenzen zum Vergleichen kann durchgeführt werden unter Verwendung des Megalign-Programms in der Lasergene suite der bioinformatics Software (DNASTAR, Inc., Madison, WI) unter Verwendung von Standardeinstellungen. Dieses Programm enthält verschiedene Schemata zur Anordnung, die in den folgenden Referenzen beschrieben sind: Dayhoff, M.O. (1978), A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (Hrsg.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC, Bd. 5, Suppl. 3, S. 345-358; Hein, J. (1990), Unified Approach to Alignment and Phylogenies, S. 626-645, Methods in Enzymology Bd. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. und Sharp, P.M. (1989), CABIOS5:151-131; Myers, E.W. und Muller, W. (1988), CABIOS4: 11-17; Robinson, E.D. (1971), Comb. Theor. 11: 105; Santou, N., Nes, M. (1987), Mol. Biol. Evol., 4:406-425; Sneath, P.H.A. und Sokal, R.R. (1973), Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy, Freeman Press, San Francisco, CA; Wilbur, W.J. und Lipman, D.J. (1983), Proc. Natl. Acad. Sci., USA 80:726-730.

[0031] Alternativ kann zum Vergleich die optimale Anordnung der Sequenzen durch den lokalen Identitätsalgorithmus von Smith und Waterman (1981), Add. APL. Math 2:482, durch den Identitätsanordnungsalgorithmus von Needleman und Wunsch (1970), J. Mol. Biol. 48:443, durch die Suche nach Ähnlichkeitsverfahren von Pearson und Lipman (1988) Proc. Natl. Acad. Sci. USA 85: 2444, durch computerisiertes Durchführung dieser Algorithmen (GAP, BESTFIT, BLAST, FASTA und TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI) oder durch Betrachtung durchgeführt werden.

[0032] Bevorzugte Beispiele von Algorithmen, die sich zur Bestimmung der prozentualen Sequenzidentität und Sequenzähnlichkeit eignen sind die BLAS- und BLAST 2.0-Algorithmen, die von Altschul et al. (1977), Nucl. Acids Res. 25: 3389-3402 und Altschul et al. (1990), J. Mol. Biol. 215:403-410 jeweils beschrieben sind. BLAST und BLAST 2.0 können zum Beispiel mit den hier beschriebenen Parametern verwendet werden, um die prozentuale Sequenzidentität der erfindungsgemäßen Nukleotide und Polypeptide zu bestimmen. Die Software zum Durchführen der BLAST-Analysen ist öffentlich zugänglich durch das National Center für Biotechnology Information: Für Aminosäuresequenzen kann eine Punktematrix verwendet werden, um den kumulativen Stand zu berechnen. Die Erweiterung des Wortes Treffer in jeder Richtung werden festgehalten wenn: der kumulative Anordnungsstand in der Menge X von seinem maximal erreichten Wert abfällt; der kumulative Stand aufgrund der Anhäufung von ein oder mehreren Restanordnungen mit negativen Punkten in Richtung Null oder darunter geht, oder das Ende einer von beiden Sequenzen erreicht wird. Die BLAST-Algorithmusparameter W, T und X bestimmen die Empfindlichkeit und die Geschwindigkeit der Anordnung.

[0033] Bei einem bevorzugten Verfahren wird der "prozentuale Anteil der Sequenzidentität" durch Vergleichen von zwei optimal angeordneten Sequenzen über ein Vergleichsfenster von mindestens 20 Positionen bestimmt, wobei der Teil der Polypeptidsequenz in dem Vergleichsfenster zur optimalen Anordnung der beiden Sequenzen Additionen oder Deletionen (z.B. Zwischenräume) von 20% oder weniger, gewöhnlich 5 bis 15% oder 10 bis 12% im Vergleich zu den Bezugssequenzen (die keine Addition oder Deletionen umfassen). Der Prozentsatz wird berechnet durch Bestimmung der Zahl der Positionen an denen der identische Aminosäurerest in beiden Sequenzen auftritt, um die Zahl der passenden Positionen zu ergeben, wobei die Zahl der passenden Positionen durch die Gesamtzahl der Positionen in der Referenzsequenz (d.h. der Fenstergröße) dividiert und das Ergebnis mit 100 multipliziert wird, um den Prozentsatz der Sequenzidentität zu ergeben.

[0034] Varianten können außerdem oder alternativ im Wesentlichen homolog zu einem nativen Ra12-Poly-

peptid (z.B. SEQ ID NO:3) oder einem Teil oder Komplement davon sein. Solche Nukleotidvarianten können unter stringenten Bedingungen an eine in der Natur vorkommende DNA-Sequenz, die ein natives Ra12-Polypeptid (oder eine Komplementärsequenz) codiert, hybridisieren.

[0035] Der Ausdruck "selektives (oder spezifisches) Hybridisieren an" betrifft die Bindung, das Doppelsträngigmachen oder Hybridisieren nur eines Moleküls an eine bestimmte Nukleotidsequenz unter stringenten Hybridisierungsbedingungen wenn die Sequenz in einer komplexen Mischung (z.B. Zellulär- oder Bibliothek-DNA oder -RNA) vorhanden ist.

[0036] Der Ausdruck "stringente Hybridisierungsbedingungen" bezieht sich auf Bedingungen, unter denen eine Sonde an seine Ziel-Subsequenz hybridisieren würde, gewöhnlich in einer komplexen Mischung aus Nukleinsäuren, aber nicht an andere Sequenzen. Stringente Bedingungen sind sequenzabhängig und unterscheiden sich bei unterschiedlichen Verhältnissen. Längere Sequenzen hybridisieren speziell bei höheren Temperaturen. Ein ausführlicher Leitfaden zur Hybridisierung von Nukleinsäuren befindet sich in Tijssen, *Techniques in Biochemistry and Molecular Biology – Hybridization with Nucleic Probes*, "Überblick der Hybridisierungsprinzipien und der Strategie von Nukleinsäureuntersuchungen" (1993). Im allgemeinen werden stringente Bedingungen ausgewählt, wobei sie bei einer definierten Ionenstärke-pH ungefähr 5-10°C geringer als der Schmelzpunkt (T_m) der spezifischen Sequenz sind. Der T_m ist die Temperatur (unter definierter Ionenstärke, pH und Nukleinkonzentration) bei der 50% der Sonden, die komplementär mit dem Ziel sind, an die Zielsequenz im Gleichgewichtszustand hybridisieren (da die Zielsequenzen im Überschuss bei T_m vorhanden sind, sind 50% der Sonden im Gleichgewichtszustand besetzt). Stringente Bedingungen sind solche, bei denen die Salzkonzentration weniger als ungefähr 1,0 M Natriumion-, gewöhnlich ungefähr 0,01 bis 1,0 M Natriumion-Konzentration (oder anderen Salzen) bei pH 7,0 bis 8,3 beträgt und die Temperatur weniger als ungefähr 30°C für kurze Sonden (z.B. 10 bis 50 Nukleotide) und mindestens ungefähr 60°C für längere Sonden (z.B. größer als 50 Nukleotide) beträgt. Stringente Bedingungen können auch durch Zugabe von Destabilisierungsagentien, wie Formamid, erreicht werden. Bei einer selektiven oder spezifischen Hybridisierung bedeutet ein positives Signal, dass es mindestens zwei-mal über der Hintergrundhybridisierung, bevorzugt 10-mal über der Hintergrundhybridisierung liegt. Beispielhafte stringente Hybridisierungsbedingungen können wie folgt sein: 50% Formamid, 5 × SSC und 1% SDS, bei 42°C inkubiert oder 5 × SSC, 1% SDS, bei 65°C inkubiert, mit einer Waschung in 0,2 × SSC und 0,1% SDS bei 65°C.

[0037] Es würde vom Fachmann anerkannt werden, dass, als Folge der Degeneration des genetischen Codes, viele Nukleotidsequenzen vorkommen, die ein Ra12-Polypeptid, wie oben beschrieben, codieren. Einige dieser Polynukleotide weisen eine Minimalhomologie zu der Nukleotidsequenz von jedem natürlichen Gen auf. Gleichwohl sind Polynukleotide, die aufgrund von Unterschieden in der Codonverwendung variieren, erfindungsgemäß besonders zu erwägen. Darüber hinaus sind Allele der Gene, umfassend die Polynukleotidsequenzen, die hier bereitgestellt werden, erfindungsgemäß umfaßt. Allele sind endogene Gene, die aufgrund von ein oder mehreren Mutationen, wie Deletionen Additionen und/oder Substitutionen von Nukleotiden geändert sind. Die resultierende mRNA oder das Protein können, müssen aber nicht, eine geänderte Struktur oder Funktion aufweisen. Allele können unter Verwendung von Standardtechniken (wie Hybridisierung, Amplifikation und/oder Sequenzvergleich in Datenbanken) identifiziert werden.

[0038] Folglich bezieht sich der Begriff "Ra12-Polynukleotid" oder "Ra12-Polynukleotidsequenz", wie er hier verwendet wird, auf native Polynukleotidsequenzen (z.B. SEQ ID NO:3), Fragmente davon oder jede Variante davon. Funktionsgemäß hat jedes Ra12-Polynukleotid die Fähigkeit ein Fusionsprotein herzustellen und seine Fähigkeit, Fusionsproteine in Wirtszellen herzustellen, kann bezüglich des nativen Ra12-Polynukleotids (z.B. SEQ ID NO:3) verbessert oder unverändert sein, oder kann bezüglich des nativen Ra12-Polynukleotids weniger als 50% und bevorzugt weniger als 20%, reduziert sein.

[0039] Nukleinsäuren, die erfindungsgemäße Ra12-Polynukleotide codieren, können durch jedes im Stand der Technik geeignete Verfahren hergestellt werden. Beispiele der Verfahren umfassen das Klonieren und die Restriktion von geeigneten Sequenzen oder die direkte chemische Synthese durch Verfahren wie das Phosphodiester-Verfahren von Narang et al. (1979), *Meth. Enzymol.* 68: 90-99; das Phosphodiester-Verfahren von Brown et al. (1979), *Meth. Enzymol.* 68: 109-151; das Diethylphosphoramidit-Verfahren von Beaucage et al. (1981), *Tetra. Lett.* 22: 1859-1862; und das trägerfixierte Verfahren des US-Patents Nr. 4,458,066.

[0040] In einer Ausführungsform wird eine MTB32A- oder Ra12-codierende Nukleinsäure durch routinemäßige Klonierungsverfahren isoliert. Nukleotidsequenzen von MTB32A oder Ra12, die hier bereitgestellt werden, werden verwendet, um Sonden bereitzustellen, die speziell an andere MTB32A- oder RA12-Nukleinsäuren in einer Genom-DNA-Probe oder an eine MTB32A-mRNA oder Ra12-mRNA in einer RNA-Gesamtprobe (z.B. in

einem Southern- oder Northern-Blot) hybridisieren. Sobald die Ziel-MTB32A- oder Ra12-Nukleinsäuren identifiziert sind, können sie gemäß Standardverfahren, die dem Fachmann bekannt sind, isoliert werden.

[0041] Die gewünschten Nukleinsäuren können außerdem durch gutbekannte Amplifikationsmaßnahmen kloniert werden. Beispiele von Protokollen, die ausreichen, um dem Fachmann durch ein In-vitro-Amplifikationsverfahren zu leiten, umfassen die Polymerase-Kettenreaktion (PCR), die Ligase-Kettenreaktion (LCR), die Q β -Replikaseamplifizierung und andere RNA-Polymerasevermittelte Maßnahmen, die zu finden sind in Berger, Sambrook und Ausubel sowie in Mullis et al. (1987), US-Patent Nr. 4,683,202; PCR Protocols A Guide to Methods and Applications (Innis et al., Hrsg.), Academic Press Inc., San Diego, CA (1990) (Innis); Arnheim & Levinson (1. Oktober 1990), C & EN 36-47; The Journal Of NIH Research (1991) 3: 81-94; (Kwoh et al. (1989), Proc. Natl. Acad. Sci. USA 86: 1173; Guatelli et al. (1990), Proc. Natl. Acad. Sci. USA 87: 1874; Lomell et al. (1988), J. Clin. Chem. 35: 1826; Landegren et al. (1988), Science 241: 1077-1080; Van Brunt (1990), Biotechnology 8: 291-294; Wu und Wallace (1989), Gene 4: 560; und Barringer et al. (1990), Gene 89: 117. Verbesserte Verfahren zum Klonieren von in vitro amplifizierten Nukleinsäuren sind im US-Patent Nr. 5,426,039 von Wallace et al. Beschrieben. Geeignete Primer, die bei der Amplifikation der erfindungsgemäßen Nukleinsäuren verwendet werden können, können auf Basis der hier bereitgestellten Sequenzen entworfen werden.

[0042] Die MTB32A- oder Ra12-Nukleinsäuren können außerdem durch Nachweisen ihres exprimierten Produktes mittels Untersuchungen auf Basis der physikalisch, chemischen oder immunologischen Eigenschaften des exprimierten Proteins kloniert werden. Zum Beispiel kann eine klonierte MTB32A- oder Ra12-Nukleinsäure identifiziert werden durch die Fähigkeit eines Polypeptids, das durch die Nukleinsäure codiert wird, an Antisera oder gereinigte Antikörpern zu binden, die gegen die hier bereitgestellten MTB32A- oder Ra12-Polypeptide gerichtet sind, und die außerdem MTB32A- oder Ra12-Homologe erkennen und selektiv daran binden,.

[0043] In einigen Ausführungsformen kann es wünschenswert sein, die erfindungsgemäßen MTB32A- oder Ra12-Nukleinsäuren zu modifizieren. Geänderte Nukleotidsequenzen, die erfindungsgemäß verwendet werden können, umfassen Deletionen, Additionen oder Substitutionen verschiedener Nukleotidreste, was zu einer Sequenz führt, die das gleiche oder ein funktional äquivalentes Genprodukt codiert. Das Genprodukt selbst kann Deletionen, Additionen oder Substitutionen von Aminosäureresten enthalten, was zu einer stillen Änderung führt, wodurch ein funktional äquivalentes antigenes Epitop hergestellt wird. Solche konservativen Aminosäuresubstitutionen können auf Basis der Ähnlichkeit der Polarität, Ladung, Löslichkeit, hydrophoben Eigenschaften, hydrophilen Eigenschaften und/oder der amphipathischen Natur der involvierten Reste erzeugt werden. Bevorzugt können Ra12-Nukleinsäuren, die in Länge kürzer als SEQ ID NO:3 sind und biologisch aktive Fusionspartner codieren, verwendet werden. Solche kleineren funktionalen Äquivalente der Ra12-Polypeptide können wünschenswert sein, um die Menge der Wirtszellquellen, die zur Herstellung von heterologen Polypeptiden von Interesse verfügbar sind, zu steigern.

[0044] Der Fachmann erkennt viele Wege der Erzeugung von Änderungen in einem gegebenen Nukleinsäurekonstrukt. Solche gut bekannten Verfahren umfassen die ortsgerichtete Mutagenese, die PCR-Amplifikation unter Verwendung von degenerierten Oligonukleotiden, die Aussetzung der Zellen, die die Nukleinsäure enthalten, mutagenen Agentien oder Bestrahlung, chemische Synthese eines gewünschten Oligonukleotids (z.B. in Verbindung Ligation und/oder Klonierung, um große Nukleinsäuren zu erzeugen) und andere bekannte Verfahren. Siehe z.B. Gilman and Smith (1979) Gene 8:81-97, Roberts et al. (1987), Nature 328: 731-734.

[0045] Rekombinante Nukleinsäuren, die ein Fusionspolypeptid codieren, das ein Ra12-Polynukleotid und ein ausgewähltes heterologes Polypeptid umfaßt, können unter Verwendung jedes im Stand der Technik bekannten Verfahrens hergestellt werden. Wie oben beschrieben werden rekombinante Nukleinsäuren erzeugt, so dass eine Ra12-Polynukleotidsequenz sich 5' zu der ausgewählten heterologen Polynukleotidsequenz befindet. Ra12 und heterologe Polynukleotidsequenzen können außerdem modifiziert werden, um ihre Fusion und anschließende Expression der Fusionspolypeptide zu erleichtern. Zum Beispiel kann das 3'-Stopp-Codon der Ra12-Polynukleotidsequenz mit einer Linkersequenz im Raster substituiert sein, die Restriktionsstellen und/oder Spaltstellen bereitstellen kann. Die rekombinanten Nukleinsäuren können außerdem andere Nukleotidsequenzen umfassen, wie Sequenzen, die Affinitätsmarkierungen codieren, um Proteinreinigungsprotokolle zu unterstützen.

