



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 693 33 807 T2 2006.02.02**

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 0 625 200 B1**

(21) Deutsches Aktenzeichen: **693 33 807.5**

(86) PCT-Aktenzeichen: **PCT/US93/01055**

(96) Europäisches Aktenzeichen: **93 904 938.3**

(87) PCT-Veröffentlichungs-Nr.: **WO 93/016185**

(86) PCT-Anmeldetag: **05.02.1993**

(87) Veröffentlichungstag

der PCT-Anmeldung: **19.08.1993**

(97) Erstveröffentlichung durch das EPA: **23.11.1994**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **11.05.2005**

(47) Veröffentlichungstag im Patentblatt: **02.02.2006**

(51) Int Cl.<sup>8</sup>: **C12N 15/13 (2006.01)**

**C12N 15/62 (2006.01)**

**C12P 21/08 (2006.01)**

**C07K 14/47 (2006.01)**

**C07K 16/18 (2006.01)**

**C12N 1/21 (2006.01)**

**G01N 33/68 (2006.01)**

**G01N 33/574 (2006.01)**

(30) Unionspriorität:

**831967 06.02.1992 US**

(73) Patentinhaber:

**Chiron Corp., Emeryville, Calif., US**

(74) Vertreter:

**Vossius & Partner, 81675 München**

(84) Benannte Vertragsstaaten:

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

(72) Erfinder:

**HUSTON, S., James, Chestnut Hill, US; HOUSTON, L., L., Oakland, US; RING, B., David, Redwood City, US; OPPERMAN, Hermann, Medway, US**

(54) Bezeichnung: **MARKER FÜR KREBS UND BIOSYNTHETISCHES BINDEPROTEIN DAFÜR**

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]** Diese Erfindung betrifft im Allgemeinen neue biosynthetische Zusammensetzungen von Substanzen und speziell biosynthetische Antikörperbindungsstellen (BABS)-Proteine und deren Konjugate. Erfindungsgemäße Zusammensetzungen sind zum Beispiel bei der Zielfindung bei Arzneistoffen und Toxinen, der Bildgebung, in der immunologischen Behandlung verschiedener Krebse und in spezifischen Bindungstests, Affinitätsreinigungsverfahren und in der Biokatalyse nützlich.

### Hintergrund der Erfindung

**[0002]** Brustkrebs ist die häufigste bösartige Erkrankung bei Frauen in Nordamerika mit 130000 neuen Fällen in 1987. Annähernd eine von 11 Frauen entwickelt in ihrer Lebenszeit Brustkrebs, was diese bösartige Erkrankung zur zweitwichtigsten Ursache von Krebstod, nach Lungenkrebs, bei Frauen in den Vereinigten Staaten macht. Obwohl die Mehrheit der Frauen mit Brustkrebs sich mit einer vollständig resezierbaren Erkrankung vorstellt, bleibt die Metastasierung ein enormes Hindernis für die Heilung. Die Verwendung von adjuvanter Chemotherapie oder hormoneller Therapie hat eindeutig einen positiven Effekt auf das erkrankungsfreie Überleben und das Gesamtüberleben in ausgewählten Untergruppen von Frauen mit vollständig reseziertem primärem Brustkrebs, aber ein wesentlicher Anteil der Frauen erleidet noch einen Rückfall durch eine Metastasierung (siehe zum Beispiel Fisher et al., J. Clin. Oncol. 4 (1986), 929 – 941; "The Scottish trial", Lancet 2 (1987), 171 – 175). Trotz der regelmäßig in entsprechend ausgewählten Patienten durch Chemotherapie und hormoneller Therapie induzierten objektiven Antworten wurde eine Heilung von metastasierendem Brustkrebs nicht erreicht (siehe zum Beispiel Aisner et al., J. Clin. Oncol. 5 (1987), 1523 – 1533). Zu diesem Zweck wurden viele innovative Behandlungsprogramme, einschließlich der Verwendung neuer Wirkstoffe, von Kombinationen von Wirkstoffen, der Hochdosistherapie (Henderson, ebenda) und einer erhöhten Dosisstärke (Kernan et al., Clin. Invest. 259 (1988), 3154 – 3157), zusammengestellt. Obwohl Verbesserungen beobachtet wurden, kam eine regelmäßige Erlangung von Vollremissionen von Metastasierungen, dem ersten Schritt in Richtung Heilung, nicht vor. Es bleibt ein dringender Bedarf nach neuen Ansätzen zur Behandlung.

**[0003]** Das Fv-Fragment eines Immunglobulinmoleküls aus IgM und in seltenen Fällen IgG oder IgA wird durch proteolytische Spaltung erzeugt und schließt ein nicht-covalentes V<sub>H</sub>-V<sub>L</sub>-Heterodimer ein, welches eine intakte Antigenbindungsstelle darstellt. Ein Einzelketten-Fv(sFv)-Polypeptid ist ein covalent verbundenes V<sub>H</sub>-V<sub>L</sub>-Heterodimer, welches von einer Genfusion, die V<sub>H</sub>- und V<sub>L</sub>-codierende Gene, verbunden durch einen ein Peptid codierenden Linker, einschließt, exprimiert wird. Siehe Huston et al., Proc. Nat. Aca. Sci. 85 (1988), 5879, hierbei durch Bezugnahme eingeschlossen.

**[0004]** Das US-Patent 4,753,894 offenbart murine monoclonale Antikörper, welche selektiv an menschliche Brustkrebszellen binden und, wenn sie an die Ricin-A-Kette konjugiert sind, eine TCID 50 % von weniger als etwa 10 nM gegenüber mindestens einer der MCF-7-, CAMA-1-, SKBR-3- oder BT-20-Zellen zeigen. Die SKBR-3-Zelllinie wird spezifisch durch den monoclonalen Antikörper 520C9 erkannt. Der als 520C9 bezeichnete Antikörper wird von einem murinen Hybridom sezerniert und ist nun bekannt dafür c-erbB-2 zu erkennen (Ring et al., Molecular Immunology 28 (1991), 915).

### Zusammenfassung der Erfindung

**[0005]** Die Erfindung zeigt die Synthese einer Klasse von neuen Proteinen, die als Einzelketten-Fv(sFv)-Polypeptide bekannt sind, welche biosynthetische Einzel-Polypeptidketten-bindungsstellen (BABS) einschließen und eine Bindungsstelle definieren, welche die immunologischen Bindungseigenschaften eines Immunglobulinmoleküls zeigt, welches c-erbB-2 oder ein mit c-erbB-2 verwandtes Tumor-Antigen bindet.

**[0006]** Das sFv schließt mindestens zwei Polypeptid-Domänen ein, welche durch einen Polypeptid-Linker verbunden sind, der den Abstand zwischen dem Carboxy-(C)-Terminus der einen Domäne und dem Amino-(N)-Terminus der anderen Domäne überbrückt, wobei die Aminosäuresequenz von jeder der Polypeptid-Domänen einen Satz von Komplementaritäts-bestimmenden Regionen (CDRs) umfasst, der in einen Satz von Framework-Regionen (FRs) eingeschoben ist, wobei die CDRs die immunologische Bindung an c-erbB-2 oder an ein mit c-erbB-2 verwandtes Tumor-Antigen bedingen.

**[0007]** In ihrer weitesten Ausdehnung zeigt diese Erfindung Einzelketten-Fv-Polypeptide einschließlich biosynthetischer Antikörperbindungsstellen, replizierbare Expressionsvektoren, die durch rekombinante DNA-Techniken hergestellt werden und welche DNA-Sequenzen, die diese Polypeptide codieren, enthalten und fähig sind diese zu exprimieren, Verfahren zur Herstellung dieser Polypeptide, Verfahren zur Darstellung

eines Tumors, der c-erbB-2 oder ein mit c-erbB-2 verwandtes Tumor-Antigen exprimiert, und Verfahren zur Behandlung eines Tumors unter Verwendung von zum Ziel lenkbaren therapeutischen Wirkstoffen mittels Konjugaten oder Fusionen mit diesen Polypeptiden.

**[0008]** Wie hier verwendet, bezieht sich der Begriff "immunologische Bindung" oder "immunologisch reaktiv" auf die nicht-covalenten Interaktionen jener Art, die zwischen einem Immunglobulinmolekül und einem Antigen, für welches das Immunglobulinmolekül spezifisch ist, auftreten; "c-erbB-2" bezieht sich auf ein Proteinantigen, welches auf der Oberfläche von Tumorzellen wie Brust- und Eierstock-Tumorzellen exprimiert wird, welches ein saures Glycoprotein mit einem Molekulargewicht von annähernd 200000 ist, einen isoelektrischen Punkt von etwa 5,3 hat und die in SEQ ID NO: 1 und 2 dargestellte Aminosäuresequenz einschließt. Ein „mit c-erbB-2 verwandtes Tumor-Antigen“ ist ein Protein, welches sich auf der Oberfläche von Tumorzellen wie Brust- und Eierstock-Tumorzellen befindet, welches bezüglich der Antigeneigenschaften mit dem c-erbB-2-Antigen verwandt ist, das heißt durch ein Immunglobulin gebunden wird, welches fähig ist, das c-erbB-2-Antigen zu binden, wobei Beispiele für solche Immunglobuline die 520C9-, 741F8- und 454C11-Antikörper sind; oder welches eine Aminosäuresequenz hat, die mindestens zu 80 %, bevorzugt zu 90 % mit der Aminosäuresequenz von c-erbB-2 homolog ist. Ein Beispiel für ein mit c-erbB-2 verwandtes Antigen ist der Rezeptor für den epidermalen Wachstumsfaktor.

**[0009]** Eine sFv-CDR, welche "im Wesentlichen homolog mit" einer Immunglobulin-CDR ist, behält mindestens 70 %, bevorzugt 80 % oder 90 % der Aminosäuresequenz der Immunglobulin-CDR bei und behält ebenfalls die immunologischen Bindungseigenschaften des Immunglobulins bei.

**[0010]** Der Ausdruck "Domäne" bezieht sich auf die Sequenz eines Polypeptids, die sich in ihrer nativen Konformation in eine einzelne globuläre Region faltet und abgegrenzte Bindungs- oder funktionelle Eigenschaften aufweisen kann. Der Ausdruck "CDR" oder Komplementaritäts-bestimmende Region bezieht sich, wie hier verwendet, auf Aminosäuresequenzen, welche zusammen die Bindungsaffinität und Bindungsspezifität der natürlichen Fv-Region einer nativen Immunglobulin-Bindungsstelle oder eines synthetischen Polypeptids, welches diese Funktion nachbildet, definieren. CDRs sind typischerweise nicht vollständig homolog zu den hypervariablen Regionen von natürlichen Fvs, sondern sie können eher auch spezifische Aminosäuren oder Aminosäuresequenzen enthalten, welche die hypervariablen Regionen flankieren und vordem als Framework, welches nicht direkt Komplementaritäts-bestimmend ist, angesehen wurden. Der Ausdruck "FR" oder Framework-Region bezieht sich, wie hier verwendet, auf Aminosäuresequenzen, welche natürlicherweise zwischen den CDRs in Immunglobulinen vorgefunden werden.

**[0011]** Einzelketten-Fv-Polypeptide, welche erfindungsgemäß hergestellt werden, schließen biosynthetisch hergestellte, neue Sequenzen von Aminosäuren ein, welche Polypeptide definieren, die dazu ausgelegt sind an ein vorher ausgewähltes c-erbB-2- oder ein verwandtes Antigen-Material zu binden. Die Struktur dieser synthetischen Polypeptide ist gegenüber der von natürlich vorkommenden Antikörpern, deren Fragmenten oder bekannten synthetischen Polypeptiden oder "chimären Antikörpern" dadurch verschieden, dass die Regionen des Einzelketten-Fv, die für die Spezifität und Affinität der Bindung (analog zu den variablen Regionen ( $V_H/V_L$ ) nativer Antikörper) verantwortlich sind, selbst chimär sein können, zum Beispiel Aminosäuresequenzen einschließen, die aus Abschnitten von mindestens zwei verschiedenen Antikörpermolekülen aus der gleichen oder aus verschiedenen Spezies abgeleitet oder zu ihnen homolog sind. Diese analogen  $V_H$ - und  $V_L$ -Regionen sind vom N-Terminus der einen bis zum C-Terminus der anderen durch ein über eine Peptidbindung gebundenes, biosynthetisches Linker-Peptid verbunden.

**[0012]** Die Erfindung stellt somit ein Einzelketten-Fv-Polypeptid bereit, welches mindestens eine vollständige Bindungsstelle definiert, die fähig ist c-erbB-2 oder ein mit c-erbB-2 verwandtes Tumor-Antigen zu binden. Eine vollständige Bindungsstelle schließt eine einzelne zusammenhängende Kette von Aminosäuren ein, welche zwei Polypeptid-Domänen, zum Beispiel  $V_H$  und  $V_L$ , verbunden durch eine Aminosäure-Linker-Region, hat. Ein sFv, welches mehr als eine vollständige Bindungsstelle hat, die zur Bindung von einem mit c-erbB-2 verwandten Antigen fähig ist, zum Beispiel zwei Bindungsstellen, wird eine einzelne, kontinuierliche Kette von Aminosäuren sein, die vier Polypeptid-Domänen hat, wobei jede kovalent durch eine Aminosäure-Linker-Region, zum Beispiel  $V_{H1}$ -Linker- $V_{L1}$ -Linker- $V_{H2}$ -Linker- $V_{L2}$ , verbunden ist. Erfindungsgemäße sFvs können jede Zahl von vollständigen Bindungsstellen ( $V_{Hn}$ -Linker- $V_{Ln}$ )<sub>n</sub> enthalten, wo  $n > 1$  ist, und können so eine einzelne, zusammenhängende Kette von Aminosäuren sein, welche  $n$  Antigenbindungsstellen und  $n \times 2$  Polypeptid-Domänen hat.

**[0013]** In einer bevorzugten Ausführungsform der Erfindung schließt das Einzelketten-Fv-Polypeptid CDRs ein, welche im Wesentlichen homolog mit mindestens einem Abschnitt der Aminosäuresequenz von CDRs aus

einer variablen Region eines Immunglobulinmoleküls aus einer ersten Spezies sind, und schließt FRs ein, welche im Wesentlichen homolog mit mindestens einem Teil der Aminosäuresequenz von FRs aus der variablen Region von einem Immunglobulinmolekül aus einer zweiten Spezies sind. Bevorzugt ist die erste Spezies die Maus und die zweite Spezies ist der Mensch.

**[0014]** Die Aminosäuresequenz von jeder der Polypeptid-Domänen schließt einen Satz von CDRs ein, welcher zwischen einen Satz von FRs eingeschoben ist. Ein "Satz von CDRs" bezieht sich, wie hier verwendet, auf 3 CDRs in jeder Domäne und ein "Satz von FRs" bezieht sich auf 4 FRs in jeder Domäne. Wegen struktureller Erwägungen kann ein ganzer Satz von CDRs aus einem Immunglobulin verwendet werden, es können aber Substitutionen von bestimmten Resten wünschenswert sein, um die biologische Aktivität zu verbessern, zum Beispiel gründend auf der Beobachtung von konservierten Resten in den CDRs von Immunglobulinarten, welche das mit c-erbB-2 verwandte Antigen binden.

**[0015]** In einem anderen bevorzugten Aspekt der Erfindung haben die CDRs der Polypeptidkette eine Aminosäuresequenz, welche im Wesentlichen homolog mit den CDRs der variablen Region des monoclonalen Antikörpers 741F8 sind.

**[0016]** In einer weiteren Ausführungsform schließt die Polypeptid-Linker-Region die Aminosäuresequenz ein, die im Sequenzprotokoll als Aminosäurerest Nummern 123 bis 137 in SEQ ID NO: 3 und 4 und als Aminosäurereste 1 – 15 in SEQ ID NO: 11 und 12 dargestellt ist. In anderen Ausführungsformen hat die Linker-Sequenz die Aminosäuresequenz, die im Sequenzprotokoll als Aminosäurereste 410 – 424 in SEQ ID NO: 9 und 10 oder als die Aminosäuresequenz der Reste 1 – 15 in SEQ ID NO: 13 und 14 dargestellt ist.

**[0017]** Die einzelne, vorstehend beschriebene Polypeptidkette kann auch eine an sie gebundene, aus der Entfernung nachweisbare Einheit einschließen, um die Abbildung oder Radioimmuntherapie von Tumoren zu erlauben, die ein c-erbB-2- oder ein damit verwandtes Tumor-Antigen tragen. "Aus der Entfernung nachweisbare" Einheit bedeutet, dass die Einheit, welche an das sFv gebunden ist, durch zur Position der Einheit äußerliche und von ihr entfernte Mittel nachgewiesen werden kann. Bevorzugt schließen aus der Entfernung nachweisbare Einheiten zur Abbildung ein radioaktives Atom wie <sup>99m</sup>Tc, einen Gammastrahler, ein. Bevorzugte Nuklide für die Hochdosis-Radioimmuntherapie schließen radioaktive Atome wie <sup>90</sup>Yttrium (<sup>90</sup>Yt), <sup>131</sup>Jod (<sup>131</sup>I) oder <sup>111</sup>Indium (<sup>111</sup>In) ein.

**[0018]** Zusätzlich kann das sFv ein Fusionsprotein enthalten, welches von einer Fusion von Genen abgeleitet ist, so dass das exprimierte sFv-Fusionsprotein ein Hilfs-Polypeptid einschließt, welches über eine Peptidbindung mit dem Bindungsstellen-Polypeptid verbunden ist. In einigen bevorzugten Aspekten hat der Abschnitt des Hilfs-Polypeptides auch eine Bindungsaffinität für ein c-erbB-2- oder ein verwandtes Antigen und kann eine dritte oder sogar eine vierte Polypeptid-Domäne einschließen, wobei jede eine Aminosäuresequenz umfasst, die zwischen FRs eingeschobene CDRs definiert, und welche zusammen eine zweite biosynthetische Einzel-Polypeptidketten-Bindungsstelle ähnlich der vorstehend beschriebenen ersten bilden.

**[0019]** In anderen Aspekten bildet die Hilfs-Polypeptidsequenz ein Toxin, welches mit dem N- oder C-Terminus des sFv verbunden ist, zum Beispiel mindestens ein toxischer Abschnitt des Pseudomonas-Exotoxins, Phytolaccins, Ricins, der Ricin-A-Kette oder des Diphtherie-Toxins oder anderer verwandter Proteine, die als Ricin-A-kettenartige, Ribosomen hemmende Proteine bekannt sind, das heißt Proteine, die fähig sind, die Proteinsynthese auf der Ebene der Ribosomen zu hemmen, wie das antivirale Protein der Kermesbeere, Gelonin und der ribosomale Proteininhibitor aus der Gerste. In noch einem weiteren Aspekt kann das sFv mindestens ein zweites Hilfs-Polypeptid oder eine Einheit, welche die Internalisierung des sFv fördert, einschließen.

**[0020]** Die Erfindung schließt auch ein Verfahren zur Herstellung von sFv ein, welches die Schritte der Bereitstellung eines replizierbaren Expressionsvektors, welcher eine DNA-Sequenz, die eine einzelne Polypeptidkette codiert, enthält und exprimiert; der Transfektion des Expressionsvektors in eine Wirtszelle, um eine Transformante herzustellen; und der Züchtung der Transformante zur Produktion des sFv-Polypeptids einschließt.

**[0021]** Die Erfindung schließt auch ein Verfahren zur Abbildung eines Tumors, welcher ein c-erbB-2- oder ein verwandtes Tumorantigen exprimiert, ein. Dieses Verfahren schließt die Schritte der Bereitstellung eines bildgebenden Stoffes, welcher ein wie vorstehend beschriebenes Einzelketten-Fv-Polypeptid einschließt, und eine daran gebundene, aus der Entfernung nachweisbare Einheit; der Verabreichung des bildgebenden Stoffes in einer Menge des bildgebenden Stoffes mit einem physiologisch verträglichen Träger, die ausreicht, um einen extrakorporalen Nachweis des Tumors zu erlauben, an einen den Tumor beherbergenden Organismus; und

des Nachweises der Position der Einheit im Individuum nach Zulassung der Bindung des Stoffes an den Tumor und der ausreichenden Ausscheidung des ungebundenen Stoffes, um die Sichtbarmachung des Tumorbildes zu erlauben, ein.

**[0022]** Die Erfindung schließt auch ein Verfahren zur Behandlung von Krebs durch Hemmung des in vivo-Wachstums eines Tumors, der ein c-erbB-2- oder ein verwandtes Antigen exprimiert, ein, wobei das Verfahren die Verabreichung einer den Tumor hemmenden Menge eines therapeutischen Wirkstoffes, welcher ein erfindungsgemäßes sFv und mindestens eine erste Einheit einschließt, die daran durch eine Peptidbindung gebunden ist und welche die Fähigkeit hat, die Vermehrung einer Tumorzelle zu begrenzen, an einen Krebs-Patienten einschließt.

**[0023]** Bevorzugt schließt die erste Einheit ein Toxin oder ein toxisches Fragment davon ein, zum Beispiel Ricin A; oder sie schließt ein Radioisotop, welches ausreichend radioaktiv ist, um die Vermehrung der Tumorzelle zu hemmen, ein, zum Beispiel  $^{90}\text{Yt}$ ,  $^{111}\text{In}$  oder  $^{131}\text{I}$ . Der therapeutische Wirkstoff kann weiterhin mindestens eine zweite Einheit einschließen, die seine Wirksamkeit verbessert.

**[0024]** Die klinische Verabreichung des Einzelketten-Fv oder angemessener, erfindungsgemäßer sFv-Fusionsproteine, welche die Aktivität von nativen, relativ kleinen Fv des entsprechenden Immunglobulins zeigen, bietet eine Vielzahl von Vorteilen gegenüber der Verwendung von größeren Fragmenten oder ganzen Antikörpermolekülen. Die erfindungsgemäßen Einzelketten-Fv- und sFv-Fusionsproteine bieten zirkulierenden proteolytischen Enzymen weniger Spaltstellen und bieten somit eine größere Stabilität. Sie erreichen ihr Zielgewebe schneller und werden schneller aus dem Körper ausgeschieden, was sie zu idealen bildgebenden Stoffen zum Nachweis von Tumoren und idealen radioimmuntherapeutischen Stoffen zur Abtötung von Tumoren macht. Sie haben auch im Vergleich zu murinen Immunglobulinen eine verringerte unspezifische Bindung und Immunogenität. Zusätzlich erleichtert ihre Expression durch einzelne Gene Anwendungen mit Zielfindung durch Fusion an andere Toxin-Proteine oder Peptidsequenzen, welche die spezifische Kopplung an andere Moleküle oder Arzneistoffe erlauben. Zudem haben einige erfindungsgemäße sFv-Analoga oder Fusionsproteine die Fähigkeit die Internalisierung von c-erbB-2 oder verwandten Antigenen, die auf der Oberfläche von Tumorzellen exprimiert werden, zu fördern, wenn sie zusammen auf der Zelloberfläche gebunden sind. Diese Verfahren erlauben die selektive Tötung von Zellen, die solche Antigene exprimieren, mit einer Einzelketten-Fv-Toxin-Fusion geeigneten Aufbaus. Erfindungsgemäße sFv-Toxin-Fusionsproteine besitzen eine 15 – 200-fach größere Aktivität bei der Tötung von Tumorzellen als Konjugate, welche ein Toxin einschließen, das chemisch mit einem ganzen Antikörper oder einem Fab vernetzt ist.

**[0025]** Übermäßige Expression von c-erbB-2 oder verwandten Rezeptoren auf bösartigen Zellen erlaubt somit die zielgerichtete Lenkung von sFv-Arten auf die Tumorzellen hin, gleich ob der Tumor gut lokalisiert oder metastasiert ist. In den vorstehenden Fällen erlaubt die Internalisierung von sFv-Toxin-Fusionsproteinen die spezifische Zerstörung von Tumorzellen, welche das übermäßig exprimierte c-erbB-2- oder ein verwandtes Antigen tragen. In anderen Fällen, abhängig von den infizierten Zellen, der Art der bösartigen Erkrankung oder anderen Faktoren, die in einem gegebenen Individuum wirken, können das gleiche c-erbB-2 oder verwandte Rezeptoren nur schlecht internalisiert werden oder sogar eine statische Tumorantigen-Population darstellen. In diesem Fall können das Einzelketten-Fv und seine Fusionsproteine auch nutzbringend, jedoch in einer anderen Weise als dies bei der Internalisierung anwendbar ist, verwendet werden. Wo c-erbB-2-Rezeptor/sFv- oder sFv-Fusionsprotein-Komplexe schlecht internalisiert werden, sind Toxine wie die Ricin-A-Kette, welche cytoplasmatisch durch die Inaktivierung von Ribosomen wirkt, bei der Tötung von Zellen nicht effektiv. Nichtsdestotrotz ist das unfusionierte Einzelketten-Fv nützlich, zum Beispiel zur Abbildung oder Radioimmuntherapie und bispezifische Einzelketten-Fv-Fusionsproteine verschiedenen Aufbaus, das heißt, solche, die zwei gesonderte Bindungsstellen auf der gleichen Polypeptidkette haben, können dazu verwendet werden, eine zielgerichtete Lenkung mittels der zwei Antigene, für die das Molekül spezifisch ist, durchzuführen. Zum Beispiel kann ein bispezifischer Einzelketten-Antikörper eine Spezifität für sowohl das c-erbB-2- als auch das CD3-Antigen haben, wobei das letztere von beiden auf cytotoxischen Lymphocyten (CTLs) vorliegt. Dieses bispezifische Molekül könnte so Antikörper-abhängige, zelluläre Cytotoxizität (ADCC) vermitteln, die eine durch CTL ausgelöste Lyse der Tumorzellen ergibt. Ähnliche Ergebnisse könnten durch Verwendung eines bispezifischen Einzelketten-Fv, der spezifisch für c-erbB-2 und den Fc $\gamma$ -Rezeptor Typ I oder II ist, erhalten werden. Andere bispezifische sFv-Formulierungen schließen Domänen mit c-erbB-2-Spezifität, gepaart mit einer Wachstumsfaktor-Domäne ein, welche für Hormone oder Wachstumsfaktorrezeptoren wie die Rezeptoren für Transferrin oder den epidermalen Wachstumsfaktor (EGF) spezifisch ist.

