



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 603 13 163 T2** 2008.01.03

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 511 768 B1**

(21) Deutsches Aktenzeichen: **603 13 163.8**

(86) PCT-Aktenzeichen: **PCT/EP03/06096**

(96) Europäisches Aktenzeichen: **03 757 056.1**

(87) PCT-Veröffentlichungs-Nr.: **WO 2003/104272**

(86) PCT-Anmeldetag: **06.06.2003**

(87) Veröffentlichungstag
der PCT-Anmeldung: **18.12.2003**

(97) Erstveröffentlichung durch das EPA: **09.03.2005**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **11.04.2007**

(47) Veröffentlichungstag im Patentblatt: **03.01.2008**

(51) Int Cl.⁸: **C07K 14/315** (2006.01)

A61K 39/385 (2006.01)

C12N 5/10 (2006.01)

C12N 15/62 (2006.01)

C12N 15/63 (2006.01)

A61K 38/16 (2006.01)

(30) Unionspriorität:

0213365 **11.06.2002** **GB**

0300914 **15.01.2003** **GB**

(73) Patentinhaber:

GlaxoSmithKline Biologicals S.A., Rixensart, BE;
Glaxo Group Ltd., Greenford, Middlesex, GB

(74) Vertreter:

HOFFMANN & EITLE, 81925 München

(84) Benannte Vertragsstaaten:

AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB,
GR, HU, IE, IT, LI, LU, MC, NL, PT, RO, SE, SI, SK,
TR

(72) Erfinder:

CABEZON SILVA, T.E.V. GlaxoSmithKI, 1330
Rixensart Brussels, BE; ELLIS, Jonathan H.
GlaxoSmithKI, Stevenage, Hertfordshire SG1 2NY,
GB; GERARD, C.M.G. GlaxoSmithKI, 1330
Rixensart Brussels, BE; HAMBLIN, Paul A.
GlaxoSmithKI, Stevenage, Hertfordshire SG1 2NY,
GB; PALMANTIER, R.M. GlaxoSmithKI, 1330
Rixensart Brussels, BE; VINALS Y DE BASSOLS,
C. GlaxoSmithKI, 1330 Rixensart Brussels, BE

(54) Bezeichnung: **IMMUNOGENE ZUSAMMENSETZUNGEN**

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

[0001] Die vorliegende Erfindung betrifft Fusionspartner, die als immunologische Fusionspartner als Expressionsverstärker wirken, und bevorzugt Fusionspartner mit beiden Funktionen. Die Erfindung betrifft auch Fusionsproteine, die sie enthalten, ihre Herstellung, ihre Verwendung in Impfstoffen und ihre Verwendung in Medikamenten. Insbesondere werden Fusionspartner bereitgestellt, die eine sogenannte Cholin-Bindungsdomäne 45enthalten, zum Beispiel Fusionen, die LytA aus *Streptococcus pneumoniae* oder das Pneumokokken-Phage-CP1-Lysozym (CPL1) umfassen, worin die Cholin-Bindungsdomäne modifiziert ist, um ein heterologes T-Helferepitop einzuschließen. Es wird gezeigt, daß solche Fusionspartner die Expressionsstärke des daran gebundenen heterologen Proteins verbessern und auch einen besonderen Nutzen finden, wenn sie an schwach immunogene Proteine oder Peptide fusioniert sind, die ansonsten nützlich als Impfstoffantigene sind. Insbesondere sind solche Fusionspartner nützlich in Konstrukten, die Selbstantigene umfassen, zum Beispiel tumorspezifische oder gewebespezifische Antigene.

Hintergrund der Erfindung

[0002] *Streptococcus pneumoniae* synthetisiert eine N-Acetyl-L-alanin-Amidase, LytA, ein Autolysin, das spezifisch das Peptidoglycangerüst der Zellwand abbaut, was schließlich zur Zelllyse führt. Seine Polypeptidkette hat zwei Domänen. Die N-terminale Domäne ist verantwortlich für die katalytische Aktivität, wohingegen die C-terminale Domäne von LytA verantwortlich für die Affinität zu Cholin und zur Verankerung in der Zellwand ist. Diese C-terminale Domäne ist dafür bekannt, an Cholin und Cholin-Analoga zu binden, und wird auch an tertiäre Amine wie DEAE (Diethylaminoethyl) binden, die herkömmlich in der Chromatographie verwendet werden.

[0003] LytA ist ein Protein mit 318 Aminosäuren, und der C-terminale Teil umfaßt ein Tandem aus sechs unvollständigen Repeats mit 20 oder 21 Aminosäuren und einem kurzen terminalen COOH-Schwanz. Die Repeats befinden sich in den folgenden Positionen:

R1: 177-191

R2: 192-212

R3: 213-234

R4: 235-254

R5: 255-275

R6: 276-298

[0004] Von diesen Repeats wird vorhergesagt, daß sie in einer beta-Turn-Konformation sind. Der C-Terminus ist verantwortlich für das Binden von Cholin. In gleicher Weise ist der C-Terminus von CPL1 verantwortlich für die Bindungsaffinität, und die aromatischen Reste im Repeat tragen zu einer solchen Bindung bei. Diese Proteine wurden als Affinitätsmarker für eine schnelle Reinigung verwendet (Sanchez Puelles, Eur. J. Biochem. 1992, 203, 153-9).

[0005] Andere Proteine mit einer Cholin-Bindungsdomäne wurden ebenfalls in *Streptococcus pneumoniae* untersucht.

[0006] Eines davon, PspA (oder Pneumokokkenoberflächenprotein A), ist ein Virulenzfaktor (J. Yother und Briles (1992), J. Bacteriol. 174(2), S. 601). Dieses Protein ist antigen und immunogen. Es hat eine C-terminale Domäne, die aus 10 Repeats mit 20 Aminosäuren besteht, homolog mit den Repeats von LytA.

[0007] CbpA (oder Cholin-Bindungsprotein A) ist an der Anhaftung des Pneumococcus an menschlichen Zellen beteiligt (Rosenow et al. (1997), Mol. Microbiol. 25(5), S. 819). Es zeigt 10 Repeats mit 20 Aminosäuren in der C-terminalen Domäne, die fast identisch mit denjenigen von PspA sind.

[0008] LytB und LytC haben eine unterschiedliche modulare Organisation gegenüber den oben genannten Proteinen, da sich ihre Cholin-Bindungsdomäne, die aus 15 bzw. 11 Repeats aufgebaut ist, am N-terminalen Ende befindet, nicht am C-terminalen Ende (P. Garcia, Mol. Microbiol. (1999), 31 (4), S. 1275 und P. Garcia et al. (1999), Mol. Microbiol. 33(1), S. 128). Ein Sequenzvergleich zeigt, daß LytB Glucosamidaseaktivität besitzt. LytC zeigt eine Aktivität vom Lysozymtyp in vitro. Zusätzlich wurden drei als PepA, PepB und PepC bezeichnete Gene 1995 kloniert. Obwohl ihre Funktion unbekannt ist, besitzen diese Gene ebenfalls eine variable Anzahl von Repeats, die homolog zu denjenigen von LytA sind.

[0009] In ihrem Infektionszyklus synthetisieren Phagen Mureinhydrolasen, die ihre Passage in das Bakterium

erleichtern. Diese Hydrolasen besitzen eine Cholin-Bindungsdomäne. Die Muramidase CPL1 des Phagen Cp-1 wurde gut untersucht. Sie zeigt 6 Repeats mit 20 Aminosäuren am C-Terminus, die an der spezifischen Erkennung von Cholin beteiligt sind (J.L. Garcia, J. Virol. 61 (8), S. 2573-80; (1987) und E. Garcia, Proc. Natl. Acad. Sci. (1988), S. 914). Ein Vergleich der LytA- und CPL1-Repeats ermöglicht es, eine anfängliche Übereinstimmung dieser Repeats vorzunehmen.

[0010] Die Mureinhydrolasen der Phagen Dp-1 (P. Garcia et al. (1983), J. Gen. Microbiol. 129 (2), S. 489), Cp1-9 (P. Garcia et al. (1989), Biochem. Biophys. Res. Commun. 158(1), S. 251), HB-3 (Romero et al. 1990, J. Bacteriol. 172 (9), S. 5064-5070) und EJ-1 (Diaz (1992), J. Bacteriol. 174 (17), S. 5516) zeigen ebenfalls die Eigenschaften von Cholin-Bindungsdomänen.

[0011] Diese Eigenschaft wird auch vom Lysozym geteilt, das durch CP-1 codiert wird, einen Pneumokokken-Phagen. WO 99/10375 beschreibt unter anderem humane Papillomavirus-Proteine E6 oder E7, die an einen His-Marker und den C-terminalen Teil von LytA (hier C-LytA) gebunden sind, und die Reinigung der Proteine durch differentielle Affinitätschromatographie. WO 99/40188 beschreibt unter anderem Fusionsproteine, die MAGE-Antigene mit einem His-Schwanz und einem C-LytA-Teil am N-Terminus des Moleküls umfassen.

[0012] Es wurde jetzt überraschend festgestellt, daß erfindungsgemäße Fusionspartner bei Fusion an ein heterologes Protein in der Lage waren, die Immunogenität der daran gebundenen heterologen Proteine zu steigern. Es wurde ebenfalls festgestellt, daß die Expressionstärke der daran gebundenen heterologen Proteine gesteigert werden kann. Die vorliegende Erfindung stellt entsprechend in einer bevorzugten Ausführungsform einen verbesserten immunologischen Fusionspartner bereit, der auch als Expressionsverstärker wirken kann.

Zusammenfassung der Erfindung

[0013] Entsprechend umfaßt die vorliegende Erfindung ein Fusionspartnermolekül, das eine Cholin-Bindungsdomäne oder ein Fragment davon oder ein Analogon davon und ein heterologes promiskues T-Helferepitop, bevorzugt ein promiskues MCH-Klasse II T-Epitop, umfaßt. Der Fusionspartner zeigt die Fähigkeit, als immunologischer Fusionspartner oder als Expressionsverstärker und bevorzugt sowohl als immunologischer Partner als auch als Expressionsverstärker zu wirken. Ein promiskues T-Helferepitop ist ein Epitop, das an mehr als ein MCH-Klasse II-Allel bindet, bevorzugt an mehr als 3 MHC-Klasse II-Allele. Insbesondere können solche Epitope eine T-Helfer-Zellreaktion in großen Anzahlen von Individuen hervorrufen, die diverse MHC-Haplotypen exprimieren. Gegebenenfalls kann das Fusionsprotein seine Fähigkeit zur Bindung an Cholin bewahren.

[0014] In einer bevorzugten Ausführungsform stammt die Cholin-Bindungseinheit aus dem C-Terminus von LytA. Bevorzugt umfassen das C-LytA oder die Derivate wenigstens vier Repeats aus einem beliebigen der Repeats R1 bis R6, die in [Fig. 1](#) dargestellt sind (SEQ ID NO:1 bis 6). In einer am meisten bevorzugten Ausführungsform erstreckt sich C-LytA über die Aminosäuren 177-298, was einen Teil des ersten Repeats und die vollständigen fünf anderen enthält.

[0015] In einem weiteren Aspekt der Erfindung wird ein Fusionspartner wie hier definiert bereitgestellt, der ferner ein heterologes Protein umfaßt. Das heterologe Protein kann an den Fusionspartner entweder chemisch konjugiert oder fusioniert sein. Bevorzugt ist das heterologe Protein ein Tumorassoziiertes Antigen oder ein immunogenes Fragment davon.

[0016] In einem weiteren Aspekt der Erfindung wird eine Nukleinsäuresequenz bereitgestellt, die die wie hier definierten Proteine codiert. Es wird ebenfalls ein Expressionsvektor bereitgestellt, der die Nukleinsäure umfaßt, und ein mit der Nukleinsäure und dem Vektor transformierter Wirt.

[0017] In einem weiteren Aspekt der Erfindung wird eine immunogene Zusammensetzung bereitgestellt, die ein Protein oder eine Nukleinsäuresequenz wie hier beschrieben und einen pharmazeutisch akzeptablen Exzipienten, Verdünnungsmittel oder Träger umfaßt. Bevorzugt umfaßt die immunogene Zusammensetzung ferner einen Th-1-induzierenden Hilfsstoff.

[0018] In noch einer weiteren Ausführungsform stellt die Erfindung die immunogene Zusammensetzung oder das Protein oder die Nukleinsäuren zur Verwendung in der Medizin bereit. Insbesondere wird ein Protein oder eine Nukleinsäure der Erfindung in der Herstellung eines Medikaments zum Hervorrufen einer Immunreaktion in einem Patienten oder zur Verwendung in der Behandlung oder Prophylaxe von infektiösen Krankheiten oder Krebskrankheiten bereitgestellt.

[0019] Die Erfindung sieht ferner Verfahren zur Behandlung eines Patienten vor, der an einer infektiösen Krankheit oder einer Krebserkrankung leidet, insbesondere Karzinom der Brust, der Lunge (insbesondere nichtkleinzelliges Lungenkarzinom), kolorektal, Eierstock, Prostata, Magen und anderweitig GI (gastrointestinal), durch Verabreichung einer sicheren und wirksamen Menge einer Zusammensetzung oder Nukleinsäure wie hier beschrieben.

[0020] In noch einer weiteren Ausführungsform stellt die Erfindung ein Verfahren zur Herstellung einer immunogenen Zusammensetzung wie hier beschrieben bereit, durch Vermischen einer Nukleinsäure oder eines Proteins der Erfindung mit einem pharmazeutisch akzeptablen Exzipienten, Verdünnungsmittel oder Träger.

Ausführliche Beschreibung der Erfindung

[0021] Wie hier beschrieben hat in einer Ausführungsform der vorliegenden Erfindung die modifizierte Cholin-Bindungsdomäne (Fusionspartner) eine Fähigkeit, als Expressionsverstärker zu wirken, wobei das resultierende Fusionsprotein in einer höheren Ausbeute in einer Wirtszelle im Vergleich mit dem unfusionierten Protein exprimiert werden wird, bevorzugt mit einer Ausbeute, die mehr als ca. 100% (2-fach höher) oder 150% oder mehr ist, gemäß Messung durch SDS-PAGE, gefolgt von Coomassie-Blue-Anfärbung oder Silber-Anfärbung, gegebenenfalls gefolgt von Gelabstung. Die modifizierte erfindungsgemäße Cholin-Bindungsdomäne besitzt ebenfalls die Fähigkeit, als immunologischer Partner zu wirken, wobei das resultierende Fusionsprotein mit einem heterologen Protein stärker immunogen in einem Wirt im Vergleich mit dem unfusionierten heterologen Protein sein wird.

[0022] In einer anderen Ausführungsform der vorliegenden Erfindung hat die modifizierte Cholin-Bindungsdomäne die Fähigkeit, als immunologischer Fusionspartner zu wirken, was den Erhalt einer gesteigerten Immunreaktion mit dem Fusionsprotein im Vergleich mit dem heterologen Protein allein erlaubt.

[0023] In einer bevorzugten Ausführungsform hat die modifizierte Cholin-Bindungsdomäne eine duale Funktion mit der Fähigkeit, sowohl als immunologischer Fusionspartner als auch als Expressionsverstärker zu wirken.

[0024] In einer bevorzugten Ausführungsform stammt die Cholin-Bindungsdomäne aus dem C-Terminus von LytA. Bevorzugt umfassen das C-LytA oder Derivate wenigstens zwei Repeats, bevorzugt wenigstens vier Repeats. In diesem Zusammenhang bezeichnen C-LytA-Derivate eine Variante von C-LytA gemäß der vorliegenden Erfindung, das heißt Varianten, die sowohl die Fähigkeit, als immunologischer Partner als auch als Expressionsverstärker zu wirken, bewahrt haben. Bevorzugte Varianten schließen zum Beispiel Peptide ein, die eine Aminosäuresequenz mit wenigstens 85% Identität, bevorzugt wenigstens 90% Identität, besonders bevorzugt wenigstens 95% Identität, am meisten bevorzugt wenigstens 97-99% Identität mit jedem der in [Fig. 1](#) dargestellten Repeats R1 bis R6 (SEQ ID NO:1 bis 6) umfassen, oder ein Peptid, das eine Aminosäuresequenz mit wenigstens 15, 20, 30, 40, 50 oder 100 zusammenhängenden Aminosäuren aus der in [Fig. 1](#) dargestellten Aminosäuresequenz (SEQ ID NO:1 bis 8) umfaßt.

[0025] Entsprechend wird in einem Aspekt der vorliegenden Erfindung ein Fusionspartnerprotein bereitgestellt, das eine modifizierte Cholin-Bindungsdomäne und ein heterologes promiskues T-Helferepitop umfaßt, worin die Cholin-Bindungsdomäne aus der Gruppe ausgewählt ist, die folgendes umfaßt:

- a) die C-terminale Domäne von LytA wie in SEQ ID NO:7 dargestellt;
- b) die Sequenz von SEQ ID NO:8;
- c) eine Peptidsequenz, die eine Aminosäuresequenz mit wenigstens 85% Identität, bevorzugt wenigstens 90% Identität, besonders bevorzugt wenigstens 95% Identität, am meisten bevorzugt wenigstens 97-99% Identität mit einem beliebigen aus SEQ ID NO:1 bis 6 umfaßt;
- d) eine Peptidsequenz, die eine Aminosäuresequenz mit wenigstens 15, 20, 30, 40, 50 oder 100 zusammenhängenden Aminosäuren aus der Aminosäuresequenz von SEQ ID NO:7 oder SEQ ID NO:8 umfaßt.

[0026] In einer am meisten bevorzugten Ausführungsform erstreckt sich C-LytA über die Aminosäuren 177-298, was einen Teil des ersten Repeats und die kompletten fünf anderen wie in [Fig. 1](#) dargestellt umfaßt.

[0027] Die zweite Komponente des Fusionspartners, das heterologe T-Zellepitop, ist bevorzugt aus der Gruppe von Epitopen ausgewählt, die an eine Anzahl von Individuen binden werden, die mehr als ein MHC II-Molekül in Menschen exprimieren. Zum Beispiel sind Epitope, die spezifisch erwogen werden, P2- und P30-Epitope aus Tetanustoxoid, Panina-Bordignon, Eur. J. Immunol. 19 (12), 2237 (1989). In einer bevorzugten Ausführungsform ist das heterologe T-Zell-Epitop P2 oder P30 aus Tetanustoxin.

[0028] Das P2-Epitop hat die Sequenz QYIKANSKFIGITE und entspricht den Aminosäuren 830-843 des Tetanustoxins. Das P30-Epitop (Reste 947-967 von Tetanustoxin) hat die Sequenz FNNFTVSFWLRVPKVSAS-HLE. Die FNNFTV-Sequenz kann gegebenenfalls deletiert sein. Andere universelle T-Epitope können aus dem Circumsporozoiten-Protein aus *Plasmodium falciparum* stammen – insbesondere die Region 378-398 mit der Sequenz DIEKKIAKMEKASSVFNVVNS (J. Alexander (1994), *Immunity* 1(9), S. 751-761). Ein anderes Epitop stammt aus dem Masernvirus-Fusionsprotein mit den Resten 288-302 mit der Sequenz LSEIKGVIVHRLEGV (C.D. Partidos, 1990, *J. Gen. Virol.* 71(9), 2099-2105). Noch ein anderes Epitop stammt aus dem Hepatitis B-Virus-Oberflächenantigen, insbesondere die Aminosäuren mit der Sequenz FFLTRILTIPQSLD. Ein anderer Satz von Epitopen stammt aus Diphtherietoxin. Vier dieser Peptide (Aminosäuren 271-290, 321-340, 331-350, 351-370) kartieren innerhalb der T-Domäne von Fragment B des Toxins, und die verbleibenden 2 kartieren in der R-Domäne (411-430, 431-450):

PVFAGANYAAWAVNVAQVI
VHHNTEEIVAQSIALSSLMV
QSIALSSLMVAQAIPLVGEL
VDIGFAAYNFVESIINLFQV
QGESGHDIKITAENTPLPIA
GVLLPTIPGKLDVNKSKTHI

(R. Raju, D. Navaneetham, D. Okita, B. Diethelm-Okita, D. McCormick, B.M. Conti-Fine (1995), *Eur. J. Immunol.* 25: 3207-14).

[0029] Das heterologe T-Epitop ist bevorzugt an C-LytA fusioniert, das wenigstens 4 Repeats enthält, bevorzugt die Repeats 2-5 einschließlich. Ein oder mehrere nachfolgende Repeats können gegebenenfalls an den C-Terminus des T-Epitops fusioniert sein. Alternativ ist das heterologe T-Epitop vorzugsweise zwischen zwei konsekutiven Repeats von C-LytA inseriert, das insgesamt wenigstens 4 Repeats enthält, oder inseriert in eines der Repeats von C-LytA, das insgesamt wenigstens 4 Repeats enthält. Besonders bevorzugt enthält das C-LytA 6 Repeats und das heterologe Epitop ist innerhalb und zu Beginn des sechsten Repeats von C-LytA inseriert.

[0030] Die vorliegende Erfindung stellt ferner in anderen Aspekten Fusionsproteine bereit, die wenigstens ein Polypeptid wie oben beschrieben umfassen, sowie Polynukleotide, die solche Fusionsproteine codieren, typischerweise in der Form von pharmazeutischen Zusammensetzungen, zum Beispiel Impfstoffzusammensetzungen, die einen physiologisch akzeptablen Träger und/oder ein Immunstimulans umfassen. So kann ein Selbstprotein oder anderes schwach immunogenes Protein an entweder das N- oder C-terminale Ende des resultierenden Fusionspartners fusioniert sein. Alternativ kann das Selbstprotein oder schwach immunogene Protein in den Fusionspartner inseriert sein. In einer optionalen Ausführungsform können ein Histidin-Marker oder wenigstens vier, bevorzugt mehr als 6 Histidinreste an das alternative Ende des schwach immunogenen Proteins fusioniert sein. Dies würde es dem Protein erlauben, durch Affinitätschromatographieschritte gereinigt zu werden, da ein Histidinschwanz, der typischerweise wenigstens vier, bevorzugt sechs oder mehr Reste umfaßt, an Metallionen bindet und deshalb für die Metall-immobilisierte Metallionen-Affinitätschromatographie (IMAC) geeignet ist. Typische Konstrukte würden deshalb umfassen:

- schwach immunogenes Protein – C-LytA-Repeats₁₋₄ – P₂-Epitop (inseriert in oder ersetzend C-LytA-Repeat₅) – C-LytA-Repeat₆;
 - C-LytA-Repeats₁₋₄ – P₂-Epitop (inseriert in oder ersetzend C-LytA-Repeat₅) – C-LytA-Repeat₆ – schwach immunogenes Protein;
 - schwach immunogenes Protein – C-LytA-Repeats₂₋₅ – P₂-Epitop (inseriert in C-LytA-Repeat₆);
 - C-LytA₂₋₅ – P₂-Epitop (inseriert in C-LytA-Repeat₆) – schwach immunogenes Protein;
 - schwach immunogenes Protein – C-LytA-Repeats₁₋₅ – P₂-Epitop (inseriert in C-LytA-Repeat₆);
 - C-LytA-Repeats₁₋₅ – P₂-Epitop (inseriert in C-LytA-Repeat₆) – schwach immunogenes Protein;
 - schwach immunogenes Protein – P₂-Epitop (inseriert in C-LytA-Repeat₁) – C-LytA-Repeats₂₋₅;
 - P₂-Epitop (inseriert in C-LytA-Repeat₁) – C-LytA-Repeats₂₋₅ – schwach immunogenes Protein;
 - schwach immunogenes Protein – P₂-Epitop (inseriert in C-LytA-Repeat₁) – C-LytA-Repeats₂₋₆;
 - P₂-Epitop (inseriert in C-LytA-Repeat₁) – C-LytA-Repeats₂₋₆ – schwach immunogenes Protein;
 - schwach immunogenes Protein – C-LytA-Repeat₁ – P₂-Epitop (inseriert in C-LytA-Repeat₂) – C-LytA-Repeats₃₋₆;
 - C-LytA-Repeat₁ – P₂-Epitop (inseriert in C-LytA-Repeat₂) – C-LytA-Repeats₃₋₆ – schwach immunogenes Protein;
- worin "inseriert in" jeden Ort im Repeat bedeutet, zum Beispiel zwischen Rest 1 und 2 oder zwischen 2 und 3 etc.

[0031] Das promiske T-Helferepitop kann in eine Repeatregion inseriert sein, zum C-LytA-Repeats₂₋₅ – C-Ly-

tA-Repeat 6a – P₂-Epitop – C-LytA-Repeat 6b, worin das P₂-Epitop im sechsten Repeat inseriert ist (siehe [Fig. 2](#)).

[0032] In anderen bevorzugten Ausführungsformen kann das C-terminale Ende von CPL1 (C-CPL1) als Alternative zu C-LytA verwendet werden.

[0033] Alternativ kann das P₂-Epitop in den obigen Konstrukten durch andere promiske T-Epitope ersetzt werden, zum Beispiel P30. In einer Ausführungsform der Erfindung sind zwei oder mehr promiske Epitope Teil des Fusionskonstrukts. Es ist jedoch bevorzugt, die Fusionspartner so klein wie möglich zu halten, wodurch die Anzahl von potentiell wechselwirkenden CD8+- und B-Epitopen beschränkt wird. Daher ist der Fusionspartner bevorzugt nicht größer als 100-140 Aminosäuren, bevorzugt nicht größer als 120 Aminosäuren, typischerweise etwa 100 Aminosäuren.

[0034] Das Antigen, an das der Fusionspartner fusioniert ist, kann aus bakteriellen, viralen, Protozoen-, pilzlichen oder Säuger-, einschließlich humanen, Quellen sein.

[0035] Die Fusionspartner der vorliegenden Erfindung sind bevorzugt an ein Selbstantigen wie an ein Tumor-assoziiertes oder gewebespezifisches Antigen wie diejenigen für Prostata-, Brust-, kolorektalen, Lungen-, Bauchspeicheldrüsen-, Eierstock-, Nieren- oder Melanomkrebs fusioniert. Fragmente der Selbst- oder Tumorantigene werden ausdrücklich als an den Fusionspartner der Erfindung fusioniert erwogen. Typischerweise wird das Fragment wenigstens 20, bevorzugt 50, besonders bevorzugt 100 zusammenhängende Aminosäuren der Sequenz voller Länge enthalten. Typischerweise werden solche Fragmente frei von einer oder mehreren Transmembrandomänen sein oder können N-terminale oder C-terminale Deletionen mit ca. 3, 5, 8, 10, 15, 20, 28, 33, 50, 54 Aminosäuren aufweisen. Solche Fragmente werden, wenn sie geeignet präsentiert werden, Immunreaktionen erzeugen können, die das Protein voller Länge erkennen. Besonders illustrative Polypeptide der vorliegenden Erfindung umfassen eine Sequenz von wenigstens 10 zusammenhängenden Aminosäuren, bevorzugt 20, besonders bevorzugt 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180 Aminosäuren eines Tumor-assoziierten oder gewebespezifischen Proteins, das an den Fusionspartner fusioniert ist.

[0036] Die Polypeptide der Erfindung sind immunogen, d.h. sie reagieren detektierbar in einem Immuntest (wie einem ELISA- oder T-Zell-Stimulationstest) mit Antiseren und/oder T-Zellen aus einem Patienten mit Crip-toexprimierendem Krebs. Die Durchmusterung auf immunogene Aktivität kann unter Verwendung von dem Fachmann wohlbekannten Techniken durchgeführt werden. Zum Beispiel können solche Durchmusterungen unter Verwendung von Verfahren wie denjenigen durchgeführt werden, die beschrieben werden in Harlow und Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In einem illustrativen Beispiel kann ein Polypeptid an einem festen Träger immobilisiert und mit Patientenseren in Kontakt gebracht werden, um die Bindung von Antikörpern in den Seren mit dem immobilisierten Polypeptid zu erlauben. Ungebundene Seren können dann entfernt und gebundene Antikörper unter Verwendung von zum Beispiel ¹²⁵I- markiertem Protein A detektiert werden. Wie der Fachmann erkennen würde, sind immunogene Teile von Tumor-assoziiertem oder tumorspezifischem Antigen ebenfalls von der vorliegenden Erfindung umfaßt. Ein "immunogener Teil" wie hier verwendet ist ein Fragment, das selbst immunologisch reaktiv gegenüber den B-zellen und/oder T-Zell-Oberflächenantigen-Rezeptoren ist (d.h. spezifisch daran bindet), die das Polypeptid erkennen. Immunogene Teile können allgemein unter Verwendung wohlbekannter Techniken identifiziert werden, wie zum Beispiel mit denjenigen, die zusammengefaßt werden in Paul, *Fundamental Immunology*, 3. Auflage, 243-247 (Rauen Press, 1993) und darin zitierten Verweisen. Solche Techniken schließen die Durchmusterung von Polypeptiden auf die Fähigkeit zur Reaktion mit antigenspezifischen Antikörpern, Antiseren und/oder T-Zelllinien oder -klonen ein. Wie hier verwendet sind Antiseren und Antikörper "antigenspezifisch", falls sie spezifisch an ein Antigen binden (d.h. mit dem Protein in einem ELISA oder anderen Immuntest reagieren und nicht detektierbar mit unverwandten Proteinen reagieren). Solche Antiseren und Antikörper können wie hier beschrieben und unter Verwendung wohlbekannter Techniken hergestellt werden. In einer bevorzugten Ausführungsform ist ein immunogener Teil eines Polypeptids ein Teil, der mit Antiseren und/oder T-Zellen in einer Stärke reagiert, die nicht substantiell weniger als die Reaktivität eines Polypeptids voller Länge ist (z.B. in einem ELISA und/oder T-Zell-Reaktivitätstest). Bevorzugt ist die Stärke der immunogenen Aktivität des immunogenen Teils wenigstens ca. 50%, bevorzugt wenigstens ca. 70% und am meisten bevorzugt größer als ca. 90% der Immunogenität des Polypeptids voller Länge. In manchen Fällen werden bevorzugte immunogene Teile identifiziert werden, die eine Stärke der immunogenen Aktivität von mehr als derjenigen des entsprechenden Polypeptids voller Länge haben, zum Beispiel mit mehr als ca. 100% oder 150% oder mehr immunogener Aktivität.

[0037] In bestimmten anderen Ausführungsformen können illustrative immunogene Teile Peptide einschlie-

ßen, in denen eine N-terminale Leitsequenz und/oder Transmembrandomäne deletiert wurden. Andere illustrative immunogene Teile werden eine kleine N- und/oder C-terminale Deletion enthalten (z.B. ca. 1-50 Aminosäuren, bevorzugt ca. 1-30 Aminosäuren, besonders bevorzugt ca. 5-15 Aminosäure), relativ zum reifen Protein.

[0038] Exemplarische Antigene oder daraus stammende Fragmente schließen MAGE 1, MAGE 3 und MAGE 4 oder andere MAGE-Antigene wie die in WO 99/40188 offenbarten, PRAME (WO 96/10577), BAGE, RAGE, LAGE 1 (WO 98/32855), LAGE 2 (auch bekannt als NY-ESO-1, WO 98/14464), XAGE (Liu et al., Cancer Res., 2000, 60:4752-4755; WO 02/18584), SAGE und HAGE (WO 99/53061) oder GAGE (Robbins und Kawakami, 1996, Current Opinions in Immunology 8, S. 628-636; Van den Eynde et al., International Journal of Clinical & Laboratory Research (eingereicht 1997); Correale et al. (1997), Journal of the National Cancer Institute 89, S. 293) ein. Tatsächlich werden diese Antigene in einer großen Reihe von Tumortypen wie Melanom, Lungenkarzinom, Sarkom und Blasenkarzinom exprimiert.

[0039] In einer bevorzugten Ausführungsform werden Prostataantigene verwendet, wie Prostata-spezifisches Antigen (PSA), PAP, PSCA, (PNAS 95(4) 1735-1740, 1998), PSMA oder das als Prostase bekannte Antigen.

[0040] In einer besonders bevorzugten Ausführungsform ist das Prostataantigen P501S oder ein Fragment davon. P501S, auch als Prostein bezeichnet (Xu et al., Cancer Res. 61, 2001, 1563-1568), ist als SEQ ID NO:113 aus WO 98/37814 bekannt und ist ein Protein mit 553 Aminosäuren. Immunogene Fragmente und Teile davon, die wenigstens 20, bevorzugt 50, besonders bevorzugt 100 zusammenhängende Aminosäuren umfassen, wie sie in der oben bezeichneten Patentanmeldung offenbart werden, werden spezifisch von der vorliegenden Erfindung erwogen. Bevorzugte Fragmente werden in WO 98/50567 (PS108-Antigen) und als Prostatakrebs-assoziiertes Protein (SEQ ID NO:9 aus WO 99/67384) offenbart. Andere bevorzugte Fragmente sind die Aminosäuren 51-553, 34-553 oder 55-553 des P501S-Proteins voller Länge. Insbesondere werden die Konstrukte 1, 2 und 3 (siehe [Fig. 2](#), SEQ ID NOs:27-32) ausdrücklich erwogen und können in Hefesystemen exprimiert werden, zum Beispiel können DNA-Sequenzen, die solche Polypeptide codieren, in einem Hefesystem exprimiert werden.

[0041] Prostase ist eine Prostata-spezifische Serinprotease (trypsinartig), 254 Aminosäuren lang, mit einer konservierten Serinprotease-katalytischen Triade H-D-S und einer Amino-terminalen Prä-Propeptidsequenz, was eine potentielle sekretorische Funktion anzeigt (P. Nelson, Lu Gan, C. Ferguson, P. Moss, R. Linas, L. Hood & K. Wand, "Molecular cloning and characterisation of prostase, an androgen-regulated serin protease with prostate restricted expression", Proc. Natl. Acad. Sci. USA (1999) 96, 3114-3119). Ein mutmaßlicher Glycosylierungsort wurde beschrieben. Die vorhergesagte Struktur ist sehr ähnlich anderen bekannten Serinproteasen, was zeigt, daß das reife Polypeptid zu einer einzelnen Domäne faltet. Das reife Protein ist 224 Aminosäuren lang, wobei gezeigt wurde, daß ein A2-Epitop natürlich gereift wird. Die Prostase-Nukleotidsequenz und abgeleitete Polypeptidsequenzen und Homologe werden offenbart in Ferguson et al. (Proc. Natl. Acad. Sci. USA 1999, 96, 3114-3119) und in WO 98/12302 (und auch im entsprechenden erteilten Patent US 5,955,306), WO 98/20117 (und auch in den entsprechenden erteilten Patenten US 5,840,871 und US 5,786,148) (Prostata-spezifisches Kallikrein) und WO 00/04149 (P703P).

[0042] Andere Prostata-spezifische Antigene sind bekannt aus WO 98/37418 und WO/004149. Ein weiteres ist STEAP (PNAS 96, 14523, 14528 7-12, 1999).

[0043] Andere, im Zusammenhang der vorliegenden Erfindung nützliche Tumorassoziierte Antigene schließen ein: Plu-1 (J. Biol. Chem. 274 (22), 15633-15645, 1999), HASH-1, HASH-2 (M. Alders et al., Hum. Mol. Genet. 1997, 6, 859-867), Cripto (Salomon et al. Bioessays 199, 21, 61-70, US-PS 5,654,140), CASB616 (WO 00/53216), Criptin (US 5,981,215). Zusätzlich umfassen Antigene, die besonders relevant für Impfstoffe in der Therapie von Krebs sind, auch Tyrosinase, Telomerase, P53, NY-Br1.1 (WO 01/47959) und Fragmente davon, wie sie in WO 00/43420 offenbart werden, B726 (WO 00/60076, SEQ ID NOs:469 und 463; WO 01/79286, SEQ ID NOs: 474 und 475); P510 (WO 01/34802, SEQ ID NO:537 und 538) und Survivin.

[0044] Die vorliegende Erfindung ist auch nützlich in Kombination mit Brustkrebs-Antigenen wie Her-2/neu, Mammaglobin (US-PS 5,668,267), B305D (WO 00/61753, SEQ ID NOs: 299, 304, 305 und 315) oder denjenigen, die offenbart werden in WO 00/52165, WO 99/33869, WO 99/19479, WO 98/45328. Her-2/neu-Antigene werden unter anderem in US-PS 5,801,005 offenbart. Bevorzugt umfaßt das Her-2/neu die gesamte extrazelluläre Domäne (umfassend ungefähr die Aminosäuren 1-645) oder Fragmente davon und wenigstens einen immunogenen Teil oder die gesamte intrazelluläre Domäne von etwa den C-terminalen 580 Aminosäuren. Insbesondere sollte der intrazelluläre Teil die Phosphorylierungsdomäne oder Fragmente davon umfassen. Sol-

che Konstrukte werden in WO 00/44899 offenbart. Ein besonders bevorzugtes Konstrukt ist als ECD-PhD bekannt, ein zweites ist als ECD deltaPhD bekannt (siehe WO 00/44899). Das Her-2/neu wie hier verwendet kann aus Ratte, Maus oder Mensch stammen.

[0045] Bestimmte Tumorantigene sind kleine Peptidantigene (d.h. weniger als ca. 50 Aminosäuren). Diese Antigene können chemisch mit dem modifizierten Cholin-Bindungsprotein der vorliegenden Erfindung konjugiert werden.

[0046] Exemplarische Peptide schließen von Mucin abgeleitete Peptide ein, wie zum Beispiel MUC-1 (siehe zum Beispiel US 5,744,144; US 5,827,666; WO 88/05054, US 4,963,484). Spezifisch erwogen werden aus MUC-1 abgeleitete Peptide, die wenigstens eine Repeat-Einheit des MUC-1-Peptids umfassen, bevorzugt wenigstens zwei solcher Repeats, und die durch den SM3-Antikörper erkannt werden (US 6,054,438). Andere, aus Mucin abgeleitete Peptide schließen das Peptid aus MUC-5 ein.

[0047] Alternativ ist das Antigen ein Interleukin wie IL13 und IL14, die bevorzugt sind, oder das Antigen kann ein Selbstpeptidhormon sein, wie zum Beispiel Gonadotrophinhormon-freisetzendes Hormon voller Länge (GnRH, WO 95/20600), ein kurzes, 10 Aminosäuren langes Peptid, das nützlich in der Behandlung von vielen Krebsarten oder in der Immunkastration ist.

[0048] Andere tumorspezifische Antigene, die zur Kopplung mit dem modifizierten Cholin-Bindungsprotein der vorliegenden Erfindung geeignet sind, schließen ohne Beschränkung tumorspezifische Ganglioside wie GM2 und GM3 ein.

[0049] Die kovalente Kopplung des Peptids an modifiziertes Cholin-Bindungsprotein kann in einer allgemein fachbekannten Weise durchgeführt werden. So ist es zum Beispiel zur direkten kovalenten Kopplung möglich, ein Carbodiimid, Glutaraldehyd oder (N-[γ -Maleimidobutyryloxy]succinimidester zu verwenden, wobei übliche handelsübliche heterobifunktionelle Linker wie CDAP und SPDP verwendet werden (unter Verwendung der Herstelleranweisungen). Nach der Kopplungsreaktion kann das Immunogen leicht mittels eines Dialyseverfahrens, eines Gelfiltrationsverfahrens, eines Fraktionierungsverfahrens etc. isoliert und gereinigt werden.

[0050] Das Antigen kann auch aus Quellen abgeleitet werden, die pathogen für Menschen sind, wie zum Beispiel: humanes Immunschwächevirus HIV-1 (wie tat, nef, reverse Transkriptase, gag, gp120 und gp160), humane Herpes simplex-Viren wie gD oder Derivate davon oder unmittelbares frühes Protein wie ICP27 aus HSV1 oder HSV2, Cytomegalovirus (speziell human) (wie gB oder Derivate davon), Rotavirus (einschließlich lebend-abgeschwächter Viren), Epstein-Barr-Virus (wie gp350 oder Derivate davon), Varizella-Zoster-Virus (wie gpl, II und IE63), oder aus einem Hepatitisvirus wie Hepatitis B-Virus (zum Beispiel Hepatitis B-Oberflächenantigen oder ein Derivat davon), Hepatitis A-Virus, Hepatitis C-Virus und Hepatitis E-Virus, oder aus anderen viralen Pathogenen wie Paramyxoviren: respiratorisches Synzytialvirus (wie F- und G-Proteine oder Derivate davon), Parainfluenzavirus, Masernvirus, Mumpsvirus, humane Papillomaviren (zum Beispiel HPV6, 11, 16, 18, ...), Flaviviren (z.B. Gelbfieberevirus, Denguevirus, zeckenbürtiges Enzephalitisvirus, Japanisches Enzephalitisvirus) oder Grippevirus (ganzes lebendes oder inaktiviertes Virus, Spaltgrippevirus, kultiviert in Eiern oder MDCK-Zellen, oder ganze Grippevirosomen (wie beschrieben von R. Gluck, Vaccine, 1992, 10, 915-920) oder gereinigte oder rekombinante Proteine davon wie HA-, NP-, NA- oder M-Proteine oder Kombinationen daraus), oder abgeleitet aus bakteriellen Pathogenen wie *Neisseria* spp., einschließlich *N. gonorrhea* und *N. meningitidis* (zum Beispiel Kapselpolysaccharide und Konjugate davon, Transferrin-bindende Proteine, Lactoferrin-bindende Proteine, Pilc, Adhäsine); *S. pyogenes* (zum Beispiel M-Proteine oder Fragmente davon, C5A-Protease, Lipoteichonsäuren), *S. agalactiae*, *S. mutans*; *H. ducreyi*; *Moraxella* spp., einschließlich *M. catarrhalis*, auch bekannt als *Branhamella catarrhalis* (zum Beispiel Adhäsine und Invasine mit hohem und niedrigem Molekulargewicht); *Bordetella* spp., einschließlich *B. Pertussis* (zum Beispiel Pertactin, Pertussistoxin oder Derivate davon, filamentöses Hämagglutinin, Adenylatcyclase, Fimbrien), *B. parapertussis* und *B. bronchiseptica*; *Mycobacterium* spp., einschließlich *M. tuberculosis* (zum Beispiel ESAT6, Antigen 85A, -B oder -C), *M. bovis*, *M. leprae*, *M. avium*, *M. paratuberculosis*, *M. smegmatis*; *Legionella* spp., einschließlich *L. pneumophila*; *Escherichia* spp., einschließlich enterotoxischem *E. coli* (zum Beispiel Kolonisierungsfaktoren, hitzelabiles Toxin oder Derivate davon, hitzestabiles Toxin oder Derivate davon), enterohämorrhagisches *E. coli*, enteropathogenes *E. coli* (zum Beispiel Shigatoxin-artiges Toxin oder Derivate davon); *Vibrio* spp., einschließlich *V. cholera* (zum Beispiel Cholera toxin oder Derivate davon); *Shigella* spp., einschließlich *S. sonnei*, *S. dysenteriae*, *S. flexnerii*; *Yersinia* spp., einschließlich *Y. enterocolitica* (zum Beispiel ein Yop-Protein), *Y. pestis*, *Y. pseudotuberculosis*; *Campylobacter* spp., einschließlich *C. jejuni* (zum Beispiel Toxine, Adhäsine und Invasine) und *C. coli*; *Salmonella* spp., einschließlich *S. typhi*, *S. paratyphi*, *S. choleraesuis*, *S. enteritidis*; *Listeria* spp., einschließlich *L. monocytogenes*; *Helicobacter* spp., einschließlich *H. pylori* (zum Beispiel Urease, Catalase,

vakuolisierendes Toxin); *Pseudomonas* spp., einschließlich *P. aeruginosa*; *Staphylococcus* spp., einschließlich *S. aureus*, *S. epidermidis*; *Enterococcus* spp., einschließlich *E. faecalis*, *E. faecium*; *Clostridium* spp., einschließlich *C. tetani* (zum Beispiel Tetanustoxin und Derivate davon), *C. botulinum* (zum Beispiel Botulinumtoxin und Derivate davon), *C. difficile* (zum Beispiel Clostridiumtoxine A oder B und Derivate davon); *Bacillus* spp., einschließlich *B. anthracis* (zum Beispiel Botulinumtoxin und Derivate davon); *Corynebacterium* spp., einschließlich *C. diphtheriae* (zum Beispiel Diphtherietoxin und Derivate davon); *Borrelia* spp., einschließlich *B. burgdorferi* (zum Beispiel OspA, OspC, DbpA, DbpB), *B. garinii* (zum Beispiel OspA, OspC, DbpA, DbpB), *B. afzelii* (zum Beispiel OspA, OspC, DbpA, DbpB), *B. andersonii* (zum Beispiel OspA, OspC, DbpA, DbpB), *B. hermsii*; *Ehrlichia* spp., einschließlich *E. equi* und der Verursacher der humanen granulozytischen Ehrlichiose; *Rickettsia* spp., einschließlich *R. rickettsii*; *Chlamydia* spp., einschließlich *C. trachomatis* (zum Beispiel MOMP, Heparin-bindende Proteine), *C. pneumoniae* (zum Beispiel MOMP, Heparin-bindende Proteine), *C. psittaci*; *Leptospira* spp., einschließlich *L. interrogans*; *Treponema* spp., einschließlich *T. pallidum* (zum Beispiel die seltenen Außenmembranproteine), *T. denticola*, *T. hyodysenteriae*; oder abgeleitet aus Parasiten wie *Plasmodium* spp., einschließlich *P. falciparum*; *Toxoplasma* spp., einschließlich *T. gondii* (zum Beispiel SAG2, SAG3, Tg34); *Entamoeba* spp., einschließlich *E. histolytica*; *Babesia* spp., einschließlich *B. microti*; *Trypanosoma* spp., einschließlich *T. cruzi*; *Giardia* spp., einschließlich *G. lamblia*; *Leshmania* spp., einschließlich *L. major*; *Pneumocystis* spp., einschließlich *P. carinii*; *Trichomonas* spp., einschließlich *T. vaginalis*; *Schistosoma* spp., einschließlich *S. mansoni*, oder abgeleitet aus Hefe wie *Candida* spp., einschließlich *C. albicans*; *Cryptococcus* spp., einschließlich *C. neoformans*.

[0051] Andere bevorzugte spezifische Antigene für *M. tuberculosis* sind zum Beispiel Tb Ra12, Tb H9, Tb Ra35, Tb38-1, Erd 14, DPV, MTI, MSL, mTTC2 und hTCCI (WO 99/51748). Proteine für *M. tuberculosis* schließen auch Fusionsproteine und Varianten davon ein, worin wenigstens zwei, bevorzugt drei Polypeptide von *M. tuberculosis* zu einem größeren Protein fusioniert sind. Bevorzugte Fusionen schließen ein: Ra12-TbH9-Ra35, Erd14-DPV-MTI, DPV-MTI-MSL, Erd14-DPV-MTI-MSL-mTCC2, Erd14-DPV-MTI-MSL, DPV-MTI-MSL-mTCC2, TbH9-DPV-MTI (WO 99/51748).

[0052] Die am meisten bevorzugten Antigene für *Chlamydia* schließen zum Beispiel das Protein mit hohem Molekulargewicht (HWMP) (WO 99/17741), ORF3 (EP 366 412) und mutmaßliche Membranproteine (Pmps) ein. Andere *Chlamydia*-Antigene der Impfstoffformulierung können aus der in WO 99/28475 beschriebenen Gruppe ausgewählt werden.

[0053] Bevorzugte bakterielle Antigene stammen aus *Streptococcus* spp., einschließlich *S. pneumoniae* (zum Beispiel Kapselpolysaccharide und Konjugate davon, PsaA, PspA, Streptolysin, Cholin-bindende Proteine) und das Proteinantigen Pneumolysin (Biochem. Biophys. Acta, 1989, 67, 1007; Rubins et al., Microbial Pathogenesis, 25, 337-342) und mutierte entgiftete Derivate davon (WO 90/06951; WO 99/03884). Andere bevorzugte bakterielle Antigene stammen aus *Haemophilus* spp., einschließlich *H. influenzae* Typ B (zum Beispiel PRP und Konjugate davon), nicht-typisierbares *H. influenzae*, zum Beispiel OMP26, Adhäsine mit hohem Molekulargewicht, P5, P6, Protein D und Lipoprotein D, und Fimbrin und Fimbrin-abgeleitete Peptide (US 5,843,464) oder Varianten mit multiplen Kopien oder Fusionsproteine davon.

[0054] Derivate von Hepatitis B-Oberflächenantigenen sind allgemein fachbekannt und schließen unter anderem die in EP-A-414 374; EP-A-0 304 578 und EP 198-474 beschriebenen PreS1, PreS2 S-Antigene ein. In einer bevorzugten Ausführungsform ist das HBV-Antigen HBV-Polymerase (Ji Hoon Jeong et al., 1996, BBRC 223, 264-271; H.J. Lee et al., Biotechnol. Lett. 15, 821-826). In einem anderen bevorzugten Aspekt ist das Antigen innerhalb der Fusion ein HIV-1-Antigen, gp120, speziell bei Expression in CHO-Zellen. In einer weiteren Ausführungsform umfaßt das Antigen gD2t wie oben definiert.

[0055] In einer bevorzugten Ausführungsform der vorliegenden Erfindung umfassen Fusionen ein aus dem humanen Papillomavirus abgeleitetes Antigen (HPV 6a, 6b, 11, 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 und 68), insbesondere diejenigen HPV-Serotypen, die als verantwortlich für Genitalwarzen betrachtet werden (HPV 6 oder HPV 11 und andere), und die für Gebärmutterhalskrebs verantwortlichen HPV-Viren (HPV16, HPV17 und andere).

[0056] Geeignete HPV-Antigene sind E1, E2, E4, E5, E6, E7, L1 und L2. Besonders bevorzugte Formen von prophylaktischen oder therapeutischen Fusionen für Genitalwarzen umfassen L1-Partikel oder Kapsomere und Fusionsproteine, die ein oder mehrere Antigene umfassen, die aus den HPV 6- und HPV 11-Proteinen E6, E7, L1 und L2 ausgewählt sind.

[0057] Die am meisten bevorzugten Formen von Fusionsprotein sind: L2E7, wie in WO 96/26277 offenbart,

und Protein D(1/3)-E7, offenbart in GB 9717953.5 (PCT/EP98/05285).

[0058] Ein bevorzugter HPV-Impfstoff für die Prophylaxe oder Therapie von zervikaler Infektion oder zervikalem Krebs kann HPV 16- oder 18-Antigene umfassen. Zum Beispiel L1- oder L2-Antigenmonomere oder L1- oder L2-Antigene, zusammen präsentiert als ein virusartiges Partikel (VLP), oder das Li-Protein, allein präsentiert in einem VLP oder einer Kapsomerstruktur. Solche Antigene, virusartigen Partikel und Kapsomere sind als solche bekannt. Siehe zum Beispiel WO 94/00152, WO 94/20137, WO 94/05792 und WO 93/02184.

[0059] Zusätzliche frühe Proteine können allein oder als Fusionsproteine eingeschlossen werden, wie zum Beispiel E7, E2 oder bevorzugt E5; besonders bevorzugte Ausführungsformen dafür schließen ein VLP ein, das L1E7-Fusionsproteine umfaßt (WO 96/11272). Besonders bevorzugte HPV 16-Antigene umfassen die frühen Proteine E6 oder E7 in Fusion mit einem Protein D-Träger unter Bildung von Protein D-E6- oder -E7-Fusionen aus HPV 16 oder Kombinationen daraus; oder Kombinationen von E6 oder E7 mit L2 (WO 96/26277).

[0060] Alternativ können die frühen Proteine E6 und E7 aus HPV 16 oder 18 in einem einzelnen Molekül präsentiert werden, bevorzugt einer Protein-D-E6/E7-Fusion. Andere Fusionen enthalten gegebenenfalls eines oder beide E6- und E7-Proteine aus HPV 18, bevorzugt in Form eines Protein D-E6- oder Protein D-E7-Fusionsproteins oder eines Protein D-E6/E7-Fusionsproteins. Fusionen können Antigene aus anderen HPV-Stämmen umfassen, bevorzugt aus den Stämmen HPV 31 oder 33.

[0061] Erfindungsgemäße Fusionen umfassen Antigene, die aus Parasiten stammen, die Malaria verursachen. Zum Beispiel schließen bevorzugte Antigene aus Plasmodia falciparum RTS,S und TRAP ein. RTS ist ein Hybridprotein, das im wesentlichen den gesamten C-terminalen Teil des Circumsporozoit-(CS)-Proteins von P. falciparum umfaßt, gebunden über vier Aminosäuren des preS2-Teils des Hepatitis B-Oberflächenantigens an das Oberflächen-(S)-Antigen des Hepatitis B-Virus. Seine vollständige Struktur wird in der internationalen Patentanmeldung Nr. PCT/EP92/02591 offenbart, veröffentlicht als WO 93/10152 unter Beanspruchung der Priorität der GB-Patentanmeldung Nr. 9124390.7. Bei Expression in Hefe wird RTS als Lipoproteinpartikel erzeugt, und bei Coexpression mit dem S-Antigen aus HBV erzeugt es ein als RTS,S bekanntes gemischtes Partikel. TRAP-Antigene werden in der internationalen Patentanmeldung PCT/GB89/00895 beschrieben, veröffentlicht als WO 90/01496. Eine bevorzugte Ausführungsform der vorliegenden Erfindung ist eine Fusion, in der die antigene Zubereitung eine Kombination aus den RTS,S- und TRAP-Antigenen umfaßt. Andere Plasmodia-Antigene, die wahrscheinlich Kandidaten als Komponenten der Fusion sind, sind P. falciparum MSP1, AMA1, MSP3, EBA, GLURP, RAP1, RAP2, Sequestrin, PfEMP1, Pf332, LSA1, LSA3, STARP, SALSA, PfEXP1, Pfs25, Pfs28, PFS27/25, Pfs16, Pfs48/45, Pfs230 und ihre Analoga in Plasmodium spp.

[0062] Die vorliegende Erfindung stellt auch ein Polynukleotid bereit, das den erfindungsgemäßen Fusionspartner codiert. Die Erfindung betrifft ferner ein Polynukleotid, das mit der hier in [Fig. 1](#) bereitgestellten Polynukleotidsequenz hybridisiert (SEQ ID NO:9 bis 16). In dieser Hinsicht betrifft die Erfindung speziell Polynukleotide, die unter stringenten Bedingungen mit dem hier beschriebenen Polynukleotid hybridisieren. Wie hier verwendet, bedeuten die Begriffe "stringente Bedingungen" und "stringente Hybridisierungsbedingungen" eine Hybridisierung, die nur auftritt, falls es wenigstens 95% und bevorzugt wenigstens 97% Identität zwischen den Sequenzen gibt. Ein spezifisches Beispiel für stringente Hybridisierungsbedingungen ist eine Inkubation über Nacht bei 42°C in einer Lösung, die folgendes umfaßt: 50% Formamid, 5 × SSC (150 mM NaCl, 15 mM Trisatriumcitrat), 50 mM Natriumphosphat (pH 7,6), 5 × Denhardt-Lösung, 10% Dextransulfat und 20 µg/ml von denaturierter, gescherter Lachssperma-DNA, gefolgt von Spülen des Hybridisierungsträgers in 0,1 × SSC bei ca. 65°C. Die Hybridisierungs- und Spülbedingungen sind allgemein bekannt und werden exemplarisch dargestellt in Sambrook et al., Molecular Cloning: A Laboratory Manual, zweite Auflage, Cold Spring Harbor, N.Y. (1989), insbesondere Kapitel 11. Die Lösungshybridisierung kann auch mit den durch die Erfindung bereitgestellten Polynukleotidsequenzen verwendet werden.

[0063] Die vorliegende Erfindung stellt auch ein Polynukleotid bereit, daß das Polypeptid codiert, das den erfindungsgemäßen Fusionspartner umfaßt, fusioniert an ein ein tumorassoziiertes Antigen oder Fragment davon. Insbesondere sieht die vorliegende Erfindung Polynukleotidsequenzen vor, die ein Fusionspartnerprotein codieren, das eine Cholin-Bindungsdomäne und ein heterologes promiskues T-Helferepitop umfaßt, bevorzugt worin die Cholin-Bindungsdomäne aus dem C-Terminus von LytA stammt. In einer besonders bevorzugten Ausführungsform umfaßt die C-LytA-Einheit der Polynukleotide gemäß der Erfindung wenigstens vier Repeats einer beliebigen aus SEQ ID NO:9-14, besonders bevorzugt die Sequenz SEQ ID NO:15, noch mehr bevorzugt die Sequenz SEQ ID NO:16. In anderen verwandten Ausführungsformen sieht die vorliegende Erfindung Polynukleotidvarianten mit substantieller Identität mit den hier in SEQ ID NOs:9-16 offenbarten Sequenzen vor, zum Beispiel mit denjenigen, die wenigstens 70% Sequenzidentität, bevorzugt wenigstens 75%, 80%, 85%,

90%, 95%, 96%, 97%, 98% oder 99% oder eine höhere Sequenzidentität umfassen, verglichen mit einer Polynukleotidsequenz dieser Erfindung unter Verwendung herkömmlicher Verfahren z.B. durch BLAST-Analyse unter Verwendung von Standardparametern. In noch einer weiteren Ausführungsform umfaßt das Polynukleotid wie beansprucht ferner ein heterologes Protein.

[0064] Solche Polynukleotidsequenzen können in einen geeigneten Expressionsvektor inseriert und in einem geeigneten Wirt exprimiert werden. Vektoren können bereitgestellt werden, die das modifizierte Cholin-bindende Protein der Erfindung codieren und die einen geeigneten Restriktionsort enthalten, in den eine DNA inseriert werden kann, die ein schwach immunogenes Protein codiert, um ein Fusionsprotein zu erzeugen. In anderen Ausführungsformen der Erfindung können Polynukleotidsequenzen oder Fragmente davon, die Polypeptidfusionen der Erfindung codieren, in rekombinanten DNA-Molekülen zur Ausrichtung der Expression eines Polypeptids in geeigneten Wirtszellen verwendet werden. Aufgrund der inhärenten Degeneriertheit des genetischen Codes können andere DNA-Sequenzen, die im wesentlichen die gleiche oder eine funktionelle Äquivalente Aminosäuresequenz codieren, hergestellt werden, und diese Sequenzen können zur Klonierung und Expression eines gegebenen Polypeptids verwendet werden.

[0065] Wie die Fachleute verstehen werden, kann es vorteilhaft in manchen Fällen sein, Polypeptid-codierende Nukleotidsequenzen zu erzeugen, die nicht-natürlich vorkommende Codonen besitzen. Der DNA-Code besitzt 4 Buchstaben (A, T, C und G) und verwendet diese zum Buchstabieren von "Codonen", die die Aminosäuren der in den Genen eines Organismus codierten Proteine darstellen. Die lineare Sequenz von Codonen entlang des DNA-Moleküls wird zur linearen Sequenz von Aminosäuren in den durch diese Gene codierten Proteinen übersetzt. Der Code ist hochgradig degeneriert, wobei 61 Codonen für die 20 natürlichen Aminosäuren codieren und drei Codonen "Stopp"-Signale darstellen. So wird für die meisten Aminosäuren durch mehr als ein Codon codiert – tatsächlich wird für mehrere durch vier oder mehr unterschiedliche Codonen codiert.

[0066] Wenn mehr als ein Codon verfügbar ist, um für eine gegebene Aminosäure zu codieren, wurde beobachtet, daß die Codon-Verwendungsmuster von Organismen hochgradig nicht zufällig sind. Unterschiedliche Arten zeigen eine unterschiedliche Betonung in ihrer Codonauswahl, und ferner kann die Verwendung von Codonen deutlich unterschiedlich in einer einzelnen Art zwischen Genen sein, die mit hohen und niedrigen Stärken exprimiert werden. Diese Betonung unterscheidet sich in Viren-, Pflanzen-, Bakterien- und Sängerezellen, und einige Arten zeigen eine stärkere Betonung weg von einer zufälligen Codonenauswahl als andere. Zum Beispiel sind Menschen und andere Säuger weniger stark unausgewogen als bestimmte Bakterien oder Viren. Aus diesen Gründen gibt es eine signifikante Wahrscheinlichkeit, daß ein in *E. coli* exprimiertes Sängergen oder ein in Sängerezellen exprimiertes virales Gen eine unangemessene Verteilung von Codonen für eine effiziente Expression haben wird. Es wird angenommen, daß die Gegenwart von Clustern von Codonen in einer heterologen DNA-Sequenz, die selten im Wirt beobachtet werden, in dem die Expression erfolgt, vorhersagend für geringe heterologe Expressionsgrade in diesem Wirt ist.

[0067] Als Ergebnis können Codonen, die von einem besonderen prokaryontischen (zum Beispiel *E. coli* oder Hefe) oder eukaryontischen Wirt bevorzugt werden, optimiert werden, d.h. ausgewählt werden zur Erhöhung der Rate der Proteinexpression, zur Erzeugung eines rekombinanten RNA-Transkripts mit wünschenswerten Eigenschaften, wie zum Beispiel einer Halbwertszeit, die länger als diejenige eines Transkripts ist, das aus der natürlich vorkommenden Sequenz erzeugt wird, oder zur Optimierung der Immunreaktion in Menschen. Der Prozeß der Codonoptimierung kann jede Sequenz einschließen, erzeugt entweder manuell oder durch Computersoftware, worin einige oder alle der Codonen der nativen Sequenz modifiziert werden. Mehrere Verfahren wurden veröffentlicht (Nakamura et al., *Nucleic Acids Research* 1996, 24:214-215; WO98/34640). Ein bevorzugtes Verfahren gemäß dieser Erfindung ist das Syngene-Verfahren, eine Modifikation des Calcgene-Verfahrens (R.S. Hale und G. Thompson (*Protein Expression and Purification*, Bd. 12, S. 185-188 (1998))).

[0068] Entsprechend hat die DNA-Sequenz des Proteins in einer bevorzugten Ausführungsform einen RS-CU-Wert ("Relative synonyms codon usage" (ebenfalls bekannt als Codon Index, CI)) von wenigstens 0,65 und weniger als 85% Identität mit der entsprechenden Wildtypregion.

[0069] Dieser Prozeß der Codonoptimierung und die resultierenden Konstrukte sind vorteilhaft, da sie einige oder alle der folgenden Vorzüge haben können: 1) Verbesserung der Expression des Genprodukts durch Ersetzen von seltenen oder nicht häufig verwendeten Codonen durch häufiger verwendete Codonen, 2) Entfernung oder Einschluß von Restriktionsenzymorten zur Erleichterung der Klonierung stromabwärts und 3) Reduzierung des Potentials für eine homologe Rekombination zwischen der Insertsequenz im DNA-Vektor und genomischen Sequenzen und 4) Verbesserung der Immunreaktion in Menschen durch Hervorrufen einer zellulären und/oder Antikörper-Reaktion (bevorzugt von beiden Reaktionen) gegen das Zielantigen. Die Sequen-

zen der vorliegenden Erfindung besitzen vorteilhaft reduziertes Rekombinationspotential, aber exprimieren mit wenigstens dem gleichen Grad wie die Wildtypsequenzen. Aufgrund der Natur des vom SynGene-Programm verwendeten Algorithmus zur Erzeugung einer Codon-optimierten Sequenz ist es möglich, eine äußerst hohe Anzahl unterschiedlicher Codon-optimierter Sequenzen zu erzeugen, die eine ähnliche Funktion durchführen werden. Kurz gesagt werden die Codonen unter Verwendung eines statistischen Verfahrens zugeordnet, um ein synthetisches Gen mit einer Codonhäufigkeit zu ergeben, die näher an derjenigen ist, die natürlich in hoch-exprimiertem *E. coli* und humanen Genen gefunden wird. Kurz gesagt werden die Codonen unter Verwendung eines statistischen Verfahrens zugeordnet, um ein synthetisches Gen mit einer Codonhäufigkeit zu ergeben, die näher an derjenigen ist, die natürlicherweise in hoch-exprimierten humanen Genen wie β -Actin gefunden wird. Illustrativ, obwohl ohne Beschränkung, werden Beispiele für geeignete Codon-optimierte Sequenzen in SEQ ID NOs:19-22 und SEQ ID NOs:24-26 angegeben.