Expressionsvektoren und Wirtszellen

[0046] Die rekombinanten Nukleinsäuren, die hier beschrieben sind, können mit einer Vielzahl von anderen Nukleotidsequenzen unter Verwendung von etablierten rekombinanten DNA-Verfahren aneinander gefügt sein. Zum Beispiel kann ein Polynukleotid in eine Vielzahl von Klonierungsvektoren, einschließlich Plasmiden,

Phagemiden, lambda-Phagen-Derivaten und Cosmiden kloniert werden. Vektoren von besonderem Interesse umfassen Expressionsvektoren, Replikationsvektoren, Sondenerzeugungsvektoren und Sequenzierungsvektoren. Im Allgemeinen enthält ein Vektor einen Replikationsursprung, der in mindestens einem Organismus funktional ist, geeignete Restriktionsendonukleasestellen und einen oder mehrere Selektionsmarker. Andere Elemente hängen von der gewünschten Verwendung ab und sind für den Fachmann ersichtlich.

[0047] DNA-Sequenzen, die Polypeptidsequenzen codieren, können getrennt zusammengefügt werden und in einen geeigneten Expressionsvektor ligiert werden. Das 3'-Ende der DNA-Sequenz, die eine Polypeptid-Komponente codiert, wird mit oder ohne eine Polynukleotidsequenz, die einen Peptidlinker codiert, an das 5'-Ende einer DNA-Sequenz, die die zweite Polypeptidkomponente codiert, ligiert, so dass die offenen Lese-rahmen der Sequenzen phasengleich sind. Dies ermöglicht die Translation in ein einziges Fusionsprotein, das die biologische Aktivität der beiden Polypeptidkomponenten beibehält.

[0048] Die ligierten DNA-Sequenzen werden operativ an geeignete Transkriptions- oder Translationsregulationselemente gebunden. Die regulatorischen Elemente, die für die DNA-Expression verantwortlich sind, befinden sich nur 5' der DNA-Sequenz, die das erste Polypeptid codiert. Gleichermaßen befinden sich Stopp-Codone, die für das Ende der Translation erforderlich sind und Transkriptionsterminationssignale ausschließlich 3' der DNA-Sequenz, die das zweite Polypeptid codiert.

[0049] In Abhängigkeit des verwendeten Wirts/Vektorsystems kann jedes einer Vielzahl von geeigneten Transkriptions- und Translationselementen, einschließlich konstitutiver und induzierbarer Promotoren, in dem Expressionsvektor verwendet werden. Zum Beispiel wenn in Bakteriensysteme kloniert wird, können induzierbare Promotoren, wie pL des Bakteriophagen λ , plac, ptrp, ptac (ptrp-lac-Hybrid-Promotor, Cytomegalovirus-Promotor) und dgl. verwendet werden; wenn in Hefezellsysteme kloniert wird, können Promotoren wie ADHI, PGK, PHO5 und der α -Faktor-Promotor verwendet werden; wenn in Insektenzellsysteme kloniert wird, können Promotoren wie der Baculovirus polyhedron-Promotor verwendet werden; wenn in Pflanzenzellsysteme kloniert wird, können Promotoren, die vom Genom von Pflanzenzellen abstammen (z.B. Hitzeschockpromotoren, der Promotor für die kleine Untereinheit von RUBISCO; der Promotor für das Chlorophyll α/β -Bindungsprotein) oder von Pflanzenviren (z.B. der 35S-RNA-Promotor von CaMV, der Hüllproteinpromotor von TMV) verwendet werden. Wenn in Säugerzellsysteme kloniert wird, können Promotoren, die vom Genom von Säugerzellen abstammen (z.B. Metallothionein-Promotor) oder von Säugerviren (z.B. der Adenovirus-späte-Promotor; der Vacciniavirus-7.5K-Promotor) verwendet werden, wenn Zelllinien erzeugt werden, die Mehrfachkopien einer Antigen codierenden Sequenz enthalten, können Vektoren auf Basis von SV40, BPV und EBV mit einem geeigneten Selektionsmarker verwendet werden.

[0050] Eine Vielzahl von Wirtsexpressionsvektorsystemen kann eingesetzt werden, um Ra12-Fusionsprotein-codierende Sequenzen zu exprimieren. Diese umfassen, sind aber nicht darauf beschränkt, Mikroorganismen, wie Bakterien (z.B. E. coli, B. subtilis), die mit rekombinanter Bakteriophagen-DNA, Plasmid-DNA oder Cosmid-DNA-Expressionsvektoren, enthaltend eine codierende Sequenz, transformiert sind; Hefe (z.B. Saccharomyces, Pichia), die mit rekombinanten Hefeexpressionsvektoren, enthaltend eine codierende Sequenz, transformiert sind; Insektenzellsysteme, die mit rekombinanten Virusexpressionsvektoren (z.B. Baculovirus), enthaltend eine codierende Sequenz, infiziert sind; Pflanzenzellsysteme, die mit rekombinanten Virusexpressionsvektoren (z.B. Blumenkohlmosaikvirus, CaMV, Tabakmosaikvirus, TMV) infiziert oder mit rekombinanten Plasmidexpressionsvektoren (z.B. Ti-Plasmid), enthaltend eine codierende Sequenz, transformiert sind; oder Säugerzellsysteme (z.B. COS-, CHO-, BHK-, 293-, 3T3-Zellen). Die Expressionselemente dieser Systeme variieren in ihrer Stärke und ihren Spezifitäten.

[0051] Bakteriensysteme sind zur Expression von Ra12-Fusionspolypeptiden bevorzugt. Gewöhnlich verwendete prokaryotische Kontrollsequenzen, die wie hier definiert Promotoren zum Transkriptionsbeginn, optional mit einem Operator zusammen mit Ribosom-Bindungsstellensequenzen umfassen, umfassen solche gewöhnlich verwendeten Promotoren, wie die beta-Lactamase (Penicillinase) und Lactose (lac)-Promotorsysteme (Change et al., Nature (1977), 198: 1056), das Tryptophan (trp)-Promotorsystem (Goeddel et al., Nucleic Acids Res. (1980) 8: 4057), den Tac-Promotor (DeBoer et al., Proc. Natl. Acad. Sci. USA (1983) 80:21-25) und den lambda abstammenden P_L-Promotor und die N-Gen-Ribosom-Bindungsstelle (Shimatake et al., Nature (1981), 292: 128). Das bestimmte Promotorsystem ist für die Erfindung unkritisch, jeder geeignete Promotor, der in Prokaryonten wirkt, kann verwendet werden.

[0052] Entweder konstitutive oder regulierende Promotoren können erfindungsgemäß verwendet werden. Regulierende Promotoren können vorteilhaft sein, da die Wirtszellen in hohen Dichten wachsen können bevor die Expression der Ra12-Fusionspolypeptide induziert wird. Die Expression von heterologen Proteinen in ho-

hen Mengen verlangsamt unter einigen Umständen das Zellwachstum. Regulierte Promotoren, die sich insbesondere zur Verwendung in *E. coli* eignen, umfassen den Bakteriophagen lambda P₁-Promotor, den Hybrid-trp-lac-Promotor (Amann et al., *Gene* (1983) 25: 167; de Boer et al., *Proc. Natl. Acad. Sci. USA* (1983), 80:21 und den Bakteriophagen T7-Promotor (Studier et al., *J. Mol. Biol.* (1986), Tabor et al. (1985). Diese Promotoren und ihre Verwendung werden in Sambrook et al. (1989), *Molecular Cloning: A Laboratory Manual*, 2. Ausgabe, Bd. 1-3, Cold Spring Harbor Laboratory, besprochen.

[0053] Zur Expression von Ra12-Fusionspolypeptiden in prokaryotischen Zellen außer *E. coli* ist ein Promotor, der in der bestimmten prokaryotischen Art wirkt, erforderlich. Solche Promotoren können aus Genen erhalten werden, die aus der Art kloniert worden sind oder heterologe Promotoren können verwendet werden. Zum Beispiel wirkt der Hybrid-trp-lac-Promotor aus *Bacillus* auch in *E. coli*.

[0054] Eine Ribosomen-Bindungsstelle (RBS) wird in geeigneter Weise in die erfindungsgemäßen Expressionskassetten eingeführt. Eine RBS in *E. coli*, zum Beispiel, besteht aus einer Nukleotidsequenz von 3-9 Nukleotiden in Länge, die sich 3-11 Nukleotide stromaufwärts des Anfangscodons befindet (Shine and Dalgarno, *Nature* (1975), 254: 34; Steitz, In *Biological regulation and development: Gene expression* (Hrsg. R.F. Goldberger), Bd. 1, 5. 349, 1979, Plenum Publishing, NY).

[0055] Wenn größere Mengen des Ra12-Fusionsproteins zu erzeugen sind, können Vektoren erwünscht sein, die die Expression von leicht zu reinigenden Fusionsproteinprodukten in hohen Mengen lenken. Solche Vektoren umfassen, sind aber nicht darauf beschränkt, den *E. coli*-Expressionsvektor pUR278 (Ruther et al. (1983), *EMBO J.* 2:1791), bei dem eine codierende Sequenz im Rahmen in dem Vektor mit dem lacZ-codierenden Bereich ligiert sein kann, so dass ein Hybridprotein erzeugt wird, pIN-Vektoren (Inouye und Inouye (1985), *Nucleic Acids Res.* 13:3101-3109; Van Heeke und Schuster (1989), *J. Biol. Chem.* 264:5503-5509); und dgl. pGEX-Vektoren können außerdem verwendet werden, um Fremdpolypeptide als Fusionsproteine mit Glutathion-S-Transferase (GST) zu exprimieren. Im Allgemeinen sind solche Fusionsproteine löslich und können leicht aus lysierten Zellen durch Adsorption an Glutathioagarosekügelchen, gefolgt von der Elution in der Gegenwart von freiem Glutathion gereinigt werden. Bei bestimmten Anwendungen kann es erwünscht sein, nach der Reinigung das heterologe Polypeptid von Interesse von dem Ra12-Fusionspolypeptid zu trennen. Dies kann durch jedes der unterschiedlichen Verfahren, die im Stand der Technik bekannt sind, erreicht werden. Zum Beispiel werden die pGEX-Vektoren so entworfen, dass sie eine Thrombin- oder Faktor Xa-Proteasespaltstelle umfassen, so dass das klonierte Fusionspolypeptid von Interesse aus dem GST-Anteil freigesetzt werden kann. Siehe z.B. Sambrook et al., *supra*; Itakura et al., *Science* (1977) 198:1056; Goeddel et al., *Proc. Natl. Acad. Sci. USA* (1979) 76:106; Naga et al., *Nature* (1984) 309:810; Sung et al., *Proc. Natl. Acad. Sci. USA* (1986) 83:561. Spaltstellen können biotechnologisch in den rekombinanten Nukleinsäuren für die Fusionsproteine an der gewünschten Spaltstelle eingeführt werden.

Fusionspolypeptide

[0056] Erfindungsgemäß umfasst ein "Fusionspolypeptid" mindestens zwei Teile: ein Ra12-Polypeptid, wie hier beschrieben, und ein heterologes Polypeptid von Interesse. Bei einem Fusionspolypeptid wird ein Ra12-Polypeptid bevorzugt an das Aminoende eines heterologen Polypeptids von Interesse direkt oder indirekt fusioniert, obgleich die Fusion an das Carboxyende des heterologen Polypeptids oder die Einführung des heterologen Polypeptids in eine Stelle innerhalb eines Ra12-Polypeptids ebenso geeignet sein kann.

[0057] Jedes heterologe Polypeptid von Interesse, entweder eukaryotischen oder prokaryotischen Ursprungs, kann als Fusionspartner für ein Ra12-Polypeptid ausgewählt werden. Diese heterologen Polypeptide umfassen, sind aber nicht darauf beschränkt, pathogene Antigene, Bakterienantigene, virale Antigene, Krebsantigene, Tumorantigene und Tumorunterdrücker. Beispiele heterologer Polypeptide umfassen DPPD, WT1, Mammaglobin, H9-32A-Polypeptide oder andere *M. tuberculosis*-Proteine. Jedes dieser Polypeptide kann allein oder in Kombination als heterologes Polypeptid, das als Fusionspartner ausgewählt werden kann, verwendet werden.

[0058] Wie oben bemerkt, kann ein Fusionspolypeptid ein natives Ra12-Polypeptid (z.B. SEQ ID NO:4), eine Variante davon oder ein Fragment davon umfassen. Eine Polypeptid-"Variante", wie hier verwendet, ist ein Polypeptid, das sich von einem nativen Ra12-Polypeptid durch eine oder mehrere Substitutionen, Deletionen, Additionen und/oder Insertionen unterscheidet, wobei die biologische Aktivität des Polypeptids nicht wesentlich vermindert ist. Mit anderen Worten kann die Fähigkeit einer Variante, ein Fusionspolypeptid in Wirtszellen herzustellen, im Verhältnis zu dem nativen Ra12-Protein verbessert oder unverändert sein oder kann um weniger als 50% und bevorzugt um weniger als 20% im Verhältnis zu dem nativen Ra12-Protein vermindert sein. Sol-

che Varianten können im allgemeinen durch Modifizieren einer der obigen Polypeptidsequenzen identifiziert werden und der Gehalt der Fusionspolypeptidherstellung in Wirtszellen kann beurteilt werden, zum Beispiel in *E. coli*. Beispiele von Varianten umfassen solche bei denen ein kleiner Teil (z.B. 1-30 Aminosäuren, bevorzugt 5-15 Aminosäuren) von dem N- und/oder C-terminalen Ende der nativen Ra12-Polypeptide entfernt worden ist. In einer Ausführungsform umfassen Varianten von nativen Ra12-Polypeptiden mindestens ungefähr 5 Aminosäuren, mindestens ungefähr 10 Aminosäuren, mindestens ungefähr 30 Aminosäuren, mindestens ungefähr 50 Aminosäuren oder mindestens ungefähr 100 Aminosäuren.

[0059] In einer Ausführungsform ist die Ra12-Polypeptidsequenz die in SEQ ID NO:4 gezeigte. In einer anderen Ausführungsform umfasst die Ra12-Polypeptidsequenz einen Teil der SEQ ID NO:4. Beispielsweise kann ein Ra12-Polypeptid, umfassend 30 Aminosäuren (z.B. Aminosäuren 1-30 der SEQ ID NO:4) oder ein Ra12-Polypeptid, umfassend 128 Aminosäuren (z.B. Aminosäuren 1-128 der SEQ ID NO:4) als Fusionspartner verwendet werden. Siehe nachfolgende Beispiele 2 und 3.

[0060] Polypeptidvarianten, entfalten bevorzugt mindestens ungefähr 70%, bevorzugter mindestens ungefähr 80% oder mindestens ungefähr 90% und am meisten bevorzugt mindestens ungefähr 95% Identität, wie oben beschrieben bestimmt, zu den identifizierten Polypeptiden. Optional besteht über einen Bereich, der mindestens ungefähr 20 bis ungefähr 50 Aminosäuren lang ist, oder optional über einen Bereich der ungefähr 75-100 Aminosäuren lang ist, Identität.

[0061] Bevorzugt enthält eine Variante konservative Substitutionen. Eine "konservative Substitution" ist eine, bei der eine Aminosäure mit einer anderen Aminosäure substituiert ist, die ähnliche Eigenschaften aufweist, so dass der Fachmann im Bereich der Peptidchemie erwarten würde, dass die Sekundärstruktur und die hydropathische Natur des Polypeptids im wesentlichen unverändert ist. Aminosäuresubstitutionen können im allgemeinen erzeugt werden auf Basis der Ähnlichkeit in Polarität, Ladung, Löslichkeit, hydrophoben Eigenschaften, hydrophilen Eigenschaften und/oder der amphipathischen Natur der Reste. Zum Beispiel umfassen negativ geladen Aminosäuren Asparinsäure und Glutaminsäure; positiv geladene Aminosäuren umfassen Lysin und Arginin und Aminosäuren mit unveränderten polaren Kopfgruppen, die ähnliche Werte der hydrophilen Eigenschaften aufweisen, umfassen Leucin, Isoleucin und Valin, Glycin und Alanin, Asparagin und Glutamin, und Serin, Threonin, Phenylalanin und Tyrosin. Andere Gruppen von Aminosäuren, die konservative Änderungen darstellen, umfassen: (1) Ala, Pro, Gly, Glu, Asp, Gln, Asn, Ser, Thr; (2) Cys, Ser, Tyr, Thr; (3) Val, Ile, Leu, Met, Ala, Phe; (4) Lys, Arg, His und (5) Phe, Tyr, Trp, His. Eine Variante kann außerdem oder alternativ nicht-konservative Änderungen enthalten. In einer bevorzugten Ausführungsform unterscheiden sich Variantpolypeptide von einer nativen Sequenz durch Substitution, Deletion oder Addition von fünf Aminosäuren oder weniger. Varianten können außerdem (oder alternativ) durch zum Beispiel Deletion oder Addition von Aminosäuren, die minimalen Einfluß auf die Immunogenität, Sekundärstruktur und hydropathische Natur des Polypeptids aufweisen, modifiziert sein.

