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

(10) **DE 692 33 137 T2** 2004.04.29

(12) # Übersetzung der europäischen Patentschrift

(97) **EP 0 586 607 B1**
(21) Deutsches Aktenzeichen: **692 33 137.9**
(86) PCT-Aktenzeichen: **PCT/US92/04295**
(96) Europäisches Aktenzeichen: **92 915 976.2**
(87) PCT-Veröffentlichungs-Nr.: **WO 92/020798**
(86) PCT-Anmeldetag: **21.05.1992**
(87) Veröffentlichungstag der PCT-Anmeldung: **26.11.1992**
(97) Erstveröffentlichung durch das EPA: **16.03.1994**
(97) Veröffentlichungstag der Patenterteilung beim EPA: **23.07.2003**
(47) Veröffentlichungstag im Patentblatt: **29.04.2004**

(51) Int Cl.[7]: **C07K 14/475**
C12P 21/02, C12N 5/10

(30) Unionspriorität:

705256	**24.05.1991**	**US**
765212	**25.09.1991**	**US**
790801	**08.11.1991**	**US**
847743	**06.03.1992**	**US**
880917	**11.05.1992**	**US**

(73) Patentinhaber:
Genentech, Inc., South San Francisco, Calif., US

(74) Vertreter:
LEINWEBER & ZIMMERMANN, 80331 München

(84) Benannte Vertragsstaaten:
BE, CH, DE, DK, ES, FR, GB, IT, LI, NL

(72) Erfinder:
VANDLEN, L., Richard, Hillsborough, US;
HOLMES, E., William, Pacifica, US

(54) Bezeichnung: **HEREGULINE (HRGS), P185?ERB2 BINDENDE PROTEINE**

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

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

Beschreibung

Gebiet der Erfindung

[0001] Die vorliegende Erfindung betrifft Polypeptidliganden, die sich an Rezeptoren binden, die am Zellwachstum beteiligt sind. Insbesondere betrifft sie Polypeptidliganden, die sich an den $p185^{HER2}$-Rezeptor binden.

Beschreibung der Hintergrunds und des verwandten Standes der Technik

[0002] Zelluläre Protoonkogene kodieren für Proteine, die – so die allgemeine Auffassung – die normale Zellvermehrung und -differenzierung regulieren. Änderungen ihrer Struktur oder Amplifikation ihrer Expression führen zu abnormalem Zellwachstum, und sie wurden mit Karzinogenese assoziiert (Bishop, J. M., Science 235, 305–311 [1978]; Rhims, J. S., Cancer Detection and Prevention 11, 139–149 [1988]; Nowell, P. C., Cancer Res. 46, 2203–2207 [1986]; Nicolson, G. L. Cancer Res. 47, 1473–1487 [1987]). Protoonkogene wurden zum ersten Mal unter Anwendung eines von zwei Verfahren identifiziert. Im ersten zeigte die molekulare Charakterisierung der Genome transformierender Retroviren, dass die für die Transformationsfähigkeit des Virus verantwortlichen Gene in vielen Fällen modifizierte Versionen von in den Genomen normaler Zelle vorkommenden Genen sind. Die normale Version ist das Protoonkogen, das durch Mutation verändert wird, um das Onkogen entstehen zu lassen. Ein Beispiel für ein solches Genpaar ist der EGF-Rezeptor und das v-erb-B-Geprodukt. Das viral kodierte v-erb-B-Genprodukt wurde Kürzung und anderen Änderungen unterzogen, die es konstitutiv aktivieren und es mit der Fähigkeit versehen, Zelltransformation zu induzieren (Yarden et al., Ann. Rev. Biochem. 57, 443–478 [1988]).

[0003] Das zweite Verfahren zum Detektieren zellulärer Transformationsgene, die sich dominant verhalten, umfasst die Transfektion zellulärer DNA aus Tumorzellen verschiedener Spezies in nichttransformierte Zielzellen einer heterologen Spezies. Am häufigsten erfolgte dies durch Transfektion von Human-, Vogel- oder Ratten-DNA in die Mäuse-NIH 3T3-Zelllinie (Bishop, J. M. Science 235, 305–311 [1987]; Rhims, J. S. Cancer Detection and Prevention 11, 139–149 [19881; Nowell, P. C. Cancer Res. 46, 2203–2207 [1986]; Nicolson, G. L., Cancer Res. 47, 1473–1487 [1987]; Yarden et al., Ann. Rev. Biochem. 57, 443–478 [1988]). Nach mehreren Zyklen genomischer DNA-Isolation und Retransfektion wurde die Human-DNA oder DNA aus anderen Spezies molekular aus dem Mäuse-Hintergrund kloniert und anschließend charakterisiert. In einigen Fällen wurden die gleichen Gene nach Transfektion und Klonierung isoliert, die durch direkte Charakterisierung transformierender Viren identifiziert worden waren. In anderen Fällen wurden neuartige Onkogene identifiziert. Ein Beispiel für ein durch diesen Transfektionstest identifiziertes neuartiges Onkogen ist das neu-Onkogen. Es wurde von Weinberg et al. in einem Transfektionsexperiment entdeckt, in dem Anfangs-DNA aus einem Karzinogen-induzierten Ratten-Neuroblastom abgeleitet wurde (Padhy et al., Cell 28, 865–871 [1982]; Schechter et al., Nature 312, 513–516 [1984]). Die Charakterisierung des Rattenneu-Onogens deckte auf, dass es die Struktur einer Wachstumsfaktorrezeptor-Tyrosin-Kinase aufweis, Homologie mit dem EGF-Rezeptor besaß und sich von seinem Gegenstück, dem neu-Protoonkogen, durch eine aktivierende Mutation in seiner Transmembrandomäne unterschied (Bargmann et al., Cell 45, 649–657 [1986]). Das Human-Gegenstück von neu ist das HER2-Protoonkogen, das auch als c-erb-B2 bezeichnet wird (Coussens et al., Science 230, 1137–1139 [1985]; WO 89/06692]).

[0004] Die Assoziation des HER2-Protoonkogens mit Krebs wurde mittels eines dritten Verfahrens festgelegt, d. h. mittels der Assoziation mit Brustkrebs beim Menschen. Das HER2-Protoonkogen wurde zuerst aufgrund seiner Homologie mit dem EGF-Rezeptor in cDNA-Bibliotheken entdeckt, mit welchem Rezeptor es durchwegs strukturelle Ähnlichkeiten aufweist (Yarden et al., Ann. Rev. Biochem. 57, 443–478 [1988]). Wenn radioaktive Sonden aus der für $p185^{HER2}$ kodierenden cDNA-Sequenz dazu dienten, DNA-Proben von Brustkrebspatienten zu screenen, wurde Amplifikation des HER2-Protoonkogens in etwa 30% der Patientenproben beobachtet (Slamon et al., Science 235, 177-182 [1987]). Weitere Studien bestätigten, diese Beobachtung und dehnten die Theorie insofern aus, dass sie eine wichtige Korrelation zwischen HER2-Protoonkogen-Amplifikation und/oder Überexpression und beeinträchtigte Prognosemöglichkeiten beim Eier-stockkrebs und nicht-kleinzelligem Lungenkrebs vorschlugen (Slamon et al., Science 244, 707–712 [1989]; Wright et al., Cancer Res. 49, 2087–2090 [1989]; Paik et al., J. Clin. Oncology 8, 103–112 [1990]; Berchuck et al., Cancer Res. 50, 4087–4091 [1990]; Kern et al., Cancer Res. 50, 5184–5191 [1990]).

[0005] Die Assoziation von HER2-Amplifikation/Überexpression mit aggressiver Malignität (s. o.) impliziert, dass sie eine wichtige Rolle im Verlauf von Krebserkrankungen beim Menschen spielen könnte; zwar wurden in der Vergangenheit zahlreiche tumorassoziierte Zelloberflächen-Antigene beschrieben, doch scheinen wenige von ihnen eine direkte Rolle in der Entstehung und im Verlauf von Erkrankungen zu spielen (Schlom et al., Cancer Res. 50, 820–827 [1990]; Szala et al., Proc. Natl. Acad. Sci. 98, 3542–3546).

[0006] Unter den Protoonkogenen finden sich jene, die für zelluläre Wachstumsfaktoren kodieren, die durch

endoplasmatische Kinase-Phosphorylierung von zytoplasmatischem Protein wirken. Das HER1-Gen (oder erb-B1) kodiert für den Epidermis-Wachstumsfaktor(EGF-) Rezeptor. Die β-Kette des Thrombozyten-Wachstumsfaktors wird vom c-sis-Gen kodiert. Der Granulozyten-Makrophagen-Kolonie-stimulierende Faktor wird vom c-fms-Gen kodiert. Das neu-Protoonkogen wurde in Ethylnitrosoharnstoff-induzierten Ratten-Neurobastomen identifiziert. Das HER2-Gen kodiert für das 1.255 Aminosäuren große Tyrosin-Kinase-Rezeptor-ähnliche Glykoprotein $p185^{HER2}$, das Homologie mit dem Human-Epidermis-Wachstumsfaktorrezeptor aufweist.

[0007] Die bekannten Rezeptor-Tyrosin-Kinasen besitzen alle das gleiche allgemeine Strukturmotiv: eine Liganden bindende extrazelluläre Domäne und eine für die Signaltransduktion und -transformation notwendige intrazelluläre Tyrosin-Kinase-Domäne. Diese zwei Domänen sind durch einen einzelnen Abschnitt von etwa 20 größtenteils hydropoben Aminosäuren, die als Transmembran-Überspannungssequenz bezeichnet werden, miteinander verbunden. Man geht davon aus, dass die Transmembran-Überspannungssequenz eine Rolle dabei spielt, das durch Ligandenbindung erzeugte Signal vom Zelläußeren in das Zellinnere zu übertragen. Das mit dieser allgemeinen Struktur übereinstimmende Human-$p185^{HER2}$-Glykoprotein, das sich auf der Zelloberfläche befindet, kann in drei Hauptabschnitte unterteilt werden: eine extrazelluläre Domäne (auch als ECD oder XCD bezeichnet); eine transmembrane Überspannungssequenz; und eine zytoplasmatische, intrazelluläre Tyrosin-Kinase-Domäne. Während man annimmt, dass die extrazelluläre Domäne ein Ligandenrezeptor ist, wurde der $p185^{HER2}$-Ligand noch nicht zuverlässig identifiziert.