[0070] In den Polynukleotiden der vorliegenden Erfindung ist das Codonverwendungsmuster von demjenigen verändert, das typisch für das Zielantigen ist, um stärker die Codonbetonung eines hoch-exprimierten Gens in einem Zielorganismus darzustellen, zum Beispiel humanes β -Actin. Der "Codon-Verwendungskoeffizient" ist ein Maß dafür, wie nahe das Codonmuster einer gegebenen Polynukleotidsequenz demjenigen einer Zielart ähnelt. Codonhäufigkeiten können aus Literaturquellen für die hoch-exprimierten Gene vieler Arten abgeleitet werden (siehe z.B. Nakamura et al., *Nucleic Acids Research* 1996, 24:214-215). Die Codonhäufigkeiten für jedes der 61 Codonen (ausgedrückt als die Anzahl von Auftritten auf 1000 Codonen der ausgewählten Klasse von Genen) werden für jede der 20 natürlichen Aminosäuren normalisiert, so daß der Wert für das am häufigsten verwendete Codon für jede Aminosäure auf 1 eingestellt ist und die Häufigkeiten für die weniger üblichen Codonen skaliert sind, so daß sie zwischen Null und 1 liegen. Somit wird jedem der 61 Codonen ein Wert von 1 oder darunter für die hoch-exprimierten Gene der Zielart zugeordnet. Zur Berechnung eines Codonverwendungskoeffizienten für ein spezifisches Polynukleotid, relativ zu den hoch-exprimierten Genen der Art, wird der skalierte Wert für jedes Codon des spezifischen Polynukleotids notiert, und der geometrische Mittelwert all dieser Werte wird herangezogen (durch Dividieren der Summe der natürlichen Logarithmen dieser Werte durch die Gesamtanzahl von Codonen und Berechnung des umgekehrten Logarithmus). Der Koeffizient wird einen Wert zwischen 0 und 1 haben, und je höher der Koeffizient ist, desto mehr Codonen im Polynukleotid sind häufig verwendete Codonen. Falls eine Polynukleotidsequenz einen Codonverwendungskoeffizienten von 1 hat, sind alle Codonen "häufigste" Codonen für hoch-exprimierte Gene der Zielart.

[0071] Erfindungsgemäß wird das Codonverwendungsmuster des Polynukleotids vorzugsweise Codonen ausschließen, die <10% der Codonen darstellen, die für eine besondere Aminosäure verwendet werden. Ein relativer synonymer Codonverwendungswert (RSCU) ist die beobachtete Anzahl von Codonen, geteilt durch die erwartete Anzahl, falls alle Codonen für die Aminosäure gleich häufig verwendet würden. Ein Polynukleotid der vorliegenden Erfindung wird bevorzugt Codonen mit einem RSCU-Wert von weniger als 0,2 in hoch-exprimierten Genen des Zielorganismus ausschließen. Ein Polynukleotid der vorliegenden Erfindung wird allgemein einen Codon-Verwendungskoeffizienten für hoch-exprimierte humane Gene von mehr als 0,6, bevorzugt von mehr als 0,65, am meisten bevorzugt von mehr als 0,7 haben. Codon-Verwendungstabellen für Menschen können auch in Genbank gefunden werden.

[0072] Im Vergleich hat ein hoch-exprimiertes beta-Actin einen RSCU-Wert von 0,747.

[0073] Die Codon-Verwendungstabelle (Tabelle) für *Homo sapiens* ist nachfolgend dargestellt:

Tabelle 1. Codonverwendung für humane (hochexprimierte) Gene 1/24/91 (human_high.cod)

Aminosäure	Codon	Anzahl	/1000	Anteil
Gly	GGG	905,00	18,76	0,24
Gly	GGA	525,00	10,88	0,14
Gly	GGT	441,00	9,14	0,12
Gly	GGC	1867,00	38,70	0,50
Glu	GAG	2420,00	50,16	0,75
Glu	GAA	792,00	16,42	0,25
Asp	GAT	592,00	12,27	0,25
Asp	GAC	1821,00	37,75	0,75
Val	GTG	1866,00	38,68	0,64
Val	GTA	134,00	2,78	0,05
Val	GTT	198,00	4,10	0,07
Val	GTC	728,00	15,09	0,25
Ala	GCG	652,00	13,51	0,17
Ala	GCA	488,00	10,12	0,13
Ala	GCT	654,00	13,56	0,17
Ala	GCC	2057,00	42,64	0,53
Arg	AGG	512,00	10,61	0,18
Arg	AGA	298,00	6,18	0,10
Ser	AGT	354,00	7,34	0,10
Ser	AGC	1171,00	24,27	0,34

DE 603 13 163 T2 2008.01.03

Aminosäure	Codon	Anzahl	/1000	Anteil
Lys	AAG	2117,00	43,88	0,82
Lys	AAA	471,00	9,76	0,18
Asn	AAT	314,00	6,51	0,22
Asn	AAC	1120,00	23,22	0,78
Met	ATG	1077,00	22,32	1,00
Ile	ATA	88,00	1,82	0,05
Ile	ATT	315,00	6,53	0,18
Ile	ATC	1369,00	28,38	0,77
Thr	ACG	405,00	8,40	0,15
Thr	ACA	373,00	7,73	0,14
Thr	ACT	358,00	7,42	0,14
Thr	ACC	1502,00	31,13	0,57
Trp	TGG	652,00	13,51	1,00
End	TGA	109,00	2,26	0,55
Cys	TGT	325,00	6,74	0,32
Cys	TGC	706,00	14,63	0,68
End	TAG	42,00	0,87	0,21
End	TAA	46,00	0,95	0,23
Tyr	TAT	360,00	7,46	0,26
Tyr	TAC	1042,00	21,60	0,74
Leu	TTG	313,00	6,49	0,06
Leu	TTA	76,00	1,58	0,02
Phe	TTT	336,00	6,96	0,20
Phe	TTC	1377,00	28,54	0,80
Ser	TCG	325,00	6,74	0,09
Ser	TCA	165,00	3,42	0,05
Ser	TCT	450,0	9,33	0,13
Ser	TCC	958,00	19,86	0,28

Aminosäure	Codon	Anzahl	/1000	Anteil
Arg	CGG	611,00	12,67	0,21
Arg	CGA	183,00	3,79	0,06
Arg	CGT	210,00	4,35	0,07
Arg	CGC	1086,00	22,51	0,37
Gln	CAG	2020,00	41,87	0,88
Gln	CAA	283,00	5,87	0,12
His	CAT	234,00	4,85	0,21
His	CAC	870,00	18,03	0,79
Leu	CTG	2884,00	59,78	0,58
Leu	CTA	166,00	3,44	0,03
Leu	CTT	238,00	4,93	0,05
Leu	CTC	1276,00	26,45	0,26
Pro	CCG	482,00	9,99	0,17
Pro	CCA	456,00	9,45	0,16
Pro	CCT	568,00	11,77	0,19
Pro	CCC	1410,00	29,23	0,48

[0074] Eine DNA-Sequenz, die die Fusionsproteine oder das modifizierte Cholin-bindende Protein der vorliegenden Erfindung codiert, kann unter Verwendung von Standard-DNA-Synthesetechniken synthetisiert werden, wie zum Beispiel durch enzymatische Ligation, wie beschrieben von D.M. Roberts et al., Biochemistry 1985, 24, 5090-5098, durch chemische Synthese, durch enzymatische Polymerisation in vitro oder durch PCR-Technologie unter Verwendung zum Beispiel einer hitzestabilen Polymerase oder durch eine Kombination dieser Techniken.

[0075] Die enzymatische Polymerisation von DNA kann in vitro unter Verwendung einer DNA-Polymerase wie DNA-Polymerase I (Klenow-Fragment) oder Taq-Polymerase in einem geeigneten Puffer durchgeführt werden, der die Nukleosidtriphosphate dATP, dCTP, dGTP und dTTP enthält, nach Bedarf bei einer Temperatur von 10-37°C, allgemein in einem Volumen von 50 µl oder weniger. Die enzymatische Ligation von DNA-Fragmenten kann unter Verwendung einer DNA-Ligase wie T4-DNA-Ligase in einem geeigneten Puffer wie 0,05 M Tris (pH 7,4), 0,01 M MgCl₂, 0,01 M Dithiothreitol, 1 mM Spermidin, 1 mM ATP und 10 mg/ml Rinderserumalbumin bei einer Temperatur von 4°C bis Umgebungstemperatur durchgeführt werden, allgemein in einem Volumen von 50 µl oder weniger. Die chemische Synthese des DNA-Polymers oder der Fragmente kann durch herkömmliche Phosphotriester-, Phosphat- oder Phosphoramiditchemie unter Verwendung von Festphasentechniken wie durch diejenigen durchgeführt werden, die beschrieben werden in "Chemical and Enzymatic Synthesis of Gene Fragments – A Laboratory Manual" (Hrsg. H.G. Gassen und A. Lang), Verlag Chemie, Weinheim (1982), oder in anderen wissenschaftlichen Veröffentlichungen, z.B. M.J. Gait, H.W.D. Matthes, M. Singh, B.S. Sproat und R.C. Titmas, Nucleic Acids Research, 1982, 10, 6243; B.S. Sproat und W. Bannwarth, Tetrahedron Letters, 1983, 24, 5771; M.D. Matteucci und M.H. Caruthers, Tetrahedron Letters, 1980, 21, 719; M.D. Matteucci und M.H. Caruthers, Journal of the American Chemical Society, 1981, 103, 3185; S.P. Adams et al., Journal of the American Chemical Society, 1983, 105, 661; N.D. Sinha, J. Biernat, J. McMannus und H. Koster, Nucleic Acids Research, 1984, 12, 4539; und H.W.D. Matthes et al., EMBO Journal, 1984, 3, 801.

[0076] Das Verfahren der Erfindung kann durch herkömmliche rekombinante Techniken durchgeführt werden, wie zum Beispiel beschrieben in Maniatis et al., Molecular Cloning – A Laboratory Manual; Cold Spring Harbor, 1982-1989.

[0077] Insbesondere kann das Verfahren die folgenden Schritte umfassen:

- i) Herstellen eines replizierbaren oder integrierenden Expressionsvektors, der in einer Wirtszelle ein DNA-Polymer exprimieren kann, das eine Nukleotidsequenz umfaßt, die das Protein oder ein immunogenes Derivat davon codiert,

- ii) Transformieren einer Wirtszelle mit dem Vektor,
- iii) Kultivieren der transformierten Wirtszelle unter Bedingungen, die die Expression des DNA-Polymers zur Erzeugung des Proteins erlauben; und
- iv) Gewinnen des Proteins.

[0078] Der Begriff "Transformieren" wird hier verwendet, um die Einführung von Fremd-DNA in eine Wirtszelle zu bezeichnen. Dies kann zum Beispiel durch Transformation, Transfektion oder Infektion mit einem geeigneten Plasmid oder viralen Vektor unter Verwendung von z.B. herkömmlichen Techniken erreicht werden, wie sie beschrieben werden in Genetic Engineering; Hrsg. S.M. Kingsman und A.J. Kingsman; Blackwell Scientific Publications; Oxford, England, 1988. Der Begriff "transformiert" oder "Transformante" wird hier nachfolgend die resultierende Wirtszelle bezeichnen, die das Fremdgen von Interesse enthält und exprimiert.

[0079] Die Expressionsvektoren sind neu und bilden ebenfalls einen Teil der Erfindung.

[0080] Die replizierbaren Expressionsvektoren können erfindungsgemäß durch Spalten eines mit der Wirtszelle kompatiblen Vektors zur Bereitstellung eines linearen DNA-Segments mit einem intakten Replikon und durch Kombinieren des linearen Segments mit einem oder mehreren DNA-Molekülen, die zusammen mit dem linearen Segment das gewünschte Produkt codieren, wie zum Beispiel das DNA-Polymer, das das Protein der Erfindung codiert, oder ein Derivat davon, unter ligierenden Bedingungen hergestellt werden.

[0081] So kann das DNA-Polymer nach Wunsch vorgebildet oder während der Konstruktion des Vektors gebildet werden.

[0082] Die Wahl des Vektors wird zum Teil durch die Wirtszelle bestimmt werden, die prokaryontisch oder eukaryontisch sein kann, aber bevorzugt E. coli-, Hefe- oder CHO-Zellen sein wird. Geeignete Vektoren schließen Plasmide, Bakteriophagen, Cosmide und rekombinante Viren ein. Expressions- und Klonierungsvektoren enthalten bevorzugt einen selektierbaren Marker, so daß nur die Wirtszellen, die den Marker exprimieren, unter selektiven Bedingungen überleben werden. Selektionsgene schließen ohne Beschränkung dasjenige ein, das ein Protein codiert, das Resistenz gegen Ampicillin, Tetracyclin oder Kanamycin verleiht. Expressionsvektoren enthalten auch Kontrollsequenzen, die kompatibel mit dem ausgewählten Wirt sind. Zum Beispiel schließen Expressionskontrollsequenzen für E. coli und allgemein für Prokaryonten Promotoren und Ribosombindungsstellen ein. Promotorsequenzen können natürlich vorkommend sein, wie das β -Lactamase- (Penicillinase) (Weisman 1981, Interferon 3 (Hrsg. L. Gresser)), Lactose-(lac) (Chang et al., Nature, 1977, 198: 1056) und Tryptophan- (trp) (Goeddel et al., Nucl. Acids Res., 1980, 8, 4057) und das lambda-abgeleitete P_L -Promotorsystem. Zusätzlich funktionieren auch synthetische Promotoren, die nicht in der Natur vorkommen, als bakterielle Promotoren. Dies ist beispielsweise der Fall für den synthetischen tac-Hybridpromotor, der aus Sequenzen der trp- und lac-Promotoren abgeleitet ist (De Boer et al., Proc. Natl. Acad. Sci. USA, 1983, 80, 21-26). Diese Systeme sind besonders geeignet mit E. coli.

[0083] Hefe-kompatible Vektoren tragen auch Marker, die die Selektion von erfolgreichen Transformanten durch Verleihen von Prototrophie für auxotrophe Mutanten oder Resistenz gegen Schwermetalle für Stämme vom Wildtyp erlauben. Expressionskontrollsequenzen für Hefevektoren schließen Promotoren für glycolytische Enzyme (Ness et al., J. Adv. Enzyme Reg. 1968, 7, 149), PHO5-Gen, das saure Phosphatase codiert, CUP1-Gen, ARG3-Gen, GAL-Genpromotoren und synthetische Promotorsequenzen ein. Andere Kontrollelemente, die in der Hefeexpression nützlich sind, sind Terminatoren und mRNA-Leitsequenzen. Die 5'-codierende Sequenz ist besonders nützlich, da sie typischerweise ein Signalpeptid codiert, das aus hydrophoben Aminosäuren zusammengesetzt ist, die die Sezernierung des Proteins aus der Zelle ausrichten. Geeignete Signalsequenzen können durch Gene für sezernierte Hefeproteine codiert werden, wie das Hefe-Invertasegen und das α -Faktor-Gen, saure Phosphatase, Killertoxin, das alpha-Paarungsfaktorgen und kürzlich die heterologe Inulinase-Signalsequenz, die aus dem INU1A-Gen von Kluyveromyces marxianus stammt. Geeignete Vektoren wurden zur Expression in Pichia pastoris und Saccharomyces cerevisiae entwickelt.

[0084] Eine Vielzahl von P. pastoris-Expressionsvektoren sind auf Basis verschiedener induzierbarer oder konstitutiver Promotoren verfügbar (Cereghino und Cregg, FEMS Microbiol. Rev. 2000, 24:45-66). Zur Erzeugung von cytosolischen und sezernierten Proteinen enthalten die am häufigsten verwendeten P. pastoris-Vektoren den sehr starken und eng regulierten Alkoholoxidase-(AOX1)-Promotor. Die Vektoren enthalten auch das P. pastoris Histidinoldehydrogenase-(HIS4)-Gen zur Selektion in his4-Wirten. Die Sezernierung von Fremdprotein erfordert die Gegenwart einer Signalsequenz, und die S. cerevisiae Präpro-alpha-Paarungsfaktor-Signalsequenz wurde weithin und erfolgreich im Pichia-Expressionssystem verwendet. Expressionsvektoren werden in das P. pastoris Genom zur Maximierung der Stabilität von Expressionsstämmen integriert. Wie in S. cerevi-

siae stimuliert die Spaltung eines *P. pastoris*-Expressionsvektors innerhalb einer vom Wirtsgenom geteilten Sequenz (AOX1 oder HIS4) homologe Rekombinationsereignisse, die effizient die Integration des Vektors auf diesen genomischen Locus ausrichten. Allgemein kann ein rekombinanter Stamm, der multiple integrierte Kopien einer Expressionskassette enthält, mehr heterologes Protein als ein Einzelkopiestamm liefern. Der effektivste Weg zum Erhalt von Transformanten mit hoher Kopiezahl erfordert die Transformation eines *Pichia*-Empfängerstamms durch die Sphäroplastentechnik (Cregg et al. 1985, Mol. Cell. Biol. 5: 3376-3385).

[0085] Die Herstellung des replizierbaren Expressionsvektors kann herkömmlich mit geeigneten Enzymen zur Restriktion, Polymerisation und Ligation der DNA durch Verfahren durchgeführt werden, die zum Beispiel in Maniatis et al., oben zitiert, beschrieben werden.

[0086] Die rekombinante Wirtszelle wird erfindungsgemäß durch Transformieren einer Wirtszelle mit einem replizierbaren Expressionsvektor der Erfindung unter transformierenden Bedingungen hergestellt. Geeignete transformierende Bedingungen sind herkömmlich und werden zum Beispiel beschrieben in Maniatis et al., oben zitiert, oder in "DNA-Cloning", Bd. II, D.M. Glover, Hrsg., IRL Press Ltd., 1985.

[0087] Die Wahl der transformierenden Bedingungen hängt von der Wahl der zu transformierenden Wirtszelle ab. Zum Beispiel wird oben die In-vitro-Transformation unter Verwendung eines lebenden viralen Vektors als transformierendes Mittel für die Polynukleotide der Erfindung beschrieben. Die bakterielle Transformation eines Wirtes wie *E. coli* kann durch direkte Aufnahme der Polynukleotide (die Expressionsvektoren sein können, die die gewünschte Sequenz enthalten) erfolgen, nachdem der Wirt mit einer CaCl_2 -Lösung (Cohen et al., Proc. Nat. Acad. Sci., 1973, 69, 2110) oder mit einer Lösung, die eine Mischung aus Rubidiumchlorid (RbCl), MnCl_2 , Kaliumacetat und Glycerin umfaßt, und dann mit 3-[N-Morpholino]-propan-sulfonsäure, RbCl und Glycerin behandelt wurde, oder durch Elektroporation. Die Transformation von niederen eukaryontischen Organismen wie Hefezellen in Kultur durch Direktaufnahme kann zum Beispiel unter Verwendung des Verfahrens von Hinnen et al. durchgeführt werden (Proc. Natl. Acad. Sci. 1978, 75:1929-1933). Sängerezellen in Kultur können unter Verwendung der Calciumphosphat-Copräzipitation der Vektor-DNA auf die Zellen transformiert werden (Graham & Van der Eb, Virology 1978, 52, 546). Andere Verfahren zur Einführung von Polynukleotiden in Sängerezellen schließen die Dextran-vermittelte Transfektion, die Polybren-vermittelte Transfektion, die Protoplastenfusion, Elektroporation, Verkapselung des Polynukleotids (der Polynukleotide) in Liposomen und die direkte Mikroinjektion der Polynukleotide in Kerne ein.

[0088] Die Erfindung erstreckt sich auch auf eine Wirtszelle, die mit einer Nukleinsäure, die das Protein der Erfindung codiert, oder einem replizierbaren Expressionsvektor der Erfindung transformiert ist.

[0089] Das Kultivieren der transformierten Wirtszelle unter Bedingungen, die Expression des DNA-Polymers erlauben, wird herkömmlich durchgeführt, wie zum Beispiel beschrieben in Maniatis et al. und "DNA Cloning", oben zitiert. So wird die Zelle bevorzugt mit Nährmittel versorgt und bei einer Temperatur unterhalb 50°C , bevorzugt zwischen 25 und 42°C , besonders bevorzugt zwischen 25 und 35°C , am meisten bevorzugt bei 30°C kultiviert. Die Inkubationszeit kann von einigen Minuten bis zu einigen Stunden gemäß dem Anteil des Polypeptids in der bakteriellen Zelle gemäß Bewertung durch SDS-PAGE oder Western-Blot variieren.

[0090] Das Produkt kann durch herkömmliche Verfahren gemäß der Wirtszelle und gemäß der Lokalisierung des Expressionsprodukts (intrazellulär oder sezerniert in das Kulturmedium oder in das Zellperiplasma) gewonnen werden. So kann die Wirtszelle, wenn sie bakteriell ist, wie zum Beispiel *E. coli*, zum Beispiel physikalisch, chemisch oder enzymatisch lysiert und das Proteinprodukt aus dem resultierenden Lysat isoliert werden. Wenn die Wirtszelle aus Säugetier stammt, kann das Produkt allgemein aus dem Nährmedium oder aus zellfreien Extrakten isoliert werden. Wenn die Wirtszelle eine Hefe wie *Saccharomyces cerevisiae* oder *Pichia pastoris* ist, kann das Produkt allgemein aus lysierten Zellen oder aus dem Kulturmedium isoliert und dann weiter unter Verwendung herkömmlicher Techniken gereinigt werden. Die Spezifität des Expressionssystems kann durch Western Blot oder durch ELISA unter Verwendung eines gegen das Polypeptid von Interesse gerichteten Antikörpers bewertet werden.

[0091] Herkömmliche Proteinisolationstechniken schließen die selektive Präzipitation, Adsorptionschromatographie und Affinitätschromatographie, einschließlich einer monoklonalen Antikörper-Affinitätssäule, ein. Wenn die Proteine der vorliegenden Erfindung mit einem Histidinschwanz (His-Tag) exprimiert werden, können sie leicht durch Affinitätschromatographie unter Verwendung einer Ionenmetall-Affinitätschromatographiesäule (IMAC) gereinigt werden. Das Metallion kann jedes geeignete Ion sein, zum Beispiel Zink, Nickel, Eisen, Magnesium oder Kupfer, aber ist bevorzugt Zink oder Nickel. Bevorzugt enthält der IMAC-Puffer Detergens, bevorzugt ein anionisches Detergens wie SDS, besonders bevorzugt ein nichtionisches Detergens wie Tween 80

oder ein zwitterionisches Detergens wie Empigen BB, da dies zu geringeren Mengen von Endotoxin im Endprodukt führen kann.

[0092] Weitere chromatographische Schritte schließen zum Beispiel einen Q-Sepharose-Schritt ein, der entweder vor oder nach IMAC-Säule durchgeführt werden kann. Bevorzugt ist der pH im Bereich von 7,5 bis 10, besonders bevorzugt von 7,5 bis 9,5, optimalerweise zwischen 8 und 9.

[0093] Die Proteine der Erfindung können somit gemäß dem folgenden Protokoll gereinigt werden. Nach der Zellaufbrechung können die das Protein enthaltenden Zellextrakte in einem Tris-Puffer mit pH 8,5, der Harnstoff (z.B. 8,0 M) und SDS (mit zum Beispiel 0,5 bis 1%) enthält, solubilisiert werden. Nach Zentrifugieren kann der resultierende Überstand dann auf eine IMAC (Nickel) Sepharose FF-Säule geladen werden, die mit Tris-Puffer von pH 8,5 äquibriert ist. Die Säule kann dann mit einem viel Salz enthaltenden Puffer (z.B. 0,75-1,5 M NaCl, 15 mM pH 8,5 Tris-Puffer) gespült werden. Die Säule kann dann gegebenenfalls erneut mit Phosphatpuffer ohne Salz gespült werden. Die Proteine der Erfindung können aus der Säule mit einer Imidazol-haltigen gepufferten Lösung eluiert werden. Die Proteine können dann einem zusätzlichen chromatographischen Schritt wie einer Anionenaustauscherchromatographie (zum Beispiel Q-Sepharose) unterworfen werden.

[0094] Die Proteine der vorliegenden Erfindung werden entweder löslich in einer flüssigen Form oder in einer lyophilisierten Form, die die bevorzugte Form ist, bereitgestellt. Es wird allgemein erwartet, daß jede humane Dosis 1 bis 1000 µg Protein und bevorzugt 30-300 µg Protein umfassen wird. Das Reinigungsverfahren kann auch einen Carboxyamidierungsschritt einschließen, durch den das Protein zuerst in Gegenwart von Glutathion reduziert und dann in Gegenwart von Iodacetamid carboxymethyliert wird. Dieser Schritt bietet den Vorteil der Kontrolle der oxidativen Aggregation des Moleküls mit sich selbst oder mit Wirtszell-Proteinverunreinigungen durch kovalente Verbrückung mit Disulfidbindungen.

[0095] Die vorliegende Erfindung stellt auch pharmazeutische und immunogene Zusammensetzungen bereit, die ein Protein der vorliegenden Erfindung in einem pharmazeutisch akzeptablen Exzipienten umfassen. Eine bevorzugte Impfstoffzusammensetzung umfaßt wenigstens ein erfindungsgemäßes Protein. Das Protein weist bevorzugt geblockte Thiol-Gruppen auf und ist hochgereinigt, zum Beispiel weist weniger als 5% Wirtszellkontamination auf. Ein solcher Impfstoff kann gegebenenfalls ein oder mehrere andere Tumorassoziierte Antigene und Derivate enthalten. Zum Beispiel schließen geeignete andere assoziierte Antigene Prostase, PAP-1, PSA (Prostata-spezifisches Antigen), PSMA (Prostata-spezifisches Membranantigen), PSCA (Prostata-Stammzellantigen) und STEAP ein.

[0096] In einer anderen Ausführungsform umfassen illustrative immunogene Zusammensetzungen, wie zum Beispiel Impfstoffzusammensetzungen, der vorliegenden Erfindung DNA, die ein oder mehrere der Fusionspolypeptide wie oben beschrieben codiert, so daß das Fusionspolypeptid in situ erzeugt wird. Wie oben festgestellt, kann das Polynukleotid mit jedem einer Vielzahl von Übertragungssystemen verabreicht werden, die den Durchschnittsfachleuten bekannt sind. Tatsächlich sind zahlreiche Genübertragungstechniken allgemein fachbekannt, wie zum Beispiel diejenigen, die beschrieben werden von Rolland, Crit. Rev. Therap. Drug Carrier Systems 15:143-198, 1998, und in darin zitierten Verweisen. Geeignete Polynukleotidexpressionssysteme werden natürlich die notwendigen regulatorischen DNA-Regulationssequenzen zur Expression in einem Patienten enthalten (wie einen geeigneten Promotor und ein Terminationssignal). Alternativ können bakterielle Übertragungssysteme die Verabreichung eines Bakteriums (wie Bacillus-Calmette-Guerrin) beinhalten, das einen immunogenen Teil des Polypeptids auf seiner Zelloberfläche exprimiert oder ein solches Epitop sezerniert.

[0097] Deshalb werden in bestimmten Ausführungsformen Polynukleotide, die die hier beschriebene immunogene Polypeptide codieren, in geeignete Säuger-Wirtszellen zur Expression unter Verwendung eines beliebigen einer Anzahl bekannter Systeme auf Virusbasis eingeführt. In einer illustrativen Ausführungsform liefern Retroviren eine zweckmäßige und wirksame Plattform für Genübertragungssysteme. Eine ausgewählte Nukleotidsequenz, die ein Polypeptid der vorliegenden Erfindung codiert, kann in einen Vektor inseriert und in retroviralen Partikeln unter Verwendung von fachbekannten Techniken verpackt werden. Das rekombinante Virus kann dann isoliert und auf einen Probanden übertragen werden. Eine Anzahl von illustrativen retroviralen Systemen wurde beschrieben (z.B. US-PS 5,219,740; Miller und Rosman (1989), BioTechniques 7:980-990; A.D. Miller (1990), Human Gene Therapy 1:5-14; Scarpa et al. (1991) Virology 180:849-852; Burns et al. (1993), Proc. Natl. Acad. Sci. USA 90:8033-8037; und Boris-Lawrie und Temin (1993), Cur. Opin. Genet. Develop. 3:102-109).

[0098] Zusätzlich wurden eine Anzahl illustrativer Systeme auf Adenovirus-Basis beschrieben. Anders als

Retroviren, die in das Wirtsgenom integrieren, dauern Adenoviren extrachromosomal fort, wodurch die mit der Insertionsmutagenese assoziierten Risiken minimiert werden (Haj-Ahmad und Graham (1986) J. Virol. 27:267-274; Bett et al. (1993), J. Virol. 67:5911-5921; Mittereder et al. (1994), Human Gene Therapy 5:717-729; Seth et al. (1994), J. Virol. 68:933-940; Barr et al. (1994), Gene Therapy 1:51-58; K.L. Berkner (1988), BioTechniques 6:616-629; und Rich et al. (1993), Human Gene Therapy 4:461-476). Da Menschen manchmal durch übliche humane Adenovirus-Serotypen wie AdHu5 infiziert werden, hat ein signifikanter Anteil der Population eine neutralisierende Antikörperreaktion auf das Adenovirus, die wahrscheinlich die Immunreaktion gegen ein heterologes Antigen in einem System auf rekombinanter Impfstoffbasis beeinflussen wird. Nicht-humane Primaten-Adenovirusvektoren wie das Schimpansen-Adenovirus 68 (AdC68, Fitzgerald et al. (2003), J. Immunol 170(3): 1416-22) mögen ein alternatives adenovirales System ohne den Nachteil einer vorbestehenden neutralisierenden Antikörperreaktion bieten.

[0099] Verschiedene Adeno-assoziierte Virus-(AAV)-Vektorsysteme wurden ebenfalls zur Polynukleotidübertragung entwickelt. AAV-Vektoren können leicht unter Verwendung von allgemein fachbekannten Techniken konstruiert werden. Siehe z.B. US-PSen 5,173,414 und 5,139,941: WO 92/01070 und WO 93/03769; Lebkowski et al. (1988), Molec. Cell. Biol. 8:3988-3996; Vincent et al. (1990), Vaccines 90 (Cold Spring Harbor Laboratory Press); B.J. Carter (1992), Current Opinion in Biotechnology 3:533-539; N. Muzyczka (1992), Current Topics in Microbiol. and Immunol. 158:97-129; R.M. Kotin (1994), Human Gene Therapy 5:793-801; Shelling und Smith (1994), Gene Therapy 1:165-169; und Zhou et al. (1994), J. Exp. Med. 179:1867-1875.

[0100] Zusätzliche virale Vektoren, die nützlich zur Übertragung der Nukleinsäuremoleküle, die die Polypeptide der vorliegenden Erfindung codieren, durch Gentransfer sind, schließen diejenigen ein, die aus der Pockenfamilie von Viren stammen, wie Vakziniavirus und Vogelpockenvirus. Beispielsweise können Vakziniavirus-Rekombinanten, die die neuen Moleküle extrahieren, wie folgt konstruiert werden. Die DNA, die ein Polypeptid codiert, wird zuerst in einen geeigneten Vektor inseriert, so daß sie benachbart zu einem Vakziniapromotor ist und Vakzinia-DNA-Sequenzen wie die Sequenz, die Thymidinkinase (TK) codiert, flankiert. Dieser Vektor wird dann zur Transfektion von Zellen verwendet, die gleichzeitig mit Vakzinia infiziert werden. Eine homologe Rekombination dient zur Insertion des Vakziniapromotors plus des Gens, das das Polypeptid von Interesse codiert, in das virale Genom. Die resultierende TK-Rekombinante kann durch Kultivieren der Zellen in Gegenwart von 5-Bromdesoxyuridin und Picken von dagegen resistenten viralen Plaques selektiert werden.

[0101] Ein Vakzinia-basiertes Infektions/Transfektionssystem kann zweckmäßig verwendet werden, um die induzierbare, transiente Expression oder Coexpression eines oder mehrerer der hier beschriebenen Polypeptide in Wirtszellen eines Organismus bereitzustellen. In diesem besonderen System werden Zellen zuerst in vitro mit einer Vakziniavirus-Rekombinante infiziert, die die Bakteriophage T7-RNA-Polymerase codiert. Diese Polymerase zeigt eine ausgezeichnete Spezifität, indem sie nur Matrizen transkribiert, die T7-Promotoren tragen. Nach der Infektion werden die Zellen mit dem Polynukleotid oder den Polynukleotiden von Interesse transfiziert, angetrieben durch einen T7-Promotor. Die im Cytoplasma aus der Vakziniavirus-Rekombinante exprimierte Polymerase transkribiert die transfizierter DNA zu RNA, die dann zu Polypeptid durch die Translationsmaschinerie des Wirtes translatiert wird. Das Verfahren liefert eine hochgradige, transiente, cytoplasmatische Produktion großer Mengen von RNA und ihrer Translationsprodukte. Siehe z.B. Elroy-Stein und Moss, Proc. Natl. Acad. Sci. USA (1990), 87:6743-6747; Fuerst et al., Proc. Natl. Acad. Sci. USA (1986) 83:8122-8126.

[0102] Alternativ können auch Vogelpockenviren wie die Geflügelpocken- und Kanarienvogelpockenviren zur Übertragung der codierenden Sequenzen von Interesse verwendet werden. Rekombinante Vogelpockenviren, die Immunogene aus Säugerpathogenen exprimieren, sind dafür bekannt, schützende Immunität zu verleihen, wenn sie an Nicht-Vogelarten verabreicht werden. Die Verwendung eines Vogelpockenvektors ("Avipox") ist besonders wünschenswert bei Menschen und anderen Säugerarten, da Mitglieder der Vogelpockengattung sich nur produktiv in anfälligen Vogelarten vermehren können und deshalb nicht infektiös in Säugerzellen sind. Verfahren zur Erzeugung von rekombinanten Vogelpockenviren sind fachbekannt und setzen die genetische Rekombination ein, wie oben in bezug auf die Produktion von Vakziniaviren beschrieben. Siehe z.B. WO 91/12882, WO 89/03429 und WO 92/03545.

[0103] Jeder einer Anzahl von Alphavirus-Vektoren kann ebenfalls zur Übertragung von Polynukleotidzusammensetzungen der vorliegenden Erfindung verwendet werden, wie zum Beispiel diejenigen Vektoren, die in US-PSen 5,843,723, 6,015,686, 6,008,035 und 6,015,694 beschrieben werden. Bestimmte Vektoren auf Basis der Venezolanischen Pferdeenzephalitis (VEE) können auch verwendet werden, für die illustrative Beispiele in US-PSen 5,505,947 und 5,643,576 gefunden werden können.

[0104] Die Zusammensetzungen der vorliegenden Erfindung können auch durch eine Anzahl von Wegen wie

intramuskulär, subkutan, intraperitoneal oder intravenös übertragen werden.

[0105] In einer anderen Ausführungsform der Erfindung wird ein Polynukleotid als "nackte" DNA verabreicht/übertragen, zum Beispiel wie beschrieben in Ulmer et al., Science 259:1745-1749, 1993 und zusammengefaßt von Cohen, Science 259:1691-1692, 1993. Die Aufnahme von nackter DNA kann durch Auftragen der DNA auf biologisch abbaubare Perlen gesteigert werden, die effizient in die Zellen transportiert werden. In einer bevorzugten Ausführungsform wird die Zusammensetzung intradermal übertragen. Insbesondere wird die Zusammensetzung mittels Verabreichungstechniken mit einer Genpistole (insbesondere Partikelbombardierung) übertragen, die die Auftragung des Vektors auf eine Perle (z.B. aus Gold) beinhalten, die dann unter Hochdruck in die Epidermis verabreicht wird, wie zum Beispiel beschrieben in Haynes et al., J. Biotechnology 44: 37-42 (1996).

[0106] In einem illustrativen Beispiel kann eine durch Gas angetriebene Partikelbeschleunigung mit Vorrichtungen wie denjenigen erreicht werden, die von Powderject Pharmaceuticals PLC (Oxford, UK) und Powderject Vaccines Inc. (Madison, WI) hergestellt werden, für die einige Beispiele beschrieben werden in US-PSen 5,846,796; 6,010,478; 5,865,796; 5,584,807; und in EP-PS 500 799. Dieser Ansatz bietet einen nadelfreien Übertragungsansatz, in dem eine trockene Pulverformulierung aus mikroskopischen Teilchen, wie zum Beispiel Polynukleotid, auf einer hohen Geschwindigkeit in einem Heliumgasstrom beschleunigt wird, der durch eine tragbare Vorrichtung erzeugt wird, wodurch die Partikel in ein Zielgewebe von Interesse, typischerweise die Haut, getrieben werden. Die Partikel sind vorzugsweise Goldperlen mit einem Durchmesser von 0,4-4,0 µm, besonders bevorzugt 0,6-2,0 µm, wobei das DNA-Konjugat darauf aufgetragen ist und diese dann in einer Kartusche oder Kassette zum Einlegen in die "Genpistole" umhüllt sind.

[0107] In einer verwandten Ausführungsform schließen andere Vorrichtungen und Verfahren, die nützlich zur gasgetriebenen nadelfreien Injektion von Zusammensetzungen der vorliegenden Erfindung sein können, diejenigen ein, die von Bioject, Inc. (Portland, OR) bereitgestellt werden, für die einige Beispiele beschrieben werden in US-PSen 4,790,824; 5,064,413; 5,312,335; 5,383,851; 5,399,163; 5,520,639 und 5,993,412.

[0108] Es ist möglich, die immunogene Komponente, die die Nukleotidsequenz umfaßt, die das antigene Peptid codiert, auf einmaliger Basis oder wiederholt zu verabreichen, zum Beispiel 1- bis 7-mal, bevorzugt 1- bis 4-mal, in Intervallen zwischen ca. 1 Tag und ca. 18 Monaten. Jedoch wird dieses Behandlungsschema in Abhängigkeit von der Größe des Patienten, der Krankheit, die behandelt wird/vor der geschützt wird, der verabreichten Menge der Nukleotidsequenz, dem Verabreichungsweg und anderen Faktoren signifikant variiert werden, die einem fachmännischen medizinischen Praktiker ersichtlich sein würden.

[0109] Es ist deshalb ein weiterer Aspekt der vorliegenden Erfindung, die Verwendung eines Proteins oder einer DNA, die das Protein codiert, wie hier beschrieben, in der Herstellung einer immunogenen Zusammensetzung zum Hervorrufen einer Immunreaktion in einem Patienten vorzusehen. Bevorzugt soll die Immunreaktion durch sequentielle Verabreichung i) des Proteins, gefolgt von der DNA-Sequenz; oder ii) der DNA-Sequenz, gefolgt vom Protein hervorgerufen werden. Besonders bevorzugt wird die DNA-Sequenz auf biologisch abbaubare Perlen aufgetragen oder über einen Partikelbombardierungsansatz übertragen. Noch mehr bevorzugt ist das Protein mit Hilfsstoff versetzt, bevorzugt mit einem TH-1-induzierenden Hilfsstoff, bevorzugt mit einer Hilfsstoffformulierung auf CpG/QS21-Basis.

[0110] Die Vektoren, die die Nukleotidsequenzen umfassen, die antigene Peptide codieren, werden in solchen Mengen verabreicht, wie sie prophylaktisch oder therapeutisch wirksam sein werden. Die zu verabreichende Menge ist allgemein im Bereich von 1 pg bis 16 mg, bevorzugt 1 pg bis 10 µg für die Partikel-vermittelte Übertragung und 10 µg bis 16 mg für andere Wege des Nukleotids pro Dosis. Die genaue Menge kann beträchtlich in Abhängigkeit vom Gewicht des zu immunisierenden Patienten und dem Verabreichungsweg variieren.

[0111] Geeignete Techniken zur Einführung des nackten Polynukleotids oder Vektors in einen Patienten schließen auch die topische Anwendung mit einem geeigneten Träger ein. Die Nukleinsäure kann topisch auf die Haut oder auf Schleimhautoberflächen durch zum Beispiel intranasale, orale, intravaginale oder intrarektale Verabreichung verabreicht werden. Das nackte Polynukleotid oder der nackte Vektor kann zusammen mit einem pharmazeutisch akzeptablen Exzipienten wie phosphatgepufferter Kochsalzlösung (PBS) vorhanden sein. Die DNA-Aufnahme kann weiter durch Verwendung von Erleichterungsmitteln wie Bupivacain, entweder separat oder eingeschlossen in der DNA-Formulierung, erleichtert werden. Andere Verfahren zur Verabreichung der Nukleinsäure direkt an einen Empfänger schließen Ultraschall, elektrische Stimulation, Elektroporation und Microseeding, das in US 5,697,901 beschrieben wird, ein.

[0112] Die Aufnahme von Nukleinsäurekonstrukten kann durch mehrere bekannte Transfektionstechniken gesteigert werden, zum Beispiel durch diejenigen, die die Verwendung von Transfektionsmitteln einschließen. Beispiele für diese Mittel schließen kationische Mittel, zum Beispiel Calciumphosphat und DEAE-Dextran, und Lipofectants, zum Beispiel Lipfectam und Transfectam, ein. Die Dosierung der zu verabreichenden Nukleinsäure kann verändert werden.

[0113] Die erfindungsgemäßen Fusionsproteine und codierenden Polypeptide können auch als pharmazeutische/immunogene Zusammensetzung formuliert werden, zum Beispiel als Impfstoff. Entsprechend stellt die vorliegende Erfindung deshalb auch eine pharmazeutische/immunogene Zusammensetzung bereit, die ein Fusionsprotein der vorliegenden Erfindung in einem pharmazeutisch akzeptablen Exzipienten umfaßt. Entsprechend wird auch ein Verfahren zur Herstellung einer erfindungsgemäßen immunogenen Zusammensetzung bereitgestellt, das das Vermischen des Fusionsproteins der Erfindung oder des codierenden Polynukleotids mit einem geeigneten Hilfsstoff, Verdünnungsmittel oder anderen pharmazeutisch akzeptablen Träger umfaßt.

[0114] Die erfindungsgemäßen Fusionsproteine werden bevorzugt wenigstens 80% rein, besonders bevorzugt 90% rein bereitgestellt, wie durch SDS-PAGE visualisiert. Bevorzugt erscheinen die Proteine als einzelne Bande gemäß SDS-PAGE.

[0115] Die Impfstoffherstellung wird allgemein beschrieben in "Vaccine Design" ("The Subunit and adjuvant approach" (Hrsg. M.F. Powell & M.J. Newman), (1995), Plenum Press New York). Die Verkapselung in Liposomen wird von Fullerton beschrieben, US-PS 4,235,877.

[0116] Die erfindungsgemäßen Fusionsproteine und codierenden Polynukleotide werden bevorzugt in der Impfstoffformulierung der Erfindung mit Hilfsstoff versetzt. Bestimmte Hilfsstoffe sind kommerziell erhältlich, wie zum Beispiel Freundsches unvollständiges Adjuvans und komplettes Adjuvans (Difco Laboratories, Detroit, MI); Merck Adjuvans 65 (Merck and Company, Inc. Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); Aluminiumsalze wie Aluminiumhydroxidgel (Alaun) oder Aluminiumphosphat; Salze von Calcium, Eisen oder Zink; eine unlösliche Suspension von acyliertem Tyrosin; acylierte Zucker; kationisch oder anionisch derivatisierte Polysaccharide; Polyphosphazene; biologisch abbaubare Mikrokügelchen; Monophosphoryllipid A und Quil A. Cytokine wie GM-CSF, Interleukin-2, -7, -12 und andere ähnliche Wachstumsfaktoren können auch als Hilfsstoffe verwendet werden.

[0117] Innerhalb bestimmter Ausführungsformen der Erfindung ist die Hilfsstoffzusammensetzung bevorzugt eine, die eine Immunreaktion vorherrschend von Th1-Typ induziert. Hohe Menge von Cytokinen vom Th1-Typ (z.B. IFN- γ , TNF α , IL-2 und IL-12) neigen dazu, die Induzierung von Zell-vermittelten Immunreaktionen gegen ein verabreichtes Antigen zu begünstigen. Im Gegensatz neigen hohe Mengen von Cytokinen vom Th2-Typ (z.B. IL-4, IL-5, IL-6 und IL-10) dazu, die Induktion von humoralen Immunreaktionen zu begünstigen. Nach Anwendung eines Impfstoffs wie hier bereitgestellt wird ein Patient eine Immunreaktion stützen, die Reaktionen vom Th1- und Th2-Typ einschließt. Innerhalb einer bevorzugten Ausführungsform, in der eine Reaktion vorherrschend vom Th1-Typ ist, wird die Menge der Cytokine vom Th1-Typ in einem stärkeren Ausmaß zunehmen als die Menge der Cytokine vom Th2-Typ. Die Mengen dieser Cytokine können leicht unter Verwendung von Standardtests bestimmt werden. Für eine Übersicht über die Familien von Cytokinen siehe Mosmann und Coffman, Ann. Rev. Immunol. 7:145-173, 1989.

[0118] Bevorzugte TH-1-induzierende Hilfsstoffe werden aus der Gruppe von Hilfsstoffen ausgewählt, die 3D-MPL, QS21, eine Mischung aus QS21 und Cholesterol und ein CpG-Oligonukleotid oder eine Mischung aus zwei oder mehr der Hilfsstoffe umfaßt. Bestimmte bevorzugte Hilfsstoffe zum Hervorrufen einer Reaktion vorherrschend vom Th1-Typ schließen zum Beispiel eine Kombination aus Monophosphoryllipid A, bevorzugt 3-des-O-acyliertem Monophosphoryllipid A, zusammen mit einem Aluminiumsalz ein. MPL[®]-Hilfsstoffe sind von Corixa Corporation erhältlich (Seattle, WA; siehe z.B. US-PSen 4,436,727; 4,877,611; 4,866,034 und 4,912,094). CpG-haltige Oligonukleotide (in denen das CpG-Dinukleotid unmethyliert ist) induzieren auch eine vorherrschend Th1-Reaktion. Solche Oligonukleotide sind wohlbekannt und werden zum Beispiel beschrieben in WO 96/02555, WO 99/33488 und US-PSen 6,008,200 und 5,856,462. Immunstimulatorische DNA-Sequenzen werden auch beschrieben, zum Beispiel von Sato et al., Science 273:352, 1996. Ein anderer bevorzugter Hilfsstoff umfaßt ein Saponin wie Quil A oder Derivate davon, einschließlich QS21 und QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA); Escin; Digitonin; oder Gypsophila- oder Chenopodium quinoa-Saponine. Andere bevorzugte Formulierungen schließen mehr als ein Saponin in den Hilfsstoffkombinationen der vorliegenden Erfindung ein, zum Beispiel Kombinationen aus wenigstens zweien der folgenden Gruppe, die QS21, QS7, Quil A, β -Escin oder Digitonin umfaßt.

[0119] Alternativ können die Saponin-Formulierungen mit Impfstoffträgern kombiniert werden, die aus Chitosan oder anderen polykationischen Polymeren, Polylactid- und Polylactid-co-Glycolid-Partikeln, Poly-N-acetylglucosamin-basierter Polymermatrix, Partikeln, die aus Polysacchariden oder chemisch modifizierten Polysacchariden zusammengesetzt sind, Liposomen und Lipid-basierten Partikeln, Partikeln, die aus Glycerinmonoestern zusammengesetzt sind, etc. zusammengesetzt sind. Die Saponine können auch in Gegenwart von Cholesterin zur Bildung von teilchenförmigen Strukturen wie Liposomen oder ISCOMs formuliert werden. Außerdem können die Saponine zusammen mit einem Polyoxyethylenether oder -ester in entweder einer nicht-teilchenförmigen Lösung oder Suspension oder in einer teilchenförmigen Struktur formuliert werden, wie zum Beispiel ein paucillamelares Liposom oder ISCOM. Die Saponine können auch mit Exzipienten wie Carbopol® zur Erhöhung der Viskosität formuliert werden oder können in einer Trockenpulverform mit einem Pulverexzipienten wie Lactose formuliert werden.

[0120] In einer bevorzugten Ausführungsform schließt das Hilfsstoffsystem die Kombination aus einem Monophosphoryllipid A und einem Saponinderivat ein, wie zum Beispiel die Kombination aus QS21 und 3D-MPL®-Hilfsstoff, wie beschrieben in WO 94/00153, oder eine weniger reaktogene Zusammensetzung, in der QS21 mit Cholesterin abgeschreckt ist, wie in WO 96/33739 beschrieben. Andere bevorzugte Formulierungen umfassen eine Öl-in-Wasser-Emulsion und Tocopherol. Eine andere, besonders bevorzugte Hilfsstoffformulierung, die QS21, 3D-MPL®-Hilfsstoff und Tocopherol in einer Öl-in-Wasser-Emulsion einsetzt, wird in WO 95/17210 beschrieben.

[0121] Ein anderes gesteigertes Hilfsstoffsystem beinhaltet die Kombination aus einem CpG-haltigen Oligonukleotid und einem Saponin-Derivat, insbesondere die Kombination aus CpG und QS21, wie in WO 00/09159 und in WO 00/62800 offenbart. Bevorzugt umfaßt die Formulierung zusätzlich eine Öl-in-Wasser-Emulsion und Tocopherol.

[0122] In noch einer weiteren Ausführungsform stellt die vorliegende Erfindung eine immunogene Zusammensetzung bereit, die ein erfindungsgemäßes Fusionsprotein umfaßt und ferner D3-MPL, ein Saponin, vorzugsweise QS21, und ein CpG-Oligonukleotid umfaßt, gegebenenfalls formuliert in einer Öl-in-Wasser-Emulsion.

[0123] Zusätzliche illustrative Hilfsstoffe zur Verwendung in den pharmazeutischen Zusammensetzungen der Erfindung schließen Montanid ISA 720 (Seppic, Frankreich), SAF (Chiron, CA, Vereinigte Staaten), ISCOMS (CSL), MF-59 (Chiron), die SBAS-Reihe von Hilfsstoffen (z.B. SBAS-2 oder SBAS-4, erhältlich von SmithKline Beecham, Rixensart, Belgien), Detox (Enhancyn®) (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) und andere Aminoalkylglucosaminid-4-phosphate (AGPs), wie die in den anhängigen US-Patentanmeldungen Nrn. 08/853,826 und 09/074,720 beschriebenen, deren Offenbarungen hier durch Verweis in ihrer Gesamtheit eingeführt wird, und Polyoxyethylenether-Hilfsstoffe, wie die in WO 99/52549A1 beschriebenen, ein.

[0124] Andere bevorzugte Hilfsstoffe schließen Hilfsstoffmoleküle der allgemeinen Formel (I) ein: $\text{HO}(\text{CH}_2\text{CH}_2\text{O})_n\text{-A-R}$, worin n 1-50 ist, A eine Bindung oder -C(O)- ist, R C₁₋₅₀-Alkyl oder Phenyl-C₁₋₅₀-alkyl ist. Eine Ausführungsform der vorliegenden Erfindung besteht aus einer Impfstoffformulierung, die einen Polyoxyethylenether der allgemeinen Formel (I) umfaßt, worin n zwischen 1 und 50 ist, bevorzugt 4-24, am meisten bevorzugt 9; die R-Komponente C₁₋₅₀ ist, bevorzugt C₄₋₂₀-Alkyl und am meisten bevorzugt C₁₂-Alkyl, und A eine Bindung ist. Die Konzentration der Polyoxyethylenether sollte im Bereich von 0,1-20% sein, bevorzugt von 0,1-10% und am meisten bevorzugt im Bereich von 0,1-1%. Bevorzugte Polyoxyethylenether werden aus der folgenden Gruppe ausgewählt: Polyoxyethylen-9-laurylether, Polyoxyethylen-9-stearylether, Polyoxyethylen-8-stearylether, Polyoxyethylen-4-laurylether, Polyoxyethylen-35-laurylether und Polyoxyethylen-23-laurylether. Polyoxyethylenether wie Polyoxyethylenlaurylether werden im Merck-Index beschrieben (12. Auflage, Eintrag 7717). Diese Hilfsstoffmoleküle werden in WO 99/52549 beschrieben. Der Polyoxyethylenether gemäß der obigen allgemeinen Formel (I) kann nach Wunsch mit einem anderen Hilfsstoff kombiniert werden. Zum Beispiel ist eine bevorzugte Hilfsstoffkombination bevorzugt mit CpG, wie in der anhängigen GB-Patentanmeldung 9820956.2 beschrieben.

[0125] Es ist eine Ausführungsform der Erfindung, daß die Antigene, einschließlich Nukleinsäurevektor, der Erfindung mit immunstimulatorischem Mittel verwendet werden. Bevorzugt wird das immunstimulatorische Mittel zur gleichen Zeit wie die Antigene der Erfindung verabreicht, und in bevorzugten Ausführungsformen werden sie zusammen formuliert. Es ist eine andere Ausführungsform der Erfindung, daß das Antigen und immunstimulatorische Mittel (oder umgekehrt) sequentiell an gleichen oder benachbarten Orten verabreicht werden, zeitlich getrennt durch Zeiträumen von 0-100 Stunden. Solche immunstimulatorischen Mittel schließen ohne Beschränkung ein: synthetische Imidazochinoline wie Imiquimod [S-26308, R-837] (Harrison et al., Vaccine 19:

1820-1826, 2001); und Resiquimod [S-28463, R-848] (Vasilakos et al., Cellular Immunology 204: 64-74, 2000); Schiffische Basen von Carbonylen und Aminen, die konstitutiv auf Antigen-präsentierenden Zell- und T-Zelloberflächen exprimiert werden, wie Tucaresol (J. Rhodes et al., Nature 377: 71-75, 1995), Cytokin-, Chemokin- und costimulatorische Moleküle als entweder Protein oder Peptid, einschließlich beispielsweise proinflammatorischer Cytokine wie Interferon, GM-CSF, IL-1-alpha, IL-1-beta, TGF-alpha und TGF-beta, Th1-Induktoren wie Interferon-gamma, IL-2, IL-12, IL-15, IL-18 und IL-21, Th2-Induktoren wie IL-4, IL-5, IL-6, IL-10 und IL-13, und andere Chemokin- und costimulatorische Gene wie MCP-1, MIP-1-alpha, MIP-1-beta, RANTES, TCA-3, CD80, CD86 und CD40L, andere immunstimulatorische Zielliganden wie CTLA-4 und L-Selectin, Apoptosestimulierende Proteine und Peptide wie Fas, (49), synthetische Lipidbasierte Hilfsstoffe wie Vaxfectin (Reges et al., Vaccine 19: 3778-3786, 2001), Squalen, alpha-Tocopherol, Polysorbat 80, DOPC und Cholesterol, Endotoxin, [LPS], (B. Beutler, Current Opinion in Microbiology 3: 23-30, 2000); CpG-Oligo- und Dinukleotide (Y. Sato et al., Science 273 (5273): 352-354, 1996; H. Hemmi et al., Nature 408: 740-745, 2000) und andere potentielle Liganden, die Toll-Rezeptoren zur Erzeugung von Th1-induzierenden Cytokinen anregen, wie synthetische mykobakterielle Lipoproteine, mykobakterielles Protein p19, Peptidoglycan, Teichonsäure und Lipid A.

[0126] Andere geeignete Hilfsstoffe schließen CT (Choleratoxin, Untereinheiten A und B) und LT (hitzelabiles Enterotoxin aus *E. coli*, Untereinheiten A und B), Hitzeschockprotein-Familie (HSPs) und LLO (Listeriolysin O; WO 01/72329) ein.

[0127] Wenn das immunstimulatorische Mittel ein Protein ist, kann das Mittel entweder als Protein oder als Polynukleotid, das das Protein codiert, verabreicht werden.

[0128] Andere geeignete Übertragungssysteme schließen Mikrokügelchen ein, in denen das antigene Material in biologisch abbaubaren Polymeren/Mikrokügelchen aufgenommen oder damit konjugiert ist, so daß das antigene Material mit einem geeigneten pharmazeutischen Träger vermischt und als Impfstoff verwendet werden kann. Der Begriff "Mikrokügelchen" wird allgemein eingesetzt, um kolloidale Partikel zu beschreiben, die im wesentlichen kugelförmig sind und einen Durchmesser im Bereich von 10 nm bis 2 µm haben. Mikrokügelchen, die aus einem sehr großen Bereich von natürlichen und synthetischen Polymeren hergestellt werden, haben Anwendung in einer Vielzahl biomedizinischer Anwendungen gefunden. Dieses Übertragungssystem ist besonders vorteilhaft für Proteine mit kurzen Halbwertszeiten in vivo, was multiple Behandlungen zur Bereitstellung von Wirksamkeit erfordert, oder die instabil in biologischen Flüssigkeiten sind oder nicht vollständig im Magen-Darm-Trakt wegen ihrer relativ hohen Molekulargewichte absorbiert werden. Mehrere Polymere wurden als Matrix zur Proteinfreisetzung beschrieben. Geeignete Polymere schließen Gelatine, Collagen, Alginate und Dextran ein. Bevorzugte Übertragungssysteme schließen biologisch abbaubare Poly(DL-milchsäure) (PLA), Poly(lactid-co-glycolid) (PLG), Poly(glykolsäure) (PGA), Poly(ε-caprolacton) (PCL) und Copolymere wie Poly(DL-milchsäure-co-glykolsäure) (PLGA) ein. Andere bevorzugte Systeme schließen heterogene Hydrogele wie Poly(etherester)-Multiblock-Copolymere ein, die Strukturblöcke auf Basis von hydrophilem Poly(ethylenglykol) PEG und hydrophobem Poly(butylenterephthalat) (PBT) enthalten, oder Poly(ethylenglykol)terephthalat/Poly(butylenterephthalat) (PEGT/PBT) (Sohier et al., Eur. J. Pharm. and Biopharm., 2003, 55, 221-228). Systeme sind bevorzugt, die eine anhaltende Freisetzung für 1 bis 3 Monate bereitstellen, wie PLGA, PLA und PEGT/PBT.

[0129] Es ist möglich, die immunogene oder Impfstoffzusammensetzung auf einmaliger Basis zu verabreichen oder vorzugsweise wiederholt zu verabreichen, so häufig wie notwendig, zum Beispiel 1- bis 7-mal, bevorzugt 1- bis 4-mal, in Intervallen zwischen ca. 1 Tag und ca. 18 Monaten, bevorzugt einem Monat. Diesem kann sich gegebenenfalls eine Dosierung in regelmäßigen Intervallen zwischen 1 und 12 Monaten für einen Zeitraum bis zum Rest des Lebens des Patienten anschließen. In einer bevorzugten Ausführungsform erhält der Patient das Antigen in unterschiedlichen Formen in einem "Prime Boost"-Schema. So wird beispielsweise das Antigen, das Fusionsprotein, zuerst als Protein-Hilfsstoff-basierte Formulierung und dann anschließend als DNA-basierter Impfstoff verabreicht. Dieser Verabreichungsmodus ist bevorzugt. Der bevorzugte Hilfsstoff ist eine Kombination aus einem CpG-haltigen Oligonukleotid und einem Saponin-Derivat, insbesondere die Kombination aus CpG und QS21 wie in WO 00/09159 und in WO 00/62800 offenbart. Die Aufnahme von nackter DNA kann durch Auftragen der DNA auf biologisch abbaubare Perlen erhöht werden, die effizient in die Zellen transportiert werden. Alternativ kann die DNA über einen Partikelbombardierungsansatz übertragen werden, zum Beispiel durch gasgetriebene Partikelbeschleunigung mit Vorrichtungen wie denjenigen, die von Powderject Pharmaceuticals PLC (Oxford, UK) und Powderject Vaccines Inc. (Madison, WI) hergestellt werden, wie hier gelehrt. Dieser Ansatz liefert einen nadelfreien Übertragungsansatz, in dem eine trockene Pulverformulierung aus mikroskopischen Partikeln, wie Polynukleotid- oder Polypeptidpartikeln, auf eine hohe Geschwindigkeit in einem Heliumgasstrom beschleunigt wird, der durch eine tragbare Vorrichtung erzeugt wird, wodurch die Partikel in ein Zielgewebe von Interesse getrieben werden.

[0130] In einer anderen bevorzugten Ausführungsform wird der DNA-basierte Impfstoff zuerst verabreicht werden, gefolgt von der Protein-Hilfsstoffbasierten Formulierung. Noch eine andere Ausführungsform wird die Übertragung des DNA-Konstrukts mittels spezialisierter Übertragungsvektoren betreffen, bevorzugt mittels eines viralen Systems, am meisten bevorzugt mittels Adenovirus-basierter Systeme. Andere geeignete Virus-basierte Systeme der DNA-Übertragung schließen Systeme auf retroviraler, lentiviraler, adeno-assoziiertes viraler, Herpes-viraler und Vakziniaviraler Basis ein.

[0131] In einer anderen bevorzugten Ausführungsform können die Protein-Hilfsstoff-Basisformulierung und der DNA-basierte Impfstoff an benachbarten oder überlappenden Orten verabreicht werden. Abhängig von der Natur der DNA-Impfstoffformulierung kann dies durch Vermischen der DNA- und Protein-Hilfsstoffformulierungen vor der Verabreichung oder durch gleichzeitige Verabreichung der DNA- und Protein-Hilfsstoffformulierung erreicht werden.

[0132] Das Behandlungsschema wird signifikant in Abhängigkeit von der Größe und Art des betreffenden Patienten, der verabreichten Menge des Nukleinsäureimpfstoff und/oder der Proteinzusammensetzung, dem Verabreichungsweg, der Wirksamkeit und Dosis etwaiger als Hilfsstoff verwendeter Verbindungen und anderen Faktoren variiert werden, die einem fähigen medizinischen Praktiker ersichtlich sein würden.

[0133] Innerhalb weiterer Aspekte stellt die vorliegende Erfindung Verfahren zur Stimulation einer Immunreaktion in einem Patienten bereit, bevorzugt einer T-Zellreaktion in einem menschlichen Patienten, umfassend das Verabreichen einer hier beschriebenen pharmazeutischen Zusammensetzung. Der Patient kann von Lungen- oder Dickdarmkrebs oder kolorektalem Krebs oder Brustkrebs befallen sein, wobei in diesem Fall die Verfahren die Behandlung für die Krankheit bereitstellen, oder der Patient, der als mit einem Risiko für eine solche Krankheit behaftet betrachtet wird, kann prophylaktisch behandelt werden.

[0134] Innerhalb weiterer Aspekte stellt die vorliegende Erfindung Verfahren zur Inhibierung der Entwicklung eines Krebses in einem Patienten bereit, umfassend das Verabreichen einer pharmazeutischen Zusammensetzung wie oben aufgeführt an einen Patienten. Der Patient kann zum Beispiel von Sarkom-, Prostata-, Eierstock-, Blasen-, Lungen-, Dickdarm-, kolorektalem oder Brustkrebs befallen sein, wobei in diesen Fällen die Verfahren eine Behandlung für die Krankheit bereitstellen, oder der Patient, der als mit einem Risiko für eine solche Krankheit behaftet betrachtet wird, kann prophylaktisch behandelt werden.