[0062] Folglich betrifft der Ausdruck "Ra12-Polypeptid" oder "Ra12-Polypeptidsequenz", wie hier verwendet, native Ra12-Polynukleotidsequenzen (z.B. SEQ ID NO:4), Fragmente davon (z.B. SEQ ID NO:17 oder 18) oder jede Variante davon. Funktionsgemäß hat ein Ra12-Polypeptid die Fähigkeit, ein Fusionsprotein zu erzeugen und seine Fähigkeit, ein Fusionsprotein in Wirtszellen zu erzeugen, kann im Verhältnis zu dem nativen Ra12-Polypeptid (z.B. SEQ ID NO:4) verbessert oder unverändert sein oder kann im Verhältnis zu dem Ra12-Polypeptid weniger als 50% und bevorzugt weniger als 20% vermindert sein.

[0063] Wie oben bemerkt, können Fusionspolypeptide an einen Linker oder an andere Sequenzen konjugiert sein, zur Erleichterung der Synthese, Reinigung oder Identifizierung des Polypeptids oder um die Bindung des Polypeptids an einen festen Träger zu verbessern. Zum Beispiel kann eine Peptid-Linkersequenz eingesetzt werden, um ein Ra12-Polypeptid und ein heterologes Polypeptid von Interesse in einem Abstand zu trennen, der ausreicht, dass jedes Polypeptid sich in seine Sekundär- und Tertiärstruktur faltet. Eine solche Peptid-Linkersequenz wird in das Fusionsprotein unter Verwendung von Standardverfahren, die im Stand der Technik bekannt sind, eingeführt. Geeignete Peptid-Linkersequenzen können auf der Basis der folgenden Faktoren ausgewählt werden: (1) ihrer Fähigkeit, eine flexible ausgedehnte Konformation anzunehmen; (2) ihrer Fähigkeit, eine Sekundärstruktur anzunehmen, die mit den funktionalen Epitopen des ersten und zweiten Polypeptids wechselwirken könnte und (3) dem Fehlen von hydrophoben oder geladenen Resten, die mit den funktionalen Epitopen des Polypeptids wechselwirken könnten. In bestimmten Ausführungsformen können die Peptid-Linkersequenzen Gly-, Asn- und Ser-Reste enthalten. Andere fast neutrale Aminosäuren, wie Thr und Ala, können außerdem in der Linkersequenz verwendet werden. Aminosäuresequenzen, die in geeigneter Weise als Linker eingesetzt werden können, umfassen solche, die offenbart sind in Maratea et al., Gene 40:39-46m 1985; Murphy et al., Proc. Natl. Acad. Sci. USA 83:8258-8262, 1986; US-Patent Nr. 4,935,233 und US-Pa-

tent-Nr. 4,751,180. Die Linkersequenz kann im allgemeinen eine Länge von 1 bis ungefähr 50 Aminosäuren betragen. Linkersequenzen sind nicht erforderlich, wenn das erste und zweite Polypeptid nicht wesentliche N-terminale Aminosäurebereiche aufweisen, die verwendet werden können, um die funktionalen Domänen zu trennen und sterische Behinderung zu vermeiden.

[0064] In einer bevorzugten Ausführungsform kann ein Linker eine spezifische Spaltstelle zwischen einem Ra12-Polypeptid und einem heterologen Polypeptid von Interesse bereitstellen. Eine solche Spaltstelle kann ein Ziel für ein proteolytisches Enzym sein, zum Beispiel für Enterokinase, Faktor Xa, Trypsin, Collagenase, Thrombin, Ubiquitin-Hydrolase oder für chemische Spaltagentien, wie zum Beispiel Cyanogenbromid oder Hydroxylamin.

[0065] Ein Fusionspolypeptid kann optional einen Affinitäts-Tag enthalten, der an das Fusionspolypeptid gebunden ist, so dass die Reinigung des rekombinanten Polypeptids vereinfacht sein kann. Zum Beispiel ermöglichen mehrere Histidinreste, die durch den Tag codiert sind, die Verwendung von Metallchelataffinitätschromatographieverfahren zur Reinigung von Fusionspolypeptiden. Andere Beispiele von Affinitätstagmolekülen umfassen Strep-Tag, PinPoint, Maltose-Bindungsprotein, Glutathion-S-Transferase usw. Siehe z.B. Glick und Pasternak (1999) *Molecular Biotechnology Principles and Applications of Recombinant DNA*, 2. Auflage, American Society for Microbiology, Washington, DC.

[0066] Fusionspolypeptide können unter Verwendung einer Vielzahl von bekannten Verfahren hergestellt werden. Rekombinante Fusionspolypeptide, die durch DNA-Sequenzen codiert sind, wie oben beschrieben, können leicht unter Verwendung einer Vielzahl von Expressionsvektoren, die dem Fachmann bekannt sind aus DNA-Sequenzen hergestellt werden. Die Expression kann in jeder geeigneten Wirtszelle erreicht werden, die mit einem Expressionsvektor transformiert oder transfiziert worden ist, der ein DNA-Molekül enthält, das ein rekombinantes Polypeptid codiert. Geeignete Wirtszellen umfassen Prokaryoten, Hefe und höher eukaryotische Zellen, die oben beschrieben sind. Bevorzugt ist die eingesetzte Wirtszelle *E. coli*. Überstände aus geeigneten Wirts/Vektorsystemen, die rekombinantes Protein oder Peptid in die Kulturmedien abgeben, können zunächst unter Verwendung eines im Handel erhältlichen Filters konzentriert werden. Nach der Konzentrierung kann das Konzentrat auf eine geeignete Reinigungsmatrix, wie eine Affinitätsmatrix oder ein Anionenaustauscherharz geladen werden. Schließlich können ein oder mehrere Umkehrphasen-HPLC-Schritte eingesetzt werden, um ein rekombinantes Polypeptid weiter zu reinigen.

[0067] Teile oder andere Varianten mit weniger als 100 Aminosäuren und im allgemeinen mit weniger als ungefähr 50 Aminosäuren können durch synthetische Mittel unter Verwendung von Verfahren, die dem Fachmann bekannt sind erzeugt werden. Zum Beispiel können solche Polypeptide unter Verwendung jedes im Handel erhältlichen Festphasenverfahrens synthetisiert werden, wie dem Merrifield-Festphasensyntheseverfahren, wobei Aminosäuren sequentiell zu einer wachsenden Aminosäurekette hinzugefügt werden. Siehe Merrifield, J. Am. Chem. Soc. 85:2149-2146, 1963. Die Ausrüstung zur automatisierten Synthese von Polypeptiden ist vom Lieferanten wie Perkin Elmer/Applied Biosystems Division (Foster City, CA) im Handel erhältlich und kann gemäß den Anweisungen des Herstellers bedient werden.

[0068] Im allgemeinen werden Polypeptide (einschließlich Fusionspolypeptide) und Polynukleotide, wie hier beschrieben, isoliert. Ein "isoliertes" Polypeptid oder Polynukleotid ist eines, das aus seiner Ursprungsumgebung entfernt worden ist. Zum Beispiel ist ein in der Natur vorkommendes Protein isoliert wenn es von einigen oder allen coexistierenden Materialien, die in dem System in der Natur vorkommen, getrennt ist. Bevorzugt weisen solche Polypeptide eine Reinheit von mindestens ungefähr 90% rein, bevorzugt von mindestens ungefähr 95% und am meisten bevorzugt von mindestens ungefähr 99% auf. Ein Polypeptid wird als isoliert angesehen, wenn es zum Beispiel in einen Vektor kloniert worden ist, der nicht Teil der natürlichen Umgebung ist.

[0069] Zusätzlich zu der Bereitstellung der Expression von Fusionspolypeptiden von Interesse in stabiler Form und mit hoher Ausbeute, können die rekombinanten Fusionsnukleinsäuren und Fusionspolypeptide der Erfindung in einer Zahl von anderen Verfahren verwendet werden. Zum Beispiel kann die erfindungsgemäße das Fusionspolypeptid codierende Sequenz verwendet werden, um ein Proteinprodukt zu codieren, das als Antigen zum Nachweis von Serumantikörpern verwendet wird. Zum Beispiel zeigt die Gegenwart von Serumantikörpern gegen *M. tuberculosis*-Antigene in einem Individuum an, dass das Individuum mit *M. tuberculosis* infiziert ist. In Standarddiagnostiktests, werden Serumantikörper gegen *M. tuberculosis* durch Verfolgen der Bindung von Serumantikörpern an *M. tuberculosis*-Proteine nachgewiesen. Die erfindungsgemäßen Fusionspolypeptide sind als Proteinquelle zum Verfolgen der Bindung von Serumantikörpern an Fusionsproteine nützlich.

[0070] Alternativ kann das Fusionspolypeptid als Immunogen eingesetzt werden, um Immunantworten zu in-

duzieren und/oder zu verbessern. Solche codierenden Sequenzen können mit einer codierenden Sequenz eines anderen Moleküls, wie eines M. tuberculosis-Antigen, eines Cytokins oder eines Adjuvans ligiert werden. Solche Polynukleotide können in vivo als DNA-Impfstoff verwendet werden (US-Patent Nrn. 5,589,466, 5,679,647 und 5,703,055). Alternativ können gereinigte oder teilweise gereinigte Fusionspolypeptide oder Fragmente als Impfstoffe oder therapeutische Zusammensetzungen eingesetzt werden. Jedes der Vielzahl von im Stand der Technik bekannten Verfahren kann eingesetzt werden, um Impfstoffe oder therapeutische Zusammensetzungen, umfassend die erfindungsgemäßen Fusionspolypeptide, herzustellen.

Proteinreinigung und Präparationen

[0071] Sobald ein rekombinantes Protein exprimiert wird, kann es durch Untersuchungen identifiziert werden, die auf den physikalischen oder funktionalen Eigenschaften des Produktes basieren, einschließlich durch radioaktive Markierung des Produktes gefolgt von der Untersuchung durch Gel-Elektrophorese, durch eine Radioimmununtersuchung, durch ELISA, biologische Untersuchungen usw.

[0072] Sobald das codierte Protein identifiziert ist, kann es durch Standardverfahren isoliert und gereinigt werden, einschließlich durch Chromatographie (z.B. Hochdruckflüssigkeitschromatographie, Ionenaustauscher-Affinitäts- und Größensortierungssäulenchromatographie), Zentrifugieren, unterschiedliche Löslichkeit oder durch jedes andere Standardverfahren zur Proteinreinigung. Siehe im allgemeinen R. Scopes, Protein Purification, Springer-Verlag, N.Y. (1982), Deutscher, Methods in Enzymology, Bd. 182: Guide to Protein Purification, Academic Press, Inc. N.Y. (1990). Die tatsächlichen Bedingungen hängen zum Teil von Faktoren ab, wie der Nettoladung, Hydrophobie, Hydrophilie usw. und sind für den Fachmann offenkundig. Die funktionalen Eigenschaften können unter Verwendung von geeigneten Untersuchungen beurteilt werden.

[0073] Die funktionalen Eigenschaften des Fusionsproteins können beurteilt werden unter Verwendung von jeder geeigneten Untersuchung wie der Antikörperbindung, der Induktion der T-Zellvermehrung, der Stimulierung der Cytokinherstellung wie von IL2, IL-4 und IFN- γ . Zur Anwendung der Erfindung ist es bevorzugt, dass jedes Fusionsprotein mindestens 80% von den anderen Proteinen gereinigt ist. Es ist bevorzugter, dass es mindestens 90% gereinigt ist. Zur in-vivo Verabreichung ist es bevorzugt, dass die Proteine mehr als 95% gereinigt sind.

[0074] Die gereinigten Proteine können außerdem vor der Verwendung weiter aufgearbeitet sein. Zum Beispiel können die Proteine mit einem spezifischen Enzym verdaut sein, um das Ra12-Polypeptid von dem heterologen Polypeptid zu trennen.

[0075] Für den Fachmann ist erkennbar, dass Modifikationen an den rekombinanten Nukleinsäuren und Fusionspolypeptiden gemacht werden können, ohne ihre biologische Aktivität zu vermindern. Solche Modifikationen können gemacht werden, um die Klonierung, Expression oder Einführung des Markierungsmoleküls in ein Fusionspolypeptid zu erleichtern. Solche Modifikationen sind dem Fachmann gut bekannt und umfassen zum Beispiel ein Methionin, das an das Aminoende gefügt wird, um eine Anfangsstelle bereitzustellen oder zusätzliche Aminosäuren (z.B. Poly-His), die an eines der beiden Enden gefügt werden, um Restriktionsstellen oder Terminationscodone oder Reinigungssequenzen an geeigneten Stellen zu erzeugen.

[0076] Die folgenden Beispiele werden angeführt zum Zweck der Erläuterung nicht aber zur Einschränkung.

Beispiele

[0077] Die folgenden Beispiele beschreiben Experimente, die erläutern, dass Ra12-Fusionskonstrukte stabile Expression von Fusionspolypeptiden mit hoher Ausbeute erzeugen. Die folgenden Beispiele erläutern außerdem, dass verschiedene Ra12-Sequenzen als Sequenzen als Fusionspartner verwendet werden können.

Beispiel 1: Die vollständige Ra12-Sequenz (SEQ ID NO:4) als Fusionspartner

A. Konstruktion von Expressionsvektoren

[0078] Codierende Sequenzen von M. tuberculosis Antigenen wurden durch PCR modifiziert, um ihre Fusion und die darauf folgende Expression des Fusionsproteins zu erleichtern. pET17b-Vektor (Novagen) wurde modifiziert, um R12, ein 14 kDa C-terminales Fragment des Serinproteaseantigens MTB32A von M. tuberculosis einzuschließen. Das 3'-Stoppcodon der Ra12-Sequenz wurde im Raster mit einer EcoRI-Stelle substituiert und das N-terminale Ende wurde biotechnologisch verändert, damit es sechs His-Markierungsreste codiert, unmit-

telbar gefolgt von dem Anfangs-Met, um ein einfaches einschrittiges Reinigungsprotokoll der rekombinanten Ra12-Proteine durch Affinitätschromatographie über eine Ni-NTA-Matrix zu erleichtern.

[0079] Speziell wurde das C-terminale Fragment des MTB32A-Antigens durch Standard-PCR-Verfahren amplifiziert unter Verwendung der Oligonukleotidprimer 5' CAA TTA CAT ATG CAT CAC CAT CAC CAT CAC ACG GCC GCG TCC GAT AAC TTC (SEQ ID NO:13) und die 3'-Oligonukleotidsequenz ist 5'-CTA ATC GAA TCC GGC CGG GGG TCC CTC GGC CAA (SEQ ID NO:14). Das 450 Bp-Produkt wurde mit NdeI und EcoRI verdaut und in den pET17b-Fusionsvektor kloniert, und auf gleiche Weise mit den gleichen Enzymen verdaut. Die Expression des rekombinanten Ra12-Proteins durch Transformation in *E. coli* BL-21 (pLyseE) Wirtszellen (Novagen) und durch Induktion mit IPTG wurde erreicht. Nach der Lyse der *E. coli*-Zellen und dem Zentrifugieren bei 10K Upm wurde rekombinantes Ra12 in der löslichen Überstandsfraction festgestellt. Protein aus dem löslichen Überstand wurde durch Affinitätschromatographie über eine Ni-NTA-Säule gereinigt, und blieb nach der Dialyse in 1 × PBS löslich. Die Menge des erhaltenen gereinigten Proteins lag routinemäßig im Bereich von 60 bis 100 mg pro Liter.

[0080] Um ein Fusionsprotein mit Ra12 zu exprimieren, wurde die DPPD-Sequenz biotechnologisch derart entworfen, das ein Oligonukleotidprimer entworfen wurde, der speziell die reife abgesonderte Form amplifiziert. Das 5'-Oligonukleotid, enthaltend eine Enterokinase-Erkennungsstelle (DDDK) weist die Oligonukleotidsequenz 5'-CAA TTA GAA TTC GAC GAC GAC GAC AAG GAT CCA CCT GAC CCG CAT CAG-3' (SEQ ID NO:15) auf und 3'-Oligonukleotidsequenz ist 5'-CAA TTA GAA TTC TCA GGG AGC GTT GGG CTG CTC (SEQ ID NO:16). Das resultierende amplifizierte PCR-Produkt wurde mit EcoRI verdaut und in die EcoRI-Stelle des pET-Ra12-Vektors subkloniert. Nach der Transformation in den *E. coli*-Wirtstamm (XL1-blue, Stratagene), wurden Klone, die das Insert in richtiger Größe enthalten, sequenziert, um solche zu identifizieren, die mit der Ra12-Fusion im Raster waren. Daraufhin wurde die DNA von Interesse ([Fig. 3](#)) in den BL-21 (pLysE)-Bakterienwirt transformiert und das Fusionsprotein durch Induktion der Kultur mit IPTG exprimiert.