[0008] Es wurde keine spezifische Ligandenbindung an $p185^{HER2}$-Ligand identifiziert, obwohl Lupu et al., Science 249, 1552–1555 (1989) ein inhibitorisches 30 kDa großes Glykoprotein beschreiben, das von Human-Brustkrebszellen sekretiert wird und angeblich ein Ligand für $p185^{HER2}$ ist. Lupu et al., Science 249, 1552–1555 (1990); Proceedings of the American Assoc. for Cancer Research, Bd. 32, Abstract 297, März 1991) berichteten über die Reinigung eines 30 kD-Faktors aus MDA-MB-231-Zellen und eines 75 kD-Faktors aus SK-BR-3-Zellen, der $p185^{HER2}$ stimuliert. Der 75 kD-Faktor induzierte der Literatur zufolge die Phosphorylierung von $p185^{HER2}$ und modulierte die Zellproliferation und Koloniebildung von SK-BR-2-Zellen, die den $p185^{HER2}$-Rezeptor überexprimieren. Der 30-kD-Faktor konkurriert mit muMab 4D5 um die Bindun an $p185^{HER2}$ g an $p185^{HER2}$, und seine Wachstumswirkung auf SK-BR-3-Zellen hing von der 30 kD-Konzentration ab (stimulierend bei niedrigen Konzentrationen und inhibitorisch bei höheren Konzentrationen). Außerdem stimulierte er das Wachstum von MDA-MB-468-Zellen (EGF-R-positiv, $p185^{HER2}$-negativ), er stimulierte die Phosphorylierung des EGF-Rezeptors, und er konnte aus SK-BR-3-Zellen erhalten werden. Im neu-Rattensystem beschreiben Yarden et al., Biochemistry 30, 3543–3550 (1991) einen 35 kDa-Glykoprotein-Kandidatenliganden für den durch ras-transformierte Fibroblasten sekretierten neu-kodierten Rezeptor. Dobashi et al., Proc. Natl. Acad. Sci. USA 88, 8582–8586 (1991); Biochem. Biophys. Res. Commun. 179, 1536–1542 (1991) beschreiben einen neu-Protein-spezifischen Aktivierungsfaktor (NAF), der durch Human-T-Zelllinien-ATL-2 sekretiert wird und ein Molekulargewicht im Bereich von 8–24 kD afweist. Ein 25 kD-Ligand aus aktivierten Makrophagen wurde ebenfalls beschrieben (Tarakhovsky et al., J. Cancer Res. 2188-2196 (1991).

[0009] Verfahren für den In-vivo-Test von Tumoren unter Verwendung HER2-spezifischer monoklonaler Antikörper und Verfahren zur Behandlung von Tumorzellen mittels HER2-spezifischer monoklonaler Antikörper sind in WO 89/06692 beschrieben.

[0010] Es besteht auf dem Gebiet der Erfindung weiterhin die Notwendigkeit, jene(n) Liganden zu identifizieren, der bzw. die $p185^{HER2}$ aktivieren, und seine bzw. ihre biologische(n) Funktionen) zu ermitteln, z. B. die Rolle im Zellwachstum und in der Zelldifferenzierung, in der Zelltransformation und ich der Schaffung maligner Neoplasmen.

[0011] Demzufolge ist es ein Ziel der Erfindung, ein oder mehrere neuartige $p185^{HER2}$-Liganden-Polypeptid(e) zu identifizieren und zu reinigen, das bzw. die $p185^{HER2}$ bindet/binden und stimuliert/stimulieren.

[0012] Ein weiteres Ziel ist die Bereitstellung von für Nucleinsäuren kodierenden neuartigen $p185^{HER2}$-Bindungsligand-Polypeptiden in rekombinanter Zellkultur für therapeutische oder diagnostische Zwecke sowie zur Produktion therapeutischer Antagonisten zur Verwendung bei bestimmten Stoffwechselstörungen, wie z. B. die Abtötung, Hemmung und/oder diagnostische Abbildung von Tumoren und Tumor-induzierenden Zellen, ohne darauf beschränkt zu sein.

[0013] Es ist ein weiteres Ziel der Erfindung, Derivate und modifizierte Formen neuartiger Glykoproteinliganden bereitzustellen, z. B. Aminosäuresequenz-Varianten, Fusionspolypeptide, die einen $p185^{HER2}$-Bindungsligandenden und ein heterologes Protein kombinieren, und kovalente Derivate eines $p185^{HER2}$-Bindungsliganden.

[0014] Es ist ein zusätzliches Ziel der Erfindung, Immunogene zur Bildung von Antikörpern gegen $p185^{HER2}$-Bindungsliganden bereitzustellen sowie Antikörper, die zur Bindung an solche Liganden fähig sind, und Antikörper, die einen $p185^{HER2}$-Bindungsliganden binden und den Liganden an der Aktivierung von $p185^{HER2}$ hindern, zu erhalten. Ein weiteres Ziel ist die Herstellung von immunogenen, die einen $p185^{HER2}$-Bindungsliganden, fusioniert mit einem immunogenen heterologen Polypeptid, umfassen.

[0015] Diese und andere Ziele der Erfindung ergeben sich für Fachleute auf dem Gebiet nach Durchsicht der gesamten Beschreibung.

Zusammenfassung der Erfindung

[0016] Gemäß den Zielen der Erfindung identifizierten und isolierten die Anmelden neuartige Ligandenfamilien, die sich an $p185^{HER2}$ binden. Diese Liganden werden als Heregulin(HRG-) Polypeptide bezeichnet; zu ihnen zählen HRG-a, HRG-β1, HRG-β2, HRG-β3 und die andere HRG-Polypeptide, die mit Antikörpern gegen diese Familienmitglieder kreuzreagieren und/oder die im Wesentlichen homolog sind (Definition siehe unten). Ein bevorzugtes HRG ist der in **Fig.** 4 gezeigte Ligand und seine Fragmente, die als HRG-a bezeichnet werden. Andere bevorzugte HRG sind die in **Fig.** 8 dargestellten Liganden und ihre Fragmente, die als HRG-β1, HRG-β2 (siehe **Fig.** 12) und HRG-β3 (siehe **Fig.** 13) bezeichnet werden.

[0017] In einem weiteren Aspekt bietet die Erfindung eine Zusammensetzung, die HRG umfasst, das aus seiner Quellenumgebung isoliert ist, insbesondere HRG, das frei von kontaminierenden Human-Polypeptiden ist. HRG wird durch Absorption an Heparin-Sepharose, Kationen- (z. B. Polyasparaginsäure) Austauschharze und RP-HPLC gereinigt.

[0018] HRG oder HRG-Fragmente (die auch durch In-vitro-Verfahren synthetisiert werden können) werden (mittels rekombinanter Expression oder einer In-vitro-Peptidyl-Bindung) mit einem immunogenen Polpeptid fusioniert und dieses Fusionspolypeptid seinerseits dazu verwendet, Antikörper gegen ein HRG-Epitop zu bilden. Anti-HRG-Antikörper werden aus dem Serum immunisierter Tiere gewonnen. Alternativ dazu werden monoklonale Antikörper aus Zellen in vitro oder aus in vivo immunisierten Tieren auf herkömmliche Weise produziert. Durch Routine-Screening identifizierte bevorzugte Antikörper binden sich an HRG, sie kreuzreagieren aber im Wesentlichen nicht mit anderen bekannten Liganden, wie z. B. EGF, und sie verhindern die Aktivierung von $p185^{HER2}$ durch HRG. Außerdem werden Anti-HRG-Antikörper selektiert, die zur spezifischen Bindung an individuelle Familienmitglieder der HRG-Familie fähig sind, z. B. HRG-α, HRG-β1, HRGβ2, HRG-β3, wodurch sie als spezifische Antagonisten davon dienen können.

[0019] HRG wird auch in vitro abgeleitet, um immobilisiertes HRG und markiertes HRG herzustellen, besonders zwecks Diagnose von HRG oder seinen Antikörpern oder für die Affinitätsreinigung von HRG-Antikörpern. Immobilisierte Anti-HRG-Antikörper eignen sich zur Diagnose (in vitro oder in vivo) oder Reinigung von HRG. In einer bevorzugten Ausführungsform wird ein Gemisch von HRG und anderen Peptiden über eine Säule geschickt, an der die Anti-HRG-Antikörper gebunden sind.

[0020] Substitutions-, Deletions- oder Insertionsvarianten von HRG werden durch In-vitro- oder Rekombinationstechniken erzeugt und z. B. auf Immunkreuzreaktivtät mit den nativen Formen von HRG sowie auf HRG-Antagonisten- oder -Agonisten-Aktivität getestet.

[0021] In einer weiteren bevorzugten Ausführungsform wird HRG zur Stimulierung der Aktivität von $p185^{HER2}$ in normalen Zellen verwendet. In einer weiteren bevorzugten Ausführungsform wird eine Variante von HRG als Antagonist verwendet, um die Stimulierung von $p185^{HER2}$ zu hemmen.

[0022] HRG, seine Derivate oder seine Antikörper sind zu physiologisch annehmbaren Vehikeln formuliert, insbesondere zur therapeutischen Nutzung. Solche Vehikel sind z. B. Depotformulierungen von HRG oder HRG-Varianten. Es werden auch eine Zusammensetzung, die HRG und einen pharmazeutisch annehmbaren Träger umfasst, und ein isoliertes Polypeptid, das mit einem heterologen Polypeptid fusioniertes HRG umfasst, bereitgestellt.

[0023] In weiteren Aspekten bietet die Erfindung eine für HRG kodierende isolierte Nucleinsäure, die mit einer detektierbaren Gruppe markiert oder unmarkiert sein kann, und eine Nucleinsäuresequenz, die zu einer für ein HRG kodierenden Nucleinsäuresequenz komplementär ist oder unter strengen Bedingungen an sie hybridisiert.

[0024] Die Nucleinsäuresequenz eignet sich auch in Hybridisierungsassays hinsichtlich HRG-Nucleinsäure und in einem Verfahren zur Bestimmung der Gegenwart eines HRG, umfassend das Hybridisieren der DNA (oder RNA), die für ein HRG kodiert oder komplementär dazu ist, an eine Testproben-Nucleinsäure und das Bestimmen der Gegenwart eines HRG. Die Erfindung betrifft auch ein Verfahren zum Amplifizieren einer Nucleinsäure-Testprobe, umfassend das Primen einer Nucleinsäure-Polymerase- (Ketten-) Reaktion mit Nucleinsäure (DNA oder RNA), die für ein HRG kodiert oder komplementär dazu ist.

[0025] In weiteren Aspekten ist die Nucleinsäure DNA und umfasst außerdem einen replizierbaren Vektor, der die Nucleinsäure umfasst, die für ein HRG kodiert und operabel mit Steuersequenzen verbunden ist, die von einem durch den Vektor transformierten Wirt erkannt werden; die mit dem Vektor transformierten Wirtszellen; und ein Verfahren zur Verwendung einer für ein HRG kodierenden Nucleinsäure, um die Produktion von HRG zu bewirken, umfassend die Expression von HRG-Nucleinsäure in einer Kultur der transformierten Wirtszellen und die Gewinnung eines HRG aus der Wirtszellenkultur.