[0135] Die vorliegende Erfindung stellt ferner innerhalb anderer Aspekte Verfahren zur Entfernung von Tumorzellen aus einer biologischen Probe bereit, umfassend das Inkontaktbringen einer biologischen Probe mit T-Zellen, die spezifisch mit einem Polypeptid der vorliegenden Erfindung reagieren, worin der Schritt des Inkontaktbringens unter Bedingungen und für eine ausreichende Dauer durchgeführt wird, um die Entfernung von Zellen, die das Protein exprimieren, aus der Probe zu erlauben.

[0136] Innerhalb verwandter Aspekte werden Verfahren zur Inhibierung der Entwicklung eines Krebses in einem Patienten bereitgestellt, umfassend das Verabreichen einer biologischen Probe, die wie oben beschrieben behandelt wurde, an einen Patienten.

[0137] Verfahren werden ferner bereitgestellt, innerhalb anderer Aspekte, zur Stimulation und/oder Vermehrung von T-Zellen, die spezifisch für ein Polypeptid der vorliegenden Erfindung sind, umfassend das Inkontaktbringen von T-Zellen mit einem oder mehreren aus: (i) einem Polypeptid wie oben beschrieben; (ii) einem Polynukleotid, das ein solches Polypeptid codiert; und/oder (iii) einer Antigen-präsentierenden Zelle, die ein solches Polypeptid exprimiert; unter Bedingungen und für eine Dauer, die ausreichend sind, um die Stimulation und/oder Vermehrung von T-Zellen zu erlauben. Isolierte T-Zellpopulationen, die T-Zellen wie oben beschrieben hergestellt umfassen, werden ebenfalls bereitgestellt.

[0138] Innerhalb weiterer Aspekte stellt die vorliegende Erfindung Verfahren zur Inhibierung der Entwicklung eines Krebses in einem Patienten bereit, umfassend das Verabreichen einer wirksamen Menge einer T-Zellpopulation wie oben beschrieben an einen Patienten. Die vorliegende Erfindung stellt ferner Verfahren zur Inhibierung der Entwicklung eines Krebses in einem Patienten bereit, umfassend die folgenden Schritte: (a) Inkubieren von CD4+ und/oder CD8+ T-Zellen, die aus einem Patienten mit einem oder mehreren aus den folgenden isoliert sind: (i) einem hier offenbarten Polypeptid; (ii) einem Polynukleotid, das ein solches Polypeptid codiert; und (iii) einer Antigen-präsentierenden Zelle, die ein solches Polypeptid exprimiert; und (b) Verabreichen einer wirksamen Menge der vermehrten T-Zellen an den Patienten und dadurch Inhibieren der Entwicklung eines Krebses im Patienten. Vermehrte Zellen können vor der Verabreichung an den Patienten kloniert werden, müssen aber nicht.

[0139] Gemäß einer anderen Ausführungsform dieser Erfindung wird eine hier beschriebene immunogene Zusammensetzung an einen Wirt über ein Antigenpräsentierende Zellen (APCs) übertragen, wie dendritische Zellen, Makrophagen, B-Zellen, Monozyten und andere Zellen, die manipuliert sein können, um effiziente APCs zu sein. Solche Zellen können genetisch modifiziert sein, aber brauchen nicht, um die Kapazität zur Präsentation des Antigens zu erhöhen, die Aktivierung und/oder Bewahrung der T-Zellreaktion zu verbessern, Antitumorwirkungen als solche aufzuweisen und/oder immunologisch mit dem Empfänger kompatibel zu sein (d.h. ein übereinstimmender HLA-Haplotyp). APCs können allgemein aus beliebigen aus einer Vielzahl biologischer Flüssigkeiten und Organe isoliert werden, die Tumor- und peritumorale Gewebe einschließen, und können autologe, allogene, syngene oder xenogene Zellen sein.

[0140] Bestimmte bevorzugte Ausführungsformen der vorliegenden Erfindung verwenden dendritische Zellen oder Vorläufer davon als Antigen-präsentierende Zellen. Dendritische Zellen sind hochpotente APCs (Banche-reau und Steinman, *Nature*, 392:245-251, 1998), und es wurde gezeigt, daß sie wirksam als physiologischer Hilfsstoff zur Hervorrufung von prophylaktischer oder therapeutischer Antitumorimmunität sind (siehe Timmerman und Levy, *Ann. Rev. Med.* 50:507-529, 1999). Allgemein können dendritische Zellen auf Basis ihrer typischen Form (sternförmig in situ, mit deutlichen cytoplasmatischen Prozessen (Dendriten), die in vitro sichtbar sind), ihrer Fähigkeit, Antigene mit hoher Effizienz aufzunehmen, zu verarbeiten und zu präsentieren, und ihrer Fähigkeit, naive T-Zellreaktionen zu aktivieren, identifiziert werden. Dendritische Zellen können natürlich manipuliert werden, um spezifische Zelloberflächenrezeptoren oder Liganden zu exprimieren, die nicht üblicherweise auf dendritischen Zellen in vivo oder ex vivo gefunden werden, und solche modifizierten dendritischen Zellen werden von der vorliegenden Erfindung erwogen. Als Alternative zu dendritischen Zellen können mit sezernierten Vesikeln Antigen-beladene dendritische Zellen (als Exosome bezeichnet) in einem Impfstoff verwendet werden (siehe Zitvogel et al., *Nature Med.* 4:594-600, 1998).

[0141] Dendritische Zellen und Vorläufer können aus peripherem Blut, Knochenmark, Tumor-infiltrierenden Zellen, peritumorales Gewebe infiltrierenden Zellen, Lymphknoten, Milz, Haut, Nabelschnurblut oder jedem anderen geeigneten Gewebe oder jeder anderen geeigneten Flüssigkeit erhalten werden. Zum Beispiel können dendritische Zellen ex vivo durch Zugabe einer Kombination von Cytokinen wie GM-CSF, IL-4, IL-13 und/oder TNF α zu Kulturen von Monozyten differenziert werden, die aus peripherem Blut geerntet wurden. Alternativ können CD34-positive Zellen, aus aus peripherem Blut, Nabelschnurblut oder Knochenmark geerntet wurden, zu dendritischen Zellen durch Zugabe von Kombinationen aus GM-CSF, IL3, TNF α , CD40-Ligand, LPS, flt3-Ligand und/oder anderen Verbindungen, die die Differenzierung, Reifung und Vermehrung von dendritischen Zellen induzieren, zum Kulturmedium differenziert werden.

[0142] Dendritische Zellen werden herkömmlich als "unreife" und "reife" Zellen kategorisiert, was eine einfache Möglichkeit zur Unterscheidung zwischen zwei gut charakterisierten Phänotypen erlaubt. Jedoch sollte diese Nomenklatur nicht als Ausschluß aller möglichen intermediären Differenzierungsstadien aufgefaßt werden. Unreife dendritische Zellen werden als APC mit einer hohen Fähigkeit zur Antigenaufnahme und -verarbeitung gekennzeichnet, was mit der hohen Expression von FC γ -Rezeptor und Mannoserezeptor korreliert. Der reife Phänotyp ist typischerweise durch eine geringere Expression dieser Marker, aber eine hohe Expression von Zelloberflächenmolekülen gekennzeichnet, die für die T-Zellaktivierung verantwortlich sind, wie Klasse I und Klasse II MHC, Adhäsionsmoleküle (z.B. CD54 und CD11) und costimulatorische Molekül (z.B. CD40, CD80, CD86 und 4-1BB).

[0143] APCs können allgemein mit einem Polynukleotid der Erfindung (oder einem Teil oder einer anderen Variante davon) transfiziert werden, so daß das codierte Polypeptid oder ein immunogener Teil davon auf der Zelloberfläche exprimiert wird. Eine solche Transfektion kann ex vivo erfolgen, und eine pharmazeutische Zusammensetzung, die solche transfizierten Zellen umfaßt, kann dann für therapeutische Zwecke wie hier beschrieben verwendet werden. Alternativ kann ein Genübertragungsträger, der auf eine dendritische oder andere Antigen-präsentierende Zelle abzielt, an einen Patienten verabreicht werden, was zu Transfektion führt, die in vivo erfolgt. In vivo- und ex vivo-Transfektion von dendritischen Zellen kann zum Beispiel allgemein unter Verwendung aller fachbekannten Verfahren durchgeführt werden, wie zum Beispiel mit denjenigen, die in WO 97/24447 beschrieben werden, oder mit dem Genpistolenansatz, beschrieben von Mahvi et al., *Immunology and Cell Biology* 75:456-460, 1997. Die Antigenbeladung von dendritischen Zellen kann durch Inkubieren von dendritischen Zellen oder Vorläuferzellen mit dem Tumorpolypeptid, DNA (nackt oder mit einem Plasmidvektor) oder RNA erreicht werden, oder mit Antigen-exprimierenden rekombinanten Bakterien oder Viren (z.B. Vakzina-, Geflügelpocken-, Adenovirus- oder Lentivirusvektoren). Vor der Beladung kann das Polypeptid kovalent mit einem immunologischen Partner konjugiert werden, der T-Zellhilfe liefert (z.B. ein Trägermolekül). Alternativ kann eine dendritische Zelle mit einem nicht-konjugierten immunologischen Partner gepulst werden, separat oder in Gegenwart des Polypeptids.

Definitionen

[0144] Ebenfalls bereitgestellt durch die Erfindung werden Verfahren zur Analyse von Kennzeichensequenzen oder Ketten, insbesondere genetischen Sequenzen oder codierten Proteinsequenzen. Bevorzugte Verfahren zur Sequenzanalyse schließen zum Beispiel Verfahren der Sequenzhomologieanalyse ein, wie Identitäts- und Ähnlichkeitsanalyse, DNA-, RNA- und Proteinstrukturanalyse, Sequenzaufbau, kladistische Analyse, Sequenzmotivanalyse, Bestimmung des offenen Leserasters, Nukleinsäure-Basenzählung, Codon-Verwendungsanalyse, Nukleinsäure-Basenzurechtstzung und sequenzierende Chromatogrammepeakanalyse ein.

[0145] Ein computerbasiertes Verfahren wird zur Durchführung der Homologieidentifikation bereitgestellt. Dieses Verfahren umfaßt die folgenden Schritte: Bereitstellen einer ersten Polynukleotidsequenz, die die Sequenz eines Polynukleotids der Erfindung in einem computerlesbaren Medium umfaßt; und Vergleichen der ersten Polynukleotidsequenz mit wenigstens einer zweiten Polynukleotid- oder Polypeptidsequenz zur Identifizierung von Homologie. Ein computerbasiertes Verfahren wird ebenfalls zur Durchführung der Homologieidentifikation bereitgestellt, wobei das Verfahren die folgenden Schritte umfaßt: Bereitstellen einer ersten Polypeptidsequenz, die die Sequenz eines Polypeptids der Erfindung umfaßt, in einem computerlesbaren Medium; und Vergleichen der ersten Polypeptidsequenz mit wenigstens einer zweiten Polynukleotid- oder Polypeptidsequenz zur Identifizierung von Homologie.

[0146] "Identität", wie es fachbekannt ist, ist eine Beziehung zwischen zwei oder mehr Polypeptidsequenzen oder zwei oder mehr Polynukleotidsequenzen, je nach Fall, gemäß Bestimmung durch Vergleich der Sequenzen. Auf diesem Gebiet bedeutet "Identität" auch den Grad der Sequenzverwandtheit zwischen Polypeptid- oder Polynukleotidsequenzen, je nach Fall, gemäß Bestimmung durch die Übereinstimmung zwischen Ketten solcher Sequenzen. "Identität" kann leicht durch bekannte Verfahren berechnet werden, die ohne Beschränkung diejenigen einschließen, die beschrieben werden in Computational Molecular Biology, A.M. Lesk, Hrsg., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, D.W. Smith, Hrsg., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Teil 1, A.M. Griffin und H.G. Griffin, Hrsg., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, G. von Heine, Academic Press, 1987; und Sequence Analysis Primer, M. Gribskov und J. Devereux, Hrsg., M. Stockton Press, New York, 1991; und H. Carillo und D. Lipman, SIAM J. Applied Math. 48: 173 (1988). Verfahren zur Bestimmung von Identität werden geschaffen, um die größte Übereinstimmung zwischen den untersuchten Sequenzen zu ergeben. Außerdem sind Verfahren zur Identitätsbestimmung in öffentlich verfügbaren Computerprogrammen codifiziert. Computerprogrammverfahren zur Bestimmung von Identität zwischen zwei Sequenzen schließen ohne Beschränkung das GAP-Programm im GCG-Programmpaket (J. Devereux et al., Nucleic Acids Research 12(1): 387 (1984)), BLASTP, BLASTN (S.F. Altschul et al., J. Molec. Biol. 215: 403-410 (1990)) und FASIA (Pearson und Lipman, Proc. Natl. Acad. Sci. USA 85: 2444-2448 (1988)) ein. Die BLAST-Familie von Programmen ist öffentlich erhältlich von NCBI und anderen Quellen (BLAST Manual, S. Altschul et al., NCBI NLM NIH Bethesda, MD 20894; S. Altschul et al., J. Mol. Biol. 215: 403-410 (1990)). Der allgemein bekannte Smith-Waterman-Algorithmus kann auch zur Identitätsbestimmung verwendet werden.

[0147] Parameter für den Polypeptid-Sequenzvergleich schließen die folgenden ein:

Algorithmus: Needleman und Wunsch, J. Mol. Biol. 48: 443-453 (1970),

Vergleichsmatrix: BLOSSUM62 von Henikoff und Henikoff, Proc. Natl.

Acad. Sci. USA, 89:10915-10919 (1992)

Lückenstrafe: 8

Lückenlängenstrafe: 2

[0148] Ein für diese Parameter nützliches Programm ist öffentlich erhältlich als das "gap"-Programm von Genetics Computer Group, Madison, WI. Die zuvor genannten Parameter sind die Standardparameter für Peptidvergleiche (neben keiner Strafe für die Endlücken).

[0149] Parameter für den Polynukleotidvergleich schließen die folgenden ein:

Algorithmus: Needleman und Wunsch, J. Mol. Biol. 48: 443-453 (1970),

Vergleichsmatrix: Übereinstimmungen = +10, Fehlübereinstimmungen = 0,

Lückenstrafe: 50

Lückenlängenstrafe: 3

[0150] Verfügbar als das "gap"-Programm von Genetics Computer Group, Madison, WI. Dies sind die Standardparameter für Nukleinsäurevergleiche.

[0151] Eine bevorzugte Bedeutung für "Identität" für Polynukleotide und Polypeptide, je nach Fall, wird nachfolgend unter (1) und (2) bereitgestellt.

(1) Polynukleotidausführungsformen schließen ferner ein isoliertes Polynukleotid ein, das eine Polynukleotidsequenz mit wenigstens einer 50, 60, 70, 80, 85, 90, 95, 97 oder 100%igen Identität mit jeder der Referenz-Sequenzen aus SEQ ID NO:9 bis SEQ ID NO:16 umfaßt, worin die Polynukleotidsequenz identisch mit jeder der Referenzsequenzen aus SEQ ID NO:9 bis SEQ ID NO: 16 sein kann oder bis zu einer bestimmten ganzen Anzahl von Nukleotidveränderungen im Vergleich zur Referenzsequenz einschließen kann, worin die Veränderungen aus der Gruppe ausgewählt sind, die aus wenigstens einer Nukleotiddeletion, -substitution, einschließlich Transition und Transversion, oder -insertion besteht, und worin die Veränderungen an den terminalen 5'- oder 3'-Positionen der Referenznukleotidsequenz oder an beliebiger Stelle zwischen diesen terminalen Positionen erfolgen können, eingestreut entweder individuell unter den Nukleotiden in der Referenzsequenz oder in einer oder mehreren zusammenhängenden Gruppen innerhalb der Referenzsequenz, und worin die Anzahl an Nukleotidveränderungen durch Multiplizieren der Gesamtanzahl von Nukleotiden in einem beliebigen aus SEQ ID NO:9 bis SEQ ID NO:16 durch die ganze Zahl, die die prozentuale Identität definiert, geteilt durch 100, und anschließendes Subtrahieren des Produkts von der Gesamtanzahl von Nukleotiden in einem beliebigen aus SEQ ID NO:9 bis SEQ ID NO:16 bestimmt wird, oder:

$$n_n \leq x_n - (x_n \cdot y),$$

worin n_n die Anzahl der Nukleotidveränderungen ist, x_n die Gesamtanzahl an Nukleotiden in einem beliebigen aus SEQ ID NO:9 bis SEQ ID NO:16 ist, y 0,50 für 50%, 0,60 für 60%, 0,70 für 70%, 0,80 für 80%, 0,85 für 85%, 0,90 für 90%, 0,95 für 95%, 0,97 für 97% oder 1,00 für 100% ist und $\cdot$ das Symbol für den Multiplikationsoperator ist, und worin jedes nicht ganzzahlige Produkt von x_n und y zur nächsten ganzen Zahl vor dem Subtrahieren von x_n abgerundet wird. Veränderungen der Polynukleotidsequenz, die die Polypeptide aus einer beliebigen aus SEQ ID NO:1 bis SEQ ID NO:8 codieren, können Unsinn-, Fehlsinn- oder Rasterverschiebungsmutationen in dieser codierenden Sequenz erzeugen und dadurch das durch das Polynukleotid nach solchen Veränderungen codierte Polypeptid verändern.

Beispielsweise kann eine Polynukleotidsequenz der vorliegenden Erfindung identisch mit jeder der Referenzsequenzen aus SEQ ID NO:9 bis SEQ ID NO:16 sein, d.h. sie kann zu 100% identisch sein, oder sie kann bis zu einer bestimmten ganzen Anzahl von Nukleinsäureveränderungen im Vergleich mit der Referenzsequenz einschließen, so daß die prozentuale Identität weniger als 100% Identität ist. Solche Veränderungen werden aus der Gruppe ausgewählt, die aus wenigstens einer Nukleinsäuredeletion, -substitution, einschließlich Transition und Transversion, oder -insertion besteht, und worin die Veränderungen an den terminalen 5'- oder 3'-Positionen der Referenzpolynukleotidsequenz oder an beliebiger Stelle zwischen diesen terminalen Positionen erfolgen können, eingestreut entweder individuell unter den Nukleinsäuren in der Referenzsequenz oder in einer oder mehreren zusammenhängenden Gruppen innerhalb der Referenzsequenz. Die Anzahl der Nukleinsäureveränderungen für eine gegebene prozentuale Identität wird durch Multiplizieren der Gesamtanzahl von Nukleinsäuren in einem beliebigen aus SEQ ID NO:9 bis SEQ ID NO:16 durch die ganze Zahl, die die prozentuale Identität definiert, geteilt durch 100, und anschließendes Subtrahieren des Produkts von der Gesamtanzahl von Nukleinsäuren in einem beliebigen aus SEQ ID NO:9 bis SEQ ID NO:16 bestimmt, oder:

$$n_n \leq x_n - (x_n \cdot y),$$

worin n_n die Anzahl der Nukleinsäureveränderungen ist, x_n die Gesamtanzahl an Nukleinsäuren in einem beliebigen aus SEQ ID NO:9 bis SEQ ID NO:16 ist, y z.B. 0,70 für 70%, 0,80 für 80%, 0,85 für 85% etc. ist, das Symbol für den Multiplikationsoperator ist, und worin jedes nicht ganzzahlige Produkt von x_n und y zur nächsten ganzen Zahl vor dem Subtrahieren von x_n abgerundet wird.

(2) Polypeptidausführungsformen schließen ferner ein isoliertes Polypeptid ein, das ein Polypeptid mit wenigstens einer 50, 60, 70, 80, 85, 90, 95, 97 oder 100%igen Identität mit der Polypeptidreferenzsequenz aus einem beliebigen aus SEQ ID NO:1 bis SEQ ID NO:8 umfaßt, worin die Polypeptidsequenz identisch mit einer beliebigen der Referenzsequenzen aus SEQ ID NO:1 bis SEQ ID NO:8 sein kann oder bis zu einer bestimmten ganzen Anzahl von Aminosäureveränderungen im Vergleich mit der Referenzsequenz einschließen kann, worin die Veränderungen aus der Gruppe ausgewählt sind, die aus wenigstens einer Aminosäuredeletion, -substitution, einschließlich konservativen und nicht-konservativen Substitution, oder -insertion besteht, und worin die Veränderungen an den Amino- oder Carboxy-terminalen Positionen der Referenzpolypeptidsequenz oder an beliebiger Stelle zwischen diesen terminalen Positionen erfolgen kann, eingestreut entweder individuell unter den Aminosäuren in der Referenzsequenz oder in einer oder mehreren zusammenhängenden Gruppen innerhalb der Referenzsequenz, und worin die Anzahl der Aminosäu-

Veränderungen durch Multiplizieren der Gesamtanzahl von Aminosäuren in einem beliebigen aus SEQ ID NO:1 bis SEQ ID NO:8 durch die ganze Zahl, die die prozentuale Identität definiert, geteilt durch 100, und anschließendes Subtrahieren des Produkts von der Gesamtanzahl von Aminosäuren in einem beliebigen aus SEQ ID NO:1 bis SEQ ID NO:8 bestimmt, oder:

$$n_a \leq x_a - (x_a \cdot y),$$

worin n_a die Anzahl der Aminosäureveränderungen ist, x_a die Gesamtanzahl der Aminosäuren in SEQ ID NO:2 ist, y 0,50 für 50%, 0,60 für 60%, 0,70 für 70%, 0,80 für 80%, 0,85 für 85%, 0,90 für 90%, 0,95 für 95%, 0,97 für 97% oder 1,00 für 100% steht und $\cdot$ das Symbol für den Multiplikationsoperator ist, und worin jedes nicht ganzzahlige Produkt von x_a und y zur nächsten ganzen Zahl vor dem Subtrahieren von x_a abgerundet wird.

[0152] Beispielsweise kann eine Polypeptidsequenz der vorliegenden Erfindung identisch mit der Referenzsequenz aus einem beliebigen aus SEQ ID NO:1 bis SEQ ID NO:8 sein, das heißt sie kann zu 100% identisch sein, oder sie kann bis zu einer bestimmten ganzen Anzahl von Aminosäureveränderungen im Vergleich mit der Referenzsequenz einschließen, so daß die prozentuale Identität weniger als 100% Identität ist. Solche Veränderungen werden aus der Gruppe ausgewählt, die aus wenigstens einer Aminosäuredeletion, -substitution, einschließlich konservativen und nicht-konservativen Substitution, oder -insertion besteht, und worin die Veränderungen an den Amino- oder Carboxy-terminalen Positionen der Referenzpolypeptidsequenz oder an beliebiger Stelle zwischen diesen terminalen Positionen erfolgen können, eingestreut entweder individuell unter den Aminosäuren in der Referenzsequenz oder in einer oder mehreren zusammenhängenden Gruppen innerhalb der Referenzsequenz. Die Anzahl der Aminosäureveränderungen für eine gegebene prozentuale Identität wird durch Multiplizieren der Gesamtanzahl an Aminosäuren in einem beliebigen aus SEQ ID NO:1 bis SEQ ID NO:8 durch die ganze Zahl, die die prozentuale Identität definiert, geteilt durch 100, und anschließendes Subtrahieren des Produkts von der gesamten Anzahl an Aminosäuren in einer beliebigen aus SEQ ID NO:1 bis SEQ ID NO:8 bestimmt, oder:

$$n_a \leq x_a - (x_a \cdot y),$$

worin n_a die Anzahl der Aminosäureveränderungen ist, x_a die Gesamtanzahl der Aminosäuren in einem beliebigen aus SEQ ID NO:1 bis SEQ ID NO:8 ist, y zum Beispiel 0,70 für 70%, 0,80 für 80%, 0,85 für 85% etc. ist und $\cdot$ das Symbol für den Multiplikationsoperator ist, und worin jedes nicht ganzzahlige Produkt von x_a und y zur nächsten ganzen Zahl vor dem Subtrahieren von x_a abgerundet wird.

Figurenbeschreibungen

[0153] [Fig. 1](#): Sequenzangaben für C-LytA. Jedes Repeat wurde auf Basis von sowohl multipler Sequenzangleichung als auch Sekundärstrukturvorhersage unter Verwendung der folgenden Angleichungsprogramme definiert: 1) MatchBox (E. Depiereux et al. (1992), Comput. Applic. Biosci. 8:501-9); 2) ClustalW (J.D. Thompson et al. (1994), Nucl. Acid Res. 22:4673-80); Block-Maker (S. Henikoff et al. (1995), Gene 163:S.17-26).

[0154] [Fig. 2](#): CPC und native Konstrukte (SEQ ID NOs:27-36)

[0155] [Fig. 3](#): Schematische Struktur von CPC-p501 His-Fusionsprotein, exprimiert in *S. cerevisiae*

[0156] [Fig. 4](#): Primärstruktur von CPC-P501 His-Fusionsprotein (SEQ ID NO:41)

[0157] [Fig. 5](#): Nukleotidsequenz von CPC P501 His(pRIT 15201) (SEQ ID NO:42)

[0158] [Fig. 6](#): Klonierungsstrategie zur Erzeugung von Plasmid pRIT 15201

[0159] [Fig. 7](#): Plasmidkarte von pRIT 15201

[0160] [Fig. 8](#): Vergleichende Expression von CPC P501 und P501 in *S. cerevisiae* Stamm DC 5

[0161] [Fig. 9](#): Erzeugung von CPC-P501S His (Y1796) im kleinen Maßstab. [Fig. 9A](#) stellt die Antigenproduktivität gemäß Abschätzung durch SDS-PAGE mit Silberanfärbung dar; [Fig. 9B](#) stellt die Antigenproduktivität gemäß Abschätzung durch Western Blot dar.

- [0162] [Fig. 10](#): Reinigungsschema von durch Y1796 erzeugtem CPC-P501-His.
- [0163] [Fig. 11](#): Muster von CPC P501 His-gereinigtem Protein (4-12% Novex Nu-Page Polyacrylamid vorgegossene Gele).
- [0164] [Fig. 12](#): Native P501S-Sequenz voller Länge (SEQ ID NO:17)
- [0165] [Fig. 13](#): Sequenz der CPC-P501S-Expressionskassette von JNW735 (SEQ ID NO:18)
- [0166] [Fig. 14](#): Zwei Codon-optimierte P501S-Sequenzen (SEQ ID NO:19-20)
- [0167] [Fig. 15](#): Zurückmanipulierte Codon-optimierte Sequenz 19 (SEQ ID NO:21)
- [0168] [Fig. 16](#): Zurückmanipulierte Codon-optimierte Sequenz 20 (SEQ ID NO:22)
- [0169] [Fig. 17](#): Die Ausgangssequenz zur Optimierung von CPC (SEQ ID NO:23)
- [0170] [Fig. 18](#): Repräsentative Codon-optimierte CPC-Sequenzen (SEQ ID NO: 24-25)
- [0171] [Fig. 19](#): Manipulierte Codon-optimierte CPC-Sequenz (SEQ ID NO:26)
- [0172] [Fig. 20](#): P501S-CPS-Fusionskandidatkonstrukte und Sequenzen (SEQ ID. NOs: 37-40 & 45-48)
- [0173] [Fig. 21](#): Western-Blot-Analyse von CHO-Zellen nach transienter Transfektion mit P501S (JNW680), CPC-P501S (JNW735) und leerer Vektorkontrolle.
- [0174] [Fig. 22](#): Anti-P501S-Antikörperreaktionen nach Immunisierung am Tag 0, 21 & 42 mit pVAC-P501S (JNW680, Mäuse B1-9) oder leerem Vektor (pVAC, Mäuse A1-6). Eine Vorblutentnahme erfolgte am Tag -1. Anschließende Blutentnahmen erfolgten an den Tagen 28 und 49 (Mäuse A1-3, B1-3) und 56 (Mäuse A4-6, B4-9). Alle Seren wurden mit einer 1/100-Verdünnung untersucht. Die Ergebnisse für die pVAC-immunisierten Mäuse wurden Bemittelt. Alle Ergebnisse für die individuellen pVAC-P501S-immunisierten Mäuse sind gezeigt. Als Positivkontrolle sind Seren aus Adeno-P501S-immunisierten Mäusen (Corixa Corp., verdünnt 1/100) eingeschlossen.
- [0175] [Fig. 23](#): Peptidbibliotheksdurchmusterung unter Verwendung von C57BL/6-Mäusen, immunisiert am Tag 0, 21, 42 und 70 mit pVAC-P501S (JNW680). Alle Peptide wurden in einer Endkonzentration von 50 µg/ml verwendet. Peptide 1-50 sind überlappende 15-20-mere, erhalten von Corixa. Peptide 51-70 sind vorhergesagte 8-9-mere Kb- und Db-Epitope und wurden von Mimotopes (GB) geordert. Proben 71-72 und 73-78 sind DMSO-Kontrollen bzw. Kontrollen ohne Peptid. Diagramm A zeigt die IFN-γ-Reaktionen, während Diagramm B die IL-2-Reaktionen zeigt. Zur Verwendung in anschließenden Immuntests ausgewählte Peptide sind schwarz gezeigt.
- [0176] [Fig. 24](#): Zelluläre Reaktionen gemäß ELISPOT am Tag 77 nach PMID-Immunisierung am Tag 0, 21, 42 und 70 mit pVAC-P501S (JNW680, B6-9) und pVAC leer (A4-6). Peptide 18, 22 & 48 wurden mit 50 µg/ml verwendet. CPC-P501S-Protein wurde mit 20 µg/ml verwendet. Diagramm A zeigt die IFN-γ-Reaktionen, während Diagramm B die IL-2-Reaktionen zeigt.
- [0177] [Fig. 25](#): Vergleich von P501S und CPC-P501S. Zelluläre Reaktionen wurden durch IL-2 ELISPOT unter Verwendung von Peptid 22 (10 µg/ml) am Tag 28 gemessen. Mäuse wurden durch PMID am Tag 0 und 21 mit pVAC leer (Kontrolle), pVAC-P501S (JNW680) und CPC-P501S (JNW735) immunisiert.
- [0178] [Fig. 26](#): Immunreaktion (Lymphproliferation an Milzzellen) nach Proteinimmunisierung mit CPC-P501S.
- [0179] [Fig. 27](#): Auswertung der Immunreaktion auf unterschiedliche CPC-P501S-Konstrukte. Zelluläre Reaktionen wurden durch IL-2-ELISPOT am Tag 28 gemessen. Mäuse wurden durch PMID am Tag 0 und 21 mit p7313, d.h. leer (Kontrolle), JNW735 und CPC-P501S-Konstrukten (JNW770, 771 und 773) immunisiert.
- [0180] [Fig. 28](#): MUC-1 CPC-Sequenzen (SEQ ID NOs: 49 & 50)

[0181] [Fig. 29](#): ss-CPC-MUC-1-Sequenzen (SEQ ID NOs: 51 & 52)

[0182] Die Erfindung wird weiter durch Verweis auf die folgenden Beispiele beschrieben:

Beispiel 1: Herstellung des rekombinanten Hefestamms Y1796, der P501-Fusionsprotein exprimiert, das ein C-LytA-P2-C-LytA (CPC) als Fusionspartner enthält

1. Proteinkonstruktion

[0183] Die Struktur des Fusionsproteins C-P2-C-p501 (alternativ als CPC-P501 bezeichnet) zur Expression in *S. cerevisiae* ist in [Fig. 3](#) gezeigt. Diese Fusion enthält die C-terminale Region von Gen LytA (Reste 187 bis 306), in die das P2-Fragment von Tetanustoxin (Reste 830-843) inseriert wurde. Das P2-Fragment wird zwischen die Reste 277 und 278 von C-Lyt-A plziert. Dem C-LytA-Fragment, das die P2-Insertion enthält, folgen P501 (Aminosäurereste 51 bis 553) und der His-Schwanz.

[0184] Die Primärstruktur des resultierenden Fusionsproteins hat die in [Fig. 4](#) beschriebene Sequenz, und die der obigen Proteinkonstruktion entsprechende codierende Sequenz ist in [Fig. 5](#).

2. Klonierungsstrategie zur Erzeugung eines Hefeplasmid, das CPC-P501 (51-553)-His-Fusionsprotein exprimiert

- Das Ausgangsmaterial ist der Hefevektor pRIT15068 (GB-Patentanmeldung 0015619.0).
- Dieser Vektor enthält den Hefe-Cupl-Promotor, die Hefe-alpha-Präprosignal-codierende Sequenz und die codierende Sequenz, die den Resten 55 bis 553 von P501S entspricht, gefolgt vom His-Schwanz.
- Die in [Fig. 6](#) umrissene Klonierungsstrategie schließt die folgenden Schritte ein:

a) Der erste Schritt ist die Insertion der P2-Sequenz (Codon-optimiert zur Hefeexpression) im Raster in die C-lytA-codierende Sequenz. Die C-lytA-codierende Sequenz wird vom Plasmid pRIT 14662 (PCT/EP99/00660) beherbergt. Die Insertion erfolgt unter Verwendung eines Adapters, der durch zwei als P21 und P22 bezeichnete komplementäre Oligonukleotide gebildet wird, in das zuvor durch NcoI geöffnete Plasmid pRIT 14662.

Die Sequenz von P21 und P22 ist:

P21 5' catgcaatacatcaaggctaactctaagtcattggtatcactgaaggcgt 3'

P22 3' gttatgtatgccgattgagattcaagtaaccatagtgactccgcagctac 5'

Nach Ligation und Transformation von *E. coli* und Charakterisierung der Transformante wird das als pRIT 15199 bezeichnete Plasmid erhalten.

b) Der zweite Schritt ist die Herstellung von C-lytA-P2-C-lytA-DNA-Fragment durch PCR-Amplifikation. Die Amplifikation wird unter Verwendung von pRIT15199 als Matrize und der als C-LytANOTATG und C-LytA-aa55 bezeichneten Oligonukleotide durchgeführt. Die Sequenz beider Oligonukleotide ist:

C-LytANOTATG

=5'aaaaccatggcgccgcttacgtacattccgacggctcttccaaaagacaag 3'

C-LytA-aa55 =5'aaacatgtacatgaactttctggcctgtctgccagtgttc 3'

Das amplifizierte Fragment wird mit den Restriktionsenzymen NcoI und AflIII zur Erzeugung der jeweiligen kohäsiven Enden behandelt.

c) Der nächste Schritt ist die Ligation des obigen Fragments mit Vektor pRIT 15068 (größtes nach NcoI-Behandlung erhaltenes Fragment) zur Erzeugung der kompletten Fusionsprotein-codierenden Sequenz. Nach Ligation und *E. coli*-Transformation wird das als pRIT15200 bezeichnete Plasmid erhalten. In diesem Plasmid enthält der verbleibende besondere NcoI-Ort das ATG, das für das Startcodon codiert.

d) Im nächsten Schritt wird ein NcoI-Fragment, das den CUP1-Promotor und einen Teil von 2 μ -Plasmidsequenzen enthält, aus Plasmid pRIT 15202 hergestellt. Plasmid pR2T 15202 ist ein Hefe 2 μ -Derivat, das den CUP1-Promotor mit einem NcoI-Ort bei ATG enthält (ATG-Sequenz: AAACC ATG).

e) Das aus pRIT 15202 isolierte NcoI-Fragment wird an pRIT 15200, zuvor mit NcoI geöffnet, in der richtigen Orientierung in einer solchen Weise ligiert, daß der pCUP1-Promotor auf der 5'-Seite der codierenden Sequenz ist. Dies führt zur Erzeugung des fertigen, als pRIT 15201 bezeichneten Expressionsplasmid (siehe [Fig. 7](#)).

3. Herstellung des rekombinanten Hefestamms Y1796 (RIX4440)

[0185] Das Plasmid pRIT 15201 wird zur Transformation des *S. cerevisiae*-Stamms DC5 (ATCC 20820) verwendet. Nach Selektion und Charakterisierung der Hefetransformanten, die das Plasmid pRIT 15201 enthalten, wird ein als Y1796 bezeichneter rekombinanter Hefestamm, der das CPC-P501-His-Fusionsprotein expri-

miert, erhalten. Das Protein nach Reduktion und Carboxyamidierung wird isoliert und durch Affinitätschromatographie (IMAC) gereinigt, gefolgt von Anionenaustauschchromatographie (Q Sepharose FF).

Beispiel II

[0186] In einer analogen Weise können Proteinkonstrukte wie in [Fig. 2](#) gezeigt unter Verwendung der darin gezeigten, entsprechenden DNA-Sequenzen exprimiert werden. Insbesondere entspricht der Hefestamm SC333 (Konstrukt 2) dem Y1796-Stamm, aber exprimiert P501₅₅₋₅₅₃ ohne den CPC-Fusionspartner. Hefestamm Y1800 (Konstrukt 3) entspricht dem Y1796-Stamm, aber umfaßt zusätzlich das native Sequenzsignal für P501S (Aminosäure 1-34), während Hefestamm Y1802 (Konstrukt 4) die alpha-Präsignalsequenz stromaufwärts der CPC-P501S-Sequenz umfaßt. Hefestamm Y1790 (Konstrukt 5) exprimiert ein P501S-Konstrukt, dem CPC fehlt und das die alpha-Präpro-Signalsequenz aufweist.

Beispiel III. Herstellung von gereinigtem CPC-P501

1. Herstellung von CPC-P501S-HIS (Y1796) im kleinen Maßstab

[0187] Für Y1796 in mit Histidin ergänztem Minimalmedium wird die Expression in der logarithmischen Phase durch Zugabe von CuSO₄ im Bereich von 100 bis 500 µM induziert, und die Kultur wird bei 30°C gehalten. Die Zellen werden nach 8 oder 24 h Induktion geerntet. Kupfer wird gerade vor der Verwendung hinzugegeben und nicht zuvor mit dem Medium vermischt.

[0188] Für die SDS-PAGE-Analyse wird die Extraktion der Hefezellen in Citrat-Phosphat-Puffer, pH 4,0 + 130 mM NaCl durchgeführt. Die Extraktion wird mit Glasperlen für eine kleine Zellmenge und mit einer French-Presse für eine höhere Zellmenge durchgeführt und dann mit Probenpuffer vermischt und durch SDS-PAGE analysiert. Die Ergebnisse der vergleichenden Analyse an SDS-PAGE der unterschiedlichen Konstrukte sind in [Fig. 8](#) dargestellt und nachfolgend in Tabelle 2 zusammengefaßt. Wie in der nachfolgenden Tabelle 2 gezeigt wird, ist der Expressionsgrad der Kultur viel höher für den Y1796-Stamm im Vergleich zur Expressionsstärke des Elternstammes SC333, eines Stammes, der das entsprechende P501S-His ohne CPC-Partner exprimiert. In gleicher Weise beeinflusst die Gegenwart einer Signalsequenz (alpha-prä) nicht die oben erörterten Ergebnisse: die Expressionsstärke der Kultur ist viel höher für den Y1802-Stamm im Vergleich mit der Expressionsstärke des entsprechenden Stammes Y1790, eines Stammes, der das entsprechende P501S-His ohne CPC-Partner exprimiert.

Tabelle 2

Rekom-bi-nanter Stamm	Plasmid	Pro-motor	Signalse-quenz	Fusi-ons-Partner	P501 Ami-nosäurese-quenzen	Expres-si-ons-stärke
SC333	Ma333	CUP1	-	-	55-553-His	θND
Y1796	pRIT 15201	CUP1	-	CPC	51-553-His	+++
Y1802	pRIT 15219	CUP1	α-prä	CPC	51-553-His	++++
Y1790	pRIT 15068	CUP1	α-präpro	-	55-553-His	+

CPC = clyta-P2-clyta

ND = nicht detektierbar, selbst im Western Blot

+ = detektierbar im Western Blot

+++ / ++++ = detektierbar im Western Blot und sichtbar in mit Silber angefärbten Gelen

2. Fermentation von Y1796 (RIX4440) im größeren Maßstab

[0189] 100 µl des Arbeitsansatzes werden auf festem Medium ausgebreitet und für ca. 24 h bei 30°C kultiviert. Diese feste Vorkultur wird dann zum Impfen einer flüssigen Vorkultur in Schüttelflaschen verwendet. Diese flüssige Vorkultur wird für 20 h bei 30°C kultiviert und in einen 20 l-Fermenter überführt. Die Zulauffermentation schließt eine Wachstumsphase von ca. 44 h und eine Induktionsphase von ca. 22 h ein.

[0190] Die Kohlenstoffquelle (Glucose) wurde zur Kultur durch kontinuierliche Zufuhr hinzugegeben. Die Restglucosekonzentration wurde sehr niedrig gehalten (≤50 mg/l), um die Ethanolerzeugung durch Fermenta-

tion zu minimieren. Dies wurde durch Begrenzen der Entwicklung des Mikroorganismus durch beschränkte Glucosezufuhr erreicht. Am Ende der Wachstumsphase wird der CUP1-Promotor durch Zugabe von CuSO_4 induziert, um das Antigen zu erzeugen.

[0191] Die Abwesenheit von Verunreinigungen wurde durch Impfen von 10^6 Zellen in Standard-TSB- und -THI-Fläschchen überprüft, die mit Nystatin ergänzt und für 14 Tage bei 20-25°C bzw. 30-35°C inkubiert wurden. Wie erwartet wurde kein Wachstum beobachtet.

3. Antigencharakterisierung und -produktivität

[0192] Zellhomogenate wurden durch French-Pressen von Fermentationsproben hergestellt, die zu unterschiedlichen Zeiten während der Induktionsphase geerntet wurden, und durch SDS-PAGE und Western-Blot analysiert. Es wurde gezeigt, daß der Großteil des Proteins von Interesse sich in der unlöslichen Fraktion befand, die aus dem Zellhomogenat nach Zentrifugieren erhalten wurde. Die SDS-PAGE- und Western Blot-Analysen, die in den nachfolgenden Figuren gezeigt sind, wurden an den Pellets durchgeführt, die nach Zentrifugieren dieser Zellhomogenate erhalten wurden.

[0193] Fig. 8A und B zeigen die Kinetik der Antigenproduktion während der Induktionsphase für Kultur PRO127. Es scheint, daß keine Antigenexpression während der Wachstumsphase erfolgte. Die spezifische Antigenproduktivität scheint von Beginn der Induktionsphase an bis zu 6 h zuzunehmen, und blieb dann relativ stabil bis zum Ende. Aber die volumetrische Produktivität erhöhte sich um einen Faktor 1,5 bis 2 aufgrund von Biomasseanreicherung, die während des gleichen Zeitraums beobachtet wurde. Die Antigenproduktivität wurde auf ca. 500 mg pro Liter Fermentationsbouillon durch Vergleich der gereinigten Referenz des Antigens und der Rohextrakte an SDS-PAGE mit Silberanfärbung (Fig. 9A) und WB-Analysen unter Verwendung eines Anti-P501S-Antikörpers abgeschätzt (ein muriner Ascite, gerichtet gegen P501S Aminosäuren 439-459, verwendet mit einer Verdünnung von 1/1000) (Fig. 9B).

Beispiel IV. Reinigung von CPC-P501 (51-553)-His-Fusionsprotein, erzeugt von Y1796

[0194] Nach dem Zellaufbruch ist das Protein mit der Pelletfraktion assoziiert. Eine Carbamidomethylierung des Moleküls wurde im Verfahren eingeführt, um mit der oxidativen Aggregation des Moleküls mit sich selbst oder mit Wirtszeil-Proteinverunreinigungen durch kovalente Verbrückung mit Disulfidbindungen zurechtzukommen. Die Verwendung von Detergentien wurde ebenfalls notwendig, um den hydrophoben Charakter dieses Proteins (12 Trans-Membrandomänen vorhergesagt) zu handhaben.

[0195] Das Reinigungsverfahren, entwickelt für den Maßstab von 1 l Kultur mit OD (optische Dichte) 120, ist in Fig. 10 beschrieben. Alle Arbeitsschritte werden bei Raumtemperatur (RT) durchgeführt. Gemäß dem DOC-TCA-BCA-Proteintest ist die globale Reinigungsausbeute 30-70 mg gereinigtes Antigen/l Kultur mit OD 120. Die Ausbeute ist mit dem Expressionsgrad der Kultur verbunden und ist höher im Vergleich mit der Reinigungsausbeute des Elternstamms, der unfusioniertes P501S-His exprimiert. Der Proteintest wird wie folgt durchgeführt: Proteine werden zuerst unter Verwendung von TCA (Trichloressigsäure) in Gegenwart von DOC (Desoxycholat) ausgefällt und dann in einem alkalischen Medium in Gegenwart von SDS gelöst. Die Proteine reagieren dann mit BCA (Bicinchoninsäure) (Pierce) unter Bildung eines löslichen violetten Komplexes, der eine hohe Extinktion bei 562 nm besitzt, die proportional zur in der Probe vorhandenen Proteinmenge ist. SDS-PAGE-Analyse von 3 gereinigten Ansätzen (Fig. 11) zeigt keinen Unterschied bei reduzierenden und nicht-reduzierenden Bedingungen (vgl. Bahnen 2, 3 und 4 gegenüber Bahnen 5, 6 und 7). Das Muster besteht aus einer Hauptbande bei 20 kDa, einer Verschmierung mit höherem Molekulargewicht und schwachen Abbaubanden. Alle Banden werden durch einen spezifischen monoklonalen Anti-P501S-Antikörper detektiert.

Beispiel V. Impfstoffherstellung unter Verwendung von CPC-P501S-His-Protein

[0196] Das Protein aus Beispiel 3 oder 4 kann zu einem Impfstoff formuliert werden, der QS21 und 3D-MPL in einer Öl-in-Wasser-Emulsion enthält.

1. Impfstoffherstellung

[0197] Das Antigen, hergestellt wie in Beispiel 1 bis 3 gezeigt, ein C-LytA-P2-P501S-His. Als Hilfsstoff umfaßt die Formulierung eine Mischung aus 3-des-O-acyliertem Monophosphoryllipid A (3D-MPL) und QS21 in einer Öl/Wasser-Emulsion. Das Hilfsstoffsystem SBAS2 wurde zuvor in WO 95/17210 beschrieben.

[0198] 3D-MPL: ist ein Immunstimulans, das aus dem Lipopolysaccharid (LPS) des Gram-negativen Bakteriums *Salmonella minnesota* stammt. MPL wurde desacyliert, und ihm fehlt eine Phosphat-Gruppe an der Lipid-A-Einheit. Diese chemische Behandlung reduziert dramatisch die Toxizität, während die immunstimulierenden Eigenschaften bewahrt werden (Ribi, 1986). Ribi Immunochemistry erzeugt und liefert MPL an SB-Biologics. Von SmithKline Beecham Biologicals durchgeführte Experimente haben gezeigt, daß 3D-MPL in Kombination mit verschiedenen Trägern die zelluläre Immunität sowohl vom humoralen als auch vom TH1-Typ stark steigert.

[0199] QS21: ist ein natürliches Saponinmolekül, das aus der Rinde des südamerikanischen Baums *Quillaja saponaria* Molina extrahiert wird. Eine zur Trennung der individuellen Saponine von den Rohextrakten der Rinde entwickelte Reinigungstechnik erlaubte die Isolierung des besonderen Saponins QS21, das ein Triterpenglycosid ist, das stärkere Hilfsstoffaktivität und geringere Toxizität im Vergleich mit der Stammkomponente zeigt. Es wurde gezeigt, daß QS21 MHC Klasse I beschränkte CTLs gegen mehrere Untereinheiten-Ags aktiviert sowie die Ag-spezifische Lymphozytenproliferation stimuliert (Kensil, 1992). Aquila (zuvor Cambridge Biotech Corporation) erzeugt und liefert QS21 an SB-Biologics.

[0200] Bei SmithKline Beecham Biologicals durchgeführte Experimente haben eine klare synergistische Wirkung der Kombinationen von MPL und QS21 in der Induzierung von zellulären Immunreaktionen vom sowohl humoralen als auch TH1-Typ gezeigt.

[0201] Die Öl/Wasser-Emulsion ist aus einer organischen Phase, die aus 2 Ölen (einem Tocopherol und Squalen) hergestellt ist, und einer wäßrigen Phase von PBS, die Tween 80 als Emulgator enthält, zusammengesetzt. Die Emulsion umfaßte 5% Squalen, 5% Tocopherol, 0,4% Tween 80 und besaß eine durchschnittliche Teilchengröße von 180 nm und ist als SB62 bekannt (siehe WO 95/17210).

[0202] Bei SmithKline Beecham Biologicals durchgeführte Experimente haben nachgewiesen, daß die Zugabe dieser O/W-Emulsion zu 3D-MPL/Q21 (SBAS2) die immunstimulatorischen Eigenschaften des letzteren gegen verschiedene Untereinheitantigene weiter steigert.

2. Herstellung von Emulsion SB62 (2-faches Konzentrat)

[0203] Tween 80 wird in phosphatgepufferter Kochsalzlösung (PBS) zum Erhalt einer 2%igen Lösung in PBS gelöst. Zur Bereitstellung von 100 ml von zweifacher Konzentratemulsion werden 5 g DL-alpha-Tocopherol und 5 ml Squalen zum sorgfältigen Vermischen verwirbelt. 90 ml PBS/Tween-Lösung werden hinzugegeben und sorgfältig vermischt. Die resultierende Emulsion wird dann durch eine Spritze geleitet und schließlich unter Verwendung einer Microfluidics-Vorrichtung M110S mikrofluidisiert. Die resultierenden Öltröpfchen haben eine Größe von ca. 180 nm.

3. Formulierungen

[0204] Eine typische Formulierung, die 3D-MPL und QS21 in einer Öl/Wasser-Emulsion enthält, wird wie folgt durchgeführt: 20 µg-25 µg C-LytA P2-P501S werden in 10-fach konzentriertem PBS pH 6,8 und H₂O verdünnt, vor der aufeinanderfolgenden Zugabe von SB62 (50 µl), MPL (20 µg), QS21 (20 µg), gegebenenfalls umfassend CpG-Oligonukleotid (100 µg) und 1 µg/ml Thiomersal als Konservierungsmittel. Die Menge jeder Komponente kann nach Bedarf variieren. Alle Inkubationen werden bei Raumtemperatur unter Rühren durchgeführt.

Beispiel VI. Codon-optimierte P501S-Sequenzen

1. Erzeugung der rekombinanten Kontrollplasmide

[0205] P501S-Sequenz voller Länge wurde in pVAC (Thomsen, Immunology, 1998; 95:510P105) kloniert, wodurch das Expressionsplasmid JNW680 erzeugt wurde. SEQ ID NO:17 stellt die humane P501S-Expressionskassette im Plasmid JNW680 dar und ist in [Fig. 12](#) veranschaulicht. Die Proteinsequenz von SEQ ID NO:17 ist im Einbuchstabenformat gezeigt, wobei die Start- und Stoppcodons fettgeschrieben gezeigt sind. Die Kozak-Sequenz ist durch die Doppelkreuzsymbole hervorgehoben. Der Codonverwendungsindex der humanen P501S-Sequenz (SEQ ID NO:17) beträgt 0,618, berechnet durch das SynGene-Programm.

SynGene-Programm

[0206] Grundsätzlich werden die Codonen unter Verwendung eines statistischen Verfahrens zugewiesen, um

ein synthetisches Gen mit einer Codonhäufigkeit zu ergeben, die näher an derjenigen ist, die natürlich in hoch-exprimierten E. coli- und humanen Genen gefunden wird.

[0207] SynGene ist eine aktualisierte Version des als CalcGene bezeichneten Visual Basic-Programms, das von R.S. Hale und G. Thompson geschrieben wurde (Protein Expression and Purification, Bd. 12, S. 185-188 (1998)). Für jeden Aminosäurerest in der ursprünglichen Sequenz wurde ein Codon auf der Basis der Wahrscheinlichkeit zugeordnet, mit der es in hoch-exprimierten E. coli-Genen erscheint. Einzelheiten des CalcGene-Programms, das unter Microsoft Windows 3.1 läuft, können von den Autoren erhalten werden. Weil das Programm ein statistisches Verfahren einsetzt, um Codonen dem synthetischen Gen zuzuweisen, sind nicht alle resultierenden Codonen die am häufigsten verwendeten im Zielorganismus. Vielmehr wird der Anteil der häufig und nicht häufig verwendeten Codonen des Zielorganismus in der synthetischen Sequenz durch Zuordnen von Codonen in den richtigen Anteilen widergespiegelt. Da es keine feste und schnelle Regel gibt, die ein besonderes Codon einer besonderen Position in der Sequenz zuordnet, wird das Programm jedoch bei jedem Mal, bei dem es läuft, ein anderes synthetisches Gen ergeben – obwohl jedes das gleiche Codonverwendungsmuster haben wird und jedes die gleiche Aminosäuresequenz codieren wird. Falls das Programm mehrere Male für eine gegebene Aminosäuresequenz und einen gegebenen Zielorganismus läuft, werden verschiedene unterschiedliche Nukleotidsequenzen erzeugt werden, die sich in der Anzahl, dem Typ und der Position von Restriktionsorten, Intron-Spleißsignalen etc. unterscheiden können, von denen einige unerwünscht sein können. Der Fachmann wird eine geeignete Sequenz zur Verwendung in der Expression des Polypeptids auf der Basis dieser Merkmale auswählen können.

[0208] Da die Codonen zufällig auf statistischer Basis zugeordnet werden, ist es außerdem möglich (obwohl vielleicht unwahrscheinlich), daß zwei oder mehr Codonen, die relativ selten im Zielorganismus verwendet werden, in großer Nähe gehäuft sein könnten. Es wird angenommen, daß solche Cluster die Translationsmaschinerie verwirren können und zu besonders niedrigen Expressionsraten führen können, so daß der Algorithmus zur Auswahl der Codonen im optimierten Gen etwaige Codonen mit einem RSCU-Wert von weniger als 0,2 für hoch-exprimierte Gene ausschließt, um die zufällige Auswahl etwaiger seltener Gencluster zu verhindern. Die Verteilung der verbleibenden Codonen wird dann gemäß den Häufigkeiten für hoch-exprimiertes E. coli zugewiesen, um eine Gesamtverteilung innerhalb des synthetischen Gens zu ergeben, die typisch für solche Gene (Koeffizient = 0,85) und auch für hoch-exprimierte humane Gene (Koeffizient = 0,50) ist.

[0209] Syngene (Peter Ert1, unveröffentlicht), eine aktualisierte Version des CalcGene-Programms, erlaubt den Ausschluß von seltenen Codonen in optionaler Weise und wird auch zur Zuordnung von Codonen gemäß dem Codonhäufigkeitsmuster von hoch-exprimierten humanen Genen verwendet.

[0210] Die Sequenz der CPC-P501S-Kassette, kloniert aus dem Vektor pRIT15201 (siehe [Fig. 7](#)) in pVAC, wodurch Plasmid JNW735 erzeugt wird, ist in SEQ ID NO:18 dargestellt und in [Fig. 13](#) veranschaulicht. Diese Sequenz ist identisch mit der pRIT15201-Sequenz mit der Ausnahme der Entfernung des His-Markers und der Addition einer Kozak-Sequenz (GCCACC) und geeigneter Restriktionsenzymorte. Die Aminosäuresequenz von SEQ ID NO:18 ist im Einbuchstabenformat gezeigt, die Start- und Stoppcodons sind in Fettschrift gezeigt. Die umrahmten Reste sind das P2-Helferepitop von Tetanustoxoid. Die unterstrichenen Reste sind der Clyta-Reinigungsmarker. Die Kozak-Sequenz ist durch die Doppelkreuzsymbole hervorgehoben.

2. Erzeugung der rekombinanten Plasmide mit P501S-Codon-optimierten Sequenzen

[0211] Obwohl der Codonkoeffizientenindex (CI) der nativen P501S-Sequenz bereits hoch ist (0,618), ist es möglich, den CI-Wert weiter zu erhöhen. Dies wird zwei potentielle Vorzüge haben – die Verbesserung der Antigenexpression und/oder Immunogenität und die Reduzierung der Möglichkeit zur Rekombination zwischen dem P501S-Vektor und genomischen Sequenzen.

[0212] Unter Verwendung des Syngene-Programms wurde eine Auswahl (SEQ ID NO:19 bis SEQ ID NO:20) der Codon-optimierten Sequenzen erhalten ([Fig. 14](#)). Die nachfolgende Tabelle 3 zeigt einen Vergleich des Codonkoeffizientenindex für die Ausgangs-P501S-Sequenz und die zwei repräsentativen Codon-optimierten Sequenzen, ausgewählt auf Basis eines geeigneten Restriktionsenzymortprofils und eines guten CI-Index.

Tabelle 3-Vergleich der Codonkoeffizientenindizes von zwei Codonoptimierten P501S-Genen

Sequenz	Codonkoeffizientenindex (CI)
P501S	0,618
SEQ ID NO:19	0,725
SEQ ID NO:20	0,755

3. Weitere Auswertung der Codon-optimierten Sequenzen

Sequenz SEQ ID NO:19

[0213] Obwohl SEQ ID NO:19 einen guten CI-Index hat (0,725), enthält es ein Dublett von seltenen Codonen in den Aminosäurepositionen 202 und 203. Diese Codonen wurden manuell durch häufigere Codonen durch Veränderung der DNA-Sequenz von TTGTTG zu CTGCTG substituiert. Zur Erleichterung der Klonierung und Expression wurden Restriktionsenzymorte und eine Kozak-Sequenz hinzugefügt. Die fertige manipulierte Sequenz (SEQ ID NO:21) ist in [Fig. 15](#) gezeigt. Das Syngene-Programm wurde zur Fragmentierung dieser Sequenz in Oligonukleotide mit einer minimalen Überlappung von 19-20 Basen verwendet. Deshalb zeigt [Fig. 15](#) die zurückmanipulierte P501S-Codon-optimierte SEQ ID NO:19. Restriktionsenzymorte sind unterstrichen, die Kozak-Sequenz ist fett geschrieben, die zurückmanipulierte DNA-Sequenz zur Entfernung eines seltenen Codon-dubletts ist umrahmt.

[0214] Unter Verwendung eines zweistufigen PCR-Protokolls wurden die überlappenden Primer, erzeugt durch das Syngene-Programm, zuerst unter Verwendung eines PCR-Konstruktionsprotokolls (nachfolgend ausgeführt) zusammengebaut. Die Konstruktionsreaktion erzeugt eine diverse Population von Fragmenten. Das korrekte Fragment voller Länge wurde unter Verwendung des PCR-Gewinnungsprotokolls und der terminalen Primer gewonnen/amplifiziert. Das resultierende PCR-Fragment wurde aus einem Agarosegel ausgeschnitten, gereinigt, mit NheI und XhoI beschränkt und in pVAC kloniert. Positive Klone wurden durch Restriktionsenzymanalyse identifiziert und durch doppelsträngige Sequenzierung bestätigt. Dies erzeugt Plasmid JNW766, das aufgrund der fehlerbehafteten Natur des PCR-Prozesses eine einzelne stille Mutation enthielt (C zu T in Position 360 von SEQ ID NO:21).

1. Konstruktionsreaktion – PCR-Bedingungen, generisches Protokoll

[0215] Reaktionmischung (Gesamtvolumen = 50 µl):

- 1x Reaktionspuffer (Pfx oder Proofstart)
- 1 µl Oligo-Sammlung (gleiche Mischung aus allen überlappenden Oligonukleotiden)
- 0,5 mM dNTPs
- DNA-Polymerase (Pfx oder Proofstart, 2,5-5U)
- +/- 1 mM MgSO₄
- +/- 1x Verstärkerlösung (Pfx-Verstärkung oder Proofstartpuffer Q)
- 1. 94°C für 120 s (nur Proofstart)
- 2. 94°C für 30 s
- 3. 40°C für 120 s
- 4. 72°C für 10 s
- 5. 94°C für 15 s
- 6. 40°C für 30 s
- 7. 72°C für 20 s + 3 s/Zyklus
- 8. Cyclieren zu Schritt 5, 25-mal
- 9. Halten bei 4°C

2. Gewinnungsreaktion – PCR-Bedingungen (generisches Protokoll)

[0216] Reaktionmischung (Gesamtvolumen = 50 µl):

- 1x Reaktionspuffer (Pfx oder Proofstart)
- 5-10 µl Konstruktionsreaktionsmischung
- 0,3-0,75 mM dNTPs
- 50 pmol Primer (5'-terminaler Primer, Sense-Orientierung)
- 50 pmol Primer (3'-terminaler Primer, Antisense-Orientierung)

- DNA-Polymerase (Pfx oder Proofstart, 2,5-5U)
- +/- 1 mM MgSO₄
- +/- 1x Verstärkerlösung (Pfx-Verstärker oder Proofstartpuffer Q)
- 1. 94°C, 120 s (nur Proofstart)
- 2. 94°C, 45 s
- 3. 60°C, 30 s
- 4. 72°C, 120 s
- 5. Cyclieren zu Schritt 2, 25-mal
- 6. 72°C, 240 s
- 7. Halten bei 4°C

Sequenz SEQ ID NO:20

[0217] Obwohl SEQ ID NO:20 einen sehr guten CI-Index hat (0,755), wurde festgestellt, daß es ein Dublett von seltenen Codonen in den Aminosäurepositionen 131 und 132 enthielt. Diese Codonen wurden manuell gegen häufigere Codonen durch Veränderung der DNA-Sequenz von TTGTTG zu CTGCTG substituiert. Zur Erleichterung der Klonierung wurde ein interner BamHI-Ort durch Mutieren von G zu C entfernt (siehe das doppelt unterstrichene Nukleotid in [Fig. 16](#)). Zur Erleichterung der Klonierung und Expression wurden Restriktionsenzymorte und eine Kozak-Sequenz hinzugefügt. Die fertige manipulierte Sequenz (SEQ ID NO:22) ist in [Fig. 16](#) gezeigt. Das Syngene-Programm wurde zur Fragmentierung dieser Sequenz in Oligonukleotide mit einer minimalen Überlappung von 19-20 Basen verwendet. [Fig. 16](#) zeigt deshalb die zurückmanipulierte P501S-Codon-optimierte Sequenz 20 (SEQ ID NO:22). Restriktionsenzymorte sind unterstrichen, die Kozak-Sequenz ist fett geschrieben, zurückmanipulierte DNA-Sequenz zur Entfernung eines selten Codondubletts ist eingerahmt und eine stumme Punktmutation zur Entfernung eines BamHI-Orts ist doppelt unterstrichen.

[0218] Unter Verwendung eines ähnlichen zweistufigen PCR-Protokolls zu dem oben beschriebenen wurde P501S-Fragment voller Länge amplifiziert und in pVAC kloniert. Positive Klone wurden durch Restriktionsenzymanalyse identifiziert und durch doppelsträngige Sequenzierung bestätigt. Dies erzeugte Plasmid JNW764. Die Sequenz der P501S-codierenden Kassette ist in [Fig. 16](#) gezeigt (SEQ ID NO: 22).

DNA-Sequenzähnlichkeit

[0219] Paarabstände nach Angleichung durch das (gewichtete) ClustalV-Verfahren sind der nachfolgenden Tabelle 3 gezeigt. Die nachfolgende Tabelle 4 zeigt die prozentuale Ähnlichkeit zwischen der humanen Ausgangs-P501S-Sequenz und den zwei Codon-optimierten Sequenzen SEQ ID NO:21 und 22, die zur weiteren Untersuchung ausgewählt wurden. Die Daten bestätigen, daß die Codon-optimierten DNA-Sequenzen ca. 80% ähnlich zur ursprünglichen P501S-Sequenz sind.