## Kurze Beschreibung der Zeichnungen

**[0026]** Die vorangehenden und andere Aufgaben dieser Erfindung, deren verschiedene Eigenschaften sowie die Erfindung selbst können durch die folgende Beschreibung vollständiger verstanden werden, wenn diese zusammen mit den begleitenden Zeichnungen gelesen wird.

**[0027]** [Fig. 1A](#) ist eine schematische Zeichnung eines DNA-Konstruktes, das ein erfindungsgemäßes sFv codiert, welche die  $V_H$ - und  $V_L$ -codierenden Domänen und die Linker-Region zeigt; [Fig. 1B](#) ist eine schematische Zeichnung der Struktur von Fv, welche  $V_H$ - und  $V_L$ -Domänen, von denen jede drei Komplementaritäts-bestimmende Regionen (CDRs) und vier Framework-Regionen (FRs) umfasst, für den monoclonalen 520C9 darstellt, einen gut bekannten und charakterisierten monoclonalen murinen Antikörper, der spezifisch für c-erbB-2 ist.

**[0028]** [Fig. 2A](#) – [Fig. 2E](#) sind schematische Darstellungen von erfindungsgemäßen Ausführungsformen, von denen jede ein biosynthetisches Einzelketten-Fv-Polypeptid umfasst, welches ein mit c-erbB-2 verwandtes Antigen erkennt: [Fig. 2A](#) ist ein sFv, welches eine anhängende Leader-Sequenz hat, [Fig. 2B](#) ist ein sFv-Toxin (oder ein anderes Hilfsprotein)-Konstrukt und [Fig. 2C](#) ist ein bivalentes oder bispezifisches sFv-Konstrukt; [Fig. 2D](#) ist ein bivalentes sFv, welches ein anhängendes Protein hat, das an das Carboxy-terminale Ende angehängt ist; [Fig. 2E](#) ist ein bivalentes sFv, welches anhängende Proteine, die sowohl am Amino- als auch am Carboxy-terminalen Ende angehängt sind, hat.

**[0029]** [Fig. 3](#) ist eine Darstellung der Konstruktion eines Plasmids, welches das 520C9-sFv-Ricin A-Immuntoxin-Fusionsgen codiert, als Diagramm; und

**[0030]** [Fig. 4](#) ist eine graphische Darstellung der Ergebnisse eines Wettbewerbsversuchs, der die c-erbB-2-Bindungsaktivität des monoclonalen 520C9-Antikörpers (spezifisch für c-erbB-2), eines Fab-Fragmentes dieses monoclonalen Antikörpers (gefüllte Punkte) und von verschiedenen, affinitätsgereinigten Fraktionen der Einzelketten-Fv-Bindungsstelle für c-erbB-2, die aus den variablen Regionen des monoclonalen 520C9-Antikörpers konstruiert ist (gesamte sFv- Probe (+), sFv gebunden und eluiert von einer Säule aus der immobilisierten, extrazellulären Domäne von c-erbB-2 (Quadrat) und Durchlauf von sFv (ungebunden, \*) mit-einander vergleicht.

## Detaillierte Beschreibung der Erfindung

**[0031]** Es werden Einzelketten-Fvs und sFv-Fusionsproteine offenbart, die eine Affinität für ein mit c-erbB-2 verwandtes Antigen haben, welches hochgradig auf Brust- und Eierstockkrebszellen und genauso auf anderen Tumorzellen in bestimmten anderen Krebsformen exprimiert wird. Die Polypeptide sind durch eine oder mehrere Sequenzen von Aminosäuren charakterisiert, die eine Region bilden, welche als eine biosynthetische Antikörperbindungsstelle fungiert. Wie in [Fig. 1](#) gezeigt, umfassen die Stellen Einzelketten aus der variablen Region der schweren Kette ( $V_H$ ) **10** und der variablen Region der leichten Ketten ( $V_L$ ) **14**, worin  $V_H$  **10** und  $V_L$  **14** durch den Polypeptid-Linker **12** verbunden sind. Die Bindungsdomänen schließen die CDRs **2**, **4**, **6** und **2'**, **4'**, **6'** von Immunglobulinmolekülen, die fähig sind ein mit c-erbB-2 verwandtes Tumorantigen zu binden, verbunden mit FRs **32**, **34**, **36**, **38** und **32'**, **34'**, **36'**, **38'**, welche aus einem gesonderten Immunglobulin abgeleitet sein können, ein. Wie in [Fig. 2A](#), [Fig. 2B](#) und [Fig. 2C](#) gezeigt, können die BABS-Einzelpolypeptidketten ( $V_H$  **10**,  $V_L$  **14** und Linker **12**) auch aus der Entfernung nachweisbare Einheiten und/oder andere Polypeptidsequenzen **16**, **18** oder **22** einschließen, welche zum Beispiel als ein Enzym, ein Toxin, eine Bindungsstelle oder eine Stelle zur Bindung an eine immobilisierende Matrix oder ein radioaktives Atom fungieren. Es werden ebenfalls Verfahren zur Herstellung der Proteine und Verfahren zu ihrer Verwendung offenbart.

**[0032]** Die erfindungsgemäßen Einzelketten-Fv-Polypeptide sind in dem Sinne biosynthetisch, als sie in einem zellulären Wirt synthetisiert und recloniert werden, der dazu gebracht wird, ein Protein zu exprimieren, das von einem Plasmid codiert wird, welches eine genetische Sequenz einschließt, die teilweise auf synthetischer DNA basiert, das ist, eine rekombinante DNA, die durch Ligierung von mehreren, chemisch synthetisierten und reclonierten Oligonucleotiden oder durch Ligierung von Fragmenten von DNA, welche aus dem Genom eines Hybridoms, reifen B-Zell-Clons oder einer cDNA-Genbank, abgeleitet aus solchen natürlichen Quellen, stammt, hergestellt wurde. Die erfindungsgemäßen Proteine sind insofern korrekt als "Antikörperbindungsstellen" charakterisiert, als diese synthetischen Einzelpolypeptidketten fähig sind, sich in eine 3-dimensionale Konformation zurückzufalten, die spezifisch dafür ausgelegt ist, eine Affinität für ein vorher ausgewähltes c-erbB-2-oder damit verwandtes Tumorantigen zu haben. Einzelketten-Fvs können, wie in der PCT-Anmeldung US88/01737 beschrieben, welche der USSN 342,449, eingereicht am 6.

**[0033]** Februar 1989 entspricht und Priorität aus der USSN 052,800, eingereicht am 21. Mai 1987, an Creative BioMolecules, Inc. übertragen, beansprucht, hierbei durch Bezugnahme eingeschlossen, produziert werden. Die erfindungsgemäßen Polypeptide sind insofern Antikörper-artig, als ihre Struktur gemäß Regionen nativer Antikörpern geprägt ist, die dafür bekannt sind, für die Erkennung von mit c-erbB-2 verwandtem Antigen verantwortlich zu sein.

**[0034]** Noch spezifischer gesagt, ist die Struktur von diesen biosynthetischen Antikörper-Bindungsstellen (BABS) in der Region, welche dem Protein die Bindungseigenschaften verleiht, analog zu der Fv-Region eines natürlichen Antikörpers gegen ein c-erbB-2- oder ein damit verwandtes Antigen. Sie schließt eine Folge von Regionen ein, die aus Aminosäuren besteht, welche mindestens drei Polypeptidabschnitte definieren, welche zusammen die molekulare Tertiärstruktur bilden, die verantwortlich für die Affinität und Bindung ist. Die CDRs werden durch Polypeptidabschnitte in geeigneter Konformation gehalten, die analog zu den Framework-Regionen des Fv-Fragments natürlicher Antikörper sind.

**[0035]** Die CDR- und FR-Polypeptidabschnitte werden empirisch, basierend auf der Sequenzanalyse der Fv-Region von bereits existierenden Antikörpern, wie die in dem US-Patent Nr. 4,753,894, hierbei durch Bezugnahme eingeschlossen, beschriebenen oder der DNA, die solche Antikörpermoleküle codiert, entworfen.

**[0036]** In einer Ausführungsform sind die Aminosäuresequenzen, welche die FRs der Einzelpolypeptidketten bilden, analog zu den FR-Sequenzen eines ersten, bereits existierenden Antikörpers, zum Beispiel eines menschlichen IgG. Die Aminosäuresequenzen, welche die CDRs bilden, sind analog zu den Sequenzen eines zweiten, unterschiedlichen, bereits existierenden Antikörpers, zum Beispiel den CDRs eines IgG aus einem Nagetier oder Menschen, welches c-erbB-2 oder damit verwandte Antigene, welche auf der Oberfläche von Eierstock- oder Brust-Tumorzellen exprimiert werden, erkennt. Alternativ können die CDRs und FRs in ihrer Gesamtheit aus einem einzelnen bereits existierenden Antikörper aus einer Zelllinie, welche instabil oder schwierig zu züchten sein kann, zum Beispiel einer ein sFv herstellenden Zelllinie, die auf einer murinen monoclonalen Antikörper, einen Antikörper aus Maus/Mensch oder Mensch produzierenden Zelllinie basiert, kopiert werden.

**[0037]** Die Anwendung der Erfindung ermöglicht den Entwurf und die Biosynthese von verschiedenen Stoffen, welche alle durch eine Region charakterisiert sind, die eine Affinität für ein vorher ausgewähltes c-erbB-2- oder ein verwandtes Antigen hat. Andere Regionen des biosynthetischen Proteins werden unter Berücksichtigung des speziellen, geplanten Nutzens des Proteins entworfen. Wenn der Stoff für die intravaskuläre Anwendung in Säugern entworfen wird, können somit die FRs Aminosäuresequenzen einschließen, die mindestens einem Teil der FR-Aminosäuren eines Antikörpers, der dieser Säugerart eigen ist, ähnlich oder dazu identisch sind. Andererseits können die Aminosäuresequenzen, welche die CDRs einschließen, analog zu einem Teil der Aminosäuresequenzen aus der hypervariablen Region (und bestimmten flankierenden Aminosäuren) eines Antikörpers sein, welcher eine bekannte Affinität und Spezifität für ein c-erbB-2- oder ein damit verwandtes Antigen hat und der zum Beispiel aus einer Maus oder einer Ratte oder einem bestimmten menschlichen Antikörper oder Immunglobulin stammt.

**[0038]** Andere Abschnitte nativer Immunglobulin-Proteinstruktur, zum Beispiel C<sub>H</sub> und C<sub>L</sub>, müssen nicht vorhanden sein und werden normaler Weise absichtlich bei den erfindungsgemäßen, biosynthetischen Proteinen weggelassen. Die erfindungsgemäßen Einzelpolypeptidketten können jedoch zusätzliche Polypeptid-Regionen einschließen, die eine Leader-Sequenz oder eine zweite Polypeptidkette, die bioaktiv ist, zum Beispiel ein Cytokin, ein Toxin, einen Liganden, ein Hormon, Immunglobulin-domäne(n) oder Enzyme oder eine Stelle, an die ein Toxin, ein Arzneistoff oder eine aus der Entfernung nachweisbare Einheit, zum Beispiel ein Radionuklid, gebunden werden kann, definieren.

**[0039]** Ein nützliches Toxin ist Ricin, ein Enzym aus dem Rizinusstrauch, das in hohem Maße toxisch ist, oder der Teil des Ricins, der Toxizität verleiht. Bei so niedrigen Konzentrationen wie 1 ng/ml hemmt Ricin effizient das Wachstum von Zellen in Kultur. Die Ricin-A-Kette hat ein Molekulargewicht von etwa 30000 und ist glycosyliert. Die Ricin-B-Kette hat eine größere Größe (ein Molekulargewicht von etwa 34000) und ist auch glycosyliert. Die B-Kette enthält zwei Galactose-Bindungsstellen, eine in jeder der zwei Domänen in der gefalteten Untereinheit. Die kristallographische Struktur von Ricin zeigt die Zeichnung des Rückgrats der A-Kette. Dort ist eine Spalte, welche wahrscheinlich das aktive Zentrum ist, das diagonal über dem Molekül verläuft. Es liegt auch eine Mischung von  $\alpha$ -Helix,  $\beta$ -Strukturen und irregulären Strukturen im Molekül vor.

**[0040]** Die A-Kette inaktiviert enzymatisch die ribosomale 60S-Untereinheit eukaryontischer Ribosomen. Die B-Kette bindet an Galactose-basierte Kohlenhydratreste auf den Oberflächen von Zellen. Sie scheint notwen-

dig zu sein, um das Toxin an die Zelloberfläche zu binden, und ermöglicht auch den Mechanismus des Eintritts des Toxins in die Zelle und ist an ihm beteiligt. Weil alle Zellen Galactose enthaltende Zelloberflächen-Rezeptoren haben, hemmt Ricin alle Arten von Säugerzellen mit einer annähernd gleichen Effizienz.

**[0041]** Die Ricin-A-Kette und die Ricin-B-Kette werden durch ein Gen codiert, das sowohl die A- als auch die B-Kette spezifiziert. Das von der von dem Gen transkribierten mRNA synthetisierte Polypeptid enthält Sequenzen der A-Kette, die durch ein "J" (für joining – verbindend)-Peptid mit den Sequenzen der B-Kette verbunden sind. Das J-Peptid-Fragment wird durch posttranslationale Modifikation entfernt, um die A- und B-Ketten freizusetzen. Die A- und B-Ketten werden jedoch immer noch durch die zwischen den Ketten liegende Disulfidbindung zusammengehalten. Die bevorzugte Form des Ricins ist die rekombinante A-Kette, da diese vollkommen frei von B-Kette ist und, wenn in *E. coli* exprimiert, nicht glycosyliert ist und somit langsamer aus dem Blut ausgeschieden wird als die glycosylierte Form. Die spezifische Aktivität der rekombinanten Ricin-A-Kette gegen Ribosomen und die der nativen A-Kette, die aus Rizinusstrauch-Ricin isoliert ist, sind gleichwertig. Eine Aminosäuresequenz und eine entsprechende Nucleinsäuresequenz der Ricin-A-Kette ist im Sequenzprotokoll als SEQ ID NO: 7 und 8 dargestellt.

**[0042]** Die rekombinante Ricin-A-Kette, die aus Pflanzen gewonnene Ricin-A-Kette, die deglycosylierte Ricin-A-Kette oder deren Abkömmlinge können mittels des erfindungsgemäßen Einzelketten-Fv-Polypeptids zielgerichtet auf eine Zelle gelenkt werden, die ein c-erbB-2- oder ein verwandtes Antigen exprimiert. Um dies zu tun, kann das sFv chemisch mit der Ricin-A-Kette oder einem aktiven Analogon davon vernetzt werden oder in einer bevorzugten Ausführungsform kann ein Einzelketten-Fv-Ricin-A-Ketten-Immuntoxin durch Fusionierung des Einzelketten-Fv-Polypeptids an eine oder mehrere Ricin-A-Ketten durch die entsprechende Fusion der Gene gebildet werden. Durch Ersatz der B-Kette des Ricins mit einer Antikörper-Bindungsstelle gegen c-erbB-2 oder verwandte Antigene, wird die A-Kette zu solchen Antigenen auf der Zelloberfläche geführt. In dieser Weise kann die selektive Tötung von Tumorzellen, die diese Antigene exprimieren, erreicht werden. Diese Selektivität wurde in vielen Fällen gegenüber in Kultur gezüchteten Zellen gezeigt. Sie hängt von der Anwesenheit oder Abwesenheit von Antigenen auf der Oberfläche der Zellen, gegen die das Immuntoxin gerichtet ist, ab.

**[0043]** Die Erfindung schließt die Verwendung von humanisierten Einzelketten-Fv-Bindungsstellen als Teil von Bildgebungsverfahren und Tumortherapien ein. Die Proteine können durch intravenöse oder intramuskuläre Injektion verabreicht werden. Wirksame Dosierungen für die Einzelketten-Fv-Konstrukte bei anti-Tumor-Therapien oder bei effektiver Darstellung von Tumoren können durch Routineexperimente unter Berücksichtigung der Zielsetzung der Behandlung bestimmt werden.

**[0044]** Die für die injizierbare Anwendung geeigneten pharmazeutischen Formulierungen schließen sterile wässrige Lösungen oder Dispersionen ein. In allen Fällen muss die Formulierung steril sein und sie muss flüchtig sein, um einfach mit einer Spritze verabreichbar zu sein. Sie muss unter den Bedingungen der Herstellung und Lagerung stabil sein und muss gegen die kontaminierende Wirkung von Mikroorganismen geschützt sein. Dies kann zum Beispiel durch Filtration durch einen sterilen 0,22 Mikron-Filter und/oder Lyophilisierung, gefolgt von einer Sterilisierung mit einer Gammastrahlenquelle, erreicht werden.

**[0045]** Sterile, injizierbare Lösungen werden durch Aufnahme des erfindungsgemäßen Einzelkettenkonstruktes in der benötigten Menge in dem geeigneten Lösemittel wie Natriumphosphat-gepufferter Kochsalzlösung, gefolgt von einer Sterilfiltration hergestellt. Wie hier verwendet, schließt "ein physiologisch verträglicher Träger" jedes und alle Lösemittel, Dispersionsmittel, antibakterielle und gegen Pilze gerichtete Wirkstoffe, die für den Menschen ungiftig sind, und ähnliches ein. Die Verwendung solcher Mittel und Wirkstoffe bei pharmazeutisch aktiven Stoffen ist im Fachgebiet gut bekannt. Das Mittel oder der Wirkstoff muss mit der Erhaltung der richtigen Konformation der Einzelpolypeptidketten und mit einer Verwendung in den therapeutischen Zusammensetzungen vereinbar sein. Ergänzende aktive Inhaltsstoffe können den Zusammensetzungen ebenfalls zugesetzt werden.

**[0046]** Ein bispezifisches Einzelketten-Fv könnte auch an ein Toxin fusioniert werden. Zum Beispiel wäre ein bispezifisches sFv-Konstrukt mit Spezifität für c-erbB-2 und den Transferrin-Rezeptor, einer Zielstruktur, die schnell internalisiert wird, aufgrund der Internalisierung des Transferrin-Rezeptor/sFv-Toxin-Komplexes ein effektiver cytolytischer Wirkstoff. Ein sFv-Fusionsprotein kann auch mehrere Protein-Domänen auf der gleichen Polypeptid-Kette einschließen, zum Beispiel EGF-sFv-Ricin A, wobei die EGF-Domäne die Internalisierung von Toxin bei Bindung des sFv durch Wechselwirkung mit dem EGF-Rezeptor fördert.

**[0047]** Die erfindungsgemäßen Einzel-Polypeptidketten können mit Radioisotopen wie zum Beispiel Jod-131,



Indium-111 und Technetium-99m markiert werden. Beta-Strahler wie Technetium-99m und Indium-111 sind bevorzugt, weil sie mit einer Gammakamera nachweisbar sind und für die Bildgebung in vivo vorteilhafte Halbwertszeiten haben. Die Einzel-Polypeptidketten können zum Beispiel mit radioaktiven Atomen und wie bei Yttrium-90, Technetium-99m oder Indium-111 über einen konjugierten Metall-Chelator (siehe zum Beispiel, Khaw et al., Science, 209 (1980), 295; Gansow et al., US-Patent Nr. 4,472,509; Hnatowich, US-Patent Nr. 4,479,930) oder durch andere, dem Fachmann bekannte Standardverfahren der Isotop-Bindung an Proteine markiert werden.

**[0048]** Die Erfindung stellt somit intakte Bindungsstellen für c-erbB-2 oder damit verwandte Antigene bereit, die analog zu  $V_H$ - $V_L$ -Dimeren sind und die über eine Polypeptidsequenz verbunden sind, um ein zusammengesetztes ( $V_H$ -Linker- $V_L$ )<sub>n</sub>- oder  $V_L$ -Linker- $V_H$ )<sub>n</sub>-Polypeptid zu erzeugen, wobei n gleich oder größer als 1 ist, welches im Wesentlichen frei vom Rest des Antikörpermoleküls sind und welches eine nachweisbare Einheit oder eine dritte Polypeptidsequenz, die entweder an  $V_H$  oder  $V_L$  gebunden ist, einschließen kann.

**[0049]** [Fig. 2A](#) – [Fig. 2E](#) veranschaulichen Beispiele von Proteinstrukturen, die Ausführungsformen der Erfindung sind und durch Befolgen der hier offenbarten Lehre hergestellt werden können. Alle sind durch mindestens ein biosynthetisches sFv-Einzelkettensegment, welches eine Bindungsstelle definiert, und durch das Enthalten von Aminosäuresequenzen, die CDRs oder FRs einschließen, oft aus verschiedenen Immunglobulinen abgeleitet, oder durch Sequenzen, die zu einem Abschnitt von CDRs und FRs aus verschiedenen Immunglobulinen homolog sind, charakterisiert.

**[0050]** [Fig. 2A](#) stellt die Einzel-Polypeptidkette sFv 100 dar, welche Polypeptid **10** umfasst, das eine Aminosäuresequenz hat, die analog zu der variablen Region der schweren Kette ( $V_H$ ) eines gegebenen monoklonalen anti-c-erbB-2 Antikörpers ist, das durch sein carboxy-terminates Ende an den Polypeptid-Linker **12** gebunden ist, welcher seinerseits an das Polypeptid **14** gebunden ist, welches eine Aminosäuresequenz hat, die analog zu der variablen Region der leichten Kette ( $V_L$ ) des anti-c-erbB-2 monoklonalen Antikörpers ist. Selbstverständlich können die Domänen der leichten und schweren Kette in umgekehrter Anordnung vorliegen. Der Linker **12** sollte mindestens lang genug sein (zum Beispiel etwa 10 bis 15 Aminosäuren oder etwa 40 Angström), um es den Ketten **10** und **14** zu erlauben, ihre richtige Konformation und Wechselbeziehung zwischen den Domänen anzunehmen.

**[0051]** Der Linker **12** kann eine Aminosäuresequenz einschließen, die zu einer von der Art, in die sie eingebracht werden soll, als "Selbst" erkannten Sequenz homolog ist, wenn eine Verwendung als Arzneimittel beabsichtigt ist. Hydrophile Aminosäuresequenzen ohne Struktur sind bevorzugt. Solche Linker-Sequenzen werden im Sequenzprotokoll als Aminosäurereste der Nummern 116 bis 135 in SEQ ID NO: 3 und 4 dargestellt, welche einen Teil der Linker-Sequenzen von 15 Aminosäuren einschließen, die im Sequenzprotokoll in SEQ ID NO: 12 und 14 dargestellt sind.

**[0052]** Andere Proteine oder Polypeptide können entweder an den Amino- oder Carboxy-Terminus des Proteins in der Art, wie sie in [Fig. 2A](#) veranschaulicht ist, angehängt werden. Als Beispiel wird die Leader-Sequenz **16**, die vom amino-terminalen Ende der  $V_H$ -Domäne **10** ausgeht, gezeigt.

**[0053]** [Fig. 2B](#) stellt eine andere Art des Stoffes **200** dar, welcher eine Einzel-Polypeptidkette **100** und ein anhängendes Protein **18** einschließt. An das Carboxyl-Ende der Polypeptidkette **100** (welche die FR- und CDR-Sequenzen einschließt, die eine Immunglobulin-Bindungsstelle bilden) ist ein anhängendes Protein **18** gebunden, welches aus zum Beispiel einem Toxin oder einem toxischen Fragment desselben, einem bindenden Protein, einem Enzym oder aktiven Enzym-Fragment oder einer Stelle zur Bindung für einen bildgebenden Stoff (zum Beispiel, um mit einem radioaktiven Ion wie Indium-111 ein Chelat zu bilden) besteht.