[0026] In weiteren Ausführungsformen bietet die Erfindung ein Verfahren zur Produktion von HRG, umfassend das Insertieren eines die Transkription modulierenden Elements in ausreichender Nähe und Orientierung zu einer HRG-Nucleinsäure, um deren Transkription zu beeinflussen (zu unterdrücken oder zu stimulieren), in die DNA einer Zelle, die die für ein HRG kodierende Nucleinsäure enthält, sowie gegebenenfalls den Schritt des Kultivierens der Zelle, die das Transkriptions-Modulationselement und eine HRG-Nucleinsäure enthält.

[0027] In weiteren Ausführungsformen bietet die Erfindung eine Zelle, umfassend die für ein HRG kodierende Nucleinsäure und ein exogenes Transkriptions-Modulationselement in ausreichender Nähe und Orientierung zu einer HRG-Nucleinsäure, um deren Transkription zu beeinflussen; und eine Wirtszelle, umfassend die Nucleinsäure, die für ein HRG kodiert und operabel mit von der Wirtszelle erkannten exogenen Steuersequenzen verbunden ist.

Kurzbeschreibung der Abbildungen

Fig. 1 Reinigung von Heregulin auf Polyasparaginsäure-Säule

[0028] Es wurde eine Polyasparaginsäure-Säulenchromatographie auf HRG-α durchgeführt und das Elutionsprofil von Proteinen anhand der A_{214} gemessen. Der 0,6 M NaCl-Pool aus dem Heparin-Sepharose-Reinigungsschritt wurde auf 0,2 M NaCl mit Wasser verdünnt und auf die in 17 mM Na-Phosphat, pH 6,8 mit 30% Ethanol äquilibrierte Polyasparaginsäure-Säule geladen. Ein linearer NaCl-Gradient aus 0,3 bis 0,6 M wurde zum Zeitpunkt 0 initiiert und war nach 30 Minuten abgeschlossen. Fraktionen wurden im HRG-Tyrosin-Autophosphorylierungsassay untersucht. Die Fraktionen, die Peak C entsprachen, wurden zwecks weiterer Reinigung auf C4-RP-HPLC gepoolt.

Fig. 2 C4-Reinigung von HRG-2 mit umgekehrten Phasen

[0029] Bild A: Pool C aus der Polyasparaginsäure-Säule wurde auf eine C4-HPLC-Säule (Syn-Chropak RP-4) aufgebracht, äquilbriert in 0,1% TFA, und die Proteine wurden mit einem linearen Acetonitrilgradienten mit 0,25%/min eluiert. Es ist die Extinktionsspur für den mit C4-17 bezeichneten Lauf dargestellt. Für den Test wurden 1 ml-Fraktionen abgenommen.
[0030] Bild B: 10 µl-Aliquoten der Fraktionen wurden in HRG-Tyrosin-Autophosphorylierungstest untersucht. Die Werte von Phosphotyrosin im $p185^{HER2}$-Protein wurden durch einen spezifischen Anti-Phosphotyrosin-Antikörper quantifiziert und in wahllosen Einheiten auf der Abszisse angezeigt.
[0031] Bild C: 10 µl-Fraktionen wurden entnommen und SDS-Gelelektrophorese auf 4–20 Acrylamid-Gradient-Gelen gemäß dem Verfahren von Laemmli (Nature 227, 680–685 [1970]) unterzogen. Die Molekulargewichte der Standardproteine sind links von der die Standards enthaltenden Spur angezeigt. Der Hauptpeak von Tyrosin-Phosphorylierungsaktivität in Fraktion 17 war mit einer prominenten 45.000 Da großen Bande (HRG-a) assoziiert.

Fig. 3 SDS-Polyacrylamidgel, das Reinigung von HRG-α aufweist

[0032] Die Molekulargewichtsmarker sind in Spur 1 zu sehen. Aliquoten aus dem mit MDA-MB-231 konditioniertem Medium (Spur 2), dem 0,6 M NaCl-Pool aus der Heparin-Sepharose-Säule (Spur 3), Spur C aus der Polyasparaginsäure-Säule (Spur 4) und Fraktion 17 aus der HPLC-Säule (C4-17) (Spur 5) wurden Elektrophorese auf einem 4–20% Gradienten-Gel unterzogen und mit Silber eingefärbt. Die Spuren 6 und 7 enthielten nur Puffer und zeigen die Gegenwart von Gel-Artefakten im 50–65 kDa-Bereich.
[0033] Die **Fig.** 4a–4d zeigen die abgeleitete Aminosäuresequenz der cDNA, die in λgt_{10}her16 (SEQ.-ID Nr. 12 und SEQ.-ID Nr. 13) enthalten ist. Die Nucleotide sind oben links von jeder Zeile nummeriert und die Aminosäuren in aus drei Buchstaben bestehenden Codes unten links von jeder Zeile nummeriert. Die Nucleotidsequenz, die mit der Sonde entspricht, umfasst Nucleotide 681–720. Die wahrscheinliche Transmembran-Domäne sind die Aminosäuren 287–309. Die 6 Cysteine des EGF-Motivs sind 226, 234, 240, 254, 256 und 265. Die 5 potenziellen aus 3 Aminosäuren bestehenden N-gebundenen Glykosylierungsstellen sind 164–166, 170–172, 208–210, 437–439 und 609–611. Die potenziellen Serin-Threonin-O-Glykosylierungsstellen sind 209–221. Potenzielle Serin-Glycin-Dipeptid-Glykosaminoglykan-Additionsstellen sind Aminosäuren 42–43, 64–65 und 151–152. Das Anfangs-Methioonin (MET) befindet sich an Position 45 in **Fig.** 4, obwohl der verarbeitete N-terminale Rest S46 ist.

Fig. 5 Northern Blot-Analyse von MDA-MB-231 und SKBR3-RNA

[0034] Markiert von links nach rechts sind: 1) MDA-MB-231 polyA minus-RNA (RNA nach Entfernung von polyA-enthaltender RNA); 2) MDA-MB-231 polyA plus-mRNA (RNA mit polyA); 3) SKBR3 polyA minus-RNA; und 4) SKBR3 polyA plus-mRNA. Die für diese Analyse verwendete Sonde war ein radioaktiv (^{32}P) markiertes internes xhol-DNA-Restriktions-Endonuclease-Fragment aus dem cDNA-Abschnitt von λgt_{10}her16.

Fig. 6 Sequenzvergleiche in der EGF-Proteinfamilie

[0035] Die Sequenzen verschiedener EGF-ähnlicher Proteine (SEQ.-ID Nr. 14, 15, 16, 17, 18 und 19) um die Cysteindomäne wurden mit der Sequenz von HRG-α ausgerichtet. Die Position in **Fig.** 6 der Cysteine und die Invarianten Glycin- und Argininreste an Positionen 238 und 264 zeigen klar auf, dass HRG-α ein Mitglieder der EGF-Familie ist. Die Region in **Fig.** 6 mit der höchsten Aminosäure-Identität der Familienmitglieder relativ zu HRG-α (30–40%) befindet sich zwischen Cys 234 und Cys 265. Die stärkste Identität (40%) befindet sich zwischen der Heparin-Bindungs-EGF (HB-EGF-) Spezies. HRG-α besitzt ein einzelnes 3-Aminosäuren-Insert zwischen Cys 240 und Cys 254. Potenzielle Transmembran-Domänen sind mit Kästchen umrahmt (287–309). Die Balken zeigen die carboxyterminalen Stellen für EGF und TGF-α an, worin proteolytische Spaltung die reifen Wachstumsfaktoren von ihren Transmembran-assoziierten Proformen abhebt. HB-EGF ist Heparinbindungs-Epidermis-Wachstumsfaktor; EGF ist Epidermis-Wachstumsfaktor; TGF-α ist transformierender Wachstumsfaktor α; und Schwannoma ist Schwannoma-abgeleiteter Wachstumsfaktor. Die restlichen Nummern in **Fig.** 6 entsprechen dem in **Fig.** 4 angewendeten System.

Fig. 7 Stimulierung von Zellwachstum durch HRG-α

[0036] Drei unterschiedliche Zelllinien wurden auf Wachstumsreaktionen auf 1 nM HRG-α untersucht. Zellprotein wurde durch Kristallviolett-Einfärbung quantifiziert, und die Reaktionen wurden auf unbehandelte Vergleichszellen normalisiert.

[0037] **Fig.** 8a–8d (SEQ.-ID Nr. 7) zeigen die gesamte potenzielle Kodierungs-DNA-Nucleotidsequenz von HRG-β1 und die abgeleitete Aminosäuresequenz der in λher11.1 dbl enthaltenen cDNA (SEQ.-ID Nr. 9). Die Nucleotide sind oben links von jeder Zeile nummeriert und die Aminosäuren in ihren aus 3 Buchstaben bestehenden Codes unten links von jeder Zeile nummeriert. Die wahrscheinliche Transmembran-Aminosäuredomäne umfasst die Aminosäuren 278–300. Die 6 Cysteine des EGF-Motifs sind 212, 220, 226, 240, 242 und 251. Die 5 potenziellen aus 3 Aminosäuren bestehenden N-gebundenen Glykosylierungsstellen sind 150–152, 156–158, 196–198, 428–430 und 600–612. Die potenziellen Serin-Threonin-O-Glykosylierungsstellen sind 195–207. Potenzielle Serin-Glycin-Dipeptid-Glykosaminoglykan-Additionsstellen sind Aminosäuren 28–29, 50–51 und 137–138. Das Anfangs-Methionin (MET) befindet sich an Position 31. HRG-β1 ist zum N-terminalen Rest S32 verarbeitet.

[0038] **Fig.** 9 zeigt einen Vergleich der Aminosäuresequenzen von HRG-α und HRG-β1. Ein Bindestrich (-) zeigt an, dass sich keine Aminosäure an dieser Position befindet (SEQ.-ID Nr. 8 und SEQ.-ID Nr. 9). Diese Abbildung verwendet das Nummerierungssystem der **Fig.** 4 und 6.

[0039] **Fig.** 10 zeigt die Stimulierung von HER2-Autophosphorylierung unter Verwendung von rekombinantem HRG-α (gemessen durch HER2-Tyrosin-Phosphorylierung).

[0040] **Fig.** 11 veranschaulicht die Nucleotid- und vermeintliche Aminosäuresequenz von λ15'her13 (SEQ.-ID Nr. 22); das Aminosäurerest-Nummerierungssystem wird nur in dieser Abbildung angewendet.

[0041] Die **Fig.** 12a–12e stellen die Nucleotidsequenz von λher 76 dar, das für HRG-β2 kodiert (SEQ.-ID Nr. 23). In dieser Abbildung beginnt die Aminosäurerest-Nummerierung mit dem exprimierten N-terminalen MET; der N-Terminus ist S2.