Tabelle 4

SEQ ID NO:	% Ähnlichkeit mit Ausgangs-P501S-Sequenz
21	79,6
22	79,4

Beispiel VII. Codon-optimierte CPC-Sequenzen

1. Ansatz

[0220] Da die ursprüngliche CPC-Sequenz ursprünglich zur optimalen Expression in Hefe geschaffen wurde, beschreibt dieser Abschnitt das Verfahren der Codon-Optimierung zur humanen Expression.

2. Sequenzkonstruktion

[0221] Die Ausgangssequenz für die Optimierung von CPC ist in [Fig. 17](#) gezeigt (SEQ ID NO:23). Diese stammt völlig aus pRIT15201 und enthält die gesamte codierende Sequenz von CPC plus vier Aminosäuren aus P501S zur Erleichterung der Klonierung stromabwärts. Unter Verwendung des Syngene-Programms wurde eine Selektion von Codon-optimierten Sequenzen erhalten, aus denen repräsentative Sequenzen in [Fig. 18](#) gezeigt sind (SEQ ID NO: 24-25). Die nachfolgende Tabelle 5 zeigt einen Vergleich des Codonkoeffizientenindex für die Ausgangs-CPC-Sequenz und die zwei repräsentativen Codon-optimierten Sequenzen.

Tabelle 5. Codonkoeffizientenindizes für zwei CPC-optimierte Sequenzen

Sequenz	Codonkoeffizientenindex (CI)
Ursprüngliches CPC = SEQ ID NO:23	0,506
SEQ ID NO:24	0,809
SEQ ID NO:25	0,800

[0222] Zusätzlich zur Codonoptimierung wurden alle Sequenzen auch auf Restriktionsenzymklonierungsorte durchmustert. Auf Basis des höchsten CI-Wertes und eines vorteilhaften Restriktionsenzymortprofils wurde SEQ ID NO:24 zur Konstruktion ausgewählt. Zur Erleichterung der Klonierung und Expression wurden 5'- und 3'-Klonierungsorte hinzugefügt, und eine Kozak-Sequenz (GCCACC) wurde 5' des einleitenden ATG-Startcodons inseriert. Diese manipulierte Sequenz ist in [Fig. 19](#) gezeigt (SEQ ID NO:26). Diese Sequenz schließt vier Aminosäuren von P501S (eingerahmt), Restriktionsenzymklonierungsorte (NheI und XhoI, unterstrichen), eine Kozak-Sequenz (fett geschrieben), ein Stoppcodon (kursiv) und 4 bp von flankierender irrelevanter DNA zur Erleichterung der Klonierung ein. Das Syngene-Programm wurde zur Fragmentierung dieser Sequenz zu 50- bis 60-meren Oligonukleotiden mit einer minimalen Überlappung von 18-20 Basen verwendet. Unter Verwendung eines ähnlichen zweistufigen PCR-Protokolls zu dem oben beschriebenen wurde das korrekte Fragment gewonnen/amplifiziert und in pVAC kloniert. Positive Klone wurde durch Restriktionsenzymanalyse identifiziert und die Sequenz unter Erzeugung von Vektor JNW759 verifiziert.

4. DNA-Ähnlichkeit

[0223] Paarabstände nach Angleichung mit ClustalV (gewichtet) sind in der nachfolgenden Tabelle 6 gezeigt. Die Tabelle zeigt die prozentuale Ähnlichkeit auf der DNA-Ebene zwischen der Ausgangssequenz von CPC und der Codonoptimierten Sequenz und bestätigt, daß die Codon-optimierten Sequenzen ca. 80% ähnlich mit der ursprünglichen CPC-Sequenz sind.

Tabelle 6

Sequenz SEQ ID NO::	% Ähnlichkeit mit Ausgangs-CPC-Sequenz
24	80,2
25	81,6

[0224] Beispiel VIII. Konstruktion des P501S-Fusionskandidaten Alle im nachfolgend gezeigten Schema gezeigten Kandidaten sind Codonoptimiert und werden unter Verwendung überlappender PCR-Methodik aus Plasmiden JNW764 und INW759 als Matrizen (SEQ ID NO:22 bzw. SEQ ID NO:26) konstruiert und in den Expressionsvektor p7313-1e kloniert.

[0225] Die vier nachfolgend schematisch gezeigten Kandidaten beruhen auf CPC-P501S. Codon-optimiertes CPC-P501S ist Konstrukt A. Kandidaten B, C und D schließen auch die Sequenz ein, die die N-terminalen 50 Aminosäuren von P501S codiert, positioniert entweder am N-Terminus von CPC-P501S (Konstrukt D), am C-Terminus von CPC-P501S (Konstrukt C) oder zwischen CPC und P501S (Konstrukt B). Eine schematische Darstellung der Konstrukte ist in [Fig. 20](#) angegeben. Die Nukleotid- und Proteinsequenzen für jedes der vier Konstrukte in in SEQ ID NO:37-40 für die Nukleotidsequenzen und SEQ ID NO:45-48 für die entsprechenden Polypeptidsequenzen gezeigt. In den Konstrukten A, C und D codiert das unterstrichene Codon vorzugsweise Tyrosin (entweder TAC oder TAT), aber die Nukleotidsequenz kann verändert werden, um Threonin zu codieren (entweder ACA, ACC, ACG oder ACT). In Konstrukt B codiert das unterstrichene Codon vorzugsweise Threonin (entweder ACA, ACC, ACG oder ACT), aber die Nukleotidsequenz kann verändert werden, um Tyrosin zu codieren (entweder TAC oder TAT). In allen Konstrukten wird die codierende Sequenz von entsprechenden Restriktionsenzymklonierungsorten (in diesem Fall NotI und BamHI) und einer Kozak-Sequenz unmittelbar stromaufwärts des einleitenden ATG flankiert. Die nachfolgende Tabelle 7 zeigt die Plasmididentifizierung für die oben erläuterten Konstrukte:

Tabelle 7

Konstrukt	Aminosäure im unterstrichenen Codon	Sequenz des Codons	Plasmid ID
A	Tyrosin	TAC	JNW771
B	Threonin	ACA	JNW773
B	Tyrosin	TAC	JNW770
C	Tyrosin	TAC	JNW777
D	Tyrosin	TAC	JNW769

[0226] Die zellulären Reaktionen nach Immunisierung mit p7313-ie (leerer Vektor), pVAC-P501S (JNW735), JNW770, JNW771 und JNW773 wurden durch ELISPOT nach einer primären Immunisierung durch PMID am Tag 0 und drei Auffrischungen am Tag 21, 42 und 70 bewertet. Tests wurden 7 Tag nach der Auffrischung durchgeführt. [Fig. 27](#) zeigt, daß gute IL-2 ELISPOT-Reaktionen in mit JNW770, JNW771 und JNW773 immunisierten Mäusen detektiert wurden.

Beispiel IX. Immunogenitätsexperimente unter Verwendung von Partikelvermittelten intradermalen Übertragungsstudien (PMID)

[0227] P501A voller Länge bei Übertragung durch eine Partikel-vermittelte intradermale Übertragung (PMID) erzeugt gute Antikörper- und zelluläre Reaktionen. Diese Daten zeigen, daß die PMID ein sehr wirksamer Übertragungsweg ist. Außerdem bestätigt der Vergleich von P501S und CPC-P501S, daß CPC-P501S eine stärkere Immunreaktion gemäß Bestimmung durch Peptid-ELISPOT induziert.

1. Materialien und Methoden

1.1 Kutane Genimmunisierung mit Genpistole

[0228] Plasmid-DNA wurde auf Goldperlen mit 2 µm Durchmesser unter Verwendung von Calciumchlorid und Spermidin präzipitiert. Beladene Perlen wurden auf Tefzel-Schlauch wie beschrieben aufgetragen (Eisenbraun et al., 1993; Pertmer et al., 1996). Die Partikelbombardierung wurde unter Verwendung des Accell-Genübertragungssystems durchgeführt (WO 95/19799). Für jedes Plasmid wurden weibliche C57BL/6-Mäuse an den Tagen 0, 21, 42 und 70 immunisiert. Jede Verabreichung bestand aus zwei Bombardierungen mit DNA/Gold, was eine Gesamtdosis von ca. 4-5 µg Plasmid lieferte.

1.2. ELISPOT-Tests für T-Zellreaktionen auf das P501S-Genprodukt

a) Herstellung von Milzzellen

[0229] Milzorgane wurden aus immunisierten Tieren an den Tagen 7-14 nach der Auffrischung erhalten. Die Milzorgane wurden durch Mahlen zwischen Glasplatten zur Erzeugung einer Zellsuspension verarbeitet. Rote Blutkörperchen wurden durch Ammoniumchloridbehandlung lysiert, und Trümmer wurden unter Zurücklassen einer feinen Suspension von Milzzellen entfernt. Die Zellen wurden auf eine Konzentration von 8×10^6 /ml in komplettem RPMI-Medium zur Verwendung in ELISPOT-Tests resuspendiert.

b) Durchmusterung der Peptidbibliothek

[0230] Eine Peptidbibliothek, die eine Mehrheit der P501s-Sequenz abdeckt, wurde von Corixa Corp. erhalten. Die Bibliothek enthielt 50 15-20-mere Peptide, die für 4-11 Aminosäuren überlappten. Die Peptide sind mit 1 bis 50 nummeriert. Zusätzlich wurde ein Vorhersageprogramm (H.G. Rammensee et al.: Immunogenetics, 1999, 50: 213-219) (<http://syfpeithi.bmiheidelberg.com/>) zur Vorhersage mutmaßlicher Kb- und Db-Epitope aus der P501S-Sequenz verwendet. Die zehn besten Epitope für Kb und Db wurden von Mimotopes (GB) bestellt und in der Bibliothek eingeschlossen. Zur Durchmusterung der Peptidbibliothek wurden Peptide in einer Endkonzentration von 50 µg/ml (ca. 25-50 µM) in IFN γ - und IL-2-ELISPOTS unter Verwendung des nachfolgend beschriebenen Protokolls verwendet. Für IFN γ -ELISPOTS wurde IL-2 zu den Tests mit 10 ng/ml hinzugegeben. Die für die Durchmusterung verwendeten Milzzellen wurden am Tag 84 aus C57BL/6-Mäusen entnommen, die am Tag 0, 21, 42 und 70 immunisiert wurden. Drei Peptide wurden aus der Bibliotheksdurchmusterung identifiziert – Peptide 18 (HCRQAYSVYAFMISLGGCLG), 22 (GLSAPSLSPHCCPCRARLAF) und 48

(VCLAAGITYVPPLLEVG). Diese Peptide wurden anschließend in den ELISPOT-Tests verwendet.

c) ELISPOT-Test

[0231] Platten wurden mit 15 µg/ml (in PBS) Ratte-Anti-Maus-INFγ oder Ratte-Anti-Maus-IL-2 (Pharmingen) beschichtet. Die Platten wurden über Nacht bei +4°C beschichtet. Vor der Verwendung wurden die Platten dreimal mit PBS gespült. Milzzellen wurden zu den Platten mit 4×10^5 Zellen/Vertiefung hinzugegeben. In der Bibliotheksdurchmusterung identifizierte Peptide wurden erneut von Genemed Synthesis bestellt und mit einer Endkonzentration von 50 µg/ml verwendet. CPC-P501S-Protein (GSKBio) wurde im Test mit 20 µg/ml verwendet. ELISPOT-Tests wurden in Gegenwart von entweder IL-2 (10 ng/ml), IL-7 (10 ng/ml) oder ohne Cytokin durchgeführt. Das Gesamtvolumen in jeder Vertiefung betrug 200 µl. Peptid-stimulierte Zellen enthaltende Platten wurden für 16 Stunden in einem befeuchteten 37°C-Inkubator inkubiert.

e) Entwicklung von ELISPOT-Testplatten

[0232] Zellen wurden aus den Platten durch einmaliges Spülen mit Wasser (mit 10 Minuten Tränken, um die Lyse der Zellen sicherzustellen) und dreimaliges Spülen mit PBS entfernt. Biotin-konjugiertes Ratte-Anti-Maus-INFγ oder IL-2 (Pharmingen) wurde mit 1 µg/ml in PBS hinzugegeben. Die Platten wurden unter Schütteln für 2 Stunden bei Raumtemperatur inkubiert. Die Platten wurden dann dreimal mit PBS vor Zugabe von Streptavidin alkalischer Phosphatase (Caltag) mit einer 1/1000-Verdünnung gespült. Nach drei Spülungen in PBS wurden Flecken durch Inkubation mit BCICP-Substrat (Biorad) für 15-45 min aufgezeigt. Das Substrat wurde unter Verwendung von Wasser abgespült, und man ließ die Platten trocknen. Flecken wurden unter Verwendung eines von Brian Hayes, Asthma Cell Biology-Einheit, GSK, entwickelten Bildanalysesystems gezählt.

1.3. ELISA-Test auf Antikörper gegen das P501S-Genprodukt

[0233] Serumproben wurden aus den Tieren durch Venenpunktion an den Tagen -1, 28, 49 und 56 erhalten und auf Gegenwart von Anti-P501S-Antikörpern getestet. Ein ELISA wurde unter Verwendung von Nunc Maxisorp-Platten durchgeführt, die über Nacht bei 4°C mit 0,5 µg/ml von CPC-P501S-Protein (GSKBio) in Natriumbicarbonatpuffer beschichtet wurden. Nach Spülen mit TBS-Tween (Tris-gepufferte Kochsalzlösung, pH 7,4, enthaltend 0,05% Tween 20) wurden die Platten mit Blockierungspuffer (3% BSA in TBS-Tween-Puffer) für 2 h bei Raumtemperatur geblockt. Alle Seren wurden mit einer 1:100 Verdünnung für 1 h bei RT in Blockierungspuffer inkubiert. Die Antikörperbindung wurde unter Verwendung von HRP-konjugierten Kaninchen-Anti-Maus-Immunglobulinen (#P0260, Dako) mit einer 1:2000 Verdünnung in Blockierungspuffer detektiert. Die Platten wurden erneut gespült und gebundenes Konjugat unter Verwendung von Fast OPD-Färbungsreagenzien (Sigma, GB) detektiert. Die Reaktion wurde durch Zugabe von 3M Schwefelsäure angehalten und das OPD-Produkt durch Messung der Extinktion bei 490 nm quantifiziert.

1.4. Transiente Transfektionstests

[0234] Humane HP501S-Expression aus verschiedenen DNA-Konstrukten wurde durch transiente Transfektion der Plasmide in CHO-Zellen (Ovarialzellen des Chinesischen Hamsters) nach Western Blotting an Gesamtzellprotein analysiert. Die transienten Transfektionen wurden mit dem Transfectam-Reagens (Promega) gemäß Herstelleranleitung durchgeführt. Kurz gesagt wurden Gewebekulturplatten mit 24 Vertiefungen mit 5×10^4 CHO-Zellen pro Vertiefung in 1 ml DMEM-Komplettmedium (DMEM, 10% FCS, 2 mM L-Glutamin, Penicillin 100 IU/ml, Streptomycin 100 µg/ml) geimpft und für 16 Stunden bei 37°C inkubiert. 0,5 µg DNA wurden zur 25 µl von 0,3 M NaCl (ausreichend für eine Vertiefung) hinzugegeben, und 2 µl Transfectam wurden zu 25 µl von Milli-Q hinzugegeben. Die DNA- und Transfectam-Lösungen wurden vorsichtig vermischt und bei Raumtemperatur für 15 Minuten inkubiert. Während dieses Inkubationsschrittes wurden die Zellen einmal in PBS gespült und mit 150 µl serumfreiem Medium (DMEM, 2 mM L-Glutamin) bedeckt. Die DNA-Transfectam-Lösung wurde zu den Zellen getropft, die Platte vorsichtig geschüttelt und bei 37°C für 4-6 Stunden inkubiert. 500 µl DMEM-Komplettmedium wurden hinzugegeben und die Zellen für weitere 48-72 Stunden bei 37°C inkubiert.

2. Western-Blot-Analyse von CHO-Zellen, die transient mit P501S-Plasmiden transfiziert wurden

[0235] Die transient transfizierten CHO-Zellen wurden mit PBS gespült und mit einer Versene (1:5000)/0,025% Trypsinlösung zur Übertragung der Zellen in Suspension behandelt. Nach Trypsinisierung wurden die CHO-Zellen pelletiert und erneut in 50 µl PBS suspendiert. Ein gleiches Volumen von 2x NP40-Lysepuffer wurde hinzugegeben und die Zellen auf Eis für 30 Minuten inkubiert. 100 µl von 2x TRIS-Glycin-SDS-Probenpuffer (Invitrogen), der 50 mM DTT enthielt, wurden hinzugegeben und die Lösung für 5 Minu-

ten auf 95°C erhitzt. 1-20 µl Probe wurden auf ein 4-20%iges TRIS-Glycin-Gel mit 1,5 mm (Invitrogen) geladen und bei konstanter Spannung (125 V) für 90 Minuten in 1x TRIS-Glycin-Puffer (Invitrogen) elektrophoretisch behandelt. Ein vorangefärbter Breitbandmarker (New England Biolabs, #P7708S) wurde zur Größenbewertung der Proben verwendet. Nach Elektrophorese wurden die Proben auf Immobilon-P PVDF-Membran (Millipore) übertragen, vorbenetzt in Methanol, unter Verwendung eines Xcell III-Blot-Moduls (Invitrogen), mit 1x Transferpuffer (Invitrogen), der 20% Methanol enthielt und mit einer konstanten Spannung von 25 V für 90 Minuten. Die Membran wurde über Nacht bei 4°C in TBS-Tween (Tris-gepufferte Kochsalzlösung, pH 7,4, enthaltend 0,05% Tween 20) geblockt, das 3% getrocknete Magermilch enthielt (Marvel). Der primäre Antikörper (10^3) wurde 1:1000 verdünnt und mit der Membran für 1 Stunde bei Raumtemperatur inkubiert. Nach umfassendem Spülen in TBS-Tween wurde der sekundäre Antikörper (HRP-konjugierte Kaninchen-anti-Maus-Immunglobuline (#P0260, Dako)) mit 1:2000 in TBS-Tween, das 3% getrocknete Magermilch enthielt, verdünnt und mit der Membran für eine Stunde bei Raumtemperatur inkubiert. Nach umfassendem Spülen wurde die Membran mit Supersignal West Pico Chemilumineszenz-Substrat (Pierce) für 5 Minuten inkubiert. Überschüssige Flüssigkeit wurde entfernt und die Membran zwischen zwei Bahnen aus Klebeband versiegelt und für 1-30 Minuten mit Hyperfilm ECL-Folie (Amersham-PharmaciaBiotech) in Kontakt gebracht.

3. Erzeugung der humanen P501S-Expressionskassette voller Länge

[0236] Der Ausgangspunkt für die Konstruktion einer P501S-Expressionskassette war das Plasmid pcDNA3.1-P501S (Corixa Corp.), das ein pcDNA3.1-Gerüst (Invitrogen) aufweist, das eine humane P501S-cDNA-Kassette voller Länge enthält, kloniert zwischen die EcoRI- und NotI-Orte. Dieser Vektor wird auch als JNW673 bezeichnet. Die Gegenwart von P501S wurde durch Fluoreszenzsequenzierung bestätigt. Die Sequenz der cDNA-Kassette wird durch die NCBI/Genbank-Sequenz (Eingangs-Nr. AY033593) angegeben. Humanes P501S wurde aus JNW673-Matrizen-DNA PCR-amplifiziert, mit XbaI und Sall beschränkt und in die NheI/XhoI-Orte von pVAC kloniert, um den Vektor JNW680 zu erzeugen. Die korrekte Orientierung des Fragments relativ zum CMV-Promotor wurde durch PCR und durch DNA-Sequenzierung bestätigt. Die Sequenz der Expressionskassette ist in [Fig. 12](#) gezeigt (SEQ ID NO: 17). Zur Konstruktion einer CPC-P501S-Expressionskassette wurde CPC-P501S aus dem Vektor pRIT15201 (siehe [Fig. 7](#)) PCR-amplifiziert, mit XbaI und Sall beschränkt und in die NheI- und XhoI-Orte von pVAC kloniert, wodurch Plasmid JNW735 erzeugt wurde. Die korrekte Orientierung wurde durch PCR und Sequenzierung bestätigt. Die Sequenz der CPC-P501S-Expressionskassette ist in [Fig. 13](#) gezeigt (SEQ ID NO:18).

4. Expression von humanem P501S aus den Plasmiden JNW680 und JNW735

[0237] Die P501S-Expressionsplasmide wurden transient in CHO-Zellen transfiziert und ein Gesamtzelllysats wie in den Methoden beschrieben hergestellt. Ein Western Blot eines Gesamtzelllysats identifizierte einzelne Banden mit ca. 55 kDa und 62 kDa für Proben, die mit JNW680 bzw. JNW735 transfiziert wurden ([Fig. 21](#)). Dies stimmt mit den vorhergesagten Molekulargewichten von 59,3 kDa und 63,3 kDa für P501S bzw. CPC-P501S überein. Die Addition des CPC-Markers beeinflusst die Expression von P501S nicht nachteilig.

5. Ergebnisse

5.1. Antikörperreaktionen auf humanes P501S nach PMID-Immunisierung

[0238] Die Antikörperreaktionen nach Immunisierung mit pVAC (leerer Vektor) und pVAC-P501S (JNW680) wurden durch ELISA nach einer primären Immunisierung durch PMID am Tag 0 und drei Auffrischungen am Tag 21 und Tag 42 und Tag 70 bewertet. [Fig. 22](#) zeigt die Antikörperreaktionen aus Seren, die am Tag -1, Tag 28 und Tag 49 (Mäuse A1-3, B1-3) und Tag 56 (Mäuse A4-6, B4-9) entnommen wurden. Obwohl es einige nicht-spezifische Reaktionen auf den leeren Vektor pVAC gab, wurden Reaktionen auf das P501S-Konstrukt in 5 von 9 Mäusen beobachtet.

5.2. Identifizierung von neuen T-Zellepitopen aus humanem P501S in C57BL/6-Mäusen durch Durchmusterung einer P501S-Peptidbibliothek

[0239] Nach Immunisierung mit JNW680 (pVAC-P501S) durch PMID am Tag 0 und drei Auffrischungen am Tag 21 und Tag 42 und Tag 70 wurden ELISPOT-Tests am Tag 84 durchgeführt. Peptide aus der P501S-Bibliothek wurden mit 50 µg/ml Endkonzentration untersucht. Aus dieser anfänglichen Durchmusterung wurde festgestellt, daß drei Peptide die INF-γ- und/oder IL-2-Sekretion stimulieren. Peptide 18, 22 und 48 ([Fig. 23](#)). Diese Peptide wurden in anschließenden zellulären Tests verwendet.

5.3 Zelluläre Reaktionen auf pVAC-P501S (JNW680) nach PMID-Immunisierung

[0240] Die zellulären Reaktionen nach Immunisierung mit pVAC (leerer Vektor) und pVAC-P501S wurden durch ELISPOT nach einer primären Immunisierung durch PMID am Tag 0 und drei Auffrischungen am Tag 21, 42 und 70 bewertet. Die Tests wurden 7 Tage nach der Auffrischung durchgeführt. Zwei unterschiedliche Testbedingungen wurden verwendet: 1) Peptide 18, 22 und 48, identifiziert in der Peptidbibliotheksdurchmusterung, verwendet mit einer Endkonzentration von 50 µg/ml, und 2) CPC-P501S-Protein, verwendet mit einer Endkonzentration von 20 µg/ml. **Fig. 24A** zeigt, daß, obwohl es keine P501S-spezifischen Reaktionen auf den leeren Vektor gab (A4-6), das pVAC-P501S-Konstrukt spezifische INF-γ-Reaktionen gegen Peptide 18 und 22 in allen Mäusen (B6-9) induzierte, während eine Maus (B7) auch eine IFN-γ-Reaktion gegen Peptid 48 zeigte. **Fig. 24B** zeigt, daß alle Mäuse spezifische IL-2-Reaktionen gegen Peptide 18, 22 und 48 zeigten. Außerdem zeigten pVAC-P501S-immunisierte Mäuse (B6-9) auch moderate IL-2-Reaktionen gegen CPC-P501S, wohingegen die mit leerem Vektor immunisierten Mäuse (A4-6) keine Reaktionen zeigten.

5.4. Vergleich der zellulären Reaktionen gegen P501S und CPC-P501S nach PMID-Immunisierung

[0241] Die zellulären Reaktionen nach Immunisierung mit pVAC (leerer Vektor), pVAC-P501S (JNW680) und CPC-P501S (JNW735) wurden durch ELISPOT nach einer primären Immunisierung durch PMID am Tag 0 und Auffrischungen am Tag 21 und 42 bewertet. Die Tests wurden 7 Tage nach der Auffrischung durchgeführt. Zwei unterschiedliche Testbedingungen wurden verwendet: 1) Peptide 18, 22 und 48, identifiziert in der Peptidbibliotheksdurchmusterung, verwendet mit einer Endkonzentration von 50 µg/ml, und 2) CPC-P501S-Protein, verwendet mit einer Endkonzentration von 20 µg/ml. **Fig. 25** zeigt, daß am Tag 28 CPC-P501S gute IL-2-Reaktionen gegen 10 µg/ml von Peptid 22 induzierte, während es keine P501S-spezifischen Reaktionen gegen weder den leeren Vektor noch das pVAC-P501S gab. Diese Ergebnisse wurden auch unter Verwendung von CPC-P501S-Protein zur Restimulation der Milzzellen beobachtet. Am Tag 49 (nach der 2. Auffrischung) waren die durch P501S und CPC-P501S induzierten Reaktionen äquivalent. Diese Daten legen nahe, daß die Addition des CPC-Markers die Kinetik und/oder Stärke der Reaktion gegen P501S verbessert.

Beispiel IX. Immunitätsexperimente in Mäusen unter Verwendung von P501S-Protein + Hilfsstoffstudien

1. Konstruktion und Hilfsstoffformulierung

[0242] Die durch Impfung unter Verwendung des rekombinanten gereinigten CPC-P501S-Proteins, formuliert in Hilfsstoffen, induzierte Immunreaktion wird in Experimenten charakterisiert, die in Mäusen durchgeführt werden. Gruppen von 5 bis 10 acht Wochen alten weiblichen C57BL/6-Mäusen werden 2- bis 6-mal intramuskulär in 2-wöchigen Intervallen mit 10 µg des CPC-P501S-Proteins geimpft, formuliert in unterschiedlichen Hilfsstoffsystemen. Das verabreichte Volumen entspricht 1/10 einer Humandosis (50 µl). Die Serologie (Gesamt-Ig-Reaktion) und zelluläre Reaktion (T-Zell-Lymphproliferation und Cytokinerzeugung) werden an Milzzellen 6-14 Tage nach der letzten Impfung unter Verwendung von Standardprotokollen analysiert, wie beschrieben in C. Gérard et al., 2001, Vaccine 19, 2583-2589.

[0243] Die Daten eines repräsentativen Experiments sind gezeigt. Es schloß 5 Gruppen von 8 weiblichen C57BL/6-Mäusen ein, die 4 intramuskuläre Injektionen von CPC-P501 (10 µg) + Hilfsstoff (A, B, C) an den Tagen 0, 14, 28, 42 erhielten. Beispiel V liefert ein experimentelles Protokoll, wie die Formulierungen durchzuführen sind. Kurz gesagt sind die Hilfsstoffformulierungen wie folgt (Mengen für eine Dosis von 100 µl angegeben):

- Hilfsstoff A: QS21 (10 µg), MPL (10 µg) und CpG7909 (100 µg), hergestellt gemäß dem in WO 00/62800 offenbarten Verfahren;
- Hilfsstoff B: Formulierung aus QS21 (20 µg), MPL (20 µg), CpG7909 (100 µg) und 50 µg SB62 Öl-in-Wasser-Emulsion (WO 95/17210);
- Hilfsstoff C: Formulierung aus QS21 (10 µg), MPL (10 µg), CpG7909 (100 µg) und 10 µl SB62 Öl-in-Wasser-Emulsion (WO 99/12565).

2. Serologie

[0244] Die durch Impfung induzierte Ig-Gesamtreaktion wurde durch ELISA unter Verwendung von entweder CPC-P501 oder RAI2-P501 (C-Terminus, das eine trunkierte Form des P501-Proteins ist, entsprechend dem C-Terminus des Proteins, fusioniert an seinem N-Terminus an ein aus TB stammendes Protein RAI2; Ra12 stammt aus MTB32A-Antigen, beschrieben in Skeiky et al., Infection and Immunity. (1999)67:3998-4007) gemessen.

[0245] Die mit Hilfsstoff versetzten CPC-P501S-Proteine ergeben eine gute Antikörperreaktion nach Impfung.

3. Zelluläre Reaktion

3.1. Lymphproliferation

[0246] 7 Tage nach der letzten Impfung wurde die Lymphproliferation an Milzzellen individuell durchgeführt. 2×10^5 Milzzellen wurden 4-fach in Mikroplatten mit 96 Vertiefungen in RPMI-Medium ausplattiert, das 1% normales Mäuseserum enthielt. Nach 72 Stunden Restimulation mit entweder Immunogen (CPC-P501) oder trunkiertem Protein (RAI2-P501) in unterschiedlichen Konzentrationen wurde 1 μ Ci 3H-Thymidin (Amersham 5 Ci/ml) hinzugegeben. Nach 16 Stunden wurden die Zellen auf Filterplatten geerntet. Aufgenommene Radioaktivität wurde in einem β -Zähler gezählt. Die Ergebnisse werden in CPM oder als Stimulationsindizes* (geometrischer Mittelwert CPM in Kulturen mit Antigen/geometrischer Mittelwert CPM in Kulturen ohne Antigen) ausgedrückt. Restimulation mit ConA (2 μ g/ml) als Positivkontrolle wurde als Positivkontrolle eingeschlossen.

[0247] Wie in [Fig. 26](#) gezeigt wird, wird eine P501-spezifische Lymphproliferation in der Milz aller Gruppen von Mäusen beobachtet, die das mit Hilfsstoff versetzte Protein nach In-vitro-Restimulation mit entweder Immunogen oder einem anderen P501-Protein, das in einem Expressionssystem (E. coli) hergestellt wurde, beobachtet, was anzeigt, daß T-Zellen in vivo durch die Impfung vorbereitet wurden.

3.2. IFN γ -Produktion, gemessen durch intrazelluläre Anfärbung von Milzzellen

[0248] Dendritische Knochenmarkszellen (BMDC), erhalten nach Kultur von Mäuse-PBL für 7 Tage in Gegenwart von GM-CSF.

[0249] 7 Tage nach der letzten Impfung werden Milz- oder PBL-Zellen aufgefangen und eine Zellsuspension hergestellt. 10^6 Zellen (1 Sammlung pro Gruppe) wurden für +/- 18 h mit 10^5 BMDC inkubiert, gepulst über Nacht mit 10 μ g/ml von entweder CPCp-501-Protein oder RAI2. Nach Behandlung mit dem 2.4.G.2-Antikörper wurden Milzzellen mit fluoreszierenden Anti-CD4- und -CD8-Antikörpern (Anti-CD4-APC und ein Anti-CD8PerCP) angefärbt. Nach einem Permeabilisierungs- und Fixierungsschritt wurden die Zellen mit einem fluoreszierenden Anti-IFN γ -FITC-Antikörper angefärbt.

[0250] In mit CPC-P501 in unterschiedlichem Hilfsstoff geimpften Mäusen wird gezeigt, daß sowohl CD4- als auch CD8-T-Zellen IFN γ als Reaktion auf DC erzeugen, gepulst mit entweder dem Immunogen oder dem C-terminalen p501, hergestellt in E. coli (wie gezeigt durch intrazelluläre Anfärbung von Milz und PBLs). Es gibt eine 4-10-fache Zunahme des Prozentwertes der Zellen, die dieses Cytokin herstellen, in den Gruppen, die das mit Hilfsstoff versetzte CPC-P501S erhalten, verglichen mit dem Protein allein, und es wird gezeigt, daß 0,1 bis 10% der CD4- oder CD8-T-Zellen IFN γ erzeugen.

[0251] Diese Daten erlauben die Schlußfolgerung, daß das mit Hilfsstoff versetzte CPC-P501-Protein immunogen in Mäusen ist. Sowohl eine P501-spezifische humorale als auch zelluläre Reaktion, einschließlich IFN γ -Erzeugung durch CD4- und CD8-T-Zellen, kann nach mehrfacher intramuskulärer Impfung mit CPC P501 in Hilfsstoffen detektiert werden.

Beispiel X. CPC-MUC-1-Konstrukte und Sequenzen

[0252] Die CPC-Sequenz wird aus Nukleotid SEQ ID NO:28 entnommen. Die MUC1-Sequenz ist aus der Genbank-Datenbank erhältlich (Eingangs-Nr. NM_002456).

1. MUC-1-CPC-Konstrukt

[0253] Aufgrund der Gegenwart einer Signalsequenz in MUC1, die posttranslational abgespalten wird, wurde das CPC-Motiv am C-Terminus plaziert. Die resultierende MUC-1-CPC-DNA-Sequenz ist in SEQ ID NO:xx ([Fig. 28A](#)) und die entsprechende MUC1-CPC-Proteinsequenz in SEQ ID NO:yy ([Fig. 28B](#)) dargestellt.

2. ss-CPC-MUC1-Konstrukt

[0254] Aufgrund der Gegenwart einer Signalsequenz in MUC1, die posttranslational abgespalten wird, wurde die MUC1-Signalsequenz durch eine heterologe Leitsequenz (aus der schweren Kette von Humanimmunoglobulin) ausgetauscht, und das CPC-Motiv wurde zwischen die heterologe Leitsequenz und die MUC1-Sequenz

inseriert, was eine als ss-CPC-MUC1 bezeichnete Sequenz erzeugte, wie in [Fig. 29](#) dargestellt.

Sequenzprotokoll

```

<110>      GlaxosmithKline Biologicals sa
           Glaxo Group Ltd.
<120>      Immunogene Zusammensetzungen
<130>      B45311
<160>      52
<170>      FastSEQ for Windows Version 4.0

<210>      1
<211>      15
<212>      PRT
<213>      Streptococcus pneumoniae

<400>      1
  Gly Trp Gln Lys Asn Asp Thr Gly Tyr Trp Tyr Val His Ser Asp
   1              5              10              15

<210>      2
<211>      21
<212>      PRT
<213>      Streptococcus pneumoniae

<400>      2
  Gly Ser Tyr Pro Lys Asp Lys Phe Glu Lys Ile Asn Gly Thr Trp Tyr
   1              5              10              15
  Tyr Phe Asp Ser Ser
                20

<210>      3
<211>      22
<212>      PRT
<213>      Streptococcus pneumoniae

<400>      3
  Gly Tyr Met Leu Ala Asp Arg Trp Arg Lys His Thr Asp Gly Asn Trp
   1              5              10              15
  Tyr Trp Phe Asp Asn Ser
                20

<210>      4
<211>      20
<212>      PRT
<213>      Streptococcus pneumoniae

<400>      4
  Gly Glu Met Ala Thr Gly Trp Lys Lys Ile Ala Asp Lys Trp Tyr Tyr
   1              5              10              15
  Phe Asn Glu Glu
                20

<210>      5
<211>      21
<212>      PRT
<213>      Streptococcus pneumoniae

```

<400> 5
 Gly Ala Met Lys Thr Gly Trp Val Lys Tyr Lys Asp Thr Trp Tyr Tyr
 1 5 10 15
 Leu Asp Ala Lys Glu
 20

<210> 6
 <211> 23
 <212> PRT
 <213> Streptococcus pneumoniae

<400> 6
 Gly Ala Met Val Ser Asn Ala Phe Ile Gln Ser Ala Asp Gly Thr Gly
 1 5 10 15
 Trp Tyr Tyr Leu Lys Pro Asp
 20

<210> 7
 <211> 142
 <212> PRT
 <213> Streptococcus pneumoniae

<400> 7
 Gly Trp Gln Lys Asn Asp Thr Gly Tyr Trp Tyr Val His Ser Asp Gly
 1 5 10 15
 Ser Tyr Pro Lys Asp Lys Phe Glu Lys Ile Asn Gly Thr Trp Tyr Tyr
 20 25 30
 Phe Asp Ser Ser Gly Tyr Met Leu Ala Asp Arg Trp Arg Lys His Thr
 35 40 45
 Asp Gly Asn Trp Tyr Trp Phe Asp Asn Ser Gly Glu Met Ala Thr Gly
 50 55 60
 Trp Lys Lys Ile Ala Asp Lys Trp Tyr Tyr Phe Asn Glu Glu Gly Ala
 65 70 75 80
 Met Lys Thr Gly Trp Val Lys Tyr Lys Asp Thr Trp Tyr Tyr Leu Asp
 85 90 95
 Ala Lys Glu Gly Ala Met Val Ser Asn Ala Phe Ile Gln Ser Ala Asp
 100 105 110
 Gly Thr Gly Trp Tyr Tyr Leu Lys Pro Asp Gly Thr Leu Ala Asp Arg
 115 120 125
 Pro Glu Phe Thr Val Glu Pro Asp Gly Leu Ile Thr Val Lys
 130 135 140

<210> 8
 <211> 112
 <212> PRT
 <213> Streptococcus pneumoniae

<400> 8
 Tyr Val His Ser Asp Gly Ser Tyr Pro Lys Asp Lys Phe Glu Lys Ile
 1 5 10 15
 Asn Gly Thr Trp Tyr Tyr Phe Asp Ser Ser Gly Tyr Met Leu Ala Asp
 20 25 30
 Arg Trp Arg Lys His Thr Asp Gly Asn Trp Tyr Trp Phe Asp Asn Ser
 35 40 45

Gly	Glu	Met	Ala	Thr	Gly	Trp	Lys	Lys	Ile	Ala	Asp	Lys	Trp	Tyr	Tyr
50						55				60					
Phe	Asn	Glu	Glu	Gly	Ala	Met	Lys	Thr	Gly	Trp	Val	Lys	Tyr	Lys	Asp
65				70					75					80	
Thr	Trp	Tyr	Tyr	Leu	Asp	Ala	Lys	Glu	Gly	Ala	Met	Val	Ser	Asn	Ala
			85					90					95		
Phe	Ile	Gln	Ser	Ala	Asp	Gly	Thr	Gly	Trp	Tyr	Tyr	Leu	Lys	Pro	Asp
		100						105				110			

<210> 9
 <211> 45
 <212> DNA
 <213> Streptococcus pneumoniae

<400> 9
ggctggcaga agaatgacac tggctactgg tacgtacatt cagac 45

<210> 10
 <211> 63
 <212> DNA
 <213> Streptococcus pneumoniae

<400> 10
ggctcttattc caaaagacaa gtttgagaaa atcaatggca cttggtacta ctttgacagt 60
tca 63

<210> 11
 <211> 66
 <212> DNA
 <213> Streptococcus pneumoniae

<400> 11
ggctatatgc ttgcagaccg ctggaggaag cacacagacg gcaactggta ctggttcgac 60
aactca 66

<210> 12
 <211> 60
 <212> DNA
 <213> Streptococcus pneumoniae

<400> 12
ggcgaaatgg ctacaggctg gaagaaaatc gctgataagt ggtactatct caacgaagaa 60

<210> 13
 <211> 63
 <212> DNA
 <213> Streptococcus pneumoniae

<400> 13
ggtgccatga agacaggctg ggtcaagtac aaggacactt ggtactactt agacgctaaa 60
gaa 63

<210> 14
 <211> 69

<212> DNA
 <213> Streptococcus pneumoniae

<400> 14
 ggcgccatgg tatcaaatgc ctttatccag tcagcggacg gaacaggctg gtactacctc 60
 aaaccagac 69

<210> 15
 <211> 429
 <212> DNA
 <213> Streptococcus pneumoniae

<400> 15
 ggctggcaga agaatgacac tggctactgg tacgtacatt cagacggctc ttatccaaaa 60
 gacaagtttg agaaaatcaa tggcacttgg tactactttg acagttcagg ctatatgctt 120
 gcagaccgct ggaggaagca cacagacggc aactgggtact ggttcgacaa ctcaggcgaa 180
 atggctacag gctggaagaa aatcgctgat aagtgggtact atttcaacga agaagggtgcc 240
 atgaagacag gctgggtcaa gtacaaggac acttgggtact acttagacgc taaagaaggc 300
 gccatgggtat caaatgcctt tatccagtca gcggacggaa caggctggta ctacctcaaa 360
 ccagacggaa cactggcaga caggccagaa ttcacagtag agccagatgg cttgattaca 420
 gtaaaataa 429

<210> 16
 <211> 336
 <212> DNA
 <213> Streptococcus pneumoniae

<400> 16
 tacgtacatt ccgacggctc ttatccaaaa gacaagtttg agaaaatcaa tggcacttgg 60
 tactactttg acagttcagg ctatatgctt gcagaccgct ggaggaagca cacagacggc 120
 aactgggtact ggttcgacaa ctcaggcgaa atggctacag gctggaagaa aatcgctgat 180
 aagtgggtact atttcaacga agaagggtgcc atgaagacag gctgggtcaa gtacaaggac 240
 acttgggtact acttagacgc taaagaaggc gccatgggtat caaatgcctt tatccagtca 300
 gcggacggaa caggctggta ctacctcaaa ccagac 336

<210> 17
 <211> 1674
 <212> DNA
 <213> Homo sapiens

<400> 17
 gccaccatgg tccagaggct gtgggtgagc cgcctgctgc ggcaccggaa agcccagctc 60
 ttgctgggtca acctgctaac ctttggcctg gaggtgtgtt tggccgcagg catcacctat 120
 gtgccgcctc tgctgctgga agtgggggta gaggagaagt tcatgaccat ggtgctgggc 180
 attgggtccag tgctgggcct ggtctgtgtc ccgctcctag gctcagccag tgacctggtg 240
 cgtggacgct atggccgccg ccggcccttc atctgggcac tgccttggg catcctgctg 300
 agcctctttc tcatcccaag ggccggctgg ctagcagggc tgctgtgccc ggatcccagg 360
 cccctggagc tggcactgct catcctgggc gtggggctgc tggacttctg tggccagggtg 420
 tgcttcactc cactggaggc cctgctctct gacctcttcc gggaccggga ccactgtcgc 480
 caggcctact ctgtctatgc cttcatgatc agtcttgggg gctgcctggg ctacctcctg 540
 cctgccattg actgggacac cagtgccttg gccccctacc tgggcacca ggaggagtgc 600

```

ctcttttgcc  tgtcacccct  catcttcttc  acctgcgtag  cagccacact  gctgggtggct  660
gaggaggcag  cgctgggccc  caccgagcca  gcagaagggc  tgtcggcccc  ctctttgtcg  720
ccccactgct  gtccatgccg  ggcccgtttg  gctttccgga  acctggggcg  cctgcttccc  780
cggctgcacc  agctgtgctg  ccgcatgccc  cgcacctgc  gccggctctt  cgtggctgag  840
ctgtgcagct  ggatggcact  catgaccttc  acgctgtttt  acacggattt  cgtgggcgag  900
gggctgtacc  agggcgtgcc  cagagctgag  ccgggcaccg  agggccggag  acactatgat  960
gaaggcgctc  ggatgggcag  cctggggctg  ttcttgcaat  gcgccatctc  cctggctctt  1020
tctctggtea  tggaccgggt  ggtgcagcga  ttccggcactc  gagcagtcta  tttggccagt  1080
gtggcagctt  tccctgtggc  tgccgggtgc  acatgcctgt  cccacagtgt  ggccgtgggtg  1140
acagcttcag  ccgccctcac  cgggttcacc  ttctcagccc  tgcagatcct  gccctacaca  1200
ctggcctccc  tctaccaccg  ggagaagcag  gtgttcctgc  ccaaataacc  aggggacact  1260
ggaggtgcta  gcagtgagga  cagcctgatg  accagcttcc  tgccaggccc  taagcctgga  1320
gctcccttcc  ctaatggaca  cgtgggtgct  ggaggcagtg  gcctgctccc  acctccacc  1380
gcgctctgcg  gggcctctgc  ctgtgatgtc  tccgtacgtg  tgggtggggg  tgagcccacc  1440
gaggccaggg  tggttccggg  ccggggcatc  tgccctggacc  tcgccatcct  ggatagtgcc  1500

ttcctgctgt  cccagggtgg  cccatccctg  tttatgggct  ccattgtcca  gctcagccag  1560
tctgtcactg  cctatatggt  gtctgcccga  ggccctgggt  tggctgccat  ttactttgct  1620
acacaggtag  tatttgacaa  gagcgacttg  gccaaatact  cagcgtaggt  cgag  1674

```

<210> 18
 <211> 1947
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> Hybriden zwischen St. pneum. C-LytA, P2 T-Helferepitop
 und humanem P501S.

<400> 18

```

gccaccatgg  cggccgctta  cgtacattcc  gacggctctt  atccaaaaga  caagtttgag  60
aaaatcaatg  gcacttggtg  ctactttgac  agttcaggct  atatgcttgc  agaccgctgg  120
aggaagcaca  cagacggcaa  ctgggtactg  ttcgacaact  caggcgaaat  ggctacaggc  180
tggaagaaaa  tcgctgataa  gtggtactat  ttcaacgaag  aaggtgccat  gaagacaggc  240
tgggtcaagt  acaaggacac  ttggtactac  ttagacgcta  aagaaggcgc  catgcaatac  300
atcaaggcta  actctaagtt  cattgggtatc  actgaaggcg  tcatggtatc  aaatgccttt  360
atccagtcag  cggacggaac  aggcgtggta  tacctcaaac  cagacggaac  actggcagac  420
aggccagaaa  agttcatgta  catggtgctg  ggcattgggt  cagtgtctgg  cctggctctg  480
gtcccgctcc  taggctcagc  cagtgaccac  tggcgtggac  gctatggccg  ccgccggccc  540
ttcatctggg  cactgtcctt  gggcatcctg  ctgagcctct  ttctcatccc  aaggggccgg  600
tggctagcag  ggctgctgtg  ccgggatccc  agggccctgg  agctggcact  gctcatcctg  660
ggcgtggggc  tgcaggactt  ctgtggccag  gtgtgcttca  ctccactgga  ggccctgctc  720
tctgacctct  tccgggaccc  ggaccactgt  cggcaggcct  actctgtcta  tgccttcagt  780
atcagtcttg  ggggctgcct  gggctacctc  ctgcctgcca  ttgactggga  caccagtgcc  840
ctggccccc  acctgggcac  ccaggaggag  tgccctcttg  gcctgctcac  cctcatcttc  900
ctcacctgcg  tagcagccac  actgctgggt  gctgaggagg  cagcgtctgg  cccaccgag  960
ccagcagaag  ggctgtcggc  cccctccttg  tcgccccact  gctgtccatg  ccggggccgc  1020
ttggctttcc  ggaacctggg  cgcctgtctt  ccccggtgc  accagctgtg  ctgccgcatg  1080
ccccgcacc  tgcgccggct  cttcgtggct  gagctgtgca  gctggatggc  actcatgacc  1140
ttcacgctgt  tttacacgga  ttctgtgggc  gaggggctgt  accagggcgt  gccagagct  1200
gagccgggca  ccgaggcccg  gagacactat  gatgaaggcg  ttccgatggg  cagcctgggg  1260
ctgttcctgc  agtgcccat  ctccctgggt  ttctctctgg  tcatggaccg  gctggtgcag  1320
cgattcggca  ctcgagcagt  ctatttggcc  agtgtggcag  ctttccctgt  ggctgccgg  1380
gccacatgcc  tgtccacag  tgtggccgtg  gtgacagctt  cagccgccct  caccgggttc  1440
accttctcag  cctgcagat  cctgccctac  acactggcct  ccctctacca  ccggggagaag  1500
caggtgttcc  tgcccaaata  ccgaggggac  actggagggt  ctacagtgga  ggacagcctg  1560
atgaccagct  tctgcccagg  ccctaagcct  ggagctccct  tccctaattg  acacgtgggt  1620
gctggaggca  gtggcctgct  cccacctcca  cccgcgtct  gcggggccct  tgccctgtgat  1680

```

```

gtctccgtac gtgtggtggt gggtagagccc accgaggcca gggtaggttcc gggccggggc 1740
atctgcctgg acctcgccat cctggatagt gccttctctg tgtcccaggt ggccccatcc 1800
ctgtttatgg gctccattgt ccagctcagc cagtctgtca ctgcctatat ggtgtctgcc 1860
gcaggcctgg gtctggtcgc catttacttt gctacacagg tagtatttga caagagcgac 1920
ttggccaaat actcagcgta ggtcagag                                     1947

```

```

<210>      19
<211>      1662
<212>      DNA
<213>      Künstliche Sequenz

```

```

<220>
<223>      Codon-optimiertes humanes P501S

```

```

<400>      19
    atggtgcagc ggctctgggt gagccgcctc ctgcggcatc gcaaggccca gctcctgctg 60
    gtgaatctgc tcacattcgg cctggagggtg tgcctggccg ccggcatcac ctacgtgccc 120
    cccctcctgc tggagggtggg agtcgaggag aagtcatga ccatggtgct gggcattggg 180
    cccgtcctgg gcctcgtgtg cgtgcctctc ctggcgagcg cttccgacca ttggcgcggc 240

    cggtatggcc gcaggagacc cttcatctgg gctctgagtc tcggcatcct gctgagcctg 300
    ttctgatcc ctcgggcccgg ctggctggcc gggtgctgt gccccgatcc tcggcccctg 360
    gagctggccc tgctgacct cggcgtgggc ctgctggact tctgcggcca ggtgtgcttc 420
    acgcccctgg aggcactgct gagcgacctg ttccgggacc ccgaccattg ccgccaggcg 480
    tacagcgtgt acgccttcat gatctccctg ggaggctgcc tgggctacct gctccccgcc 540
    atcgattggg acaccagcgc actcgcccc tatctcggaa cacaggagga atgcctgttc 600
    ggattgttga cgtcatctt cctcacgtgc gtcgcggcca ccctgttggg ggccgaggag 660
    gccgcccctg gggccaccga gccggccgag ggactgagcg ccccgagcct gagtccacac 720
    tgetgccctt gccgggcccg cctggccttc cgtaatctgg gcgccctcct gcctcggtctc 780
    catcagctgt gttgcagaat gcctaggacg ctgcggcgcc tgttcgtcgc tgagtgtgc 840
    tcctggatgg ctctcatgac cttcaccctg ttttatacgg acttcgtcgg ggagggcctg 900
    taccaggggg tgccgcgcgc cgagcccggg acagaggcgc gccgccacta cgacgaggga 960
    gtgcgtatgg gctcccctggg cctcttcttg cagtgcgcca tcagtctggt tttctctctg 1020
    gtcattggaa ggctggtgca gcgcttcgga acccgggcgg tgtacctggc gacgtggcc 1080
    gccttccccg tggtgcccgg cgccacctgc ctctctcact cggtggccgt ggtcaccgcc 1140
    agcgccgccc tgaccgggtt caccctctct gccctgcaga ttctgcctta caccctggcc 1200
    agcctgtacc atcgcgagaa acaggtgttt ctccccaaagt acagaggcga caccgggggc 1260
    gcctccagcg aggacagcct catgacctcc ttctgcctg gccccaaagg ccgccccct 1320
    ttccccaaag ggcacgtggg cgccggcggg agtgggctcc tgcccccccc tcctgcgctg 1380
    tgccggggcca gcgcctgcga cgtgagcgtg cgcgtggtgg tgggcgagcc caccgaggcc 1440
    cgcgtggtgc cgggcagagg catttgtctg gacctggcca tcctcgactc cgccttcttc 1500
    ctcagccagg tggccccgtc cctcttcatg ggctctatcg tccagctgtc tcagagcgtc 1560
    accgcttaca tgggtgtcgc tgetggactg ggcttggtgg ctattttatt cgccaccag 1620
    gtggtgttcg acaagagcga cctggccaaa tactccgcct ga                                     1662

```

```

<210>      20
<211>      1662
<212>      DNA
<213>      Künstliche Sequenz

```

```

<220>
<223>      Codon-optimiertes humanes P501S

```

<400> 20

```

atggtgcagc ggctgtgggt gtccccgctg ctgcgccata gaaaggccca gttgctgctg 60
gtgaacctgc tgactttcgg actggagggtg tgcctggctg ccgggatcac gtacgtgcc 120
cccctgctgc tggagggtggg cgtggaggag aagttcatga caatgggtgct gggcatcggc 180
cccgtcctgg gcctcgtgtg tgtgccccctc ctcgggagtg cgtccgatca ttggcggggc 240
cgctacggcc gccgcagacc gttcatctgg gccctgagcc tggggatcct gctctctctc 300
ttcctgatcc cccgggcccgg ctggctggcc ggctgctgt gtcccagacc ccgccctctg 360
gagctggccc tcctgatcct gggcgtgggc ttgttggaact tctgcggcca ggtgtgtttc 420
actcccctgg aggtctctgt ctccgacctc ttccgcgacc ccgacctg taggcaggct 480
tacagcgtgt acgccttcat gatcagctctg gggggatgcc tgggctatct gctgcccgct 540
atcgactggg acaccagcgc cctggcccc tacctgggga ctcaggagga gtgcctgttc 600
ggcctgctca ccttgatctt cctgacgtgc gtgcgcgcca ccctgctggt ggccgaggag 660
gcggccctgg gggccaccga gccgcgcgag ggctgagcg ctcccagcct gagcccccat 720
tgctgcccggt gcagggctag gctcgccttc aggaatctgg gcgctttgt gccccgctg 780
catcagctgt gctgtcgcgt gccctgcacc ctgcgcgcgc tgttcgtcgc tgagctctgt 840
tcctggatgg ccctgatgac gttaccctc ttctacaccg acttcgtggg ggagggcctg 900
taccaggggc tgcccagggc cgagcccggc accgaggcta ggccgcatca cgacgagggc 960
gtcaggatgg gctctctggg cctcttcctg cagtgcgcca tcagtctggt gttctctctg 1020
gtgatggacc ggctggtgca gcgcttcggc acccgggccc tgtaacctgc ctctgtggcg 1080
gctttccccg tcgccgcccg cgcgacctgc ctgtctcatt ctgtcgcctg ggtgaccgcc 1140
agcgccgccc tgaccggctt caccttcagt gcgtccaga ttctgcccta caccctggcg 1200
tctctgtacc atcgcgagaa gcagggtgtc ctgcccaagt accgcgggga cacaggggga 1260
gcttctctg aggacagcct gatgaccagc ttcttgcccg gcccgaagcc gggggcccct 1320
ttccccaacg gccatgtcgg ggccggcgcc agcggcctgc tccctcccc cccgcctctg 1380
tgcgcgcta gtgcctcgga cgtgagcgtg cgggtgggtg tgggggagcc caccgaggct 1440
agggtcgtgc ctggccgggg gatctgcctg gacctggcca tcctcgactc cgccttctctg 1500
ctctcccagg tggcgcccag cctgttcctg ggcagtatcg tgcagctgag ccagagcgtg 1560
accgcctaca tgggtgagcgc cgcgggcctg ggggtgggtg ccatctactt tgccaccag 1620

gtcgtgttcg acaagagcga tctcgccaag tatagcgct ga 1662

```

<210> 21

<211> 1688

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Codon-optimiertes humanes P501S

<400> 21

```

gacggctagc gccaccatgg tgcagcggct ctgggtgagc cgcctcctgc ggcacgcgaa 60
ggcccagctc ctgctggtga atctgctcac attcggcctg gaggtgtgcc tggccgcccg 120
catcacctac gtgccccccc tcctgctgga ggtgggagtc gaggagaagt tcatgaccat 180
ggtgctgggc attgggcccg tcctgggcct cgtgtgcgtg cctctcctcg gcagcgttc 240
cgaccattgg cgcggccggg atggccgcag gagacccttc atctgggctc tgagtctcgg 300
catcctgctg agcctgttcc tgatccctcg ggccggctgg ctggccgggc tgctgtgcc 360
cgatcctcgg cccctggagc tggccctgct gatcctcggc gtgggcctgc tggacttctg 420
cggccagggt tgcttcacgc ccctggaggc actgctgagc gacctgttcc gggaccccg 480
ccattgccgc caggcgtaca gcgtgtacgc cttcatgatc tcctggggag gctgcctggg 540
ctacctgctc cccgccatcg attgggacac cagcgcactc gcccctatc tcggaacaca 600
ggaggaatgc ctgttcggac tgetgacgct catcttcttc acgtgcgtcg cggccacctc 660
gttggtggcc gaggaggccg ccctggggcc caccgagccg gccgagggac tgagcgcccc 720
gagcctgagt ccacactgct gcccttgccg gcccgcctg gccttcgta atctggcgcg 780
cctcctgcct cggctccatc agctgtgttg cagaatgcct aggacgctgc ggcgctgtt 840
cgctcgtgag ttgtgctcct ggatggctct catgaccttc accctgtttt atacggactt 900
cgtcggggag ggcctgtacc aggggggtgc gcgcgcggag cccgggacag aggcgcgccg 960
ccactacgac gagggagtgc gtatgggctc cctgggcctc ttcttgagc gcgccatcag 1020
tctggttttc tctctggtca tggacaggct ggtgcagcgc ttcggaaccc gggcgggtgta 1080
cctggcgagc gtggccgctt tcccgtggc tgccggcgcc acctgcctct ctactcgggt 1140
ggcctggtc accgccagcg ccgccctgac cgggttcacc ttctctgccc tgcagattct 1200

```

```

gccttacacc ctggccagcc tgtaccatcg cgagaaacag gtgtttctcc ccaagtacag 1260
aggcgacacc gggggcgccct ccagcgagga cagcctcatg acctccttcc tgcctggccc 1320
caagcccggc gcccctttcc ccaacgggca cgtgggcgcc ggcgggagtg ggctcctgcc 1380
ccccctcct gcgctgtgcg gggccagcgc ctgcgacgtg agcgtgcgcg tgggtgggtgg 1440
cgagcccacc gagggccgcg tgggtgccggg cagaggcatt tgtctggacc tggccatcct 1500
cgactccgcc ttctctctca gccagggtgg cccgtccctc ttcatgggct ctatcgtcca 1560
gctgtctcag agcgtcaccg cttacatggt gtccgctgct ggactgggct tgggtggctat 1620
ttatttcgcc acccagggtg tggtcgacaa gagcgacctg gccaaatact ccgcctgact 1680
cgaggcag                                     1688

```

```

<210>      22
<211>      1688
<212>      DNA
<213>      Künstliche Sequenz

<220>
<223>      Codon-optimiertes humanes P501S

```

```

<400>      22

gacggctagc gccaccatgg tgcagcggct gtgggtgtcc cggctgctgc gccatagaaa 60
ggcccagttg ctgctggtga acctgctgac ttctggactg gaggtgtgcc tggctgccgg 120
gatcacgtac gtgccccccc tgctgctgga ggtgggcgtg gaggagaagt tcatgacaat 180
ggtgctgggg atcgcccccg tcctgggcct cgtgtgtgtg cccctcctcg ggagtgcgtc 240
cgatcattgg cggggccgct acggccgccg cagaccgttc atctgggccc tgagcctggg 300
cctcctgctc tctctcttcc tgatcccccg ggccggctgg ctggccggcc tgetgtgtcc 360
cgacccccgc cctctggagc tggccctcct gatcctgggc gtgggcctgc tggacttctg 420
cggccaggtg tgtttcactc ccctggaggc tctgctctcc gacctcttcc gcgaccccca 480
ccactgtagg caggcttaca gcgtgtacgc cttcatgata agtctggggg gatgcctggg 540
ctatctgctg cccgctatcg actgggacac cagcgccctg gccccctacc tggggactca 600
ggaggagtgc ctgttcggcc tgctcacctt gatcttctcg acgtgcgtcg ccgccaccct 660

gctgggtggc gaggaggcgg ccctggggcc caccgagccc gccgagggcc tgagcgctcc 720
cagcctgagc cccatttget gcccgctgag ggctaggctc gccttcagga atctgggcgc 780
tttgctgccc cgcctgcata agctgtgctg tcgcatgcct cgcaccctgc gccgcctgtt 840
cgctcgctgag ctctgttcct ggatggccct gatgacgttc accctcttct acaccgactt 900
cgtggggggg ggccctgtacc agggcgtgcc cagggccgag cccggcaccg aggcctaggcg 960
ccattacgac gagggcgtca ggatgggctc tctgggcctc ttctgacagt gcgccatcag 1020
tctgggtgtt tctctggtga tggaccggct ggtgcagcgc ttcggcaccg gggccgtgta 1080
cctcgccctc gtggcggtt tccccgtcgc cgcggcgcg acctgcctgt ctcatctctg 1140
cgccgtggtg accgccagcg ccgccctgac cggcttcacc ttcagtgcgc tccagattct 1200
gccctacacc ctggcgctct tgtaccatcg cgagaagcag gtgttcctgc ccaagtaccg 1260
cggggacaca gggggagctt cctctgagga cagcctgatg accagcttct tggccggccc 1320
caagccgggg gcccctttcc ccaacggcca tgtcggggcg ggcggcagcg gcctgctccc 1380
tcccccccc gccctgtgcg gcgctagtgc ctgcgacgtg agcgtgcggg tgggtgggtgg 1440
ggagcccacc gaggctaggg tcgtgcctgg ccgggggatc tgccctggacc tggccatcct 1500
cgactccgcc ttctgctct cccagggtgg gccagcctg ttcatgggca gtatcgtgca 1560
gctgagccag agcgtgaccg cctacatggt gagcgccgcc ggccctgggg tgggtggccat 1620
ctactttgcc acccaggctc tggtcgacaa gagcgatctc gccaaagata gcgcctgact 1680
cgaggcag                                     1688

```

```

<210>      23
<211>      435
<212>      DNA
<213>      Künstliche Sequenz

<220>
<223>      Hybridgen zwischen St. pneum. C-LytA, P2 T-Helferepitop
und einem kleinen Teil des 5'-Endes von humanem P501S.