B. Expression und Reinigung von Fusionsproteinen

[0081] Das rekombinante Ra12-DPPD-Fusionsprotein mit His-Tag wurde aus 500 ml IPTG-induzierten Kulturchargen von dem löslichen Überstand durch Affinitätschromatographie unter Verwendung der einschrittigen QIAexpressNiNTA-Agarosematrix (QIAGEN, Chatsworth, CA) in Gegenwart von 8 M Harnstoff gereinigt. Kurz gesagt wurden 20 ml einer über Nacht-gesättigten BL21-Kultur, enthaltend das pET-Konstrukt, in 500 ml 2 × YT-Media, enthaltend 50 µg/ml Ampicillin und 34 µg/ml Chloramphenicol gegeben, und bei 37°C unter Schütteln stehengelassen. Die Bakterienkulturen wurden mit 2 mM IPTG bei einer OD 560 von 0,3 induziert und weitere 3 h wachsen gelassen (OD 1,3 bis 1,9). Die Zellen wurden aus 500 ml Kulturchargen durch Zentrifugieren geerntet und in 20 ml Bindungspuffer (0,1 M Natriumphosphat, pH 8,0, 10 mM Tris-HCl, pH 8,0), enthaltend 2 mM PMSF und 20 µg/ml Leupeptin wieder suspendiert. *E. coli* wurde durch Zugabe von 15 mg Lysozym lysiert und 30 min bei 4°C geschüttelt gefolgt von einer Ultraschallbehandlung (4 × 30 s). Lysierte Zellen wurden bei 12 K Upm 30 min gedreht und Harnstoff wurde direkt zu dem Überstand auf eine Endkonzentration von 8 M gegeben.

[0082] Der Überstand wurde an Ni-NTA-Agarosegel (5 ml Harz pro 500 ml Induktionen) durch Schütteln bei R/T 1 h gebunden und die Matrice über eine Säule geleitet. Der Durchfluss wurde 2-mal über die gleiche Säule geleitet, gefolgt wenn drei Waschungen jeweils mit 30 ml Waschpuffer (0,1 M Natriumphosphat und 10 mM Tris-HCl, pH 6,3), außerdem 8 M Harnstoff enthaltend. Gebundenes Protein wurde mit 30 ml 100 mM Imidazol in Waschpuffer eluiert und in 5 ml-Fractionen gesammelt. Fractionen, enthaltend das rekombinante Antigen wurden zusammengeführt, gegen 10 mM Tris-HCl (pH 8,0) dialysiert, nochmals an die Ni-NTA-Matrix gebunden und in 1 × PBS (pH 7,4) oder 10 mM Tris-HCl (pH 7,8) eluiert und dialysiert. Die Ausbeute des gereinigten rekombinanten Fusionsproteins betrug 50 bis 75 mg pro Liter induzierter Bakterienkultur mit einer Reinheit von größer als 95% und stellte eine einzelne Bande dar. Rekombinante Proteine wurden auf Endotoxinverunreinigung unter Verwendung der Limulus-Untersuchung (BioWhittaker) untersucht und enthielten <10 E.U./mg (al ng LPS/mg).

C. Erzeugung von Antiserum

[0083] Das gereinigte Fusionsprotein (100 µg) wurde mit 100 µg Muramyldipeptid gemischt, auf 1 ml mit 1 × PBS gebracht und mit 1 ml IFA (unvollständiges Freund-Adjuvans, Life Technologies) emulgiert. Die Emulsion wurde an mehrere Stellen s.c. in ein weibliches Neuseeland-Kaninchen (R&R Rabbitry, Stanwood, WA) injiziert. Das Kaninchen erhielt in einem Abstand von 6 Wochen zwei darauffolgende Booster (100 µg Antigen in IFA), und ein i.v. Endschuß mit 100 µg rekombinanten Protein wurde wieder nach 6 Wochen gegeben. Eine

Woche nach dem letzten Boost wurde das Kaninchen geopfert und Serum wurde gesammelt und bei -20°C aufbewahrt.

D. Untersuchung durch Immun-Blotting

[0084] *M. tuberculosis* (Stamm H37Rv) Gesamtlysate oder PPD (2,5 µg jeweils) und 25 ng des gereinigten rekombinanten Ra12-DPPD-Fusionsproteins wurden durch Elektrophorese auf 16%igen SDS-PAGE-Gels getrennt und auf Nitrocellulose unter Verwendung einer halbtrockenen Übertragungsvorrichtung (BioRad) übertragen. Blots, in doppelter Ausfertigung, wurden mindestens 1 h mit PBS/0,1% Tween blockiert und mit polyklonalen Sera aus dem gleichen Kaninchen vor der Immunisierung oder nach der Immunisierung mit dem gereinigten rekombinanten Fusionsprotein (verdünnt 1:500 in PBS/0,1% Tween20) sondiert. Die Reaktivität wurde wie zuvor unter Verwendung von $[^{125}\text{I}]$ -Protein A, gefolgt von Autoradiographie untersucht.

E. Ergebnisse

[0085] Verschiedene Expressionssysteme wurden anfänglich zur Expression von DPPD in *E. coli* bewertet. Dies umfaßte das Subklonieren der DPPD-codierenden Sequenz als Nicht-Fusionskonstrukte in a) pET 17b (Novagen) und pQ30 (Qiagen, Santa Clarita, CA) oder 2) als Fusionskonstrukte unter Verwendung von pET32A (Novagen, Madison, WI) oder pGEX-2T (Pharmacia Biotech, Piscataway, NJ). In all diesen Systemen wurde sehr wenig wenn überhaupt DPPD exprimiert und gereinigt.

[0086] Im Gegensatz dazu, wenn die DPPD-codierende Sequenz 3' an die Ra12-Sequenz in einem Expressionsvektor eingeführt wurde und in *E. coli* transformiert wurde, wurde eine große Menge des Ra12-DPPD-Fusionsproteins hergestellt. Die Nukleotidsequenz (SEQ ID NO:5) und die Aminosäuresequenz (SEQ ID NO:6) von Ra12-DPPD sind in [Fig. 3](#) offenbart. Die Immunogenität von DPPD wurde beibehalten, wie ausweislich gezeigt durch die Fähigkeit des Antiserums mit dem gereinigten Protein bei Immun-Blotting-Untersuchungen eine Reaktion einzugehen. Darüber hinaus wurden drei andere Proteine eukaryotischen oder prokaryotischen Ursprungs (siehe [Fig. 4-6](#)) erfolgreich durch die Ra12-Fusionskonstrukte ebenfalls exprimiert. Folglich ist die Ra12-codierende Sequenz als Fusionspartner in einem Expressionskonstrukt nützlich, um die Expression einer heterologen Sequenz zu erleichtern.

Beispiel 2: Kurzes Ra12-Polypeptid (SEQ ID NO:17) als Fusionspartner

[0087] In diesem Beispiel wurde ein Ra12-Polypeptid, umfassend die Aminosäuren 1-30 der SEQ ID NO:4, als Fusionspartner verwendet, um mit dem vollständigen humanen Mammaglobin-Gen verbunden zu werden. Diese kurze Form des Ra12-Polypeptids hat die Aminosäuresequenz, die in SEQ ID NO:17 gezeigt ist und wird hiernach als "Ra12 (kurz)" bezeichnet.

[0088] Wie in [Fig. 9](#) gezeigt, wird das 3'-Ende der Ra12 (kurzen)-Sequenz an das vollständige humane Mammaglobin-Gen fusioniert. Speziell wurde das humane Mammaglobin-Gen durch Standard-PCR-Verfahren unter Verwendung der folgenden Oligonukleotidprimer amplifiziert: der 5'-Primer, HindIII-Stelle: 5'-gcgaagcttATGAAGTTGCTGATGGTCCTCATGC-3' (SEQ ID NO:19), der 3'-Primer, XhoI-Stelle: 5'-cggctcgagTAAAAATAATCACAAAGACTGCTGTC-3' (SEQ ID NO:20). Die 5'-HindIII- und 3'-XhoI-Stellen wurden hinzugefügt, um das Subklonieren in einen Vektor zu unterstützen. Das N-terminale Ende des Fusionskonstruktes wurde biotechnologisch hergestellt, so dass sechs His-Tags sofort gefolgt von dem Met codiert werden, um Reinigungsprotokolle zu erleichtern. Die Expression des Fusionsproteins wurde durch Transformation in *E. coli* unter Verwendung von Verfahren erreicht, die ähnlich zu dem in Beispiel 1 beschriebenen sind. Im Vergleich mit einem Konstrukt ohne eine Ra12 (kurze)-Sequenz erhöhte das Fusionskonstrukt mit einer Ra12 (kurzen)-Sequenz die Expression des Fusions-Ra12 (kurzen)-Mammaglobinproteins wesentlich.

Beispiel 3: Längeres Ra12-Polypeptid (SEQ ID NO:18) als Fusionspartner

[0089] In diesem Beispiel wurde ein Ra12-Polypeptid, umfassend die Aminosäuren 1-128 der SEQ ID NO:4 als Fusionspartner verwendet, um mit dem vollständigen humanen Mammaglobin-Gen verbunden zu werden. Diese lange Form des Ra12-Polypeptids weist die in SEQ ID NO:18 gezeigte Sequenz auf und wird hier als "Ra12 (lang)" bezeichnet. Die Klonierungs- und Expressionsverfahren waren ähnlich zu denen, die in Beispiel 2 verwendet wurden. Im Vergleich mit einem Konstrukt ohne eine Ra12 (lange)-Sequenz, erhöhte das Fusionskonstrukt mit einer Ra12 (langen)-Sequenz die Expression des Ra12 (lang)-Mammaglobin-Fusionsproteins wesentlich.

[0090] Die Erfindung ist nicht durch die beispielhaft dargelegten Ausführungsformen beschränkt, die zur Erläuterung von erfindungsgemäßen Aspekten dienen und andere Klone, Nukleotid- oder Aminosäuresequenzen, die funktional äquivalent sind, können ebenso verwendet werden. In der Tat sind für den Fachmann verschiedene Modifikationen der Erfindung zusätzlich zu den hier beschriebenen aufgrund der vorgenannten Beschreibung und der beigefügten Abbildungen offensichtlich. Es ist ebenfalls nachvollziehbar, dass alle Basenpaargrößen, die für Nukleotide angegeben sind, ungefähr sind und zum Zweck der Beschreibung verwendet wurden.

SEQUENZLISTE

<110> Corixa Corporation

<120> Verfahren unter Verwendung einer Mycobacterium tuberculosis-codierenden Sequenz, um die stabile Expression heterologer Proteine mit hoher Ausbeute zu erleichtern

<130> AHB/FP6050272

<140> EP 00970619.3

<141> 06.10.2000

<150> PCT/US 00/27652

<151> 06.10.2000

<150> US 60/158,585

<151> 07.10.1999

<160> 22

<170> Patentin Ver. 2.1

<210> 1

<211> 1872

<212> DNA

<213> Mycobacterium tuberculosis

<220>

<223> 32 KD Serinprotease MTB32A

<220>

<221> CDS

<222> (89) .. (1156)

<223> MTB32A

<220>

<221> sig-Peptid

<222> (89) .. (184)

<223> N-terminale hydrophobe Sekretionssignalsequenz

<220>

<221> mat_Peptid

<222> (185)..(1153)

<400> 1

```

gactacgttg.gtgtagaaaa atcctgccgc ccggaccctt aaggctggga caatttctga 60
tagctacccc gacacaggag gttacggg atg agc aat tcg cgc cgc cgc tca 112
                               Met Ser Asn Ser Arg Arg Arg Ser
                               -30                               -25

ctc agg tgg tca tgg ttg ctg agc gtg ctg gct gcc gtc ggg ctg ggc 160
Leu Arg Trp Ser Trp Leu Leu Ser Val Leu Ala Ala Val Gly Leu Gly
                               -20                               -15                               -10

ctg gcc acg gcg ccg gcc cag gcg gcc ccg ccg gcc ttg tcg cag gac 208
Leu Ala Thr Ala Pro Ala Gln Ala Ala Pro Pro Ala Leu Ser Gln Asp
                               -5                               -1    1                               5

cgg ttc gcc gac ttc ccc gcg ctg ccc ctc gac ccg tcc gcg atg gtc 256
Arg Phe Ala Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala Met Val

```

10	15	20	
gcc caa gtg ggg cca cag gtg gtc aac atc aac acc aaa ctg ggc tac			304
Ala Gln Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr			
25	30	35	40
aac aac gcc gtg ggc gcc ggg acc ggc atc gtc atc gat ccc aac ggt			352
Asn Asn Ala Val Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly			
45	50	55	
gtc gtg ctg acc aac aac cac gtg atc gcg ggc gcc acc gac atc aat			400
Val Val Leu Thr Asn Asn His Val Ile Ala Gly Ala Thr Asp Ile Asn			
60	65	70	
gcg ttc agc gtc ggc tcc ggc caa acc tac ggc gtc gat gtg gtc ggg			448
Ala Phe Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly			
75	80	85	
tat gac cgc acc cag gat gtc gcg gtg ctg cag ctg cgc ggt gcc ggt			496
Tyr Asp Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly			
90	95	100	
ggc ctg ccg tcg gcg gcg atc ggt ggc ggc gtc gcg gtt ggt gag ccc			544
Gly Leu Pro Ser Ala Ala Ile Gly Gly Gly Val Ala Val Gly Glu Pro			
105	110	115	120
gtc gtc gcg atg ggc aac agc ggt ggg cag ggc gga acg ccc cgt gcg			592
Val Val Ala Met Gly Asn Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala			
125	130	135	
gtg cct ggc agg gtg gtc gcg ctc ggc caa acc gtg cag gcg tcg gat			640
Val Pro Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp			
140	145	150	
tcg ctg acc ggt gcc gaa gag aca ttg aac ggg ttg atc cag ttc gat			688
Ser Leu Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp			
155	160	165	
gcc gcg atc cag ccc ggt gat tcg ggc ggg ccc gtc gtc aac ggc cta			736
Ala Ala Ile Gln Pro Gly Asp Ser Gly Gly Pro Val Val Asn Gly Leu			
170	175	180	
gga cag gtg gtc ggt atg aac acg gcc gcg tcc gat aac ttc cag ctg			784
Gly Gln Val Val Gly Met Asn Thr Ala Ala Ser Asp Asn Phe Gln Leu			
185	190	195	200
tcc cag ggt ggg cag gga ttc gcc att ccg atc ggg cag gcg atg gcg			832
Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala			
205	210	215	
atc gcg ggc cag atc cga tcg ggt ggg ggg tca ccc acc gtt cat atc			880
Ile Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile			
220	225	230	
ggg cct acc gcc ttc ctc ggc ttg ggt gtt gtc gac aac aac ggc aac			928
Gly Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn			
235	240	245	
ggc gca cga gtc caa cgc gtg gtc ggg agc gct ccg gcg gca agt ctc			976
Gly Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu			
250	255	260	

ggc atc tcc acc ggc gac gtg atc acc gcg gtc gac ggc gct ccg atc 1024
 Gly Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile
 265 270 275 280

aac tcg gcc acc gcg atg gcg gac gcg ctt aac ggg cat cat ccc ggt 1072
 Asn Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly
 285 290 295

gac gtc atc tcg gtg acc tgg caa acc aag tcg ggc ggc acg cgt aca 1120
 Asp Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr
 300 305 310

ggg aac gtg aca ttg gcc gag gga ccc ccg gcc tga tttcgtcgcg 1166
 Gly Asn Val Thr Leu Ala Glu Gly Pro Pro Ala
 315 320

gataccaccc gccggccggc caattggatt ggcgccagcc gtgattgccg cgtgagcccc 1226

cgagttccgt ctcccgtgcg cgtggcatcg tggagcaat gaacgaggca gaacacagcg 1286

tcgagcacc tcccgtgcag ggcagtcacg tcgaaggcgg tgtggtcgag catccggatg 1346

ccaaggactt cggcagcgcc gccgccctgc ccgccgatcc gacctggttt aagcacgccg 1406

tcttctacga ggtgctggtc cgggcgttct tcgacgccag cgcggacggg tccggcgatc 1466

tgcgtggact catcgatcgc ctcgactacc tgcagtggct tggcatcgac tgcattctgg 1526

tgcgcgcgtt ctacgactcg ccgctgcgcg acggcgggta cgacattcgc gacttctaca 1586

aggtgctgcc cgaattcggc accgtcgacg atttcgtcgc cctggtcgac gccgctcacc 1646

ggcgaggtat ccgcatcatc accgacctgg tgatgaatca cacctcggag tcgcacccct 1706

ggtttcagga gtcccgcgcg gaccagacg gaccgtacgg tgactattac gtgtggagcg 1766

acaccagcga gcgctacacc gacgcccgga tcattctcgt cgacaccgaa ggtcgaact 1826

ggtcattcga tcctgtccgc cgacagttct actggcaccg attctt 1872

<210> 2

<211> 355

<212> PRT

<213> Mycobacterium tuberculosis

<223> 32 KD-Serinprotease MTB32A

<400> 2

```

Met Ser Asn Ser Arg Arg Arg Ser Leu Arg Trp Ser Trp Leu Leu Ser
 1          5          10          15
Val Leu Ala Ala Val Gly Leu Gly Leu Ala Thr Ala Pro Ala Gln Ala
          20          25          30
Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
          35          40          45
Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
          50          55          60
Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
65          70          75          80
Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
          85          90          95