[0042] Die **Fig.** 13a–13c zeigen die Nucleotidsequenz von λher78 (kodiert für HRG-β3; SEQ.-ID Nr. 24). Diese Abbildung wendet das Aminosäure-Nummerierungssystem von **Fig.** 12 an; S2 ist der verarbeitete N-Terminus.

[0043] Die **Fig.** 14a–14d veranschaulichen die Nucleotidsequenz von λher84 (kodiert für HRGβ2-ähnliches Polypeptid; SEQ.-ID Nr. 25). die Abbildung verwendet das Aminosäure-Nummerierungssystem von **Fig.** 12; S2 ist der verarbeitete N-Terminus.

[0044] Die **Fig.** 15a–15c zeigen Aminosäure-Homologien zwischen den bekannten Heregulinen (α, β1, β2, β2-ähnlich und β3 in absteigender Reihenfolge) und veranschaulichen Aminosäure-Insertionen, -Deletionen oder -Substitutionen, die die unterschiedlichen Formen ausmachen (SEQ.-ID Nr. 26–30). Diese Abbildung wendet das Aminosäure-Nummerierungssystem der **Fig.** 12–14 an.

Ausführliche Beschreibung der bevorzugten Ausführungsformen

1. Definitionen

[0045] Im Allgemeinen besitzen die folgenden Wörter oder Phrasen bei Verwendung in der vorliegenden Beschreibung, in den Beispielen und in den Patentansprüchen die angeführte Definition.

[0046] Heregulin („HRG") ist hierin als jede isolierte Polypeptidsequenz definiert, die die biologische Aktivität eines in **Fig.** 4, 8, 12, 13 oder 15 dargestellten Polypeptids und von Fragmenten, Allelen oder tierischen Anlogen davon aufweist. HRG schließt jedes bislang identifizierte Polypeptid aus, z. B. jedes bekannte Polypeptid,

das gemäß 35 U. S. C. 102 denkbar wäre, insbesondere EFG, TFG-α, Amphiregulin (Plowman et al., Mol. Cell. Biol. 10, 1969 [1990], HB-EGF (Higashimaya et al., Science 251, 936 [1991]), Schwannoma-Faktor oder Polypeptide, die für Fachleute auf dem Gebiet der Erfindung offenkundig sind.

[0047] Unter „biologischer Aktivität" ist hierin ein In-vivo-Effektor oder eine antigene Funktion zu verstehen, die direkt durch ein HRG-Polypeptid (in seiner nativen oder denaturierten Konformation) oder durch jede Subsequenz davon durchgeführt wird. Effektorfunktionen umfassen Rezeptorbindung oder -Aktivierung, Induktion von Differenzierung, mitogener oder wachstumsfördernder Aktivität, Immunmodulation, DNA-Regulationsfunktionen u. dgl., ob sie nun derzeit bekannt oder inhärent sind. Antigene Funktionen sind z. B. der Besitz einer epitopen oder antigenen Stelle, die zur Kreuzreaktion mit Antikörpern gegen ein natürlich vorkommendes oder denaturierten HRG-Polypeptid oder Fragment davon fähig ist.

[0048] Biologisch aktives HRG umfasst Polypeptide mit Effektor- und antigener Funktion oder nur einer dieser Funktionen. HRG umfasst antagonistische Polypeptide gegen HRG, sofern solche Antagonisten ein Epitop eines native HRG enthalten. Eine wesentliche bekannte Effektorfunktion von HRG ist seine Fähigkeit, sich an $p185^{HER2}$ zu binden und die Rezeptor-Tyrosin-Kinase zu aktivieren.

[0049] HRG umfasst die translatierte Aminosäuresequenz von Human-HRG voller Länge (pro HRG; hierin in den Abbildungen dargestellt); deglykosylierte oder unglykosylierte Derivate; Aminosäuresequenz-Varianten; und kovalente Derivate von HRG, sofern sie biologische Aktivität aufweisen. Zwar ist die native Proform von HRG wahrscheinlich ein Membran-gebundenes Polypeptid, doch sind auch lösliche Formen, z. B. Formen ohne funktionale Transmembrandomäne (proHRG oder seine Fragmente), in dieser Definition enthalten.

[0050] Die Definition von HRG gilt auch für Fragmente von intaktem HRG. Zwei Hauptdomänen sind innerhalb der Fragmente enthalten. Es handelt sich um die Wachstumsfaktordomäne („GFD"), die homolog zur EGF-Familie ist und sich etwa an Resten 5216-A227 bis N268-R286 (**Fig.** 9, HRG-α; die GFD-Domänen für andere HRG [**Fig.** 15] sind homologe Sequenzen) befindet. Vorzugsweise sind die GFD für HRG-α, HRG-β1, HRG-β2, HRG-β2-ähnlich und HRG-β3 G175-K241, G175-K246, G175-K238, G175-K238 bzw. G175-E241 (**Fig.** 15).

[0051] Ein weiteres Fragment von Interesse ist die N-terminale Domäne („NTD"). Die NTD erstreckt sich vom N-Terminus des verarbeiteten HRG (S2) bis zum Rest, der an einen Nterminalen Rest der GFD angrenzt, d. h. etwa T172-C182 (**Fig.** 15), vorzugsweise T174. Eine zusätzliche Gruppe von Fragmenten umfasst die NTD-GFD-Domänen, die das Äquivalent der extrazellulären Domänen von HRG-α und β_1-β_2 sind. Ein weiteres Fragment ist das C-terminale Peptid („CTP"), das sich etwa 20 Reste N-terminal zum ersten Rest der Transmembran-Domäne befindet (entweder alleine oder in Kombination mit dem C-terminalen Rest des HRG).

[0052] In bevorzugten Ausführungsformen ist antigenisch aktives HRG ein Polypeptid, das sich mit einer Affinität von zumindest etwa 10^7 l/Mol an einen Antikörper gegen eine natürlich vorkommende HRG-Sequenz bindet. Üblicherweise bindet sich das Polypeptid mit einer Affinität von zumindest etwa 10^8 l/mol. Am bevorzugtesten ist das antigenisch aktive HRG ein Polypeptid, das sich an einen Antikörper gegen eines der HRG in seiner nativen Konformation bindet. HRG in seiner nativen Konformation ist im Allgemeinen HRG, wie es in der Natur vorkommt – nicht denaturiert durch chaotrope Mittel, Wärme oder eine andere Behandlung, die die dreidimensionale Struktur von HRG substanziell verändert (bestimmt z. B. durch Migration nicht-reduzierender, nicht-denaturierender Größenklassierungsgele). In dieser Bestimmung verwendeter Antikörper ist polyklonaler Kaninchen-Antikörper, der durch Formulierung von nativem HRG aus einer Nicht-Kaninchen-Spezies in komplettem Freund-Adjuvans, subkutane Injektion der Formulierung in Kaninchen und Wiederauffrischen der Immunreaktion durch intraperitoneale Injektion der Formulierung, bis der Titer des Anti-HRG-Antikörpers ein Plateau erreicht, gebildet wird).

[0053] Üblicherweise besitzt biologisch aktives HRG eine Aminosäuresequenz von zumindest 75% Aminosäuresequenz-Identität mit einer HRG-Sequenz, noch bevorzugter mit zumindest 80%, noch bevorzugter von zumindest 90%, am bevorzugtesten von zumindest 95% Identität. Identität oder Homologie in Bezug auf ein HRG-Sequenz ist hierin als Prozentsatz von Aminosäureresten in der Kandidatensequenz definiert, die mit HRG-Resten in **Fig.** 15 identisch sind – nach dem Ausrichten von Sequenzen von Einführen von Zwischenräumen, falls sich dies als notwendig erweist, um maximale prozentuelle Homologie zu erzielen, wobei man konservative Substitutionen nicht als identische Reste auffasst. Keine der N-terminalen, C-terminalen oder internen Verlängerungen, Deletionen oder Insertionen in die HRG-Sequenz soll so ausgelegt werden, dass sie die Homologie beeinflusst.

[0054] Somit umfassen die biologisch aktiven HRG-Polypeptide der Erfindung jeweils exprimierte oder verarbeitete HRG-Sequenz; Fragmente davon mit einer konsekutiven Sequenz von zumindest 15, 20, 25, 30 oder 40 Aminosäureresten; Aminosäuresequenz-Varianten von HRG, worin ein Aminosäurerest N- oder C-terminal zur oder innerhalb der HRG-Sequenz bzw. zum oder innerhalb ihres Fragments insertiert wurde (Definition siehe oben); Aminosäuresequenz-Varianten von HRG-Sequenz oder ihrem Fragment (siehe oben), worin ein Rest durch einen anderen Rest substituiert wurde. HRG-Polypeptide umfassen jene, die vorbestimmte Mutationen besitzen, z. B. durch stellengerichtete oder PCR-Mutagenese. HRG umfasst HRG aus einer Spezies wie z. B. Kaninchen, Ratten, Schweinen, Primaten (ausschließlich des Menschen), Pferden, Mäusen und

Schafen sowie Allele oder andere natürlich vorkommende Varianten der obigen Substanzen; Derivate von HRG oder seiner Fragmente (Definition siehe oben), worin HRG oder seine Fragmente durch Substitution, chemische, enzymatische oder andere geeignete Techniken mit einer anderen Gruppe als einer natürlich vorkommenden Aminosäure kovalent modifiziert wurde(n) (z. B. mit einer detektierbaren Gruppe wie etwa einem Enzym oder Radioisotop); Glykosylierungsvarianten von HRG (Insertion einer Glykosylierungsstelle oder Deletion einer beliebigen Glykosylierungsstelle durch Deletion, Insertion oder Substitution eines geeigneten Rests); und lösliche Formen von HRG, z. B. HRG-GFD oder jene, die keine funktionale transmembrane Domäne aufweisen.

[0055] Von besonderem Interesse sind Fusionsproteine, die HRG-NTD enthalten, aber frei von GFD sind, die üblicherweise mit der vorliegenden HRG-NTD assoziiert sind. Die ersten 23 Aminosäuren der NTD sind von geladenen Resten dominiert und enthalten eine Sequenz (GKKKER; Reste 13–18, **Fig.** 15), die eine starke Ähnlichkeit zum Konsenssequenz-Motiv für das nucleare Targeting aufweist (Roberts, Biochim. Biophys. Acta 1008, 263 [1989]). Demzufolge umfasst das HRG Fusionen, in denen die NTD oder zumindest ein Polypeptid mit den ersten etwa 23 Resten an einem Terminus mit einem Nicht-HRG-Polypeptid oder einer GFD eines anderen HRG-Familienmitglieds fusioniert ist. Das Nicht-HRG-Polypeptid ist in der vorliegenden Ausführungsform ein regulatorisches Protein, ein Wachstumsfaktor wie z. B. EGF oder TGF-α oder ein Polypeptidligand, der sich an einen Zellrezeptor bindet, insbesondere einen Zelloberflächenrezeptor, der sich auf der Oberfläche einer Zelle befindet, deren Regulation erwünscht ist, z. B. einer Krebszelle.