**[0054]** [Fig. 2C](#) veranschaulicht das Einzelketten-Polypeptid **300**, welches das zweite, erfindungsgemäße Einzelketten-Polypeptid **110** einschließt, das die gleiche oder eine andere Spezifität hat und über den Peptid-Linker **22** mit der ersten Einzel-Polypeptidkette **100** verbunden ist.

**[0055]** [Fig. 2D](#) veranschaulicht das Einzelketten-Polypeptid **400**, welches die Einzel-Polypeptidketten **110** und **100**, die durch den Linker **22** miteinander verbunden sind, und das anhängende Protein **18** einschließt, welches an das Carboxyl-Ende der Kette **110** gebunden ist.

**[0056]** [Fig. 2E](#) veranschaulicht die Einzel-Polypeptidkette **500**, welche die Kette **400** der [Fig. 2D](#) und das anhängende Protein **20** (EGF), welches an den Amino-Terminus der Kette **400** gebunden ist, einschließt.

**[0057]** Wie aus den [Fig. 2A](#) – [Fig. 2E](#) hervorgeht, können die erfindungsgemäßen Einzelketten-Proteine durch das Einschließen mehrerer biosynthetischer Bindungsstellen Perlen auf einer Kette ähnlich sehen, wobei jede Bindungsstelle eine einzigartige Spezifität oder wiederholte Stellen der gleichen Spezifität hat, die die Avidität des Proteins steigern. Wie dies aus dem vorstehenden belegt ist, stellt die Erfindung eine große Familie von Stoffen bereit, die Proteine umfasst, von denen mindestens ein Abschnitt eine Bindungsstelle definiert, die gemäß der variablen Region oder Regionen von Immunglobulinen gegen c-erbB-2 oder damit verwandten Antigenen gemustert ist.

**[0058]** Die erfindungsgemäßen Einzelketten-Polypeptide werden auf der DNA-Ebene entworfen. Die synthetischen DNAs werden dann in einem geeigneten Wirtssystem exprimiert und die exprimierten Proteine gesammelt und falls nötig renaturiert.

**[0059]** Die Fähigkeit, die erfindungsgemäßen Einzelpolypeptidketten zu entwerfen, hängt von der Fähigkeit ab, interessierende monoclonale Antikörper zu identifizieren und dann die Sequenz der Aminosäuren in der variablen Region dieser Antikörper oder der diese codierenden DNA zu bestimmen. Die Hybridom-Technologie ermöglicht die Herstellung von Zelllinien, die Antikörper gegen im Grunde jede erwünschte Substanz sezernieren, die eine Immunantwort auslöst. Zum Beispiel beschreibt das US-Patent Nr. 4,753,894 einige monoclonale Antikörper von Interesse, welche mit c-erbB-2 verwandte Antigene auf Brustkrebs-Zellen erkennen, und erklärt, wie solche Antikörper gewonnen wurden. Ein monoclonaler Antikörper, der für diesen Zweck besonders nützlich ist, ist 520C9 (Bjorn et al., *Cancer Res.* 45 (1985), 124-1221; US-Patent Nr. 4,753,894). Dieser Antikörper erkennt spezifisch das c-erbB-2-Antigen, welches auf der Oberfläche von verschiedenen Tumor-Zelllinien exprimiert wird, und zeigt sehr wenig Bindung an normale Gewebe. Alternative Quellen für sFv-Sequenzen mit der gewünschten Spezifität können den Vorteil von Phagen-Antikörper- und kombinatorischen Genbank-Methoden nutzen. Solche Sequenzen würden auf der cDNA von Mäusen basieren, welche mit Tumorzellmembranen oder c-erbB-2 oder mit c-erbB-2 verwandten Antigenfragmenten oder Peptiden präimmunisiert wurden (siehe zum Beispiel Clackson et al., *Nature* 352 (1991), 624 – 628).

**[0060]** Das Verfahren des Entwurfs von DNA, welche die interessierende Einzel-Polypeptidkette codiert, kann wie folgt bewerkstelligt werden. Die leichten und schweren Ketten des gewünschten Immunglobulins codierende RNA kann aus dem Cytoplasma des Hybridoms gewonnen werden, welches das Immunglobulin herstellt. Die mRNA kann dazu verwendet werden, die cDNA für die nachfolgende Isolierung der  $V_H$ - und  $V_L$ -Gene durch im Fachgebiet bekannte PCR-Verfahrensweisen herzustellen (Sambrook et al., Hrsg., *Molecular Cloning* (1989), Cold Spring Harbor Laboratories Press, NY). Die N-terminale Aminosäuresequenz der H- und L-Kette kann unabhängig durch automatisierte Edman-Sequenzierung bestimmt werden; wenn nötig, können weitere Bereiche der CDRs und flankierender FRs durch Aminosäure-Sequenzierung der V-Region-Fragmente der H- und L-Ketten bestimmt werden. Eine solche Sequenzanalyse wird nun routinemäßig durchgeführt. Dieses Wissen erlaubt einem synthetische Primer zur Isolierung von  $V_H$ - und  $V_L$ -Genen aus Hybridom-Zellen, die monoclonale Antikörper herstellen, von denen bekannt ist, dass sie c-erbB-2 oder ein damit verwandtes Antigen binden, zu entwerfen. Diese V-Gene werden die Fv-Region codieren, die im parentalen Antikörper c-erbB-2 bindet.

**[0061]** Noch ein anderer Ansatz schließt den Entwurf und die Konstruktion von synthetischen V-Genen ein, die eine Fv-Bindungsstelle codieren werden, die spezifisch für c-erbB-2 oder verwandte Rezeptoren ist. Zum Beispiel kann man mit Hilfe eines Computerprogramms wie zum Beispiel Compugene und bekannten DNA-Sequenzen von variablen Regionen native oder annähernd native FR-Sequenzen eines ersten Antikörper-Moleküls und CDR-Sequenzen eines zweiten Antikörper-Moleküls entwerfen und direkt synthetisieren. Die vorstehend beschriebenen  $V_H$ - und  $V_L$ -Sequenzen sind miteinander durch eine Aminosäurekette oder einen Linker direkt verbunden, der den C-Terminus einer Kette mit dem N-Terminus der anderen verbindet.

**[0062]** Diese Gene können, einmal synthetisiert, mit oder ohne zusätzliche DNA-Sequenzen, die zum Beispiel ein Leader-Peptid codieren, welches die Sekretion oder intrazelluläre Stabilität eines Fusionspolypeptids fördert, oder eine Leader- oder nachfolgende Sequenz, die ein zweites Polypeptid codiert, cloniert werden. Die Gene können dann direkt in einer geeigneten Wirtszelle exprimiert werden.

**[0063]** Durch direkte Sequenzierung eines Antikörpers gegen ein c-erbB-2- oder ein damit verwandtes Antigen oder durch Bezug der Sequenz aus der Literatur kann der Durchschnittsfachmann angesichts dieser Offenbarung einen Einzelketten-Fv herstellen, der jede gewünschte CDR und FR umfasst. Zum Beispiel kann unter Verwendung der DNA-Sequenz für den monoclonalen Antikörper 520C9, dargestellt im Sequenzprotokoll als SEQ ID NO: 3, ein Einzelketten-Polypeptid hergestellt werden, das eine Bindungsaffinität für ein mit c-erbB-2 verwandtes Antigen hat. Exprimierte Sequenzen können auf Bindung getestet werden und empirisch

durch den Austausch ausgewählter Aminosäuren in vergleichsweise konservierten Regionen basierend auf Beobachtungen von Tendenzen in Aminosäuresequenzdaten und/oder Computermodell-Techniken verbessert werden. Eine ausgeprägte Flexibilität im Entwurf von  $V_H$  und  $V_L$  ist möglich, weil Änderungen in der Aminosäuresequenz auf der DNA-Ebene durchgeführt werden können.

**[0064]** Dementsprechend kann der Zusammenbau von DNAs, welche die erfindungsgemäßen Einzelketten-Fv und sFv-Fusionsproteine codieren, unter Verwendung bekannter Verfahren, welche die Verwendung von verschiedenen Restriktionsenzymen, welche sequenzspezifische Spaltungen in der DNA bewirken, um glatte Enden oder kohäsive Enden zu erzeugen, von DNA-Ligasen, Techniken, welche die enzymatische Auffüllung von klebrigen Enden zu DNA mit glatten Enden ermöglichen, des Zusammenbaus von synthetischen DNAs durch Verbindung von Oligonucleotiden kurzer oder mittlerer Länge, von Verfahren der cDNA-Synthese und synthetischen Sonden zur Isolierung von Immunglobulin-Genen einschließen, durchgeführt werden. Verschiedene Promotor-Sequenzen und andere regulatorische RNA-Sequenzen, die zur Erreichung der Expression verwendet werden, und verschiedene Arten von Wirtszellen sind ebenfalls bekannt und verfügbar. Herkömmliche Transfektionsverfahren und genauso herkömmliche Verfahren zur Clonierung und Subclonierung von DNA sind für die Umsetzung dieser Erfindung nützlich und dem Durchschnittsfachmann bekannt. Verschiedene Arten von Vektoren wie Plasmide und Viren, einschließlich tierspezifischer Viren und Bakteriophagen, können verwendet werden. Die Vektoren können verschiedene Markierungs-Gene nutzen, welche einer erfolgreich transfizierten Zelle eine nachweisbare phänotypische Eigenschaft verleihen, die dazu verwendet werden kann, festzustellen, welcher aus einer Familie von Clonen die rekombinante DNA des Vektors erfolgreich eingebaut hat.

**[0065]** Natürlich sind die Verfahren zur Manipulierung, Vervielfältigung und Rekombinierung von DNA, welche interessierende Aminosäuresequenzen codiert, im Allgemeinen im Fachgebiet wohl bekannt und deshalb hier nicht im Detail beschrieben. Verfahren zur Identifizierung der isolierten V-Gene, die interessierende Antikörper-Fv-Regionen codieren, sind gut verstanden und in der Patentliteratur und anderer Literatur beschrieben. Im Allgemeinen schließen die Verfahren die Auswahl von genetischem Material ein, das entsprechend dem genetischen Code bei reverser Transkription Aminosäuresequenzen codiert, welche die interessierenden CDRs und FRs definieren.

**[0066]** Ein Verfahren zur Gewinnung von das hier offenbarte Einzelketten-Fv codierender DNA erfolgt durch den Zusammenbau von synthetischen Oligonucleotiden, die in einem herkömmlichen, automatisierten Polynucleotid-Synthesegerät hergestellt wurden, gefolgt von der Ligierung mit geeigneten Ligasen. Zum Beispiel können überlappende, complementäre DNA-Fragmente, die 15 Basen umfassen, halbmanuell unter Verwendung einer Phosphoramidit-Chemie synthetisiert werden, wobei die Endabschnitte nicht-phosphoryliert belassen werden, um eine Polymerisierung während der Ligierung zu verhindern. Ein Ende der synthetischen DNA wird mit einem "klebrigen Ende" belassen, welches der Wirkungsstelle einer bestimmten Restriktionsendonuclease entspricht, und das andere Ende wird mit einem Ende belassen, welches der Wirkungsstelle einer anderen Restriktionsendonuclease entspricht. Alternativ kann dieser Ansatz voll automatisiert werden. Die das Einzelketten-Polypeptid codierende DNA kann durch Synthetisierung längerer Einzelstrang-Fragmente (zum Beispiel 50 – 100 Nucleotide lang) in zum Beispiel einem Biosearch-Oligonucleotid-Synthesegerät und nachfolgender Ligierung der Fragmente erzeugt werden.

**[0067]** Zusätzliche Nucleotidsequenzen, die zum Beispiel die Aminosäuren einer konstanten Region oder ein bioaktives Molekül codieren, können auch mit der Gensequenz verbunden werden, um ein bifunktionelles Protein herzustellen.

**[0068]** Zum Beispiel können die, wie vorstehend beschrieben, entworfenen, synthetischen Gene und DNA-Fragmente durch Zusammenbau chemisch synthetisierter Oligonucleotide hergestellt werden. 15 – 100mer-Oligonucleotide könne mit einem Biosearch DNA Model 8600 Synthesizer synthetisiert und durch Polyacrylamid-Gel-Elektrophorese (PAGE) in Tris-Borat-EDTA-Puffer (TBE) aufgereinigt werden. Die DNA wird dann durch Elektroelution aus dem Gel eluiert. Überlappende Oligomere können durch die T4-Polynucleotid-Kinase phosphoryliert und zu größeren Blöcken ligiert werden, welche ebenfalls durch PAGE aufgereinigt werden können.

**[0069]** Die Blöcke oder die Paare längerer Oligonucleotide können unter Verwendung eines geeigneten Clonierungs-Vektors, zum Beispiel pUC, in *E. coli* cloniert werden. Anfänglich kann dieser Vektor durch Einzelstrang-Mutagenese verändert werden, um verbleibende, an sechs Basen veränderte Stellen zu entfernen. Zum Beispiel kann  $V_H$  in Form fünf ursprünglicher Blöcke, welche sich zwischen den folgenden Restriktionschnittstellen erstrecken: (1) EcoRI zur ersten Narf-Stelle; (2) erste Narf zur XbaI; (3) XbaI zur Sall; (4) Sall zur

NcoI; und (5) NcoI zur BamHI; synthetisiert und in pUC cloniert werden. Diese clonierten Fragmente können dann isoliert und durch mehrere Ligierungen dreier Fragmente und Clonierungsschritte in dem pUC8-Plasmid zusammengesetzt werden. Gewünschte, durch PAGE ausgewählte Ligierungen werden dann in zum Beispiel in den E. coli -Stamm JM83 transformiert und entsprechend Standardverfahren auf LB-Ampicillin + XGal-Platten ausplattiert. Die Gen-Sequenz kann nach der Clonierung durch Supercoil-Sequenzierung oder nach einer Sub-Clonierung in M13 durch das Dideoxy-Verfahren nach Sanger (Molecular Cloning, Sambrook et al., Hrsg., 2. Ausgabe (1989), Band 2, Cold Spring Harbor Laboratory Press, NY) bestätigt werden.

**[0070]** Die konstruierten Gene können in geeigneten prokaryontischen Wirten wie verschiedenen E. coli-Stämmen und in eukaryontischen Wirten wie Ovarzellen Chinesischer Hamster (CHO), Maus-Myelom-, Hybridom-, Transfectom- und menschlichen Myelomzellen exprimiert werden.

**[0071]** Wenn das Gen in E. coli exprimiert werden soll, kann es zuerst in einen Expressionsvektor cloniert werden. Dies wird bewerkstelligt, indem das konstruierte Gen stromabwärts von einer Promotorsequenz wie Trp oder Tac und einem Gen, welches ein Leader-Polypeptid wie das Fragment B (FB) des Proteins A der Staphylokokken codiert, positioniert wird. Das sich ergebende, exprimierte Fusionsprotein sammelt sich in lichtbrechenden Körperchen im Cytoplasma der Zellen und kann nach dem Aufbrechen der Zellen durch eine French-Press oder Ultraschallbehandlung geerntet werden. Die lichtbrechenden Körperchen werden löslich gemacht und die exprimierten Fusionsproteine werden durch die bereits für viele andere rekombinante Proteine etablierten Verfahren (Huston et al., 1988, vorstehend) gespalten und rückgefaltet oder bei direkten Expressionsverfahren gibt es keinen Leader und die Einschlusskörperchen können ohne Spaltung rückgefaltet werden (Huston et al., Methods in Enzymology, Band 203 (1991), Seiten 46 – 88).

**[0072]** Zum Beispiel kann eine nachfolgende proteolytische Spaltung der isolierten sFv von ihren Leader-Sequenz-Fusionen durchgeführt werden, um freie sFvs zu ergeben, welche renaturiert werden können, um eine intakte biosynthetische Hybrid-Antikörper-Bindungsstelle zu erhalten. Die Spaltstelle ist bevorzugt unmittelbar benachbart zum sFv-Polypeptid und schließt eine Aminosäure oder Sequenz von Aminosäuren ein, die keine Aminosäure oder Aminosäuresequenz enthält, die in der Aminosäure-Struktur der Einzel-Polypeptidkette zu finden ist.

**[0073]** Die Spaltstelle ist bevorzugt für die spezifische Spaltung durch ein ausgewähltes Mittel entworfen. Endopeptidasen sind bevorzugt, obwohl nicht-enzymatische (chemische) Spaltmittel verwendet werden können. Viele nützliche Spaltmittel, zum Beispiel Cyanogenbromid, verdünnte Säure, Trypsin, Staphylococcus aureus-V-8-Protease, Post-Prolin-Spaltenzym, Blutgerinnungsfaktor Xa, Enterokinase und Renin, erkennen und spalten bevorzugt oder ausschließlich an bestimmten Spaltstellen. Ein zur Zeit bevorzugtes Peptidsequenz-Spaltmittel ist die V-8-Protease. Die zur Zeit bevorzugte Spaltstelle ist an einem Glu-Rest. Andere nützliche Enzyme erkennen mehrere Reste als eine Spaltstelle, zum Beispiel Faktor Xa (Ile-Glu-Gly-Arg) oder Enterokinase (Asp-Asp-Asp-Asp-Lys). Verdünnte Säure spaltet bevorzugt die Peptidbindung zwischen Asp-Pro-Resten und CNBr in Säure spaltet nach Met, sofern es nicht von Tyr gefolgt wird.

**[0074]** Wenn das konstruierte Gen in eukaryontischen Hybridom-Zellen, dem herkömmlichen Expressionssystem für Immunglobuline, exprimiert werden soll, wird es zuerst in einen Expressionsvektor eingefügt, der zum Beispiel den Immunglobulin-Promotor, ein Sekretionssignal, Immunglobulinhancer und verschiedene Introns enthält. Dieses Plasmid kann auch Sequenzen enthalten, die ein anderes Polypeptid codieren, wie die gesamte oder einen Teil der konstanten Region, was die Expression eines gesamten Teils der schweren oder leichten Kette ermöglicht, oder wenigstens einen Teil eines Toxins, Enzyms, Cytokins oder Hormons. Das Gen wird über die etablierte Elektroporations- oder Protoplastenfusions-Verfahren in Myelom-Zellen transfiziert. So transfizierte Zellen können dann  $V_H$ -Linker- $V_L$  oder  $V_L$ -Linker- $V_H$ -Einzelketten-Fv-Polypeptide exprimieren, wobei jedes von ihnen in den verschiedenen, vorstehend diskutierten Arten mit einer Proteindomäne, welche eine weitere Funktion (zum Beispiel Cytotoxizität) hat, verbunden sein kann.

**[0075]** Zum Zusammenbau einer einzelnen, zusammenhängenden Kette von Aminosäuren, welche mehrere Bindungsstellen spezifiziert, werden Restriktionsschnittstellen an den Grenzen von DNA, die eine einzelne Bindungsstelle (das heißt,  $V_H$ -Linker- $V_L$ ) codiert, verwendet oder geschaffen, wenn sie nicht bereits vorliegen. DNAs, die einzelne Bindungsstellen codieren, werden ligiert und in Shuttle-Plasmide cloniert, aus denen sie weiter zusammengesetzt und in das Expressionsplasmid cloniert werden können. Die Anordnung der Domänen wird variiert werden und die Spacer zwischen den Domänen stellen die für die unabhängige Faltung der Domänen benötigte Flexibilität bereit. Die optimale Architektur in Bezug auf Expressionsniveaus, Rückfaltung und funktionelle Aktivität wird empirisch bestimmt werden. Um bivalente sFvs herzustellen, wird zum Beispiel das Stoppcodon in dem die erste Bindungsstelle codierenden Gen zu einem offenem Leserahmen geändert

und es werden am richtigen Platz mehrere Glycin- und Serincodons, einschließlich einer Restriktionsschnittstelle wie BamHI (die Gly-Ser codiert) oder XhoI (die Gly-Ser-Ser codiert), eingesetzt. Das zweite sFv-Gen wird an seinem 5'-Ende, die gleiche Restriktionsschnittstelle im gleichen Leserahmen erhaltend, ähnlich verändert. Die Gene werden an dieser Stelle kombiniert, um das bivalente sFv-Gen zu ergeben.

**[0076]** Linker, die den C-Terminus einer Domäne mit dem N-Terminus der nächsten verbinden, umfassen im Allgemeinen hydrophile Aminosäuren, welche in physiologischen Lösungen eine nicht-strukturierte Konfiguration annehmen und bevorzugt frei von Resten sind, die große Seitengruppen haben, welche mit der richtigen Faltung der  $V_H$ - ,  $V_L$ - oder angehängter Ketten interferieren könnten. Ein nützlicher Linker hat die Aminosäuresequenz  $[(Gly)_4Ser]_3$  (siehe SEQ ID NO: 9 und 10, Reste Nr. 410 – 424). Ein zur Zeit bevorzugter Linker hat die Aminosäuresequenz, welche 2 oder 3 Wiederholungen von  $[(Ser)_4Gly]$ , wie  $[(Ser)_4Gly]_2$  oder  $[(Ser)_4Gly]_3$  (siehe SEQ ID NO: 3 und 4), umfasst.

**[0077]** Die Erfindung wird weiter durch die folgenden, nicht-begrenzenden Beispiele veranschaulicht.

## BEISPIELE

### 1. Antikörper gegen mit c-erbB-2 verwandte Antigene

**[0078]** Monoklonale Antikörper gegen Brustkrebs wurden unter Verwendung menschlicher Brustkrebszellen oder von Membran-Extrakten der Zellen für die Immunisierung von Mäusen, wie in Frankel et al., J. Biol. Resp. Modif. 4 (1985), 273-286 beschrieben, hierbei durch Bezugnahme eingeschlossen, entwickelt. Es wurden Hybridome hergestellt und unter Verwendung einer Reihe von normalen und Brustkrebszellen auf die Herstellung von Antikörpern hin ausgewählt. Eine Reihe von 8 Normalgewebe-Membranen, eine Fibroblasten-Zelllinie und Gefrierschnitte von Brustkrebsgeweben wurden bei der Durchmusterung verwendet. Kandidaten, welche die erste Durchmusterung bestanden hatten, wurden auf 16 Schnitten von Normalgewebe, 5 Arten normaler Blutzellen, 11 Schnitten von nicht aus der Brust stammenden Neoplasmen, 21 Brustkrebschnitten und 14 Brustkrebszelllinien weiter untersucht. Aus dieser Auswahl wurden 127 Antikörper ausgewählt. In Kontrollexperimenten wurden irrelevante Antikörper und nicht aus der Brust stammende Krebszelllinien verwendet.

**[0079]** Es wurde herausgefunden, dass nützliche monoklonale Antikörper 520C9, 454C11, (A.T.C.C. Nr. HB8696 bzw. HB8484) und 741F8 einschließen. In dieser Durchmusterung als für Brustkrebs selektiv identifizierte Antikörper waren gegen 5 verschiedene Antigene reaktiv. Die Größen der Antigene, welche die Antikörper erkennen, waren jeweils: 200 kD; eine Serie von Proteinen, die wahrscheinlich Abbauprodukte sind, mit Molekulargewichten von 200 kD, 93 kD, 60 kD und 37 kD; 180 kD (Transferrinrezeptor); 42 kD und 55 kD. Von den gegen die fünf Arten von Antigenen gerichteten Antikörpern sind die spezifischsten diejenigen, welche gegen das 200 kD -Antigen gerichtet sind, wobei 520C9 ein repräsentativer Antikörper gegen diese Art von Antigen ist. 520C9 reagiert mit weniger Brustkrebsgeweben (etwa 20 – 70 %, abhängig von den Versuchsbedingungen) und er reagiert mit der geringsten Zahl von Normalgeweben als jeder andere der Antikörper. 520C9 reagiert mit Nierentubuli (wie dies viele monoklonale Antikörper tun), aber, neben einigen der untersuchten Gewebe, nicht mit Pankreas, Ösophagus, Lunge, Colon, Magen, Hirn, Tonsillen, Leber, Herz, Ovar, Haut, Knochen, Uterus, Blase oder normalem Brustgewebe.

### 2. Herstellung einer den Antikörper 520C9 codierenden cDNA-Genbank

**[0080]** Polyadenylierte RNA wurde aus annähernd  $1 \times 10^8$  (520C9-Hybridom) Zellen unter Verwendung des "Fast Track" mRNA-Isolationskits von Invitrogen (San Diego, CA) isoliert. Das Vorhandensein von RNA der schweren Kette des Immunglobulins wurde durch eine Northern-Untersuchung (Sambrook et al., Hrsg., Molecular Cloning, 2. Auflage (1989), Cold Spring Harbor Laboratory Press, NY) unter Verwendung einer rekombinanten Sonde, welche die verschiedenen J-Regionen der genomischen DNA der schweren Kette enthält, bestätigt. Unter Verwendung von jeweils 6 µg RNA wurde cDNA unter Verwendung des Invitrogen-cDNA-Synthese-Systems mit entweder Zufalls- oder Oligo-dT-Primern präpariert. Der Synthese nachfolgend wurde die cDNA der Größe nach durch eine einer Agarosegel-Elektrophorese nachfolgenden Isolierung von 0,5 – 3,0 Kilobasen (Kb)-Fragmenten selektiert. Nach einer Optimierung des Verhältnisses von cDNA zu Vektor wurden diese Fragmente dann mit dem pcDNA II-Clonierungsvektor von Invitrogen ligiert.