```

<400> 23

```

atggcgcccg cttacgtaca ttccgacggc tcttatccaa aagacaagtt tgagaaaatc 60
aatggcactt ggtactactt tgacagttca ggctatatgc ttgcagaccg ctggaggaag 120
cacacagacg gcaactggta ctggttcgac aactcaggcg aaatggctac aggctggaag 180
aaaatcgctg ataagtggta ctatttcaac gaagaagggtg ccatgaagac aggctgggtc 240
aagtacaagg acacttggta ctacttagac gctaaagaag gcgccatgca atacatcaag 300
gctaactcta agttcattgg tatcactgaa ggcgtcatgg tatcaaagtc ctttatccag 360
tcagcggacg gaacaggctg gtactacctc aaaccagacg gaacactggc agacaggcca 420
gaaaagttca tgtac                                     435

```

<210> 24

<211> 435

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Hybridgen zwischen St. pneum. C-LytA, P2 T-Helferepitop
und und einem kleinen Teil des 5'-Endes von humanem P501S -
Codon-optimiert

<400> 24

```

atggccgccc cctacgtgca tagcgacggg agctacccca aggacaagtt cgagaagatc 60
aacgggacat ggtactactt cgactcctcc ggctacatgc tcgccgaccg ctggcggaag 120
cacaccgacg gcaactggta ctggttcgat aactcgggag agatggccac cggctggaag 180
aagatcgccg acaagtggta ctatttcaac gaggaggggc ccatgaagac cggctgggtg 240
aagtataagg acacctggta ctacctcgac gccaaaggagg gcgccatgca gtatatcaag 300
gccaacagca agttcatcgg catcaccgag ggagtgatgg tcagcaacgc ctttatccag 360
agcgcggacg gcaccggatg gtactacttg aagccggacg gcaccctcgc ggatcgcccc 420
gagaagttca tgtac                                     435

```

<210> 25

<211> 435

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Hybridgen zwischen St. pneum. C-LytA, P2 T-Helferepitop
und einem kleinen Teil des 5'-Endes von humanem P501S -
Codon-optimiert

<400> 25

```

atggccgccc cctacgtgca cagcgacggg tcctacccaa aggacaagtt cgagaagatc 60
aacggcacgt ggtactatct cgacagcagc ggctacatgc tcgccgatcg ctggcgcaag 120
cacaccgacg ggaactggta ctggttcgac aactctggcg agatggctac ggggtggaag 180
aagatcgccc acaagtggta ctacttcaac gaggaggggc ccatgaagac cgggtgggtg 240
aagtacaagg acacctggta ctacctggac gctaaggagg gcgccatgca gtacatcaag 300
gccaactcga agttcatcgg gatcaccgag ggcgtgatgg tcagtaacgc tttcatccag 360
agcgcggacg gcacaggctg gtattacctg aagcccgatg gcaccctggc ggacagacct 420
gagaaattca tgtac                                     435

```

<210> 26

<211> 464

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Hybridgen zwischen St. pneum. C-LytA, P2 T-Helferepitop
und einem kleinen Teil des 5'-Endes von humanem P501S -
Codon-optimiert

<400> 26

```

gacgggctagc gccaccatgg ccgcccgccta cgtgcatagc gacggggagct accccaagga 60
caagtctcgag aagatcaacg ggacatggta ctacttcgac tcctccgggt acatgctcgc 120
cgaccgctgg cggaagcaca ccgacggcaa ctggtactgg ttcgataact cgggagagat 180
ggccaccggc tggagaaga tcgcggacaa gtggtactat ttcaacgagg agggcgccat 240
gaagaccggc tgggtgaagt ataaggacac ctggtactac ctcgacgcca aggagggcgc 300
catgcagtat atcaaggcca acagcaagtt catcggcac accgagggag tgatggtcag 360
caacgccttt atccagagcg ccgacggcac cggatgggtac tacttgaagc cggacggcac 420
cctcgcggat cggcccgcaga agttcatgta ctgactcgag gcag 464

```

<210> 27

<211> 652

<212> PRT

<213> Künstliche Sequenz

<220>

<223> Hybridgen zwischen St. pneum. C-LytA, P2 T-Helferepitop
und Aminosäuren 51-553 von humanem P501S

<400> 27

```

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

```

Gln	Tyr	Ile	Lys	Ala	Asn	Ser	Lys	Phe	Ile	Gly	Ile	Thr	Glu	Gly	Val	100	105	110
Met	Val	Ser	Asn	Ala	Phe	Ile	Gln	Ser	Ala	Asp	Gly	Thr	Gly	Trp	Tyr	115	120	125
Tyr	Leu	Lys	Pro	Asp	Gly	Thr	Leu	Ala	Asp	Arg	Pro	Glu	Lys	Phe	Met	130	135	140
Tyr	Met	Val	Leu	Gly	Ile	Gly	Pro	Val	Leu	Gly	Leu	Val	Cys	Val	Pro	145	150	155
Leu	Leu	Gly	Ser	Ala	Ser	Asp	His	Trp	Arg	Gly	Arg	Tyr	Gly	Arg	Arg	165	170	175
Arg	Pro	Phe	Ile	Trp	Ala	Leu	Ser	Leu	Gly	Ile	Leu	Leu	Ser	Leu	Phe	180	185	190
Leu	Ile	Pro	Arg	Ala	Gly	Trp	Leu	Ala	Gly	Leu	Leu	Cys	Pro	Asp	Pro	195	200	205
Arg	Pro	Leu	Glu	Leu	Ala	Leu	Leu	Ile	Leu	Gly	Val	Gly	Leu	Leu	Asp	210	215	220
Phe	Cys	Gly	Gln	Val	Cys	Phe	Thr	Pro	Leu	Glu	Ala	Leu	Leu	Ser	Asp	225	230	235
Leu	Phe	Arg	Asp	Pro	Asp	His	Cys	Arg	Gln	Ala	Tyr	Ser	Val	Tyr	Ala	245	250	255
Phe	Met	Ile	Ser	Leu	Gly	Gly	Cys	Leu	Gly	Tyr	Leu	Leu	Pro	Ala	Ile	260	265	270
Asp	Trp	Asp	Thr	Ser	Ala	Leu	Ala	Pro	Tyr	Leu	Gly	Thr	Gln	Glu	Glu	275	280	285
Cys	Leu	Phe	Gly	Leu	Leu	Thr	Leu	Ile	Phe	Leu	Thr	Cys	Val	Ala	Ala	290	295	300
Thr	Leu	Leu	Val	Ala	Glu	Glu	Ala	Ala	Leu	Gly	Pro	Thr	Glu	Pro	Ala	305	310	315
Glu	Gly	Leu	Ser	Ala	Pro	Ser	Leu	Ser	Pro	His	Cys	Cys	Pro	Cys	Arg	325	330	335
Ala	Arg	Leu	Ala	Phe	Arg	Asn	Leu	Gly	Ala	Leu	Leu	Pro	Arg	Leu	His	340	345	350
Gln	Leu	Cys	Cys	Arg	Met	Pro	Arg	Thr	Leu	Arg	Arg	Leu	Phe	Val	Ala	355	360	365
Glu	Leu	Cys	Ser	Trp	Met	Ala	Leu	Met	Thr	Phe	Thr	Leu	Phe	Tyr	Thr	370	375	380
Asp	Phe	Val	Gly	Glu	Gly	Leu	Tyr	Gln	Gly	Val	Pro	Arg	Ala	Glu	Pro	385	390	395
Gly	Thr	Glu	Ala	Arg	His	Tyr	Asp	Glu	Gly	Val	Arg	Met	Gly	Ser		405	410	415
Leu	Gly	Leu	Phe	Leu	Gln	Cys	Ala	Ile	Ser	Leu	Val	Phe	Ser	Leu	Val	420	425	430
Met	Asp	Arg	Leu	Val	Gln	Arg	Phe	Gly	Thr	Arg	Ala	Val	Tyr	Leu	Ala	435	440	445
Ser	Val	Ala	Ala	Phe	Pro	Val	Ala	Ala	Gly	Ala	Thr	Cys	Leu	Ser	His	450	455	460
Ser	Val	Ala	Val	Val	Thr	Ala	Ser	Ala	Ala	Leu	Thr	Gly	Phe	Thr	Phe	465	470	475
Ser	Ala	Leu	Gln	Ile	Leu	Pro	Tyr	Thr	Leu	Ala	Ser	Leu	Tyr	His	Arg	485	490	495
Glu	Lys	Gln	Val	Phe	Leu	Pro	Lys	Tyr	Arg	Gly	Asp	Thr	Gly	Gly	Ala	500	505	510
Ser	Ser	Glu	Asp	Ser	Leu	Met	Thr	Ser	Phe	Leu	Pro	Gly	Pro	Lys	Pro	515	520	525
Gly	Ala	Pro	Phe	Pro	Asn	Gly	His	Val	Gly	Ala	Gly	Gly	Ser	Gly	Leu	530	535	540
Leu	Pro	Pro	Pro	Pro	Ala	Leu	Cys	Gly	Ala	Ser	Ala	Cys	Asp	Val	Ser	545	550	555
Val	Arg	Val	Val	Val	Gly	Glu	Pro	Thr	Glu	Ala	Arg	Val	Val	Pro	Gly	565	570	575
Arg	Gly	Ile	Cys	Leu	Asp	Leu	Ala	Ile	Leu	Asp	Ser	Ala	Phe	Leu	Leu			

	580		585		590										
Ser	Gln	Val	Ala	Pro	Ser	Leu	Phe	Met	Gly	Ser	Ile	Val	Gln	Leu	Ser
	595		600		605										
Gln	Ser	Val	Thr	Ala	Tyr	Met	Val	Ser	Ala	Ala	Gly	Leu	Gly	Leu	Val
	610		615		620										
Ala	Ile	Tyr	Phe	Ala	Thr	Gln	Val	Val	Phe	Asp	Lys	Ser	Asp	Leu	Ala
	625		630		635									640	
Lys	Tyr	Ser	Ala	Gly	Gly	His	His	His	His	His	His				
	645		650												

<210> 28
 <211> 1959
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> DNA, die das Hybridprotein zwischen St. pneum. C-LytA,
 P2 T-Helferepitop und Aminosäuren 51-553 von humanem P501S
 codiert

<400> 28

```

atggcggccg cttacgtaca ttccgacggc tcttatccaa aagacaagtt tgagaaaatc 60
aatggcactt ggtactactt tgacagttca ggctatatgc ttgcagaccg ctggaggaag 120
cacacagacg gcaactggta ctgggtcgac aactcaggcg aaatggctac aggctggaag 180
aaaatcgctg ataagtggta ctatttcaac gaagaagggtg ccatgaagac aggctgggtc 240
aagtacaagg acacttggtg ctacttagac gctaaagaag gcgccatgca atacatcaag 300
gctaactcta agttcattgg tatcactgaa ggcgtcatgg tatcaaagtc ctttatccag 360
tcagcggacg gaacaggctg gtactacctc aaaccagacg gaacactggc agacaggcca 420
gaaaagttca tgtacatggg gctgggcatt ggtccagtgc tgggcctggg ctgtgtcccg 480
ctcctaggct cagccagtga ccactggcgt ggacgctatg gccgccggcg gcccttcac 540
tgggcactgt ccttgggcat cctgctgagc ctctttctca tcccaagggc cggctgggta 600
gcagggctgc tgtgcccggg tcccaggccc ctggagctgg cactgctcat cctgggcgtg 660
gggctgtctg acttctgtgg ccagggtgtg ttactccac tggaggccct gctctctgac 720
ctcttccggg acccggaacca ctgtcgccag gcctactctg tctatgcctt catgatcagt 780
cttgggggct gcctgggcta cctcctgcct gccattgact gggacaccag tgccctggcc 840
ccctacctgg gcacccagga ggagtgcctc tttggcctgc tcacctcat cttcctcacc 900
tgcgtagcag ccacactgct ggtggctgag gaggcagcgc tgggccccac cgagccagca 960
gaagggctgt cggccccctc cttgtcgccc cactgctgtc catgccgggc ccgcttggct 1020
ttccggaacc tgggcgcctt gcttccccgg ctgcaccagc tgtgctgccg catgccccgc 1080
accctgcgcc ggctcttcgt ggctgagctg tgcagctgga tggcactcat gaccttcacg 1140
ctgttttaca cggatttcgt gggcgagggg ctgtaccagg gcgtgcccag agctgagccg 1200
ggcaccgagg cccggagaca ctatgatgaa ggcgttcgga tgggcagcct ggggctgttc 1260
ctgcagtgcg ccatctccct ggtcttctct ctggtcattg accggtggtg gcagcgattc 1320
ggcactcgag cagtctatct ggccagtgtg gcagctttcc ctgtggctgc cgggtgccaca 1380
tgccctgtccc acagtgtggc cgtggtgaca gcttcagccg ccctcaccgg gttcaccttc 1440
tcagccctgc agatccctgc ctacacactg gcctccctct accaccggga gaagcaggtg 1500
ttcctgcccc aataccgagg ggacactgga ggtgctagca gtgaggacag cctgatgacc 1560
agcttccctg caggccctaa gcctggagct ccctcccta atggacacgt ggggtgctgga 1620
ggcagtgggc tgcctccacc tccacccgcg cctctgcctg tgatgtctcc 1680
gtacgtgtgg tgggtgggtga gcccaccgag gccagggtgg ttccggggcg gggcatctgc 1740
ctggacctcg ccatcctgga tagtgcttc ctgctgtccc aggtggcccc atccctgttt 1800
atgggctcca ttgtccagct cagccagtct gtcactgcct atatggtgtc tgccgcaggc 1860
ctgggtctgg tcgccattta ctttgctaca caggtagtat ttgacaagag cgacttggcc 1920
aaatactcag cgggtggaca ccatcaccat caccattaa 1959

```

<210> 29
 <211> 507
 <212> PRT
 <213> Künstliche Sequenz

<220>

<223> Humanes P501S (Aminosäuren 55-553), fusioniert an
6 Histidinreste

<400> 29

```

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

```

Gln Val Ala Pro Ser Leu Phe Met Gly Ser Ile Val Gln Leu Ser Gln
 450 455 460
 Ser Val Thr Ala Tyr Met Val Ser Ala Ala Gly Leu Gly Leu Val Ala
 465 470 475 480
 Ile Tyr Phe Ala Thr Gln Val Val Phe Asp Lys Ser Asp Leu Ala Lys
 485 490 495
 Tyr Ser Ala Gly Gly His His His His His His
 500 505

<210> 30
 <211> 1524
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> DNA, die humanes P501S (Aminosäuren 55-553), fusioniert
 an 6 Histidinreste, codiert

<400> 30
 atggtgctgg gcattggtcc agtgctgggc ctgggtctgtg tcccgtctct aggetcagcc 60
 agtgaccact ggcgtggacg ctatggccgc cgccggccct tcatctgggc actgtccttg 120
 ggcatacctgc tgagcctctt tctcatccca agggccggct ggctagcagg gctgctgtgc 180
 ccgcatccca ggcccctgga gctggcactg ctcatcctgg gcgtggggct gctggacttc 240
 tgtggccagg tgtgtctcac tccactggag gccctgctct ctgacctctt ccgggacccg 300
 gaccactgtc gccaggccta ctctgtctat gccttcatga tcagtcttgg gggctgcctg 360
 ggctacctcc tgccctgcat tgactgggac accagtgcc tggcccccta cctgggcacc 420
 caggaggagt gcctctttgg cctgtctacc ctcatcttcc tcacctgcgt agcagccaca 480
 ctgctgggtgg ctgaggaggc agcgtgggc cccaccgagc cagcagaagg gctgtcggcc 540
 cctccttctg cgcctcactg ctgtccatgc cgggcccgtt tggctttccg gaacctgggc 600
 gccctgcttc cccggctgca ccagctgtgc tgccgcatgc cccgcacct gcgcgggctc 660
 ttcgtggctg agctgtgcag ctggatggca ctcatgacct tcacgctgtt ttacacggat 720
 ttcgtgggag aggggctgta ccagggcgtg cccagagctg agccgggcac cgaggcccg 780
 agacactatg atgaaggcgt tcggatgggc agcctggggc tgctcctgca gtgcgccatc 840
 tccctggctc tctctctggt catggaccgg ctgggtgcag gattcggcac tcgagcagtc 900
 tatttggcca gtgtggcagc ttccctgtg gctgcccgtg ccacatgcct gtcccacagt 960
 gtggccgtgg tgacagcttc agccgccctc accgggttca ccttctcagc cctgcagatc 1020
 ctgccctaca cactggcctc cctctaccac cgggagaagc aggtgttctt gcccaaatac 1080
 cgaggggaca ctggagggtg tagcagttag gacagcctga tgaccagctt cctgccaggc 1140
 cctaagcctg gagctccctt ccctaattga cacgtgggtg ctggaggcag tggcctgctc 1200
 ccacctccac ccgcgtctg cggggcctct gcctgtgatg tctccgtacg tgtgggtggtg 1260
 ggtgagccca ccgaggccag ggtggttccg ggccggggca tctgcctgga cctcgccatc 1320
 ctggatagtg ccttcctgct gtcccagggt gccccatccc tgtttatggg ctccattgtc 1380
 cagctcagcc agtctgtcac tgccatatat gtgtctgccg caggcctggg tctggctgcc 1440
 atttactttg ctacacaggt agtatttgac aagagcgact tggccaaata ctacgcgggt 1500
 ggacaccatc accatcacca ttaa 1524

<210> 31
 <211> 685
 <212> PRT
 <213> Künstliche Sequenz

<220>
 <223> Humanes P501S (Aminosäuren 1-34, fusioniert an 55-553),
 fusioniert an 6 Histidinreste

<400> 31
 Met Ala Ala Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg
 1 5 10 15
 Lys Ala Gln Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val
 20 25 30

Cys Leu Ala Ala Ala Tyr Val His Ser Asp Gly Ser Tyr Pro Lys Asp
 35 40 45
 Lys Phe Glu Lys Ile Asn Gly Thr Trp Tyr Tyr Phe Asp Ser Ser Gly
 50 55 60
 Tyr Met Leu Ala Asp Arg Trp Arg Lys His Thr Asp Gly Asn Trp Tyr
 65 70 75 80
 Trp Phe Asp Asn Ser Gly Glu Met Ala Thr Gly Trp Lys Lys Ile Ala
 85 90 95
 Asp Lys Trp Tyr Tyr Phe Asn Glu Glu Gly Ala Met Lys Thr Gly Trp
 100 105 110
 Val Lys Tyr Lys Asp Thr Trp Tyr Tyr Leu Asp Ala Lys Glu Gly Ala
 115 120 125
 Met Gln Tyr Ile Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu Gly
 130 135 140
 Val Met Val Ser Asn Ala Phe Ile Gln Ser Ala Asp Gly Thr Gly Trp
 145 150 155 160
 Tyr Tyr Leu Lys Pro Asp Gly Thr Leu Ala Asp Arg Pro Glu Lys Phe
 165 170 175
 Met Tyr Met Val Leu Gly Ile Gly Pro Val Leu Gly Leu Val Cys Val
 180 185 190
 Pro Leu Leu Gly Ser Ala Ser Asp His Trp Arg Gly Arg Tyr Gly Arg
 195 200 205
 Arg Arg Pro Phe Ile Trp Ala Leu Ser Leu Gly Ile Leu Leu Ser Leu
 210 215 220
 Phe Leu Ile Pro Arg Ala Gly Trp Leu Ala Gly Leu Leu Cys Pro Asp
 225 230 235 240
 Pro Arg Pro Leu Glu Leu Ala Leu Leu Ile Leu Gly Val Gly Leu Leu
 245 250 255
 Asp Phe Cys Gly Gln Val Cys Phe Thr Pro Leu Glu Ala Leu Leu Ser
 260 265 270
 Asp Leu Phe Arg Asp Pro Asp His Cys Arg Gln Ala Tyr Ser Val Tyr
 275 280 285
 Ala Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr Leu Leu Pro Ala
 290 295 300
 Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu Gly Thr Gln Glu
 305 310 315 320
 Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu Thr Cys Val Ala
 325 330 335
 Ala Thr Leu Leu Val Ala Glu Glu Ala Ala Leu Gly Pro Thr Glu Pro
 340 345 350
 Ala Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His Cys Cys Pro Cys
 355 360 365
 Arg Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala Leu Leu Pro Arg Leu
 370 375 380
 His Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg Arg Leu Phe Val
 385 390 395 400
 Ala Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe Thr Leu Phe Tyr
 405 410 415
 Thr Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val Pro Arg Ala Glu
 420 425 430
 Pro Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly Val Arg Met Gly
 435 440 445
 Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu Val Phe Ser Leu
 450 455 460
 Val Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu
 465 470 475 480
 Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys Leu Ser
 485 490 495
 His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly Phe Thr
 500 505 510
 Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu Tyr His

515	520	525
Arg Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly Asp Thr Gly Gly		
530	535	540
Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu Pro Gly Pro Lys		
545	550	555
Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly Ser Gly		560
	565	570
Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala Cys Asp Val		575
	580	585
Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala Arg Val Val Pro		590
	595	600
Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp Ser Ala Phe Leu		605
610	615	620
Leu Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser Ile Val Gln Leu		
625	630	635
Ser Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala Gly Leu Gly Leu		640
	645	650
Val Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp Lys Ser Asp Leu		655
	660	665
Ala Lys Tyr Ser Ala Gly Gly His His His His His His		670
675	680	685

<210> 32
 <211> 2058
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> DNA, die humanes P501S (Aminosäuren 1-34, fusioniert an 55-553), fusioniert an 6 Histidinrest, codiert

<400> 32

```

atggcgggcgc tgcagaggct atgggtatcg agactgctaa gacaccgcaa agctcagttg 60
ttgttgggtta acttgttgac cttcgggctg gaagtctgtt tggcgggcgc ttactgtacat 120
tccgacggcct cttatccaaa agacaagttt gagaaaatca atggcacttg gtactacttt 180
gacagttcag gctatatgct tgcagaccgc tggaggaagc acacagacgg caactgggtac 240
tgggttcgaca actcaggcga aatggctaca ggctggaaga aaatcgctga taagtgggtac 300
tatttcaacg aagaagggtgc catgaagaca ggctgggtca agtacaagga cacttgggtac 360
tacttagacg ctaaagaagg cgccatgcaa tacatcaagg ctaactctaa gttcattgggt 420
atcactgaag gcgtcatggt atcaaatgcc tttatccagt cagcggacgg aacaggctgg 480
tactacctca aaccagacgg aacactggca gacaggccag aaaagtccat gtacatgggtg 540
ctgggcattg gtccagtgtc gggcctggtc tgtgtcccgc tcctaggctc agccagtgtg 600
cactggcggtg gacgttatgg ccgccggcgg cccttcattc gggcactgtc cttggggcatc 660
ctgctgagcc tctttctcat cccaagggcc ggctgggctag cagggtgtgt gtgcccggat 720
cccaggcccc tggagctggc actgtctatc ctgggcgtgg ggctgtgtga cttctgtggc 780
cagggtgtgt tcaactccat ggaggccctg ctctctgacc tcttccggga cccgggaccac 840
tgtcgccagg cctactctgt ctatgccttc atgatcagtc ttgggggctg cctggggctac 900
ctcctgcctg ccattgactg ggacaccagt gccctggccc cctacctggg caccaggag 960
gagtgcctct ttggcctgct caccctcatc ttcctcacct gcgtagcagc cacactgtgt 1020
gtggctgagg aggcagcgtt gggccccacc gagccagcag aagggtgtgt gggccccctc 1080
ttgtcgcccc actgtgttcc atgcccgggc cgcttggttt tccggaacct gggcgccctg 1140
cttccccggc tgcaccagct gtgctgccc atgccccgca ccctgcgccg gctcttctgt 1200
gctgagctgt gcagctggat ggcactcatg accttcacgc tgttttacac ggatttctgt 1260
ggcgagggggc tgtaccaggc cgtgccaga gctgagccgg gcaccgaggg ccggagacac 1320
tatgatgaag gcgttcggat gggcagcctg gggctgttcc tgcagtgcgc catctccctg 1380
gtcttctctc tgggtcatgga ccggctgggt cagcgattcg gcactcgagc agtctatttg 1440
gccagtgtgg cagctttccc tgtggctgcc ggtgccacat gcctgtccca cagtgtggcc 1500
gtggtgacag cttcagccgc cctaccggg ttcaccttct cagccctgca gatcctgccc 1560
tacacactgg cctccctcta ccaccgggag aagcaggtgt tcctgcccc ataccgaggg 1620
gacactggag gtgctagcag tgaggacagc ctgatgacca gcttccctgcc aggccctaag 1680

```

```

cctggagctc ccttccttaa tggacacgtg ggtgctggag gcagtggcct gctccacact 1740
ccacccgcgc tctgcggggc ctctgcctgt gatgtctccg tacgtgtggt ggtgggtgag 1800
cccaccgagg ccaggggtggt tccggggcgg ggcattctgcc tggacctcgc catcctggat 1860
agtgccttcc tgctgtccca ggtggcccca tccctgttta tgggctccat tgtccagctc 1920
agccagtctg tcaactgccta tatgggtgtct gccgcaggcc tgggtctggt cgccatttac 1980
tttgctacac aggtagtatt tgacaagagc gacttggcca aatactcagc ggggtggacac 2040
catcaccatc accattaa                                     2058

```

<210> 33
 <211> 671
 <212> PRT
 <213> Künstliche Sequenz

<220>
 <223> St. pneum. C.-LytA-Teil, fusioniert an P2 T-Helferepitop,
 fusioniert an humanes P501S (Aminosäuren 55-553), fusioniert
 an 6 Histidinreste stromabwärts von Hefe-alpha-präpro-
 Signalsequenz

<400> 33

```

Met Ala Ala Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala
 1           5           10           15
Ser Ser Ala Leu Ala Ala Tyr Val His Ser Asp Gly Ser Tyr Pro
          20          25          30
Lys Asp Lys Phe Glu Lys Ile Asn Gly Thr Trp Tyr Tyr Phe Asp Ser
        35          40          45
Ser Gly Tyr Met Leu Ala Asp Arg Trp Arg Lys His Thr Asp Gly Asn
       50          55          60
Trp Tyr Trp Phe Asp Asn Ser Gly Glu Met Ala Thr Gly Trp Lys Lys
      65          70          75          80
Ile Ala Asp Lys Trp Tyr Tyr Phe Asn Glu Glu Gly Ala Met Lys Thr
          85          90          95
Gly Trp Val Lys Tyr Lys Asp Thr Trp Tyr Tyr Leu Asp Ala Lys Glu
        100         105         110
Gly Ala Met Gln Tyr Ile Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr
       115         120         125
Glu Gly Val Met Val Ser Asn Ala Phe Ile Gln Ser Ala Asp Gly Thr
      130         135         140
Gly Trp Tyr Tyr Leu Lys Pro Asp Gly Thr Leu Ala Asp Arg Pro Glu
     145         150         155         160
Lys Phe Met Tyr Met Val Leu Gly Ile Gly Pro Val Leu Gly Leu Val
          165         170         175
Cys Val Pro Leu Leu Gly Ser Ala Ser Asp His Trp Arg Gly Arg Tyr
        180         185         190
Gly Arg Arg Arg Pro Phe Ile Trp Ala Leu Ser Leu Gly Ile Leu Leu
       195         200         205
Ser Leu Phe Leu Ile Pro Arg Ala Gly Trp Leu Ala Gly Leu Leu Cys
      210         215         220
Pro Asp Pro Arg Pro Leu Glu Leu Ala Leu Leu Ile Leu Gly Val Gly
     225         230         235         240
Leu Leu Asp Phe Cys Gly Gln Val Cys Phe Thr Pro Leu Glu Ala Leu
          245         250         255
Leu Ser Asp Leu Phe Arg Asp Pro Asp His Cys Arg Gln Ala Tyr Ser
        260         265         270
Val Tyr Ala Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr Leu Leu
       275         280         285
Pro Ala Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu Gly Thr
      290         295         300
Gln Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu Thr Cys
     305         310         315         320
Val Ala Ala Thr Leu Leu Val Ala Glu Glu Ala Ala Leu Gly Pro Thr

```

```

      325      330      335
Glu Pro Ala Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His Cys Cys
      340      345      350
Pro Cys Arg Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala Leu Leu Pro
      355      360      365
Arg Leu His Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg Arg Leu
      370      375      380
Phe Val Ala Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe Thr Leu
385      390      395      400
Phe Tyr Thr Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val Pro Arg
      405      410      415
Ala Glu Pro Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly Val Arg
      420      425      430
Met Gly Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu Val Phe
      435      440      445
Ser Leu Val Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val
      450      455      460
Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys
465      470      475      480
Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly
      485      490      495
Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu
      500      505      510
Tyr His Arg Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly Asp Thr
      515      520      525
Gly Gly Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu Pro Gly
      530      535      540
Pro Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly
545      550      555      560
Ser Gly Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala Cys
      565      570      575
Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala Arg Val
      580      585      590
Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp Ser Ala
      595      600      605
Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser Ile Val
      610      615      620
Gln Leu Ser Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala Gly Leu
625      630      635      640
Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp Lys Ser
      645      650      655
Asp Leu Ala Lys Tyr Ser Ala Gly Gly His His His His His His
      660      665      670

```

<210> 34
 <211> 2477
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> DNA, die St. pneum. C-LytA-Teil codiert, fusioniert an P2
 T-Helferepitop, fusioniert an humanes P501S (Aminosäuren
 55-553), fusioniert an 6 Histidinreste stromabwärts der Hefe-
 alpha-präpro-Signalsequenz

<400> 34
 tacgtacatt cgcacggctc ttatccaaaa gacaagtttg agaaaatcaa tggcacttgg 60
 tactactttg acagttcagg ctatatgctt gcagaccgct ggaggaagca cacagacggc 120
 aactggtact ggttcgacaa ctcaggcgaa atggctacag gctggaagaa aatcgctgat 180
 aagtgttact atttcaacga agaaggtgcc atgaagacag gctgggtcaa gtacaaggac 240

```

acttggtact acttagacgc taaagaaggg gccatgcaat acatcaaggc taactctaag 300
ttcattggta tcaactgaagg cgtcatggta tcaaatgcct ttatccagtc agcggacgga 360
acaggctggg actacctcaa accagacgga acactggcag acaggccaga aatggcggcc 420
agatttcctt caatttttac tgcagtttta ttgcagcat cctccgcatt agcggccgct 480
tacgtacatt ccgacgggtc ttatccaaaa gacaagtttg agaaaatcaa tggcacttgg 540
tactactttg acagttcagg ctatatgctt gcagaccgct ggaggaagca cacagacggc 600
aactgggtact ggttcgacaa ctcaggcgaa atggctacag gctggaagaa aatcgctgat 660
aagtgggtact atttcaacga agaaggtgcc atgaagacag gctgggtcaa gtacaaggac 720
acttggtact acttagacgc taaagaaggg gccatgcaat acatcaaggc taactctaag 780
ttcattggta tcaactgaagg cgtcatggta tcaaatgcct ttatccagtc agcggacgga 840
acaggctggg actacctcaa accagacgga acactggcag acaggccaga agctgggtatt 900
acttacgttc caccattgtt gttggaagtt ggtgttgaag aaaagttcat gtacatgggtg 960
ctgggcattg gtccagtgtc gggcctgggtc tgtgtccgc tcctaggctc agccagtgtg 1020
cactggcggtg gacgctatgg ccgccgcgg cccttcattt gggcactgtc cttgggcatc 1080
ctgctgagcc tctttctcat cccaagggcc ggctggctag cagggtgtgt gtgcccggat 1140
cccaggcccc tggagctggc actgtctatc ctgggcgtgg ggtgtgtgga cttctgtggc 1200
caggtgtgtc tcaactccact ggaggccctg ctctctgacc tcttccggga cccggaccac 1260
tgtcgccagg cctactctgt ctatgcttca tgatcagtct tgggggctgc ctgggctacc 1320
tcctgcctgc cattgactgg gacaccagtg ccctggcccc ctacctgggc acccaggagg 1380
agtgcctctt tggcctgtc accctcatct tcctcacctg cgtagcagcc acactgtgtg 1440
tggctgagga ggcagcgtg ggccccaccg agccagcaga agggctgtcg gccccctct 1500
tgtcgcccca ctgctgtcca tgccggggccc gcttggcttt ccggaacctg ggcgcctgc 1560
ttccccgggt gcaccagtg tgctgccgca tgccccgcac cctgcgccgg ctcttcgtgg 1620
ctgagctgtg cagctggatg gcactcatga cttcacgtt gttttacacg gatttcgtgg 1680
gcgagggggc gtaccagggc gtgccagag ctgagccggg caccgaggcc cgagacact 1740
atgatgaagg cgttcggatg ggcagcctgg ggctgttctt gcagtgcgcc atctccctgg 1800
tcttctctct ggcatggac cggctgggtc agcgattcgg cactcgagca gtctatttgg 1860
ccagtgtggc agctttccct gtggctgccg gtgccacatg cctgtccac agtgtggccg 1920
tggtgacagc ttcagccgcc ctcaccgggt tcacctctc agccctgcag atcctgcct 1980
acacactggc ctccctctac caccgggaga agcaggtgtt cctgcccata taccgagggg 2040
acactggagg tgctagcagt gaggacagcc tgatgaccag cttcctgcca ggccctaagc 2100
ctggagctcc ctccctaata ggacacgtgg gtgctggagg cagtggcctg ctccacctc 2160
caccgcgct ctgccccggc tctgcctgtg atgtctccgt acgtgtggtg gtgggtgagc 2220
ccaccgaggc caggggtggt ccggggccgg gcactctgct ggacctgcc atcctggata 2280
gtgccttctt gctgtcccag gtggccccat ccctgtttat gggctccatt gtccagctca 2340
gccagtctgt cactgcctat atggtgtctg ccgcaggcct gggctctggt gccatttact 2400
ttgctacaca ggtagtattt gacaagagcg acttggccaa atactcagcg ggtggacacc 2460
atcaccatca ccattaa
2477

```

<210> 35
 <211> 595
 <212> PRT
 <213> Künstliche Sequenz

<220>
 <223> Humanes P501S (Aminosäuren 55-553), fusioniert an 6 Histidin-
 reste stromabwärts der Hefe-alpha-präpro-Signalsequenz

<400> 35

```

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

```

85					90					95					
Leu	Gly	Leu	Val	Cys	Val	Pro	Leu	Leu	Gly	Ser	Ala	Ser	Asp	His	Trp
100					105					110					
Arg	Gly	Arg	Tyr	Gly	Arg	Arg	Arg	Pro	Phe	Ile	Trp	Ala	Leu	Ser	Leu
115					120					125					
Gly	Ile	Leu	Leu	Ser	Leu	Phe	Leu	Ile	Pro	Arg	Ala	Gly	Trp	Leu	Ala
130					135					140					
Gly	Leu	Leu	Cys	Pro	Asp	Pro	Arg	Pro	Leu	Glu	Leu	Ala	Leu	Leu	Ile
145					150					155					
Leu	Gly	Val	Gly	Leu	Leu	Asp	Phe	Cys	Gly	Gln	Val	Cys	Phe	Thr	Pro
165					170					175					
Leu	Glu	Ala	Leu	Leu	Ser	Asp	Leu	Phe	Arg	Asp	Pro	Asp	His	Cys	Arg
180					185					190					
Gln	Ala	Tyr	Ser	Val	Tyr	Ala	Phe	Met	Ile	Ser	Leu	Gly	Gly	Cys	Leu
195					200					205					
Gly	Tyr	Leu	Leu	Pro	Ala	Ile	Asp	Trp	Asp	Thr	Ser	Ala	Leu	Ala	Pro
210					215					220					
Tyr	Leu	Gly	Thr	Gln	Glu	Glu	Cys	Leu	Phe	Gly	Leu	Leu	Thr	Leu	Ile
225					230					235					
Phe	Leu	Thr	Cys	Val	Ala	Ala	Thr	Leu	Leu	Val	Ala	Glu	Glu	Ala	Ala
245					250					255					
Leu	Gly	Pro	Thr	Glu	Pro	Ala	Glu	Gly	Leu	Ser	Ala	Pro	Ser	Leu	Ser
260					265					270					
Pro	His	Cys	Cys	Pro	Cys	Arg	Ala	Arg	Leu	Ala	Phe	Arg	Asn	Leu	Gly
275					280					285					
Ala	Leu	Leu	Pro	Arg	Leu	His	Gln	Leu	Cys	Cys	Arg	Met	Pro	Arg	Thr
290					295					300					
Leu	Arg	Arg	Leu	Phe	Val	Ala	Glu	Leu	Cys	Ser	Trp	Met	Ala	Leu	Met
305					310					315					
Thr	Phe	Thr	Leu	Phe	Tyr	Thr	Asp	Phe	Val	Gly	Glu	Gly	Leu	Tyr	Gln
325					330					335					
Gly	Val	Pro	Arg	Ala	Glu	Pro	Gly	Thr	Glu	Ala	Arg	Arg	His	Tyr	Asp
340					345					350					
Glu	Gly	Val	Arg	Met	Gly	Ser	Leu	Gly	Leu	Phe	Leu	Gln	Cys	Ala	Ile
355					360					365					
Ser	Leu	Val	Phe	Ser	Leu	Val	Met	Asp	Arg	Leu	Val	Gln	Arg	Phe	Gly
370					375					380					
Thr	Arg	Ala	Val	Tyr	Leu	Ala	Ser	Val	Ala	Ala	Phe	Pro	Val	Ala	Ala
385					390					395					
Gly	Ala	Thr	Cys	Leu	Ser	His	Ser	Val	Ala	Val	Val	Thr	Ala	Ser	Ala
405					410					415					
Ala	Leu	Thr	Gly	Phe	Thr	Phe	Ser	Ala	Leu	Gln	Ile	Leu	Pro	Tyr	Thr
420					425					430					
Leu	Ala	Ser	Leu	Tyr	His	Arg	Glu	Lys	Gln	Val	Phe	Leu	Pro	Lys	Tyr
435					440					445					
Arg	Gly	Asp	Thr	Gly	Gly	Ala	Ser	Ser	Glu	Asp	Ser	Leu	Met	Thr	Ser
450					455					460					
Phe	Leu	Pro	Gly	Pro	Lys	Pro	Gly	Ala	Pro	Phe	Pro	Asn	Gly	His	Val
465					470					475					
Gly	Ala	Gly	Gly	Ser	Gly	Leu	Leu	Pro	Pro	Pro	Pro	Ala	Leu	Cys	Gly
485					490					495					
Ala	Ser	Ala	Cys	Asp	Val	Ser	Val	Arg	Val	Val	Val	Gly	Glu	Pro	Thr
500					505					510					
Glu	Ala	Arg	Val	Val	Pro	Gly	Arg	Gly	Ile	Cys	Leu	Asp	Leu	Ala	Ile
515					520					525					
Leu	Asp	Ser	Ala	Phe	Leu	Leu	Ser	Gln	Val	Ala	Pro	Ser	Leu	Phe	Met
530					535					540					
Gly	Ser	Ile	Val	Gln	Leu	Ser	Gln	Ser	Val	Thr	Ala	Tyr	Met	Val	Ser
545					550					555					
Ala	Ala	Gly	Leu	Gly	Leu	Val	Ala	Ile	Tyr	Phe	Ala	Thr	Gln	Val	Val
565					570					575					

Phe Asp Lys Ser Asp Leu Ala Lys Tyr Ser Ala Gly Gly His His His
 580 585 590
 His His His
 595

<210> 36
 <211> 1788
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> DNA, die humanes P501S (Aminosäuren 55-553) codiert, fusioniert an 6 Histidinreste stromabwärts der Hefe-alpha-präpro-Signalsequenz

<400> 36
 atgagtttcc tcaattttac tgcagtttta ttgcgagcat cctccgcatt agctgctcca 60
 gtcaacacta caacagaaga tgaaacggca caaattccgg ctgaagctgt catcggttac 120
 tcagatttag aaggggattt cgatgttgct gttttgccat tttccaacag cacaaataac 180
 gggttattgt ttataaatac tactattgcc agcattgctg cttaaagaaga aggggtatct 240
 ctcgagaaaa gagaggctga agccatggtg ctgggcattg gtccagtgtc gggcctggtc 300
 tgtgtcccg ccttaggctc agccagtgc cactggcgtg gacgctatgg ccgccgccgg 360
 cccttcatct gggcactgtc cttgggcatc ctgctgagcc tctttctcat cccaagggcc 420
 ggctggctag cagggtgtgt gtgcccggat cccaggcccc tggagctggc actgctcatc 480
 ctgggcgtgg ggctgtctga cttctgtggc cagggtgtgt tcactccact ggaggccctg 540
 ctctctgacc tcttccggga cccggaccac tgtcgccagg cctactctgt ctatgccttc 600
 atgatcagtc ttgggggctg cctgggctac ctctgcctg ccattgactg ggacaccagt 660
 gccctggccc cctacctggg caccagggag gaggcctctc ttggcctgct caccctcatc 720
 ttcctcacct gcgtagcagc cacactgctg gtggctgagg aggcagcgtc gggccccacc 780
 gagccagcag aagggtgtgc ggccccctcc ttgtcgcccc actgctgtcc atgccgggcc 840
 cgcttggtct tccggaacct gggcgccctg cttccccggc tgcaccagct gtgctgccgc 900
 atgccccgca ccctgcgcgc gctcttcgtg gctgagctgt gcagctggat ggcactcatg 960
 accttcacgc tgttttacac ggatttcgtg ggcgaggggc tgtaccaggc cgtgcccaga 1020
 gctgagccgg gcaccgaggg ccggagacac tatgatgaag gcgttcggat gggcagcctg 1080
 gggctgttcc tgcagtgcgc catctccctg gtcttctctc tggtcatgga ccggctgggtg 1140
 cagcgattcg gcactcgagc agtctatttg gccagtgtgg cagctttccc tgtggtgctc 1200
 ggtgccacat gcctgtccca cagtgtggcc gtggtgacag cttcagccgc cctcaccggg 1260
 ttcaccttct cagccctgca gatcctgccc tacacactgg cctccctcta ccaccgggag 1320
 aagcaggtgt tcctgcccac ataccgaggg gacactggag gtgctagcag tgaggacagc 1380
 ctgatgacca gcttcttgcc aggccttaag cctggagctc ccttccctaa tggacacgtg 1440
 ggtgctggag gcactggcct gctcccacct ccaccgcgc tctgcggggc ctctgcctgt 1500
 gatgtctccg tacgtgtggt ggtgggtgag cccaccgagg ccagggtggt tccgggcccg 1560
 ggcattctgc tggacctgc catcctggat agtgccctcc tgctgtccca ggtggcccca 1620
 tccctgttta tgggtcccat tgtccagctc agccagtctg tcaactgccta tatggtgtct 1680
 gccgcaggcc tgggtctggg cgccatttac tttgctacac aggtagtatt tgacaagagc 1740
 gacttggcca aatactcagc ggggtggacac catcaccatc accattaa 1788

<210> 37
 <211> 1995
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> DNA, die Codon-optimiertes humanes P501S (Aminosäuren 51-553) codiert, fusioniert an St. pneum. C-LytA P2-Helferepitop C-Lyta

<400> 37
 gcggccgcgc caccatggcc gccgcctacg tgcatagcga cgggagctac cccaaggaca 60
 agttcgagaa gatcaacggg acatgggtact acttcgactc ctccggctac atgctcgccc 120

```

accgctggcg gaagcacacc gacggcaact ggtactggtt cgataactcg ggagagatgg 180
ccaccggctg gaagaagatc gcggaacaagt ggtactatct caacgaggag ggcgccatga 240
agaccggctg ggtgaagtat aaggacacct ggtactacct cgacgccaag gagggcgcca 300
tgcagtatat caaggccaac agcaagttca tcggcatcac cgagggagtg atggtcagca 360
acgcctttat ccagagcgcc gacggcaccg gatggtacta cttgaagccg gacggcaccc 420
tcgcggtatcg gcccagagaag ttcatgtaca tgggtgctggg catcggtccc gtccctgggcc 480
tcgtgtgtgt gcccctcttc gggagtgcgt ccgatacttg gcggggccgc tacggccgcc 540
gcagaccgtt catctgggcc ctgagcctgg gcatactgct ctctctcttc ctgatcccc 600
gggcccggctg gctggccggc ctgctgtgtc ccgacccccg ccctctggag ctggccctcc 660
tgatcctggg cgtgggcctg ctggacttct gcggccaggt gtgtttcact cccctggagg 720
ctctgtcttc cgacctcttc cgcgaccccc accactgtag gcaggcttac agcgtgtacg 780
ccttcattgat cagtctgggg ggatgcctgg gctatctgct gccgctatc gactgggaca 840
ccagcgccct ggccccctac ctggggactc agggaggatg cctgttcggc ctgctcacct 900
tgatcttctt gacgtgcgtc gccgccacct tgctggtggc cgaggaggcg gccctggggc 960
ccaccgagcc cgcgagggc ctgagcgctc ccagcctgag ccccatattg tgcccggtga 1020
gggctaggct cgccttcagg aatctgggcg ctttctgctc ccgctgcat cagctgtgct 1080
gtgcgatgcc tcgcacctg cgcgcctgt tcgtcgtga gctctgttcc tggatggccc 1140
tgatgacgtt caccctcttc tacaccgact tcgtggggga gggcctgtac caggcggtgc 1200
ccaggggcga gcccggcacc gaggctaggc gccattacga cgaggcgctc aggatgggct 1260
ctctgggctt ctctctgcag tgcgccatca gtctggtgtt ctctctggtg atggaccggc 1320
tgggtgcagcg cttcggcacc cgggcccgtg acctcgctc tgtggcggct tccccgctc 1380
ccgcccggcg gacctgcctg tctcattctg tcgcccgtgt gaccgccagc gccgccctga 1440
ccggtctcac cttcagtgcg ctccagattc tgccctacac cctggcgctc ctgtaccatc 1500
gcgagaagca ggtgttcttg cccaagtacc gcggggacac agggggagct tcctctgagg 1560
acagcctgat gaccagcttc ttgcccggcc ccaagccggg ggcccccttc cccaacggcc 1620
atgtcggggc gggcggcagc ggctgctcc ccccccccc cgccctgtgc ggcgctagt 1680
cctgcgacgt gagcgtgcg gtggtggtgg gggagccac cgaggctagg gtcgtgctg 1740
gccgggggat ctgcctggac ctggccatcc tcgactccgc cttcctgctc tcccaggtg 1800
cgcccagcct gttcatgggc agtatcgtgc agctgagcca gagcgtgacc gcctacatg 1860
tgagcgccgc cggcctgggg ttggtggcca tctactttgc caccagggtc gtgttcgaca 1920
agagcgatct cgccaagtat agcgcctgag gatcc 1955

```

<210> 38

<211> 2045

<212> DNA

<213> Künstliche Sequenz

<220>

<223> DNA, die Codon-optimiertes humanes P501S (Aminosäuren 1-553) codiert, fusioniert an St. pneum. C-LytA P2 Helfer-Epitop C-LytA

<400> 38

```

gcggccgcgc caccatggcc gccgcctacg tgcatagcga cgggagctac cccaaggaca 60
agttcgagaa gatcaacggg acatggtact acttcgactc ctccggctac atgctcgccg 120
accgctggcg gaagcacacc gacggcaact ggtactggtt cgataactcg ggagagatgg 180
ccaccggctg gaagaagatc gcggaacaagt ggtactatct caacgaggag ggcgccatga 240
agaccggctg ggtgaagtat aaggacacct ggtactacct cgacgccaag gagggcgcca 300
tgcagtatat caaggccaac agcaagttca tcggcatcac cgagggagtg atggtcagca 360
acgcctttat ccagagcgcc gacggcaccg gatggtacta cttgaagccg gacggcaccc 420
tcgcggtatc gcccagatg gtgcagcgcc tgtgggtgtc ccgctgctg cgccatagaa 480
agggccagtt gctgctgggtg aacctgctga ctttcggact ggaggtgtgc ctggctgccc 540
tgggtgctggg catcggtccc gtccctgggcc tcgtgtgtgt gcccctcttc gggagtgcgt 600
ccgatacttg gcggggccgc tacggccgcc gcagaccgtt catctgggcc ctgagcctgg 660
gcatactgct ctctctcttc ctgatcccc gggccggctg gctggccggc ctgctgtgtc 720
ccgacccccg ccctctggag ctggccctcc tgatcctggg cgtgggcctg ctggacttct 780
gcggccaggt gtgtttcact cccctggagg ctctgtcttc cgacctcttc cgcgaccccc 840
accactgtag gcaggcttac agcgtgtacg ccttcattgat cagtctgggg ggatgcctgg 900
gctatctgct gccgctatc gactgggaca ccagcgccct ggccccctac ctggggactc 960
aggaggatg cctgttcggc ctgctcacct tgatcttctt gacgtgcgtc gccgccacct 1020
tgctggtggc cgaggaggcg gccctggggc ccaccgagcc cgcgagggc ctgagcgctc 1080
ccagcctgag ccccatattg tgcccggtga gggctaggct cgccttcagg aatctgggcg 1140

```

```

cttttgetgcc ccgcctgcat cagctgtgct gtcgcatgcc tcgcaccctg cgcgcctgt 1200
tcgtcgctga gctctgttcc tggatggccc tgatgacgtt caccctcttc tacaccgact 1260
tcgtggggga gggcctgtac cagggcgtgc ccagggccga gcccggcacc gaggctaggc 1320
gccattacga cgagggcgctc aggatgggct ctctgggccc ctctctgcag tgcgccatca 1380
gtctggtgtt ctctctgggtg atggaccggc tgggtgcagcg cttcggcacc cgggcccgtgt 1440
acctcgcttc tgtggcggtt ttccccgtcg ccgcggcgcc gacctgcctg tctcattctg 1500
tcgcctgggt gaccgccagc gccgccctga ccggcttcac cttcagtgcg ctccagattc 1560
tgccctacac cctggcgctc ctgtaccatc gcgagaagca ggtgttcctg cccaagtacc 1620
gcggggacac agggggagct tcctctgagg acagcctgat gaccagcttc ttgccgggcc 1680
ccaagccggg gggcccttcc cccaacggcc atgtcggggc gggcggcagc ggccgtctcc 1740
ctcccccccc cgccctgtgc ggcgctagt cctgcgacgt gacgctgcgg gtggtggtgg 1800
gggagcccac cgaggctagg gtcgtgcctg gccgggggat ctgcctggac ctggccatcc 1860
tcgactccgc cttctgtctc tcccaggtgg cggccagcct gttcatgggc agtatcgtgc 1920
agctgagcca gacgctgacc gcctacatgg tgagcgccgc cggcctgggg ttggtggcca 1980
tctactttgc caccaggtc gtgttcgaca agagcgatct cgccaagtat agcgcttag 2040
gatcc 2045

```

<210> 39

<211> 2105

<212> DNA

<213> Künstliche Sequenz

<220>

<223> DNA, die St. pneum. C-LytA P2-Helferepitop C-LytA codiert,
fusioniert an humanes P501S (Aminosäuren 51-553), fusioniert
an humanes P501S (Aminosäuren 1-50) - Codon-optimiert

<400> 39

```

gcggccgcgc caccatggcc gccgcctacg tgcatagcga cgggagctac cccaaggaca 60
agttcgagaa gatcaacggg acatgggtact acttcgactc ctccggctac atgctcgccg 120
accgctggcg gaagcacacc gacggcaact ggtactgggt cgataactcg ggagagatgg 180
ccaccggctg gaagaagatc gcggacaagt ggtactattt caacgaggag ggcgccatga 240
agaccggctg ggtgaagtat aaggacacct ggtactacct cgacgccaag gagggcgcca 300
tgagtatata caaggccaac agcaagttca tcggcatcac cgagggagtg atggtcagca 360
acgcctttat ccagagcgcc gacggcaccg gatggtaacta cttgaagccg gacggcacc 420
tcgcggatcg gcccgagaag ttcattgtaca tgggtgctggg catcgcccc gtccctggcc 480
tcgtgtgtgt gccctcctc gggagtgcgt ccgatcattg gcggggccgc tacggccgcc 540
gcagaccgtt catctggggc ctgagcctgg gcacccctgt ctctctcttc ctgatecccc 600
gggcccggctg gctggccggc ctgctgtgtc ccgacccccg cctctggag ctggccctcc 660
tgatcctggg cgtgggcctg ctggacttct gcggccaggt gtgtttcact cccctggagg 720
ctctgctctc cgacctcttc cgcgaccccc accactgtag gcaggcttac agcgtgtacg 780
ccttcattat cagtctgggg ggatgcctgg gctatctgct gcccgctatc gactgggaca 840
ccagcgccct ggccccctac ctggggactc aggaggagt cctgttcggc ctgctcacct 900
tgatcttcct gacgtgcgtc gccgccaccc tgcgtggtggc cgaggaggcg gccctggggc 960
ccaccgagcc cgcggagggc ctgagcgctc ccagcctgag cccccattgc tgcccggtga 1020
gggctaggct cgccttcagg aatctgggcg ctttgctgcc ccgcctgcat cagctgtgct 1080
gtcgcatgcc tcgcaccctg cgcgcctgt tgcgtcgtga gctctgttcc tggatggccc 1140
tgatgacgtt caccctcttc tacaccgact tcgtggggga gggcctgtac cagggcgtgc 1200
ccagggccga gcccggcacc gaggctaggc gccattacga cgagggcgctc aggatgggct 1260
ctctgggccc ctctctgcag tgcgccatca gtctggtgtt ctctctgggt atggaccggc 1320
tgggtgcagcg cttcggcacc cgggcccgtgt acctcgcttc tgtggcggtt ttccccgtcg 1380
ccgcggcgcc gacctgcctg tctcattctg tcgcctgggt gaccgccagc gccgccctga 1440
ccggcttcac cttcagtgcg ctccagattc tgccctacac cctggcgctc ctgtaccatc 1500
gcgagaagca ggtgttcctg cccaagtacc gcggggacac agggggagct tcctctgagg 1560
acagcctgat gaccagcttc ttgccgggcc ccaagccggg gggcccttcc cccaacggcc 1620
atgtcggggc gggcggcagc ggccgtctcc ctcccccccc cgccctgtgc ggcgctagt 1680
cctgcgacgt gacgctgcgg gtggtggtgg gggagccac cgaggctagg gtcgtgcctg 1740
gccgggggat ctgcctggac ctggccatcc tcgactccgc cttctgtctc tcccaggtgg 1800
cggccagcct gttcatgggc agtatcgtgc agctgagcca gacgctgacc gcctacatgg 1860
tgagcgccgc cggcctgggg ttggtggcca tctactttgc caccaggtc gtgttcgaca 1920

```

```

agagcgatct cgccaagtat agcgccatgg tgcagcgget gtgggtgtcc cggctgctgc 1980
gccatagaaa ggcccagttg ctgctggtga acctgctgac ttccggactg gaggtgtgcc 2040
tggtgcccgg gatcacgtac gtgccccccc tgctgctgga ggtgggcgtg gaggagttag 2100
gatcc                                           2105

```

<210> 40
 <211> 2105
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> DNA, die humanes P501S (Aminosäuren 1-50) codiert, fusioniert
 an St. pneum. C-LyTA P2-Helferepitop C-Lyta, fusioniert an
 humanes P501S (Aminosäuren 51-553) - Codon-optimiert

```

<400> 40
gcggccgcgc caccatggtg cagcggtgtg ggggtgtccc gctgctgcgc catagaaagg 60
cccagtgtgt gctggtgaac ctgctgactt tccgactgga ggtgtgcctg gctgccggga 120
tcacgtacgt gccccccctg ctgctggagg tgggcgtgga ggagatggcc gccgcctacg 180
tgcatacgca cgggagctac cccaaggaca agttcgagaa gatcaacggg acatggtact 240
acttcgactc ctccggctac atgctcgccg accgctggcg gaagcacacc gacggcaact 300
ggtactggtt cgataactcg ggagagatgg ccaccggctg gaagaagatc gcggacaagt 360
ggtactatct caacgaggag ggcgccatga agaccggctg ggtgaagtat aaggacacct 420
ggtactacct cgacgccaaag gagggcgcca tgcagtatat caaggccaac agcaagttca 480
tcggcatcac cgaggagtg atggtcagca acgcctttat ccagagcgcc gacggcaccg 540
gatggtacta cttgaagccg gacggcaccc tcgcggatcg gcccgagaag ttcattgtaca 600
tggtgctggg catcgcccc gtccgtgggc tcgtgtgtgt gcccctcctc gggagtgcgt 660
ccgatcattg gcggggccgc tacggccgcc gcagaccgtt catctgggcc ctgagcctgg 720
gcattcctgt ctctctcttc ctgatcccc gggccggctg gctggccggc ctgctgtgtc 780
ccgacccccg ccctctggag ctggccctcc tgatcctggg cgtgggcctg ctggacttct 840
gcggccaggt gtgtttcact cccctggagg ctctgctctc cgacctctc cgcgaccccc 900
accactgtag gcaggcttac agcgtgtacg ccttcattgat cagtctgggg ggatgcctgg 960
gctatctgct gcccgtatc gactgggaca ccagcgccct gggccccctac ctggggactc 1020
aggaggagtg cctgttcggc ctgctcacct tgatcttctt gacgtgcgtc gccgccaccc 1080
tgctggtggc cgaggaggcg gccctggggc ccaccgagcc cgcggagggc ctgagcgctc 1140
ccagcctgag ccccatctgc tgccctgca gggctaggct cgcttcagg aatctgggcg 1200
ctttgctgcc ccgctgcat cagctgtgct gtgcgatgcc tcgcacctg cgccgcctgt 1260
tcgtcgctga gctctgttcc tggatggccc tgatgacgtt cacctcttc tacaccgact 1320
tcgtggggga gggcctgtac cagggcgtgc ccagggccga gcccggcacc gaggctaggc 1380
gccattacga cgagggcgtc aggatgggct ctctgggctt ctctctgcag tgcgccatca 1440
gtctggtgtt ctctctgggt atggaccggc tgggtgcagc cttcggcacc cgggccgtgt 1500
acctcgctc tgtggcggtt ttccccgtcg ccgcccggcg gacctgctg tctcattctg 1560
tcgccgtggt gaccgccagc gccgccctga ccggcttcac cttcagtgcg ctccagattc 1620
tgccctacac cctggcgtct ctgtaccatc gcgagaagca ggtgttctct cccaagtacc 1680
gcggggacac agggggagct tcctctgagg acagcctgat gaccagcttc ttgcccggcc 1740
ccaagccggg ggcccccttc cccaacggcc atgtcggggc gggcggcagc ggcctgctcc 1800
ctcccccccc cgccctgtgc ggcgctagt cctgcgacgt gagcgtgcgg gtggtggtgg 1860
gggagcccac cgaggctagg gtcgtgcctg gccgggggat ctgcctggac ctggccatcc 1920
tcgactccgc ctctctgtct tcccagggtg cgcacagcct gttcatgggc agtatcgtgc 1980
agctgagcca gagcgtgacc gcctacatgg tgagcgccgc cggcctgggg ttggtggcca 2040
tctactttgc caccaggtc gtgttcgaca agagcgatct cgccaagtat agcgccctgag 2100
gatcc                                           2105

```

<210> 41
 <211> 652
 <212> PRT
 <213> Künstliche Sequenz

<220>
 <223> St. pneum. C-LyTA P2-Helferepitop C-Lyta, fusioniert an

humanes P501S

<400>

41

```

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

```

```

Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly Phe Thr Phe
465                               470                               475                               480
Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu Tyr His Arg
                               485                               490                               495
Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly Asp Thr Gly Gly Ala
                               500                               505                               510
Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu Pro Gly Pro Lys Pro
                               515                               520                               525
Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly Ser Gly Leu
                               530                               535                               540
Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala Cys Asp Val Ser
545                               550                               555                               560
Val Arg Val Val Val Gly Glu Pro Thr Glu Ala Arg Val Val Pro Gly
                               565                               570                               575
Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp Ser Ala Phe Leu Leu
                               580                               585                               590
Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser Ile Val Gln Leu Ser
                               595                               600                               605
Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala Gly Leu Gly Leu Val
610                               615                               620
Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp Lys Ser Asp Leu Ala
625                               630                               635                               640
Lys Tyr Ser Ala Gly Gly His His His His His His
                               645                               650

```

<210> 42
 <211> 1959
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> DNA, die St. pneum. C-LytA P2-Helferepitop C-Lyta codiert,
 fusioniert an humanes P501A (plus His-Marker)

<400> 42

```

atggcgggccg cttacgtaca ttccgacggc tcttatccaa aagacaagtt tgagaaaatc 60
aatggcactt ggtactactt tgacagttca ggctatatgc ttgcagaccg ctggaggaag 120
cacacagacg gcaactggta ctgggttcgac aactcaggcg aaatgggtac aggctggaag 180
aaaatcgctg ataagtggta ctatttcaac gaagaagggtg ccatgaagac aggctgggtc 240
aagtacaagg acacttggta ctacttagac gctaaagaag gcgccatgca atacatcaag 300
gctaactcta agttcattgg tatcactgaa ggcgtcatgg tatcaaatgc ctttatccag 360
tcagcggacg gaacaggctg gtactacctc aaaccagacg gaacactggc agacaggcca 420
gaaaagttca tgtacatggt gctgggcatt ggtccagtgc tgggcctggc ctgtgtcccg 480
ctcctaggct cagccagtga ccactggcgt ggacgctatg gccgccgccg gcccttcac 540
tgggcactgt ccttgggcat cctgctgagc ctctttctca tccaagggc cggctggcta 600
gcagggctgc tgtgcccgga tcccaggccc ctggagctgg cactgctcat cctgggcgtg 660
gggctgctgg acttctgtgg ccagggtgtgc ttactccac tggaggccct gctctctgac 720
ctcttcggg acccggaaca ctgtcgccag gcctactctg tctatgcct catgatcagt 780
cttgggggct gcctgggcta cctcctgcct gccattgact gggacaccag tgccctggc 840
ccctacctg gcacccagga ggagtgcctc tttggcctgc tcaccctcat ctctctacc 900
tgctgagcag ccacactgct ggtggctgag gaggcagcgc tgggcccac cgagccagca 960
gaagggctgt cggccccctc cttgtcgccc cactgctgtc catgccgggc ccgcttggct 1020
ttccggaacc tgggcgcctt gcttcccccg ctgcaccagc tgtgctgccg catgccccgc 1080
accctgcgcc ggtcttctgt ggctgagctg tgcagctgga tggcactcat gaccttcacg 1140
ctgttttaca cggatttctt gggcgagggg ctgtaccagg gcgtgcccag agctgagccg 1200
ggcaccgagg cccggagaca ctatgatgaa ggcgttcgga tgggcagcct ggggctgttc 1260
ctgcagtgcg ccctctccct ggtcttctct ctggctcatg accggctggc gcagcgattc 1320
ggcactcgag cagtctatct ggccagtgtg gcagctttcc ctgtggtgct cggtgccaca 1380
tgcctgtccc acagtgtggc cgtggtgaca gcttcagccg ccctcaccgg gttcaccttc 1440
tcagccctgc agatcctgcc ctacacactg gcctccctct accaccggga gaagcaggtg 1500

```

```

ttcctgcccc aataccgagg ggacactgga ggtgctagca gtgaggacag cctgatgacc 1560
agcttctctgc caggccctaa gcctggagct cccttcccta atggacacgt ggggtgctgga 1620
ggcagtggcc tgctccacc tccaccgcg ctctgcgggg cctctgcctg tgatgtctcc 1680
gtacgtgtgg tgggtgggtga gcccaccgag gccagggtgg ttccggggccg gggcatctgc 1740
ctggacctcg ccactctgga tagtgcttc ctgctgtccc aggtggcccc atccctgttt 1800
atgggctcca ttgtccagct cagccagtct gtcactgcct atatggtgtc tgccgcaggc 1860
ctgggtctgg tcgccattta ctttgctaca caggtagtat ttgacaagag cgacttggcc 1920
aaatactcag cgggtggaca ccataccat caccattaa 1959