[0056] In einer weiteren Ausführungsform ist/sind ein oder mehr Reste 13–18 unabhängig voneinander variiert, um eine Sequenz zu produzieren, die für das nucleare Targeting nicht in Frage kommt. Beispielsweise ist G13 zu einem anderen natürlich vorkommenden Rest mutiert, z. B. P, L, I, V, A, M, F, K, D oder S; ein oder mehrere von K14-K16 ist/sind zu anderen natürlich vorkommenden Resten mutiert, z. B. D, R, K, N, N oder Q; und R18 ist zu einem anderen natürlich vorkommenden Rest mutiert, z. B. K, N, D, E, N oder Q. Alle oder beliebige dieser Reste 13–18 sind auch deletiert, oder es sind externe Reste angrenzend an diese Reste insertiert; z. B. Reste, die angrenzend an Reste 13–18 insertiert sind, die die gleichen wie die oben angeführten Substitutionen für die Reste selbst sind.

[0057] In einer weiteren Ausführungsform sind Enzyme oder ist ein nucleares Regulationsprotein wie z. B. ein transkriptionaler Regulationsfaktor mit HRG-NTR, HRG-NTD-GFD oder HRG-GFD fusioniert. Das Enzym oder der Faktor ist mit dem N- oder C-Terminus fusioniert, zwischen der NTD- und GFD-Domäne insertiert oder für die Region von NTD zwischen den ersten etwa 23 Resten und der GFD substituiert.

[0058] „Isoliertes" HRG bezieht sich hierin auf HG, das identifiziert wurde und frei von Komponenten seiner natürlichen Umgebung ist. Kontaminierende Komponenten seiner natürlichen Umgebung sind z. B. Materialien, die die diagnostischen oder therapeutischen Nutzungen von HRG beeinträchtigen würden, beispielsweise Proteine, Hormone und andere Substanzen. In bevorzugten Ausführungsformen ist HRG solcherart gereinigt, dass (1) es mehr als 95 Gew.-%, am bevorzugtesten mehr als 99 Gew.-%, Protein enthält (bestimmt durch das Lowry-Verfahren oder ein anderes validiertes Proteinbestimmungs-Verfahren); (2) man einen ausreichenden Reinigungsgrad erzielt, um zumindest 15 Reste N-terminaler oder interner Aminosäuresequenz unter Verwendung des bis dato besten im Handel verfügbaren Aminosäure-Sequenators zu erhalten; oder (3) Homogenität mittels SDS-PAGE und Coomassieblau oder vorzugsweise Silbereinfärbung erzielt wird. Zu isoliertem HRG zählt HRG in situ innerhalb heterologer rekombinanter Zellen, da zumindest eine Komponente aus der natürlichen Umgebung von HRG nicht vorhanden ist. Isoliertes HRG umfasst HRG aus einer Spezies in rekombinanter Zellkultur einer anderen Spezies, da HRG unter diesen Umständen keine Quellen-Polypeptide aufweist. Üblicherweise wird jedoch isoliertes HRG durch zumindest einen Reinigungsschritt erzeugt.

[0059] Gemäß der Erfindung ist HRG-Nucleinsäure RNA oder DNA mit mehr als 10 Basen, die für biologisch oder antigenisch aktives HRG kodiert, komplementär zur Nucleinsäuresequenz ist, die für ein solches HRG kodiert, oder an Nucleinsäuresequenz hybridisiert, die für solches HRG kodiert, und unter strengen Bedingungen stabil daran gebunden bleibt.

[0060] Vorzugsweise kodiert HRG-Nucleinsäure für ein Polypeptid, das zumindest 75%, noch bevorzugter zumindest 80%, noch bevorzugter zumindest 85%, noch bevorzugter zumindest 90%, am bevorzugtesten zumindest 95%, Sequenzidentität mit einer HRG-Sequenz aufweist (siehe **Fig.** 4, 8, 12, 13 oder 15). Vorzugsweise enthält die hybridisierende HRG-Nucleinsäure etwa 20, noch bevorzugter zumindest etwa 40, am bevorzugtesten zumindest etwa 90, Basen. Eine derartige hybridisierende oder komplementäre Nucleinsäure schließt jedoch Nucleinsäure aus, die für EGF, TGF-α, Amphiregulin, HB-EGF, Schwannoma-Faktor oder Fragmente oder Varianten davon kodiert, die für Fachleute auf dem Gebiet offenkundig sind.

[0061] Isolierte NRG-Nucleinsäure ist z. B. Nucleinsäure, die frei von zumindest einer kontaminierenden Nucleinsäure ist, mit der sie üblicherweise in der natürlichen Quelle von HRG-Nucleinsäure assoziiert ist. Isolierte HRG-Nucleinsäure ist somit in einer anderen Form oder Umgebung vorhanden, in der sie in der Natur vorliegt. Isolierte HRG-kodierende Nucleinsäure enthält jedoch NRG-Nucleinsäure in üblicherweise HRG-exprimierenden Zellen, worin die Nucleinsäure an einer chromosomalen Position vorliegt, die sich von jener natürlicher Zellen unterscheidet, oder durch eine andere DNA-Sequenz flankiert ist als in der Natur. HRG-kodierende Nucle-

insäure kann in spezifischen Hybridisierungstests verwendet werden, insbesondere jene Abschnitte von HRG-kodierender Sequenz, die mit anderen bekannten DNA-Sequenzen nicht hybridisieren, z. B. jene, die für die EGF-ähnlichen Moleküle von **Fig.** 6 kodieren.

[0062] „Strenge Bedingungen" sind jene, die (1) geringe Ionenstärke und hohe Temperatur für das Waschen von z. B. 0,015 M NaCl/0,0015 M Natriumcitrat/0,1% $NaDodSO_4$ bei 50 °C vorsehen; (2) während der Hybridisierung ein Denaturierungsmittel wie z. B. Formamid vorsehen, z. B. 50% (v/v) Formamid mit 0,1% Rinderserumalbumin, 0,1% Ficoll, 0,1% Polyvinylpyrrolidon, 50 mM Natriumphosphatpuffer, pH 6,5 mit 750 mM NaCl, 75 mM Natriumcitrat bei 42°C; oder (3) 50% Formamid, 5 × SSC (0,75 M NaCl, 0,075 M Natriumcitrat), 50 mM Natriumphosphat (pH 6,8), 0,1 % Natriumpyrophosphat, 5 × Denhardt-Lösung, ultraschallbehandelte Lachssperma-DNA (50 g/ml), 0,1% SDS und 10% Dextransulfat bei 42°C vorsehen, wobei die Waschungen bei 42°C in 0,2 × SSC und 0,1% SDS erfolgen.

[0063] Beispiele für HRG-α-Nucleinsäuren sind Nucleinsäuren oder Oligonucletoide, die aus einer aus **Fig.** 4a–4d ausgewählten Nucleotidsequenz bestehen, umfassend mehr als 17 Basen (ausschließlich Nucleinsäuresequenzen kleiner polydisperser zirkulärer Human-DNA [HUMPC125], Hühner-c-mos-Protoonkogen-Homolog (CHKMOS), Basismembran-Heparinsulfat-Proteoglykan [HUMBMHSP] und Human-Lipocortin 2-Pseudogen (komplette cds-Region, HUMLIP2B]), üblicherweise mehr als 20 Basen, vorzugsweise mehr als 25 Basen, gemeinsam mit den komplementären Sequenzen davon.

[0064] Beispiele für HRG-β_1-, -β_2- oder -β_3-Nucleinsäuren sind Nucleinsäuren oder Oligonucleotide, die aus einer aus den **Fig.** 8a–8d, 12a–12e oder 13a–13c ausgewählten Nucleotidsequenz bestehen, umfassend mehr als 20 Basen (unter Ausschuss der polyA-Sequenz am 3'-Ende jedes Gens, siehe Abbildungen), gemeinsam mit den Komplementen solcher Sequenzen. Vorzugsweise enthält die Sequenz mehr als 25 Basen. HRG-β-Sequenzen können auch die kleine polydisperse zirkuläre DNA-Sequenz (HUMP-C125) ausschließen.

[0065] In anderen Ausführungsform enthält die HRG-Nucleotidsequenz eine aus 15 oder mehr Basen bestehende HRG-Sequenz und ist ausgewählt aus der Sequenz, die für die HRG-Domäne kodiert (erstreckt sich vom N-Terminus der GFD bis zum N-Terminus der Transmembran-Sequenz oder dem Komplement dieser Nucleinsäuresequenz). In Bezug auf HRG-α ist z. B. die Nucletoidsequenz aus der Sequenz 678–869 (**Fig.** 4b) ausgewählt und enthält eine Sequenz von 15 oder mehr Basen aus diesem Abschnitt der HRG-Nucleinsäure.

[0066] In anderen Ausführungsformen ist die HRG-Nucleinsäuresequenz mehr als 14 Basen groß und ist aus einer Nucleotidsequenz ausgewählt, die für jeden Subtyp charakteristisch ist, z. B. aus einer Nucleinsäuresequenz, die für eine Aminosäuresequenz kodiert, die für jeden der HRG-Subtypen charakteristisch ist (oder das Komplement, dieser Nucleinsäuresequenz). Diese Sequenzen eignen sich in diagnostischen Assays für die Expression verschiedener Subtypen sowie für die spezifische Amplifikation der Subtyp-DNA. Beispielsweise würde die HRG-α-Sequenz von Interesse aus der Sequenz ausgewählt sein, die für den einzigen N-Terminus oder die GFD-Transmembran-Verbindungssequenz kodiert, z. B. etwa bp771-860. Ebenso ist eine einzigartige HRG-β_1-Sequenz eine, die für die letzten 15 C-terminalen Aminosäurereste kodiert; diese Sequenz findet sich in HRG-α nicht.

[0067] Im Allgemeinen ist die Länge der HRG-α- oder HRG-β-Sequenz über die oben angeführte Basenanzahl hinaus unwesentlich, da alle derartigen Nucleinsäuren als Sonden oder Amplifikationsprimer geeignet sind. Die ausgewählte HRG-Sequenz kann zusätzliche HRG-Sequenz (entweder die normale Flankierungssequenz oder andere Regionen der HRG-Nucleinsäure) sowie andere Nucleinsäuresequenzen enthalten. Für die Zwecke der Hybridisierung ist nur die HRG-Sequenz entscheidend.