### 3. Isolierung von $V_H$ - und $V_L$ -Domänen

**[0081]** Nach Transformierung der Bakterien mit Plasmid-Genbank-DNA wurde eine Hybridisierung der Kolonien unter Verwendung von Sonden für die konstante (C) Region und die Joining (J)-Region der Antikörper

nach Genen entweder der leichten oder der schweren Kette durchgeführt. Siehe Orlandi, R., et al., Proc. Nat. Aca. Sci., 86 (1989), 3833. Die Sonde für die konstante Region der Antikörper kann aus jeder der Nucleotidsequenzen der leichten oder schweren Kette eines Immunglobulin-Gens unter Verwendung bekannter Verfahren gewonnen werden. Einige mögliche positive Clone wurde für die Gene der schweren und der leichten Kette identifiziert und diese wurden, nach Aufreinigung durch eine zweite Durchmusterungsrunde, sequenziert. Ein Clon (M207) enthielt die Sequenz einer nicht-funktionellen Kappa-Kette, welche ein Tyrosin enthält, das ein konserviertes Cystein ersetzt und auch vorzeitig aufgrund einer Deletion über 4 Basen, welche eine Mutation durch Verschiebung des Leserahmens im Übergang von der variablen zur J-Region bedingt, endet. Ein zweiter Clon der leichten Kette (M230) enthielt, außer den letzten 18 Aminosäuren der konstanten Region und annähernd der Hälfte der Signal-Sequenz, faktisch das gesamte Gen der leichten Kette von 520C9. Die variable Region der schweren Kette von 520C9 lag in einem Clon von annähernd 1100 Basenpaaren (F320) vor, welcher nahe dem Ende der CH2-Domäne endete.

#### 4. Mutagenese von $V_H$ - und $V_L$

**[0082]** Um das sFv zusammzusetzen, wurden sowohl die variable Region der schweren als auch die der leichten Kette mutiert, um geeignete Restriktionsschnittstellen einzufügen (Kunkel, T. A., Proc. Nat. Acad. Sci. USA 82 (1985), 1373). Der Clon der schweren Kette (F320) wurde mutiert, um eine BamHI-Stelle am 5'-Ende von  $V_H$  (F321) einzufügen. Die leichte Kette wurde ebenfalls durch gleichzeitiges Einfügen einer EcoRV-Stelle am 5'-Ende und einer PstI-Stelle mit einem Translations-Stoppocodon am 3'-Ende der variablen Region (M231) mutiert.

#### 5. Sequenzierung

**[0083]** Die leichte und die schwere Kette codierende cDNA-Clone wurden unter Verwendung externer Standard-Primer für pUC und einiger spezifischer interner Primer, welche auf der Basis der für die schwere Kette erhaltenen Sequenzen hergestellt worden waren, sequenziert. Die Nucleotidsequenzen wurden in einer Genbank-Homologie-Suche (Nucscan-Programm von DNA-star) analysiert, um endogene Immunglobulin-Gene auszuschließen. Die Translation in Aminosäuresequenzen wurde mit Aminosäuresequenzen im von E. Kabat herausgegebenen NIH-Atlas überprüft.

**[0084]** Vom Immunglobulin 520C9 abgeleitete Aminosäuresequenzen bestätigten die Identität von diesen  $V_H$ - und  $V_L$ -cDNA-Clonen. Der Clon der schweren Kette pF320 begann 6 Nucleotide stromaufwärts vom ersten ATG-Codon und dehnte sich bis in die die CH2 codierende Region aus, es fehlten ihm aber die letzten 9 Aminosäure-Codons der konstanten Domäne CH2 und alle der die CH3 codierenden Region sowie die nicht-translatierte 3'-Region und der Poly-A-Schwanz. Von einem weiteren kurzen Clon der schweren Kette, der nur die CH2 und CH3 codierenden Regionen und den Poly-A-Schwanz enthielt, wurde anfänglich angenommen, dass er den fehlenden Teil der schweren Kette von 520C9 darstellte. Die Überlappung beider Sequenzen war jedoch nicht identisch. Der Clon (pF320) von 520C9 codiert die CH1- und CH2-Domäne des murinen IgG1, während der kurze Clon pF315 CH2 und CH3 von IgG2b codiert.

#### 6. Entwurf von Genen

**[0085]** Eine Nucleinsäuresequenz, die eine zusammengesetzte 520C9 sFv-Region codiert, die eine Einzelketten-Fv-Bindungsstelle enthält, welche mit c-erbB-2 verwandte Tumorantigene erkennt, wurde mit Hilfe der Compugene-Software entworfen. Das Gen enthält Nucleinsäuresequenzen, welche die  $V_H$ - und  $V_L$ -Regionen des vorstehend beschriebenen 520C9-Antikörpers codieren, die miteinander durch ein doppelsträngiges, synthetisches Oligonucleotid verbunden sind, das ein Peptid mit der im Sequenzprotokoll als Aminosäurereste 116 bis 133 in SEQ ID NO: 3 und 4 dargestellten Aminosäuresequenz codiert. Dieses Linker-Oligonucleotid enthält die helfenden Clonierungsstellen EcoRI und BamHI und wurde so entworfen, dass es als Stellen für den Zusammenbau SacI und EcoRV in der Nähe seines 5'- beziehungsweise 3'-Endes enthält. Diese Stellen ermöglichen das Zusammenfügen und die Ligierung an die 3'- und 5'-Enden von 520C9- $V_H$  bzw. - $V_L$ , welche diese Stellen ebenfalls enthalten ( $V_H$ -Linker- $V_L$ ). Die Reihenfolge der Anbindung an das Oligonucleotid kann jedoch in diesem oder jedem erfindungsgemäßen sFv umgekehrt werden ( $V_L$ -Linker- $V_H$ ). Andere Restriktionsschnittstellen wurden in das Gen eingebaut, um alternative Stellen für den Zusammenbau bereitzustellen. Eine das FB-Fragment des Protein A codierende Sequenz wurde als Leader verwendet.

**[0086]** Die Erfindung beinhaltet auch ein humanisiertes Einzelketten-Fv, das heißt es enthält menschliche Framework-Sequenzen und CDR-Sequenzen, welche die Bindung an c-erbB-2 spezifizieren, zum Beispiel wie die CDRs des Antikörpers 520C9. Das humanisierte Fv ist somit fähig c-erbB-2 zu binden, während es eine

geringe oder keine Immunantwort auslöst, wenn es einem Patienten verabreicht wird. Eine Nucleinsäuresequenz, die ein humanisiertes sFv codiert, kann wie folgt entworfen und zusammengesetzt werden. Zwei Strategien zum Entwurf eines sFv sind besonders nützlich. Es kann eine Suche nach Homologien in der Datenbank Gen-Bank nach den am nächsten verwandten menschlichen Framework-Regionen durchgeführt werden und die FR-Regionen des sFv können gemäß den in der Suche identifizierten Sequenzen durch Mutation verändert werden, um die entsprechende menschliche Sequenz zu reproduzieren; oder es kann die Information aus auf Röntgenstrukturen basierenden Modellen von Fab-Fragmenten im Computer verwendet werden (Amit et al., Science 233 (1986), 747 – 753; Colman et al., Nature 326 (1987), 358 – 363; Sheriff et al., Proc. Nat. Aca. Sci. 84 (1987), 8075 – 8079; und Satow et al., J. Mol. Biol. 190 (1986), 593 – 604, von denen alle hierbei durch Bezugnahme eingeschlossen sind). In einem bevorzugten Fall können die am stärksten homologen menschlichen  $V_H$ - und  $V_L$ -Sequenzen aus einer Sammlung von durch PCR clonierten menschlichen V-Regionen ausgewählt werden. Die FRs werden synthetisch hergestellt und an CDRs fusioniert, um schrittweise zunehmend vollständige V-Regionen durch auf PCR basierende Ligierung herzustellen, bis die vollständige humanisierte  $V_L$  und  $V_H$  fertig gestellt sind. Zum Beispiel kann ein humanisiertes sFv, welches ein Hybrid aus den CDRs des murinen Antikörpers 520C9 und den FRs des menschlichen Myelom-Proteins NEW ist, in der Art entworfen werden, dass jede variable Region die murine Bindungsstelle innerhalb eines menschlichen Frameworks aufweist (FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4). Die Kristallstruktur des Fab NEW (Saul et al., J. Biol. Chem. 253 (1978), 585 – 597) kann auch dazu verwendet werden, die Position der FRs in den variablen Regionen vorherzusagen. Sind diese Regionen einmal vorhergesagt, kann die Aminosäuresequenz oder die entsprechende Nucleotidsequenz der Regionen bestimmt werden und die Sequenzen können synthetisiert und in ein Shuttle-Plasmid cloniert werden, aus dem heraus sie weiter zusammengesetzt und in ein Expressionsplasmid cloniert werden können; alternativ können die FR-Sequenzen des sFv 520C9 direkt durch Mutation verändert werden und die Veränderungen durch Supercoil-Sequenzierung mit internen Primern überprüft werden (Chen et al., DNA 4 (1985), 165 – 170).

## 7. Herstellung und Aufreinigung des sFv 520C9

### A. Solubilisierung von Einschlusskörperchen

**[0087]** Das sFv 520C9-Plasmid, welches auf einem  $T_7$ -Promotor und Vektor basiert, wurde durch direkte Expression der fusionierten Gensequenz, die im Sequenzprotokoll als SEQ ID NO: 3 dargestellt ist, in *E. coli* hergestellt. Einschlusskörperchen (15,8 g) aus einer 2,0 Liter-Fermentation wurden mit 25 mM Tris, 10 mM EDTA, pH 8,0 (TE) und 1M Guanidinhydrochlorid (GuHCl) gewaschen. Die Einschlusskörperchen wurden in TE, 6M GuHCl, 10 mM Dithiothreitol (DTT), pH 9,0 solubilisiert und ergaben 3825  $A_{280}$ -Einheiten Material. Dieses Material wurde mit Ethanol gefällt, mit TE, 3 M Harnstoff gewaschen, dann in TE, 8 M Harnstoff, 10 mM DTT, pH 8,0, resuspendiert. Dieser Fällungsschritt bereitete das Protein auf die Aufreinigung des denaturierten sFv durch einen Ionenaustauscher vor.

### B. Ionenaustauscher-Chromatographie

**[0088]** Die solubilisierten Einschlusskörperchen wurden einer Ionenaustauscher-Chromatographie unterzogen, um dadurch die kontaminierenden Nucleinsäuren und die Proteine von *E. coli* vor der Renaturierung des sFv zu entfernen. Die solubilisierten Einschlusskörperchen in 8 M Harnstoff wurden mit TE auf eine Endkonzentration des Harnstoffs von 6 M verdünnt, dann durch 100 ml Fast-Flow-DEAE-Sepharose in einer Radial-Fluss-Säule geleitet. Das sFv wurde in der ungebundenen Fraktion gewonnen (69 % der Ausgangsprobe).

**[0089]** Der pH-Wert dieser sFv-Lösung ( $A_{280} = 5,7$ ; 290 ml) wurde mit 1 M Essigsäure auf 5,5 eingestellt, um sie für den Auftrag auf eine Fast-Flow-S-Sepharose-Säule vorzubereiten. Als sich der pH-Wert unter 6,0 bewegte, bildete sich jedoch ein Niederschlag in der Probe. Die Probe wurde geklärt; 60 % der Probe waren im Sediment und 40 % im Überstand. Der Überstand wurde durch 100 ml Fast-Flow-S-Sepharose geleitet und der sFv in der ungebundenen Fraktion gewonnen. Das Sediment wurde in TE, 6M GuHCl, 10 mM DTT, pH 9,0 resolubilisiert und es wurde herausgefunden, dass es auch hauptsächlich sFv in einer Ansammlung mit einem Volumen von 45 ml mit einer Absorption von 20 Absorptionseinheiten bei 280 nm enthielt. Mit diesem eingengten sFv-Konzentrat wurden die verbleibenden Schritte der Aufreinigung durchgeführt.

### C. Renaturierung des sFv

**[0090]** Die Renaturierung des sFv wurde unter Anwendung eines Disulfid-beschränkten Rückfaltungsansatzes, bei dem die Disulfide oxidiert wurden, während der sFv vollständig denaturiert war, gefolgt von der Entfernung des denaturierenden Agens und der Rückfaltung, bewerkstelligt. Die Oxidierung der sFv-Proben wurde



in TE, 6 M GuHCl, 1 mM oxidiertem Glutathion (GSSG), 0,1 mM reduziertem Glutathion (GSH), pH 9,0 durchgeführt. Das sFv wurde in dem Oxidationspuffer auf eine End- $A_{280} = 0,075$  des Proteins in einem Volumen von 4000 ml verdünnt und über Nacht bei Raumtemperatur inkubiert. Nach einer Oxidation über Nacht wurde diese Lösung gegen 10 mM Natriumphosphat, 1 mM EDTA, 150 mM NaCl, 500 mM Harnstoff, pH 8,0 (PENU) [ $4 \times (20 \text{ Liter} \times 24 \text{ Stunden})$ ] dialysiert. In der rückgefalteten Probe wurden niedrige Aktivitätsspiegel nachgewiesen.

#### D. Membranfraktionierung und Konzentrierung von aktivem sFv

**[0091]** Um aggregiertes und fehlgefaltetes Material vor jeglichem Konzentrationsschritt zu entfernen, wurde das dialysierte und rückgefaltete sFv 520C9 (5050 ml) durch eine 100K-MWCO-Membran (Ausschluss: Mol. Gew. 100000) ( $4 \times 60 \text{ cm}^2$ ) unter Verwendung eines Minitan-Ultrafiltrationsapparates (Millipore) filtriert. Dieser Schritt benötigte, hauptsächlich aufgrund der Bildung von Niederschlag im Rückstand und einem Verstopfen der Membran mit Zunahme der Konzentration im Rückstand, eine bemerkenswerte Zeitspanne (9 Stunden). 95 % des Proteins, mit 79 % in der Form nicht-löslichen Materials, in der zurückgefalteten Probe wurde durch die 100K-Membranen zurückgehalten. Das 100K-Rückstand hatte eine sehr geringe Aktivität und wurde verworfen.

**[0092]** Das 100K-Filtrat enthielt den Großteil der löslichen sFv-Aktivität bei der c-erbB-2-Bindung und es wurde als nächstes unter Verwendung von 10K-MWCO-Membranen (Ausschluss: Mol. Gew. 10000) ( $4 \times 60 \text{ cm}^2$ ) in dem Minitan auf ein Volumen von 100 ml ( $50 \times$ ) konzentriert. Dieses Material wurde unter Verwendung einer YM10-10K-MWCO-Membran in einer gerührten 50ml-Amicon-Zelle auf ein Endvolumen von 5,2 ml ( $1000 \times$ ) weiter konzentriert. Nur eine geringe Menge von Niederschlag bildete sich während der zwei 10K-Konzentrationsschritte. Die spezifische Aktivität dieses konzentrierten Materials wurde in Bezug auf die anfängliche dialysierte Rückfaltung wesentlich erhöht.

#### E. Größenausschluss-Chromatographie von konzentriertem sFv

**[0093]** Als das rückgefaltete sFv durch Größenausschluss-Chromatographie fraktioniert wurde, wurde bestimmt, dass die gesamte Aktivität des sFv 520C9 an der Position des gefalteten Monomers eluierte. Um eine Anreicherung aktiven Monomers durchzuführen, wurde die  $1000 \times$  konzentrierte sFv-Probe auf einer Sephacryl-S200-HR-Säule ( $2,5 \times 40 \text{ cm}$ ) in PBSA (2,7 mM KCl, 1,1 mM  $\text{KH}_2\text{PO}_4$ , 138 mM NaCl, 8,1 mM  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ , 0,02 %  $\text{NaN}_3$ ) + 0,5 M Harnstoff fraktioniert. Das Elutionsprofil der Säule und die SDS-PAGE-Untersuchung der Fraktionen zeigte zwei sFv-Monomer-Peaks. Die zwei Peak-Fraktionen der sFv-Monomere wurden vereinigt (insgesamt 10 ml) und zeigten Bindungsaktivität gegenüber c-erbB-2 in Konkurrenzversuchen.

#### F. Affinitätsreinigung des sFv 520C9

**[0094]** Die extrazelluläre Domäne (ECD) von c-erbB-2 wurde in mit Baculoviren infizierten Insektenzellen exprimiert. Dieses Protein (ECD c-erbB-2) wurde auf einer Agarose-Affinitätsmatrix immobilisiert. Der Monomerpeak des sFv wurde gegen PBSA dialysiert um den Harnstoff zu entfernen und dann auf eine  $0,7 \times 4,5 \text{ cm}$  ECD c-erbB-2-Agarose-Affinitätssäule in PBSA aufgetragen. Die Säule wurde bis auf die  $A_{280}$ -Grundlinie gewaschen, dann mit PBSA + 3 M LiCl, pH = 6,1 eluiert. Die Peak-Fraktionen wurden vereinigt (4 ml) und gegen PBSA dialysiert, um das LiCl zu entfernen. 72  $\mu\text{g}$  aufgereinigten sFv wurden aus 750  $\mu\text{g}$  der S-200-Monomerfraktionen gewonnen. Die Aktivitätsmessungen mit den Säulenfraktionen wurden durch einen Konkurrenzversuch bestimmt. Kurz gesagt wurde den sFv-Fraktionen der Affinitätsreinigung und HRP-konjugierten Fab-Fragmenten von 520C9 erlaubt, um die Bindung an Membranen von SK-BR-3 zu konkurrieren. Eine erfolgreiche Bindung der sFv-Präparation hinderte das HRP-520C9-Fab-Fragment an der Bindung auf die Membranen und verringerte oder verhinderte somit auch die Verwendung des HRP-Substrates und so die Farbtwicklung (siehe nachstehend für Details des Konkurrenzversuchs). Die Ergebnisse zeigten, dass faktisch die gesamte sFv-Aktivität durch die Säule gebunden und im eluierten Peak gewonnen wurde ([Fig. 4](#)). Wie erwartet wurde die spezifische Aktivität des eluierten Peaks im Vergleich zur Säulen-Probe gesteigert und schien im Wesentlichen, innerhalb des experimentellen Fehlers dieser Messungen, gleich mit der Kontrolle des parentalen Fab zu sein.

#### 9. Ausbeute nach der Aufreinigung

**[0095]** Die Tabelle 1 zeigt die Ausbeute von verschiedenen Präparationen von 520C9 während des Aufreinigungsprozesses. Die Protein-Konzentration ( $\mu\text{g}/\text{ml}$ ) wurde mit dem BioRad-Proteintest bestimmt. Unter "Ge-



samte Ausbeute" stellen 300 AU des denaturierten sFv-Vorrats 3,15 g Einschlusskörperchen aus 0,4 Litern aus der Fermentation dar. Der Oxidationspuffer war 25 mM Tris, 10 mM EDTA, 6 M GdnHCl, 1 mM GSSG, 0,1 mM GSH, pH 9,0. Die Oxidation wurde bei Raumtemperatur über Nacht durchgeführt. Die oxidierte Probe wurde gegen 10 mM Natriumphosphat, 1 mM EDTA, 150 mM NaCl, 500 mM Harnstoff, pH 8,0 dialysiert. Alle nachfolgenden Schritte wurden, mit Ausnahme der Affinitätschromatographie, welche in PBSA durchgeführt wurde, in diesem Puffer durchgeführt.

Tabelle 1

<u>Probe</u>	<u>Volumen</u>	<u>Proteinkonzentration</u>	<u>Gesamte Ausbeute</u>	<u>% Ausbeute</u>
1. Rückfaltung III (Oxidierung)	4000 ml	0,075 A <sub>280</sub>	300 AU	-
2. Dialysierte Rück- faltung III	5050 ml	38 µg/ml	191,9 mg	100
3. Minitan-100K-Filtrat	5000 ml	2 µg/ml	10,0 mg	5,4
4. Minitan-10K-Rück- stand	100 ml	45 µg/ml	4,5 mg	2,3
6. YM10-10K-Rück- stand	5,2 ml	600 µg/ml	3,1 mg	1,6
7. S-200-sFv-Mono- merpeak	10,0 ml	58 µg/ml	0,58 mg	0,3
8. Affinitätsgereinigtes sFv	5,5 ml	13 µg/ml	0,07 mg	0,04

## 10. Zusammenbau eines Immuntoxins

**[0096]** Das das fusionierte Ricin A-520C9-Einzelketten-Immuntoxin (SEQ. ID NO: 7) codierende Gen wurde durch die Isolierung des das Ricin A codierenden Gens auf einem HindIII- bis BamHI-Fragment aus pPL229 (Cetus Corporation, Emeryville, CA) und dessen Verwendung stromaufwärts des sFv 520C9 in pH777, wie in [Fig. 3](#) gezeigt, zusammengesetzt. Diese Fusion enthält den natürlichen Linker mit 122 Aminosäuren, der zwischen den A- und B-Domänen des Ricins vorliegt. Im originalen Expressionsvektor pRAP229 war jedoch das Codon für Aminosäure 268 des Ricins zu einem TAA-Stoppocodon der Translation umgewandelt worden, so dass die Expression des sich ergebenden Gens nur Ricin A ergibt. Deshalb wurde, um das Stoppocodon der Translation zu entfernen, eine zielgerichtete Mutation durchgeführt, um das TAA zu entfernen und das natürliche Serin-Codon wiederherzustellen. Dieses erlaubt dann der Translation, sich über das gesamte Immuntoxin-Gen fortzusetzen.

**[0097]** Um das Immuntoxin zurück in die Expressionsvektoren pPL229 und pRAP229 einzusetzen, musste die PstI-Stelle am Ende des Immuntoxin-Gens zu einer Sequenz umgewandelt werden, die mit der BamHI-Stelle im Vektor kompatibel war. Ein synthetischer Oligonucleotid-Adapter, welcher eine zwischen PstI-Enden platzierte BclI-Stelle enthielt, wurde eingefügt. BclI- und BamHI-Enden sind kompatibel und können zu einer hybriden BclI/BamHI-Stelle kombiniert werden. Da die BclI-Nulcease empfindlich gegenüber einer dam-Methylierung ist, wurde das Konstrukt zuerst in einen dam (-) E. coli-Stamm, Gm48, transformiert, um die Plasmid-DNA mit BclI (und HindIII) zu spalten, dann das gesamte Immuntoxin-Gen auf einem HindIII/BclI-Fragment zurück in beide mit HindIII/BamHI gespaltene Expressionsvektoren einzufügen.

**[0098]** Wenn das native 520C9-IgG1 mit der nativen Ricin A-Kette oder der rekombinanten Ricin A-Kette konjugiert wird, ist das sich ergebende Immuntoxin fähig, die Proteinsynthese bei einer Konzentration von etwa  $0,4 \times 10^{-9}$  M bei SK-Br-3-Zellen um 50 % zu hemmen. Zusätzlich zur Reaktivität mit SK-Br-3-Brustkrebszellen

hemmt natives 520C9-IgG1-Immuntoxin auch eine Eierstockkrebszelllinie, OVCAR-3, mit einer  $ID_{50}$  von  $2,0 \times 10^{-9}$  M.

**[0099]** Im vorstehend beschriebenen Ricin A-sFv-Fusionsprotein wirkt Ricin als Leader für die Expression, das heißt, es ist an den Amino-Terminus des sFv fusioniert. Der direkten Expression folgend wurde für das lösliche Protein gezeigt, dass es mit Antikörpern gegen natives Fab 520C9 reagiert und auch die enzymatische Aktivität der Ricin A-Kette zeigt.

**[0100]** In einem anderen Entwurf ist die Ricin A-Kette an den Carboxy-Terminus des sFv fusioniert. Das sFv 520C9 kann mittels der Signalsequenz PelB sezerniert werden, wobei die Ricin A-Kette mit dem C-Terminus des sFv verbunden ist. Für dieses Konstrukt werden Sequenzen, welche die Signalsequenz PelB, sFv und Ricin codieren, in einem Bluescript-Plasmid über eine HindIII-Stelle, die dem sFv direkt folgt (in unseren Expressionsplasmiden), und die dem Ricin-Gen vorausgehende HindIII-Stelle in einer Zusammensetzung dreier Teile (RI-HindIII-BamHI) verbunden. Eine neue PstI-Stelle, die dem Ricin-Gen folgt, wird über den Polylinker des Bluescript gewonnen. Die Mutation dieser DNA entfernt das Stoppcodon und die ursprüngliche PstI-Stelle am Ende des sFv und setzt einige Serin-Reste zwischen die Gene des sFv und des Ricins. Diese neue Fusion von Genen, Signalsequenz PelB/sFv/Ricin A, kann als EcoRI/PstI-Fragment in Expressionsvektoren eingefügt werden.