```

```

Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln
          100          105          110
Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
          115          120          125
Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
          130          135          140
Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
145          150          155          160
Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
          165          170          175
Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
          180          185          190
Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser
          195          200          205
Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
          210          215          220
Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe Ala
225          230          235          240
Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser Gly
          245          250          255
Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu
          260          265          270
Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val Val
          275          280          285
Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val Ile
          290          295          300
Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala Asp
305          310          315          320
Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr Trp Gln
          325          330          335
Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly
          340          345          350
Pro Pro Ala
          355

```

```

<210>      3
<211>     396
<212>     DNA

```

<213> Mycobacterium tuberculosis
 <220>
 <223> 14 KD-C-terminales Fragment von MTB32A Ra12
 <220>
 <221> CDS
 <222> (1) .. (396)
 <223> Ra12
 <400> 3

```

acg gcc gcg tcc gat aac ttc cag ctg tcc cag ggt ggg cag gga ttc 48
Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
  1             5             10             15

gcc att ccg atc ggg cag gcg atg gcg atc gcg ggc cag atc cga tcg 96
Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser
          20             25             30

ggg ggg ggg tca ccc acc gtt cat atc ggg cct acc gcc ttc ctc ggc 144
Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly
      35             40             45

ttg ggt gtt gtc gac aac aac ggc aac ggc gca cga gtc caa cgc gtg 192
Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val
      50             55             60

gtc ggg agc gct ccg gcg gca agt ctc ggc atc tcc acc ggc gac gtg 240
Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val
      65             70             75             80

atc acc gcg gtc gac ggc gct ccg atc aac tcg gcc acc gcg atg gcg 288
Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala
          85             90             95

gac gcg ctt aac ggg cat cat ccc ggt gac gtc atc tcg gtg acc tgg 336
Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr Trp
      100             105             110

caa acc aag tcg ggc ggc acg cgt aca ggg aac gtg aca ttg gcc gag 384
Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu
      115             120             125

gga ccc ccg gcc 396
Gly Pro Pro Ala
      130

```

<210> 4
 <211> 132
 <212> PRT
 <213> Mycobacterium tuberculosis
 <223> 14 KD-C-terminales Fragment von MTB32A Ra12
 <400> 4

```

  Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
   1          5          10          15
  Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser
           20          25          30
  Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly
           35          40          45
  Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val
   50          55          60
  Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val
   65          70          75          80
  Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala
           85          90          95
  Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr Trp
           100          105          110
  Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu
           115          120          125
  Gly Pro Pro Ala
           130
  
```

<210> 5
 <211> 702
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <221> CDS
 <222> (4)..(696)
 <223> Ra12-DPPD Fusionspolypeptid
 <220>
 <223> Beschreibung der künstlichen Sequenz: Ra12-DPPD-
 Fusionspolypeptid
 <400> 5

cat atg cat cac cat cac cat cac acg gcc gcg tcc gat aac ttc cag	48
Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln	
1 5 10 15	
ctg tcc cag ggt ggg cag gga ttc gcc att ccg atc ggg cag gcg atg	96
Leu Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met	
20 25 30	
gcg atc gcg ggc cag atc cga tcg ggt ggg ggg tca ccc acc gtt cat	144
Ala Ile Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His	
35 40 45	
atc ggg cct acc gcc ttc ctc ggc ttg ggt gtt gtc gac aac aac ggc	192
Ile Gly Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly	
50 55 60	
aac ggc gca cga gtc caa cgc gtg gtc ggg agc gct ccg gcg gca agt	240
Asn Gly Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser	
65 70 75	
ctc ggc atc tcc acc ggc gac gtg atc acc gcg gtc gac ggc gct ccg	288
Leu Gly Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro	
80 85 90 95	
atc aac tcg gcc acc gcg atg gcg gac gcg ctt aac ggg cat cat ccc	336
Ile Asn Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro	
100 105 110	
ggt gac gtc atc tcg gtg acc tgg caa acc aag tcg ggc ggc acg cgt	384
Gly Asp Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg	
115 120 125	
aca ggg aac gtg aca ttg gcc gag gga ccc ccg gcc gaa ttc gac gac	432
Thr Gly Asn Val Thr Leu Ala Glu Gly Pro Pro Ala Glu Phe Asp Asp	
130 135 140	
gac gac aag gat cca cct gac ccg cat cag ccg gac atg acg aaa ggc	480
Asp Asp Lys Asp Pro Pro Asp Pro His Gln Pro Asp Met Thr Lys Gly	
145 150 155	
tat tgc ccg ggt ggc cga tgg ggt ttt ggc gac ttg gcc gtg tgc gac	528
Tyr Cys Pro Gly Gly Arg Trp Gly Phe Gly Asp Leu Ala Val Cys Asp	
160 165 170 175	
ggc gag aag tac ccc gac ggc tcg ttt tgg cac cag tgg atg caa acg	576
Gly Glu Lys Tyr Pro Asp Gly Ser Phe Trp His Gln Trp Met Gln Thr	
180 185 190	
tgg ttt acc ggc cca cag ttt tac ttc gat tgt gtc agc ggc ggt gag	624
Trp Phe Thr Gly Pro Gln Phe Tyr Phe Asp Cys Val Ser Gly Gly Glu	
195 200 205	
ccc ctc ccc ggc ccg ccg cca ccg ggt ggt tgc ggt ggg gca att ccg	672
Pro Leu Pro Gly Pro Pro Pro Pro Gly Gly Cys Gly Gly Ala Ile Pro	

210	215	220	702
tcc gag cag ccc aac gct ccc tga gaattc			
Ser Glu Gln Pro Asn Ala Pro			
225	230		

<210> 6

<211> 230

<212> PRT

<213> Künstliche Sequenz

<223> Beschreibung der künstlichen Sequenz: Ra12-DPPD-Fusions-
polypeptid

<400> 6

Met	His	His	His	His	His	His	Thr	Ala	Ala	Ser	Asp	Asn	Phe	Gln	Leu
1				5					10					15	
Ser	Gln	Gly	Gly	Gln	Gly	Phe	Ala	Ile	Pro	Ile	Gly	Gln	Ala	Met	Ala
		20						25					30		
Ile	Ala	Gly	Gln	Ile	Arg	Ser	Gly	Gly	Ser	Pro	Thr	Val	His	Ile	
		35					40				45				
Gly	Pro	Thr	Ala	Phe	Leu	Gly	Leu	Gly	Val	Val	Asp	Asn	Asn	Gly	Asn
	50					55				60					
Gly	Ala	Arg	Val	Gln	Arg	Val	Val	Gly	Ser	Ala	Pro	Ala	Ala	Ser	Leu
65					70				75					80	
Gly	Ile	Ser	Thr	Gly	Asp	Val	Ile	Thr	Ala	Val	Asp	Gly	Ala	Pro	Ile
			85					90					95		
Asn	Ser	Ala	Thr	Ala	Met	Ala	Asp	Ala	Leu	Asn	Gly	His	His	Pro	Gly
			100					105				110			
Asp	Val	Ile	Ser	Val	Thr	Trp	Gln	Thr	Lys	Ser	Gly	Gly	Thr	Arg	Thr
	115						120					125			
Gly	Asn	Val	Thr	Leu	Ala	Glu	Gly	Pro	Pro	Ala	Glu	Phe	Asp	Asp	Asp
	130					135				140					
Asp	Lys	Asp	Pro	Pro	Asp	Pro	His	Gln	Pro	Asp	Met	Thr	Lys	Gly	Tyr
145					150					155					160
Cys	Pro	Gly	Gly	Arg	Trp	Gly	Phe	Gly	Asp	Leu	Ala	Val	Cys	Asp	Gly
			165					170					175		
Glu	Lys	Tyr	Pro	Asp	Gly	Ser	Phe	Trp	His	Gln	Trp	Met	Gln	Thr	Trp
		180						185					190		
Phe	Thr	Gly	Pro	Gln	Phe	Tyr	Phe	Asp	Cys	Val	Ser	Gly	Gly	Glu	Pro
	195						200					205			
Leu	Pro	Gly	Pro	Pro	Pro	Pro	Gly	Gly	Cys	Gly	Gly	Ala	Ile	Pro	Ser
	210					215					220				
Glu	Gln	Pro	Asn	Ala	Pro										
225					230										

<210> 7
<211> 1746
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: Ra12-WT1-Fusion

<220>
<221> CDS
<222> (4) .. (1740)
<223> Ra12-WT1-Fusionspolypeptid

<400> 7

cat atg cat cac cat cac cat cac acg gcc gcg tcc gat aac ttc cag	48
Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln	
1 5 10 15	
ctg tcc cag ggt ggg cag gga ttc gcc att ccg atc ggg cag gcg atg	96
Leu Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met	
20 25 30	
gcg atc gcg ggc cag atc cga tcg ggt ggg ggg tca ccc acc gtt cat	144
Ala Ile Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His	
35 40 45	
atc ggg cct acc gcc ttc ctc ggc ttg ggt gtt gtc gac aac aac ggc	192
Ile Gly Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly	
50 55 60	
aac ggc gca cga gtc caa cgc gtg gtc ggg agc gct ccg gcg gca agt	240
Asn Gly Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser	
65 70 75	
ctc ggc atc tcc acc ggc gac gtg atc acc gcg gtc gac ggc gct ccg	288
Leu Gly Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro	
80 85 90 95	
atc aac tcg gcc acc gcg atg gcg gac gcg ctt aac ggg cat cat ccc	336
Ile Asn Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro	
100 105 110	
ggt gac gtc atc tcg gtg acc tgg caa acc aag tcg ggc ggc acg cgt	384
Gly Asp Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg	
115 120 125	
aca ggg aac gtg aca ttg gcc gag gga ccc ccg gcc gaa ttc ccg ctg	432
Thr Gly Asn Val Thr Leu Ala Glu Gly Pro Pro Ala Glu Phe Pro Leu	
130 135 140	
gtg ccg cgc ggc agc ccg atg ggc tcc gac gtt ccg gac ctg aac gca	480
Val Pro Arg Gly Ser Pro Met Gly Ser Asp Val Arg Asp Leu Asn Ala	
145 150 155	
ctg ctg ccg gca gtt ccg tcc ctg ggt ggt ggt ggt ggt tgc gca ctg	528
Leu Leu Pro Ala Val Pro Ser Leu Gly Gly Gly Gly Gly Cys Ala Leu	
160 165 170 175	
ccg gtt agc ggt gca gca cag tgg gct ccg gtt ctg gac ttc gca ccg	576
Pro Val Ser Gly Ala Ala Gln Trp Ala Pro Val Leu Asp Phe Ala Pro	
180 185 190	
ccg ggt gca tcc gca tac ggt tcc ctg ggt ggt ccg gca ccg ccg ccg	624
Pro Gly Ala Ser Ala Tyr Gly Ser Leu Gly Gly Pro Ala Pro Pro Pro	
195 200 205	
gca ccg ccg ccg ccg ccg ccg ccg ccg ccg cac tcc ttc atc aaa cag	672
Ala Pro Pro Pro Pro Pro Pro Pro Pro Pro His Ser Phe Ile Lys Gln	
210 215 220	
gaa ccg agc tgg ggt ggt gca gaa ccg cac gaa gaa cag tgc ctg agc	720
Glu Pro Ser Trp Gly Gly Ala Glu Pro His Glu Glu Gln Cys Leu Ser	
225 230 235	
gca ttc acc gtt cac ttc tcc ggc cag ttc act ggc aca gcc gga gcc	768

Ala Phe Thr Val His Phe Ser Gly Gln Phe Thr Gly Thr Ala Gly Ala 240 245 250 255	
tgt cgc tac ggg ccc ttc ggt cct cct ccg ccc agc cag gcg tca tcc Cys Arg Tyr Gly Pro Phe Gly Pro Pro Pro Ser Gln Ala Ser Ser 260 265 270	816
ggc cag gcc agg atg ttt cct aac gcg ccc tac ctg ccc agc tgc ctc Gly Gln Ala Arg Met Phe Pro Asn Ala Pro Tyr Leu Pro Ser Cys Leu 275 280 285	864
gag agc cag ccc gct att cgc aat cag ggt tac agc acg gtc acc ttc Glu Ser Gln Pro Ala Ile Arg Asn Gln Gly Tyr Ser Thr Val Thr Phe 290 295 300	912
gac ggg acg ccc agc tac ggt cac acg ccc tcg cac cat gcg gcg cag Asp Gly Thr Pro Ser Tyr Gly His Thr Pro Ser His His Ala Ala Gln 305 310 315	960
ttc ccc aac cac tca ttc aag cat gag gat ccc atg ggc cag cag ggc Phe Pro Asn His Ser Phe Lys His Glu Asp Pro Met Gly Gln Gln Gly 320 325 330 335	1008
tcg ctg ggt gag cag cag tac tcg gtg ccg ccc ccg gtc tat ggc tgc Ser Leu Gly Glu Gln Gln Tyr Ser Val Pro Pro Val Tyr Gly Cys 340 345 350	1056
cac acc ccc acc gac agc tgc acc ggc agc cag gct ttg ctg ctg agg His Thr Pro Thr Asp Ser Cys Thr Gly Ser Gln Ala Leu Leu Leu Arg 355 360 365	1104
acg ccc tac agc agt gac aat tta tac caa atg aca tcc cag ctt gaa Thr Pro Tyr Ser Ser Asp Asn Leu Tyr Gln Met Thr Ser Gln Leu Glu 370 375 380	1152
tgc atg acc tgg aat cag atg aac tta gga gcc acc tta aag ggc cac Cys Met Thr Trp Asn Gln Met Asn Leu Gly Ala Thr Leu Lys Gly His 385 390 395	1200
agc aca ggg tac gag agc gat aac cac aca acg ccc atc ctc tgc gga Ser Thr Gly Tyr Glu Ser Asp Asn His Thr Thr Pro Ile Leu Cys Gly 400 405 410 415	1248
gcc caa tac aga ata cac acg cac ggt gtc ttc aga ggc att cag gat Ala Gln Tyr Arg Ile His Thr His Gly Val Phe Arg Gly Ile Gln Asp 420 425 430	1296
gtg cga cgt gtg cct gga gta gcc ccg act ctt gta cgg tcg gca tct Val Arg Arg Val Pro Gly Val Ala Pro Thr Leu Val Arg Ser Ala Ser 435 440 445	1344
gag acc agt gag aaa cgc ccc ttc atg tgt gct tac tca ggc tgc aat Glu Thr Ser Glu Lys Arg Pro Phe Met Cys Ala Tyr Ser Gly Cys Asn 450 455 460	1392
aag aga tat ttt aag ctg tcc cac tta cag atg cac agc agg aag cac Lys Arg Tyr Phe Lys Leu Ser His Leu Gln Met His Ser Arg Lys His 465 470 475	1440
act ggt gag aaa cca tac cag tgt gac ttc aag gac tgt gaa cga agg Thr Gly Glu Lys Pro Tyr Gln Cys Asp Phe Lys Asp Cys Glu Arg Arg 480 485 490	1488

480	485	490	495	
ttt ttt cgt tca gac cag ctc aaa aga cac caa agg aga cat aca ggt				1536
Phe Phe Arg Ser Asp Gln Leu Lys Arg His Gln Arg Arg His Thr Gly				
	500	505	510	
gtg aaa cca ttc cag tgt aaa act tgt cag cga aag ttc tcc cgg tcc				1584
Val Lys Pro Phe Gln Cys Lys Thr Cys Gln Arg Lys Phe Ser Arg Ser				
	515	520	525	
gac cac ctg aag acc cac acc agg act cat aca ggt gaa aag ccc ttc				1632
Asp His Leu Lys Thr His Thr Arg Thr His Thr Gly Glu Lys Pro Phe				
	530	535	540	
agc tgt cgg tgg cca agt tgt cag aaa aag ttt gcc cgg tca gat gaa				1680
Ser Cys Arg Trp Pro Ser Cys Gln Lys Lys Phe Ala Arg Ser Asp Glu				
	545	550	555	
tta gtc cgc cat cac aac atg cat cag aga aac atg acc aaa ctc cag				1728
Leu Val Arg His His Asn Met His Gln Arg Asn Met Thr Lys Leu Gln				
	560	565	570	575
ctg gcg ctt tga gaattc				1746
Leu Ala Leu				