[0068] Der Ausdruck „Steuersequenzen" bezieht sich hierin auf DNA-Sequenzen, die für die Expression einer operabel verbundenen Kodiersequenz in einem konkreten Wirtsorganismus notwendig sind. Die Steuersequenzen, die sich für Prokaryoten eignen, sind z. B. eine Promotor-, gegebenenfalls eine Operatorsequenz, eine Ribosomen-Bindungsstelle und möglicherweise noch relativ unbekannte andere Sequenzen. Man weiß, dass eukaryotische Zellen Promotoren, Polyadenylationssignale und Enhancer verwenden.

[0069] Nucleinsäure ist „operabel verbunden", wenn sie in eine funktionale Beziehung mit einer anderen Nucleinsäuresequenz gebracht wird. Beispielsweise ist DNA für eine Prosequenz oder einen sekretorischen Leader operabel mit DNA für ein Polypeptid verbunden, wenn sie als Vorprotein exprimiert ist, das an der Sekretion des Polypeptids beteiligt ist; ein Promotor oder Enhancer (Verstärker) ist operabel mit einer Kodiersequenz verbunden, wenn er die Transkription der Sequenz beeinfluss T; oder eine Ribosomen-Bindungsstelle ist operabel mit einer Kodiersequenz verbunden, wenn er so positioniert ist, dass er die Translation erleichtert. Im Allgemeinen bedeutet „operabel verbunden", dass die verbundenen DNA-Sequenzen zusammenhängend und – im Fall eines sekretorischen Leaders – zusammenhängend und in Lesephase angeordnet sind. Enhancer müssen allerdings nicht zusammenhängend sein. Die Verbindung erfolgt durch Ligation an geeigneten Restriktionsstellen. Wenn solche Stellen nicht existieren, werden synthetische Oligonucleotid-Adaptoren oder -Linker in Einklang mit konventionellen Verfahren verwendet.

[0070] Ein „exogenes" Element ist hierin als Nucleinsäuresequenz definiert, die für die Zelle fremd oder zur Helle homolog ist, aber sich an einer Position innerhalb der Wirtszellen-Nucleinsäure befindet, an der das Element üblicherweise nicht vorkommt.

[0071] Die Ausdrücke „Zelle", „Zelllinie" und „Zellkultur" werden hierin austauschbar verwendet, wobei alle diese Termini auch die Nachkommenschaft umfassen. Die Ausdrücke „Transformanten" und „transformierte Zellen" umfassen hierin die primären daraus abgeleiteten Zellen und Kulturen (ohne Berücksichtigung der Anzahl an Transfers). Es ist ferner zu beachten, dass die gesamte Nachkommenschaft hinsichtlich des DNA-Gehalts möglicherweise nicht genau identisch ist, was auf beabsichtigte oder unbeabsichtigte Mutationen zurückzuführen ist. Mutanten in der Nachkommenschaft, die die gleiche Funktion oder biologische Aktivität aufweisen, wie sie in der ursprünglich transformierten Zelle gescreent wurden, sind ebenfalls enthalten. Es ergibt sich aus dem Kontext, wo voneinander unterschiedliche Bezeichnungen beabsichtigt sind.

[0072] „Plasmide" werden durch ein klein geschriebenes „p" identifiziert, vor und/oder nach dem Großbuchstaben und/oder Zahlen stehen. Die Ausgangsplasmide sind hierin im Handel erhältlich, öffentlich uneingeschränkt verfügbar oder können aus solchen erhältlichen Plasmiden in Einklang mit veröffentlichten Verfahren konstruiert werden. Außerdem sind andere gleichwertige Plasmide auf dem Gebiet der Erfindung bekannt und für Fachleute offenkundig.

[0073] „Restriktionsenzym-Verdau" von DNA bezieht sich auf die katalytische Spaltung der DNA mit einem Enzym, das nur an bestimmten Positionen in der DNA wirkt. Solche Enzyme werden als Restriktions-Endonucleasen bezeichnet, und die Stellen, für die jede spezifisch ist, nennt man eine Restriktionsstelle. Die verschiedenen hierin verwendeten Restriktionsenzyme sind im Handel erhältlich, wobei die von den Enzymherstellern angegebenen Reaktionsbedingungen, Co-Faktoren und anderen Bedingungen berücksichtigt werden. Restriktionsenzyme sind üblicherweise durch Abkürzungen gekennzeichnet, die aus einem Großbuchstaben, gefolgt von anderen Buchstaben, die den Mikroorganismus repräsentieren, aus dem jedes Restriktionsenzym ursprünglich erhalten wurde, und anschließend einer das jeweilige Enzym bezeichnenden Zahl bestehen. Im Allgemeinen wird etwa 1 µg Plasmid oder DNA-Fragment mit etwa 1–2 Enzymeinheiten in etwa 20 µl Pufferlösung verwendet. Geeignete Puffer und Substratmengen für bestimmte Restriktionsenzyme werden vom Hersteller angegeben. Üblicherweise erfolgt etwa einstündige Inkubation bei 37°C, doch dies kann je nach Anweisungen des Herstellers variieren. Nach der Inkubation wird Protein oder Polypeptid durch Extraktion mit Phenol und Chloroform entfernt und die verdaute Nucleinsäure aus der wässrigen Fraktion durch Ethanol-Fällung gewonnen. An den Verdau mit einem Restriktionsenzym kann sich bakterielle alkalische Phosphatase-Hydrolyse der terminalen 5'-Phosphate anschließen, um zu verhindern, dass die zwei restriktionsgespaltenen Enden eines DNA-Fragments eine geschlossene Schleife bilden („Ringschließung"), die die Insertion anderer DNA-Fragmente an der Restriktionsstelle erschweren würde. Sofern nicht anders angeführt schließt sich an den Verdau von Plasmiden keine 5'-terminale Dephosphorylierung an. Die Verfahren und Reagenzien für die Dephosphorylierung sind konventionell und sind in Kapiteln 1.56–1.61 von Sambrook et al., Molecular Cloning: A Laboratory Manual, New York, Cold Spring Harbor Laboratory Press (1989) beschrieben.

[0074] „Ligation" bezieht sich auf das Verfahren zur Bildung von Phosphodiester-Bindungen zwischen zwei Nucleinsäurefragmenten. Um die DNA-Fragmente miteinander zu ligieren, müssen die Enden der DNA-Fragmente miteinander kompatibel sein. In einigen Fällen sind die Enden nach Endonuclease-Verdau direkt kompatibel. Es kann jedoch notwendig sein, zuerst die versetzten Enden, die nach Endonuclease-Verdau gemeinsam produziert wurden, in stumpfe Enden umzuwandeln, um sie ligationskompatibel zu machen. Um die Enden abzustumpfen, wird die DNA in einem geeigneten Puffer zumindest 15 Minuten lang bei 15°C mit etwa 10 Einheiten Klenow-Fragment von DNA-Polymerase I oder T4 DNA-Polymerase in Gegenwart der vier Desoxyribonucleotidtriphosphate behandelt. Die DNA wird dann durch Phenol-Chloroform-Extraktion und Ethanol-Fällung gereinigt. Die zu ligierenden DNA-Fragmente werden dann in etwa äquimolaren Mengen in Lösung eingebracht. Die Lösung enthält auch ATP, Ligasepuffer und eine Ligase wie z. B. T4 DNA-Ligase (etwa 10 Einheiten pro 0,5 µg DNA). Wenn die DNA in einen Vektor zu ligieren ist, wird der Vektor zunächst durch Verdau mit der bzw. den geeigneten Restriktions-Endonuclease(n) linearisiert. Das linearisierte Fragment wird anschließend mit bakterieller alkalischer Phosphatase oder Kalbsdarm-Phosphatase behandelt, um während des Ligationsschritts Selbstligation zu verhindern.

[0075] Die Technik der „Polymerase-Kettenreaktion" oder „PCR" bezieht sich hierin allgemein auf ein Verfahren, in dem winzige Mengen eines spezifischen Stücks Nucleinsäure, RNA und/oder DNA amplifiziert werden (siehe US-Patent 4.683.195 vom 28. Juli 1987). Im Allgemeinen muss Sequenzinformation von den Enden der Region von Interesse oder darüber hinaus vorliegen, sodass Oligonucleotid-Primer konstruiert werden können; diese Primer sind hinsichtlich ihrer Sequenz mit gegenüberliegenden Strängen des zu amplifizierenden Templats identisch oder ähneln ihnen. Die 5'-terminalen Nucleotide der zwei Primer können mit den Enden des amplifizierten Materials zusammenfallen. PCR kann dazu dienen, spezifische RNA-Sequenzen, spezifische DNA-Sequenzen aus genomischer Voll-DNA und cDNA, transkribiert aus zellulärer Voll-RNA, Bakteriophagen oder Plasmidsequenzen usw. zu amplifizieren. Siehe für einen allgemeinen Überblick Mullis et al., Cold Spring Harbor Symp. Quant. Biol. 51, 263 (1987); Erlich, Hg., PCR Technology, Stockton Press, NY, 1989. PCR wird hierin als ein – aber nicht als einziges – Beispiel für ein Nucleinsäure-Polymerase-Reaktionsverfahren zum Amplifizieren einer Nucleinsäure-Testprobe betrachtet, das die Verwendung einer bekannten Nucleinsäure (DNA oder RNA) als Primer umfasst; es wird ferner eine Nucleinsäure-Polymerase verwendet, um ein spezifisches

Stück Nucleinsäure zu amplifizieren oder zu produzieren oder um ein spezifisches Stück Nucleinsäure zu amplifizieren oder zu produzieren, das zu einer bestimmten Nucleinsäure komplementär sein kann.