**[0101]** In einem anderen Entwurf wird das zu der Ricin A-Kette analoge Exotoxin aus Pseudomonas, PE40, mit dem Carboxy-Terminus des anti-c-erbB-2-741F8-sFv fusioniert (SEQ ID NO: 15 und 16). Das sich ergebende 741F8-sFv-PE40 ist ein Einzelketten-Fv-Toxin-Fusionsprotein, welches mit einem 18 Reste kurzen FB-Leader zusammengesetzt wurde, der am Anfang an dem Protein belassen wurde.

**[0102]** Die Expression dieses Proteins in E. coli ergab Einschlusskörperchen, die in einem 3 M Harnstoff-Glutathion/Redox-Puffer rückgefaltet wurden. Von dem sich ergebenden sFv-PE40 wurde gezeigt, dass es vollständiger und mit offensichtlich besserer Cytotoxizität als das entsprechende quervernetzte Immuntoxin spezifisch c-erbB-2 tragende Zellen in Kultur abtötet. Das sFv-Toxin-Protein ebenso wie das sFv 741 F8 können mit guten Ausbeuten durch diese Verfahren hergestellt werden und können als therapeutische oder diagnostische Wirkstoffe für Tumoren, die das c-erbB-2- oder damit verwandte Antigene tragen, wie Brust- und Eierstockkrebs, verwendet werden.

## 11. Experimente

### A. Kompetitions-ELISA

**[0103]** Extrakt aus SK-Br-3 wird als Quelle für das c-erbB-2-Antigen wie folgt präpariert. Die Brustkrebszellen SK-Br-3 (Ring et al., Cancer Research 49 (1989), 3070-3080) werden bis fast zur Konfluenz in Iscove-Medium (Gibco BRL, Gaithersburg, Md.) mit 5 % fötalem Rinderserum und 2 mM Glutamin gezüchtet. Das Medium wird abgesaugt und die Zellen werden mit 10 ml fötalem Rinderserum (FBS) mit Calcium und Magnesium gespült. Die Zellen werden mit einem Gummischaber in 10 ml FBS mit Calcium und Magnesium abgeschabt und die Flasche wird mit weiteren 5 ml dieses Puffers ausgespült. Die Zellen werden dann bei 100 UpM zentrifugiert. Der Überstand wird abgesaugt und die Zellen werden zu  $10^7$  Zellen/ml in 10 mM NaCl, 0,5 % NP40, pH 8 (TNN-Puffer) resuspendiert und auf- und abpipettiert, um das Sediment aufzulösen. Die Lösung wird dann bei 1000 UpM zentrifugiert, um Zellkerne und andere unlösliche Trümmer zu entfernen. Der Extrakt wird durch 0,45 Millex HA- und 0,2 Millex Gv-Filter filtriert. Der TNN-Extrakt wird in Aliquots in Wheaton-Einfrierröhrchen bei  $-70^{\circ}\text{C}$  gelagert.

**[0104]** Ein frisches Röhrchen des TNN-Extrakts von SK-Br-3 wird aufgetaut und 200-fach in entionisiertem Wasser verdünnt. Unmittelbar danach werden 40  $\mu\text{g}$  pro Vertiefung in eine PVC-Platte von Dynatech mit 96 Vertiefungen gegeben, welche über Nacht in einem Trockeninkubator bei  $37^{\circ}\text{C}$  belassen wird. Die Platten werden viermal mit phosphat-gepufferter Kochsalzlösung (PBS), 1 % Magermilch, 0,05 % Tween 20 gewaschen.

**[0105]** Die unspezifischen Bindungsstellen werden wie folgt blockiert. Wenn die Platte trocken ist, wird 100  $\mu\text{g}$  PBS, das 1 % Magermilch enthält, pro Vertiefung zugefügt und die Inkubation wird für eine 1 Stunde bei Raumtemperatur laufen gelassen.

**[0106]** Die Testproben des Einzelketten-Fv und Standardverdünnungen des ganzen Antikörpers 520C9 werden dann wie folgt zugefügt. Der Antikörper 520C9 und die Testproben werden in einem Verdünnungspuffer (PBS + 1 % Magermilch) in Serie mit anfangs 50  $\mu\text{g}/\text{ml}$  mit 2-fach Schritten verdünnt, wobei mindestens 10

Verdünnungen für den Standard 520C9 gemacht werden. Eine Kontrolle, die nur Verdünnungspuffer enthält, wird eingeschlossen. Die verdünnten Proben und Standards werden mit 50 µl pro Vertiefung zugefügt und über 30 Minuten bei Raumtemperatur inkubiert.

**[0107]** Die 520C9-Meerrettich-Peroxidase (HRP)-Sonde wird wie folgt zugefügt. Das 520C9-HRP-Konjugat (Zymed Labs., Süd-San Francisco, Kalifornien) wird auf 14 µg/ml mit 1 % Magermilch im Verdünnungspuffer verdünnt. Die optimalen Verdünnungen müssen für jede neue Charge des Peroxidase-Konjugates bestimmt werden, ohne die vorangehenden Schritte wegzulassen. Pro Vertiefung wurden 20 µl der Sonde zugefügt und über 1 Stunde bei Raumtemperatur inkubiert. Die Platte wird dann viermal mit PBS gewaschen. Das Peroxidase-Substrat wird dann zugefügt. Die Substratlösung sollte für jede Anwendung durch eine Verdünnung der Tetramethylbenzidin-Vorratslösung (TMB; 2 mg/ml in 100 % Ethanol) 1 : 20 und der 3%igen Wasserstoffperoxid-Vorratslösung 1:2200 in Substratpuffer (10 mM Natriumacetat, 10 mM Na-EDTA, pH 5,0) frisch hergestellt werden. Dies wird über 30 Minuten bei Raumtemperatur inkubiert. Die Vertiefungen werden dann mit 100 µl 0,8 M H<sub>2</sub>SO<sub>4</sub> pro Vertiefung gestoppt und die Absorption bei 150 nm abgelesen.

**[0108]** [Fig. 4](#) vergleicht die Bindungsfähigkeit des parentalen, rückgefalteten, aber nicht aufgereinigten monoclonalen Antikörpers 520C9, der Fab-Fragmente von 520C9 und der sFv-Einzelketten-Bindungsstelle von 520C9 nach Bindung und Elution von einer Affinitätsäule (eluiert) oder der ungebundenen Durchlaufraction (durchgelaufen). In [Fig. 4](#) zeigt der vollständig aufgereinigte sFv 520C9 eine Affinität für c-erbB-2, die von dem parentalen monoclonalen Antikörper, innerhalb des Fehlers der Messung der Proteinkonzentration, nicht zu unterscheiden ist.

#### B. Testung in vivo

**[0109]** Immunttoxine, die starke Hemmstoffe der Proteinsynthese bei in Kultur gezüchteten Brustkrebszellen sind, können auf ihre Wirksamkeit in vivo getestet werden. Das Experiment in vivo wird typischerweise in einem Nacktmaus-Modell unter Verwendung von Xeno-Transplantaten menschlicher MX-1-Brustkrebszellen durchgeführt. Den Mäusen werden entweder PBS (Kontrolle) oder verschiedene Konzentrationen des sFv-Toxin-Immuntoxins injiziert und eine konzentrationsabhängige Hemmung des Tumorwachstums wird beobachtet werden. Es wird erwartet, dass höhere Dosen des Immuntoxins eine bessere Wirkung hervorbringen werden.

**[0110]** Die Erfindung kann in anderen spezifischen Formen ausgeführt werden, ohne deren Geist oder Umfang zu verlassen. Die vorliegenden Ausführungsformen sind deshalb in allen Gesichtspunkten als veranschaulichend und nicht als beschränkend zu verstehen, wobei der Umfang der Erfindung vielmehr durch die anhängenden Ansprüche als durch die vorausgegangene Beschreibung aufgezeigt wird, und es ist beabsichtigt, dass alle Änderungen, die innerhalb der Äquivalenzbedeutung und des Äquivalenzumfangs der Ansprüche aufkommen, darin umfasst sein sollen.

SEQUENZPROTOKOLL

(1) Allgemeine Information

(i) Anmelder

(A) NAME: Creative BioMolecules, Inc.

(B) STRASSE: 35 South Street

(C) STADT: Hopkinton

(D) STAAT: Massachusetts

(E) LAND: Vereinigte Staaten

(F) POSTLEITZAHL (PLZ): 01748

(G) TELEFON: 1-508-435-9001

(H) TELEFAX: 1-508-435-0454

(I) TELEX:

(A) NAME: Cetus Corporation

(B) STRASSE: 1400 Fifty-Third Street

(C) STADT: Emeryville

(D) STAAT: Kalifornien

(E) LAND: Vereinigte Staaten

(F) POSTLEITZAHL (PLZ): 94608

(G) TELEFON:

(H) TELEFAX:

(I) TELEX:

(ii) TITEL DER ERFINDUNG: Marker für Krebs und biosynthetisches Bindeprotein dafür

(iii) ZAHL DER SEQUENZEN: 14

(iv) KORRESPONDENZADRESSE:

(A) ADRESSAT: Creative BioMolecules, Inc.

(B) STRASSE: 35 South Street

(C) STADT: Hopkinton

(D) STAAT: Massachusetts

(E) LAND: Vereinigte Staaten

(F) PLZ: 01748

(v) IM COMPUTER LESBARE FORM:

(A) MEDIUMTYP: Diskette

(B) COMPUTER: IMB-PC kompatibel

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

(D) SOFTWARE: PatentIn Ausgabe #1,0, Version #1,25

(vi) DATEN DER AKTUELLEN ANMELDUNG:

(A) ANMELDUNGSNUMMER: PCT/US 93/01055

(B) EINREICHUNGSDATUM: 5. Februar 1993

(C) KLASSIFIZIERUNG:

(viii) ANWALT/VERTRETERINFORMATION:

(A) NAME: Pitcher, Edmund R.

(B) REGISTRIERUNGSNUMMER: 27829

(C) REFERENZ/AKTENNUMMER: 2054/22

(ix) TELEKOMMUNIKATIONSINFORMATION:

(A) TELEFON: (617) 248-7000

(B) TELEFAX: (617) 248-7100

2) INFORMATION FÜR SEQ ID NO: 1:

(i) SEQUENZCHARAKTERISTIK:

(A) LÄNGE: 4299 Basenpaare

(B) ART: Nucleinsäure

(C) STRANGFORM: einzel

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: cDNA

(ix) EIGENSCHAFT:

(A) NAME / SCHLÜSSEL: CDS

(B) POSITION: 1 ... 4299

(D) ANDERE INFORMATION: /Anmerkung "Produkt = "c-erb-b-2""

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

ATG	GAG	CTG	GCG	GCC	TTG	TGC	CGC	TGG	GGG	CTC	CTC	CTC	GCC	CTC	TTG	48
Met	Glu	Leu	Ala	Ala	Leu	Cys	Arg	Trp	Gly	Leu	Leu	Leu	Ala	Leu	Leu	
1				5					10					15		
CCC	CCC	GGA	GCC	GCG	AGC	ACC	CAA	GTG	TGC	ACC	GGC	ACA	GAC	ATG	AAG	96
Pro	Pro	Gly	Ala	Ala	Ser	Thr	Gln	Val	Cys	Thr	Gly	Thr	Asp	Met	Lys	
			20					25					30			
CTG	CGG	CTC	CCT	GCC	AGT	CCC	GAG	ACC	CAC	CTG	GAC	ATG	CTC	CGC	CAC	144
Leu	Arg	Leu	Pro	Ala	Ser	Pro	Glu	Thr	His	Leu	Asp	Met	Leu	Arg	His	
		35					40					45				
CTC	TAC	CAG	GGC	TGC	CAG	GTG	GTG	CAG	GGA	AAC	CTG	GAA	CTC	ACC	TAC	192
Leu	Tyr	Gln	Gly	Cys	Gln	Val	Val	Gln	Gly	Asn	Leu	Glu	Leu	Thr	Tyr	
	50					55					60					
CTG	CCC	ACC	AAT	GCC	AGC	CTG	TCC	TTC	CTG	CAG	GAT	ATC	CAG	GAG	GTG	240
Leu	Pro	Thr	Asn	Ala	Ser	Leu	Ser	Phe	Leu	Gln	Asp	Ile	Gln	Glu	Val	
65					70					75					80	
CAG	GGC	TAC	GTG	CTC	ATC	GCT	CAC	AAC	CAA	GTG	AGG	CAG	GTC	CCA	CTG	288
Gln	Gly	Tyr	Val	Leu	Ile	Ala	His	Asn	Gln	Val	Arg	Gln	Val	Pro	Leu	
			85						90					95		
CAG	AGG	CTG	CGG	ATT	GTG	CGA	GGC	ACC	CAG	CTC	TTT	GAG	GAC	AAC	TAT	336
Gln	Arg	Leu	Arg	Ile	Val	Arg	Gly	Thr	Gln	Leu	Phe	Glu	Asp	Asn	Tyr	
			100					105					110			
GCC	CTG	GCC	GTG	CTA	GAC	AAT	GGA	GAC	CCG	CTG	AAC	AAT	ACC	ACC	CCT	384
Ala	Leu	Ala	Val	Leu	Asp	Asn	Gly	Asp	Pro	Leu	Asn	Asn	Thr	Thr	Pro	
		115					120					125				
GTC	ACA	GGG	GCC	TCC	CCA	GGA	GGC	CTG	CGG	GAG	CTG	CAG	CTT	CGA	AGC	432
Val	Thr	Gly	Ala	Ser	Pro	Gly	Gly	Leu	Arg	Glu	Leu	Gln	Leu	Arg	Ser	
	130					135					140					
CTC	ACA	GAG	ATC	TTG	AAA	GGA	GGG	GTC	TTG	ATC	CAG	CGG	AAC	CCC	CAG	480
Leu	Thr	Glu	Ile	Leu	Lys	Gly	Gly	Val	Leu	Ile	Gln	Arg	Asn	Pro	Gln	
145					150				155						160	
CTC	TGC	TAC	CAG	GAC	ACG	ATT	TTG	TGG	AAG	GAC	ATC	TTC	CAC	AAG	AAC	528
Leu	Cys	Tyr	Gln	Asp	Thr	Ile	Leu	Trp	Lys	Asp	Ile	Phe	His	Lys	Asn	
			165					170					175			
AAC	CAG	CTG	GCT	CTC	ACA	CTG	ATA	GAC	ACC	AAC	CGC	TCT	CGG	GCC	TGC	576
Asn	Gln	Leu	Ala	Leu	Thr	Leu	Ile	Asp	Thr	Asn	Arg	Ser	Arg	Ala	Cys	
			180					185					190			

CAC	CCC	TGT	TCT	CCG	ATG	TGT	AAG	GGC	TCC	CGC	TGC	TGG	GGA	GAG	AGT	624
His	Pro	Cys	Ser	Pro	Met	Cys	Lys	Gly	Ser	Arg	Cys	Trp	Gly	Glu	Ser	
		195					200					205				
TCT	GAG	GAT	TGT	CAG	AGC	CTG	ACG	CGC	ACT	GTC	TGT	GCC	GGT	GGC	TGT	672
Ser	Glu	Asp	Cys	Gln	Ser	Leu	Thr	Arg	Thr	Val	Cys	Ala	Gly	Gly	Cys	
	210					215					220					
GCC	CGC	TGC	AAG	GGG	CCA	CTG	CCC	ACT	GAC	TGC	TGC	CAT	GAG	CAG	TGT	720
Ala	Arg	Cys	Lys	Gly	Pro	Leu	Pro	Thr	Asp	Cys	Cys	His	Glu	Gln	Cys	
225					230					235					240	
GCT	GCC	GGC	TGC	ACG	GGC	CCC	AAG	CAC	TCT	GAC	TGC	CTG	GCC	TGC	CTC	768
Ala	Ala	Gly	Cys	Thr	Gly	Pro	Lys	His	Ser	Asp	Cys	Leu	Ala	Cys	Leu	
				245					250					255		
CAC	TTC	AAC	CAC	AGT	GGC	ATC	TGT	GAG	CTG	CAC	TGC	CCA	GCC	CTG	GTC	816
His	Phe	Asn	His	Ser	Gly	Ile	Cys	Glu	Leu	His	Cys	Pro	Ala	Leu	Val	
			260					265					270			
ACC	TAC	AAC	ACA	GAC	ACG	TTT	GAG	TCC	ATG	CCC	AAT	CCC	GAG	GGC	CGG	864
Thr	Tyr	Asn	Thr	Asp	Thr	Phe	Glu	Ser	Met	Pro	Asn	Pro	Glu	Gly	Arg	
		275					280					285				
TAT	ACA	TTC	GGC	GCC	AGC	TGT	GTG	ACT	GCC	TGT	CCC	TAC	AAC	TAC	CTT	912
Tyr	Thr	Phe	Gly	Ala	Ser	Cys	Val	Thr	Ala	Cys	Pro	Tyr	Asn	Tyr	Leu	
	290					295					300					
TCT	ACG	GAC	GTG	GGA	TCC	TGC	ACC	CTC	GTC	TGC	CCC	CTG	CAC	AAC	CAA	960
Ser	Thr	Asp	Val	Gly	Ser	Cys	Thr	Leu	Val	Cys	Pro	Leu	His	Asn	Gln	
305					310					315					320	
GAG	GTG	ACA	GCA	GAG	GAT	GGA	ACA	CAG	CGG	TGT	GAG	AAG	TGC	AGC	AAG	1008
Glu	Val	Thr	Ala	Glu	Asp	Gly	Thr	Gln	Arg	Cys	Glu	Lys	Cys	Ser	Lys	
				325					330					335		
CCC	TGT	GCC	CGA	GTG	TGC	TAT	GGT	CTG	GGC	ATG	GAG	CAC	TTG	CGA	GAG	1056
Pro	Cys	Ala	Arg	Val	Cys	Tyr	Gly	Leu	Gly	Met	Glu	His	Leu	Arg	Glu	
			340					345					350			
GTG	AGG	GCA	GTT	ACC	AGT	GCC	AAT	ATC	CAG	GAG	TTT	GCT	GGC	TGC	AAG	1104
Val	Arg	Ala	Val	Thr	Ser	Ala	Asn	Ile	Gln	Glu	Phe	Ala	Gly	Cys	Lys	
		355					360					365				
AAG	ATC	TTT	GGG	AGC	CTG	GCA	TTT	CTG	CCG	GAG	AGC	TTT	GAT	GGG	GAC	1152
Lys	Ile	Phe	Gly	Ser	Leu	Ala	Phe	Leu	Pro	Glu	Ser	Phe	Asp	Gly	Asp	
	370					375					380					
CCA	GCC	TCC	AAC	ACT	GCC	CCG	CTC	CAG	CCA	GAG	CAG	CTC	CAA	GTG	TTT	1200
Pro	Ala	Ser	Asn	Thr	Ala	Pro	Leu	Gln	Pro	Glu	Gln	Leu	Gln	Val	Phe	
385					390					395				400		
GAG	ACT	CTG	GAA	GAG	ATC	ACA	GGT	TAC	CTA	TAC	ATC	TCA	GCA	TGG	CCG	1248
Glu	Thr	Leu	Glu	Glu	Ile	Thr	Gly	Tyr	Leu	Tyr	Ile	Ser	Ala	Trp	Pro	
				405					410					415		

GAC Asp	AGC Ser	CTG Leu	CCT Pro	GAC Asp	CTC Leu	AGC Ser	GTC Val	TTC Phe	CAG Gln	AAC Asn	CTG Leu	CAA Gln	GTA Val	ATC Ile	CGG Arg	1296
			420				425						430			
GGA Gly	CGA Arg	ATT Ile	CTG Leu	CAC His	AAT Asn	GGC Gly	GCC Ala	TAC Tyr	TCG Ser	CTG Leu	ACC Thr	CTG Leu	CAA Gln	GGG Gly	CTG Leu	1344
			435				440						445			
GGC Gly	ATC Ile	AGC Ser	TGG Trp	CTG Leu	GGG Gly	CTG Leu	CGC Arg	TCA Ser	CTG Leu	AGG Arg	GAA Glu	CTG Leu	GGC Gly	AGT Ser	GGA Gly	1392
			450				455				460					
CTG Leu	GCC Ala	CTC Leu	ATC Ile	CAC His	CAT His	AAC Asn	ACC Thr	CAC His	CTC Leu	TGC Cys	TTC Phe	GTG Val	CAC His	ACG Thr	GTG Val	1440
			465			470				475					480	
CCC Pro	TGG Trp	GAC Asp	CAG Gln	CTC Leu	TTT Phe	CGG Arg	AAC Asn	CCG Pro	CAC His	CAA Gln	GCT Ala	CTG Leu	CTC Leu	CAC His	ACT Thr	1488
				485					490					495		
GCC Ala	AAC Asn	CGG Arg	CCA Pro	GAG Glu	GAC Asp	GAG Glu	TGT Cys	GTG Val	GGC Gly	GAG Glu	GGC Gly	CTG Leu	GCC Ala	TGC Cys	CAC His	1536
			500				505						510			
CAG Gln	CTG Leu	TGC Cys	GCC Ala	CGA Arg	GGG Gly	CAC His	TGC Cys	TGG Trp	GGT Gly	CCA Pro	GGG Gly	CCC Pro	ACC Thr	CAG Gln	TGT Cys	1584
			515				520					525				
GTC Val	AAC Asn	TGC Cys	AGC Ser	CAG Gln	TTC Phe	CTT Leu	CGG Arg	GGC Gly	CAG Gln	GAG Glu	TGC Cys	GTG Val	GAG Glu	GAA Glu	TGC Cys	1632
			530			535					540					
CGA Arg	GTA Val	CTG Leu	CAG Gln	GGG Gly	CTC Leu	CCC Pro	AGG Arg	GAG Glu	TAT Tyr	GTG Val	AAT Asn	GCC Ala	AGG Arg	CAC His	TGT Cys	1680
					550					555					560	
TTG Leu	CCG Pro	TGC Cys	CAC His	CCT Pro	GAG Glu	TGT Cys	CAG Gln	CCC Pro	CAG Gln	AAT Asn	GGC Gly	TCA Ser	GTG Val	ACC Thr	TGT Cys	1728
				565					570					575		
TTT Phe	GGA Gly	CCG Pro	GAG Glu	GCT Ala	GAC Asp	CAG Gln	TGT Cys	GTG Val	GCC Ala	TGT Cys	GCC Ala	CAC His	TAT Tyr	AAG Lys	GAC Asp	1776
				580				585					590			
CCT Pro	CCC Pro	TTC Phe	TGC Cys	GTG Val	GCC Ala	CGC Arg	TGC Cys	CCC Pro	AGC Ser	GGT Gly	GTG Val	AAA Lys	CCT Pro	GAC Asp	CTC Leu	1824
			595				600					605				
TCC Ser	TAC Tyr	ATG Met	CCC Pro	ATC Ile	TGG Trp	AAG Lys	TTT Phe	CCA Pro	GAT Asp	GAG Glu	GAG Glu	GGC Gly	GCA Ala	TGC Cys	CAG Gln	1872
			610			615					620					
CCT Pro	TGC Cys	CCC Pro	ATC Ile	AAC Asn	TGC Cys	ACC Thr	CAC His	TCC Ser	TGT Cys	GTG Val	GAC Asp	CTG Leu	GAT Asp	GAC Asp	AAG Lys	1920
			625			630				635					640	



GGC TGC CCC GCC GAG CAG AGA GCC AGC CCT CTG ACG TCC ATC ATC TCT Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Ile Ser 645 650 655	1968
GCG GTG GTT GGC ATT CTG CTG GTC GTG GTC TTG GGG GTG GTC TTT GGG Ala Val Val Gly Ile Leu Leu Val Val Leu Gly Val Val Phe Gly 660 665 670	2016
ATC CTC ATC AAG CGA CGG CAG CAG AAG ATC CGG AAG TAC ACG ATG CGG Ile Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg 675 680 685	2064
AGA CTG CTG CAG GAA ACG GAG CTG GTG GAG CCG CTG ACA CCT AGC GGA Arg Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly 690 695 700	2112
GCG ATG CCC AAC CAG GCG CAG ATG CGG ATC CTG AAA GAG ACG GAG CTG Ala Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu 705 710 715 720	2160
AGG AAG GTG AAG GTG CTT GGA TCT GGC GCT TTT GGC ACA GTC TAC AAG Arg Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys 725 730 735	2208
GGC ATC TGG ATC CCT GAT GGG GAG AAT GTG AAA ATT CCA GTG GCC ATC Gly Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile 740 745 750	2256
AAA GTG TTG AGG GAA AAC ACA TCC CCC AAA GCC AAC AAA GAA ATC TTA Lys Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu 755 760 765	2304
GAC GAA GCA TAC GTG ATG GCT GGT GTG GGC TCC CCA TAT GTC TCC CGC Asp Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg 770 775 780	2352
CTT CTG GGC ATC TGC CTG ACA TCC ACG GTG CAG CTG GTG ACA CAG CTT Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu 785 790 795 800	2400
ATG CCC TAT GGC TGC CTC TTA GAC CAT GTC CGG GAA AAC CGC GGA CGC Met Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg 805 810 815	2448
CTG GGC TCC CAG GAC CTG CTG AAC TGG TGT ATG CAG ATT GCC AAG GGG Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly 820 825 830	2496
ATG AGC TAC CTG GAG GAT GTG CGG CTC GTA CAC AGG GAC TTG GCC GCT Met Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala 835 840 845	2544
CGG AAC GTG CTG GTC AAG AGT CCC AAC CAT GTC AAA ATT ACA GAC TTC Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe 850 855 860	2592