<210> 8

<211> 578

<212> PRT

<213> Künstliche Sequenz

<223> Beschreibung der künstlichen Sequenz: Ra12-WT1-Fusions-
polypeptid

<400> 8

Met	His	His	His	His	His	His	Thr	Ala	Ala	Ser	Asp	Asn	Phe	Gln	Leu
1				5					10					15	
Ser	Gln	Gly	Gly	Gln	Gly	Phe	Ala	Ile	Pro	Ile	Gly	Gln	Ala	Met	Ala
		20						25					30		
Ile	Ala	Gly	Gln	Ile	Arg	Ser	Gly	Gly	Gly	Ser	Pro	Thr	Val	His	Ile
		35					40					45			
Gly	Pro	Thr	Ala	Phe	Leu	Gly	Leu	Gly	Val	Val	Asp	Asn	Asn	Gly	Asn
	50					55					60				
Gly	Ala	Arg	Val	Gln	Arg	Val	Val	Gly	Ser	Ala	Pro	Ala	Ala	Ser	Leu
	65				70					75					80
Gly	Ile	Ser	Thr	Gly	Asp	Val	Ile	Thr	Ala	Val	Asp	Gly	Ala	Pro	Ile
			85						90					95	
Asn	Ser	Ala	Thr	Ala	Met	Ala	Asp	Ala	Leu	Asn	Gly	His	His	Pro	Gly
		100						105					110		
Asp	Val	Ile	Ser	Val	Thr	Trp	Gln	Thr	Lys	Ser	Gly	Gly	Thr	Arg	Thr
		115					120					125			
Gly	Asn	Val	Thr	Leu	Ala	Glu	Gly	Pro	Pro	Ala	Glu	Phe	Pro	Leu	Val
	130					135					140				
Pro	Arg	Gly	Ser	Pro	Met	Gly	Ser	Asp	Val	Arg	Asp	Leu	Asn	Ala	Leu
145					150					155					160
Leu	Pro	Ala	Val	Pro	Ser	Leu	Gly	Gly	Gly	Gly	Gly	Cys	Ala	Leu	Pro
			165						170					175	
Val	Ser	Gly	Ala	Ala	Gln	Trp	Ala	Pro	Val	Leu	Asp	Phe	Ala	Pro	Pro
		180						185				190			
Gly	Ala	Ser	Ala	Tyr	Gly	Ser	Leu	Gly	Gly	Pro	Ala	Pro	Pro	Pro	Ala
	195						200					205			
Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	His	Ser	Phe	Ile	Lys	Gln	Glu

210	215	220
Pro Ser Trp Gly Gly	Ala Glu Pro His Glu Glu	Gln Cys Leu Ser Ala
225	230	235
Phe Thr Val His Phe	Ser Gly Gln Phe Thr Gly Thr	Ala Gly Ala Cys
245	250	255
Arg Tyr Gly Pro Phe	Gly Pro Pro Pro Ser	Gln Ala Ser Ser Gly
260	265	270
Gln Ala Arg Met Phe	Pro Asn Ala Pro Tyr Leu	Pro Ser Cys Leu Glu
275	280	285
Ser Gln Pro Ala Ile Arg	Asn Gln Gly Tyr Ser Thr	Val Thr Phe Asp
290	295	300
Gly Thr Pro Ser Tyr Gly	His Thr Pro Ser His His	Ala Ala Gln Phe
305	310	315
Pro Asn His Ser Phe Lys	His Glu Asp Pro Met Gly	Gln Gln Gly Ser
325	330	335
Leu Gly Glu Gln Gln Tyr	Ser Val Pro Pro Pro Val	Tyr Gly Cys His
340	345	350
Thr Pro Thr Asp Ser Cys	Thr Gly Ser Gln Ala Leu	Leu Leu Arg Thr
355	360	365
Pro Tyr Ser Ser Asp Asn	Leu Tyr Gln Met Thr Ser	Gln Leu Glu Cys
370	375	380
Met Thr Trp Asn Gln Met	Asn Leu Gly Ala Thr Leu	Lys Gly His Ser
385	390	395
Thr Gly Tyr Glu Ser Asp	Asn His Thr Thr Pro Ile	Leu Cys Gly Ala
405	410	415
Gln Tyr Arg Ile His Thr	His Gly Val Phe Arg Gly	Ile Gln Asp Val
420	425	430
Arg Arg Val Pro Gly Val	Ala Pro Thr Leu Val Arg	Ser Ala Ser Glu
435	440	445
Thr Ser Glu Lys Arg Pro	Phe Met Cys Ala Tyr Ser	Gly Cys Asn Lys
450	455	460
Arg Tyr Phe Lys Leu Ser	His Leu Gln Met His Ser	Arg Lys His Thr
465	470	475
Gly Glu Lys Pro Tyr Gln	Cys Asp Phe Lys Asp Cys	Glu Arg Arg Phe
485	490	495
Phe Arg Ser Asp Gln Leu	Lys Arg His Gln Arg Arg	His Thr Gly Val
500	505	510
Lys Pro Phe Gln Cys Lys	Thr Cys Gln Arg Lys Phe	Ser Arg Ser Asp
515	520	525
His Leu Lys Thr His Thr	Arg Thr His Thr Gly Glu	Lys Pro Phe Ser
530	535	540
Cys Arg Trp Pro Ser Cys	Gln Lys Lys Phe Ala Arg	Ser Asp Glu Leu
545	550	555
Val Arg His His Asn Met	His Gln Arg Asn Met Thr	Lys Leu Gln Leu
565	570	575
Ala Leu		

<210> 9
 <211> 672
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: Ra12-menschliches
 Mammaglobin-Fusion
 <220>
 <221> CDS

<222> (4)..(666)

<223> Ra12-menschliches Mammaglobin-Fusionspolypeptid

<400> 9

```

cat atg cat cac cat cac cat cac acg gcc gcg tcc gat aac ttc cag 48
Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln
1 5 10 15

ctg tcc cag ggt ggg cag gga ttc gcc att ccg atc ggg cag gcg atg 96
Leu Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met
20 25 30

gcg atc gcg ggc cag atc cga tcg ggt ggg ggg tca ccc acc gtt cat 144
Ala Ile Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His
35 40 45

atc ggg cct acc gcc ttc ctc ggc ttg ggt gtt gtc gac aac aac ggc 192
Ile Gly Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly
50 55 60

aac ggc gca cga gtc caa cgc gtg gtc ggg agc gct ccg gcg gca agt 240
Asn Gly Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser
65 70 75

ctc ggc atc tcc acc ggc gac gtg atc acc gcg gtc gac ggc gct ccg 288
Leu Gly Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro
80 85 90 95

atc aac tcg gcc acc gcg atg gcg gac gcg ctt aac ggg cat cat ccc 336
Ile Asn Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro
100 105 110

ggt gac gtc atc tcg gtg acc tgg caa acc aag tcg ggc ggc acg cgt 384
Gly Asp Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg
115 120 125

aca ggg aac gtg aca ttg gcc gag gga ccc ccg gcc gaa ttc atc gag 432
Thr Gly Asn Val Thr Leu Ala Glu Gly Pro Pro Ala Glu Phe Ile Glu
130 135 140

gga agg ggc tct ggc tgc ccc tta ttg gag aat gtg att tcc aag aca 480
Gly Arg Gly Ser Gly Cys Pro Leu Leu Glu Asn Val Ile Ser Lys Thr
145 150 155

atc aat cca caa gtg tct aag act gaa tac aaa gaa ctt ctt caa gag 528
Ile Asn Pro Gln Val Ser Lys Thr Glu Tyr Lys Glu Leu Leu Gln Glu
160 165 170 175

ttc ata gac gac aat gcc act aca aat gcc ata gat gaa ttg aag gaa 576
Phe Ile Asp Asp Asn Ala Thr Thr Asn Ala Ile Asp Glu Leu Lys Glu
180 185 190

tgt ttt ctt aac caa acg gat gaa act ctg agc aat gtt gag gtg ttt 624
Cys Phe Leu Asn Gln Thr Asp Glu Thr Leu Ser Asn Val Glu Val Phe
195 200 205

atg caa tta ata tat gac agc agt ctt tgt gat tta ttt taa gaattc 672
Met Gln Leu Ile Tyr Asp Ser Ser Leu Cys Asp Leu Phe
210 215 220

```

<210> 10
 <211> 220
 <212> PRT
 <213> Künstliche Sequenz
 <223> Beschreibung der künstlichen Sequenz: Ra12-menschliches
 Mammaglobin-Fusionspolypeptid
 <400> 10

Met	His	His	His	His	His	His	Thr	Ala	Ala	Ser	Asp	Asn	Phe	Gln	Leu
1				5					10					15	
Ser	Gln	Gly	Gly	Gln	Gly	Phe	Ala	Ile	Pro	Ile	Gly	Gln	Ala	Met	Ala
		20						25					30		
Ile	Ala	Gly	Gln	Ile	Arg	Ser	Gly	Gly	Gly	Ser	Pro	Thr	Val	His	Ile
		35					40					45			
Gly	Pro	Thr	Ala	Phe	Leu	Gly	Leu	Gly	Val	Val	Asp	Asn	Asn	Gly	Asn
	50					55					60				
Gly	Ala	Arg	Val	Gln	Arg	Val	Val	Gly	Ser	Ala	Pro	Ala	Ala	Ser	Leu
	65				70					75					80
Gly	Ile	Ser	Thr	Gly	Asp	Val	Ile	Thr	Ala	Val	Asp	Gly	Ala	Pro	Ile
				85					90					95	
Asn	Ser	Ala	Thr	Ala	Met	Ala	Asp	Ala	Leu	Asn	Gly	His	His	Pro	Gly
			100					105					110		
Asp	Val	Ile	Ser	Val	Thr	Trp	Gln	Thr	Lys	Ser	Gly	Gly	Thr	Arg	Thr
		115					120					125			
Gly	Asn	Val	Thr	Leu	Ala	Glu	Gly	Pro	Pro	Ala	Glu	Phe	Ile	Glu	Gly
	130					135					140				
Arg	Gly	Ser	Gly	Cys	Pro	Leu	Leu	Glu	Asn	Val	Ile	Ser	Lys	Thr	Ile
	145				150					155					160
Asn	Pro	Gln	Val	Ser	Lys	Thr	Glu	Tyr	Lys	Glu	Leu	Leu	Gln	Glu	Phe
				165					170					175	
Ile	Asp	Asp	Asn	Ala	Thr	Thr	Asn	Ala	Ile	Asp	Glu	Leu	Lys	Glu	Cys
			180					185					190		
Phe	Leu	Asn	Gln	Thr	Asp	Glu	Thr	Leu	Ser	Asn	Val	Glu	Val	Phe	Met
		195					200					205			
Gln	Leu	Ile	Tyr	Asp	Ser	Ser	Leu	Cys	Asp	Leu	Phe				
	210					215					220				

<210> 11
 <211> 2191
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: Ra12-H9-32A-Fusion
 (Ra12-MTB39-MTB32A(N-ter)-Fusion)
 <220>
 <221> CDS
 <222> (1)..(2190)
 <223> Ra12-H9-32A (Ra12-MTH39-MTH32A(N-ter))-Fusionspolypeptid

<400> 11
 atg cat cac cat cac cat cac acg gcc gcg tcc gat aac ttc cag ctg 48
 Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
 1 5 10 15
 tcc cag ggt ggg cag gga ttc gcc att ccg atc ggg cag gcg atg gcg 96
 Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
 20 25 30
 atc gcg ggc cag atc cga tcg ggt ggg ggg tca ccc acc gtt cat atc 144
 Ile Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile

35	40	45	
ggg cct acc gcc ttc ctc ggc ttg ggt gtt gtc gac aac aac ggc aac Gly Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn 50 55 60			192
ggc gca cga gtc caa cgc gtg gtc ggg agc gct ccg gcg gca agt ctc Gly Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu 65 70 75 80			240
ggc atc tcc acc ggc gac gtg atc acc gcg gtc gac ggc gct ccg atc Gly Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile 85 90 95			288
aac tcg gcc acc gcg atg gcg gac gcg ctt aac ggg cat cat ccc ggt Asn Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly 100 105 110			336
gac gtc atc tcg gtg acc tgg caa acc aag tcg ggc ggc acg cgt aca Asp Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr 115 120 125			384
ggg aac gtg aca ttg gcc gag gga ccc ccg gcc gaa ttc atg gtg gat Gly Asn Val Thr Leu Ala Glu Gly Pro Pro Ala Glu Phe Met Val Asp 130 135 140			432
ttc ggg gcg tta cca ccg gag atc aac tcc gcg agg atg tac gcc ggc Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly 145 150 155 160			480
ccg ggt tcg gcc tcg ctg gtg gcc gcg gct cag atg tgg gac agc gtg Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gln Met Trp Asp Ser Val 165 170 175			528
gcg agt gac ctg ttt tcg gcc gcg tcg gcg ttt cag tcg gtg gtc tgg Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp 180 185 190			576
ggt ctg acg gtg ggg tcg tgg ata ggt tcg tcg gcg ggt ctg atg gtg Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val 195 200 205			624
gcg gcg gcc tcg ccg tat gtg gcg tgg atg agc gtc acc gcg ggg cag Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln 210 215 220			672
gcc gag ctg acc gcc gcc cag gtc cgg gtt gct gcg gcg gcc tac gag Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala Ala Tyr Glu 225 230 235 240			720
acg gcg tat ggg ctg acg gtg ccc ccg ccg gtg atc gcc gag aac cgt Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg 245 250 255			768
gct gaa ctg atg att ctg ata gcg acc aac ctc ttg ggg caa aac acc Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr 260 265 270			816
ccg gcg atc gcg gtc aac gag gcc gaa tac ggc gag atg tgg gcc caa Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln 275 280 285			864

gac gcc gcc gcg atg ttt ggc tac gcc gcg gcg acg gcg acg gcg acg	912
Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr	
290 295 300	
gcg acg ttg ctg ccg ttc gag gag gcg ccg gag atg acc agc gcg ggt	960
Ala Thr Leu Leu Pro Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly	
305 310 315 320	
ggg ctc ctc gag cag gcc gcc gcg gtc gag gag gcc tcc gac acc gcc	1008
Gly Leu Leu Glu Gln Ala Ala Ala Val Glu Glu Ala Ser Asp Thr Ala	
325 330 335	
gcg gcg aac cag ttg atg aac aat gtg ccc cag gcg ctg caa cag ctg	1056
Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu Gln Gln Leu	
340 345 350	
gcc cag ccc acg cag ggc acc acg cct tct tcc aag ctg ggt ggc ctg	1104
Ala Gln Pro Thr Gln Gly Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu	
355 360 365	
tgg aag acg gtc tgc ccg cat cgg tgc ccg atc agc aac atg gtg tgc	1152
Trp Lys Thr Val Ser Pro His Arg Ser Pro Ile Ser Asn Met Val Ser	
370 375 380	
atg gcc aac aac cac atg tgc atg acc aac tgc ggt gtg tgc atg acc	1200
Met Ala Asn Asn His Met Ser Met Thr Asn Ser Gly Val Ser Met Thr	
385 390 395 400	
aac acc ttg agc tgc atg ttg aag ggc ttt gct ccg gcg gcg gcc gcc	1248
Asn Thr Leu Ser Ser Met Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala	
405 410 415	
cag gcc gtg caa acc gcg gcg caa aac ggg gtc ccg gcg atg agc tgc	1296
Gln Ala Val Gln Thr Ala Ala Gln Asn Gly Val Arg Ala Met Ser Ser	
420 425 430	
ctg ggc agc tgc ctg ggt tct tgc ggt ctg ggc ggt ggg gtg gcc gcc	1344
Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu Gly Gly Gly Val Ala Ala	
435 440 445	
aac ttg ggt cgg gcg gcc tgc gtc ggt tgc ttg tgc gtg ccg cag gcc	1392
Asn Leu Gly Arg Ala Ala Ser Val Gly Ser Leu Ser Val Pro Gln Ala	
450 455 460	
tgg gcc gcg gcc aac cag gca gtc acc ccg gcg gcg cgg gcg ctg ccg	1440
Trp Ala Ala Ala Asn Gln Ala Val Thr Pro Ala Ala Arg Ala Leu Pro	
465 470 475 480	
ctg acc agc ctg acc agc gcc gcg gaa aga ggg ccc ggg cag atg ctg	1488
Leu Thr Ser Leu Thr Ser Ala Ala Glu Arg Gly Pro Gly Gln Met Leu	
485 490 495	
ggc ggg ctg ccg gtg ggg cag atg ggc gcc agg gcc ggt ggt ggg ctc	1536
Gly Gly Leu Pro Val Gly Gln Met Gly Ala Arg Ala Gly Gly Gly Leu	
500 505 510	
agt ggt gtg ctg cgt gtt ccg ccg cga ccc tat gtg atg ccg cat tct	1584
Ser Gly Val Leu Arg Val Pro Pro Arg Pro Tyr Val Met Pro His Ser	
515 520 525	

```

ccg gca gcc ggc gat atc gcc ccg ccg gcc ttg tcg cag gac cgg ttc 1632
Pro Ala Ala Gly Asp Ile Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe
530 535 540

gcc gac ttc ccc gcg ctg ccc ctc gac ccg tcc gcg atg gtc gcc caa 1680
Ala Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala Met Val Ala Gln
545 550 555 560

gtg ggg cca cag gtg gtc aac atc aac acc aaa ctg ggc tac aac aac 1728
Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn
565 570 575

gcc gtg ggc gcc ggg acc ggc atc gtc atc gat ccc aac ggt gtc gtg 1776
Ala Val Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val
580 585 590

ctg acc aac aac cac gtg atc gcg ggc gcc acc gac atc aat gcg ttc 1824
Leu Thr Asn Asn His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe
595 600 605

agc gtc ggc tcc ggc caa acc tac ggc gtc gat gtg gtc ggg tat gac 1872
Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr Asp
610 615 620

cgc acc cag gat gtc gcg gtg ctg cag ctg cgc ggt gcc ggt ggc ctg 1920
Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu
625 630 635 640

ccg tcg gcg gcg atc ggt ggc ggc gtc gcg gtt ggt gag ccc gtc gtc 1968
Pro Ser Ala Ala Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val Val
645 650 655

gcg atg ggc aac agc ggt ggg cag ggc gga acg ccc cgt gcg gtg cct 2016
Ala Met Gly Asn Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val Pro
660 665 670

ggc agg gtg gtc gcg ctc ggc caa acc gtg cag gcg tcg gat tcg ctg 2064
Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu
675 680 685

acc ggt gcc gaa gag aca ttg aac ggg ttg atc cag ttc gat gcc gcg 2112
Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala
690 695 700

atc cag ccc ggt gat tcg ggc ggg ccc gtc gtc aac ggc cta gga cag 2160
Ile Gln Pro Gly Asp Ser Gly Gly Pro Val Val Val Asn Gly Leu Gly Gln
705 710 715 720

gtg gtc ggt atg aac acg gcc gcg tcc tag g 2191
Val Val Gly Met Asn Thr Ala Ala Ser
725 730