[0076] Der „HRG-Tyrosin-Autophösphorylierungstest" zum Nachweis der Gegenwart von HRG-Liganden dient dazu, die Reinigung eines liganden für den $p185^{HER2}$-Rezeptor zu überwachen. Dieser Assay beruht auf der Annahme, dass ein spezifischer Ligand für den $p185^{HER2}$-Rezeptor die Autophosphorylierung des Rezeptors in Analogie zu EGF und seiner Stimulierung von EGF-Rezeptor-Autophosphorylierung stimuliert. MDA-MB-453-Zellen oder MCF7-Zellen, die hohe Werte an $p185^{HER2}$-Rezeptoren, aber sehr kleine Mengen an Human-EGF-Rezeptoren enthalten, stammten von der American Type Culture Collection, Rockville, MD (ATCC Nr. HTB-131) und wurden in Gewebekultur mit 10 Kalbsfötenserum in DMEM/Hams F12- (1 : 1) Medium gehalten. Für den Test wurden die Zellen trypsiniert und mit etwa 150.000 Zellen/Napf in 24-Napf-Schalen (Costar) ausplattiert. Nach der Inkubation mit Serum-enthaltendem Medium über Nacht wurden die Zellen 2–18 Stunden vor dem Test in serumfreies Medium gelegt. Testproben von 100 µl-Aliquoten wurden jedem Napf zugesetzt. Die Zellen wurden 5–30 Minuten lang, typischerweise 30 Minuten lang, bei 37°C inkubiert und das Medium entfernt. Die Zellen in jedem Napf wurden mit 100 µl SDS-Gel-Denaturierungspuffer (Seprosol, Enpotech, Inc.) behandelt und die Platten 5 Minuten lang bei 100°C erhitzt, um die Zellen zu lösen und die Proteine zu denaturieren. Allquote aus jedem Napf wurden Elektrophorese auf 5–20% Gradient-SDS-Gelen (Novex, Encinitas, CA, USA) gemäß den Anweisungen der Herstellerfirma unterzogen. Nachdem die Farbstofffront den Gelboden erreicht hatte, wurde die Elektrophorese abgeschlossen und eine Folie aus PVDF-Membran (ProBlott, ABI) auf das Gel gelegt; die Proteine wurden in einer Blotting-Kammer (BioRad) bei 200 mAmp 30–60 Minuten lang vom Gel auf die Membran übertragen. Nach dem Blotting wurden die Membranen mit Tris-gepufferter Salzlösung, die 0,1 Tween 20-Detergenspuffer mit 5% BSA enthielt, 2–18 h lang inkubiert, um nichtspezifische Bindung zu blockieren, und dann mit einem Mäuse-Anti-Phosphotyrosin-Antikörper (Upstate Biological Inc. NY, USA) behandelt. Anschließend wurden die Membranblots mit Ziegen-Anti-Mäuse-Antikörper (konjugiert an alkalische Phosphatase) behandelt. Die Gele wurden unter Anwendung des ProtoBlot-Systems von Promega entwickelt. Nach dem Trocknen der Membranen wurde die $p185^{HER2}$ entsprechende Dichte der Banden in jeder Probenspur mit einem Hewlett Packard ScanJet Plus Scanner (angeschlossen an einen Macintosh-Computer) quantifiziert. Die Anzahl an Rezeptoren pro Zelle in den MDA-MB-453- oder MCF-7-Zellen ist solcherart, dass unter diesen experimentellen Bedingungen das $p185^{HER2}$-Rezeptorprotein das Hauptprotein ist, das markiert ist.

[0077] „Protein-Mikrosequenzieren" erfolgte gemäß dem folgenden Verfahren. Protein aus dem letzten HPLC-Schritt wurden entweder durch automatisierten Edman-Abbau mit einem Modell 470A Gasphasen-Sequenzierer von Applied Biosystems (ausgestattet mit einem 120A PTH-Aminosäure-Analysator) direkt oder nach Verdau mit verschiedenen Chemikalien oder Enzymen sequenziert. PTH-Aminosäuren wurden unter Anwendung des ChromPerfect-Datensystems (Justice Innovations, Palo Alto, CA, USA) integriert. Die Sequenzinterpretation erfolgte auf dem Computer VAX 11/785 der Digital Equipment Corporation (Henze) et al., J. Chromatography 404, 41–52 [1987]). In einige Fällen wurden Aliquoten der HPLC-Fraktionen auf 5–20% SDS-Polyacrylamidgelen Elektrophorese unterzogen, auf eine PVDF-Membran elektrotransferiert (ProBlott, ABI, Foster City, CA, USA) und mit Coomassie Brilliant-Blau (Matsudaira, P., J. Biol. Chem. 262, 10035-10038 [1987]) eingefärbt. Das spezifische Protein wurde aus dem Blot für N-terminales Sequenzieren ausgeschnitten. Um interne Proteinsequenzen zu bestimmen, wurden HPLC-Fraktionen unter Vakuum (SpeedVac) getrocknet, in geeigneten Puffern resuspendiert und mit Bromcyan, dem Lysin-spezifischen Enzym Lys-C (Wako Chemicals, Richmond, VA, USA) oder Asp-N (Boehringer Mannheim, Indianapolis, IN, USA) verdaut. Nach dem Verdau wurden die resultierenden Peptide als Gemisch sequenziert oder durch HPLC auf einer C4-Säule (entwickelt mit einem Propanol-Gradienten in 0,1 TFA) vor dem Sequenzieren gelöst (siehe oben).

II. Verwendung und Herstellung von HRG-Polypeptiden

1. Herstellung von HRG-Polypeptiden einschließlich Varianten

[0078] Das zur Herstellung von HRG-Polypeptiden zu verwendende System hängt von der jeweils gewählten HRG-Sequenz ab. Wenn die Sequenz ausreiched klein ist, wird HRG durch synthetische In-vitro-Polypeptidverfahren hergestellt. Am häufigsten jedoch wird HRD in rekombinanter Zellkultur unter Anwendung des unten beschriebenen Wirtsvektorsystems erzeugt.

[0079] Im Allgemeinen werden Säugetierwirtszellen verwendet, wobei solche Wirte post-translationale Systeme zur herkömmlichen Verarbeitung von HRG-Prosequenzen enthalten können oder nicht. Wenn die Wirtszellen solche Systeme enthalten, ist es möglich, natürliche Subdomänenfragmente wie z. B. HRG-GFD oder HRG-NTD-GFD aus den Kulturen zu gewinnen. Wenn nicht, kann die richtige Verarbeitung durch Transformieren der Wirte mit dem bzw. den erforderlichen Enzymen) oder durch Spalten des Vorläufers in vitro erfolgen. Es ist jedoch nicht notwendig, Zellen mit DNA zu tansformieren, die für die gesamte Prosequenz für ein ausgewähltes HRG kodiert, wenn es erwünscht ist, nur Fragmente von HRG-Sequenzen wie z. B. HRG-GFD zu

produzieren. Um beispielsweise HRG-GFD zu erzeugen, wird ein Startcodon an das 5'-Ende einer für HRG-GDF kodierenden DNA ligiert, diese DNA wird dazu verwendet, die Wirtszellen zu transformieren, und das Produkt wird direkt als die Met-N-terminale Form exprimiert (auf Wunsch kann das externe Met in vitro oder durch endogene N-terminate Demethionylasen entfernt werden). Alternativ dazu wird HRG-GFD als Fusion mit einer Signalsequenz exprimiert, die durch die Wirtszelle erkannt wird; sie verarbeitet und sekretiert dann die reife HRG-GFD, wie dies weiter unten beschrieben ist. Aminosäuresequenz-Varianten nativer HRG-GFD-Sequenzen werden in gleicher Weise produziert.

[0080] HRG-NTD wird in gleicher Weise als Molekül voller Länge erzeugt, jedoch aus der Expression von nur für HRG-NTD kodierender DNA, wobei das Stopcodon nach einem von S172-C182 angeordnet ist (**Fig.** 15).

[0081] Außerdem werden HRG-Varianten aus für Protein kodierender DNA exprimiert, worin sowohl die GFD- als auch die NTD-Domäne in ihrer richtigen Orientierung vorliegen, aber eine Aminosäure-Insertion, -Deletion oder -Substitution an der NTD-GFD-Verbindungsstelle aufweisen (z. B. innerhalb der Sequenz S172-C182). In einer weiteren Ausführungsform ist ein Stopcodon am 3'-Ende der NTD-GFD-kodierenden Sequenz positioniert (nach den Resten T/Q222-T245 von **Fig.** 15). Das Ergebnis ist eine lösliche Form von HRG-α, -β_1 oder -β_2, ohne Transmembran-Sequenz (diese Sequenz kann auch eine interne Signalsequenz sein, wird aber als Transmembran-Sequenz bezeichnet). In weiteren Variationen dieser Ausführungsform befindet sich eine interne Signalsequenz eines anderen Polypeptids an der Stelle der nativen HRG-Transmembran-Domäne, oder eine Zytoplasmadomäne eines weiteren Zellmembran-Polypeptids, z. B. Rezeptor-Kinase, ist für das HRG-α- oder zytoplasmatische HRG-β_1-β_2-Peptid substituiert.

[0082] In einer weiteren Ausführungsform sind die NTD-, GFD- und Transmembran-Domänen von HRG und anderen EGF-Familienmitgliedern füreinander substituiert, z. B. ist die NTD-äquivalente Region von EGF für die NTD von HRG substituiert oder die GFD von HRG für EGF in der verarbeiteten löslichen Vorform von EGF substituiert. Alternativ dazu ist eine HRG- oder EGF-Familienmitglied-Transmembran-Domäne mit dem C-terminalen E236 von HRG-β_3 fusioniert.

[0083] In einer weiteren Variante ist die HRG-Sequenz, die sich von K241 bis zum C-Terminus erstreckt, an ihrem N-Terminus mit dem C-Terminus eines Nicht-HRG-Polypeptids fusioniert.

[0084] Eine weitere Ausführungsform umfasst die funktionelle oder strukturelle Deletion der proteolytischen Verarbeitungsstelle in CTP, die GFD-Transmembran-überspannende Domäne. Beispielsweise ist das vermeintliche C-terminate Lysin (K241) von verarbeitetem HRG-α oder β_1-β_2 deletiert, mit einem anderen Rest substituiert, ein anderer Rest als K oder R ist zwischen K241 und R242 insertiert, oder es wird eine andere inaktivierende Mutation in der Prosequenz vorgenommen.

[0085] In einer weiteren Ausführungsform ist die Domäne eines beliebigen EGF-Familienmitglieds, die sich von (a) ihrem Cystein (entspricht (b) C221) bis zum C-terminalen Rest des Familienmitglieds erstreckt, für die analoge Domäne von HRG-α, -β_1 oder -β_2 substituiert (oder mit dem C-Terminus von HRG-β_3 fusioniert). Solche Varianten sind in gleicher Weise wie das Familienmitglied (und nicht das Stamm-HRG) frei von Wirtszellen verarbeitet. In komplexeren Ausführungsformen sind andere spezifische Spaltungsstellen (z. B. Protease-Setellen) in die CTP- oder GFD-Transmembran-überspannende Domäne substituiert (etwa Reste T/Q222-T 245 in **Fig.** 15). Beispielsweise ist die Amphiregulin-Sequenz E84-K99 oder die TGF-α-Sequenz E44-K58 für die HRG-α-Reste E221-K241 substituiert.

[0086] In einer weiteren Ausführungsform wird eine Variante (mit der Bezeichnung HRG-NTDxGFD) hergestellt, in der (1) der Lysinrest in der NTF-GFD-Verbindungssequenz VKC (Reste 180–182, **Fig.** 15) deletiert oder (vorzugsweise) durch einen anderen Rest als R wie z. B. N, A, T oder S substituiert ist und in der (2) ein Stopcodon in der Sequenz RCT oder RCQ (Reste 220–222, **Fig.** 15) anstelle von C, T (für HRG-α) oder Q (für HRG(β) eingefügt ist.