```

<210> 43
 <211> 553
 <212> PRT
 <213> Homo sapiens

<400> 43

```

Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys Ala
1      5      10      15
Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val Cys Leu
20     25     30
Ala Ala Gly Ile Thr Tyr Val Pro Leu Leu Leu Glu Val Gly Val
35     40     45
Glu Glu Lys Phe Met Thr Met Val Leu Gly Ile Gly Pro Val Leu Gly
50     55     60
Leu Val Cys Val Pro Leu Leu Gly Ser Ala Ser Asp His Trp Arg Gly
65     70     75     80
Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp Ala Leu Ser Leu Gly Ile
85     90     95
Leu Leu Ser Leu Phe Leu Ile Pro Arg Ala Gly Trp Leu Ala Gly Leu
100    105    110
Leu Cys Pro Asp Pro Arg Pro Leu Glu Leu Ala Leu Leu Ile Leu Gly
115    120    125
Val Gly Leu Leu Asp Phe Cys Gly Gln Val Cys Phe Thr Pro Leu Glu
130    135    140
Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro Asp His Cys Arg Gln Ala
145    150    155    160
Tyr Ser Val Tyr Ala Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr
165    170    175
Leu Leu Pro Ala Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu
180    185    190
Gly Thr Gln Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu
195    200    205
Thr Cys Val Ala Ala Thr Leu Leu Val Ala Glu Glu Ala Ala Leu Gly
210    215    220
Pro Thr Glu Pro Ala Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His
225    230    235    240
Cys Cys Pro Cys Arg Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala Leu
245    250    255
Leu Pro Arg Leu His Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg
260    265    270
Arg Leu Phe Val Ala Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe
275    280    285
Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val
290    295    300
Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly
305    310    315    320
Val Arg Met Gly Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu
325    330    335
Val Phe Ser Leu Val Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg
340    345    350
Ala Val Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala
355    360    365

```

```

Thr Cys Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu
 370                      375                      380
Thr Gly Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala
385                      390                      395                      400
Ser Leu Tyr His Arg Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly
                      405                      410                      415
Asp Thr Gly Gly Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu
                      420                      425                      430
Pro Gly Pro Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala
                      435                      440                      445
Gly Gly Ser Gly Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser
                      450                      455                      460
Ala Cys Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala
465                      470                      475                      480
Arg Val Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp
                      485                      490                      495
Ser Ala Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser
                      500                      505                      510
Ile Val Gln Leu Ser Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala
                      515                      520                      525
Gly Leu Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp
                      530                      535                      540
Lys Ser Asp Leu Ala Lys Tyr Ser Ala
545                      550

```

<210> 44
 <211> 644
 <212> PRT
 <213> Künstliche Sequenz

<220>
 <223> St. pneum. C-LytA P2-Helferepitop C-Lyta, fusioniert an
 humanes P501S

<400> 44

```

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

```

```

      195              200              205
Arg Pro Leu Glu Leu Ala Leu Leu Ile Leu Gly Val Gly Leu Leu Asp
  210              215              220
Phe Cys Gly Gln Val Cys Phe Thr Pro Leu Glu Ala Leu Leu Ser Asp
  225              230              235              240
Leu Phe Arg Asp Pro Asp His Cys Arg Gln Ala Tyr Ser Val Tyr Ala
      245              250              255
Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr Leu Leu Pro Ala Ile
      260              265              270
Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu Gly Thr Gln Glu Glu
      275              280              285
Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu Thr Cys Val Ala Ala
  290              295              300
Thr Leu Leu Val Ala Glu Glu Ala Ala Leu Gly Pro Thr Glu Pro Ala
  305              310              315              320
Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His Cys Cys Pro Cys Arg
      325              330              335
Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala Leu Leu Pro Arg Leu His
      340              345              350
Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg Arg Leu Phe Val Ala
  355              360              365
Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe Thr Leu Phe Tyr Thr
  370              375              380
Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val Pro Arg Ala Glu Pro
  385              390              395              400
Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly Val Arg Met Gly Ser
      405              410              415
Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu Val Phe Ser Leu Val
      420              425              430
Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu Ala
      435              440              445
Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys Leu Ser His
  450              455              460
Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly Phe Thr Phe
  465              470              475              480
Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu Tyr His Arg
      485              490              495
Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly Asp Thr Gly Gly Ala
  500              505              510
Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu Pro Gly Pro Lys Pro
  515              520              525
Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly Ser Gly Leu
  530              535              540
Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala Cys Asp Val Ser
  545              550              555              560
Val Arg Val Val Val Gly Glu Pro Thr Glu Ala Arg Val Val Pro Gly
      565              570              575
Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp Ser Ala Phe Leu Leu
      580              585              590
Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser Ile Val Gln Leu Ser
  595              600              605
Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala Gly Leu Gly Leu Val
  610              615              620
Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp Lys Ser Asp Leu Ala
  625              630              635              640
Lys Tyr Ser Ala

```

```

<210>      45
<211>      644
<212>      PRT
<213>      Künstliche Sequenz

```

<220>

<223>

Codon-optimiertes Hybridprotein zwischen St. pneum.
C-LytA P2-Helferepitop C-Lyta, fusioniert an humanes
P501S (Aminosäuren 51-53)

<400>

45

```

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

```

```

Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu Val Phe Ser Leu Val
      420                      425                      430
Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu Ala
      435                      440                      445
Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys Leu Ser His
      450                      455                      460
Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly Phe Thr Phe
      465                      470                      475                      480
Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu Tyr His Arg
      485                      490                      495
Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly Asp Thr Gly Gly Ala
      500                      505                      510
Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu Pro Gly Pro Lys Pro
      515                      520                      525
Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly Ser Gly Leu
      530                      535                      540
Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala Cys Asp Val Ser
      545                      550                      555                      560
Val Arg Val Val Val Gly Glu Pro Thr Glu Ala Arg Val Val Pro Gly
      565                      570                      575
Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp Ser Ala Phe Leu Leu
      580                      585                      590
Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser Ile Val Gln Leu Ser
      595                      600                      605
Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala Gly Leu Gly Leu Val
      610                      615                      620
Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp Lys Ser Asp Leu Ala
      625                      630                      635                      640
Lys Tyr Ser Ala

```

<210> 46
 <211> 694
 <212> PRT
 <213> Künstliche Sequenz

<220>
 <223> St. pneum. C-LytA P2-Helferepitop C-Lyta, fusioniert an
 humanes P501S (Aminosäuren 1-553) - Codon-optimiert

<400> 46

```

Met Ala Ala Ala Tyr Val His Ser Asp Gly Ser Tyr Pro Lys Asp Lys
  1      5      10      15
Phe Glu Lys Ile Asn Gly Thr Trp Tyr Tyr Phe Asp Ser Ser Gly Tyr
      20      25      30
Met Leu Ala Asp Arg Trp Arg Lys His Thr Asp Gly Asn Trp Tyr Trp
      35      40      45
Phe Asp Asn Ser Gly Glu Met Ala Thr Gly Trp Lys Lys Ile Ala Asp
      50      55      60
Lys Trp Tyr Tyr Phe Asn Glu Glu Gly Ala Met Lys Thr Gly Trp Val
      65      70      75      80
Lys Tyr Lys Asp Thr Trp Tyr Tyr Leu Asp Ala Lys Glu Gly Ala Met
      85      90      95
Gln Tyr Ile Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu Gly Val
      100     105     110
Met Val Ser Asn Ala Phe Ile Gln Ser Ala Asp Gly Thr Gly Trp Tyr
      115     120     125
Tyr Leu Lys Pro Asp Gly Thr Leu Ala Asp Arg Pro Glu Met Val Gln
      130     135     140
Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys Ala Gln Leu Leu

```

```

145          150          155          160
Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val Cys Leu Ala Ala Gly
          165          170          175
Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu Val Gly Val Glu Glu Lys
          180          185          190
Phe Met Thr Met Val Leu Gly Ile Gly Pro Val Leu Gly Leu Val Cys
          195          200          205
Val Pro Leu Leu Gly Ser Ala Ser Asp His Trp Arg Gly Arg Tyr Gly
          210          215          220
Arg Arg Arg Pro Phe Ile Trp Ala Leu Ser Leu Gly Ile Leu Leu Ser
225          230          235          240
Leu Phe Leu Ile Pro Arg Ala Gly Trp Leu Ala Gly Leu Leu Cys Pro
          245          250          255
Asp Pro Arg Pro Leu Glu Leu Ala Leu Leu Ile Leu Gly Val Gly Leu
          260          265          270
Leu Asp Phe Cys Gly Gln Val Cys Phe Thr Pro Leu Glu Ala Leu Leu
          275          280          285
Ser Asp Leu Phe Arg Asp Pro Asp His Cys Arg Gln Ala Tyr Ser Val
          290          295          300
Tyr Ala Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr Leu Leu Pro
305          310          315          320
Ala Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu Gly Thr Gln
          325          330          335
Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu Thr Cys Val
          340          345          350
Ala Ala Thr Leu Leu Val Ala Glu Glu Ala Ala Leu Gly Pro Thr Glu
          355          360          365
Pro Ala Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His Cys Cys Pro
          370          375          380
Cys Arg Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala Leu Leu Pro Arg
385          390          395          400
Leu His Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg Arg Leu Phe
          405          410          415
Val Ala Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe Thr Leu Phe
          420          425          430
Tyr Thr Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val Pro Arg Ala
          435          440          445
Glu Pro Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly Val Arg Met
          450          455          460
Gly Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu Val Phe Ser
465          470          475          480
Leu Val Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val Tyr
          485          490          495
Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys Leu
          500          505          510
Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly Phe
          515          520          525
Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu Tyr
          530          535          540
His Arg Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly Asp Thr Gly
545          550          555          560
Gly Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu Pro Gly Pro
          565          570          575
Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly Ser
          580          585          590
Gly Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala Cys Asp
          595          600          605
Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala Arg Val Val
          610          615          620
Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp Ser Ala Phe
625          630          635          640

```

Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser Ile Val Gln
 645 650 655
 Leu Ser Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala Gly Leu Gly
 660 665 670
 Leu Val Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp Lys Ser Asp
 675 680 685
 Leu Ala Lys Tyr Ser Ala
 690

<210> 47
 <211> 694
 <212> PRT
 <213> Künstliche Sequenz

<220>

<223> St. pneum. C-LytA P2 Helferepitop C-Lyta, fusioniert an
 humanes P501S (Aminosäuren 51-553), fusioniert an humanes
 P501S (Aminosäuren 1-50) - Codon-optimiert

<400> 47

Met Ala Ala Ala Tyr Val His Ser Asp Gly Ser Tyr Pro Lys Asp Lys
 1 5 10 15
 Phe Glu Lys Ile Asn Gly Thr Trp Tyr Tyr Phe Asp Ser Ser Gly Tyr
 20 25 30
 Met Leu Ala Asp Arg Trp Arg Lys His Thr Asp Gly Asn Trp Tyr Trp
 35 40 45
 Phe Asp Asn Ser Gly Glu Met Ala Thr Gly Trp Lys Lys Ile Ala Asp
 50 55 60
 Lys Trp Tyr Tyr Phe Asn Glu Glu Gly Ala Met Lys Thr Gly Trp Val
 65 70 75 80
 Lys Tyr Lys Asp Thr Trp Tyr Tyr Leu Asp Ala Lys Glu Gly Ala Met
 85 90 95
 Gln Tyr Ile Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu Gly Val
 100 105 110
 Met Val Ser Asn Ala Phe Ile Gln Ser Ala Asp Gly Thr Gly Trp Tyr
 115 120 125
 Tyr Leu Lys Pro Asp Gly Thr Leu Ala Asp Arg Pro Glu Lys Phe Met
 130 135 140
 Tyr Met Val Leu Gly Ile Gly Pro Val Leu Gly Leu Val Cys Val Pro
 145 150 155 160
 Leu Leu Gly Ser Ala Ser Asp His Trp Arg Gly Arg Tyr Gly Arg Arg
 165 170 175
 Arg Pro Phe Ile Trp Ala Leu Ser Leu Gly Ile Leu Leu Ser Leu Phe
 180 185 190
 Leu Ile Pro Arg Ala Gly Trp Leu Ala Gly Leu Leu Cys Pro Asp Pro
 195 200 205
 Arg Pro Leu Glu Leu Ala Leu Leu Ile Leu Gly Val Gly Leu Leu Asp
 210 215 220
 Phe Cys Gly Gln Val Cys Phe Thr Pro Leu Glu Ala Leu Leu Ser Asp
 225 230 235 240
 Leu Phe Arg Asp Pro Asp His Cys Arg Gln Ala Tyr Ser Val Tyr Ala
 245 250 255
 Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr Leu Leu Pro Ala Ile
 260 265 270
 Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu Gly Thr Gln Glu Glu
 275 280 285
 Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu Thr Cys Val Ala Ala
 290 295 300
 Thr Leu Leu Val Ala Glu Ala Ala Leu Gly Pro Thr Glu Pro Ala
 305 310 315 320

Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His Cys Cys Pro Cys Arg
 325 330 335
 Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala Leu Leu Pro Arg Leu His
 340 345 350
 Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg Arg Leu Phe Val Ala
 355 360 365
 Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe Thr Leu Phe Tyr Thr
 370 375 380
 Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val Pro Arg Ala Glu Pro
 385 390 395 400
 Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly Val Arg Met Gly Ser
 405 410 415
 Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu Val Phe Ser Leu Val
 420 425 430
 Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu Ala
 435 440 445
 Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys Leu Ser His
 450 455 460
 Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly Phe Thr Phe
 465 470 475 480
 Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu Tyr His Arg
 485 490 495
 Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly Asp Thr Gly Gly Ala
 500 505 510
 Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu Pro Gly Pro Lys Pro
 515 520 525
 Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly Ser Gly Leu
 530 535 540
 Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala Cys Asp Val Ser
 545 550 555 560
 Val Arg Val Val Val Gly Glu Pro Thr Glu Ala Arg Val Val Pro Gly
 565 570 575
 Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp Ser Ala Phe Leu Leu
 580 585 590
 Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser Ile Val Gln Leu Ser
 595 600 605
 Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala Gly Leu Gly Leu Val
 610 615 620
 Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp Lys Ser Asp Leu Ala
 625 630 635 640
 Lys Tyr Ser Ala Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg
 645 650 655
 His Arg Lys Ala Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu
 660 665 670
 Glu Val Cys Leu Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu
 675 680 685
 Glu Val Gly Val Glu Glu
 690

<210> 48

<211> 694

<212> PRT

<213> Künstliche Sequenz

<220>

<223> Humanes P501S (Aminosäuren 1-50), fusioniert an St. pneum.
 C-LytA P2-Helferepitop C-Lyta, fusioniert an humanes
 P501S (Aminosäuren 51-553) - Codon-optimiert

<400> 48

```

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

```

	485		490		495										
Leu	Ala	Ser	Val	Ala	Ala	Phe	Pro	Val	Ala	Ala	Gly	Ala	Thr	Cys	Leu
	500						505						510		
Ser	His	Ser	Val	Ala	Val	Val	Thr	Ala	Ser	Ala	Ala	Leu	Thr	Gly	Phe
	515						520					525			
Thr	Phe	Ser	Ala	Leu	Gln	Ile	Leu	Pro	Tyr	Thr	Leu	Ala	Ser	Leu	Tyr
	530					535					540				
His	Arg	Glu	Lys	Gln	Val	Phe	Leu	Pro	Lys	Tyr	Arg	Gly	Asp	Thr	Gly
545				550					555						560
Gly	Ala	Ser	Ser	Glu	Asp	Ser	Leu	Met	Thr	Ser	Phe	Leu	Pro	Gly	Pro
			565						570					575	
Lys	Pro	Gly	Ala	Pro	Phe	Pro	Asn	Gly	His	Val	Gly	Ala	Gly	Gly	Ser
	580							585					590		
Gly	Leu	Leu	Pro	Pro	Pro	Pro	Ala	Leu	Cys	Gly	Ala	Ser	Ala	Cys	Asp
	595						600				605				
Val	Ser	Val	Arg	Val	Val	Val	Gly	Glu	Pro	Thr	Glu	Ala	Arg	Val	Val
	610					615					620				
Pro	Gly	Arg	Gly	Ile	Cys	Leu	Asp	Leu	Ala	Ile	Leu	Asp	Ser	Ala	Phe
625				630					635						640
Leu	Leu	Ser	Gln	Val	Ala	Pro	Ser	Leu	Phe	Met	Gly	Ser	Ile	Val	Gln
			645						650					655	
Leu	Ser	Gln	Ser	Val	Thr	Ala	Tyr	Met	Val	Ser	Ala	Ala	Gly	Leu	Gly
	660						665						670		
Leu	Val	Ala	Ile	Tyr	Phe	Ala	Thr	Gln	Val	Val	Phe	Asp	Lys	Ser	Asp
	675					680						685			
Leu	Ala	Lys	Tyr	Ser	Ala										
	690														

<210> 49
 <211> 1971
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> DNA, die humanes MUC-1 codiert, fusioniert an St. pneum.
 C-LytA P2-Helferepitop C-Lyta

<400> 49

```

atgacaccgg gcaccagtc tcctttcttc ctgctgctgc tcctcacagt gcttacagtt 60
gttacagggt ctggtcatgc aagctctacc ccagggtggag aaaaggagac ttcggctacc 120
cagagaagtt cagtgccag ctctactgag aagaatgctg tgagtatgac cagcagcgta 180
ctctccagcc acagccccgg ttcaggctcc tccaccactc agggacagga tgtcactctg 240
gccccggcca cggaaccagc ttcagggttc gctgccacct ggggacagga tgtcacctcg 300
gtcccgatca ccaggccagc cctgggctcc accaccccg cagcccacga tgtcacctca 360
gccccggaca acaagccagc cccgggctcc accgcccccc cagcccacgg tgtcacctcg 420
gccccggaca ccaggccgcc cccgggctcc accgcccccc cagcccacgg tgtcacctcg 480
gccccggaca ccaggccgcc cccgggctcc accgcccccc cagcccacgg tgtcacctcg 540
gccccggaca ccaggccggc cccgggctcc accgcccccc cagcccacgg tgtcacctcg 600
gccccggaca acaggccggc cttggcgctc accgcccccc cagtccacaa tgtcacctcg 660
gcttcagggt ctgcatcagg ctccagcttc actctggtgc acaacggcac ctctgccagg 720
gctaccacaa cccagccag caagagcact ccattctcaa tcccagcca ccaactctgat 780
actcctacca cccttgccag ccatagcacc aagactgatg ccagtagcac tcaccatagc 840
acggtacctc ctctcacctc ctccaatcac agcactcttc cccagttgtc tactggggtc 900
tctttctttt tcctgtcttt tcacatttca aacctccagt ttaattctc tctggaagat 960
cccagcaccg actactacca agagctgcag agagacattt ctgaaatgtt tttgcagatt 1020
tataaacaag ggggttttct gggcctctcc aatattaagt tcaggccagg atctgtggtg 1080
gtacaattga ctctggcctt ccgagaaggt accatcaatg tccacgacgt ggagacacag 1140
ttcaatcagt ataaaacgga agcagcctct cgatataacc tgacgatctc agacgtcagc 1200
gtgagtgatg tgccatttcc tttctctgcc cagtctgggg ctgggggtgcc aggctggggc 1260
atcgcgctgc tgggtgctgg ctgtgttctg gttgcgctgg ccattgtcta tctcattgcc 1320

```

```

ttggctgtct gtcagtgccg ccgaaagaac tacgggcagc tggacatctt tccagcccgg 1380
gatacctacc atcctatgag cgagtacccc acctaccaca cccatgggcg ctatgtgccc 1440
cctagcagta ccgatcgtag cccctatgag aagggttctg caggtaatgg tggcagcagc 1500
ctctcttaca caaaccacgc agtggcagcc acttctgcca acttgatggc ggccgcttac 1560
gtacattccg acggctctta tccaaaagac aagtttgaga aaatcaatgg cacttggtac 1620
tactttgaca gttcaggcta tatgcttgca gaccgctgga ggaagcacac agacggcaac 1680
tggtactggg tcgacaactc aggcgaaatg gctacaggct ggaagaaaat cgctgataag 1740
tggtactatt tcaacgaaga aggtgccatg aagacaggct gggtaagta caaggacact 1800
tggtactact tagacgctaa agaaggcgcc atgcaatata tcaaggctaa ctctaagttc 1860
attggtatca ctgaaggcgt catggtatca aatgccttta tccagtcagc ggacggaaca 1920
ggctggtact acctcaaacc agacggaaca ctggcagaca ggccagaatg a 1971

```

<210> 50
 <211> 656
 <212> PRT
 <213> Künstliche Sequenz

<220>

<223> Humanes MUC-1, fusioniert an St. pneum. C-LytA P2-Helfer-
 epitop C-Lyta

<400> 50

```

Met Thr Pro Gly Thr Gln Ser Pro Phe Phe Leu Leu Leu Leu Leu Thr
 1          5          10          15
Val Leu Thr Val Val Thr Gly Ser Gly His Ala Ser Ser Thr Pro Gly
 20          25          30
Gly Glu Lys Glu Thr Ser Ala Thr Gln Arg Ser Ser Val Pro Ser Ser
 35          40          45
Thr Glu Lys Asn Ala Val Ser Met Thr Ser Ser Val Leu Ser Ser His
 50          55          60
Ser Pro Gly Ser Gly Ser Ser Thr Thr Gln Gly Gln Asp Val Thr Leu
 65          70          75          80
Ala Pro Ala Thr Glu Pro Ala Ser Gly Ser Ala Ala Thr Trp Gly Gln
 85          90          95
Asp Val Thr Ser Val Pro Val Thr Arg Pro Ala Leu Gly Ser Thr Thr
100          105          110
Pro Pro Ala His Asp Val Thr Ser Ala Pro Asp Asn Lys Pro Ala Pro
115          120          125
Gly Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser Ala Pro Asp Thr
130          135          140
Arg Pro Pro Pro Gly Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser
145          150          155          160
Ala Pro Asp Thr Arg Pro Pro Pro Gly Ser Thr Ala Pro Ala Ala His
165          170          175
Gly Val Thr Ser Ala Pro Asp Thr Arg Pro Ala Pro Gly Ser Thr Ala
180          185          190
Pro Pro Ala His Gly Val Thr Ser Ala Pro Asp Asn Arg Pro Ala Leu
195          200          205
Ala Ser Thr Ala Pro Pro Val His Asn Val Thr Ser Ala Ser Gly Ser
210          215          220
Ala Ser Gly Ser Ala Ser Thr Leu Val His Asn Gly Thr Ser Ala Arg
225          230          235          240
Ala Thr Thr Thr Pro Ala Ser Lys Ser Thr Pro Phe Ser Ile Pro Ser
245          250          255
His His Ser Asp Thr Pro Thr Thr Leu Ala Ser His Ser Thr Lys Thr
260          265          270
Asp Ala Ser Ser Thr His His Ser Thr Val Pro Pro Leu Thr Ser Ser
275          280          285
Asn His Ser Thr Ser Pro Gln Leu Ser Thr Gly Val Ser Phe Phe Phe
290          295          300
Leu Ser Phe His Ile Ser Asn Leu Gln Phe Asn Ser Ser Leu Glu Asp

```

305		310		315		320									
Pro	Ser	Thr	Asp	Tyr	Tyr	Gln	Glu	Leu	Gln	Arg	Asp	Ile	Ser	Glu	Met
			325						330					335	
Phe	Leu	Gln	Ile	Tyr	Lys	Gln	Gly	Gly	Phe	Leu	Gly	Leu	Ser	Asn	Ile
		340						345					350		
Lys	Phe	Arg	Pro	Gly	Ser	Val	Val	Val	Gln	Leu	Thr	Leu	Ala	Phe	Arg
		355					360					365			
Glu	Gly	Thr	Ile	Asn	Val	His	Asp	Val	Glu	Thr	Gln	Phe	Asn	Gln	Tyr
	370					375					380				
Lys	Thr	Glu	Ala	Ala	Ser	Arg	Tyr	Asn	Leu	Thr	Ile	Ser	Asp	Val	Ser
385					390					395				400	
Val	Ser	Asp	Val	Pro	Phe	Pro	Phe	Ser	Ala	Gln	Ser	Gly	Ala	Gly	Val
			405						410					415	
Pro	Gly	Trp	Gly	Ile	Ala	Leu	Leu	Val	Leu	Val	Cys	Val	Leu	Val	Ala
		420					425						430		
Leu	Ala	Ile	Val	Tyr	Leu	Ile	Ala	Leu	Ala	Val	Cys	Gln	Cys	Arg	Arg
		435					440					445			
Lys	Asn	Tyr	Gly	Gln	Leu	Asp	Ile	Phe	Pro	Ala	Arg	Asp	Thr	Tyr	His
	450					455					460				
Pro	Met	Ser	Glu	Tyr	Pro	Thr	Tyr	His	Thr	His	Gly	Arg	Tyr	Val	Pro
465					470					475				480	
Pro	Ser	Ser	Thr	Asp	Arg	Ser	Pro	Tyr	Glu	Lys	Val	Ser	Ala	Gly	Asn
			485						490					495	
Gly	Gly	Ser	Ser	Leu	Ser	Tyr	Thr	Asn	Pro	Ala	Val	Ala	Ala	Thr	Ser
		500						505					510		
Ala	Asn	Leu	Met	Ala	Ala	Ala	Tyr	Val	His	Ser	Asp	Gly	Ser	Tyr	Pro
		515					520					525			
Lys	Asp	Lys	Phe	Glu	Lys	Ile	Asn	Gly	Thr	Trp	Tyr	Tyr	Phe	Asp	Ser
	530					535					540				
Ser	Gly	Tyr	Met	Leu	Ala	Asp	Arg	Trp	Arg	Lys	His	Thr	Asp	Gly	Asn
545					550					555				560	
Trp	Tyr	Trp	Phe	Asp	Asn	Ser	Gly	Glu	Met	Ala	Thr	Gly	Trp	Lys	Lys
			565						570					575	
Ile	Ala	Asp	Lys	Trp	Tyr	Tyr	Phe	Asn	Glu	Glu	Gly	Ala	Met	Lys	Thr
		580					585						590		
Gly	Trp	Val	Lys	Tyr	Lys	Asp	Thr	Trp	Tyr	Tyr	Leu	Asp	Ala	Lys	Glu
		595					600					605			
Gly	Ala	Met	Gln	Tyr	Ile	Lys	Ala	Asn	Ser	Lys	Phe	Ile	Gly	Ile	Thr
	610					615					620				
Glu	Gly	Val	Met	Val	Ser	Asn	Ala	Phe	Ile	Gln	Ser	Ala	Asp	Gly	Thr
625					630					635				640	
Gly	Trp	Tyr	Tyr	Leu	Lys	Pro	Asp	Gly	Thr	Leu	Ala	Asp	Arg	Pro	Glu
			645						650					655	

<210> 51

<211> 2037

<212> DNA

<213> Künstliche Sequenz

<220>

<223> DNA, die St. pneum. C-LytA P2-Helferepitop C-Lyta codiert,
fusioniert an humanes MUC-1

<400> 51

```

atgggatgga gctgtatcat cctcttcttg gtagcaacag ctacaggtgt ccactcccag 60
gtccaaatgg cgccgctta cgtacattcc gacggctctt atccaaaaga caagtttgag 120
aaaatcaatg gcacttggtg ctactttgac agttcaggct atatgcttgc agaccgctgg 180
aggaagcaca cagacggcaa ctggtactgg ttcgacaact caggcgaaat ggctacaggc 240
tggaagaaaa tcgctgataa gtggtactat ttcaacgaag aaggtgccat gaagacaggc 300
tgggtcaagt acaaggacac ttggtactac ttagacgcta aagaaggcgc catgcaatac 360

```

```

atcaaggcta actctaagtt cattggtatc actgaaggcg tcatggtatc aaatgccttt 420
atccagtcag cggacggaac aggetggtac tacctcaaac cagacggaac actggcagac 480
aggccagaaa tgacaccggg caccacgtct cttttcttcc tgctgctgct cctcacagtg 540
cttacagttg ttacaggttc tggatcatgca agctctaccc caggtggaga aaaggagact 600
tcggctaccc agagaagtcc agtgcccagc tctactgaga agaatgctgt gagtatgacc 660
agcagcgtac tctccagcca cagccccggt tcaggtcctt ccaccactca gggacaggat 720
gtcactctgg ccccggccac ggaaccagct tcaggttcag ctgccacctg gggacaggat 780
gtcacctcgg tcccagtcac caggccagcc ctgggctcca ccacccgcc agcccacgat 840
gtcacctcag ccccggacaa caagccagcc ccgggctcca ccgccccccc agcccacggt 900
gtcacctcgg ccccggacac caggccgccc ccgggctcca ccgccccccc agcccacggt 960
gtcacctcgg ccccggacac caggccgccc ccgggctcca ccgccccccc agcccacggt 1020
gtcacctcgg ccccggacac caggccgccc ccgggctcca ccgccccccc agcccacggt 1080
gtcacctcgg ccccggacaa caggcccgcc ttggcgtcca ccgccccccc agtccacaat 1140
gtcacctcgg cctcaggttc tgcacagggc tcagcttcta ctctggtgca caacggcacc 1200
tctgccaggg ctaccacaac cccagccagc aagagcactc cattctcaat tcccagccac 1260
cactctgata ctctaccac ccttgccagc catagcacca agactgatgc cagtagcact 1320
caccatagca cggtagctcc tctcacctcc tccaatcaca gcacttctcc ccagttgtct 1380
actggggtct ctttcttttt cctgtctttt cacatttcaa acctccagtt taattcctct 1440
ctggaagatc ccagcaccca ctactaccaa gagctgcaga gagacatttc tgaaatgttt 1500
ttgcagattt ataaacaagg gggttttctg ggcctctcca atattaagtt caggccagga 1560
tctgtggtgg tacaattgac tctggccttc cgagaaggta ccatcaatgt ccacgacgtg 1620
gagacacagt tcaatcagta taaaacggaa gcagcctctc gatataacct gacgatctca 1680
gacgtcagcg tgagtgatgt gccatttctt ttctctgccc agtctggggc tggggtgcca 1740
ggctggggca tcgcgtgctt ggtgctggtc tgtgttctgg ttgcgctggc cattgtctat 1800
ctcattgcct tggctgtctg tcagtgcgcg cgaaagaact acgggcagct ggacatcttt 1860
ccagcccggg atacctacca tcctatgagc gagtacccca cctaccacac ccattgggcgc 1920
tatgtgcccc ctagcagtag cgatcgtagc ccctatgaga aggtttctgc aggtaatggt 1980
ggcagcagcc tctcttacac aaaccagca gtggcagcca cttctgcca cttgtag 2037

```

<210> 52

<211> 678

<212> PRT

<213> Künstliche Sequenz

<220>

<223> St. pneum. C-LytA P2-Helferepitop C-Lyta, fusioniert an
humanes MUC-1

<400> 52

```

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

```

Leu Leu Thr Val Leu Thr Val Val Thr Gly Ser Gly His Ala Ser Ser
 180 185 190
 Thr Pro Gly Gly Glu Lys Glu Thr Ser Ala Thr Gln Arg Ser Ser Val
 195 200 205
 Pro Ser Ser Thr Glu Lys Asn Ala Val Ser Met Thr Ser Ser Val Leu
 210 215 220
 Ser Ser His Ser Pro Gly Ser Gly Ser Ser Thr Thr Gln Gly Gln Asp
 225 230 235 240
 Val Thr Leu Ala Pro Ala Thr Glu Pro Ala Ser Gly Ser Ala Ala Thr
 245 250 255
 Trp Gly Gln Asp Val Thr Ser Val Pro Val Thr Arg Pro Ala Leu Gly
 260 265 270
 Ser Thr Thr Pro Pro Ala His Asp Val Thr Ser Ala Pro Asp Asn Lys
 275 280 285
 Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser Ala
 290 295 300
 Pro Asp Thr Arg Pro Pro Gly Ser Thr Ala Pro Pro Ala His Gly
 305 310 315 320
 Val Thr Ser Ala Pro Asp Thr Arg Pro Pro Gly Ser Thr Ala Pro
 325 330 335
 Ala Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg Pro Ala Pro Gly
 340 345 350
 Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser Ala Pro Asp Asn Arg
 355 360 365
 Pro Ala Leu Ala Ser Thr Ala Pro Pro Val His Asn Val Thr Ser Ala
 370 375 380
 Ser Gly Ser Ala Ser Gly Ser Ala Ser Thr Leu Val His Asn Gly Thr
 385 390 395 400
 Ser Ala Arg Ala Thr Thr Pro Ala Ser Lys Ser Thr Pro Phe Ser
 405 410 415
 Ile Pro Ser His His Ser Asp Thr Pro Thr Thr Leu Ala Ser His Ser
 420 425 430
 Thr Lys Thr Asp Ala Ser Ser Thr His His Ser Thr Val Pro Pro Leu
 435 440 445
 Thr Ser Ser Asn His Ser Thr Ser Pro Gln Leu Ser Thr Gly Val Ser
 450 455 460
 Phe Phe Phe Leu Ser Phe His Ile Ser Asn Leu Gln Phe Asn Ser Ser
 465 470 475 480
 Leu Glu Asp Pro Ser Thr Asp Tyr Tyr Gln Glu Leu Gln Arg Asp Ile
 485 490 495
 Ser Glu Met Phe Leu Gln Ile Tyr Lys Gln Gly Gly Phe Leu Gly Leu
 500 505 510
 Ser Asn Ile Lys Phe Arg Pro Gly Ser Val Val Val Gln Leu Thr Leu
 515 520 525
 Ala Phe Arg Glu Gly Thr Ile Asn Val His Asp Val Glu Thr Gln Phe
 530 535 540
 Asn Gln Tyr Lys Thr Glu Ala Ala Ser Arg Tyr Asn Leu Thr Ile Ser
 545 550 555 560
 Asp Val Ser Val Ser Asp Val Pro Phe Pro Phe Ser Ala Gln Ser Gly
 565 570 575
 Ala Gly Val Pro Gly Trp Gly Ile Ala Leu Leu Val Leu Val Cys Val
 580 585 590
 Leu Val Ala Leu Ala Ile Val Tyr Leu Ile Ala Leu Ala Val Cys Gln
 595 600 605
 Cys Arg Arg Lys Asn Tyr Gly Gln Leu Asp Ile Phe Pro Ala Arg Asp
 610 615 620
 Thr Tyr His Pro Met Ser Glu Tyr Pro Thr Tyr His Thr His Gly Arg
 625 630 635 640
 Tyr Val Pro Pro Ser Ser Thr Asp Arg Ser Pro Tyr Glu Lys Val Ser
 645 650 655
 Ala Gly Asn Gly Gly Ser Ser Leu Ser Tyr Thr Asn Pro Ala Val Ala
 660 665 670
 Ala Thr Ser Ala Asn Leu
 675

Patentansprüche

1. Fusionspartnerprotein, das eine Cholin-Bindungsdomäne und ein heterologes promiskues T-Helferepitop umfaßt.
2. Fusionspartnerprotein gemäß Anspruch 1, worin die Cholin-Bindungsdomäne der C-Terminus von LytA oder ein Derivat davon ist, worin das Derivat des C-Terminus von LytA sowohl die Fähigkeit zur Funktion als immunologischer Partner als auch als Expressionsverstärker bewahrt.
3. Fusionspartnerprotein gemäß Anspruch 2, worin das C-LytA oder Derivat davon wenigstens vier Repeats aus einem beliebigen aus SEQ ID NO:1 bis 6 umfaßt.
4. Fusionspartnerprotein gemäß einem der Ansprüche 1 bis 3, worin die Cholin-Bindungsdomäne aus der Gruppe ausgewählt ist, die folgendes umfaßt:
 - a) die C-terminale Domäne von LytA wie in SEQ ID NO: 7 dargestellt; oder
 - b) die Sequenz von SEQ ID NO: 8; oder
 - c) eine Peptidsequenz, die eine Aminosäuresequenz mit wenigstens 85% Identität, bevorzugt wenigstens 90% Identität, besonders bevorzugt wenigstens 95% Identität, am meisten bevorzugt wenigstens 97-99% Identität mit einem beliebigen aus SEQ ID NO: 1 bis 6 umfaßt; oder
 - d) eine Peptidsequenz, die eine Aminosäuresequenz mit wenigstens 15, 20, 30, 40, 50 oder 100 zusammenhängenden Aminosäuren aus der Aminosäuresequenz von SEQ ID NO: 7 oder SEQ ID NO: 8 umfaßt.
5. Fusionsprotein, das ein Fusionspartnerprotein gemäß einem der Ansprüche 1 bis 4 und ein heterologes Protein umfaßt.
6. Fusionsprotein gemäß Anspruch 5, worin das heterologe Protein chemisch an den Fusionspartner konjugiert ist.
7. Fusionsprotein gemäß Anspruch 5 oder 6, worin das heterologe Protein aus einem Organismus stammt, der aus der folgenden Gruppe ausgewählt ist: humanes Immundefizienzvirus HIV-1, humane Herpes simplex-Viren, Cytomegalovirus, Rotavirus, Epstein-Barr-Virus, Varicella Zoster-Virus, aus einem Hepatitisvirus wie Hepatitis B-Virus, Hepatitis A-Virus, Hepatitis C-Virus und Hepatitis E-Virus, aus respiratorischem Synzytialvirus, Parainfluenzavirus, Masernvirus, Mumpsvirus, humane Papillomaviren, Flaviviren oder Influenzavirus, aus *Neisseria* spp., *Moraxella* spp., *Bordetella* spp., *Mycobacterium* spp., einschließlich *M. tuberculosis*; *Escherichia* spp., einschließlich enterotoxisches *E. coli*; *Salmonella* spp.; *Listeria* spp.; *Helicobacter* spp.; *Staphylococcus* spp.; einschließlich *S. aureus*, *S. epidermidis*; *Borrelia* spp.; *Chlamydia* spp., einschließlich *C. trachomatis*, *C. pneumoniae*; *Plasmodium* spp., einschließlich *P. falciparum*; *Toxoplasma* spp., *Candida* spp.
8. Fusionsprotein gemäß Anspruch 5 oder 6, worin das heterologe Protein ein Tumor-assoziiertes Protein oder gewebespezifisches Protein oder immunogenes Fragment davon ist.
9. Fusionsprotein gemäß Anspruch 8, worin das heterologe Protein oder Fragment davon ausgewählt ist aus MAGE 1, MAGE 3, MAGE 4, PRAME, BAGE, LAGE 1, LAGE 2, SAGE, HAGE, XAGE, PSA, PAP, PSCA, Prostein, P501S, HASH2, Cripto, B726, NY-BR1.1, P510, MUC-1, Prostase, STEAP, Tyrosinase, Telomerase, Survivin, CASB616, P53 oder her 2 neu.
10. Fusionsprotein gemäß einem der Ansprüche 6 bis 9, das ferner einen Affinitätsmarker mit wenigstens 4 Histidinresten umfaßt.
11. Nukleinsäuresequenz, die ein Protein gemäß Anspruch 1 bis 10 codiert.
12. Expressionsvektor, der eine Nukleinsäuresequenz gemäß Anspruch 11 umfaßt.
13. Wirtszelle, die mit einer Nukleinsäuresequenz gemäß Anspruch 11 oder mit einem Expressionsvektor gemäß Anspruch 12 transformiert ist.
14. Immunogene Zusammensetzung, die ein Protein gemäß einem der Ansprüche 1 bis 10 oder eine DNA-Sequenz gemäß Anspruch 11 und einen pharmazeutisch akzeptablen Exzipienten umfaßt.
15. Immunogene Zusammensetzung gemäß Anspruch 14, die zusätzlich einen TH-1-induzierenden Hilfs-

stoff umfaßt.

16. Immunogene Zusammensetzung gemäß Anspruch 15, worin der TH-1-induzierende Hilfsstoff aus der Gruppe von Hilfsstoffen ausgewählt ist, die 3D-MPL, QS21, eine Mischung aus QS21 und Cholesterol, ein CpG-Oligonukleotid oder eine Mischung aus zwei oder mehreren der Hilfsstoffe umfaßt.

17. Verfahren zur Herstellung einer immunogenen Zusammensetzung gemäß einem der Ansprüche 14 bis 16, das das Vermischen des Fusionsproteins gemäß einem der Ansprüche 6 bis 10 oder des codierenden Polynukleotids gemäß Anspruch 11 mit einem geeigneten Hilfsstoff, Verdünnungsmittel oder anderen pharmazeutisch akzeptablen Träger umfaßt.

18. Verfahren zur Herstellung eines Fusionsproteins gemäß einem der Ansprüche 1 bis 10, das das Kultivieren einer Wirtszelle gemäß Anspruch 13 unter Bedingungen, die ausreichend zur Herstellung des Fusionsproteins sind, und das Gewinnen des Fusionsproteins aus dem Kulturmedium umfaßt.

19. Protein gemäß einem der Ansprüche 1 bis 10 oder DNA-Sequenz gemäß Anspruch 11 zur Verwendung in der Medizin.

20. Verwendung eines Proteins gemäß einem der Ansprüche 1 bis 10 oder einer DNA-Sequenz gemäß Anspruch 11 in der Herstellung einer immunogenen Zusammensetzung zum Hervorrufen einer Immunreaktion in einem Patienten.

21. Verwendung gemäß Anspruch 20, worin die Immunreaktion durch aufeinanderfolgende Verabreichung i) des Proteins, gefolgt von der DNA-Sequenz; oder ii) der DNA-Sequenz, gefolgt vom Protein hervorgerufen wird.

22. Verwendung gemäß Anspruch 21, worin die DNA-Sequenz auf biologisch abbaubaren Perlen aufgetragen ist oder über einen Partikelbombardierungsansatz übertragen wird.

23. Verwendung gemäß Anspruch 21 oder 22, worin das Protein mit Hilfsstoff versetzt ist.

24. Verwendung eines Proteins gemäß einem der Ansprüche 1 bis 10 oder einer DNA-Sequenz gemäß Anspruch 11 in der Herstellung einer immunogenen Zusammensetzung zur immuntherapeutischen Behandlung eines Patienten, der an Krebs leidet oder dafür anfällig ist.

25. Verwendung gemäß Anspruch 24, worin der Krebs Prostatakrebs, Darmkrebs, Lungenkrebs, Brustkrebs oder Melanom ist.

Es folgen 45 Blatt Zeichnungen

Anhängende Zeichnungen

Fig. 1 - Sequenzangaben für C-LytA.

SEQ ID NO:1 - Aminosäuresequenz von C-LytA Repeat 1

GWQKNDTGYWVHSD 15

SEQ ID NO:2 - Aminosäuresequenz von C-LytA Repeat 2

GSYPKDKFEKINGTWYFDSS 21

SEQ ID NO:3 - Aminosäuresequenz von C-LytA Repeat 3

GYMLADRWRKHTDGNWYFDNS 22

SEQ ID NO:4 - Aminosäuresequenz von C-LytA Repeat 4

GEMATGWKKIADKWYFNEE 20

SEQ ID NO:5 - Aminosäuresequenz von C-LytA Repeat 5

GAMKTGWVKYKDTWYLLDAKE 21

SEQ ID NO:6 - Aminosäuresequenz von C-LytA Repeat 6

GAMVSNAFIQSADGTWYLLKPD 23

SEQ ID NO:7 - Aminosäuresequenz von C-LytA Cholin-Bindungsdomäne

GWQKNDTGYWVHSDGSYPKDKFEKINGTWYFDSSGYMLADRWRKHTDGNWYFDNSGEMATGWKKIADKWYFNEEGAMKTGWVKYKDTWYLLDAKEGAMVSNAFIQSADGTWYLLKPD 60PDGTLADRPEFTVEPDGLITVK 142

SEQ ID NO:8 - Aminosäuresequenz von C-LytA-Domäne aus trunkiertem Repeat 1 bis Repeat 6 (als Teil unserer in Figur 2 gezeigten Konstrukte)

YVHSDGSYPKDKFEKINGTWYFDSSGYMLADRWRKHTDGNWYFDNSGEMATGWKKIADKWYFNEEGAMKTGWVKYKDTWYLLDAKEGAMVSNAFIQSADGTWYLLKPD

SEQ ID NO:9 - DNA-Sequenz, die die Aminosäuresequenz von SEQ ID NO:1 codiert

ggctggcaga agaatgacac tggctactgg tacgtacatt cagac

SEQ ID NO:10 - DNA-Sequenz, die die Aminosäuresequenz von SEQ ID NO:2 codiert

ggctcttata caaaagacaa gtttgagaaa atcaatggca cttggtacta ctttgacagt tca

SEQ ID NO:11 - DNA-Sequenz, die die Aminosäuresequenz von SEQ ID NO:3 codiert

ggctatatgc ttgcagaccg ctggaggaag cacacagacg gcaactggta ctggttcgac aactca

SEQ ID NO:12 - DNA-Sequenz, die die Aminosäuresequenz von SEQ ID NO:4 codiert

ggcgaaatgg ctacaggctg gaagaaaatc gctgataagt ggtactattt caacgaagaa

SEQ ID NO:13 - DNA-Sequenz, die die Aminosäuresequenz von SEQ ID NO:5 codiert

Ggtgccatga agacaggctg ggtcaagtac aaggacactt ggtactactt agacgctaaa gaa

SEQ ID NO:14 - DNA-Sequenz, die die Aminosäuresequenz von SEQ ID NO:6 codiert

**Ggcgccatgg tatcaaatgc ctttatccag tcagcggacg gaacaggctg gtactacctc
aaaccagac**

SEQ ID NO:15 - DNA-Sequenz, die die Aminosäuresequenz von SEQ ID NO:7 codiert

**ggctggcaga agaatgacac tggctactgg tacgtacatt cagacggctc ttatccaaaa 60
gacaagtttg agaaaatcaa tggcacttgg tactactttg acagttcagg ctatatgctt 120
gcagaccgct ggaggaagca cacagacggc aactgggtact ggttcgacaa ctcaggcgaa 180
atggctacag gctggaagaa aatcgctgat aagtgggtact atttcaacga agaaggtgcc 240
atgaagacag gctgggtcaa gtacaaggac acttgggtact acttagacgc taaagaaggc 300
gccatggtat caaatgcctt tatccagtca gcggacggaa caggctggta ctacctcaaa 360
ccagacggaa cactggcaga caggccagaa ttcacagtag agccagatgg cttgattaca 420
gtaaaataa 429**

SEQ ID NO:16 - DNA-Sequenz, die die Aminosäuresequenz von SEQ ID NO:8 codiert

**TACGTACATTCCGACGGCTCTTATCCAAAAGACAAGTTTGAGAAAATCAATGGCACTTGGTACTACTTTGACA
GTTTCAGGCTATATGCTTGCAGACCGCTGGAGGAAGCACACAGACGGCAACTGGTACTGGTTCGACAACTCAGG
CGAAATGGCTACAGGCTGGAAGAAAATCGCTGATAAGTGGTACTATTTCAACGAAGAAGGTGCCATGAAGACA
GGCTGGGTCAAGTACAAGGACACTTGGTACTACTTAGACGCTAAAGAAGGCGCCATGGTATCAAATGCCTTTA
TCCAGTCAGCGGACGGAACAGGCTGGTACTACCTCAAACCAGAC**

Fig. 2 CPC und native Konstrukte

Konstrukt 1 - codierende Sequenz von CPC-P501₅₁₋₅₅₃ (siehe Plasmid aus Figur 7 - Y1796)

Proteinsequenz (SEQ ID NO:27)

R1 R2 R3 R4
 MAAAYVHSDGSYPKDKFEKINGTWYYFDSSGYMLADRWRKHTDGNWYWFDNSGEMATG
 R5 P2 R6
 WKKIADKWWYFNEEGAMKTGWVKYKDTWYYLDAKEGAMQYIKANSKEIGITEGV MVSNAFIQS
 ADGTGWYYLKPDGTLADRPEKFMVMVLGIGPVLGLVCVPLLGSASDHWGRGRYGRRRPFTWALS
 GILLSLFLPRAGWLAGLLCPDRPLELALLILGVLLDFCGQVCFTPLEALLSDLFRDPDHCRAYSV
 YAFMISLGGCLGYLLPAIDWDTALAPYLGTEECFLGLLTLIFLTCVAATLLVAEEAALGPTEPAEG
 LSAPSLSPHCCPCRARLAFRNLGALLPRLHLQCCRMPTLRRLFVAELCSWMALMTFTLFYTDVGE
 GLYQGVPRAEPTARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVAAFPV
 AAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGGASSEDLSMTSF
 LPGPKPGAPFPNGHVAGGSGLLPPPALCGASACDVSVRVVVGEPTARVVPGRGICLDLAILDSAF
 LLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQVVFDDKSDLAKYSAGGHHHHHH

R1 (normal): Aminosäuren 5-9 (Fragment)	R4 (fett): Aminosäuren 53-72	P2 (unterstrichen): 97-110
R2 (fett): Aminosäuren 10-30	R5 (normal): Aminosäuren 73-93	
R3 (normal): Aminosäuren 31-52	R6a (fett): Aminosäuren 94-95	R6b (fett): 113-133

Nukleotidsequenz (SEQ ID NO:28)

ATGgcggcgctTACGTACATTCCGACGGCTCTTATCCAAAAGACAAGTTTGAGAAAATCAATGGCACTTGGT
 ACTACTTTGACAGTTTCAGGCTATATGCTTGACAGCCGCTGGAGGAAGCACACAGACGGCACTGGTACTGGTT
 CGACAACCTCAGGCGAAATGGCTACAGGCTGGAAGAAAATCGCTGATAAGTGGTACTATTTCAACGAAGAAGGT
 GCCATGAAGACAGGCTGGGTCAAGTACAAGGACACTTGGTACTACTTAGACGCTAAAGAAGGCGCCatgcaat
 acatcaaggctaactctaagttcattggtatcactgaaggcgctATGGTATCAAATGCCTTTATCCAGTCAGC
 GGACGGAACAGGCTGGTACTACCTCAAACCAGACGGAACACTGGCAGACAGGCCAGAAAagttcatgtaCatg
 GTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTG
 GACGCTATGGCCGCCGCCGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCC
 AAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCGGATCCCAGGCCCTGGAGCTGGCACTGCTCATCCTGGGC
 GTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGG
 ACCCGGACCACTGTGCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCT
 CCTGCCTGCCATTGACTGGGACACCACTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCTCTTTGGC
 CTGCTCACCTCATCTTCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCA
 CCGAGCCAGCAGAAGGGCTGTGCGCCCCCTCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTT

CCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTC
 TTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTTCGTGGGCGAGG
 GGCTGTACCAGGGCGTGCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTTCGGAT
 GGGCAGCCTGGGGCTGTTCTCTGCACTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAG
 CGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT
 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCC
 CTACACACTGGCCTCCCTCTACCACCGGGAAGCAGGTGTTCTGCCCCAAATACCGAGGGGACACTGGAGGT
 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGAC
 ACGTGGGTGCTGGAGGCAGTGGCCTGCTCCACCTCCACCCGCGCTCTGCGGGGCTCTGCCTGTGATGTCTC
 CGTACGTGTGGTGGTGGGTGAGCCACCGAGGCCAGGGTGGTTCGGGGCCGGGGCATCTGCCTGGACCTCGCC
 ATCCTGGATAGTGCCTTCTGCTGTCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCC
 AGTCTGTCACTGCCTATATGGTGTCTGCCGAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT
 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGggtggacaccatcaccatcaccattaa

Konstrukt 2 - codierende Sequenz von P501₅₁₋₅₅₃-HIS (Kontrolle) (Hefestamm SC333)

Proteinsequenz (SEQ ID NO:29)

MVLGIGPVLG	LVCVPLLGSA	SDHWRGRYGR	RRPFIWALS	L GILLSLFLIP	RAGWLAGLLC	60
PDPRPLELAL	LILGVGLLDF	CGQVCFTPLE	ALLSDLFRDP	DHCRQAYSVY	AFMISLGGLC	120
GYLLPAIDWD	TSALAPYLGT	QEECLFGLLT	LIFLTCVAAT	LLVAEEAALG	PTEPAEGLSA	180
PSLSPHCCPC	RARLAFRNLG	ALLPRLHQLC	CRMPRTLRL	FVAELCSWMA	LMTFTLFYTD	240
FVGEGLYQGV	PRAEPGTEAR	RHYDEGVRMG	SLGLFLQCAI	SLVFSLVMDR	LVQRFQTRAV	300
YLASVAAPFV	AAGATCLSHS	VAVVTASAAL	TGFTFSALQI	LPYTLASLYH	REKQVFLPKY	360
RGDTGGASSE	DSLMTSFLPG	PKPGAPFPNG	HVGAGGSGLL	PPPPALCGAS	ACDVSVRVVV	420
GEPTEARVVP	GRGICLDLAI	LDSAFLLSQV	APSLFMGSIV	QLSQSVTAYM	VSAAGLGLVA	480
IYFATQVVFD	KSDLAKYSAG	GHHHHHH				507

Nukleotidsequenz (SEQ ID NO:30)

atgGTGCTGG	GCATTGGTCC	AGTGTGGGC	CTGGTCTGTG	TCCCGCTCCT	AGGCTCAGCC	60
AGTGACCACT	GGCGTGGACG	CTATGGCCGC	CGCCGGCCCT	TCATCTGGGC	ACTGTCTTG	120
GGCATCCTGC	TGAGCCTCTT	TCTCATCCCA	AGGGCCGGCT	GGCTAGCAGG	GCTGCTGTGC	180
CCGGATCCCA	GGCCCCTGGA	GCTGGCACTG	CTCATCCTGG	GCGTGGGGCT	GCTGGACTTC	240
TGTGGCCAGG	TGTGCTTCAC	TCCACTGGAG	GCCCTGCTCT	CTGACCTCTT	CCGGGACCCG	300
GACCACTGTC	GCCAGGCCTA	CTCTGTCTAT	GCCTTCATGA	TCAGTCTTGG	GGGCTGCCTG	360
GGCTACCTCC	TGCCTGCCAT	TGACTGGGAC	ACCAGTGCCC	TGGCCCCCTA	CCTGGGCACC	420
CAGGAGGAGT	GCCTCTTTGG	CCTGCTCACC	CTCATCTTCC	TCACCTGCGT	AGCAGCCACA	480
CTGCTGGTGG	CTGAGGAGGC	AGCGCTGGGC	CCCACCGAGC	CAGCAGAAGG	GCTGTGCGCC	540
CCCTCCTTGT	CGCCCCACTG	CTGTCCATGC	CGGGCCCGCT	TGGCTTTCCG	GAACCTGGGC	600

```

GCCCTGCTTC CCCGGCTGCA CCAGCTGTGC TGCCGCATGC CCCGCACCCT GCGCCGGCTC 660
TTCGTGGCTG AGCTGTGCAG CTGGATGGCA CTCATGACCT TCACGCTGTT TTACACGGAT 720
TTCGTGGGCG AGGGGCTGTA CCAGGGCGTG CCCAGAGCTG AGCCGGGCAC CGAGGCCCGG 780
AGACACTATG ATGAAGGCGT TCGGATGGGC AGCCTGGGGC TGTTCTTGCA GTGCGCCATC 840
TCCCTGGTCT TCTCTCTGGT CATGGACCGG CTGGTGCAGC GATTCGGCAC TCGAGCAGTC 900
TATTTGGCCA GTGTGGCAGC TTTCCCTGTG GCTGCCGGTG CCACATGCCT GTCCACAGT 960
GTGGCCGTGG TGACAGCTTC AGCCGCCCTC ACCGGGTTC CTTCTCAGC CCTGCAGATC 1020
CTGCCCTACA CACTGGCCTC CCTCTACCAC CGGGAGAAGC AGGTGTTCTT GCCCAAATAC 1080
CGAGGGGACA CTGGAGGTGC TAGCAGTGAG GACAGCCTGA TGACCAGCTT CCTGCCAGGC 1140
CCTAAGCCTG GAGCTCCCTT CCCTAATGGA CACGTGGGTG CTGGAGGCAG TGGCCTGCTC 1200
CCACCTCCAC CCGCGCTCTG CGGGGCCTCT GCCTGTGATG TCTCCGTACG TGTGGTGGTG 1260
GGTGAGCCCA CCGAGGCCAG GGTGGTTCCG GGCCGGGGCA TCTGCCTGGA CCTCGCCATC 1320
CTGGATAGTG CCTTCTGCT GTCCCAGGTG GCCCATCCC TGTTTATGGG CTCCATTGTC 1380
CAGCTCAGCC AGTCTGTCAC TGCCTATATG GTGTCTGCCG CAGGCCTGGG TCTGGTCGCC 1440
ATTACTTTG CTACACAGGT AGTATTTGAC AAGAGCGACT TGGCCAAATA CTCAGCGgt 1500
ggacaccatc accatcacca ttaa 1524

```

Konstrukt 3 - codierende Sequenz von natssP501₁₋₃₄P501₅₁₋₅₅₃HIS (Hefestamm Y1800)

Proteinsequenz (SEQ ID NO:31)

```

                                R1      R2
MAAVQRLWVSRLLRHRKAQLLLNNLLTFGLEVCLAAAYVHSDGSYPKDKFEKINGTW
                                R3      R4      R5
YYFDSSGYMLADRWRKHTDGNWYWF DNSGEMATGWKKIADKWYFNEEGAMKTGWVK
                                P2      R6
YKDTWYYLDAKEGAMOYIKANSKEFIGITEGV MVSNAFIQSADGTGWYYLKPDTLADRPEKFMY
MVLGIGPVLGLVCVPLLGSASDHWGRYGRRRPFIWALSLGILLSLFLIPRAGWLAGLLCPDPRPLEL
ALLILGVGLLDFCGQVCFTPLEALLSDLFRDPDHCRQAYSVYAFMISLGGCLGYLLPAIDWDT SALAP
YLGTQEECLFGLLTLIFLTCVAATLLVAEEAALGPTEPAEGLSAPSLPHCCPCRARLAFRNLGALLPR
LHQLCCRMPTLRRLFVAELCSWMALMTFTLFYTFVGEGLYQGVPRAEPTARRHYDEGVRMG
SLGLFLQCAISLVFSLVMDRLVQRFGTRAVYLASVAAPVAAGATCLSHSVAVVTASAALTGFTFSA
LQILPYTLASLYHREKQVFLPKYRGDTGGASSEDLSMTSFLPGPKPGAPFPNGHVGAGGSGLLPPPPA
LCGASACDVSVRVVVGEPTARVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVS
AAGLGLVAIYFATQVVFDKSDLAKYSAGGHHHHHH

```

R1 (normal): Aminosäuren 38-42 (Fragment)	R4 (fett): Aminosäuren 77-106	P2 (unterstrichen): 130-1430
R2 (fett): Aminosäuren 43-64	R5 (normal): Aminosäuren 107-126	
R3 (normal): Aminosäuren 65-76	R6a (fett): Aminosäuren 127-128	R6b (fett): Aminosäuren 146-166

natss steht für native Signalsequenz

Nukleotidsequenz (SEQ ID NO:32)

ATGgcGGCCGTGCAGAGGCTATGGGTATCGAGACTGCTAAGACACCGCAAAGCTCAGTTGTTGTTGGTTAACT
 TGTTGACCTTCGGGCTGGAAGTCTGTTTGGCggcgcgtTACGTACATTCCGACGGCTCTTATCCAAAAGACAA
 GTTTGAGAAAATCAATGGCACTTGGTACTACTTTGACAGTTTCAAGGCTATATGCTTGCAGACCGCTGGAGGAAG
 CACACAGACGGCAACTGGTACTGGTTCGACAACTCAGGCGAAATGGCTACAGGCTGGAAGAAAATCGCTGATA
 AGTGGTACTATTTCAACGAAGAAGGTGCCATGAAGACAGGCTGGGTCAAGTACAAGGACACTTGGTACTACTT
 AGACGCTAAAGAAGGCGCCatgcaatacatcaaggctaactctaagttcattggtatcactgaaggcgtcATG
 GTATCAAATGCCTTTATCCAGTCAGCGGACGGAAACAGGCTGGTACTACCTCAAACCAGACGGAACACTGGCAG
 ACAGGCCAGAAaagttcatgtaCatgGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCCTCCT
 AGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGCCCTTCATCTGGGCCTGTCTTGGGC
 ATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCC
 TGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGA
 GGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTGCCAGGCCTACTCTGTCTATGCCTTCATGATC
 AGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGG
 GCACCCAGGAGGAGTGCTCTTTGGCCTGCTCACCTCATCTTCCTCACCTGCGTAGCAGCCCACTGCTGGT
 GGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTGGCCCCCTCCTTGTGCCCCCACTGC
 TGTCCATGCCGGGCCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCC
 GCATGCCCCGCACTTGCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCT
 GTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCAGAGCTGAGCCGGGCACCGAGGCCCCGG
 AGACACTATGATGAAGGCGTTCCGGATGGGCAGCCTGGGGCTGTTCTCTGCAGTGCGCCATCTCCCTGGTCTTCT
 CTCTGGTCATGGACCGCTGGTGACGATTCCGGCACTCGAGCAGTCTATTGGCCAGTGTTGGCAGCTTTCCC
 TGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTC
 ACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCAACGGGAGAAGCAGGTGTTCTCTGC
 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCTCTGCCAGGCCCTAA
 GCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCACCTCCACCCGCGCTC
 TGCGGGGCTCTGCCTGTGAtGTCTCCGTACGTGTGGTGGTGGGTGAGCCACCGAGGCCAGGTGGTTCCGG
 GCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCTTCCCTGCTGTCCCAGGTGGCCCCATCCCTGTT
 TATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGAGGCCTGGGTCTGGTC
 GCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGggtggacaccatc
 accatcaccattaa

Konstrukt 4 - Codierende Sequenz von alpha-prä-CPC-P501₅₁₋₅₅₃-HIS (Hefestamm Y1802)

Proteinsequenz (SEQ ID NO:33)

Alpha-pre	signal	R1	R2	R3
MAARFPSIFTAVLFAASSALAAA		YVHSDGGBYPKDKFEKINGTWYFDSSGYMLADRWRKHTDGNWYWF		
		R4	R5	P2
NSGEMATGWKKIADKWYFNEEGAMKTGWVKYKDTWYYLDAKEGAM		MOYIKANSKFIGITEGVMSNAFI		

R6

QSADGTGWYYLK**P**DTLADRPEKFMVMVLGIGPVLGLVCVPLLGSASDHWRGRYGRRRPFIWALSILGILLSLF
 LI PRAGWLAGLLCPDPRPLELALLILGVGLLDFCGQVCFTPLEALLSDLFRDPDHCRQAYSVYAFMISLGGCL
 GYLLPAIDWDTSALAPYLGTQBECLFGLLTLIFLTCVAATLLVAEEAALGPTEPAEGLSAPSLSPHCCPCRAR
 LAFRNLGALLPRLHQLCCMRPRTLRLFLVAELCSWMALMTFTLFYTDVFVGEGLYQGVPRAPGTEARRHYDEG
 VRMGSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVAAFVVAAGATCLSHSVAVVTASAALTGFTFSALQ
 ILPYTLASLYHREKQVFLPKYRGDTGGASSEDLSMTSFLPGPKPGAPFPNGHVGAGGSGLLPPPPALCGASAC
 DVSVRVVVGEPTEARVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFAT
 QVVFDKSDLAKYSAGGHHHHH

Alpha-prä-Signal (fett): Aminosäuren 4-22

R1 (normal): Aminosäuren 24-28 (Fragment)	R4 (fett): Aminosäuren 72-91	P2 (unterstrichen): 116-129
R2 (fett): Aminosäuren 29-49	R5 (normal): Aminosäuren 92-112	
R3 (normal): Aminosäuren 50-71	R6a (fett): Aminosäuren 113-114	R6b (fett): Aminosäuren 132-152

Alphapre steht für alpha-prä-Signalsequenz

Nukleotidsequenz (SEQ ID NO:34)