```

```

<210>      12
<211>      729
<212>      PRT
<213>      Künstliche Sequenz
<223>      Beschreibung der künstlichen Sequenz: Ra12-H9-32A-Fusions
polypeptid (Ra12-MTB39-MTB32A(N-ter)-Fusionspolypeptid)
<400>      12

```

Met	His	His	His	His	His	His	Thr	Ala	Ala	Ser	Asp	Asn	Phe	Gln	Leu
1				5				10						15	
Ser	Gln	Gly	Gly	Gln	Gly	Phe	Ala	Ile	Pro	Ile	Gly	Gln	Ala	Met	Ala
		20						25					30		
Ile	Ala	Gly	Gln	Ile	Arg	Ser	Gly	Gly	Gly	Ser	Pro	Thr	Val	His	Ile
		35					40					45			
Gly	Pro	Thr	Ala	Phe	Leu	Gly	Leu	Gly	Val	Val	Asp	Asn	Asn	Gly	Asn
	50					55					60				
Gly	Ala	Arg	Val	Gln	Arg	Val	Val	Gly	Ser	Ala	Pro	Ala	Ala	Ser	Leu
65					70					75					80
Gly	Ile	Ser	Thr	Gly	Asp	Val	Ile	Thr	Ala	Val	Asp	Gly	Ala	Pro	Ile
				85				90						95	
Asn	Ser	Ala	Thr	Ala	Met	Ala	Asp	Ala	Leu	Asn	Gly	His	His	Pro	Gly
			100					105					110		
Asp	Val	Ile	Ser	Val	Thr	Trp	Gln	Thr	Lys	Ser	Gly	Gly	Thr	Arg	Thr
		115					120					125			
Gly	Asn	Val	Thr	Leu	Ala	Glu	Gly	Pro	Pro	Ala	Glu	Phe	Met	Val	Asp
	130					135					140				
Phe	Gly	Ala	Leu	Pro	Pro	Glu	Ile	Asn	Ser	Ala	Arg	Met	Tyr	Ala	Gly
145					150					155					160
Pro	Gly	Ser	Ala	Ser	Leu	Val	Ala	Ala	Ala	Gln	Met	Trp	Asp	Ser	Val
				165				170						175	
Ala	Ser	Asp	Leu	Phe	Ser	Ala	Ala	Ser	Ala	Phe	Gln	Ser	Val	Val	Trp
			180					185					190		
Gly	Leu	Thr	Val	Gly	Ser	Trp	Ile	Gly	Ser	Ser	Ala	Gly	Leu	Met	Val
		195					200					205			
Ala	Ala	Ala	Ser	Pro	Tyr	Val	Ala	Trp	Met	Ser	Val	Thr	Ala	Gly	Gln
	210					215					220				
Ala	Glu	Leu	Thr	Ala	Ala	Gln	Val	Arg	Val	Ala	Ala	Ala	Ala	Tyr	Glu
225					230					235					240
Thr	Ala	Tyr	Gly	Leu	Thr	Val	Pro	Pro	Pro	Val	Ile	Ala	Glu	Asn	Arg
				245				250						255	
Ala	Glu	Leu	Met	Ile	Leu	Ile	Ala	Thr	Asn	Leu	Leu	Gly	Gln	Asn	Thr
			260					265					270		
Pro	Ala	Ile	Ala	Val	Asn	Glu	Ala	Glu	Tyr	Gly	Glu	Met	Trp	Ala	Gln
		275					280					285			
Asp	Ala	Ala	Ala	Met	Phe	Gly	Tyr	Ala	Ala	Ala	Thr	Ala	Thr	Ala	Thr
	290					295					300				
Ala	Thr	Leu	Leu	Pro	Phe	Glu	Glu	Ala	Pro	Glu	Met	Thr	Ser	Ala	Gly
305					310					315					320
Gly	Leu	Leu	Glu	Gln	Ala	Ala	Ala	Val	Glu	Glu	Ala	Ser	Asp	Thr	Ala
				325				330						335	
Ala	Ala	Asn	Gln	Leu	Met	Asn	Asn	Val	Pro	Gln	Ala	Leu	Gln	Gln	Leu
			340					345					350		
Ala	Gln	Pro	Thr	Gln	Gly	Thr	Thr	Pro	Ser	Ser	Lys	Leu	Gly	Gly	Leu
		355					360					365			
Trp	Lys	Thr	Val	Ser	Pro	His	Arg	Ser	Pro	Ile	Ser	Asn	Met	Val	Ser
	370					375						380			
Met	Ala	Asn	Asn	His	Met	Ser	Met	Thr	Asn	Ser	Gly	Val	Ser	Met	Thr
385					390					395					400
Asn	Thr	Leu	Ser	Ser	Met	Leu	Lys	Gly	Phe	Ala	Pro	Ala	Ala	Ala	Ala
				405				410					415		
Gln	Ala	Val	Gln	Thr	Ala	Ala	Gln	Asn	Gly	Val	Arg	Ala	Met	Ser	Ser
			420					425				430			
Leu	Gly	Ser	Ser	Leu	Gly	Ser	Ser	Gly	Leu	Gly	Gly	Gly	Val	Ala	Ala
		435					440					445			
Asn	Leu	Gly	Arg	Ala	Ala	Ser	Val	Gly	Ser	Leu	Ser	Val	Pro	Gln	Ala
	450					455					460				
Trp	Ala	Ala	Ala	Asn	Gln	Ala	Val	Thr	Pro	Ala	Ala	Arg	Ala	Leu	Pro
465					470					475					480
Leu	Thr	Ser	Leu	Thr	Ser	Ala	Ala	Glu	Arg	Gly	Pro	Gly	Gln	Met	Leu

[illegible]

```

<210>      13
<211>      51
<212>      DNA
<213>      Künstliche Sequenz
<220>
<223>      Beschreibung der künstlichen Sequenz: Oligonukleotidprimer für
die
      PCR-Amplifikation des Ra12-C-terminalen Fragments von MTB32A
<400>      13
      caattacata tgcataccca tcaccatcac acgqccgcqct ccgataactt c      51

```

```

<210>          14
<211>          33
<212>          DNA
<213>          Künstliche Sequenz
<220>
<223>          Beschreibung der künstlichen Sequenz: Oligonukleotidprimer für
die          PCR-Amplifikation des Ra12C-terminalen Fragments von MTB32A
<400>          14
          ctaatcgaat tcggccgggg gtccctcggc caa          33
<210>          15

```

<211> 48
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: 5'-Oligonukleotidprimer
 enthaltend die Enterokinase-Erkennungsstelle für die
 PCR-Amplifikation der reifen, abgesonderten DPPD-Form

<400> 15
 caattagaat tcgacgacga cgacaaggat ccacctgacc cgcatcag 48

<210> 16
 <211> 33
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: 3'-Oligonukleotidprimer
 enthaltend die Enterokinase-Erkennungsstelle für die
 PCR-Amplifikation der reifen, abgesonderten DPPD-Form

<400> 16
 caattagaat tctcagggag cggtgggctg ctc 33

<210> 17
 <211> 30
 <212> PRT
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: Ra12 (kurzes)-Polypeptid

<400> 17
 Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
 1 5 10 15
 Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile
 20 25 30

<210> 18
 <211> 128
 <212> PRT
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: Ra12 (langes)-Polypeptid

<400> 18
 Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
 1 5 10 15
 Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Lys Leu
 20 25 30
 Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu Gly Val Val

	35		40		45										
Asp	Asn	Asn	Gly	Asn	Gly	Ala	Arg	Val	Gln	Arg	Val	Val	Gly	Ser	Ala
	50					55					60				
Pro	Ala	Ala	Ser	Leu	Gly	Ile	Ser	Thr	Gly	Asp	Val	Ile	Thr	Ala	Val
65					70					75					80
Asp	Gly	Ala	Pro	Ile	Asn	Ser	Ala	Thr	Ala	Met	Ala	Asp	Ala	Leu	Asn
			85						90					95	
Gly	His	His	Pro	Gly	Asp	Val	Ile	Ser	Val	Thr	Trp	Gln	Thr	Lys	Ser
			100					105					110		
Gly	Gly	Thr	Arg	Thr	Gly	Asn	Val	Thr	Leu	Ala	Glu	Gly	Pro	Pro	Ala
		115					120					125			

<210> 19
 <211> 34
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: 5'-Oligonukleotidprimer, HindIII-Stelle für die PCR-Amplifikation von menschlichem Mammaglobin

<400> 19
 gcgaagctta tgaagttgct gatggtcctc atgc 34

<210> 20
 <211> 36
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: 3'-Oligonukleotidprimer, XhoI-Stelle für die PCR-Amplifikation von menschlichem Mammaglobin

<400> 20
 cggctcgagt taaaataaat cacaaagact gctgtc 36

<210> 21
 <211> 7
 <212> PRT
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: Met-His Tag 6aa

<400> 21
Met His His His His His His
 1 5

<210> 22
 <211> 4
 <212> PRT
 <213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: 5'-Oligonukleotidprimer
enthaltend die Enterokinase-Erkennungsstelle

<400> 22

Asp Asp Asp Lys**1.**

Patentansprüche

1. Verfahren zur stabilen Expression mit hoher Ausbeute eines rekombinanten heterologen Polypeptids, wobei man in dem Verfahren ein für ein Fusionspolypeptid codierendes rekombinantes Nukleinsäuremolekül in einer Wirtszelle exprimiert, wobei das Fusionspolypeptid ein Ra12-Polypeptid und ein heterologes Polypeptid umfaßt, wobei das Ra12-Polypeptid von einer Ra12-Polynukleotidsequenz codiert wird, die unter stringen-ten Bedingungen an eine zu SEQ ID NO:3 komplementäre Sequenz hybridisiert, und wobei die Ra12-Polynu-
kleotidsequenz 5' zu der für das heterologe Polypeptid codierenden heterologen Polynukleotidsequenz lokali-
siert ist.

2. Verfahren nach Anspruch 1, wobei das Fusionspolypeptid ferner ein Affinitäts-Tag umfaßt, das mit dem Fusionspolypeptid verknüpft ist.

3. Verfahren nach Anspruch 1 oder Anspruch 2, wobei das Fusionspolypeptid aus der Wirtszelle auf gerei-
nigt wird.

4. Verfahren nach einem der Ansprüche 1 bis 3, wobei es sich bei der Wirtszelle um E. coli handelt.

5. Verfahren nach einem der Ansprüche 1 bis 4, wobei das rekombinante Nukleinsäuremolekül ferner eine Polynukleotidsequenz umfaßt, die für ein Linkerpeptid zwischen der Ra12-Polynukleotidsequenz und der he-
terologen Polynukleotidsequenz codiert.

6. Verfahren nach Anspruch 5, wobei das Linkerpeptid eine Spaltstelle enthält.

7. Verfahren nach Anspruch 1, wobei die heterologe Nukleinsäuresequenz für ein DPPD-, ein WT1-, ein Mammaglobin- oder ein H9-32A-Polypeptid codiert.

8. Verfahren nach einem der Ansprüche 1 bis 7, wobei die Ra12-Polynukleotidsequenz wenigstens 30 Nu-
kleotide umfaßt.

9. Verfahren nach Anspruch 8, wobei die Ra12-Polynukleotidsequenz wenigstens 60 Nukleotide umfaßt.

10. Verfahren nach Anspruch 8, wobei die Ra12-Polynukleotidsequenz wenigstens 100 Nukleotide umfaßt.

11. Verfahren nach einem der Ansprüche 1 bis 10, wobei die Ra12-Polynukleotidsequenz für ein wie in SEQ ID NO: 17 dargestelltes Ra12-Polypeptid codiert.

12. Verfahren nach einem der Ansprüche 1 bis 10, wobei die Ra12-Polynukleotidsequenz für ein wie in SEQ ID NO: 18 dargestelltes Ra12-Polypeptid codiert.

13. Verfahren nach einem der Ansprüche 1 bis 10, wobei es sich bei der Ra12-Polynukleotidsequenz um die wie in SEQ ID NO: 3 dargestellte Sequenz handelt.

14. Verfahren nach einem der Ansprüche 1 bis 10, wobei die Ra12-Polynukleotidsequenz für ein wie in SEQ ID NO: 4 dargestelltes Ra12-Polypeptid oder eine Variante mit wenigstens 90% Aminosäuresequenzi-
dentität dazu codiert.

15. Verfahren nach Anspruch 14, wobei die Ra12-Polynukleotidsequenz für ein wie in SEQ ID NO: 4 dargestelltes Ra12-Polypeptid oder eine Variante mit wenigstens 95% Aminosäuresequenzidentität dazu codiert.
16. Verfahren nach einem der Ansprüche 1 bis 7, wobei das Ra12-Polypeptid wenigstens 10 Aminosäuren umfaßt.
17. Verfahren nach Anspruch 16, wobei das Ra12-Polypeptid wenigstens 30 Aminosäuren umfaßt.
18. Verfahren nach Anspruch 16, wobei das Ra12-Polypeptid wenigstens 100 Aminosäuren umfaßt.
19. Verfahren nach einem der Ansprüche 1 bis 7, wobei das Ra12-Polypeptid eine wie in SEQ ID NO: 4 dargestellte Sequenz aufweist.
20. Verfahren nach einem der Ansprüche 1 bis 7, wobei das Ra12-Polypeptid eine wie in SEQ ID NO: 17 dargestellte Sequenz aufweist.
21. Verfahren nach einem der Ansprüche 1 bis 7, wobei das Ra12-Polypeptid eine wie in SEQ ID NO: 18 dargestellte Sequenz aufweist.
22. Verfahren nach einem der Ansprüche 1 bis 21, bei dem man ferner das Fusionspolypeptid zwischen dem Ra12-Polypeptid und dem heterologen Polypeptid spaltet.
23. Verwendung einer Ra12-Polynukleotidsequenz, die unter stringenten Bedingungen an eine zu SEQ ID NO: 3 komplementäre Sequenz hybridisiert, in einem Verfahren zur stabilen Expression mit hoher Ausbeute eines rekombinanten heterologen Polypeptids, wobei man in dem Verfahren ein für ein Fusionspolypeptid codierendes rekombinantes Nukleinsäuremolekül in einer Wirtszelle exprimiert, wobei das Fusionspolypeptid ein von der Ra12-Polynukleotidsequenz codiertes Ra12-Polypeptid und ein heterologes Polypeptid umfaßt, wobei die Ra12-Polynukleotidsequenz 5' zu der für das heterologe Polypeptid codierenden heterologen Polynukleotidsequenz lokalisiert ist.