GGG Gly 865	CTG Leu	GCT Ala	CGG Arg	CTG Leu	CTG Leu	GAC Asp	ATT Ile	GAC Asp	GAG Glu	ACA Thr	GAG Glu	TAC Tyr	CAT His	GCA Ala	GAT Asp	2640
870									875						880	
GGG Gly	GGC Gly	AAG Lys	GTG Val	CCC Pro	ATC Ile	AAG Lys	TGG Trp	ATG Met	GCC Ala	CTG Leu	GAG Glu	TCC Ser	ATT Ile	CTC Leu	CGC Arg	2688
885				885					890						895	
CGG Arg	CGG Arg	TTC Phe	ACC Thr	CAC His	CAG Gln	AGT Ser	GAT Asp	GTG Val	TGG Trp	AGT Ser	TAT Tyr	GGT Gly	GTG Val	ACT Thr	GTG Val	2736
			900					905					910			
TGG Trp	GAG Glu	CTG Leu	ATG Met	ACT Thr	TTT Phe	GGG Gly	GCC Ala	AAA Lys	CCT Pro	TAC Tyr	GAT Asp	GGG Gly	ATC Ile	CCA Pro	GCC Ala	2784
		915					920					925				
CGG Arg	GAG Glu	ATC Ile	CCT Pro	GAC Asp	CTG Leu	CTG Leu	GAA Glu	AAG Lys	GGG Gly	GAG Glu	CGG Arg	CTG Leu	CCC Pro	CAG Gln	CCC Pro	2832
	930						935				940					
CCC Pro	ATC Ile	TGC Cys	ACC Thr	ATT Ile	GAT Asp	GTC Val	TAC Tyr	ATG Met	ATC Ile	ATG Met	GTC Val	AAA Lys	TGT Cys	TGG Trp	ATG Met	2880
945					950				955						960	
ATT Ile	GAC Asp	TCT Ser	GAA Glu	TGT Cys	CGG Arg	CCA Pro	AGA Arg	TTC Phe	CGG Arg	GAG Glu	TTG Leu	GTG Val	TCT Ser	GAA Glu	TTC Phe	2928
				965					970						975	
TCC Ser	CGC Arg	ATG Met	GCC Ala	AGG Arg	GAC Asp	CCC Pro	CAG Gln	CGC Arg	TTT Phe	GTG Val	GTC Val	ATC Ile	CAG Gln	AAT Asn	GAG Glu	2976
			980					985					990			
GAC Asp	TTG Leu	GGC Gly	CCA Pro	GCC Ala	AGT Ser	CCC Pro	TTG Leu	GAC Asp	AGC Ser	ACC Thr	TTC Phe	TAC Tyr	CGC Arg	TCA Ser	CTG Leu	3024
		995					1000						1005			
CTG Leu	GAG Glu	GAC Asp	GAT Asp	GAC Asp	ATG Met	GGG Gly	GAC Asp	CTG Leu	GTG Val	GAT Asp	GCT Ala	GAG Glu	GAG Glu	TAT Tyr	CTG Leu	3072
	1010					1015					1020					
GTA Val	CCC Pro	CAG Gln	CAG Gln	GGC Gly	TTC Phe	TTC Phe	TGT Cys	CCA Pro	GAC Asp	CCT Pro	GCC Ala	CCG Pro	GGC Gly	GCT Ala	GGG Gly	3120
1025					1030					1035					1040	
GGC Gly	ATG Met	GTC Val	CAC His	CAC His	AGG Arg	CAC His	CGC Arg	AGC Ser	TCA Ser	TCT Ser	ACC Thr	AGG Arg	AGT Ser	GGC Gly	GGT Gly	3168
				1045					1050					1055		
GGG Gly	GAC Asp	CTG Leu	ACA Thr	CTA Leu	GGG Gly	CTG Leu	GAG Glu	CCC Pro	TCT Ser	GAA Glu	GAG Glu	GAG Glu	GCC Ala	CCC Pro	AGG Arg	3216
			1060					1065					1070			
TCT Ser	CCA Pro	CTG Leu	GCA Ala	CCC Pro	TCC Ser	GAA Glu	GGG Gly	GCT Ala	GGC Gly	TCC Ser	GAT Asp	GTA Val	TTT Phe	GAT Asp	GGT Gly	3264
		1075					1080					1085				

GAC CTG GGA ATG GGG GCA GCC AAG GGG CTG CAA AGC CTC CCC ACA CAT Asp Leu Gly Met Gly Ala Ala Lys Gly Leu Gln Ser Leu Pro Thr His 1090 1095 1100	3312
GAC CCC AGC CCT CTA CAG CGG TAC AGT GAG GAC CCC ACA GTA CCC CTG Asp Pro Ser Pro Leu Gln Arg Tyr Ser Glu Asp Pro Thr Val Pro Leu 1105 1110 1115 1120	3360
CCC TCT GAG ACT GAT GGC TAC GTT GCC CCC CTG ACC TGC AGC CCC CAG Pro Ser Glu Thr Asp Gly Tyr Val Ala Pro Leu Thr Cys Ser Pro Gln 1125 1130 1135	3408
CCT GAA TAT GTG AAC CAG CCA GAT GTT CGG CCC CAG CCC CCT TCG CCC Pro Glu Tyr Val Asn Gln Pro Asp Val Arg Pro Gln Pro Pro Ser Pro 1140 1145 1150	3456
CGA GAG GGC CCT CTG CCT GCT GCC CGA CCT GCT GGT GCC ACT CTG GAA Arg Glu Gly Pro Leu Pro Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu 1155 1160 1165	3504
AGG CCC AAG ACT CTC TCC CCA GGG AAG AAT GGG GTC GTC AAA GAC GTT Arg Pro Lys Thr Leu Ser Pro Gly Lys Asn Gly Val Val Lys Asp Val 1170 1175 1180	3552
TTT GCC TTT GGG GGT GCC GTG GAG AAC CCC GAG TAC TTG ACA CCC CAG Phe Ala Phe Gly Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln 1185 1190 1195 1200	3600
GGA GGA GCT GCC CCT CAG CCC CAC CCT CCT CCT GCC TTC AGC CCA GCC Gly Gly Ala Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala 1205 1210 1215	3648
TTC GAC AAC CTC TAT TAC TGG GAC CAG GAC CCA CCA GAG CGG GGG GCT Phe Asp Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala 1220 1225 1230	3696
CCA CCC AGC ACC TTC AAA GGG ACA CCT ACG GCA GAG AAC CCA GAG TAC Pro Pro Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr 1235 1240 1245	3744
CTG GGT CTG GAC GTG CCA GTG TGA ACC AGA AGG CCA AGT CCG CAG AAG Leu Gly Leu Asp Val Pro Val * Thr Arg Arg Pro Ser Pro Gln Lys 1250 1255 1260	3792
CCC TGA TGT GTC CTC AGG GAG CAG GGA AGG CCT GAC TTC TGC TGG CAT Pro * Cys Val Leu Arg Glu Gln Gly Arg Pro Asp Phe Cys Trp His 1265 1270 1275 1280	3840
CAA GAG GTG GGA GGG CCC TCC GAC CAC TTC CAG GGG AAC CTG CCA TGC Gln Glu Val Gly Gly Pro Ser Asp His Phe Gln Gly Asn Leu Pro Cys 1285 1290 1295	3888

CAG GAA CCT GTC CTA AGG AAC CTT CCT TCC TGC TTG AGT TCC CAG ATG Gln Glu Pro Val Leu Arg Asn Leu Pro Ser Cys Leu Ser Ser Gln Met 1300 1305 1310	3936
GCT GGA AGG GGT CCA GCC TCG TTG GAA GAG GAA CAG CAC TGG GGA GTC Ala Gly Arg Gly Pro Ala Ser Leu Glu Glu Glu Gln His Trp Gly Val 1315 1320 1325	3984
TTT GTG GAT TCT GAG GCC CTG CCC AAT GAG ACT CTA GGG TCC AGT GGA Phe Val Asp Ser Glu Ala Leu Pro Asn Glu Thr Leu Gly Ser Ser Gly 1330 1335 1340	4032
TGC CAC AGC CCA GCT TGG CCC TTT CCT TCC AGA TCC TGG GTA CTG AAA Cys His Ser Pro Ala Trp Pro Phe Pro Ser Arg Ser Trp Val Leu Lys 1345 1350 1355 1360	4080
GCC TTA GGG AAG CTG GCC TGA GAG GGG AAG CGG CCC TAA GGG AGT GTC Ala Leu Gly Lys Leu Ala * Glu Gly Lys Arg Pro * Gly Ser Val 1365 1370 1375	4128
TAA GAA CAA AAG CGA CCC ATT CAG AGA CTG TCC CTG AAA CCT AGT ACT * Glu Gln Lys Arg Pro Ile Gln Arg Leu Ser Leu Lys Pro Ser Thr 1380 1385 1390	4176
GCC CCC CAT GAG GAA GGA ACA GCA ATG GTG TCA GTA TCC AGG CTT TGT Ala Pro His Glu Glu Gly Thr Ala Met Val Ser Val Ser Arg Leu Cys 1395 1400 1405	4224
ACA GAG TGC TTT TCT GTT TAG TTT TTA CTT TTT TTG TTT TGT TTT TTT Thr Glu Cys Phe Ser Val * Phe Leu Leu Phe Leu Phe Cys Phe Phe 1410 1415 1420	4272
AAA GAT GAA ATA AAG ACC CAG GGG GAG Lys Asp Glu Ile Lys Thr Gln Gly Glu 1425 1430	4299

## (2) INFORMATION FÜR SEQ ID NO: 2:

- (i) SEQUENZCHARAKTERISTIK:  
 (A) LÄNGE: 1433 Aminosäuren  
 (B) ART: Aminosäure  
 (D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

Met	Glu	Leu	Ala	Ala	Leu	Cys	Arg	Trp	Gly	Leu	Leu	Leu	Ala	Leu	Leu
1				5					10					15	
Pro	Pro	Gly	Ala	Ala	Ser	Thr	Gln	Val	Cys	Thr	Gly	Thr	Asp	Met	Lys
			20					25						30	

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His  
 35 40 45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr  
 50 55 60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val  
 65 70 75 80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu  
 85 90 95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr  
 100 105 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro  
 115 120 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser  
 130 135 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln  
 145 150 155 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn  
 165 170 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys  
 180 185 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser  
 195 200 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys  
 210 215 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys  
 225 230 235 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu  
 245 250 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val  
 260 265 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg  
 275 280 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu  
 290 295 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln  
 305 310 315 320

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

Ser Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln  
 610 615 620  
 Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys  
 625 630 635 640  
 Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Ile Ser  
 645 650 655  
 Ala Val Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly  
 660 665 670  
 Ile Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg  
 675 680 685  
 Arg Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly  
 690 695 700  
 Ala Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu  
 705 710 715 720  
 Arg Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys  
 725 730 735  
 Gly Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile  
 740 745 750  
 Lys Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu  
 755 760 765  
 Asp Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg  
 770 775 780  
 Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu  
 785 790 795 800  
 Met Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg  
 805 810 815  
 Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly  
 820 825 830  
 Met Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala  
 835 840 845  
 Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe  
 850 855 860  
 Gly Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp  
 865 870 875 880  
 Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg  
 885 890 895

Arg Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val  
 900 905 910  
 Trp Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala  
 915 920 925  
 Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro  
 930 935 940  
 Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met  
 945 950 955 960  
 Ile Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe  
 965 970 975  
 Ser Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu  
 980 985 990  
 Asp Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu  
 995 1000 1005  
 Leu Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr Leu  
 1010 1015 1020  
 Val Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly Ala Gly  
 1025 1030 1035 1040  
 Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg Ser Gly Gly  
 1045 1050 1055  
 Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu Ala Pro Arg  
 1060 1065 1070  
 Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser Asp Val Phe Asp Gly  
 1075 1080 1085  
 Asp Leu Gly Met Gly Ala Ala Lys Gly Leu Gln Ser Leu Pro Thr His  
 1090 1095 1100  
 Asp Pro Ser Pro Leu Gln Arg Tyr Ser Glu Asp Pro Thr Val Pro Leu  
 1105 1110 1115 1120  
 Pro Ser Glu Thr Asp Gly Tyr Val Ala Pro Leu Thr Cys Ser Pro Gln  
 1125 1130 1135  
 Pro Glu Tyr Val Asn Gln Pro Asp Val Arg Pro Gln Pro Pro Ser Pro  
 1140 1145 1150  
 Arg Glu Gly Pro Leu Pro Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu  
 1155 1160 1165  
 Arg Pro Lys Thr Leu Ser Pro Gly Lys Asn Gly Val Val Lys Asp Val  
 1170 1175 1180



Phe Ala Phe Gly Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln  
 1185 1190 1195 1200  
 Gly Gly Ala Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala  
 1205 1210 1215  
 Phe Asp Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala  
 1220 1225 1230  
 Pro Pro Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr  
 1235 1240 1245  
 Leu Gly Leu Asp Val Pro Val \* Thr Arg Arg Pro Ser Pro Gln Lys  
 1250 1255 1260  
 Pro \* Cys Val Leu Arg Glu Gln Gly Arg Pro Asp Phe Cys Trp His  
 1265 1270 1275 1280  
 Gln Glu Val Gly Gly Pro Ser Asp His Phe Gln Gly Asn Leu Pro Cys  
 1285 1290 1295  
 Gln Glu Pro Val Leu Arg Asn Leu Pro Ser Cys Leu Ser Ser Gln Met  
 1300 1305 1310  
 Ala Gly Arg Gly Pro Ala Ser Leu Glu Glu Glu Gln His Trp Gly Val  
 1315 1320 1325  
 Phe Val Asp Ser Glu Ala Leu Pro Asn Glu Thr Leu Gly Ser Ser Gly  
 1330 1335 1340  
 Cys His Ser Pro Ala Trp Pro Phe Pro Ser Arg Ser Trp Val Leu Lys  
 1345 1350 1355 1360  
 Ala Leu Gly Lys Leu Ala \* Glu Gly Lys Arg Pro \* Gly Ser Val  
 1365 1370 1375  
 \* Glu Gln Lys Arg Pro Ile Gln Arg Leu Ser Leu Lys Pro Ser Thr  
 1380 1385 1390  
 Ala Pro His Glu Glu Gly Thr Ala Met Val Ser Val Ser Arg Leu Cys  
 1395 1400 1405  
 Thr Glu Cys Phe Ser Val \* Phe Leu Leu Phe Leu Phe Cys Phe Phe  
 1410 1415 1420  
 Lys Asp Glu Ile Lys Thr Gln Gly Glu  
 1425 1430

## (2) INFORMATION FÜR SEQ ID NO: 3:

## (i) SEQUENZCHARAKTERISTIK:

- (A) LÄNGE: 739 Basenpaare
- (B) ART: Nucleinsäure
- (C) STRANGFORM: einzel

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: cDNA

(ix) EIGENSCHAFT:

(A) NAME/SCHLÜSSEL: CDS

(B) POSITION: 1 .. 739

(D) ANDERE INFORMATION: /Anmerkung= "Produkt =  
"520C9sFv/Aminosäureinformation: 520C9sFv-Protein"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

GAG ATC CAA TTG GTG CAG TCT GGA CCT GAG CTG AAG AAG CCT GGA GAG	48
Glu Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu	
1 5 10 15	
ACA GTC AAG ATC TCC TGC AAG GCT TCT GGA TAT ACC TTC GCA AAC TAT	96
Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Ala Asn Tyr	
20 25 30	
GGA ATG AAC TGG ATG AAG CAG GCT CCA GGA AAG GGT TTA AAG TGG ATG	144
Gly Met Asn Trp Met Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp Met	
35 40 45	
GGC TGG ATA AAC ACC TAC ACT GGA CAG TCA ACA TAT GCT GAT GAC TTC	192
Gly Trp Ile Asn Thr Tyr Thr Gly Gln Ser Thr Tyr Ala Asp Asp Phe	
50 55 60	
AAG GAA CGG TTT GCC TTC TCT TTG GAA ACC TCT GCC ACC ACT GCC CAT	240
Lys Glu Arg Phe Ala Phe Ser Leu Glu Thr Ser Ala Thr Thr Ala His	
65 70 75 80	
TTG CAG ATC AAC AAC CTC AGA AAT GAG GAC TCG GCC ACA TAT TTC TGT	288
Leu Gln Ile Asn Asn Leu Arg Asn Glu Asp Ser Ala Thr Tyr Phe Cys	
85 90 95	
GCA AGA CGA TTT GGG TTT GCT TAC TGG GGC CAA GGG ACT CTG GTC AGT	336
Ala Arg Arg Phe Gly Phe Ala Tyr Trp Gly Gln Gly Thr Leu Val Ser	
100 105 110	
GTC TCT GCA TCG ATA TCG AGC TCC TCC GGA TCT TCA TCT AGC GGT TCC	384
Val Ser Ala Ser Ile Ser Ser Ser Ser Gly Ser Ser Ser Ser Gly Ser	
115 120 125	
AGC TCG AGT GGA TCC GAT ATC CAG ATG ACC CAG TCT CCA TCC TCC TTA	432
Ser Ser Ser Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu	
130 135 140	
TCT GCC TCT CTG GGA GAA AGA GTC AGT CTC ACT TGT CGG GCA AGT CAG	480
Ser Ala Ser Leu Gly Glu Arg Val Ser Leu Thr Cys Arg Ala Ser Gln	
145 150 155 160	
GAC ATT GGT AAT AGC TTA ACC TGG CTT CAG CAG GAA CCA GAT GGA ACT	528
Asp Ile Gly Asn Ser Leu Thr Trp Leu Gln Gln Glu Pro Asp Gly Thr	
165 170 175	

ATT	AAA	CGC	CTG	ATC	TAC	GCC	ACA	TCC	AGT	TTA	GAT	TCT	GGT	GTC	CCC	576
Ile	Lys	Arg	Leu	Ile	Tyr	Ala	Thr	Ser	Ser	Leu	Asp	Ser	Gly	Val	Pro	
		180						185					190			
AAA	AGG	TTC	AGT	GGC	AGT	CGG	TCT	GGG	TCA	GAT	TAT	TCT	CTC	ACC	ATC	624
Lys	Arg	Phe	Ser	Gly	Ser	Arg	Ser	Gly	Ser	Asp	Tyr	Ser	Leu	Thr	Ile	
		195					200					205				
AGT	AGC	CTT	GAG	TCT	GAA	GAT	TTT	GTA	GTC	TAT	TAC	TGT	CTA	CAA	TAT	672
Ser	Ser	Leu	Glu	Ser	Glu	Asp	Phe	Val	Val	Tyr	Tyr	Cys	Leu	Gln	Tyr	
		210				215					220					
GCT	ATT	TTT	CCG	TAC	ACG	TTC	GGA	GGG	GGG	ACC	AAC	CTG	GAA	ATA	AAA	720
Ala	Ile	Phe	Pro	Tyr	Thr	Phe	Gly	Gly	Gly	Thr	Asn	Leu	Glu	Ile	Lys	
225					230					235					240	
CGG	GCT	GAT	TAA	TCT	GCA	G										739
Arg	Ala	Asp	*	Ser	Ala											
				245												

## (2) INFORMATION FÜR SEQ ID NO: 4:

## (i) SEQUENZCHARAKTERISTIK:

- (A) LÄNGE: 246 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: linear

## (ii) MOLEKÜLTYP: Protein

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

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

(2) INFORMATION FÜR SEQ ID NO: 5: gelöscht

(2) INFORMATION FÜR SEQ ID NO: 6: gelöscht

(2) INFORMATION FÜR SEQ ID NO: 7:

(i) SEQUENZCHARAKTERISTIK:

- (A) LÄNGE: 807 Basenpaare
- (B) ART: Nucleinsäure
- (C) STRANGFORM: einzel
- (D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (genomisch)

(ix) EIGENSCHAFT: -

- (A) NAME/SCHLÜSSEL: CDS
- (B) POSITION: 1 .. 807
- (D) ANDERE INFORMATION: /Anmerkung "Produkt = "Gen der Ricin A-Kette/ Aminosäureinformation: Protein der Ricin A-Kette""

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

ATG	ATA	TTC	CCC	AAA	CAA	TAC	CCA	ATT	ATA	AAC	TTT	ACC	ACA	GCG	GGT	48
Met	Ile	Phe	Pro	Lys	Gln	Tyr	Pro	Ile	Ile	Asn	Phe	Thr	Thr	Ala	Gly	
1				5				10						15		
GCC	ACT	GTG	CAA	AGC	TAC	ACA	AAC	TTT	ATC	AGA	GCT	GTT	CGC	GGT	CGT	96
Ala	Thr	Val	Gln	Ser	Tyr	Thr	Asn	Phe	Ile	Arg	Ala	Val	Arg	Gly	Arg	
			20					25						30		

TTA	ACA	ACT	GGA	GCT	GAT	GTG	AGA	CAT	GAA	ATA	CCA	GTG	TTG	CCA	AAC	144
Leu	Thr	Thr	Gly	Ala	Asp	Val	Arg	His	Glu	Ile	Pro	Val	Leu	Pro	Asn	
		35					40					45				
AGA	GTT	GGT	TTG	CCT	ATA	AAC	CAA	CGG	TTT	ATT	TTA	GTT	GAA	CTC	TCA	192
Arg	Val	Gly	Leu	Pro	Ile	Asn	Gln	Arg	Phe	Ile	Leu	Val	Glu	Leu	Ser	
	50					55					60					
AAT	CAT	GCA	GAG	CTT	TCT	GTT	ACA	TTA	GCG	CTG	GAT	GTC	ACC	AAT	GCA	240
Asn	His	Ala	Glu	Leu	Ser	Val	Thr	Leu	Ala	Leu	Asp	Val	Thr	Asn	Ala	
	65				70					75					80	
TAT	GTG	GTA	GGC	TAC	CGT	GCT	GGA	AAT	AGC	GCA	TAT	TTC	TTT	CAT	CCT	288
Tyr	Val	Val	Gly	Tyr	Arg	Ala	Gly	Asn	Ser	Ala	Tyr	Phe	Phe	His	Pro	
			85					90						95		
GAC	AAT	CAG	GAA	GAT	GCA	GAA	GCA	ATC	ACT	CAT	CTT	TTC	ACT	GAT	GTT	336
Asp	Asn	Gln	Glu	Asp	Ala	Glu	Ala	Ile	Thr	His	Leu	Phe	Thr	Asp	Val	
		100						105					110			
CAA	AAT	CGA	TAT	ACA	TTC	GCC	TTT	GGT	GGT	AAT	TAT	GAT	AGA	CTT	GAA	384
Gln	Asn	Arg	Tyr	Thr	Phe	Ala	Phe	Gly	Gly	Asn	Tyr	Asp	Arg	Leu	Glu	
		115					120					125				
CAA	CTT	GCT	GGT	AAT	CTG	AGA	GAA	AAT	ATC	GAG	TTG	GGA	AAT	GGT	CCA	432
Gln	Leu	Ala	Gly	Asn	Leu	Arg	Glu	Asn	Ile	Glu	Leu	Gly	Asn	Gly	Pro	
	130					135					140					
CTA	GAG	GAG	GCT	ATC	TCA	GCG	CTT	TAT	TAT	TAC	AGT	ACT	GGT	GGC	ACT	480
Leu	Glu	Glu	Ala	Ile	Ser	Ala	Leu	Tyr	Tyr	Tyr	Ser	Thr	Gly	Gly	Thr	
145					150					155					160	
CAG	CTT	CCA	ACT	CTG	GCT	CGT	TCC	TTT	ATA	ATT	TGC	ATC	CAA	ATG	ATT	528
Gln	Leu	Pro	Thr	Leu	Ala	Arg	Ser	Phe	Ile	Ile	Cys	Ile	Gln	Met	Ile	
				165					170						175	
TCA	GAA	GCA	GCA	AGA	TTC	CAA	TAT	ATT	GAG	GGA	GAA	ATG	CGC	ACG	AGA	576
Ser	Glu	Ala	Ala	Arg	Phe	Gln	Tyr	Ile	Glu	Gly	Glu	Met	Arg	Thr	Arg	
			180					185					190			
ATT	AGG	TAC	AAC	CGG	AGA	TCT	GCA	CCA	GAT	CCT	AGC	GTA	ATT	ACA	CTT	624
Ile	Arg	Tyr	Asn	Arg	Arg	Ser	Ala	Pro	Asp	Pro	Ser	Val	Ile	Thr	Leu	
		195					200					205				
GAG	AAT	AGT	TGG	GGG	AGA	CTT	TCC	ACT	GCA	ATT	CAA	GAG	TCT	AAC	CAA	672
Glu	Asn	Ser	Trp	Gly	Arg	Leu	Ser	Thr	Ala	Ile	Gln	Glu	Ser	Asn	Gln	
	210					215					220					
GGA	GCC	TTT	GCT	AGT	CCA	ATT	CAA	CTG	CAA	AGA	CGT	AAT	GGT	TCC	AAA	720
Gly	Ala	Phe	Ala	Ser	Pro	Ile	Gln	Leu	Gln	Arg	Arg	Asn	Gly	Ser	Lys	
225					230					235					240	
TTC	AGT	GTG	TAC	GAT	GTG	AGT	ATA	TTA	ATC	CCT	ATC	ATA	GCT	CTC	ATG	768
Phe	Ser	Val	Tyr	Asp	Val	Ser	Ile	Leu	Ile	Pro	Ile	Ile	Ala	Leu	Met	
				245					250						255	