TACGTACATTCCGACGGCTCTTATCCAAAAGACAAGTTTGAGAAAATCAATGGCACTTGGTACTACTTTGACA
 GTTCAGGCTATATGCTTGCAGACCGCTGGAGGAAGCACACAGACGGCAACTGGTACTGGTTCGACAACCTCAGG
 CGAAATGGCTACAGGCTGGAAGAAAATCGCTGATAAGTGGTACTATTTCAACGAAGAAGGTGCCATGAAGACA
 GGCTGGGTCAAGTACAAGGACACTTGGTACTACTTAGACGCTAAAGAAGGCGCCatgcaatacatcaaggcta
actctaagttcattggtatcactgaaggcgctATGGTATCAAATGCCTTTATCCAGTCAGCGGACGGAACAGG
 CTGGTACTACCTCAAACCAGACGGAACACTGGCAGACAGGCCAGAA

ATGgcGGCCAGATTTCTTCAATTTTTACTGCAGTTTTATTGCGAGCATCCTCCGCATTAGCggccgctTACG
 TACATTCCGACGGCTCTTATCCAAAAGACAAGTTTGAGAAAATCAATGGCACTTGGTACTACTTTGACAGTTC
 AGGCTATATGCTTGCAGACCGCTGGAGGAAGCACACAGACGGCAACTGGTACTGGTTCGACAACCTCAGGCGAA
 ATGGCTACAGGCTGGAAGAAAATCGCTGATAAGTGGTACTATTTCAACGAAGAAGGTGCCATGAAGACAGGCT
 GGGTCAAGTACAAGGACACTTGGTACTACTTAGACGCTAAAGAAGGCGCCatgcaatacatcaaggctaactc
taagttcattggtatcactgaaggcgctATGGTATCAAATGCCTTTATCCAGTCAGCGGACGGAACAGGCTGG
 TACTACCTCAAACCAGACGGAACACTGGCAGACAGGCCAGAAgctggtattacttacgttccaccattggtgt
 tggaagttggtgttgaagaaagttcatgtaCatgGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGT
 CCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCCGGCCCTTCATCTGGGCACTG
 TCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATC
 CCAGGCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCAC
 TCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTGCGCCAGGCCTACTCTGTCTATGCT
 TCATGATCAGTCTTGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCC
 CTACCTGGGCACCCAGGAGGAGTGCCCTCTTGGCCTGCTCACCTCATCTTCTCACCTGCGTAGCAGCCACA

CTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTGCGCCCCCTCCTTGTCGC
 CCCACTGCTGTCCATGCCGGGCGCGCTTGGCTTTTCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCT
 GTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACC
 TTCACGCTGTTTTACACGGATTTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCAGAGCTGAGCCGGGCACCG
 AGGCCCCGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCTGCAGTGCGCCATCTCCCT
 GGTCTTCTCTCTGGTCATGGACCGGCTGGTGACGCGATTTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA
 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCA
 CCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGT
 GTTCCTGCCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCTCTGCCA
 GGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCAC
 CCGCGCTCTGCGGGGCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT
 GGTTCGGGCGGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCTCTGCTGTCCCAGGTGGCCCCA
 TCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGG
 GTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGggtgg
 acaccatcaccatcaccattaa

Konstrukt 5 - Codierende Sequenz von alpha-präpro-P501₅₁₋₅₅₃-HIS (in Plasmid pRIT 15068 und Hefestamm Y1790)

Proteinsequenz (SEQ ID NO:35)

MSFLNFTAVL	FAASSALAAP	VNTTTEDETA	QIPAEAVIGY	SDLEGDFDVA	VLPFSNSTNN	60
GLLFINTTIA	SIAAKEEGVS	LEKREAEAMV	LGIGPVLGLV	CVPLLGSASD	HWRGRYGRRR	120
PFIWALSIGI	LLSLFLIPRA	GWLAGLLCPD	PRPLELALLI	LGVGLLDFCG	QVCFTPLEAL	180
LSDLFRDPDH	CRQAYSVYAF	MISLGGCLGY	LLPAIDWDTS	ALAPYLGTQE	ECLFGLLTLI	240
FLTCVAATL	VAEEAALGPT	EPAEGLSAPS	LSPHCCPCRA	RLAFRNLGAL	LPRLHQLCCR	300
MPRTLRLRFV	AELCSWMALM	TFTLFYTDFV	GEGLYQGVP	AEPGTEARRH	YDEGVRMGSL	360
GLFLQCAISL	VFSLVMDRLV	QRFQTRAVYL	ASVAAPVAA	GATCLSHSVA	VVTASAALTG	420
FTFSALQILP	YTLASLYHRE	KQVFLPKYRG	DTGGASSEDS	LMTSFLPGPK	PGAPFPNGHV	480
GAGGSGLLPP	PPALCGASAC	DVSVRVVVE	PTEARVVPGR	GICLDLAILD	SAFLLSQVAP	540
SLFMGSIVQL	SQSVTAYMVS	AAGLGLVAIY	FATQVVFDDK	DLAKYSAGGH	HHHHH	595

Nukleotidsequenz (SEQ ID NO:36)

ATGAGTTTCC	TCAATTTTAC	TGCAGTTTTA	TTGCGCAGCAT	CCTCCGCATT	AGCTGCTCCA	60
GTCAACACTA	CAACAGAAGA	TGAAACGGCA	CAAATTCGGG	CTGAAGCTGT	CATCGGTTAC	120
TCAGATTTAG	AAGGGGATTT	CGATGTTGCT	GTTTTGCCAT	TTTCCAACAG	CACAAATAAC	180
GGGTTATTGT	TTATAAATAC	TACTATTGCC	AGCATTGCTG	CTAAAGAAGA	AGGGGTATCT	240
CTCGAGAAAA	GAGAGGCTGA	AGCCatgGTG	CTGGGCATTG	GTCCAGTGCT	GGGCCTGGTC	300
TGTGTCCCGC	TCCTAGGCTC	AGCCAGTGAC	CACTGGCGTG	GACGCTATGG	CCGCCGCCGG	360

```

CCCTTCATCT GGGCACTGTC CTTGGGCATC CTGCTGAGCC TCTTTCTCAT CCCAAGGGCC 420
GGCTGGCTAG CAGGGCTGCT GTGCCCCGAT CCCAGGCCCC TGGAGCTGGC ACTGCTCATC 480
CTGGGCGTGG GGCTGCTGGA CTTCTGTGGC CAGGTGTGCT TCACTCCACT GGAGGCCCTG 540
CTCTCTGACC TCTTCCGGGA CCCGGACCAC TGTCGCCAGG CCTACTCTGT CTATGCCTTC 600
ATGATCAGTC TTGGGGGCTG CCTGGGCTAC CTCCTGCCTG CCATTGACTG GGACACCACT 660
GCCCTGGCCC CCTACCTGGG CACCCAGGAG GAGTGCCTCT TTGGCCTGCT CACCCTCATC 720
TTCCTCACCT GCGTAGCAGC CACACTGCTG GTGGCTGAGG AGGCAGCGCT GGGCCCCACC 780
GAGCCAGCAG AAGGGCTGTC GGCCCCCTCC TTGTGCGCCC ACTGCTGTCC ATGCCGGGCC 840
CGCTTGGCTT TCCGGAACCT GGGCGCCCTG CTTCCCCGGC TGCACCAGCT GTGCTGCCGC 900
ATGCCCGCA CCCTGCGCCG GCTCTTCGTG GCTGAGCTGT GCAGCTGGAT GGGACTCATG 960
ACCTTCACGC TGTTTTACAC GGATTTCGTG GCGAGGGGC TGTACCAGG CGTGCCCAGA 1020
GCTGAGCCGG GCACCGAGGC CCGGAGACAC TATGATGAAG GCGTTCGGAT GGGCAGCCTG 1080
GGGCTGTTCC TGCAGTGCGC CATCTCCCTG GTCTTCTCTC TGGTCATGGA CCGGCTGGTG 1140
CAGCGATTCT GCACTCGAGC AGTCTATTTG GCCAGTGTGG CAGCTTTCCC TGTGGCTGCC 1200
GGTGCCACAT GCCTGTCCCA CAGTGTGGCC GTGGTGACAG CTTCAGCCGC CCTCACCGGG 1260
TTCACCTTCT CAGCCCTGCA GATCCTGCCC TACACACTGG CCTCCCTCTA CCACCGGGAG 1320
AAGCAGGTGT TCCTGCCCAA ATACCGAGGG GACACTGGAG GTGCTAGCAG TGAGGACAGC 1380
CTGATGACCA GCTTCCTGCC AGGCCCTAAG CCTGGAGCTC CTTCCCTAA TGGACACGTG 1440
GGTGCTGGAG GCAGTGGCCT GCTCCACCT CCACCCGCGC TCTGCGGGGC CTCTGCCTGT 1500
GAtGTCTCCG TACGTGTGGT GGTGGGTGAG CCCACCGAGG CCAGGGTGGT TCCGGGCCGG 1560
GGCATCTGCC TGGACCTCGC CATCCTGGAT AGTGCCTTCC TGCTGTCCCA GGTGGCCCCA 1620
TCCCTGTTTA TGGGCTCCAT TGTCCAGCTC AGCCAGTCTG TCACTGCCTA TATGGTGTCT 1680
GCCGAGGCC TGGGTCTGGT CGCCATTAC TTTGCTACAC AGGTAGTATT TGACAAGAGC 1740
GACTTGGCCA AATACTCAGC Gggtggacac catcaccatc accattaa 1788

```

Fig. 3. Struktur von CPC-p501-His-Fusionsprotein, exprimiert in *S. cerevisiae*

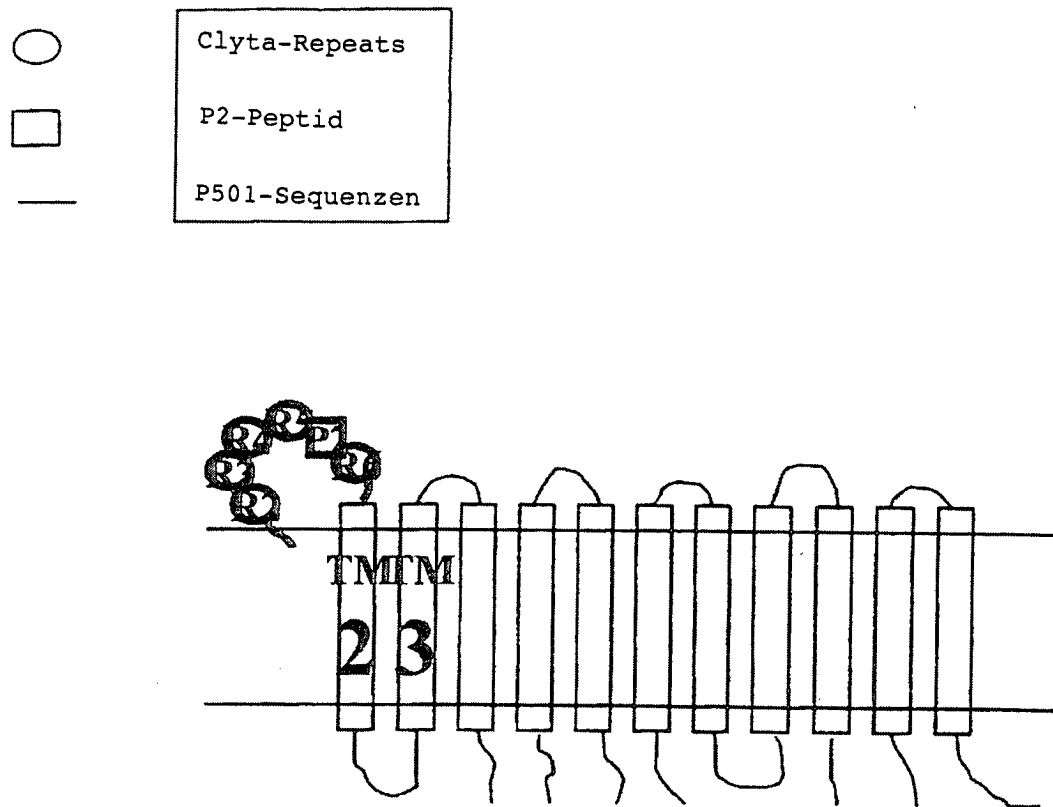


Fig. 4. Primärstruktur von CPC-P501-His-Fusionsprotein (SEQ ID NO:41)

```

MAAAYVHSDG SYPKDKFEKI NGTWYYFDSS GYMLADRWRK HTDGNWYWFD NSGEMATGWK 60
KIADKWYYFN EEGAMKTGWV KYKDTWYYLD AKEGAMQYIK ANSKFIGITE GVMVSNAFIQ 120
SADGTGWYYL KPDGTLADRP EKFMVMVLGI GPVLGLVCVP LLGSASDHW R GRYGRRRPFI 180
WALSLGILLS LFLIPRAGWL AGLLCPDPRP LELÄLLILGV GLLDFCGQVC FTPLEALLSD 240
LFRDPDHCRQ AYSVYAFMIS LGGCLGYLLP AIDWDTSALA PYLGTQEECL FGLLTLIFLT 300
CVAATLLVAE EAALGPTEPA EGLSAPSLSP HCCPCRARLA FRNLGALLPR LHQLCCRMPR 360
TLRRLFVAEL CSWMALMTFT LFYTDFVGEG LYQGVPRAP GTEARRHYDE GVRMGSLGLF 420
LQCAISLVFS LVMDRLVQRF GTRAVYLASV AAFPVAAGAT CLSHSVAVVT ASAALTGFTF 480
SALQILPYTL ASLYHREKQV FLPKYRGDTG GASSEDSLMT SFLPGPKPGA PFPNGHVGAG 540
GSGLLPPPPA LCGASACDVS VRVVVGEPTE ARVVPGRGIC LDLAILDSAF LLSQVAPSLF 600
MGSIVQLSQS VTAYMVSAAG LGLVAIYFAT QVVFDKSDLA KYSAGGHHHH HH 652

```

Fig. 5. Nukleotidsequenz von CPC-P501-His (pRIT15201) (SEQ ID NO:42)

```

ATGGCGGCCG CTTACGTACA TTCCGACGGC TCTTATCCAA AAGACAAGTT TGAGAAAATC 60
AATGGCACTT GGTACTACTT TGACAGTTCA GGCTATATGC TTGCAGACCG CTGGAGGAAG 120
CACACAGACG GCAACTGGTA CTGGTTTCGAC AACTCAGGCG AAATGGCTAC AGGCTGGAAG 180
AAAATCGCTG ATAAGTGGTA CTATTTCAAC GAAGAAGGTG CCATGAAGAC AGGCTGGGTC 240
AAGTACAAGG ACACCTGGTA CTACTTAGAC GCTAAAGAAG GCGCCATGCA ATACATCAAG 300
GCTAACTCTA AGTTCATTGG TACTACTGAA GCGTCATGG TATCAAATGC CTTTATCCAG 360
TCAGCGGACG GAACAGGCTG GTACTACCTC AAACCAGACG GAACACTGGC AGACAGGCCA 420
GAAAAGTTCA TGTACATGGT GCTGGGCATT GGTCCAGTGC TGGGCCTGGT CTGTGTCCCG 480
CTCCTAGGCT CAGCCAGTGA CCACTGGCGT GGACGCTATG GCCGCCGCCG GCCCTTCATC 540
TGGGCACTGT CTTGGGGCAT CCTGCTGAGC CTCTTTCTCA TCCCAAGGGC CGGCTGGCTA 600
GCAGGGCTGC TGTGCCCCGA TCCCAGGCCC CTGGAGCTGG CACTGCTCAT CCTGGGCGTG 660
GGGCTGCTGG ACTTCTGTGG CCAGGTGTGC TTTACTCCAC TGGAGGCCCT GCTCTCTGAC 720
CTCTTCCGGG ACCCGGACCA CTGTCGCCAG GCCTACTCTG TCTATGCCCT CATGATCAGT 780
CTTGGGGGCT GCCTGGGCTA CCTCCTGCCT GCCATTGACT GGGACACCAG TGCCCTGGCC 840
CCCTACCTGG GCACCCAGGA GGAGTGCCTC TTTGGCCTGC TCACCCTCAT CTTCTCACC 900
TGCGTAGCAG CCACACTGCT GGTGGCTGAG GAGGCAGCGC TGGGCCCCAC CGAGCCAGCA 960
GAAGGGCTGT CGGCCCCCTC CTTGTGCCCC CACTGCTGTC CATGCCGGGC CCGCTTGGCT 1020
TTCCGGAACC TGGGCGCCCT GCTTCCCCGG CTGCACCAGC TGTGCTGCCG CATGCCCCGC 1080
ACCCTGCGCC GGCTCTTCGT GGCTGAGCTG TGCAGCTGGA TGGCACTCAT GACCTTCACG 1140
CTGTTTTACA CGGATTTCGT GGGCGAGGGG CTGTACCAGG GCGTGCCCAG AGCTGAGCCG 1200
GGCACCGAGG CCCGGAGACA CTATGATGAA GCGTTTCGGA TGGGCAGCCT GGGGCTGTTC 1260
CTGCAGTGCG CCATCTCCCT GGTCTTCTCT CTGGTCATGG ACCGGCTGGT GCAGCGATTC 1320
GGCACTCGAG CAGTCTATTT GGCCAGTGTG GCAGCTTTCC CTGTGGCTGC CGGTGCCACA 1380
TGCCTGTCCC ACAGTGTGGC CGTGGTGACA GCTTCAGCCG CCCTCACC GGTTACCTTC 1440
TCAGCCCTGC AGATCCTGCC CTACACACTG GCCTCCCTCT ACCACCGGGA GAAGCAGGTG 1500
TTCCTGCCCC AATACCGAGG GGACACTGGA GGTGCTAGCA GTGAGGACAG CCTGATGACC 1560
AGCTTCCTGC CAGGCCCTAA GCCTGGAGCT CCCTTCCCTA ATGGACACGT GGGTGTCTGA 1620
GGCAGTGGCC TGCTCCCACC TCCACCCGCG CTCTGCGGGG CCTCTGCCTG TGATGTCTCC 1680
GTACGTGTGG TGGTGGGTGA GCCCACCAGG GCCAGGGTGG TTCCGGGGCC GGGCATCTGC 1740
CTGGACCTCG CCATCCTGGA TAGTGCTTC CTGCTGTCCC AGGTGGCCCC ATCCCTGTTT 1800
ATGGGCTCCA TTGTCCAGCT CAGCCAGTCT GTCAGTGCCT ATATGGTGTC TGCCGCAGGC 1860
CTGGGTCTGG TCGCCATTTA CTTTGCTACA CAGGTAGTAT TTGACAAGAG CGACTTGGCC 1920
AAATACTCAG CGGGTGGACA CCATCACCAT CACCATTAA 1959

```

Fig. 6. Klonierungsstrategie zur Erzeugung von Plasmid pRIT 15201

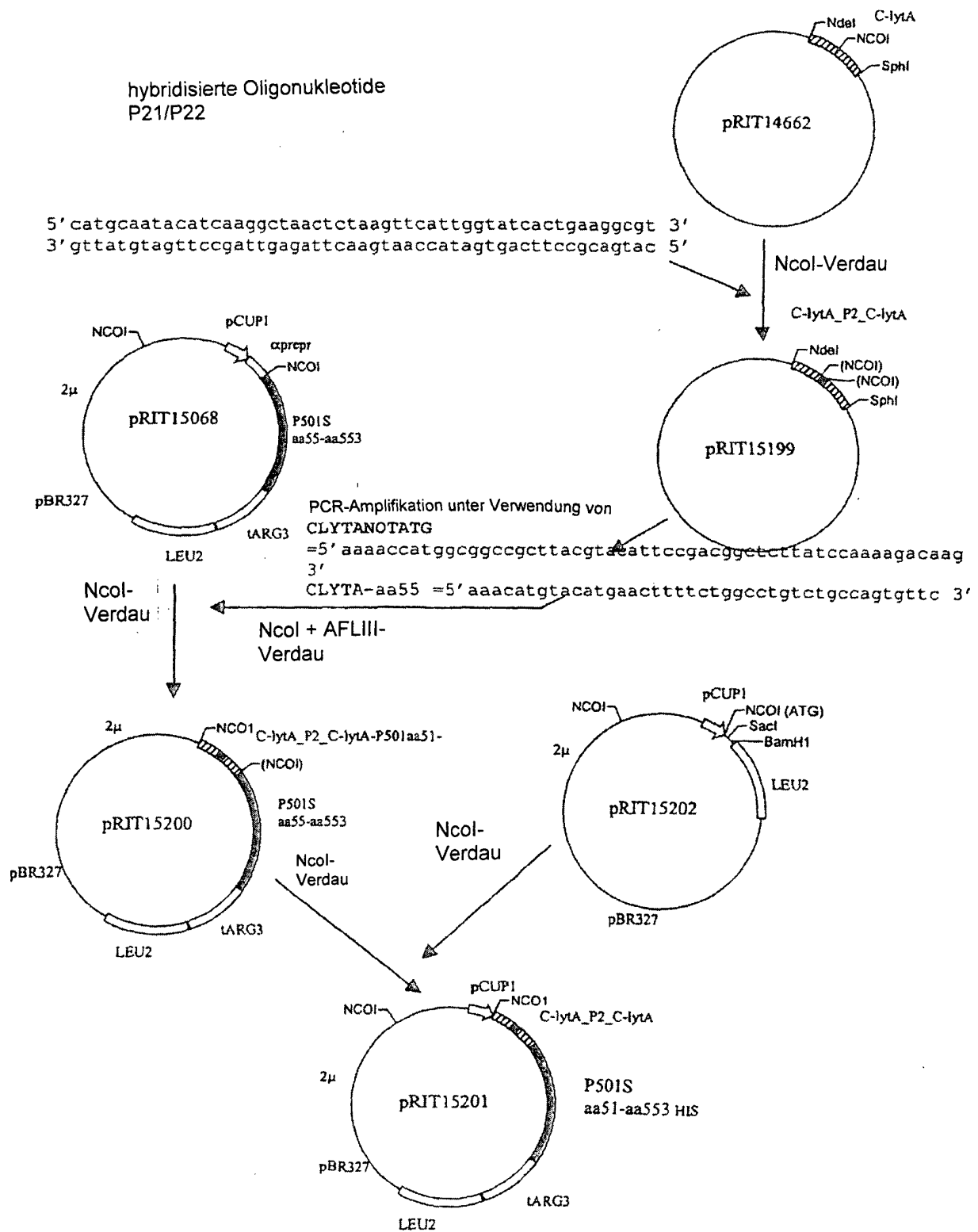
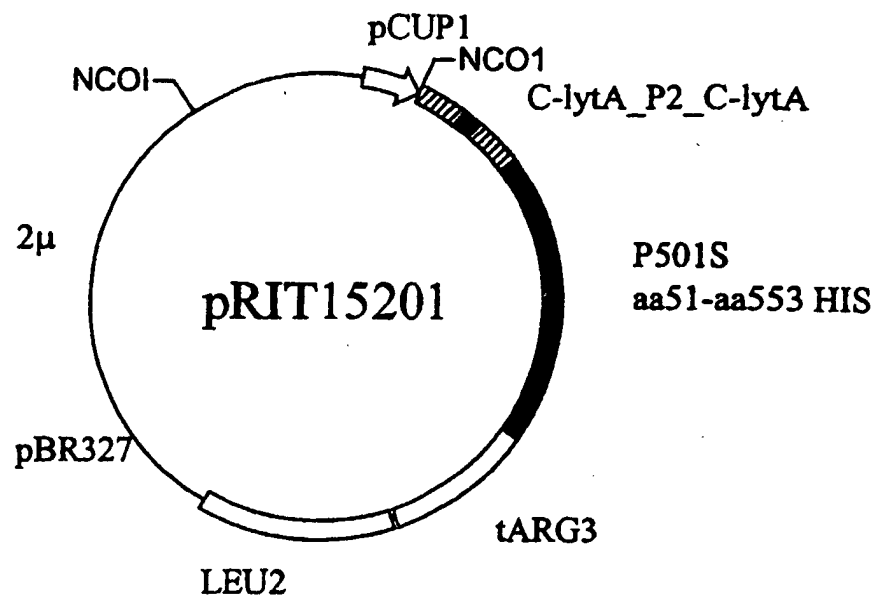
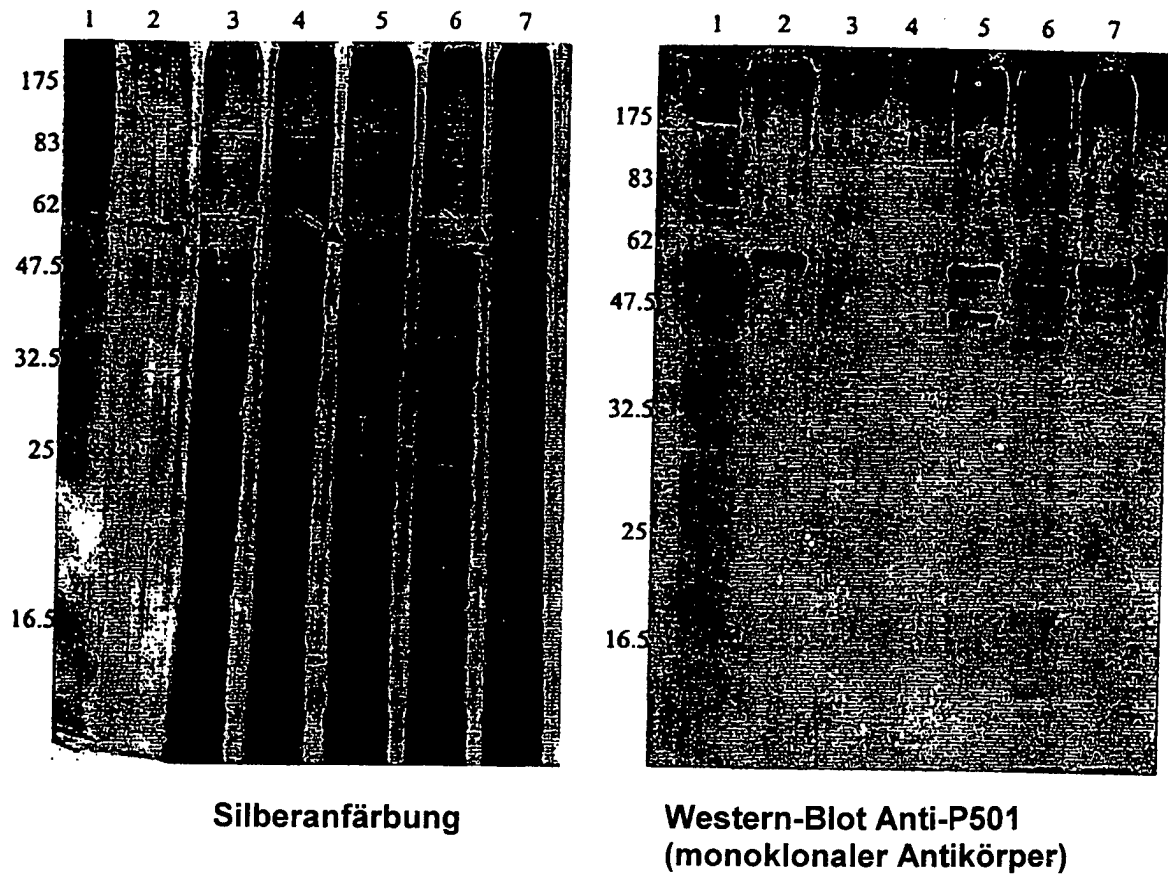


Fig. 7. Plasmidkarte von pRIT15201



Figur 8. Vergleichende Expression von CPC-P501 und P501 in *S. cerevisiae* Stamm DC 5 (Gel-Laemmli 10 %)



- 1 MW Biolabs (175/83/62/47,5/32,5/16,5 kDa)
- 2 Y1796 gereinigt
- 3 Y1795 Rohextrakt (Negativkontrolle)
- 4 SC333 Rohextrakt
- 5 Y1796 Rohextrakt
- 6 Y1790 Rohextrakt
- 7 Y1802 Rohextrakt

FIG. 9A.

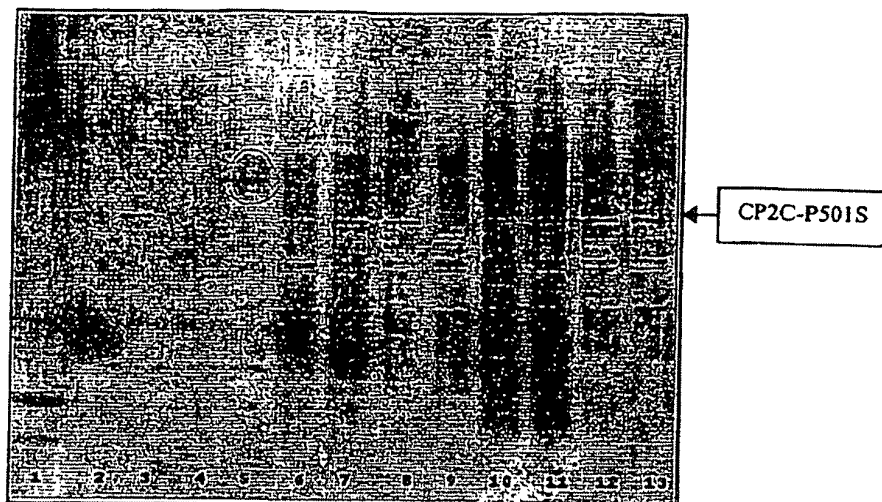
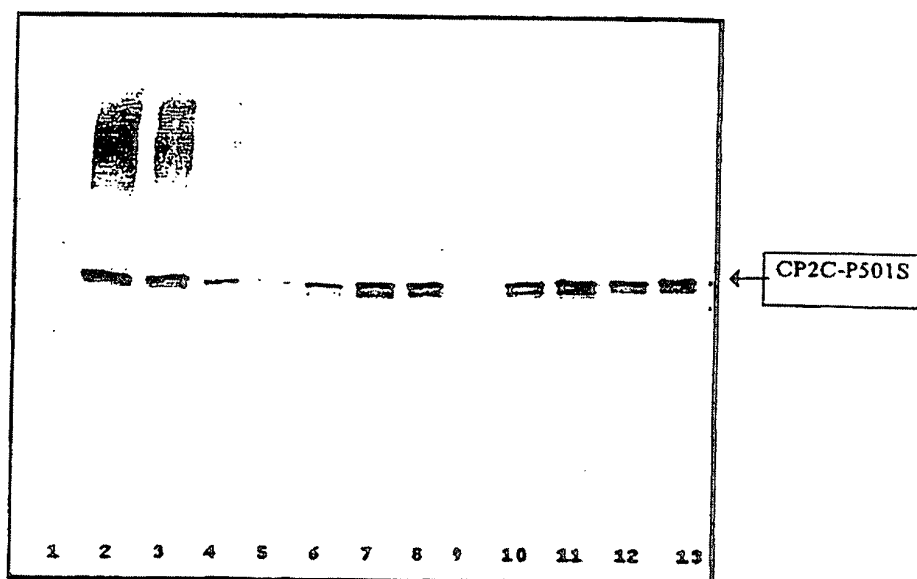


FIG. 9B.



- 1 - Molekulargewichtsmarker (Biolabs - Grow Range) 175; 83; 62; 47,5; 32,5; 25; 16,5; 6,5 kD - 10
- 2 - Gereinigte Referenz CP2CP501S/12 135 ng
- 3 - Gereinigte Referenz CP2CP501S/12 67,8 ng
- 4 - Gereinigte Referenz CP2CP501S/12 33,9 ng
- 5 - Gereinigte Referenz CP2CP501S/12 16,9 ng
- 6 - Fermentation PRO119-21h30
- 7 - Fermentation PRO124-21h30
- 8 - Fermentation PRO124-22h30
- 9 - Fermentation PRO127-0 h
- 10 - Fermentation PRO127-4 h
- 11 - Fermentation PRO127-6 h
- 12 - Fermentation PRO127-22h20
- 13 - Fermentation PRO127-22h45

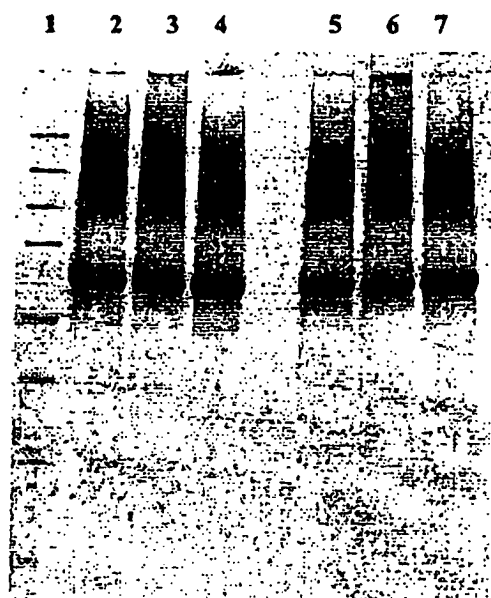
Fig. 10. Reinigungsschema von CPC-P501-His, erzeugt von Y1796

S. cerevisiae-Zellen	
↓	
Dynomill-Aufschluß	OD 120/2 Passagen/20 mM Tris pH 8,5 - 5 mM EDTA
↓	
Zentrifugieren	12,000 g/RT/90 min (Überstand verworfen)
↓	
Pelletwaschschritt 1	20 mM Tris pH 8,5 - 0,15 M NaCl - 2,0 M Guanidin-HCl - 0,1 % Empingen (30 min/RT)
↓	
Zentrifugieren	12,000 g/RT/60 min (Überstand verworfen)
↓	
Pelletwaschschritt 2	20 mM Tris pH 8,5 - 0,15 M NaCl - 4,0 M Harnstoff
↓	
Zentrifugieren	12,000 g/RT/60 min (Überstand verworfen)
↓	
Solubilisierung/Reduktion	20 mM Tris pH 8,5 - 0,15 M NaCl - 8,0 M Harnstoff - 1 % SDS - 0,2 M Glutathion (60 min/RT)
↓	
Zentrifugieren	12,000 g/RT/30 min (Pellet verworfen)
↓	
Carbamidomethylierung	0,3 M Iodacetamid (30 min/RT/Dunkelheit)/ pH eingestellt auf 8,5 (mit 5 M NaOH-Lösung) vor Inkubation
↓	
R/C-Überstand	
↓	
10-fache Verdünnung und pH-Einstellung (8,5)	<u>Verdünnungspuffer</u> : 20 mM Tris pH 8,5 - 1 M NaCl - 8,0 M Harnstoff
↓	

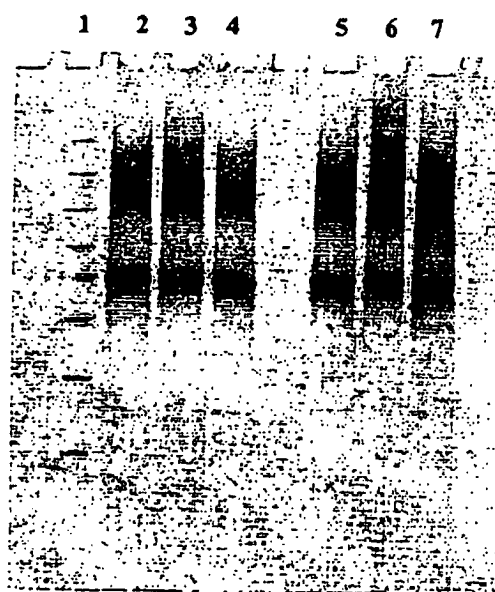
Fig. 10 (Fortsetzung)

Immobilisierte Metallionen-Affinitätschromatographie an Ni⁺⁺-komplexierender Sepharose FF (Amersham) (10x25 cm-Säule - 2000 ml)	<u>Äquilibrationpuffer:</u> 20 mM Tris pH 8,5 - 0,9 M NaCl - 8,0 M Harnstoff - 0,1 % SDS <u>Waschpuffer:</u> 1) Äquilibrationpuffer 2) 20 mM Tris pH 8,5 - 0,15 M NaCl - 8,0 M Harnstoff - 0,1 % SDS 3) 20 mM Tris pH 8,5 - 8,0 M Harnstoff - 0,1 % Tween 80 <u>Elutionspuffer:</u> 20 mM Tris pH 8,5 - 8,0 M Harnstoff - 0,1 % Tween 80 - 0,5 M Imidazol
↓	
2-fache Verdünnung und pH-Einstellung (10,0)	20 mM Piperazin pH 10,0 - 8,0 M Harnstoff - 0,1 % Tween 80
↓	
Anionenaustauscherchromatographie an Q-Sepharose FF (Amersham) (2,6 x 6,5 cm-Säule - 35 ml)	<u>Äquilibrationpuffer:</u> 20 mM Piperazin pH 10,0 - 8,0 M Harnstoff - 0,1 % Tween 80 <u>Waschpuffer:</u> 1) Äquilibrationpuffer 2) 20 mM Tris pH 8,5 - 8,0 M Harnstoff - 0,1 % Tween 80 <u>Elutionspuffer:</u> 20 mM Tris pH 7,5 - 8,0 M Harnstoff - 0,1 % Tween 80 - 0,5 M NaCl
↓	
Aufkonzentrieren/Diafiltration (Pall-Omega 10 kDa - 200 cm²)	+/- 3-faches Aufkonzentrieren <u>Diafiltrationspuffer:</u> Tris 20 mM pH 7,5
↓	
Sterilfiltration (Millipore - Millex GV 0,22 µm)	
↓	
gereinigtes Volumen	<u>Endpuffer:</u> 20 mM Tris pH 7,5 - +/- 0,3 % Tween 80
↓	
Lagerung -20°C	

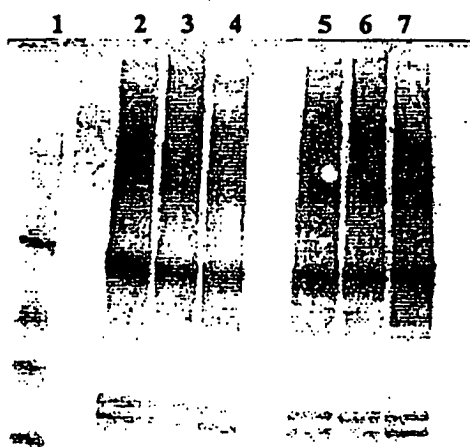
Fig. 11. Muster von gereinigtem CPC-P501-His-Protein (4-12 % Novex Nu-Page-vorgelassene Polyacrylamidgele)



Coomassie Blue R250



Daiichi-Silberanfärbung



**Western-Blot Anti-P501S
(Monoklonaler Antikörper)**

- 1: MW (250/150/75/50/37/25/15/10 kDa)
- 2: Gereinigter Ansatz A (reduzierende Bedingungen)
- 3: Gereinigter Ansatz B (reduzierende Bedingungen)
- 4: Gereinigter Ansatz C (reduzierende Bedingungen)
- 5: Gereinigter Ansatz A (nicht-reduzierende Bedingungen)
- 6: Gereinigter Ansatz B (nicht-reduzierende Bedingungen)
- 7: Gereinigter Ansatz C (nicht-reduzierende Bedingungen)

Fig. 12. Native P501S-Sequenz voller Länge (SEQ ID NO:17 & 43)

Nukleotidsequenz: SEQ ID NO:17

Polypeptidsequenz: SEQ ID NO:43

```

#####
GCCACCATGGTCCAGAGGCTGTGGGTGAGCCGCTGCTGCGGCACCGG
      M V Q R L W V S R L L R H R   14

AAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCA
K A Q L L L V N L L T F G L E V C L A A   34

GGCATCACCTATGTGCCGCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACC
G I T Y V P P L L L E V G V E E K F M T   54

ATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCC
M V L G I G P V L G L V C V P L L G S A   74

AGTGACCACTGGCGTGGACGCTATGGCCGCGCGCGGCCCTTCATCTGGGCACTGTCTTTC
S D H W R G R Y G R R R P F I W A L S L   94

GGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGC
G I L L S L F L I P R A G W L A G L L C   114

CCGGATCCCAGGCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTC
P D P R P L E L A L L I L G V G L L D F   134

TGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCG
C G Q V C F T P L E A L L S D L F R D P   154

GACCACTGTGCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTG
D H C R Q A Y S V Y A F M I S L G G C L   174

GGCTACCTCCTGCCTGCCATTGACTGGGACACCACTGCCCTGGCCCCCTACCTGGGCACC
G Y L L P A I D W D T S A L A P Y L G T   194

CAGGAGGAGTGCTCTTTGGCCTGCTCACCTCATCTTCCTCACCTGCGTAGCAGCCACA
Q E E C L F G L L T L I F L T C V A A T   214

CTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTGGGCC
L L V A E E A A L G P T E P A E G L S A   234

CCCTCCTTGTGCCCCACTGCTGTCCATGCCGGGCCCCGCTTGGCTTTCGGGAACCTGGGC
P S L S P H C C P C R A R L A F R N L G   254

GCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTC
A L L P R L H Q L C C R M P R T L R R L   274

TTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGAT
F V A E L C S W M A L M T F T L F Y T D   294

TTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCAGAGCTGAGCCGGGCACCGAGGCCCGG
F V G E G L Y Q G V P R A E P G T E A R   314

AGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTTCCTGCAGTGCGCCATC
R H Y D E G V R M G S L G L F L Q C A I   334

TCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCCGGCACTCGAGCAGTC
S L V F S L V M D R L V Q R F G T R A V   354

```

TATTTGGCCAGTGTGGCAGCTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGT
Y L A S V A A F P V A A G A T C L S H S 374

GTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCAGGTTACCTTCTCAGCCCTGCAGATC
V A V V T A S A A L T G F T F S A L Q I 394

CTGCCCTACACACTGGCCTCCCTCTACCACCGGAGAAGCAGGTGTTCCCTGCCCAAATAC
L P Y T L A S L Y H R E K Q V F L P K Y 414

CGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGC
R G D T G G A S S E D S L M T S F L P G 434

CCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTC
P K P G A P F P N G H V G A G G S G L L 454

CCACCTCCACCCGCGCTCTGCGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTG
P P P P A L C G A S A C D V S V R V V V 474

GGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATC
G E P T E A R V V P G R G I C L D L A I 494

CTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTC
L D S A F L L S Q V A P S L F M G S I V 514

CAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCC
Q L S Q S V T A Y M V S A A G L G L V A 534

ATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAG
I Y F A T Q V V F D K S D L A K Y S A * 554

GTCGAG

Fig. 13. Sequenz der CPC-P501S-Expressionskassette von JNW735 (SEQ ID NO:18 & 44)

Nukleotidsequenz: SEQ ID NO:18

Polypeptidsequenz: SEQ ID NO:44

```

#####
GCCACCATGGCGGCCGCTTACGTACATTCCGACGGCTCTTATCCAAAA
      M A A A Y V H S D G S Y P K 14

GACAAGTTTGAGAAAATCAATGGCACTTGGTACTACTTTGACAGTTCAGGCTATATGCTT
D K F E K I N G T W Y Y F D S S G Y M L 34

GCAGACCGCTGGAGGAAGCACACAGACGGCAACTGGTACTGGTTCGACAACTCAGGCGAA
A D R W R K H T D G N W Y W F D N S G E 54

ATGGCTACAGGCTGGAAGAAAATCGCTGATAAGTGGTACTATTTCAACGAAGAAGGTGCC
M A T G W K K I A D K W Y Y F N E E G A 74

ATGAAGACAGGCTGGGTCAAGTACAAGGACACTTGGTACTACTTAGACGCTAAAGAAGGC
M K T G W V K Y K D T W Y Y L D A K E G 94

GCCATGCAATACATCAAGGCTAACTCTAAGTTTATTGGTATCACTGAAGGCGTCATGGTA
A M Q Y I K A N S K F I G I T E G V M V 114

TCAAATGCCTTTATCCAGTCAGCGGACGGAACAGGCTGGTACTACCTCAAACCAGACGGA
S N A F I Q S A D G T G W Y Y L K P D G 134

ACACTGGCAGACAGGCCAGAAAAGTTTCATGTACATGGTGCTGGGCATTGGTCCAGTGCTG
T L A D R P E K F M Y M V L G I G P V L 154

GGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGC
G L V C V P L L G S A S D H W R G R Y G 174

CGCCGCCGCGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTCTCATC
R R R P F I W A L S L G I L L S L F L I 194

CCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCAGGCCCCCTGGAGCTGGCA
P R A G W L A G L L C P D P R P L E L A 214

CTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTG
L L I L G V G L L D F C G Q V C F T P L 234

GAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTGCGCCAGGCCTACTCTGTC
E A L L S D L F R D P D H C R Q A Y S V 254

TATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGG
Y A F M I S L G G C L G Y L L P A I D W 274

GACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTC
D T S A L A P Y L G T Q E E C L F G L L 294

ACCCTCATCTTCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTG
T L I F L T C V A A T L L V A E E A A L 314

GGCCCCACCGAGCCAGCAGAAGGGCTGTGCGGCCCTCCTTGTGCGCCCCACTGCTGTCCA
G P T E P A E G L S A P S L S P H C C P 334

```

TGCCGGGCCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTG
 C R A R L A F R N L G A L L P R L H Q L 354

TGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATG
 C C R M P R T L R R L F V A E L C S W M 374

GCACTCATGACCTTCACGCTGTTTTACACGGATTTTCGTGGGCGAGGGGCTGTACCAGGGC
 A L M T F T L F Y T D F V G E G L Y Q G 394

GTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTTCGGATG
 V P R A E P G T E A R R H Y D E G V R M 414

GGCAGCCTGGGGCTGTTCCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGCTCATGGAC
 G S L G L F L Q C A I S L V F S L V M D 434

CGGCTGGTGCAGCGATTTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCT
 R L V Q R F G T R A V Y L A S V A A F P 454

GTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCC
 V A A G A T C L S H S V A V V T A S A A 474

CTCACCGGGTTACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTAC
 L T G F T F S A L Q I L P Y T L A S L Y 494

CACCGGGAGAAGCAGGTGTTCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGT
 H R E K Q V F L P K Y R G D T G G A S S 514

GAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAAT
 E D S L M T S F L P G P K P G A P F P N 534

GGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCACCTCCACCCGCGCTCTGCGGGGCC
 G H V G A G G S G L L P P P P A L C G A 554

TCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCACCGAGGCCAGGGTGGTT
 S A C D V S V R V V V G E P T E A R V V 574

CCGGGCCGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCCTGCTGTCCCAG
 P G R G I C L D L A I L D S A F L L S Q 594

GTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTAT
 V A P S L F M G S I V Q L S Q S V T A Y 614

ATGGTGTCTGCCGAGGCCTGGGTCTGGTGCCTATTTACTTTGCTACACAGGTAGTATTT
 M V S A A G L G L V A I Y F A T Q V V F 634

GACAAGAGCGACTTGGCCAAATACTCAGCGTAGGTCGAG
 D K S D L A K Y S A * 645

Fig. 14 - Codon-optimierte P501S-Sequenzen (SEQ ID NO:19-20)

SEQ ID NO:19

ATGGTGCAGCGGCTCTGGGTGAGCCGCTCCTGCGGCATCGCAAGGCCAGCTCCTGCTGGTGAATCTGCTCA
 CATTGCGCCTGGAGGTGTGCCTGGCCGCCGCGCATCACCTACGTGCCCCCCTCCTGCTGGAGGTGGGAGTCGA
 GGAGAAGTTCATGACCATGGTGCTGGGCATTGGGCCCGTCTGGGCCTCGTGTGCGTGCCCTCTCCTCGGCAGC
 GCTTCCGACCATTTGGCGCGGCCGGTATGGCCGCAGGAGACCCTTCATCTGGGCTCTGAGTCTCGGCATCCTGC
 TGAGCCTGTTCCTGATCCCTCGGGCCGGCTGGCTGGCCGGGCTGCTGTGCCCCGATCCTCGGCCCCCTGGAGCT
 GGCCCTGCTGATCCTCGGCGTGGGCCTGCTGGACTTCTGCGGCCAGGTGTGCTTCACGCCCCCTGGAGGCACTG
 CTGAGCGACCTGTTCCGGGACCCCGACCATTTGCCGCCAGGCGTACAGCGTGTACGCCTTCATGATCTCCCTGG
 GAGGCTGCCTGGGCTACCTGCTCCCGCCATCGATTGGGACACCAGCGCACTCGCCCCCTATCTCGGAACACA
 GGAGGAATGCCTGTTCGGATTGTTGACGCTCATCTTCCTCACGTGCGTCGCGGCCACCCTGTTGGTGGCCGAG
 GAGGCCGCCCTGGGGCCCAACGAGCCGGCCGAGGGACTGAGCGCCCCGAGCCTGAGTCCACACTGCTGCCCTT
 GCCGGGCCCCGCTGGCCTTCCGTAATCTGGGCGCCCTCCTGCCTCGGCTCCATCAGCTGTGTTGCAGAATGCC
 TAGGACGCTGCGGCGCCTGTTGTCGCTGAGTTGTGCTCCTGGATGGCTCTCATGACCTTCACCCCTGTTTTAT
 ACGGACTTCGTCGCGGGAGGGCCTGTACCAGGGGTGCCGCGCGCCGAGCCCGGGACAGAGGCGCGCCGCGCACT
 ACGACGAGGGAGTGCATATGGGCTCCCTGGGCCTCTTCTTGCACTGCGCCATCAGTCTGGTTTTCTCTCTGGT
 CATGGACAGGCTGGTGCAGCGCTTCGGAACCCGGGCGGTGTACCTGGCGAGCGTGGCCGCCCTTCCCCGTGGCT
 GCCCGCGCCACCTGCCTCTCTCACTCGGTGGCCGTGGTCAACGCCAGCGCCGCCCTGACCGGGTTACACCTTCT
 CTGCCCTGCAGATTCTGCCTTACACCTGGCCAGCCTGTACCATCGCGAGAAACAGGTGTTTTCTCCCCAAGTA
 CAGAGGCGACACCGGGGGCGCCTCCAGCGAGGACAGCCTCATGACCTCCTTCCCTGCCTGGCCCCAAGCCCGGC
 GCCCCTTTCCCCAACGGGCACGTGGGCGCCGGCGGGAGTGGGCTCCTGCCCCCCCCCTCCTGCGCTGTGCGGGG
 CCAGCGCCTGCGACGTGAGCGTGCCTGGTGGTGGGCGAGCCACCGAGGCCCGCGTGGTGGCGGGCAGAGG
 CATTTGTCTGGACCTGGCCATCCTCGACTCCGCCTTCTCCTCAGCCAGGTGGCCCCGTCCCTCTTCATGGGC
 TCTATCGTCCAGCTGTCTCAGAGCGTCACCGCTTACATGGTGTCCGCTGCTGGACTGGGCTTGGTGGCTATTT
 ATTTGCGCCACCCAGGTGGTGTTCGACAAGAGCGACCTGGCCAAATACTCCGCCTGA

SEQ ID NO:20

ATGGTGCAGCGGCTGTGGGTGTCCCGCTGCTGCGCCATAGAAAGGCCAGTTGCTGCTGGTGAACCTGCTGA
 CTTTCGGACTGGAGGTGTGCCTGGCTGCCGGGATCACGTACGTGCCCCCCTGCTGCTGGAGGTGGGCGTGGA
 GGAGAAGTTCATGACAATGGTGCTGGGCATCGGCCCGTCTGGGCCTCGTGTGTGTGCCCCCTCCTCGGGAGT
 GCGTCCGATCATTGGCGGGGGCGCTACGGCCGCCGAGACCGTTTCATCTGGGCCCTGAGCCTGGGGATCCTGC
 TCTCTCTCTCCTGATCCCCCGGGCCGGCTGGCTGGCCGGCCTGCTGTGTCCCGACCCCGCCCTCTGGAGCT
 GGCCCTCCTGATCCTGGGCGTGGGCTTGTGGACTTCTGCGGCCAGGTGTGTTTCACTCCCTGGAGGCTCTG
 CTCTCCGACCTCTTCCGCGACCCCGACCACTGTAGGCAGGCTTACAGCGTGTACGCCTTCATGATCAGTCTGG

GGGGATGCCTGGGCTATCTGCTGCCCCGCTATCGACTGGGACACCAGCGCCCTGGCCCCCTACCTGGGGACTCA
GGAGGAGTGCCTGTTTCGGCCTGCTCACCTTGATCTTCCTGACGTGCGTCGCCGCCACCCCTGCTGGTGGCCGAG
GAGGCGGCCCTGGGGCCCACCGAGCCCCGCCGAGGGCCTGAGCGCTCCAGCCTGAGCCCCCATTGCTGCCCCGT
GCAGGGCTAGGCTCGCCTTCAGGAATCTGGGCGCTTTGCTGCCCCGCCTGCATCAGCTGTGCTGTGCGCATGCC
TCGCACCCCTGCGCCGCCCTGTTTCGTCGCTGAGCTCTGTTTCCTGGATGGCCCTGATGACGTTACACCTCTTCTAC
ACCGACTTCGTGGGGGAGGGCCTGTACCAGGGCGTGCCAGGGCCGAGCCCGGCACCGAGGCTAGGCGCCATT
ACGACGAGGGCGTCAGGATGGGCTCTCTGGGCCTCTTCCTGCAGTGCGCCATCAGTCTGGTGTTCCTCTCTGGT
GATGGACCGGCTGGTGCAGCGCTTCGGCACCCGGGCCGTGTACCTCGCCTCTGTGGCGGCTTTCCCCGTGCCC
GCCGGCGCGACCTGCCTGTCTCATTCTGTGCGCGTGGTGACCGCCAGCGCCGCCCTGACCGGCTTCACCTTCA
GTGCGCTCCAGATTCTGCCCTACACCCCTGGCGTCTCTGTACCATCGCGAGAAGCAGGTGTTCTTGCCCAAGTA
CCGCGGGGACACAGGGGGAGCTTCCTCTGAGGACAGCCTGATGACCAGCTTCTTGCCCGGCCCAAGCCGGGG
GCCCCCTTTCCCCAACGGCCATGTGCGGGCGGGCGGCAGCGGCCTGCTCCCTCCCCCCCCCGCCCTGTGCGGCG
CTAGTGCCCTGCGACGTGAGCGTGCGGGTGGTGGTGGGGGAGCCACCGAGGCTAGGGTCGTGCCTGGCCGGGG
GATCTGCCTGGACCTGGCCATCCTCGACTCCGCCTTCTCTGCTCTCCCAGGTGGCGCCAGCCTGTTTCATGGGC
AGTATCGTGCGAGCTGAGCCAGAGCGTGACCGCCTACATGGTGAGCGCCGCCCGGCTGGGGTTGGTGGCCATCT
ACTTTGCCACCCAGGTCGTGTTTCGACAAGAGCGATCTCGCCAAGTATAGCGCCTGA

Fig. 15 - Zurückmanipulierte Codon-optimierte Sequenz 19 (SEQ ID NO:21)

GACGGCTAGCGCCACCATGGTGCAGCGGCTCTGGGTGAGCCGCCTCCTGCGGCATCGCAAGGCCCCAGCTCCTG
 CTGGTGAATCTGCTCACATTGCGCCTGGAGGTGTGCCCTGGCCGCCGGCATCACCTACGTGCCCCCCCCCTCCTGC
 TGGAGGTGGGAGTCGAGGAGAAGTTCATGACCATGGTGTGGGCATTGGGCCCCGTCCTGGGCCTCGTGTGCGT
 GCCTCTCCTCGGCAGCGCTTCCGACCATTGGCGCGGCCGGTATGGCCGCAGGAGACCCCTTCATCTGGGCTCTG
 AGTCTCGGCATCCTGCTGAGCCTGTTCTGATCCCTCGGGCCGGCTGGCTGGCCGGGCTGCTGTGCCCCGATC
 CTCGGCCCCCTGGAGCTGGCCCTGCTGATCCTCGGCGTGGGCCTGCTGGACTTCTGCGGCCAGGTGTGCTTCAC
 GCCCCCTGGAGGCACCTGCTGAGCGACCTGTTCCGGGACCCCGACCATTGGCCGCCAGGCGTACAGCGTGTACGCC
 TTCATGATCTCCCTGGGAGGCTGCCTGGGCTACCTGCTCCCCGCCATCGATTGGGACACCAGCGCACTCGCCC
 CCTATCTCGGAACACAGGAGGAATGCCTGTTTCGGA[]TG[]TGACGCTCATCTTCCTCACGTGCGTCGCGGCCAC
 CCTGTTGGTGGCCGAGGAGGCCGCCCTGGGGCCCCACCGAGCCGGCCGAGGACTGAGCGCCCCGAGCCTGAGT
 CCACACTGCTGCCCCTTGCCGGGCCCCGCTGGCCTTCCGTAATCTGGGCGCCCTCCTGCCTCGGCTCCATCAGC
 TGTGTTGCAGAAATGCCTAGGACGCTGCGGCGCCTGTTTCGTGCTGAGTTGTGCTCCTGGATGGCTCTCATGAC
 CTTACCCCTGTTTTATACGGACTTCGTGCGGGAGGGCCTGTACCAGGGGGTGCCGCGCGCCGAGCCCCGGGACA
 GAGGCGCGCCGCACTACGACGAGGGAGTGCGTATGGGCTCCCTGGGCCTCTTCTTGCAGTGCGCCATCAGTC
 TGGTTTTCTCTCTGGTTCATGGACAGGCTGGTGCAGCGCTTCGGAACCCGGGCGGTGTACCTGGCGAGCGTGGC
 CGCCTTCCCCGTGGCTGCCGGCGCCACCTGCCTCTCTCACTCGGTGGCCGTGGTCAACGCCAGCGCCGCCCTG
 ACCGGGTTACCTTCTCTGCCCTGCAGATTCTGCCCTTACACCCTGGCCAGCCTGTACCATCGCGAGAAACAGG
 TGTTTTCTCCCAAGTACAGAGGCGACACCGGGGGCGCCTCCAGCGAGGACAGCCTCATGACCTCCTTCCTGCC
 TGGCCCCAAGCCCGGCGCCCCCTTTCCCCAACGGGCACGTGGGCGCCGGCGGGAGTGGGCTCCTGCCCCCCCCCT
 CCTGCGCTGTGCGGGGCCAGCGCCTGCGACGTGAGCGTGCGCGTGGTGGTGGGCGAGCCACCGAGGCCCGCG
 TGGTGGCCGGGAGAGGCATTGTCTGGACCTGGCCATCCTCGACTCCGCCTTCCTCCTCAGCCAGGTGGCCCC
 GTCCCTCTTCATGGGCTCTATCGTCCAGCTGTCTCAGAGCGTCACCGCTTACATGGTGTCCGCTGCTGGACTG
 GGCTTGGTGGCTATTTATTTGCCACCCAGGTGGTGTTCGACAAGAGCGACCTGGCCAAATACTCCGCCTGAC
TCGAGGCAG

Fig. 16 - Zurückmanipulierte Codon-optimierte Sequenz 20 (SEQ ID NO:22)

GACGGCTAGCGCCACCATGGTGCAGCGGCTGTGGGTGTCCCGGCTGCTGCGCCATAGAAAGGCCAGTTGCTG
 CTGGTGAACCTGCTGACTTTCGGACTGGAGGTGTGCCTGGCTGCCGGGATCACGTACGTGCCCCCCTGCTGC
 TGGAGGTGGGCGTGGAGGAGAAGTTCATGACAATGGTGTGCGGCATCGGCCCCGTCTGGGCCTCGTGTGTGT
 GCCCCCTCTCGGGAGTGCCTCCGATCATTGGCGGGGCCGCTACGGCCGCCGAGACCGTTTCATCTGGGCCCTG
 AGCCTGGGCGATCCTGCTCTCTCTCTCTCTGATCCCCCGGGCCGGCTGGCTGGCCGGCCTGCTGTGTCCCGACC
 CCCGCCCTCTGGAGCTGGCCCTCCTGATCCTGGGCGTGGGCGTGGACTTCTGCGGCCAGGTGTGTTTCAC
 TCCCCCTGGAGGCTCTGCTCTCCGACCTCTTCCGCGACCCCGACCACTGTAGGCAGGCTTACAGCGTGTACGCC
 TTCATGATCAGTCTGGGGGGATGCCTGGGCTATCTGCTGCCCGTATCGACTGGGACACCAGCGCCCTGGCCC
 CCTACCTGGGGACTCAGGAGGAGTGCCCTGTTCCGGCCTGCTCACCTTGATCTTCCCTGACGTGCGTCGCCGCCAC
 CCTGCTGGTGGCCGAGGAGGCGGCCCTGGGGCCACCGAGCCCGCCGAGGGCCTGAGCGCTCCCAGCCTGAGC
 CCCCATTGCTGCCCCGTGCAGGGCTAGGCTCGCCTTCAGGAATCTGGGCGCTTTGCTGCCCCGCCCTGCATCAGC
 TGTGCTGTGCGATGCCTCGCACCCCTGCGCCGCTGTTCTGTCGCTGAGCTCTGTTCTTGATGGCCCTGATGAC
 GTTCAACCTCTTCTACACCGACTTCGTGGGGGAGGGCCTGTACCAGGGCGTGCCAGGGCCGAGCCCGGCACC
 GAGGCTAGGCGCCATTACGACGAGGGCGTCAGGATGGGCTCTCTGGGCCTCTTCTGTCAGTGCGCCATCAGTC
 TGGTGTCTCTCTGCTGATGGACCGGCTGGTGCAGCGCTTCGGCACCCGGGCGGTGTACCTCGCCTCTGTGGC
 GGCTTTCCCCGTGCGCGCCGGCGCGACCTGCCTGTCTCATTCTGTCGCCGTGGTGACCGCCAGCGCCGCCCTG
 ACCGGCTTACCTTCAGTGCCTCCAGATTCTGCCCTACACCCTGGCGTCTCTGTACCATCGCGAGAAGCAGG
 TGTTCTGCCCCAAGTACCGCGGGACACAGGGGAGCTTCTCTGAGGACAGCCTGATGACCAGCTTCTTGCC
 CGGCCCCAAGCCGGGGGCCCCCTTTCCCCAACGGCCATGTGCGGGCGGGCGGCAGCGGCCTGCTCCCTCCCCC
 CCGCCCTGTGCGGCGCTAGTGCTGCGACGTGAGCGTGCGGGTGGTGGTGGGGAGCCCCACCGAGGCTAGGG
 TCGTGCTGGCCGGGGGATCTGCCTGGACCTGGCCATCCTCGACTCCGCCTTCTGCTCTCCAGGTGGCGCC
 CAGCCTGTTTCATGGGCAGTATCGTGCAGCTGAGCCAGAGCGTGACCGCCTACATGGTGAGCGCCCGGCCCTG
 GGGTTGGTGGCCATCTACTTTGCCACCCAGGTCGTGTTTCGACAAGAGCGATCTCGCCAAGTATAGCGCCTGAC
 TCGAGGCAG

Fig. 17 - Die Ausgangssequenz zur Optimierung von CPC (SEQ ID NO:23)

Vier Aminosäuren von P501S-Sequenz sind umrahmt

ATGGCGCCGCTTACGTACATTCCGACGGCTCTTATCCAAAAGACAAGTTTGAGAAAATCAATGGCACTTGGT
 ACTACTTTGACAGTTCAGGCTATATGCTTGACACCGCTGGAGGAAGCACACAGACGGCAACTGGTACTGGTT
 CGACAACCTCAGGCGAAATGGCTACAGGCTGGAAGAAAATCGCTGATAAGTGGTACTATTTCAACGAAGAAGGT
 GCCATGAAGACAGGCTGGGTCAAGTACAAGGACACTTGGTACTACTTAGACGCTAAAGAAGGCGCCATGCAAT
 ACATCAAGGCTAACTCTAAGTTCATTGGTATCACTGAAGGCGTCATGGTATCAAATGCCTTTATCCAGTCAGC
 GGACGGAACAGGCTGGTACTACCTCAAACCAGACGGAACACTGGCAGACAGGCCAGAAAGTTCATGTAC