Es folgen 12 Blatt Zeichnungen

FIG. 1A.

```

GCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGGATTCCGCATTCCGATCGGGCAGGGCGATGGCGATCGGGGCCAGATCCGATCGGG 855
-----MTB322A-----
A A S D N F Q L S Q G G Q G F A I P I G Q A M A I A G Q I R S G
TGGGGGTCAACCCACCGTTTCATATCGGGCCTACCGCCTTCTCCGCTTGGGTGTTCTCGACAAACAGGCAACGGCGACGAGTCCAACGCGTGG 950
-----MTB322A-----
G G S P T V H I G P T A F L G L G V V D N N G N F A R V Q R V
TCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGGACGTGATCACCGGCTCGACGGGCTCCGATCAACTCGGCCACCGCGATGGCG 1045
-----MTB322A-----
V G S A P A A S L G I S T G D V I T A V D G A P I N S A T A M A
GACGCGTTAACGGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGGCGGCGGTACAGGGAACGTGACATTTGGCCGA 1140
-----MTB322A-----
D A L N G H H P G D V I S V T W Q T K S G G T R T G N V T L A
GGGACCCCGCGCTGATTTCGTGCGGATACACCCCGCGCGGCCCAATTGGATTGGCGCCAGCCCGTGTGCGCGTCCGAGTTCGG 1235
---MTB322A--->
E G P P A

TCTCCGTGCGGTGGCATCGTGAAGCAATGAACGAGGCGAGACACAGCGTCGAGCACCCCTCCCTGCAGGGCAGTCACTCGAAGCGGTGTG 1330
GTCGAGCATCCGGATGCCAAGGACTTCGGAGCGCCCGCCCTGCCGCGATCCGACCTGGTTTAAGCAGCGCCGCTCTTACGAGGTGCTGGT 1425
CCGGCGTCTTCGACGCGAGCGGACGGTTCGGCGATCTGCGTGGACTATCGATCGCTCGACTACCTGCAGTGGCTTGGCATCGACTGCA 1520
TCTGGTTCGGCCGCTTACGACTCGCCGCTGCGCGACGGCGGTACGACATTCGCGACTTCTACAAGGTGCTGCCCGAATTCGGCACCGTGGAC 1615
GATTTCTGCGCCCTGCTGACCGCGCTCACCGCGGAGGTATCCGCATCATACCGACCTGGTGATGAATCACACCTCGGAGTTCGCACCCCTGGTT 1710
TCAGGAGTCCCGCGGACCCAGACGGACCGTACGTTGAGTACGTTGAGGCGACACCGAGGAGCGGTACACCGACGCGCGGATCATCTTCG 1805
TCGACACCGAAGTCAAGTGGTCAATTCGATCCTGTCCGCGCACAGTTCTACTGGCACCCGATCTT

```

FIG. 1B.

```

ACGGCGCGGTCCGATAAATTCCAGCTGTCCAGGTGGGCAGGGATTCCCATTCGGGCGGCGGATGGCGA
-----Ra12-----
T A A S D N F Q L S Q G G Q G F A I P I G G A M A

TCGGGGCCAGATCCGATCGGGTGGGGGTACCCACCTTCATATCGGGCTACCGCTTCCTCGGCTTGGGTGTTGTCGACAAACAGGCAACGGCGC
-----Ra12-----
I A G Q I R S G G S P T V H I G P T A F L G L G V V D N N G N G A

ACGAGTCCAAACGCGTGTGGGAGCGCTCCGGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTCCGACGGCGCTCCGATCAACTCGGCC
-----Ra12-----
R V Q R V V G S A P A A S L G I S T G D V I T A V D G A P I N S A

ACCGGATGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGGGTACAGGACGTGACAT
-----Ra12-----
T A M A D A L N G H H P G D V I S V T W Q T K S G G T R T G N V T

TGGCCGAGGGACCCCGGCC
-----Ra12-----}
L A E G P F A

```

FIG. 2.

```

CAPATGCATCACCATCACACGGCCGGTCCGATAACTTCCAGCTGTCCAGGGTGGGCAGGGATTGCGCCATTCCGATCGGCAGGCGAT 95
|-----Met (6xHis)-----|-----Ral2-----
M H H H H H H T A A S D N F Q L S Q G G Q G F A I P I G Q A M
GGCGATCGGGCCAGATCCGATCGGGTGGGGGTACCCACCGTTATATCGGGCTACCGCTTCTCGGTTGGGTGTTCGACAACAACG 190
-----Ral2-----
A I A G Q I R S G G S P T V H I G P T A F L G L G V D N N
GCAACGGCAGTCCAGTCCAGCGGTGGTGGAGGCTCCGGGGCAAGTCTCGGCATCTCCACGGGACGTGATCAACCGGGTCGACGGCGCT 285
-----Ral2-----
G N G A R V Q R V V G S A P A A S L G I S T G O V I T A V D G A
CCGATCAACTCGGCCACCGGATGGCGGACGGCTTAACGGGCATCATCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGGGCAC 380
-----Ral2-----
P I N S A T A M A D A L N G H H P G O V I S V T W Q T K S G T
GCGTACAGGGAACGTGACATTGGCCGAGGGACCCCGGCCGAAATTCGACGACGACAGGATCCACCTGACCGCATCAGCGGACATGACGA 475
|-----Ral2-----|EcoR|Enterokinase|-----DPPD-----
R T G N V T L A E G P P A E F D D D D K D P P D P H Q P D M T
AAGGTATGCCGGTGGCCGATGGGGTTTGGCGACTTGGCCGTGTGGACGGGAGAGTACCCCGACGGCTGTTTGGCACCAGTGGATG 570
-----DPPD-----
K G Y C P G G R W G F G D L A V C D G E K Y P D G S F W H Q W M
CAAACGTGGTTTACGGGCCACAGTTTACTTCGATTGTGTACGGCGGTGAGCCCTCCCGGCCCGCCGACCGGGTGTTCGGTGGGCG 665
-----DPPD-----
Q T W F T G P Q F Y F D C V S G G E P L P G P P P G C G G A
AATTCGTCCGAGCAGCCCAACGCTCCCTGAGAATTC 702
|-----DPPD-----|-----→
I P S E Q P N A P

```

FIG. 3.

FIG. 4A.

```

CAAACAGGAACCGAGCTGGGGTGGTGCAGAACCGCACGAAGAACAGTGCCTGAGCGCAATTCACCGTTCACTTCTCCGGCCAGTTCACCTGGCACAG 760
-----WT1-----
K Q E P S W G G A E P H E E Q C L S A F T V H F S G Q F T G T

CCGAGCCTGTCGCTACGGGCCCTTCGGTCTCTCCGCCAGCCAGCGGTTCATCCGGCCAGGCCAGGATGTTTCCTAACGGCCCTACCTGCCCC 855
-----WT1-----
A G A C R Y G P F G P P P S Q A S S G Q A R M F P N A P Y L P

AGCTGCCTGAGAGCCCGCTATTCCGCAATCAGGGTTACAGCACGGTCACCTTCGACGGGACGCCAGCTACGGTCACACGCCCTCGCACCA 950
-----WT1-----
S C L E S Q P A I R N Q G Y S T V T F D G T P S Y G H T P S H H

TGCGGGCAGTTCGCCAACCACTCATTCAGCATGAGGATCCCATGGGCCAGCAGGGCTCGCTGGTGAGCAGCAGTACTCGGTGCCGCCCCCGG 1045
-----WT1-----
A A Q F P N H S F K H E D P M G - Q G S L G E Q Q Y S V P P P

TCTATGGCTGCCACACCCCCACGACAGCTGCACCGCAGCCAGGCTTTCCTGCTAAGGACGCCCTACAGCAGTGACAAATTATACCAATGACA 1140
-----WT1-----
V Y G C H T P T D S C T G S Q A L L L R T P Y S S. D N L Y Q M T

TCCCAGCTTGNATGCATGACCTGGAATCAGATGAACCTTAGAGCCACCTTAAGGGCCACAGCACAGGGTACGAGAGCGGTAACCCACACACGCC 1235
-----WT1-----
S Q L E C M T W N Q M N L G A T L K G H S T G Y E S D N H T T P

CATCCTCTCGGAGGCCCAATACAGATACACACGCGGTGCTTCAGAGGCAATTCAGGATGTGCACCGTGTCCCTGGAGTAGCCCCGACTCTTG 1330
-----WT1-----
I L C G A Q Y R I H T H G V F R G I Q D V A R V P G V A P T L

```

FIG. 4B.

```

TACGGTCGGCATCTGAGACCAAGTGAAGAACGGCCCTTCATGTGTGCTTACTCAGGCTGCAATAAGAGATAATTAAAGCTGTCCCACTTACAGATG 1425
-----WT1-----
V R S A S E T S E K R P F M C A Y S G C N K R Y F K L S H L Q M
CACAGCAGGAAGCACTGGTGAGAAACCATACCAAGTGTGACTTCAAGGACTGTGAACGAAGGTTTTTCGTTCAAAACCAGCTCAAAAGCACCA 1520
-----WT1-----
H S R K H T G E K P Y Q C D F K D C E R R F E R S D Q L K R H Q
AAGGACATACAGGTGTGAACCATTCAGTGTANAACTGTACGCGAAGTTCTCCCGGTCCGACCACCTGAAGACCCACACAGGACTCATA 1615
-----WT1-----
R R H T G V K P F Q C K T C Q R K F S R S D H L K T A T R T H
CAGGTGAAAAGCCCTTCAGCTGTGCGTGGCCCAAGTTGTGAGAAAAGTTTGCCCGGTGAGATGAATTAGTCCGCCATCACAACATACATCAGAGA 1710
-----WT1-----
T G E K P F S C R W P S C Q K K F A R S D E L V R H H N M H Q R
AACATGACCAAACTCCAGCTGGCGGTTTGAGAAATTC
-----WT1-----→
N M T K L Q L A L

```

1746

FIG. 4C.

CATATGCATCACCATCACCGCCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCAGGGATTCCGCCATTCCGATCGGGCAGGCGAT 95
 |-----Met (6xHis)-----||-----Ra12-----
 M H H H H H H T A A S D N F Q L S Q G G Q G F A I P I G Q A M
 GCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCAGTTCATATCGGGCCFACCGCCTTCTCGGCTGGGTGTTGTCGACAACAACG 190
 -----Ra12-----
 A I A G Q I R S G G S P T V H I G P T A F L G L G V V D N N
 GCACGGCGCAGAGTCCAACGGTGGTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGATCACC GGCTCGACGGCGCT 285
 -----Ra12-----
 G N G A R V Q R V V G S A P A A S L G I S T G O V I T A V D G A
 CCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGCGGCAC 380
 -----Ra12-----
 P I N S A T A M A D A L N G H H P G O V I S V T W Q T K S G G T
 GCGTACAGGGAACGTGACATTGGCCGAGGACCCCGCGGATTCATCGAGGGAAGGGCTCTGGCTGCCCTTATTGGAGAATGTGATTCCA 475
 -----Ra12-----
 R T G N V T L A E G P P A E F I E G R G S G C P L L E N V I S
 AGACAATCAATCCCAAGTGTCTAAGACTGAATACAAGAACTCTTCAAGAGTTTCATAGACGACAATGCCACTACAATGCCATAGATGATTG 570
 -----MAMMAGLOBIN-----
 K T I N P Q V S K T E Y K E L L Q E F I D D N A T T N A I D E L
 AAGGAATGTTTTCTTAACCAACGGATGAATCTCTGAGCAATGTTGAGGTGTTTATGCAATTAATATATACAGCAGCTTTGTGATTATTTA 665
 -----MAMMAGLOBIN-----
 K E C F L N Q T D E T L S N V E V F M Q L I Y D S S L C D L F
 AGAATTTC

672

FIG. 5.

ATGCATCACCATCACCACGGCCGCGTCCGATAACTTCAGCTGTCCCGGGTGGCAGGANTTCGCCATTCCGATCGGCGAGCGATGGC 95
 |-----Met (6xHis)-----||-----Rai2-----
 M H H H H H T A A S D N F Q L S Q G G Q G F A I P I G Q A M A

 GATCGGGCCAGATCCGATCGGGTGGGGGTCAACCCGCTCATATCGGGCTACCGCCTTCTCGGCTTGGGTGTGTGCGACAACACGGCA 190
 -----Rai2-----
 I A G Q I R S G G G S P T V H I G P T A F L G L G V V D N N G

 ACGGCGCAGAGTCCACGCGTGGTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTGCGACGGCGCTCCG 285
 -----Rai2-----
 N G A R V Q R V V G S A P A A S L G I S T G O V I T A V D G A P

 ATCAACTCGGCCACCGGATGGCGGCGCTTACGGGCGATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCG 380
 -----Rai2-----
 I N S A T A M A D A L N G H H P G O V I S V T W Q T K S G G T R

 TACAGGAACGTGACATTTGGCCGAGGACCCCGCCGGAATTCAATGTTGATTTCGGGGCGTTACCAACCGGAGATCAACTCCCGGAGGATGTACG 475
 -----Rai2-----|EcoRI|-----MTB39-----
 T G N V T L A E G P P A E F M V D F G A L P P E I N S A R M Y

 CCGGCCGGGTTCCGGCTCGGTGGTGGCGGCTCAGATGTGGACAGCGTGGCGAGTGACCTGTTTCGGCGCGCTCGGCGTTTCAGTCCGTG 570
 -----MTB39-----
 A G P G S A S L V A A A Q M W D S V A S D L F S A A S A F Q S V

 GTCTGGGGTCTGACGTTGGGTCTGATAGGTTCTGTCGGCGGGTCTGATGGTGGCGGCGGCTCGCCGTATGTGGGTGATGAGCGTCACCGC 665
 -----MTB39-----
 V W G L T V G S W I G S S A G L M V A A A S P Y V A W M S V T A

 GGGGCGGCGGAGCTEACCGCGCCAGGTCCGGGTTGCTGGCGGCGCTACGAGACGGCGTATGGGTGACGGTGCCCGCGCGGTGATCGCGG 760
 -----MTB39-----
 G Q A E L T A A Q V R V A A A A Y E T A Y G L T V P P P V I A

FIG. 6A.

AGAACCGTGTGAAGTATCTGTAGTAGCGAACCACTCTTGGGGCAAAACACCCCGCGGATCGCGGTCAACGAGCGCGGAATACGGCGAGATG 855
 -----MTB39-----
 E N R A E L M I L I A T N L L G Q N T P A I A V N E A E Y G E M
 TGGGCCAAGACGCCCGCGGATGTTGGCTACGCCGCGGACGGCGGACGGTTCGCGGTTCGAGGAGCGCGCGGAGATGAC 950
 -----MTB39-----
 W A Q D A A A M F G Y A A A T A T A T L L P F E E A P E M T
 CAGCGCGGTGGGCTCCTCGAGCAGCGCGCGGTTCGAGGAGCGCTCGACACCGCGCGCGGAACCACTGATGAACAATGTCGCCAGGCGC 1045
 -----MTB39-----
 S A G G L L E Q A A A V E E A S D T A A A N Q L M N N V P Q A
 TGCAACAGCTGGCCCGCAGCGCAGGCGCACCGCTTCTTCAAGCTGGGTGGCTGTGGAAGACGGTCTCGCGCATCGGTGCGCGATCAGC 1140
 -----MTB39-----
 L Q Q L A Q P T Q G T T P S S K L G G L W K T V S P H R S P I S
 ARCATGGTGTGATGGCCAAACACCATGTGCGATGACCACTCGGTGTGTCTGATGACCAACACCTTGAGCTCGATGTTGAAGGGCTTTGCTCC 1235
 -----MTB39-----
 N M V S M A N N H M S M T N S G V S M T N T L S S M L K G F A P
 GGCGGGCGCGCGCGCGTGCATAACCGCGCGCAAAACGGGTCCGGCGGATGAGCTCGCTGGGCAGCTCGCTGGTTCCTTCGGGTCTGGCG 1330
 -----MTB39-----
 A A A A Q A V Q T A A Q N G V R A M S S L G S S L G S S G L G
 GTGGGTGGCGCGCAACTTGGGTGGCGGCGCTCGGTTCGTTGCTGCTGCGCAGGCTGGCGCGCGCGCAACCGAGCAGTCAACCCCGCGC 1425
 -----MTB39-----
 G G V A A N L G R A A S V G S L S V P Q A W A A A N Q A V T P A
 GCGCGGCGCTGCCGCTGACCAAGCGCGGGAAGAGGGCCCGGCGAGATGCTGGCGGGCTGCGGTGGGCGAGATGGGCGCCAG 1520
 -----MTB39-----
 A R A L P L T S L T S A A E R G P G Q M L G G L P V G D M G A R

FIG. 6B.

FIG. 6C.

Ra12 (kurzes) Polypeptid (SEQ ID NO:17)
TAASDNFQLSQGGQGFaIPiGQAMaIAGQI

FIG. 7.

Ra12 (langes) Polypeptid (SEQ ID NO:18)
TAASDNFQLSQGGQGFaIPiGQAMaIAGQIKLPTVhIGPTaFLGLGVVDnNGARV
QRVVGSApAASLGISTGdVITaVDGaPiNSaTAMaDALnGHHPGDViSVTWQTKSG
GTRTGNVTLAEGPPA

FIG. 8.



FIG. 9.