GTG TAT AGA TGC GCA CCT CCA CCA TCG TCA CAG TTT TAA  
Val Tyr Arg Cys Ala Pro Pro Pro Ser Ser Gln Phe  
260 265

807

(2) INFORMATION FÜR SEQ ID NO: 8:

(i) SEQUENZCHARAKTERISTIK:

(A) LÄNGE: 268 Aminosäuren

(B) ART: Aminosäure

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8:

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

(2) INFORMATION FÜR SEQ ID NO: 9:

(i) SEQUENZCHARAKTERISTIK:

(A) LÄNGE: 1605 Basenpaare



(B) ART: Nucleinsäure

(C) STRANGFORM: einzel

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (genomisch)

(ix) EIGENSCHAFT:

(A) NAME/SCHLÜSSEL: CDS

(B) POSITION: 1 .. 1605

(D) ANDERE INFORMATION: /Anmerkung "Produkt = "G-FIT""

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 9:

AAG	CTT	ATG	ATA	TTC	CCC	AAA	CAA	TAC	CCA	ATT	ATA	AAC	TTT	ACC	ACA	48
Lys	Leu	Met	Ile	Phe	Pro	Lys	Gln	Tyr	Pro	Ile	Ile	Asn	Phe	Thr	Thr	
1				5					10					15		
GCG	GGT	GCC	ACT	GTG	CAA	AGC	TAC	ACA	AAC	TTT	ATC	AGA	GCT	GTT	CGC	96
Ala	Gly	Ala	Thr	Val	Gln	Ser	Tyr	Thr	Asn	Phe	Ile	Arg	Ala	Val	Arg	
			20					25					30			
GGT	CGT	TTA	ACA	ACT	GGA	GCT	GAT	GTG	AGA	CAT	GAA	ATA	CCA	GTG	TTG	144
Gly	Arg	Leu	Thr	Thr	Gly	Ala	Asp	Val	Arg	His	Glu	Ile	Pro	Val	Leu	
		35					40					45				
CCA	AAC	AGA	GTT	GGT	TTG	CCT	ATA	AAC	CAA	CGG	TTT	ATT	TTA	GTT	GAA	192
Pro	Asn	Arg	Val	Gly	Leu	Pro	Ile	Asn	Gln	Arg	Phe	Ile	Leu	Val	Glu	
	50					55					60					
CTC	TCA	AAT	CAT	GCA	GAG	CTT	TCT	GTT	ACA	TTA	GCG	CTG	GAT	GTC	ACC	240
Leu	Ser	Asn	His	Ala	Glu	Leu	Ser	Val	Thr	Leu	Ala	Leu	Asp	Val	Thr	
65					70				75					80		
AAT	GCA	TAT	GTG	GTA	GGC	TAC	CGT	GCT	GGA	AAT	AGC	GCA	TAT	TTC	TTT	288
Asn	Ala	Tyr	Val	Val	Gly	Tyr	Arg	Ala	Gly	Asn	Ser	Ala	Tyr	Phe	Phe	
			85						90					95		

CAT	CCT	GAC	AAT	CAG	GAA	GAT	GCA	GAA	GCA	ATC	ACT	CAT	CTT	TTC	ACT	336
His	Pro	Asp	Asn	Gln	Glu	Asp	Ala	Glu	Ala	Ile	Thr	His	Leu	Phe	Thr	
			100					105					110			
GAT	GTT	CAA	AAT	CGA	TAT	ACA	TTC	GCC	TTT	GGT	GGT	AAT	TAT	GAT	AGA	384
Asp	Val	Gln	Asn	Arg	Tyr	Thr	Phe	Ala	Phe	Gly	Gly	Asn	Tyr	Asp	Arg	
		115					120					125				
CTT	GAA	CAA	CTT	GCT	GGT	AAT	CTG	AGA	GAA	AAT	ATC	GAG	TTG	GGA	AAT	432
Leu	Glu	Gln	Leu	Ala	Gly	Asn	Leu	Arg	Glu	Asn	Ile	Glu	Leu	Gly	Asn	
	130					135					140					
GGT	CCA	CTA	GAG	GAG	GCT	ATC	TCA	GCG	CTT	TAT	TAT	TAC	AGT	ACT	GGT	480
Gly	Pro	Leu	Glu	Glu	Ala	Ile	Ser	Ala	Leu	Tyr	Tyr	Tyr	Ser	Thr	Gly	
145					150					155					160	
GGC	ACT	CAG	CTT	CCA	ACT	CTG	GCT	CGT	TCC	TTT	ATA	ATT	TGC	ATC	CAA	528
Gly	Thr	Gln	Leu	Pro	Thr	Leu	Ala	Arg	Ser	Phe	Ile	Ile	Cys	Ile	Gln	
				165				170						175		
ATG	ATT	TCA	GAA	GCA	GCA	AGA	TTC	CAA	TAT	ATT	GAG	GGA	GAA	ATG	CGC	576
Met	Ile	Ser	Glu	Ala	Ala	Arg	Phe	Gln	Tyr	Ile	Glu	Gly	Glu	Met	Arg	
			180					185					190			
ACG	AGA	ATT	AGG	TAC	AAC	CGG	AGA	TCT	GCA	CCA	GAT	CCT	AGC	GTA	ATT	624
Thr	Arg	Ile	Arg	Tyr	Asn	Arg	Arg	Ser	Ala	Pro	Asp	Pro	Ser	Val	Ile	
		195					200					205				
ACA	CTT	GAG	AAT	AGT	TGG	GGG	AGA	CTT	TCC	ACT	GCA	ATT	CAA	GAG	TCT	672
Thr	Leu	Glu	Asn	Ser	Trp	Gly	Arg	Leu	Ser	Thr	Ala	Ile	Gln	Glu	Ser	
	210					215					220					
AAC	CAA	GGG	GCC	TTT	GCT	AGT	CCA	ATT	CAA	CTG	CAA	AGA	CGT	AAT	GGT	720
Asn	Gln	Gly	Ala	Phe	Ala	Ser	Pro	Ile	Gln	Leu	Gln	Arg	Arg	Asn	Gly	
225					230					235					240	
TCC	AAA	TTC	AGT	GTG	TAC	GAT	GTG	AGT	ATA	TTA	ATC	CCT	ATC	ATA	GCT	768
Ser	Lys	Phe	Ser	Val	Tyr	Asp	Val	Ser	Ile	Leu	Ile	Pro	Ile	Ile	Ala	
				245					250					255		
CTC	ATG	GTG	TAT	AGA	TGC	GCA	CCT	CCA	CCA	TCG	TCA	CAG	TTT	TCT	CTT	816
Leu	Met	Val	Tyr	Arg	Cys	Ala	Pro	Pro	Pro	Ser	Ser	Gln	Phe	Ser	Leu	
			260					265					270			
CTT	ATA	AGG	CCA	GTG	GTA	CCA	AAT	TTT	AAT	GCT	GAT	GTT	TGT	ATG	GAT	864
Leu	Ile	Arg	Pro	Val	Val	Pro	Asn	Phe	Asn	Ala	Asp	Val	Cys	Met	Asp	
		275					280					285				
CCT	GAG	ATC	CAA	TTG	GTG	CAG	TCT	GGA	CCT	GAG	CTG	AAG	AAG	CCT	GGA	912
Pro	Glu	Ile	Gln	Leu	Val	Gln	Ser	Gly	Pro	Glu	Leu	Lys	Lys	Pro	Gly	
	290					295					300					

GAG Glu 305	ACA Thr	GTC Val	AAG Lys	ATC Ile	TCC Ser 310	TGC Cys	AAG Lys	GCT Ala	TCT Ser	GGA Gly 315	TAT Tyr	ACC Thr	TTC Phe	GCA Ala	AAC Asn 320	960
TAT Tyr	GGA Gly	ATG Met	AAC Asn	TGG Trp 325	ATG Met	AAG Lys	CAG Gln	GCT Ala 330	CCA Pro	GGA Gly	AAG Lys	GGT Gly	TTA Leu	AAG Lys 335	TGG Trp	1008
ATG Met	GGC Gly	TGG Trp	ATA Ile 340	AAC Asn	ACC Thr	TAC Tyr	ACT Thr	GGA Gly 345	CAG Gln	TCA Ser	ACA Thr	TAT Tyr	GCT Ala 350	GAT Asp	GAC Asp	1056
TTC Phe	AAG Lys	GAA Glu	CGG Arg 355	TTT Phe	GCC Ala	TTC Phe	TCT Ser 360	TTG Leu	GAA Glu	ACC Thr	TCT Ser	GCC Ala 365	ACC Thr	ACT Thr	GCC Ala	1104
CAT His	TTG Leu 370	CAG Gln	ATC Ile	AAC Asn	AAC Asn	CTC Leu 375	AGA Arg	AAT Asn	GAG Glu	GAC Asp	TCG Ser 380	GCC Ala	ACA Thr	TAT Tyr	TTC Phe	1152
TGT Cys 385	GCA Ala	AGA Arg	CGA Arg	TTT Phe	GGG Gly 390	TTT Phe	GCT Ala	TAC Tyr	TGG Trp	GGC Gly 395	CAA Gln	GGG Gly	ACT Thr	CTG Leu	GTC Val 400	1200
AGT Ser	GTC Val	TCT Ser	GCA Ala	TCG Ser 405	ATA Ile	TCG Ser	AGC Ser	TCT Ser	GGT Gly 410	GGC Gly	GGT Gly	GGC Gly	TCG Ser	GGC Gly 415	GGT Gly	1248
GGT Gly	GGG Gly	TCG Ser	GGT Gly 420	GGC Gly	GGC Gly	GGA Gly	TCG Ser	GAT Asp 425	ATC Ile	CAG Gln	ATG Met	ACC Thr	CAG Gln 430	TCT Ser	CCA Pro	1296
TCC Ser	TCC Ser	TTA Leu 435	TCT Ser	GCC Ala	TCT Ser	CTG Leu	GGA Gly 440	GAA Glu	AGA Arg	GTC Val	AGT Ser	CTC Leu 445	ACT Thr	TGT Cys	CGG Arg	1344
GCA Ala 450	AGT Ser	CAG Gln	GAC Asp	ATT Ile	GGT Gly 455	AAT Asn	AGC Ser	TTA Leu	ACC Thr	TGG Trp	CTT Leu	TCA Ser	CAG Gln	GAA Glu	CCA Pro	1392
GAT Asp 465	GGA Gly	ACT Thr	ATT Ile	AAA Lys	CGC Arg 470	CTG Leu	ATC Ile	TAC Tyr	GCC Ala 475	ACA Thr	TCC Ser	AGT Ser	TTA Leu	GAT Asp 480	TCT Ser	1440
GGT Gly	GTC Val	CCC Pro	AAA Lys	AGG Arg 485	TTC Phe	AGT Ser	GGC Gly	AGT Ser	CGG Arg 490	TCT Ser	GGG Gly	TCA Ser	GAT Asp	TAT Tyr 495	TCT Ser	1488
CTC Leu	ACC Thr	ATC Ile	AGT Ser 500	AGC Ser	CTT Leu	GAG Glu	TCT Ser	GAA Glu 505	GAT Asp	TTT Phe	GTA Val	GTC Val	TAT Tyr 510	TAC Tyr	TGT Cys	1536
CTA Leu	CAA Gln	TAT Tyr 515	GCT Ala	ATT Ile	TTT Phe	CCG Pro	TAC Tyr 520	ACG Thr	TTC Phe	GGA Gly	GGG Gly	GGG Gly	ACC Thr	AAC Asn	CTG Leu	1584

GAA ATA AAA CGG GCT GAT TAA  
 Glu Ile Lys Arg Ala Asp  
 530 535

1605

## (2) INFORMATION FÜR SEQ ID NO: 10:

## (i) SEQUENZCHARAKTERISTIK:

- (A) LÄNGE: 534 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: linear

## (ii) MOLEKÜLTYP: Protein

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:

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

Thr Leu Glu Asn Ser Trp Gly Arg Leu Ser Thr Ala Ile Gln Glu Ser  
 210 215 220

Asn Gln Gly Ala Phe Ala Ser Pro Ile Gln Leu Gln Arg Arg Asn Gly  
 225 230 235 240

Ser Lys Phe Ser Val Tyr Asp Val Ser Ile Leu Ile Pro Ile Ile Ala  
 245 250 255

Leu Met Val Tyr Arg Cys Ala Pro Pro Pro Ser Ser Gln Phe Ser Leu  
 260 265 270

Leu Ile Arg Pro Val Val Pro Asn Phe Asn Ala Asp Val Cys Met Asp  
 275 280 285

Pro Glu Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly  
 290 295 300

Glu Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Ala Asn  
 305 310 315 320

Tyr Gly Met Asn Trp Met Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp  
 325 330 335

Met Gly Trp Ile Asn Thr Tyr Thr Gly Gln Ser Thr Tyr Ala Asp Asp  
 340 345 350

Phe Lys Glu Arg Phe Ala Phe Ser Leu Glu Thr Ser Ala Thr Thr Ala  
 355 360 365

His Leu Gln Ile Asn Asn Leu Arg Asn Glu Asp Ser Ala Thr Tyr Phe  
 370 375 380

Cys Ala Arg Arg Phe Gly Phe Ala Tyr Trp Gly Gln Gly Thr Leu Val  
 385 390 395 400

Ser Val Ser Ala Ser Ile Ser Ser Ser Gly Gly Gly Ser Gly Gly  
 405 410 415

Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro  
 420 425 430

Ser Ser Leu Ser Ala Ser Leu Gly Glu Arg Val Ser Leu Thr Cys Arg  
 435 440 445

Ala Ser Gln Asp Ile Gly Asn Ser Leu Thr Trp Leu Ser Gln Glu Pro  
 450 455 460

Asp Gly Thr Ile Lys Arg Leu Ile Tyr Ala Thr Ser Ser Leu Asp Ser  
 465 470 475 480

Gly Val Pro Lys Arg Phe Ser Gly Ser Arg Ser Gly Ser Asp Tyr Ser  
 485 490 495

Leu Thr Ile Ser Ser Leu Glu Ser Glu Asp Phe Val Val Tyr Tyr Cys  
 500 505 510  
 Leu Gln Tyr Ala Ile Phe Pro Tyr Thr Phe Gly Gly Gly Thr Asn Leu  
 515 520 525  
 Glu Ile Lys Arg Ala Asp  
 530

(2) INFORMATION FÜR SEQ ID NO: 11:

(i) SEQUENZCHARAKTERISTIK:

- (A) LÄNGE: 45 Basenpaare
- (B) ART: Nucleinsäure
- (C) STRANGFORM: einzel
- (D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (genomisch)

(ix) EIGENSCHAFT:

- (A) NAME/SCHLÜSSEL: CDS
- (B) POSITION: 1 .. 45
- (D) ANDERE INFORMATION: /Anmerkung "Produkt = "neuer Linker/  
 Information: neuer Linker""

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:

TCG AGC TCC TCC GGA TCT TCA TCT AGC GGT TCC AGC TCG AGT GGA 45  
 Ser Ser Ser Ser Gly Ser Ser Ser Ser Gly Ser Ser Ser Ser Gly  
 1 5 10 15

(2) INFORMATION FÜR SEQ ID NO: 12:

(i) SEQUENZCHARAKTERISTIK:

- (A) LÄNGE: 15 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:

Ser Ser Ser Ser Gly Ser Ser Ser Ser Gly Ser Ser Ser Ser Gly  
1 5 10 15

(2) INFORMATION FÜR SEQ ID NO: 13:

(i) SEQUENZCHARAKTERISTIK:

- (A) LÄNGE: 45 Basenpaare
- (B) ART: Nucleinsäure
- (C) STRANGFORM: einzel
- (D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (genomisch)

(ix) EIGENSCHAFT:

- (A) NAME/SCHLÜSSEL: CDS
- (B) POSITION: 1 .. 45
- (D) ANDERE INFORMATION: /Anmerkung "Produkt = "alter Linker/  
Proteininformation: alter Linker""

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:

GGA GGA GGA GGA TCT GGA GGA GGA GGA TCT GGA GGA GGA GGA TCT 45  
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser  
1 5 10 15

(2) INFORMATION FÜR SEQ ID NO: 14:

(i) SEQUENZCHARAKTERISTIK:

- (A) LÄNGE: 15 Aminosäuren
- (B) ART: Aminosäure

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 14:

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser  
 1 5 10 15

(2) INFORMATION FÜR SEQ ID NO: 15:

(i) SEQUENZCHARAKTERISTIK:

(A) LÄNGE: 2001 Basenpaare

(B) ART: Nucleinsäure

(C) STRANGFORM: einzel

(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: DNA (genomisch)

(ix) EIGENSCHAFT:

(A) NAME/SCHLÜSSEL: CDS

(B) POSITION: 1 .. 2001

(D) ANDERE INFORMATION: /Anmerkung "Produkt = "741sFv-PE40""

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 15:

GAT CCT GAG ATC CAA TTG GTG CAG TCT GGA CCT GAG CTG AAG AAG CCT 48  
 Asp Pro Glu Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro  
 1 5 10 15

GGA GAG ACA GTC AAG ATC TCC TGC AAG GCT TCT GGG TAT ACC TTC ACA 96  
 Gly Glu Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr  
 20 25 30



AAC TAT GGA ATG AAC TGG GTG AAG CAG GCT CCA GGA AAG GGT TTA AAG Asn Tyr Gly Met Asn Trp Val Lys Gln Ala Pro Gly Lys Gly Leu Lys 35 40 45	144
TGG ATG GGC TGG ATA AAC ACC AAC ACT GGA GAG CCA ACA TAT GCT GAA Trp Met Gly Trp Ile Asn Thr Asn Thr Gly Glu Pro Thr Tyr Ala Glu 50 55 60	192
GAG TTC AAG GGA CGG TTT GCC TTC TCT TTG GAA ACC TCT GCC AGC ACT Glu Phe Lys Gly Arg Phe Ala Phe Ser Leu Glu Thr Ser Ala Ser Thr 65 70 75 80	240
GCC TAT TTG CAG ATC AAC AAC CTC AAA AAT GAG GAC ACG GCT ACA TAT Ala Tyr Leu Gln Ile Asn Asn Leu Lys Asn Glu Asp Thr Ala Thr Tyr 85 90 95	288
TTC TGT GGA AGG CAA TTT ATT ACC TAC GGC GGG TTT GCT AAC TGG GGC Phe Cys Gly Arg Gln Phe Ile Thr Tyr Gly Gly Phe Ala Asn Trp Gly 100 105 110	336
CAA GGG ACT CTG GTC ACT GTC TCT GCA TCG AGC TCC TCC GGA TCT TCA Gln Gly Thr Leu Val Thr Val Ser Ala Ser Ser Ser Ser Gly Ser Ser 115 120 125	384
TCT AGC GGT TCC AGC TCG AGC GAT ATC GTC ATG ACC CAG TCT CCT AAA Ser Ser Gly Ser Ser Ser Ser Asp Ile Val Met Thr Gln Ser Pro Lys 130 135 140	432
TTC ATG TCC ACG TCA GTG GGA GAC AGG GTC AGC ATC TCC TGC AAG GCC Phe Met Ser Thr Ser Val Gly Asp Arg Val Ser Ile Ser Cys Lys Ala 145 150 155 160	480
AGT CAG GAT GTG AGT ACT GCT GTA GCC TGG TAT CAA CAA AAA CCA GGG Ser Gln Asp Val Ser Thr Ala Val Ala Trp Tyr Gln Gln Lys Pro Gly 165 170 175	528
CAA TCT CCT AAA CTA CTG ATT TAC TGG ACA TCC ACC CGG CAC ACT GGA Gln Ser Pro Lys Leu Leu Ile Tyr Trp Thr Ser Thr Arg His Thr Gly 180 185 190	576
GTC CCT GAT CCG TTC ACA GGC AGT GGA TCT GGG ACA GAT TAT ACT CTC Val Pro Asp Pro Phe Thr Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu 195 200 205	624
ACC ATC AGC AGT GTG CAG GCT GAA GAC CTG GCA CTT CAT TAC TGT CAG Thr Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Leu His Tyr Cys Gln 210 215 220	672
CAA CAT TAT AGA GTG GCC TAC ACG TTC GGA AGG GGG ACC AAG CTG GAG Gln His Tyr Arg Val Ala Tyr Thr Phe Gly Arg Gly Thr Lys Leu Glu 225 230 235 240	720
ATA AAA CGG GCT GAT GCT GCA CCA ACT GTA TCC ATC TTC CCA CCA TCC Ile Lys Arg Ala Asp Ala Ala Pro Thr Val Ser Ile Phe Pro Pro Ser 245 250 255	768

AGT Ser	GAG Glu	CAG Gln	TTT Phe 260	GAG Glu	GGC Gly	GGC Gly	AGC Ser	CTG Leu	GCC Ala	GCG Ala	CTG Leu	AAC Asn	GCG Ala	CAC His	CAG Gln	816
GCT Ala	TGC Cys	CAC His 275	CTG Leu	CCG Pro	CTG Leu	GAG Glu	ACT Thr 280	TTC Phe	ACC Thr	CGT Arg	CAT His	CGC Arg	CAG Gln	CCG Pro	CGC Arg	864
GGC Gly	TGG Trp 290	GAA Glu	CAA Gln	CTG Leu	GAG Glu	CAG Gln 295	TGC Cys	GGC Gly	TAT Tyr	CCG Pro	GTG Val	CAG Gln	CGG Arg	CTG Leu	GTC Val	912
GCC Ala 305	CTC Leu	TAC Tyr	CTG Leu	GCG Ala	GCG Ala	CGG Arg 310	CTG Leu	TCG Ser	TGG Trp	AAC Asn 315	CAG Gln	GTC Val	GAC Asp	CAG Gln	GTC Val 320	960
ATC Ile	CGC Arg	AAC Asn	GCC Ala 325	CTG Leu	GCC Ala	AGC Ser	CCC Pro	GGC Gly	AGC Ser	GGC Gly	GGC Gly	GAC Asp	CTG Leu	GGC Gly	GAA Glu 335	1008
GCG Ala	ATC Ile	CGC Arg	GAG Glu 340	CAG Gln	CCG Pro	GAG Glu	CAG Gln 345	GCC Ala	CGT Arg	CTG Leu	GCC Ala	CTG Leu	ACC Thr	CTG Leu	GCC Ala	1056
GCC Ala	GCC Ala	GAG Glu 355	AGC Ser	GAG Glu	CGC Arg	TTC Phe 360	GTC Val	CGG Arg	CAG Gln	GGC Gly	ACC Thr	GGC Gly	AAC Asn	GAC Asp	GAG Glu	1104
GCC Ala 370	GGC Gly	GCG Ala	GCC Ala	AAC Asn	GCC Ala	GAC Asp 375	GTG Val	GTG Val	AGC Ser	CTG Leu	ACC Thr	TGC Cys	CCG Pro	GTC Val	GCC Ala	1152
GCC Ala 385	GGT Gly	GAA Glu	TGC Cys	GCG Ala	GGC Gly 390	CCG Pro	GCG Ala	GAC Asp	AGC Ser	GGC Gly 395	GAC Asp	GCC Ala	CTG Leu	CTG Leu	GAG Glu 400	1200
CGC Arg	AAC Asn	TAT Tyr	CCC Pro	ACT Thr 405	GGC Gly	GCG Ala	GAG Glu	TTC Phe	CTC Leu	GGC Gly	GAC Asp	GGC Gly	GGC Gly	GAC Asp	GTC Val 415	1248
AGC Ser	TTC Phe	AGC Ser	AAC Asn 420	CGC Arg	GGC Gly	ACG Thr	CAG Gln	AAC Asn 425	TGG Trp	ACG Thr	GTG Val	GAG Glu	CGG Arg	CTG Leu	CTC Leu	1296
CAG Gln	GCG Ala	CAC His 435	CGC Arg	CAA Gln	CTG Leu	GAG Glu	GAG Glu 440	CGC Arg	GGC Gly	TAT Tyr	GTG Val	TTC Phe	GTC Val	GGC Gly	TAC Tyr	1344
CAC His 450	GGC Gly	ACC Thr	TTC Phe	CTC Leu	GAA Glu	GCG Ala	GCG Ala	CAA Gln 455	AGC Ser	ATC Ile	GTC Val	TTC Phe	GGC Gly	GGG Gly	GTC Val	1392
CGC Arg 465	GCG Ala	CGC Arg	AGC Ser	CAG Gln	GAC Asp 470	CTC Leu	GAC Asp	GCG Ala	ATC Ile	TGG Trp 475	CGC Arg	GGT Gly	TTC Phe	TAT Tyr	ATC Ile 480	1440