Fig. 18 - Repräsentative Codon-optimierte CPC-Sequenzen (SEQ ID NO:24-25)**SEQ ID NO:24**

ATGGCCGCGCCTACGTGCATAGCGACGGGAGCTACCCCAAGGACAAGTTCGAGAAGATCAACGGGACATGGT
 ACTACTTCGACTCCTCCGGCTACATGCTCGCCGACCGCTGGCGGAAGCACACCGACGGCAACTGGTACTGGTT
 CGATAACTCGGGAGAGATGGCCACCGGCTGGAAGAAGATCGCGGACAAGTGGTACTATTTCAACGAGGAGGGC
 GCCATGAAGACCGGCTGGGTGAAGTATAAGGACACCTGGTACTACCTCGACGCCAAGGAGGGCGCCATGCAGT
 ATATCAAGGCCAACAGCAAGTTCATCGGCATCACCGAGGGAGTGATGGTCAGCAACGCCTTTATCCAGAGCGC
 CGACGGCACCGGATGGTACTACTTGAAGCCGGACGGCACCCCTCGCGGATCGGCCCGAGAAGTTCATGTAC

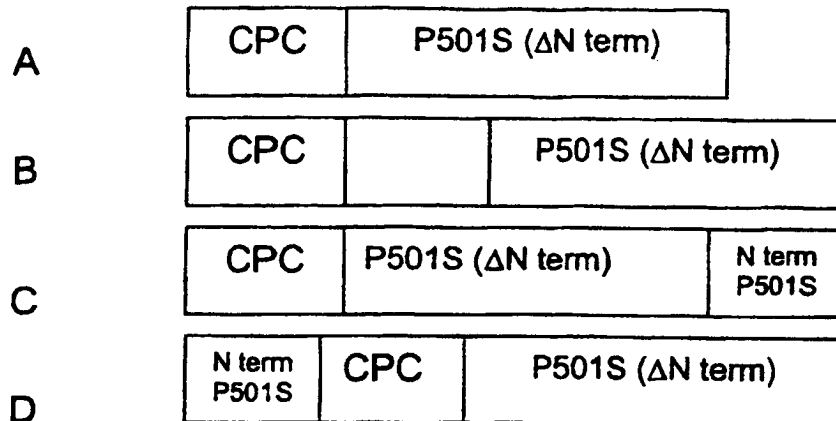
SEQ ID NO:25

ATGGCCGCGCCTACGTGCACAGCGACGGGTCCTACCCAAAGGACAAGTTCGAGAAGATCAACGGCACGTGGT
 ACTATTTTCGACAGCAGCGGTACATGCTCGCCGATCGCTGGCGCAAGCACACCGACGGGAAGTGGTACTGGTT
 CGACAACCTCGGCGAGATGGCTACGGGGTGAAGAAGATCGCCGACAAGTGGTACTACTTCAACGAGGAGGGC
 GCCATGAAGACCGGGTGGGTGAAGTACAAGGACACCTGGTACTACCTGGACGCTAAGGAGGGCGCCATGCAGT
 ACATCAAGGCCAACTCGAAGTTCATCGGGATCACCGAGGGCGTGATGGTCAGTAACGCTTTCATCCAGAGCGC
 GGACGGCACAGGCTGGTATTACCTGAAGCCCGATGGCACCCCTGGCGGACAGACCTGAGAAAATTCATGTAC

Fig. 19 - Manipulierte Codon-optimierte CPC-Sequenz (SEQ ID NO:26)**SEQ ID NO:26**

GACGGCTAGCGCCACCATGGCCGCGCCTACGTGCATAGCGACGGGAGCTACCCCAAGGACAAGTTCGAGAAG
 ATCAACGGGACATGGTACTACTTCGACTCCTCCGGCTACATGCTCGCCGACCGCTGGCGGAAGCACACCGACG
 GCAACTGGTACTGGTTCGATAACTCGGGAGAGATGGCCACCGGCTGGAAGAAGATCGCGGACAAGTGGTACTA
 TTTCAACGAGGAGGGCGCCATGAAGACCGGCTGGGTGAAGTATAAGGACACCTGGTACTACCTCGACGCCAAG
 GAGGGCGCCATGCAGTATATCAAGGCCAACAGCAAGTTCATCGGCATCACCGAGGGAGTGATGGTCAGCAACG
 CCTTTATCCAGAGCGCCGACGGCACCGGATGGTACTACTTGAAGCCGGACGGCACCCCTCGCGGATCGGCCGA
 GAAGTTCATGTACTGACTCGAGGCAG

Fig. 20 - P501S-CPC-Fusionskandidatkonstrukte und Sequenzen



Konstrukt A = SEQ ID NO:37 (Nukleotid) & 45 (Polypeptid)

GCGGCCGCGCCACCATGGCCGCCGCTACGTGCATAGCGACGGGAGCTACCCCAAGGACA
 M A A A Y V H S D G S Y P K D K
 AGTTCGAGAAGATCAACGGGACATGGTACTACTTCGACTCCTCCGGCTACATGCTCGCCG
 F E K I N G T W Y Y F D S S G Y M L A D
 ACCGCTGGCGGAAGCACACCGACGGCAACTGGTACTGGTTCGATAACTCGGGAGAGATGG
 R W R K H T D G N W Y W F D N S G E M A
 CCACCGGCTGGAAGAAGATCGCGGACAAGTGGTACTATTTCAACGAGGAGGGCGCCATGA
 T G W K K I A D K W Y Y F N E E G A M K
 AGACCGGCTGGGTGAAGTATAAGGACACCTGGTACTACCTCGACGCCAAGGAGGGCGCCA
 T G W V K Y K D T W Y Y L D A K E G A M
 TGCAGTATATCAAGGCCAACAGCAAGTTCATCGGCATCACCGAGGGAGTGATGGTCAGCA
 Q Y I K A N S K F I G I T E G V M V S N
 ACGCCTTTATCCAGAGCGCCGACGGCACCGGATGGTACTACTTGAAGCCGGACGGCACCC
 A F I Q S A D G T G W Y Y L K P D G T L
 TCGCGGATCGGCCCGAGAAGTTCATGTACATGGTGCTGGGCATCGGCCCCGTCTGGGCC
 A D R P E K F M Y M V L G I G P V L G L
 TCGTGTGTGTGCCCCCTCCTCGGGAGTGCGTCCGATCATTGGCGGGGCCGCTACGGCCGCC
 V C V P L L G S A S D H W R G R Y G R R
 GCAGACCGTTCATCTGGGCCCTGAGCCTGGGCATCCTGCTCTCTCTCTCTCTGATCCCCC
 R P F I W A L S L G I L L S L F L I P R
 GGGCCGGCTGGCTGGCCGGCCTGCTGTGTCCCGACCCCCGCCCTCTGGAGCTGGCCCTCC
 A G W L A G L L C P D P R P L E L A L L
 TGATCCTGGGCGTGGGCCTGCTGGACTTCTGCGGCCAGGTGTGTTTCACTCCCCTGGAGG

I L G V G L L D F C G Q V C F T P L E A
 CTCTGCTCTCCGACCTCTTCCGCGACCCCGACCACTGTAGGCAGGCTTACAGCGTGTACG
 L L S D L F R D P D H C R Q A Y S V Y A
 CCTTCATGATCAGTCTGGGGGGATGCCTGGGCTATCTGCTGCCCGCTATCGACTGGGACA
 F M I S L G G C L G Y L L P A I D W D T
 CCAGCGCCCTGGCCCCCTACCTGGGGACTCAGGAGGAGTGCCTGTTCCGGCCTGCTCACCT
 S A L A P Y L G T Q E E C L F G L L T L
 TGATCTTCCTGACGTGCGTCGCCGCCACCCTGCTGGTGGCCGAGGAGGCGGCCCTGGGGC
 I F L T C V A A T L L V A E E A A L G P
 CCACCGAGCCCGCCGAGGGCCTGAGCGCTCCCAGCCTGAGCCCCCATTGCTGCCCCGTGCA
 T E P A E G L S A P S L S P H C C P C R
 GGGCTAGGCTCGCCTTCAGGAATCTGGGCGCTTTGCTGCCCCGCTGCATCAGCTGTGCT
 A R L A F R N L G A L L P R L H Q L C C
 GTCGCATGCCTCGCACCCCTGCGCCGCTGTTGCTGCTGAGCTCTGTTCTTGATGGCCC
 R M P R T L R R L F V A E L C S W M A L
 TGATGACGTTACCCCTCTTCTACACCGACTTCGTGGGGGAGGGCCTGTACCAGGGCGTGC
 M T F T L F Y T D F V G E G L Y Q G V P
 CCAGGGCCGAGCCCGGCACCGAGGCTAGGCGCCATTACGACGAGGGCGTCAGGATGGGCT
 R A E P G T E A R R H Y D E G V R M G S
 CTCTGGGCCTCTTCTGTCAGTGCGCCATCAGTCTGGTGTCTCTCTGGTGTATGGACCGGC
 L G L F L Q C A I S L V F S L V M D R L
 TGGTGCAGCGCTTCGGCACCCGGGCGCTGTACCTCGCCTCTGTGGCGGCTTTCCCCGTCG
 V Q R F G T R A V Y L A S V A A P P V A
 CCGCCGGCGCGACCTGCCTGTCTATTCTGTGCGCGTGGTGACCGCCAGCGCCGCCCTGA
 A G A T C L S H S V A V V T A S A A L T
 CCGGCTTCACCTTCAGTGCCTCCAGATTCTGCCCTACACCCTGGCGTCTCTGTACCATC
 G F T F S A L Q I L P Y T L A S L Y H R
 GCGAGAAGCAGGTGTTCTGCCCCAAGTACCGCGGGGACACAGGGGGAGCTTCTCTGAGG
 E K Q V F L P K Y R G D T G G A S S E D
 ACAGCCTGATGACCAGCTTCTTGCCCGGCCCAAGCCGGGGGCCCCCTTTCCCCAACGGCC
 S L M T S F L P G P K P G A P F P N G H
 ATGTCGGGGCGGGCGGCAGCGGCCTGCTCCCTCCCCCCCCCGCCCTGTGCGGCGCTAGTG
 V G A G G S G L L P P P P A L C G A S A
 CCTGCGACGTGAGCGTGCGGGTGGTGGTGGGGGAGCCCACCGAGGCTAGGGTCGTGCCTG
 C D V S V R V V V G E P T E A R V V P G
 GCCGGGGGATCTGCCTGGACCTGGCCATCCTCGACTCCGCCTTCTGCTCTCCAGGTGG
 R G I C L D L A I L D S A F L L S Q V A
 CGCCAGCCTGTTTCATGGGCAGTATCGTGCAGCTGAGCCAGAGCGTGACCGCCTACATGG
 P S L F M G S I V Q L S Q S V T A Y M V

TGAGCGCCCGCCGCTGGGGTTGGTGGCCATCTACTTTGCCACCCAGGTCGTGTTGACA
S A A G L G L V A I Y F A T Q V V F D K

AGAGCGATCTCGCCAAGTATAGCGCCTGAGGATCC
S D L A K Y S A *

Konstrukt B = SEQ ID NO:38 (Nukleotid) & 46 (Polypeptid)

GCGGCCGCGCCACCATGGCCGCGCCTACGTGCATAGCGACGGGAGCTACCCCAAGGACA
M A A A Y V H S D G S Y P K D K

AGTTCGAGAAGATCAACGGGACATGGTACTACTTCGACTCCTCCGGCTACATGCTCGCCG
F E K I N G T W Y Y F D S S G Y M L A D

ACCGCTGGCGGAAGCACACCGACGGCAACTGGTACTGGTTCGATAACTCGGGAGAGATGG
R W R K H T D G N W Y W F D N S G E M A

CCACCGGCTGGAAGAAGATCGCGGACAAGTGGTACTATTTCAACGAGGAGGGCGCCATGA
T G W K K I A D K W Y Y F N E E G A M K

AGACCGGCTGGGTGAAGTATAAGGACACCTGGTACTACCTCGACGCCAAGGAGGGCGCCA
T G W V K Y K D T W Y Y L D A K E G A M

TGCAGTATATCAAGGCCAACAGCAAGTTCATCGGCATCACCGAGGGAGTGATGGTCAGCA
Q Y I K A N S K F I G I T E G V M V S N

ACGCCTTTATCCAGAGCGCCGACGGCACCGGATGGTACTACTTGAAGCCGGACGGCACCC
A F I Q S A D G T G W Y Y L K P D G T L

TCGCGGATCGGCCCGAGATGGTGCAGCGGCTGTGGGTGTCCCGGCTGCTGCGCCATAGAA
A D R P E M V Q R L W V S R L L R H R K

AGGCCAGTTGCTGCTGGTGAACCTGCTGACTTTTCGGACTGGAGGTGTGCCTGGCTGCCG
A Q L L L V N L L T F G L E V C L A A G

GGATCACGTACGTGCCCCCCTGCTGCTGGAGGTGGGCGTGGAGGAGAAGTTCATGACAA
I T Y V P P L L L E V G V E E K F M T M

TGGTGCTGGGCATCGGCCCGCTCCTGGGCCTCGTGTGTGTGCCCCCTCCTCGGGAGTGCGT
V L G I G P V L G L V C V P L L G S A S

CCGATCATTGGCGGGGCGCTACGGCCGCGCAGACCGTTCATCTGGGCCCTGAGCCTGG
D H W R G R Y G R R R P F I W A L S L G

GCATCCTGCTCTCTCTCTTCCTGATCCCCCGGGCGGCTGGCTGGCCGGCCTGCTGTGTC
I L L S L F L I P R A G W L A G L L C P

CCGACCCCGCCCTCTGGAGCTGGCCCTCCTGATCCTGGGCGTGGGCCTGCTGGACTTCT
D P R P L E L A L L I L G V G L L D F C

GCGGCCAGGTGTGTTTCACTCCCTGGAGGCTCTGCTCTCCGACCTCTTCCGCGACCCCG
G Q V C F T P L E A L L S D L F R D P D

ACCACTGTAGGCAGGCTTACAGCGTGTACGCCTTCATGATCAGTCTGGGGGGATGCCTGG
H C R Q A Y S V Y A F M I S L G G C L G

GCTATCTGCTGCCCCGCTATCGACTGGGACACCAGCGCCCTGGCCCCCTACCTGGGGACTC
Y L L P A I D W D T S A L A P Y L G T Q

AGGAGGAGTGCTGTTTCGGCCTGCTCACCTTGATCTTCCTGACGTGCGTCGCCGCCACCC
 E E C L F G L L T L I F L T C V A A T L
 TGCTGGTGGCCGAGGAGGCGGCCCTGGGGCCACCGAGCCCGCCGAGGGCCTGAGCGCTC
 L V A E E A A L G P T E P A E G L S A P
 CCAGCCTGAGCCCCCATTGCTGCCCCGTGCAGGGCTAGGCTCGCCTTCAGGAATCTGGGCG
 S L S P H C C P C R A R L A F R N L G A
 CTTTGCTGCCCCGCCTGCATCAGCTGTGCTGTGCGATGCCTCGCACCCCTGCGCCGCCTGT
 L L P R L H Q L C C R M P R T L R R L F
 TCGTCGCTGAGCTCTGTTCTGATGGCCCTGATGACGTTACCCCTCTTCTACACCGACT
 V A E L C S W M A L M T F T L F Y T D F
 TCGTGGGGGAGGGCCTGTACCAGGGCGTGCCAGGGCCGAGCCCGGCACCGAGGCTAGGC
 V G E G L Y Q G V P R A E P G T E A R R
 GCCATTACGACGAGGGCGTCAGGATGGGCTCTCTGGGCCTCTTCTGCAGTGCGCCATCA
 H Y D E G V R M G S L G L F L Q C A I S
 GTCTGGTGTCTCTCTGATGGACCGGCTGGTGCAGCGCTTCGGCACCCGGGCGGTGT
 L V F S L V M D R L V Q R F G T R A V Y
 ACCTCGCCTCTGTGGCGGCTTTCCCCGTCGCCGCCGGCGCGACCTGCCTGTCTCATTCTG
 L A S V A A F P V A A G A T C L S H S V
 TCGCCGTGGTGACCGCCAGCGCCGCCCTGACCGGCTTCACCTTCAGTGCGCTCCAGATTC
 A V V T A S A A L T G F T F S A L Q I L
 TGCCCTACACCCTGGCGTCTCTGTACCATCGCGAGAAGCAGGTGTTCTGCCCCAAGTACC
 P Y T L A S L Y H R E K Q V F L P K Y R
 GCGGGGACACAGGGGGAGCTTCCTCTGAGGACAGCCTGATGACCAGCTTCTTGCCCGGCC
 G D T G G A S S E D S L M T S F L P G P
 CCAAGCCGGGGGCCCTTTCCCAACGGCCATGTGCGGGCGGGCGGCAGCGGCCTGCTCC
 K P G A P F P N G H V G A G G S G L L P
 CTCCCCCCCCCGCCCTGTGCGGCGCTAGTGCCTGCGACGTGAGCGTGCGGGTGGTGGTGG
 P P P A L C G A S A C D V S V R V V V G
 GGGAGCCCACCGAGGCTAGGGTCGTGCCTGGCCGGGGGATCTGCCTGGACCTGGCCATCC
 E P T E A R V V P G R G I C L D L A I L
 TCGACTCCGCCTTCCTGCTCTCCAGGTGGCGCCAGCCTGTTTCATGGGCAGTATCGTGC
 D S A F L L S Q V A P S L F M G S I V Q
 AGCTGAGCCAGAGCGTGACCGCCTACATGGTGAGCGCCGCCGGCCTGGGGTTGGTGGCCA
 L S Q S V T A Y M V S A A G L G L V A I
 TCTACTTTGCCACCCAGGTCGTGTTGACAAGAGCGATCTCGCCAAGTATAGCGCCTGAG
 Y F A T Q V V F D K S D L A K Y S A *
 GATCC

Konstrukt C = SEQ ID NO:39 (Nukleotid) & 47 (Polypeptid)

GCGGCCGCGCCACCATGGCCGCGCCTACGTGCATAGCGACGGGAGCTACCCCAAGGACA
 M A A A Y V H S D G S Y P K D K
 AGTTCGAGAAGATCAACGGGACATGGTACTACTTCGACTCCTCCGGCTACATGCTCGCCG
 F E K I N G T W Y Y F D S S G Y M L A D
 ACCGCTGGCGGAAGCACACCGACGGCAACTGGTACTGGTTCGATAACTCGGGAGAGATGG
 R W R K H T D G N W Y W F D N S G E M A
 CCACCGGCTGGAAGAAGATCGCGGACAAGTGGTACTATTTCAACGAGGAGGGCGCCATGA
 T G W K K I A D K W Y Y F N E E G A M K
 AGACCGGCTGGGTGAAGTATAAGGACACCTGGTACTACCTCGACGCCAAGGAGGGCGCCA
 T G W V K Y K D T W Y Y L D A K E G A M
 TGCAGTATATCAAGGCCAACAGCAAGTTCATCGGCATCACCGAGGGAGTGATGGTCAGCA
 Q Y I K A N S K F I G I T E G V M V S N
 ACGCCTTTATCCAGAGCGCCGACGGCACC GGATGGTACTACTTGAAGCCGGACGGCACCC
 A F I Q S A D G T G W Y Y L K P D G T L
 TCGCGGATCGGCCCCGAGAAGTTCATGTACATGGTGCTGGGCATCGGCCCCGCTCCTGGGCC
 A D R P E K F M Y M V L G I G P V L G L
 TCGTGTGTGTGCCCCCTCCTCGGGAGTGCCTCCGATCATTTGGCGGGGCCGCTACGGCCGCC
 V C V P L L G S A S D H W R G R Y G R R
 GCAGACCGTTTCATCTGGGCCCTGAGCCTGGGCATCCTGCTCTCTCTCTCCTGATCCCC
 R P F I W A L S L G I L L S L F L I P R
 GGGCCGGCTGGCTGGCCGGCCTGCTGTGTCCCGACCCCCGCCCTCTGGAGCTGGCCCTCC
 A G W L A G L L C P D P R P L E L A L L
 TGATCCTGGGCGTGGGCCTGCTGGACTTCTGCGGCCAGGTGTGTTTCACTCCCCTGGAGG
 I L G V G L L D F C G Q V C F T P L E A
 CTCTGCTCTCCGACCTCTTCCGCGACCCCCGACCACTGTAGGCAGGCTTACAGCGTGTACG
 L L S D L F R D P D H C R Q A Y S V Y A
 CCTTCATGATCAGTCTGGGGGGATGCCTGGGCTATCTGCTGCCCCGCTATCGACTGGGACA
 F M I S L G G C L G Y L L P A I D W D T
 CCAGCGCCCTGGCCCCCTACCTGGGGACTCAGGAGGAGTGCCTGTTCCGGCCTGCTCACCT
 S A L A P Y L G T Q E E C L F G L L T L
 TGATCTTCCTGACGTGCGTCGCCGCCACCCTGCTGGTGGCCGAGGAGGCGGCCCTGGGGC
 I F L T C V A A T L L V A E E A A L G P
 CCACCGAGCCCCGCGAGGGCCTGAGCGCTCCAGCCTGAGCCCCCATTGCTGCCCCGTGCA
 T E P A E G L S A P S L S P H C C P C R
 GGGCTAGGCTCGCCTTCAGGAATCTGGGCGCTTTGCTGCCCCGCTGCATCAGCTGTGCT
 A R L A F R N L G A L L P R L H Q L C C
 GTCGCATGCCTCGCACCCCTGCGCCGCTGTTCTGCTGAGCTCTGTTCTGATGGCCCC
 R M P R T L R R L F V A E L C S W M A L
 TGATGACGTTACCCCTCTTCTACACCGACTTCGTGGGGGAGGGCCTGTACCAGGGCGTGC
 M T F T L F Y T D F V G E G L Y Q G V P

CCAGGGCCGAGCCCGGCACCGAGGCTAGGCGCCATTACGACGAGGGCGTCAGGATGGGCT
 R A E P G T E A R R H Y D E G V R M G S
 CTCTGGGCCTCTTCCTGCAGTGCGCCATCAGTCTGGTGTCTCTCTGGTGATGGACCGGC
 L G L F L Q C A I S L V F S L V M D R L
 TGGTGACGCGCTTCGGCACCCGGGCGGTGTACCTCGCCTCTGTGGCGGCTTTCCCCGTG
 V Q R F G T R A V Y L A S V A A F P V A
 CCGCCGGCGCGACCTGCCTGTCTCATTCTGTGCGCGTGGTGACCGCCAGCGCCGCCCTGA
 A G A T C L S H S V A V V T A S A A L T
 CCGGCTTCACCTTCAGTGCCTCCAGATTCTGCCCTACACCCTGGCGTCTCTGTACCATC
 G F T F S A L Q I L P Y T L A S L Y H R
 GCGAGAAGCAGGTGTTCTGCCCAAGTACCGCGGGGACACAGGGGAGCTTCCTCTGAGG
 E K Q V F L P K Y R G D T G G A S S E D
 ACAGCCTGATGACCAGCTTCTTGCCCGGCCCAAGCCGGGGGCCCTTTCCCAACGGCC
 S L M T S F L P G P K P G A P F P N G H
 ATGTGCGGGCGGGCGGCGAGCGGCTGCTCCCTCCCCCCCCCGCCCTGTGCGGCGCTAGTG
 V G A G G S G L L P P P P A L C G A S A
 CCTGCGACGTGAGCGTGCGGGTGGTGGTGGGGAGCCACCGAGGCTAGGGTGTGCCTG
 C D V S V R V V V G E P T E A R V V P G
 GCCGGGGGATCTGCCTGGACCTGGCCATCCTCGACTCCGCCTTCTGTCTCTCCAGGTGG
 R G I C L D L A I L D S A F L L S Q V A
 CGCCAGCCTGTTTCATGGGCAGTATCGTGACGTGAGCCAGAGCGTGACCGCTACATGG
 P S L F M G S I V Q L S Q S V T A Y M V
 TGAGCGCCGCGGCCTGGGGTGGTGGCCATCTACTTTGCCACCCAGGTGCTGTTTCGACA
 S A A G L G L V A I Y F A T Q V V F D K
 AGAGCGATCTCGCCAAGTATAGCGCCATGGTGCAGCGGCTGTGGGTGTCCCGGCTGCTGC
 S D L A K Y S A M V Q R L W V S R L L R
 GCCATAGAAAGGCCAGTTGCTGCTGGTGAACCTGCTGACTTTCGGACTGGAGGTGTGCC
 H R K A Q L L L V N L L T F G L E V C L
 TGGCTGCCGGGATCACGTACGTGCCCCCCTGCTGCTGGAGGTGGGCGTGGAGGAGTGAG
 A A G I T Y V P P L L L E V G V E E *
 GATCC

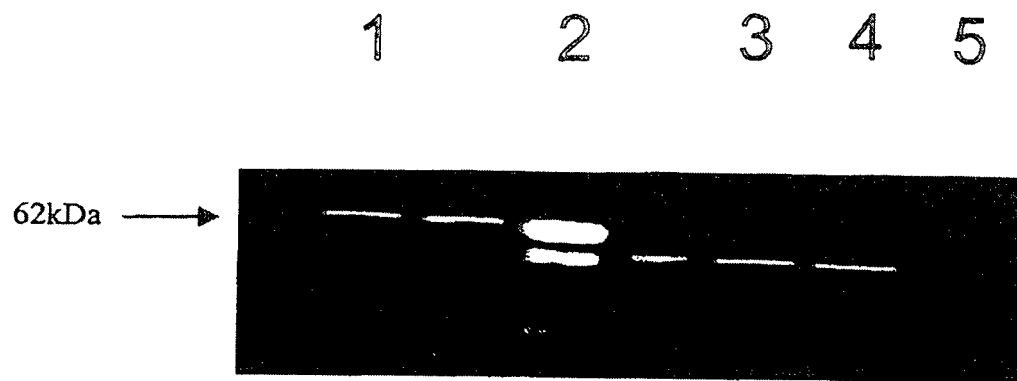
Konstrukt D = SEQ ID NO:40 (Nukleotid) & 48 (Polypeptid)

GCGGCCGCGCCACCATGGTGCAGCGGCTGTGGGTGTCCCGGCTGCTGCGCCATAGAAAGG
 M V Q R L W V S R L L R H R K A
 CCCAGTTGCTGCTGGTGAACCTGCTGACTTTCGGACTGGAGGTGTGCCTGGCTGCCGGGA
 Q L L L V N L L T F G L E V C L A A G I
 TCACGTACGTGCCCCCCTGCTGCTGGAGGTGGGCGTGGAGGAGATGGCCGCCGCTACG
 T Y V P P L L L E V G V E E M A A A Y V

TGCATAGCGACGGGAGCTACCCCAAGGACAAGTTCGAGAAGATCAACGGGACATGGTACT
 H S D G S Y P K D K F E K I N G T W Y Y
 ACTTCGACTCCTCCGGCTACATGCTCGCCGACCGCTGGCGGAAGCACACCGACGGCAACT
 F D S S G Y M L A D R W R K H T D G N W
 GGTACTGGTTCGATAACTCGGGAGAGATGGCCACCGGCTGGAAGAAGATCGCGGACAAGT
 Y W F D N S G E M A T G W K K I A D K W
 GGTACTATTTCAACGAGGAGGGCGCCATGAAGACCGGCTGGGTGAAGTATAAGGACACCT
 Y Y F N E E G A M K T G W V K Y K D T W
 GGTACTACCTCGACGCCAAGGAGGGCGCCATGCAGTATATCAAGGCCAACAGCAAGTTCA
 Y Y L D A K E G A M Q Y I K A N S K F I
 TCGGCATCACCAGGGGAGTGATGGTCAGCAACGCCTTTATCCAGAGCGCCGACGGCACCG
 G I T E G V M V S N A F I Q S A D G T G
 GATGGTACTACTTGAAGCCGGACGGCACCCCTCGCGGATCGGCCCCGAGAAGTTCATGTACA
 W Y Y L K P D G T L A D R P E K F M Y M
 TGGTGCTGGGCATCGGCCCCGTCCTGGGCCTCGTGTGTGTGCCCCCTCCTCGGGAGTGCGT
 V L G I G P V L G L V C V P L L G S A S
 CCGATCATTTGGCGGGGCCGCTACGGCCGCCGAGACCGTTTCATCTGGGCCCTGAGCCTGG
 D H W R G R Y G R R R P F I W A L S L G
 GCATCCTGCTCTCTCTCTCTCTGATCCCCCGGGCCGGCTGGCTGGCCGGCCTGCTGTGTC
 I L L S L P L I P R A G W L A G L L C P
 CCGACCCCCGCCCTCTGGAGCTGGCCCTCCTGATCCTGGGCGTGGGCCTGCTGGACTTCT
 D P R P L E L A L L I L G V G L L D F C
 GCGGCCAGGTGTGTTTCACTCCCCTGGAGGCTCTGCTCTCCGACCTCTTCCGCGACCCCG
 G Q V C F T P L E A L L S D L F R D P D
 ACCACTGTAGGCAGGCTTACAGCGTGTACGCCTTCATGATCAGTCTGGGGGGATGCCTGG
 H C R Q A Y S V Y A F M I S L G G C L G
 GCTATCTGCTGCCCCGCTATCGACTGGGACACCAGCGCCCTGGCCCCCTACCTGGGGACTC
 Y L L P A I D W D T S A L A P Y L G T Q
 AGGAGGAGTGCTGTTTCGGCCTGCTCACCTTGATCTTCCTGACGTGCGTCCGCCACCC
 E E C L F G L L T L I F L T C V A A T L
 TGCTGGTGGCCGAGGAGGGCGGCCCTGGGGCCACCGAGCCCGCCGAGGGCCTGAGCGCTC
 L V A E E A A L G P T E P A E G L S A P
 CCAGCCTGAGCCCCCATTTGCTGCCCCGTGCAGGGCTAGGCTCGCCTTCAGGAATCTGGGCG
 S L S P H C C P C R A R L A F R N L G A
 CTTTGCTGCCCCGCCTGCATCAGCTGTGCTGTGCGCATGCCTCGCACCCCTGCGCCGCTGT
 L L P R L H Q L C C R M P R T L R R L F
 TCGTCGCTGAGCTCTGTTCTGGATGGCCCTGATGACGTTACCCCTCTTCTACACCGACT
 V A E L C S W M A L M T F T L F Y T D F
 TCGTGGGGGAGGGCCTGTACCAGGGCGTGCCAGGGCCGAGCCCGGCACCGAGGCTAGGC

V G E G L Y Q G V P R A E P G T E A R R
GCCATTACGACGAGGGCGTCAGGATGGGCTCTCTGGGCCTCTTCCTGCAGTGCGCCATCA
H Y D E G V R M G S L G L F L Q C A I S
GTCTGGTGTTCTCTCTGGTGATGGACCGGCTGGTGCAGCGCTTCGGCACC CGGGCCGTGT
L V F S L V M D R L V Q R F G T R A V Y
ACCTCGCCTCTGTGGCGGCTTTCCCCGTCGCCGCCGGCGCGACCTGCCTGTCTCATTCTG
L A S V A A F P V A A G A T C L S H S V
TCGCCGTGGTGACCGCCAGCGCCGCCCTGACCGGCTTCACCTTCAGTGCCTCCAGATTC
A V V T A S A A L T G F T F S A L Q I L
TGCCCTACACCCTGGCGTCTCTGTACCATCGCGAGAAGCAGGTGTTCTGCCCCAAGTACC
P Y T L A S L Y H R E K Q V F L P K Y R
GCGGGGACACAGGGGAGCTTCCTCTGAGGACAGCCTGATGACCAGCTTCTTGCCCCGGCC
G D T G G A S S E D S L M T S F L P G P
CCAAGCCGGGGGCCCCCTTTCCCCAACGGCCATGTCGGGGCGGGCGGCAGCGGCCTGCTCC
K P G A P F P N G H V G A G G S G L L P
CTCCCCCCCCCGCCCTGTGCGGCGCTAGTGCCTGCGACGTGAGCGTGCGGGTGGTGGTGG
P P P A L C G A S A C D V S V R V V V G
GGGAGCCCACCGAGGCTAGGGTCGTGCCTGGCCGGGGGATCTGCCTGGACCTGGCCATCC
E P T E A R V V P G R G I C L D L A I L
TCGACTCCGCCTTCCTGCTCTCCCAGGTGGCGCCCAGCCTGTTTCATGGGCAGTATCGTGC
D S A F L L S Q V A P S L F M G S I V Q
AGCTGAGCCAGAGCGTGACCGCCTACATGGTGAGCGCCGCCCGGCCTGGGGTTGGTGGCCA
L S Q S V T A Y M V S A A G L G L V A I
TCTACTTTGCCACCCAGGTGCTGTTTCGACAAGAGCGATCTCGCCAAGTATAGCGCCTGAG
Y F A T Q V V F D K S D L A K Y S A *
GATCC

Fig. 21 - Western-Blot-Analyse von CHO-Zellen nach transients Transfektion mit P501S (JNW680), CPC-P501S (JNW735) und leerer Vektorkontrolle.



Bahn	Probe
1	CPC-P501S (JNW735)
2	CPC P501S-Protein (62,5 ng)
3	P501S (JNW680)
4	P501S (JNW680)
5	Leere Vektorkontrolle

Fig. 22 - Anti-P501S-Antikörperreaktionen nach Immunisierung am Tag 0, 21 & 42 mit pVAC-P501S (JNW680, Mäuse B1-9) oder leerem Vektor (pVAC, Mäuse A1-6)

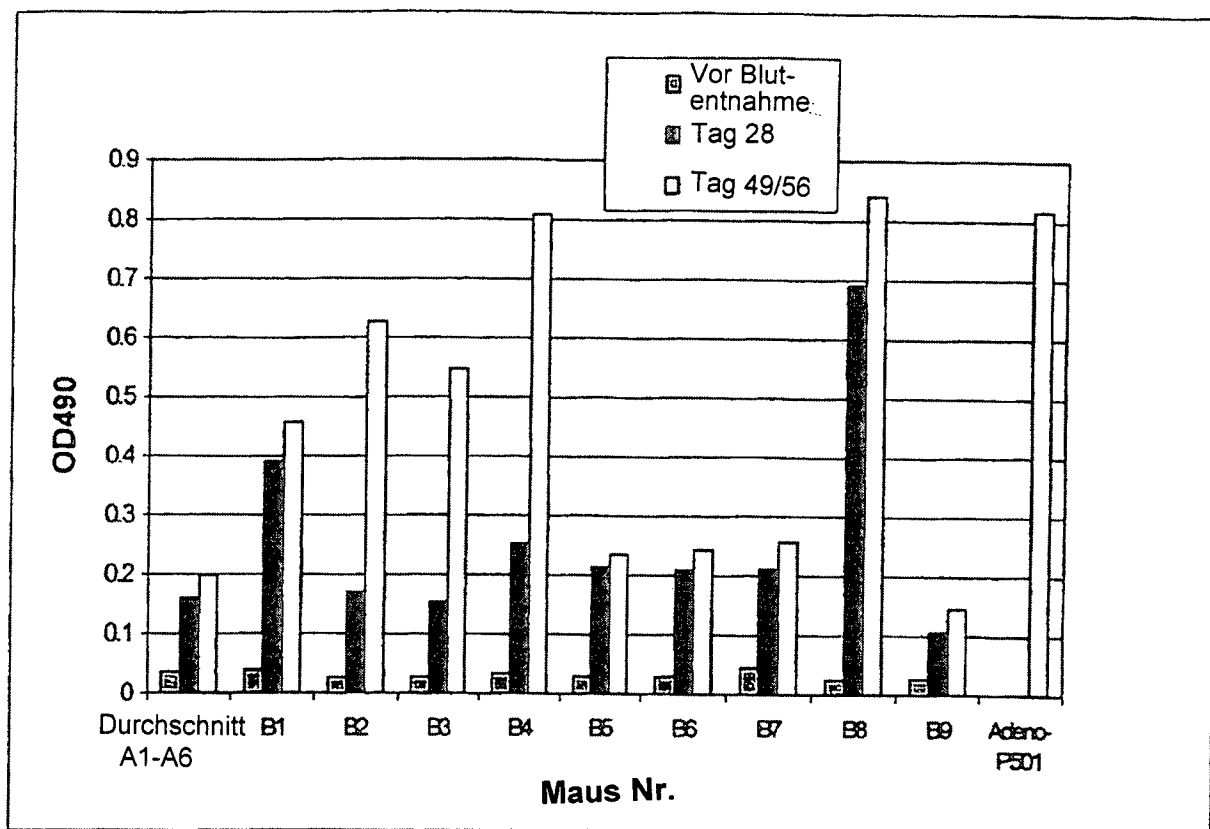


Fig. 23 - Peptidbibliotheksdurchmusterung unter Verwendung von C57BL/6-Mäusen, immunisiert am Tag 0, 21, 42 und 70 mit pVAC-P501S (JNW680)

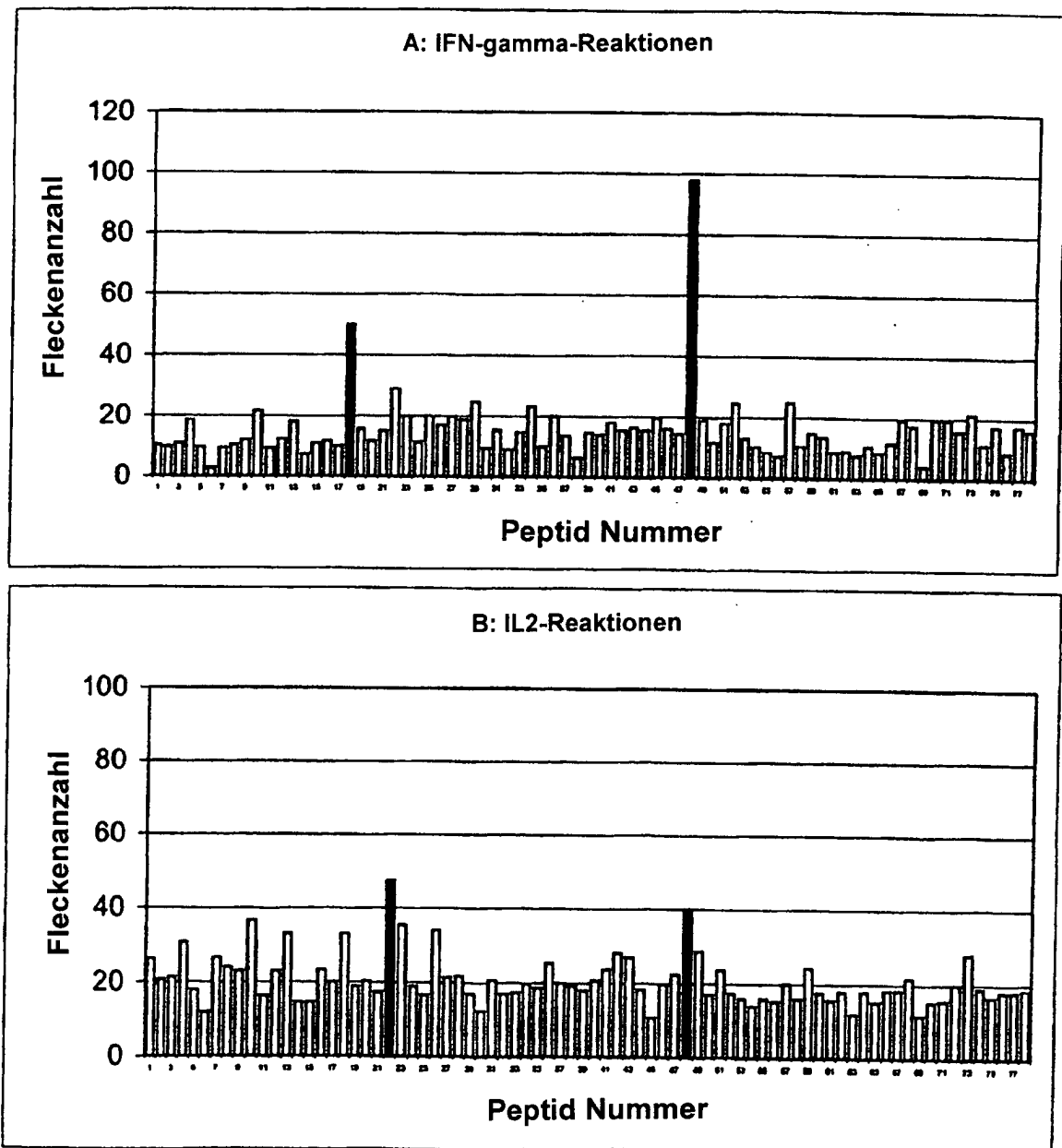


Fig. 24 - Zelluläre Reaktionen gemäß ELISPOT am Tag 77 nach PMID-Immunisierung am Tag 0, 21, 42 und 70 mit pVAC-P501S (JNW680, B6-9) und pVAC leer (A4-6).

Diagramm A zeigt die IFN- γ -Reaktionen, während Diagramm B die IL-2-Reaktionen zeigt

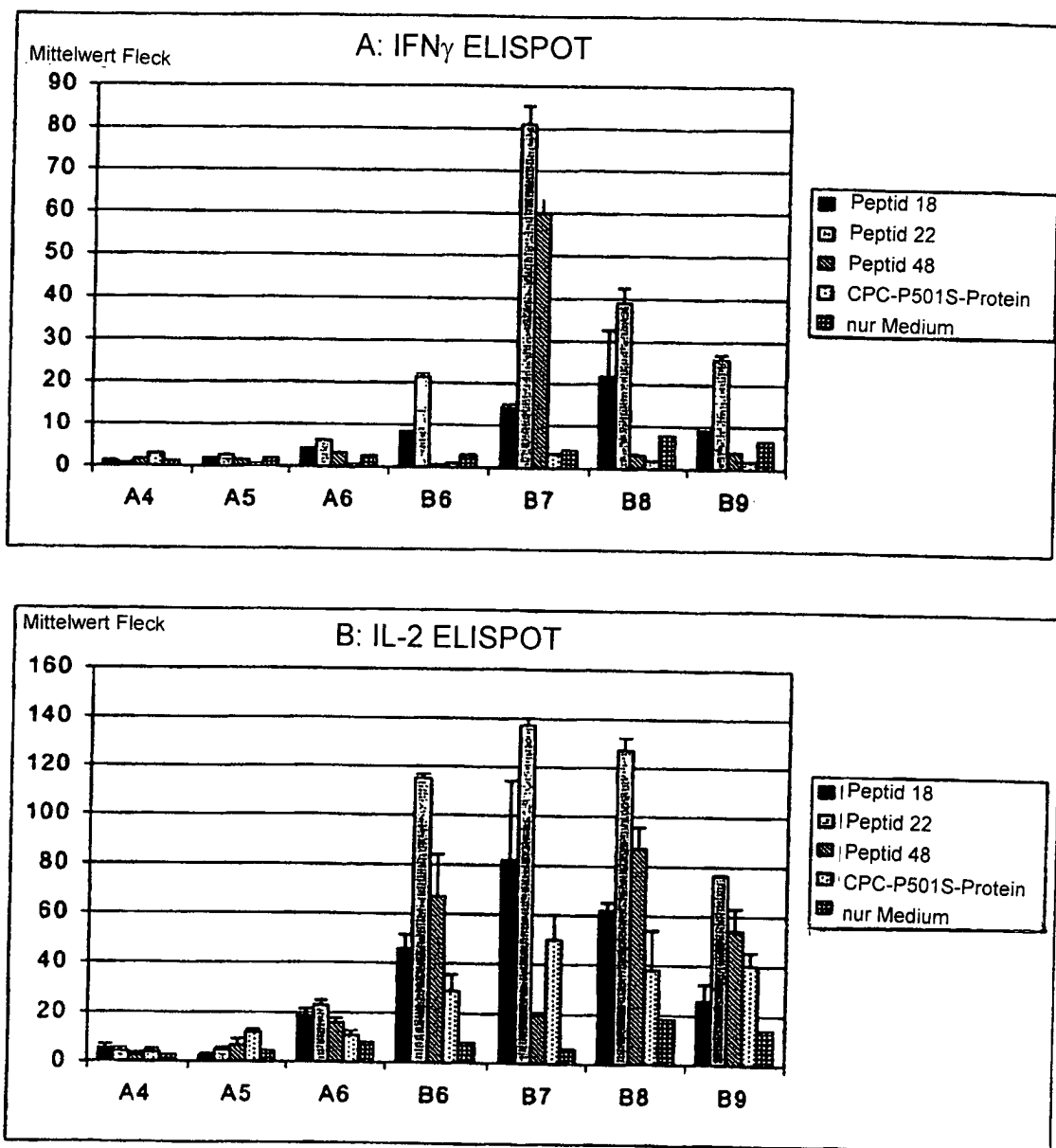


Fig. 25 - Vergleich von P501S und CPC-P501S

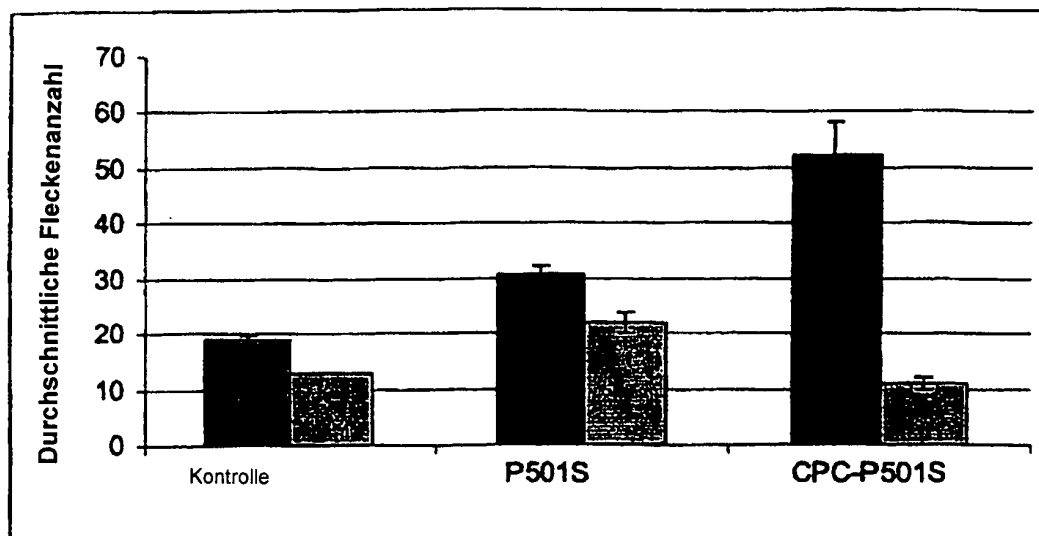


Fig. 26 - Immunreaktion (Lymphproliferation an Milzzellen) nach Protein-immunisierung mit CPC-P501S

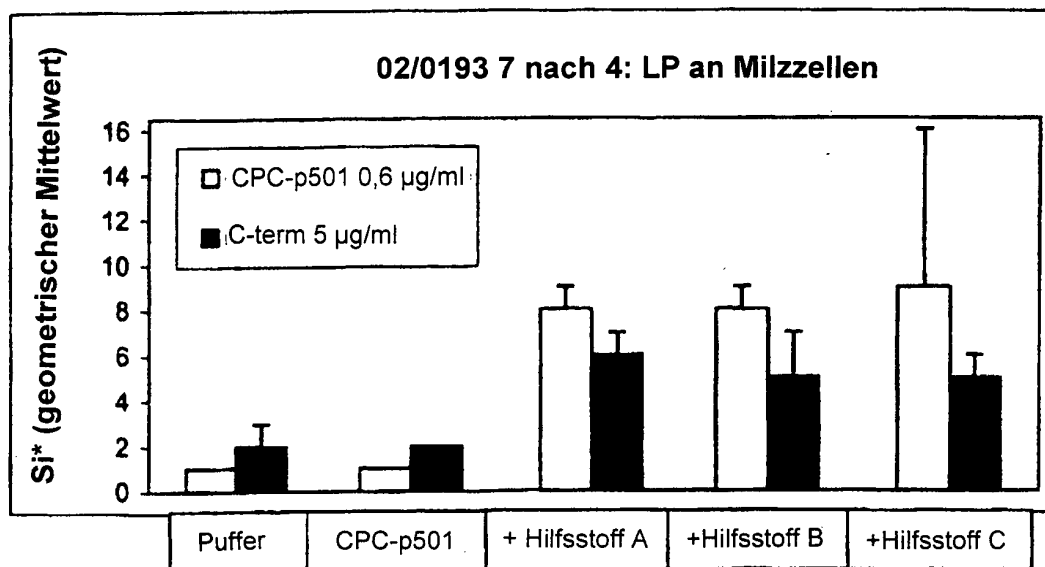


Fig. 27 - Auswertung der Immunreaktion auf unterschiedliche CPC-P501S-Konstrukte

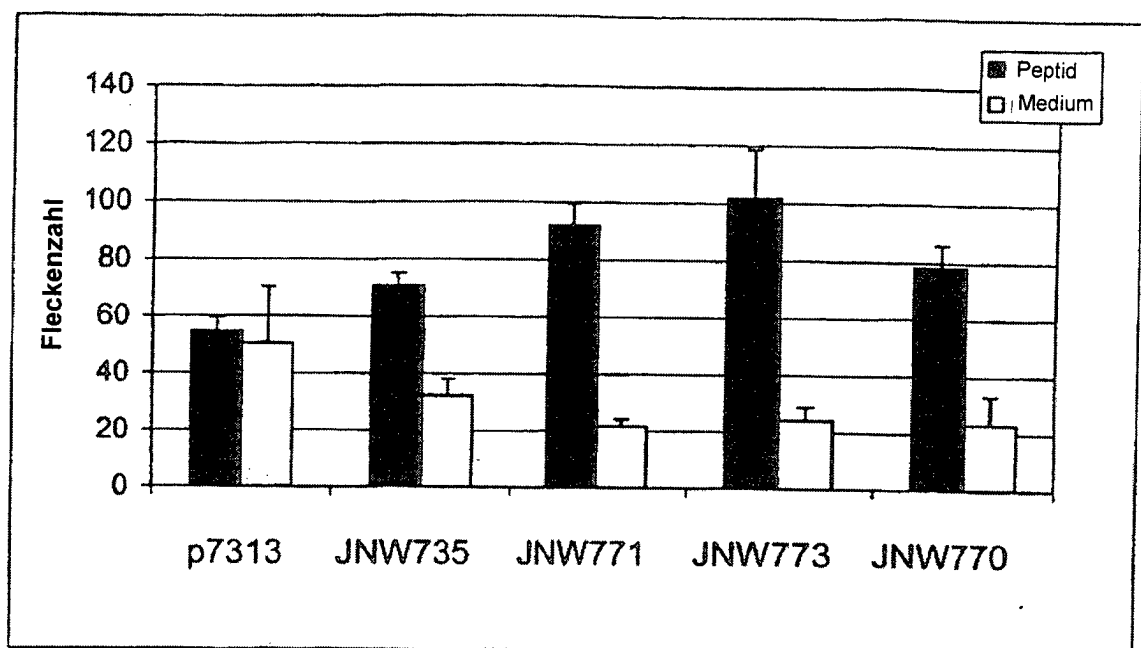


Fig. 28. MUC1-CPC-DNA und Polypeptidsequenzen

Fig. 28A. DNA-Sequenz (SEQ ID NO:49)

ATGACACCGGGCACCCAGTCTCCTTTCTTCTGCTGCTGCTCCTCACAGTGCTTACAGTTGTTACAGGTTCTG
 GTCATGCAAGCTCTACCCAGGTGGAGAAAAGGAGACTTCGGCTACCCAGAGAAGTTCAGTGCCAGCTCTAC
 TGAGAAGAATGCTGTGAGTATGACCAGCAGCGTACTCTCCAGCCACAGCCCCGGTTCAGGCTCCTCCACCCT
 CAGGGACAGGATGTCACTCTGGCCCCGGCCACGGAACAGCTTCAGGTTCACTGCCCACCTGGGGACAGGATG
 TCACCTCGGTCCCAGTCACCAGGCCAGCCCTGGGCTCCACCACCCCGCCAGCCCACGATGTCACTCAGCCCC
 GGACAACAAGCCAGCCCCGGGCTCCACCGCCCCCCCCAGCCCACGGTGTCACTCGGCCCCGGACACCAGGCCGCCGCCGGGCTCCA
 CCGCGCCCCGAGCCCCACGGTGTCACTCGGCCCCGGACACCAGGCCGGCCCCGGGCTCCACCGCCCCCCCCAGC
 CCATGGTGTCACTCGGCCCCGGACAACAGGCCCGCCTTGGCGTCCACCGCCCCCTCAGTCCACAATGTCACT
 TCGGCCTCAGGCTCTGCATCAGGCTCAGCTTCTACTCTGGTGACAAACGGCACCTCTGCCAGGGCTACCAAA
 CCCCAGCCAGCAAGAGCACTCCATTCTCAATTCCAGCCACCCTCTGATACTCCTACCACCCTTGCCAGCCA
 TAGCACCAGACTGATGCCAGTAGCACTCACCATAGCACGGTACCTCCTCTCACCTCCTCCAATCACAGCACT
 TCTCCCCAGTTGTCTACTGGGGTCTCTTTCTTTTCTGCTCTTTTACATTTCAAACCTCCAGTTTAATTCCT
 CTCTGGAAGATCCCAGCACCGACTACTACCAAGAGCTGCAGAGAGACATTTCTGAAATGTTTTGAGATTTA
 TAAACAAGGGGTTTTCTGGGCTCTCCAATATTAAGTTCAAGGCCAGGATCTGTGGTGGTACAATTGACTCTG
 GCCTTCCGAGAAGGTACCATCAATGTCCAGACGTGGAGACACAGTTCAATCAGTATAAAACGGAAGCAGCCT
 CTCGATATAACCTGACGATCTCAGACGTGAGCGTGAAGTATGTGCCATTTCTTTCTCTGCCAGTCTGGGGC
 TGGGGTGCCAGGCTGGGGCATCGCGCTGCTGGTGTGGTCTGTGTTCTGGTTGCGCTGGCCATTGTCTATCTC
 ATTGCCTTGGCTGTCTGTCACTGCCGCCGAAAGAACTACGGGCAGCTGGACATCTTTCCAGCCCCGGGATACCT
 ACCATCCTATGAGCGAGTACCCACCTACCACACCCATGGGCGCTATGTGCCCCCTAGCAGTACCGATCGTAG
 CCCCTATGAGAAGGTTTTCTGCAGGTAATGGTGGCAGCAGCCTCTCTTACACAAACCCAGCAGTGGCAGCCACT
 TCTGCCAAGTTGATGGCGGCCGCTTACGTACATTCGACGGCTCTTATCCAAAAGACAAGTTTGAGAAAATCA
 ATGGCACTTGGTACTACTTTGACAGTTCAAGGCTATATGCTTGACAGCCGCTGGAGGAAGCACACAGACGGCAA
 CTGGTACTGGTTCGACAACCTCAGGCGAAATGGCTACAGGCTGGAAGAAAATCGCTGATAAGTGGTACTATTTT
 AACGAAGAAGGTGCCATGAAGACAGGCTGGGTCAAGTACAAGGACACTTGGTACTACTTAGACGCTAAAGAAG
 GCGCCATGCAATACATCAAGGCTAACTCTAAGTTCAATTGGTATCACTGAAGGCGTCATGGTATCAAATGCCTT
 TATCCAGTCAGCGGACGGAACAGGCTGGTACTACCTCAAACCAGACGGAACACTGGCAGACAGGCCAGAATGA

Fig. 28B. MUC1-CPC-Polypeptidsequenz (SEQ ID NO:50)

MTPGTQSPFFLLLLLLTVLTVVTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMSTSSVLSSHSPGSGSSTT
 QGQDVTLPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVTSAPDNKPAPGSTAPPAHGVTSAPDTRP
 PPGSTAPPAHGVTSAPDTRPPPGSTAPAAHGVTSAPDTRPAPGSTAPPAHGVTSAPDNRPALASTAPPVHNVT
 SASGSASGSASTLVHNGTSARATTTTPASKSTPFSPSHHSDTPTTLASHSTKTDASSTHHSTVPPLTSSNHST
 SPQLSTGVSFFLSFHISNLQFNSSLEDPSDYYQELQORDISEMFLQIYKQGGFLGLSNIKFRPGSVVVQLTL

AFREGTINVHDTVETQFNQYKTEAASRYNLTI SDVSVDVFPFSAQSGAGVPGWGIALLLVLCVLVALAIVYL
IALAVCQCRRKNYGQLDIFPARDTYHPMSEYPTYHTHGRIYVPPSSSTRSPYEKVSAGNGGSSLSYTNPAVAAT
SANLMAAAVHSDGSYPKDKFEKINGTWYYFDSSGYMLADRWRKHTDGNWYWF DNSGEMATGWKKIADKWYYF
NEEGAMKTGWVKYKDTWYYLDAKEGAMQYIKANSKFIGITEGVMVSNAFIQSADGTGWYYLKP DGT LADRPE

Fig. 29. ss-CPC-MUC1 Konstrukt und Sequenz

Fig. 29A. DNA-Sequenz (SEQ ID NO:51)

ATGGGATGGAGCTGTATCATCTCTTCTTGGTAGCAACAGCTACAGGTGTCCACTCCCAGGTCCAAATGGCGG
CCGCTTACGTACATTCCGACGGCTCTTATCCAAAAGACAAGTTTGAGAAAATCAATGGCACTTGGTACTACTT
TGACAGTTT CAGGCTATATGCTTG CAGACCGCTGGAGGAAGCACACAGACGGCAACTGGTACTGGTTCGACAAC
TCAGGCGAAATGGCTACAGGCTGGAAGAAAATCGCTGATAAGTGGTACTATTTCAACGAAGAAGGTGCCATGA
AGACAGGCTGGGTCAAGTACAAGGACACTTGGTACTACTTAGACGCTAAAGAAGGCGCCATGCAATACATCAA
GGCTAACTCTAAGTTTATTGGTATCACTGAAGGCGTCATGGTATCAAATGCCTTTATCCAGTCAGCGGACGGA
ACAGGCTGGTACTACCTCAAACCAGACGGAACACTGGCAGACAGGCCAGAAATGACACCGGGCACCCAGTCTC
CTTTCTTCTCTGCTGCTGCTCCTCACAGTGCTTACAGTTGTTACAGGTTCTGGTTCATGCAAGCTCTACCCCAGG
TGGAGAAAAGGAGACTTCGGCTACCCAGAGAAGTTT CAGTGCCAGCTCTACTGAGAAGAATGCTGTGAGTATG
ACCAGCAGCGTACTCTCCAGCCACAGCCCCGGTT CAGGCTCCTCCACCACTCAGGGACAGGATGTCACTCTGG
CCCCGGCCACGGAACCAGCTT CAGGTT CAGCTGCCACCTGGGGACAGGATGTCACCTCGGTCCCAGTCACCAG
GCCAGCCCTGGGCTCCACCACCCCGCCAGCCACGATGTACCTCAGCCCCGACACAAGCCAGCCCCGGGC
TCCACCGCCCCCCCCAGCCACGGTGTACCTCGGCCCCGGACACCAGGCCGCCCCCGGGCTCCACCGCCCCCG
CAGCCACGGTGTACCTCGGCCCCGGACACCAGGCCGCCCCCGGGCTCCACCGCGCCCGCAGCCACGGTGT
CACCTCGGCCCCGGACACCAGGCCGCCCCCGGGCTCCACCGCCCCCCCCAGCCATGGTGTACCTCGGCCCCG
GACAACAGGCCCGCCTTGGCGTCCACCGCCCCCTCAGTCCACAATGTACCTCGGCCCTCAGGCTCTGCATCAG
GCTCAGCTTCTACTCTGGTGCAACGGCACCTCTGCCAGGGCTACCACAACCCAGCCAGCAAGAGCACTCC
ATTCTCAATTCCAGCCACCACTCTGATACTCCTACCACCTTGCCAGCCATAGCACCAAGACTGATGCCAGT
AGCACTCACCATAGCACGGTACCTCCTCTCACCTCCTCCAATCACAGCACTTCTCCCCAGTTGTCTACTGGGG
TCTCTTTCTTTTCTCTGTCTTTTCAATTTCAAACCTCCAGTTTAATTCTCTCTGGAAGATCCCAGCACCGA
CTACTACCAAGAGCTGCAGAGAGACATTTCTGAAATGTTTTGTCAGATTTATAAACAAGGGGGTTTTCTGGGC
CTCTCCAATATTAAGTT CAGGCCAGGATCTGTGGTGGTACAATTGACTCTGGCCTTCCGAGAAGGTACCATCA
ATGTCCACGACGTGGAGACACAGTTCAATCAGTATAAAACGGAAGCAGCCTCTCGATATAACCTGACGATCTC
AGACGT CAGCGTGAGTGATGTGCCATTTCTTTCTCTGCCCAGTCTGGGGCTGGGGTGCCAGGCTGGGGCATC
GCGCTGCTGGTGTGGTCTGTGTTCTGGTTGCGCTGGCCATTGTCTATCTCATTGCCTTGGCTGTCTGTCACT
GCCGCCGAAAGAACTACGGGCAGCTGGACATCTTTCCAGCCCCGGGATACCTACCATCCTATGAGCGAGTACCC
CACCTACCACACCCATGGGCGCTATGTGCCCCCTAGCAGTACCGATCGTAGCCCTATGAGAAGGTTTTCTGCA
GGTAATGGTGGCAGCAGCCTCTCTTACACAAACCAGCAGTGGCAGCCACTTCTGCCAAGTTGTAG



Fig. 29B. ss-CPC-MUC1-Proteinsequenz Polypeptidsequenz (SEQ ID NO:52)

MGWSCIIILFLVATATGVHSQVQMAAAYVHSDGSYPKDKFEKINGTWYYFDSSGYMLADRWRKHTDGNWYWFDN
SGEMATGWKKIADKWYYFNEEGAMKTGWVKYKDTWYYLDAKEGAMQYIKANSKFIGITEGVMVSNAFIQSADG
TGWYYLKP DGT LADRPEMTPGTQSPFFLLLLLLTVLTVVTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSM
TSSVLSSHSPGSGSSTTQGQDVT LAPATEPASGSAATWGQDVTSPVTRPALGSTTTPAH DVT SAPDNKPAPG
STAPPAHGVTSAPDTRPPPGSTAPPAHGVTSAPDTRPPPGSTAPAAHGVTSAPDTRPAPGSTAPPAHGVTSAP
DNRPALASTAPPVHNVT SASGSASGSASTLVHNGTSARATTT PASKSTPFSIPSHHSDTPTTLASHSTKTDAS
STHHSTVPPLTSSNHSTSPQLSTGVSFFFLSFHISNLQFNSSLED PSTDYYQELQORDISEMFLQIYKQGGFLG
LSNIKFRPGSVVVQLTLAFREGTINVHDVETQFNQYKTEAASRYNLTISDVSVSDVPPFFSAQSGAGVPGWGI
ALLVLCVLVALAIVYLIALAVCQCRKKNYGQLDIFPARDTYHPMSEYPTYHTHGRYVPPSSSTRSPYEKVSA
GNGGSSLSYTNPAVAATSANL