GCC Ala	GGC Gly	GAT Asp	CCG Pro	GCG Ala 485	CTG Leu	GCC Ala	TAC Tyr	GGC Gly	TAC Tyr 490	GCC Ala	CAG Gln	GAC Asp	CAG Gln	GAA Glu 495	CCC Pro	1488
GAC Asp	GCA Ala	CGC Arg	GGC Gly 500	CGG Arg	ATC Ile	CGC Arg	AAC Asn	GGT Gly 505	GCC Ala	CTG Leu	CTG Leu	CGG Arg	GTC Val 510	TAT Tyr	GTG Val	1536
CCG Pro	CGC Arg	TCG Ser 515	AGC Ser	CTG Leu	CCG Pro	GGC Gly	TTC Phe 520	TAC Tyr	CGC Arg	ACC Thr	AGC Ser	CTG Leu 525	ACC Thr	CTG Leu	GCC Ala	1584
GCG Ala 530	CCG Pro	GAG Glu	GCG Ala	GCG Ala	GGC Gly	GAG Glu 535	GTC Val	GAA Glu	CGG Arg	CTG Leu	ATC Ile 540	GGC Gly	CAT His	CCG Pro	CTG Leu	1632
CCG Pro 545	CTG Leu	CGC Arg	CTG Leu	GAC Asp 550	GCC Ala	ATC Ile	ACC Thr	GGC Gly	CCC Pro 555	GAG Glu 555	GAG Glu	GAA Glu	GGC Gly	GGG Gly	CGC Arg 560	1680
CTG Leu	GAG Glu	ACC Thr	ATT Ile	CTC Leu 565	GGC Gly	TGG Trp	CCG Pro	CTG Leu	GCC Ala 570	GAG Glu	CGC Arg	ACC Thr	GTG Val 575	GTG Val	ATT Ile	1728
CCC Pro	TCG Ser	GCG Ala	ATC Ile 580	CCC Pro	ACC Thr	GAC Asp	CCG Pro	CGC Arg	AAC Asn 585	GTC Val	GGC Gly	GGC Gly	GAC Asp 590	CTC Leu	GAC Asp	1776
CCG Pro	TCC Ser	AGC Ser 595	ATC Ile	CCC Pro	GAC Asp	AAG Lys	GAA Glu 600	CAG Gln	GCG Ala	ATC Ile	AGC Ser	GCC Ala 605	CTG Leu	CCG Pro	GAC Asp	1824
TAC Tyr 610	GCC Ala	AGC Ser	CAG Gln	CCC Pro	GGC Gly	AAA Lys 615	CCG Pro	CCG Pro	CGC Arg	GAG Glu	GAC Asp 620	CTG Leu	AAG Lys	TAA *	CTG Leu	1872
CCG Pro 625	CGA Arg	CCG Pro	GCC Ala	GGC Gly 630	TCC Ser	CTT Leu	CGC Arg	AGG Arg	AGC Ser	CGG Arg 635	CCT Pro	TCT Ser	CGG Arg	GGC Gly	CTG Leu 640	1920
GCC Ala	ATA Ile	CAT His	CAG Gln	GTT Val 645	TTC Phe	CTG Leu	ATG Met	CCA Pro	GCC Ala	CAA Gln 650	TCG Ser	AAT Asn	ATG Met	AAT Asn 655	TGA *	1968
TCC Ser	TCT Ser	AGA Arg	GTC Val 660	GAC Asp	CTG Leu	CAG Gln	GCA Ala	TGC Cys 665	AAG Lys	CTT Leu						2001

## (2) INFORMATION FÜR SEQ ID NO: 16:

## (i) SEQUENZCHARAKTERISTIK:

(A) LÄNGE: 667 Aminosäuren

(B) ART: Aminosäure

(D) TOPOLOGIE: linear

## (ii) MOLEKÜLTYP: Protein

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 16:

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

Ala Cys His Leu Pro Leu Glu Thr Phe Thr Arg His Arg Gln Pro Arg  
275 280 285

Gly Trp Glu Gln Leu Glu Gln Cys Gly Tyr Pro Val Gln Arg Leu Val  
290 295 300

Ala Leu Tyr Leu Ala Ala Arg Leu Ser Trp Asn Gln Val Asp Gln Val  
305 310 315 320

Ile Arg Asn Ala Leu Ala Ser Pro Gly Ser Gly Gly Asp Leu Gly Glu  
325 330 335

Ala Ile Arg Glu Gln Pro Glu Gln Ala Arg Leu Ala Leu Thr Leu Ala  
340 345 350

Ala Ala Glu Ser Glu Arg Phe Val Arg Gln Gly Thr Gly Asn Asp Glu  
355 360 365

Ala Gly Ala Ala Asn Ala Asp Val Val Ser Leu Thr Cys Pro Val Ala  
370 375 380

Ala Gly Glu Cys Ala Gly Pro Ala Asp Ser Gly Asp Ala Leu Leu Glu  
385 390 395 400

Arg Asn Tyr Pro Thr Gly Ala Glu Phe Leu Gly Asp Gly Gly Asp Val  
405 410 415

Ser Phe Ser Asn Arg Gly Thr Gln Asn Trp Thr Val Glu Arg Leu Leu  
420 425 430

Gln Ala His Arg Gln Leu Glu Glu Arg Gly Tyr Val Phe Val Gly Tyr  
435 440 445

His Gly Thr Phe Leu Glu Ala Ala Gln Ser Ile Val Phe Gly Gly Val  
450 455 460

Arg Ala Arg Ser Gln Asp Leu Asp Ala Ile Trp Arg Gly Phe Tyr Ile  
465 470 475 480

Ala Gly Asp Pro Ala Leu Ala Tyr Gly Tyr Ala Gln Asp Gln Glu Pro  
485 490 495

Asp Ala Arg Gly Arg Ile Arg Asn Gly Ala Leu Leu Arg Val Tyr Val  
500 505 510

Pro Arg Ser Ser Leu Pro Gly Phe Tyr Arg Thr Ser Leu Thr Leu Ala  
515 520 525

Ala Pro Glu Ala Ala Gly Glu Val Glu Arg Leu Ile Gly His Pro Leu  
530 535 540

Pro Leu Arg Leu Asp Ala Ile Thr Gly Pro Glu Glu Glu Gly Gly Arg  
545 550 555 560

Leu Glu Thr Ile Leu Gly Trp Pro Leu Ala Glu Arg Thr Val Val Ile  
 565 570 575  
 Pro Ser Ala Ile Pro Thr Asp Pro Arg Asn Val Gly Gly Asp Leu Asp  
 580 585 590  
 Pro Ser Ser Ile Pro Asp Lys Glu Gln Ala Ile Ser Ala Leu Pro Asp  
 595 600 605  
 Tyr Ala Ser Gln Pro Gly Lys Pro Pro Arg Glu Asp Leu Lys \* Leu  
 610 615 620  
 Pro Arg Pro Ala Gly Ser Leu Arg Arg Ser Arg Pro Ser Arg Gly Leu  
 625 630 635 640  
 Ala Ile His Gln Val Phe Leu Met Pro Ala Gln Ser Asn Met Asn \*  
 645 650 655  
 Ser Ser Arg Val Asp Leu Gln Ala Cys Lys Leu  
 660 665

### Patentansprüche

1. Einzelketten-Fv (sFv)-Polypeptid, das eine Bindungsstelle definiert, an welche ein c-erbB-2-Tumor-Antigen bindet, wobei das Fv mindestens zwei Polypeptid-Domänen umfasst, die durch einen Polypeptid-Linker verknüpft sind, der den Abstand zwischen dem C-Terminus der einen Domäne und dem N-Terminus der anderen überbrückt, wobei die Aminosäuresequenz von jeder der Polypeptid-Domänen einen Satz von Komplementaritäts-bestimmenden Regionen (CDRs) umfasst, welche mindestens 70 % der Aminosäuresequenz der CDRs eines c-erbB-2-bindenden monoclonalen Antikörpers 741 F8 beibehalten.
2. Einzelketten-Fv (sFv)-Polypeptid nach Anspruch 1, wobei die Komplementaritäts-bestimmenden Regionen (CDRs) mindestens 90 % der Aminosäuresequenz der CDRs eines c-erbB-2-bindenden monoclonalen Antikörpers 741 F8 beibehalten.
3. Einzelketten-Fv (sFv)-Polypeptid nach Anspruch 2, wobei die Komplementaritäts-bestimmenden Regionen (CDRs) 100 % der Aminosäuresequenz der CDRs eines c-erbB-2-bindenden monoclonalen Antikörpers 741 F8 beibehalten.
4. Einzelketten-Fv-Polypeptid nach einem der Ansprüche 1 bis 3, wobei der Polypeptid-Linker die Aminosäuresequenz umfasst, wie sie im Sequenzprotokoll als Aminosäurereste Nummern 116 bis einschließlich 133 in SEQ ID NO: 4 dargestellt ist.
5. Einzelketten-Fv-Polypeptid nach einem der Ansprüche 1 bis 3, wobei der Polypeptid-Linker eine Aminosäuresequenz umfasst, die ausgewählt ist aus der Gruppe von Sequenzen dargestellt als Aminosäurereste 122-135 in SEQ ID NO: 15 und den Aminosäuresequenzen wie dargestellt in SEQ ID NO: 12 und SEQ ID NO: 14.
6. Einzelketten-Fv-Polypeptid nach einem der Ansprüche 1 bis 3, das weiterhin eine daran gebundene, aus der Entfernung nachweisbare Einheit umfasst, um die Darstellung einer Zelle, die das c-erbB-2-Tumor-Antigen trägt, zu ermöglichen.
7. Einzelketten-Fv-Polypeptid nach Anspruch 6, wobei die aus der Entfernung nachweisbare Einheit ein radioaktives Atom umfasst.
8. Einzelketten-Fv-Polypeptid nach einem der Ansprüche 1 bis 3, das weiterhin, verknüpft mit dem N- oder C-Terminus der verknüpften Domäne, eine dritte Polypeptid-Domäne umfasst, die eine Aminosäuresequenz umfasst, welche zwischen Framework Regionen (FRs) eingeschobenene CDRs und eine zweite immunologisch aktive Stelle definiert.
9. Einzelketten-Fv-Polypeptid nach Anspruch 8, das weiterhin eine vierte Polypeptid-Domäne umfasst, wobei die dritte und vierte Polypeptid-Domäne gemeinsam eine zweite Stelle umfassen, welche ein c-erbB-2-Tu-

mor-Antigen immunologisch bindet.

10. DNA-Sequenz, die die Polypeptidkette nach einem der Ansprüche 1 bis 3 codiert.

11. Verfahren zur Herstellung eines Einzelketten-Polypeptids, das Spezifität für ein c-erbB-2-Tumor-Antigen aufweist, wobei das Verfahren die Schritte umfasst:

(a) Transfizieren der DNA nach Anspruch 10 in eine Wirtszelle zur Herstellung einer Transformante; und  
(b) Züchten der Transformante zur Herstellung des Einzelketten-Polypeptids.

12. Wirtszelle, transfiziert mit einer DNA nach Anspruch 10.

13. Polypeptid nach einem der Ansprüche 1 bis 9 zur Verwendung in einem Diagnoseverfahren.

14. Verwendung eines Polypeptids nach einem der Ansprüche 1 bis 9 für die Herstellung eines bildgebenden Stoffes zur Darstellung eines c-erbB-2-Antigenexprimierenden Tumors.

Es folgen 5 Blatt Zeichnungen

Anhängende Zeichnungen

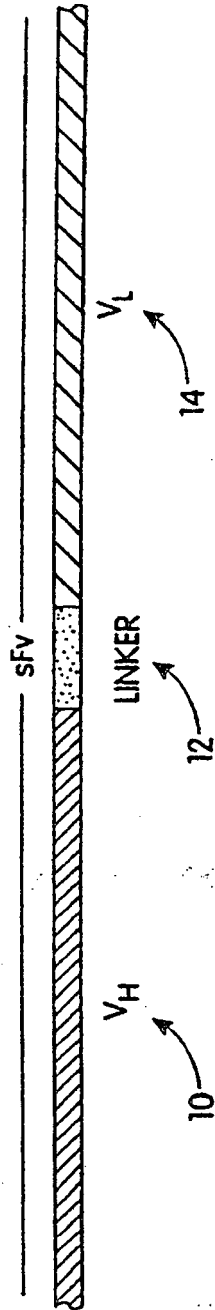


Fig. 1A



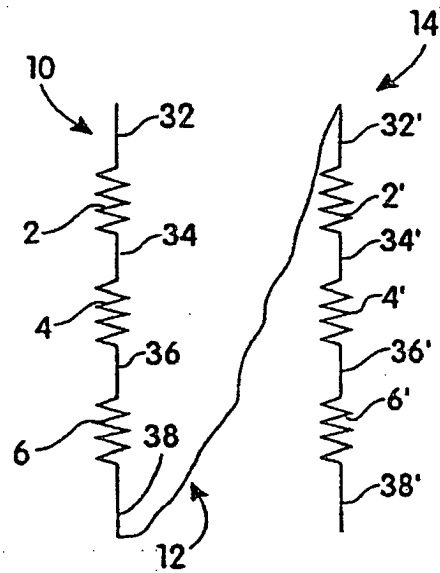


Fig. 1B

Fig. 2A

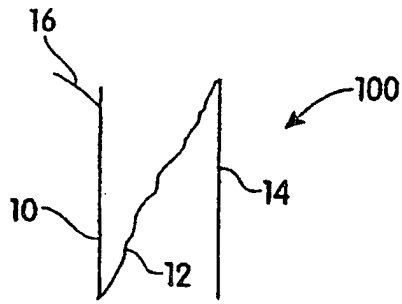


Fig. 2B

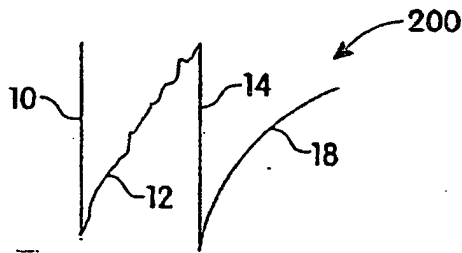


Fig. 2C

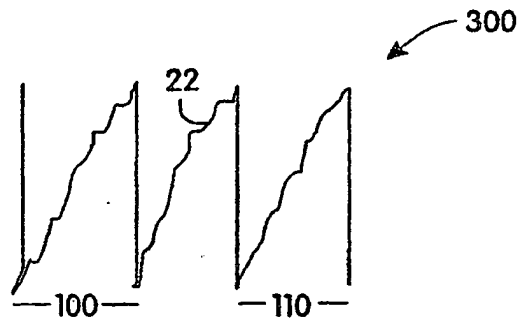


Fig. 2D

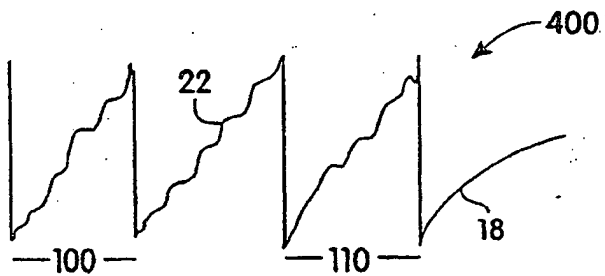
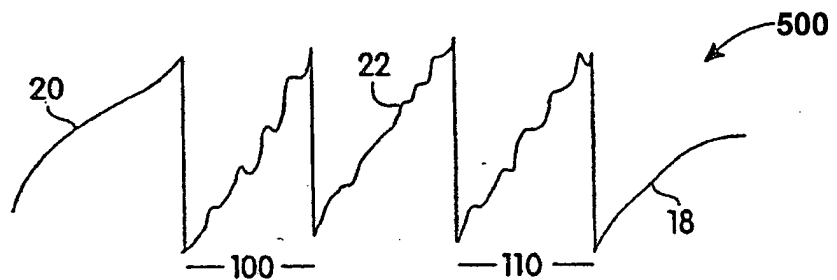


Fig. 2E



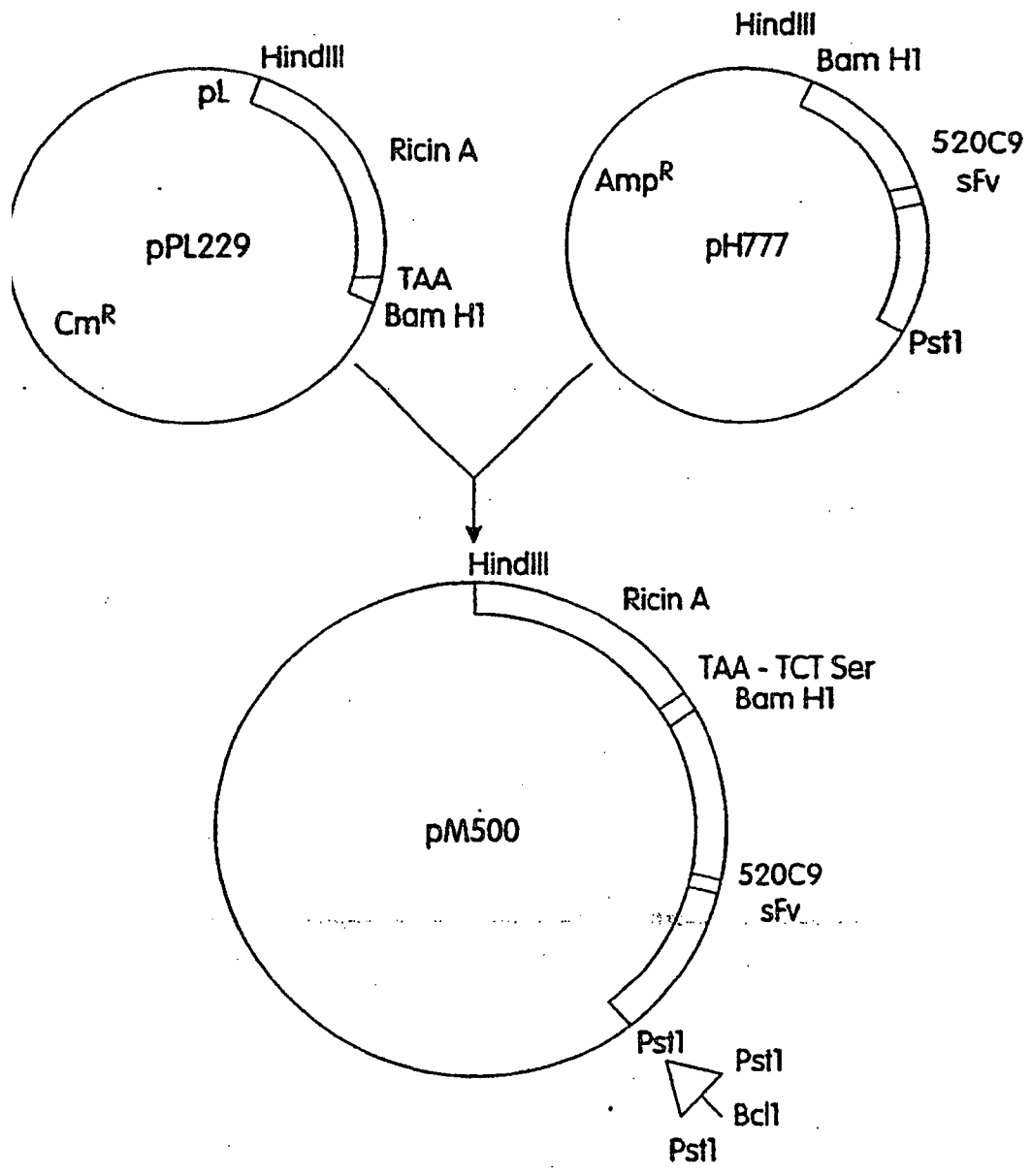
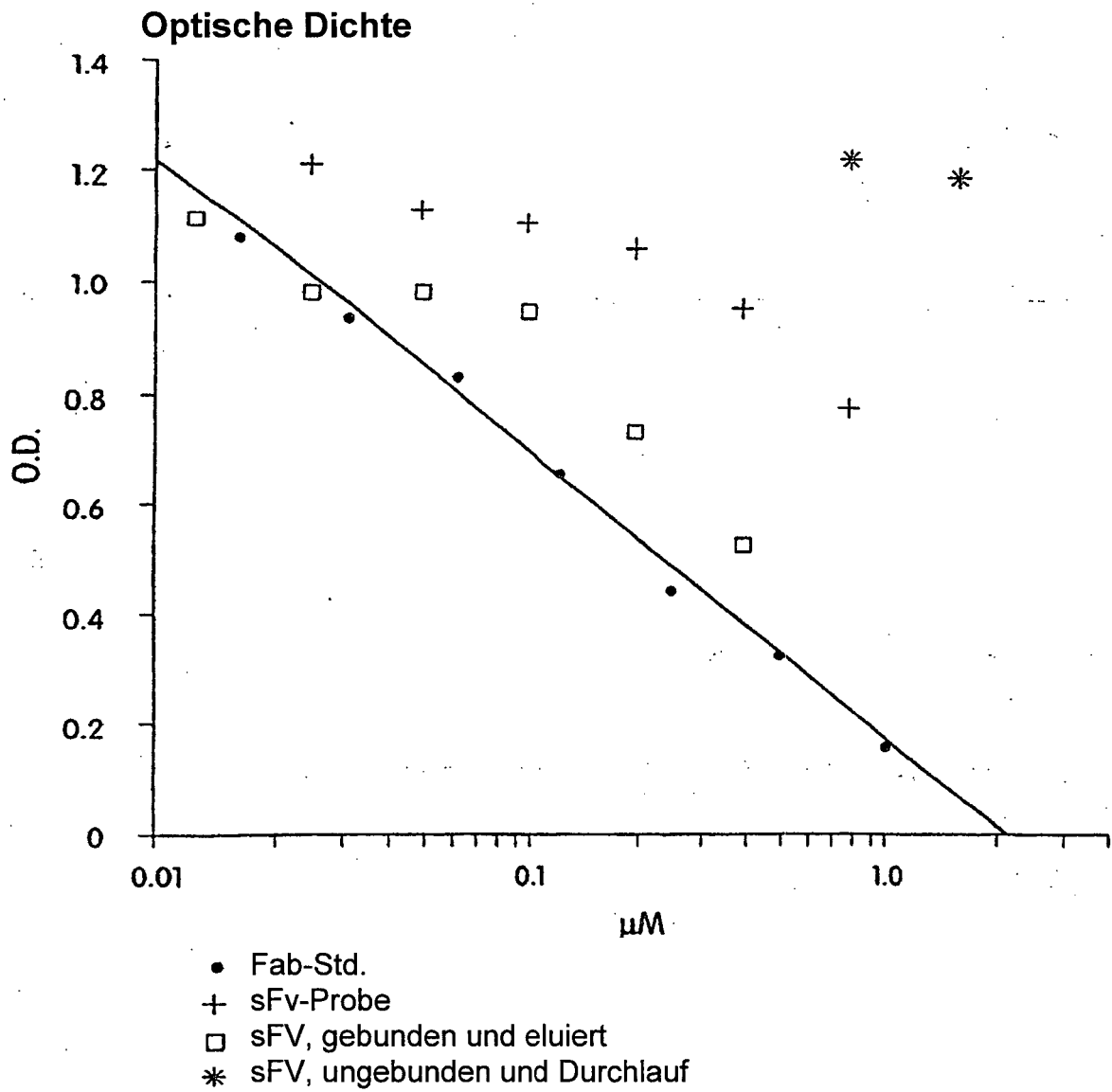


Fig. 3



**Fig. 4**