[0087] Ein bevorzugter HRG-α-Ligand mit Bindungsaffinität für $p185^{HER2}$ umfasst die Aminosäuren 226–265 von **Fig.** 4. Diese HRG-α-Ligand kann zusätzlich bis zu 1–20 Aminosäuren vor Aminosäure 226 in **Fig.** 4 und 1–20 Aminosäuren nach Aminosäure 265 in **Fig.** 4 umfassen. Ein bevorzugter HRG-β-Ligand mit Bindungsaffinität für $p185^{HER2}$ umfasst die Aminosäuren 226–265 von **Fig.** B. Dieser HRG-β-Ligand kann zusätzlich bis zu 1–20 Aminosäuren vor der Aminosäure 226 in **Fig.** 8 und 1–20 Aminosäuren nach der Aminosäure 265 in **Fig.** 8 enthalten.

[0088] GFD-Sequenzen sind jene, worin ein oder mehrere Reste, die einem anderen Mitglied der EGF-Familie entsprechen, deletiert oder substituiert sind oder einen angrenzend daran insertierten Rest aufweisen. Beispielsweise ist F216 von HRG durch Y, L202 mit E und F189 mit Y substituiert oder S203-P205 deletiert.

[0089] HRG enthält auch NTD-GFD mit ihrem C-Terminus in einem der ersten etwa 1–3 extrazellulären Domänreste (QKR, Reste 240-243, HRG-α, **Fig.** 15) oder den ersten etwa 1–2 Transmembran-Regionresten. Außerdem sind in einigen HRG-GFD-Varianten die Codons an der GFD-Transmembran-Proteolysestelle durch Substitution, Insertion oder Deletion modifziert. Die GFD-Proteolysestelle ist die Domäne, die den C-terminalen GFD-Rest und etwa 5 Reste N- und 5 Reste C-terminal von diesem Rest enthält. Bislang wurde weder der natürliche C-terminale Rest für HRG-α noch für HRG-β identifiziert, obwohl bekannt ist, dass Met-227 terminal und Va1-229 terminal HRG-α-GFD biologisch aktiv sind. Der native C-Terminus für HRG-α-GFD ist wahr-

scheinlich Met-227, Lys-228, Val-229, Gln-230, Asn-231 oder Gln-232, und für HRG-β_1-β_2-GFD ist er wahrscheinlich Met-226, Ala-227, Ser-228, Phe-229, Trp-230, Lys-231 oder (für HRG-β_1) K240 oder (für HRG-β_2) K 246. Der native C-Terminus wird durch C-terminales Sequenzieren problemlos bestimmt, obwohl es nicht entscheidend ist, dass HRG-GFD den nativen Terminus aufweist, sofern die GFD-Sequenz die erwünschte Aktivität besitzt. In einigen Ausführungsformen der HRG-GFD-Varianten wird bzw. werden die Aminosäure-Modifikation(en) im CTP hinsichtlich ihrer Fähigkeit gescreent, Proteolyse in vitro standzuhalten und die Protease, die für die Erzeugung von HRG-GFD verantwortlich ist, zu hemmen.

[0090] Falls es erwünscht ist, die HRG-Polypeptide voller Länge herzustellen und die 5'- oder 3'-Enden des konkreten HRG hierin nicht beschrieben sind, kann es notwendig sein, Nucleinsäuren zu produzieren, in denen die fehlenden Domänen durch homologe Regionen aus vollständigeren HRG-Nucleinsäuren gebildet werden. Alternativ dazu können die fehlenden Domänen durch Sondieren von Bibliotheken unter Einsatz der in den Abbildungen gezeigten DNA oder ihrer Fragmente erhalten werden.

A. Isolierung von für Heregulin kodierender DNA

[0091] Die für HRG kodierende DNA kann aus jeder cDNA-Bibliothek stammen, die aus Gewebe gebildet ist, das vermutlich HRG-mRNA besitzt und sie in detektierbaren Mengen exprimiert. HRG-DNA stammt auch aus einer genomischen Bibliothek.

[0092] Bibliotheken werden mit Sonden oder analytischen Werkzeugen sondiert, um das Gen von Interesse oder das Protein, für das es kodiert, zu identifizieren. Für cDNA-Expressions-Bibliotheken sind geeignete Sonden monoklonale oder polyklonale Antikörper, die HRG erkennen und sich spezifisch daran binden; Oligonucleotide einer Länge von etwa 20–80 Basen, die für bekannte oder vermutete Abschnitte von HRG-cDNA aus der gleichen oder einer anderen Spezies kodieren; und/oder komplementäre oder homologe cDNA oder Fragmente davon, die für das gleiche oder ein hybridisierendes Gen kodieren. Zweckmäßige Sonden für das Screenen genomischer DNA-Bibliotheken sind z. B. Oligonucleotide; cDNA oder Fragmente davon, die für die gleiche oder hybridisierende DNA kodieren; und/oder homologe genomische DNA oder Fragmente davon. Das Screenen der cDNA oder genomischen Bibliothek mit der ausgewählten Sonde kann unter Anwendung von Standardverfahren erfolgen, wie sie z. B. in Kapiteln 10–12 von Sambrook et al., s. o., beschrieben sind.

[0093] Eine andere Möglichkeit zur Isolierung des für HRG kodierenden Gens ist die Anwendung von PCR-Methodologie, wie sie in Kapitel 14 von Sambrook et al., s. o., beschrieben ist. Dieses Verfahren erfordert die Verwendung von Oligonucleotid-Sonden, die an HRG hybridisieren. Strategien für die Auswahl von Oligonucleotiden sind nachstehend erläutert.

[0094] Ein alternatives Verfahren zum Erhalten des Gens von Interesse ist seine chemische Synthetisierung unter Anwendung eines der Verfahren von Engels et al., Angew. Chem. Int. Ed. Engl. 28, 716–734 (1989). Diese Verfahren umfassen Triester-, Phosphit-, Phosphoramidit- und N-Phosphonat-Techniken, PCR und andere Autoprimer-Verfahren sowie die Oligonucleotid-Synthese auf festen Trägern. Diese Verfahren können angewendet werden, wenn die gesamte Nucleinsäuresequenz des Gens bekannt ist oder wenn die Sequenz der zum Kodierungsstrang komplementären Nucleinsäure verfügbar ist; wenn alternativ dazu die Ziel-Aminosäuresequenz bekannt ist, kann man potenzielle Nucleinsäuresequenzen unter Einsatz bekannter und bevorzugter Kodierungsreste für jeden Aminosäurerest ableiten.

[0095] Ein bevorzugtes Verfahren zur Durchführung der vorliegenden Erfindung ist die Verwendung sorgfältig ausgewählter Oligonucleotidsequenzen zum Screenen von cDNA-Bibliotheken aus verschiedenen Geweben, vorzugsweise aus Brust-, Darm-, Speicheldrüsen-, Plazenta-, Föten-, Hirn- und Karzinom-Zelllinien des Menschen. Andere biologische Quellen von für einen HRG-ähnlichen Liganden kodierender DNA sind andere Säugetiere und Vögel. Unter den bevorzugten Säugetieren finden sich Rinder, Schafe, Pferde, Mäuse und Nagetiere.

[0096] Die als Sonden ausgewählten Oligonucleotidsequenzen sollten ausreichende Länge besitzen und ausreichend eindeutig sein, um falsche Positive zu minimieren. Die konkrete(n) Nucleotidsequenz(en) basiert/basieren auf konservierten oder hochgradig homologen Nucleotidsequenzen oder Regionen von HRG-α. Die Oligonucleotide können an einer oder mehreren Positionen degeneriert sein. Die Verwendung degenerierter Oligonucleotide kann von besonderer Bedeutung sein, wenn eine Bibliothek aus einer Spezies gescreent ist, in der bevorzugte Codon-Verwendung in dieser Spezies unbekannt ist. Das Oligonucleotid muss markiert sein, sodass es nach Hybridisierung an DNA in der gescreenten Bibliothek detektiert werden kann. Das bevorzugte Markierungsverfahren ist die Verwendung von ^{32}P-markiertem ATP mit Polynucleotid-Kinase, wie des auf dem Gebiet allgemein bekannt ist, um das Oligonucleotid radiomarkieren zu können. Es können jedoch auch andere Verfahren zum Markieren des Oligonucleotids angewendet werden, z. B. Biotinylierung oder Enzymmarkierung.

[0097] Von besonderem Interesse ist HRG-Nucleinsäure, die für Polypeptid voller Länge kodiert. In einigen bevorzugten Ausführungsformen enthält die Nucleinsäuresequenz die native HRG-Signal-Transmembran-Sequenz. Nucleinsäure mit der gesamten Proteinkodiersequenz wird durch Screenen ausgewählter cDNA- oder

genomischer Biobliotheken und – falls erforderlich – unter Einsatz herkömmlicher Primer-Verlängerungsverfahren, wie sie in Kapitel 7.79 von Sambrook et al., s. o., beschrieben sind, erhalten, um Vorläufer und Verarbeitungs-Zwischenprodukte von mRNA zu detektieren, die möglicherweise nicht revers in cDNA transkribiert wurden.

[0098] Für HRG kodierende DNA dient dazu, DNA zu isolieren, die für den analogen Liganden aus anderen Tierspezies kodiert; dies erfolgt mittels Hybridisierung unter Einsatz der oben beschriebenen Verfahren. Die bevorzugten Tiere sind Säugetiere, insbesondere Rinder, Schafe, Pferde, Katzen, Hunde und Nagetiere, vor allem Ratten, Mäuse und Kaninchen.

B. Aminosäuresequenz-Varianten von Heregulin

[0099] Aminosäuresequenz-Varianten von HRG werden durch geeignete Nucleotid-Änderungen in HRG-DNA oder durch In-vitro-Synthese des erwünschten HRG-Polypeptids produziert. Zu solchen Varianten zählen z. B. Deletionen aus – oder Insertionen oder Substitutionen von – Resten innerhalb der Aminosäuresequenz für Human-HRG-Sequenzen. Jede beliebige Kombination von Deletion, Insertion und Substitution ist möglich, um zum Endkonstrukt zu gelangen, sofern dieses die erwünschten Eigenschaften aufweist. Die Aminosäure-Modifikationen können auch post-translationale Prozesse von HRG-α verändern – sie bewirken z. B. eine Änderung der Anzahl und Position von Glykosylierungsstellen, der Membran-Verankerungseigenschaften, der intrazellulären Position von HRG durch Insertieren, Deletieren oder eine andere Beeinflussung der Transmembran-Sequenz von nativem HRG, oder der Suszeptibilität gegenüber proteolytischer Spaltung.

[0100] Bei der Konstruktion von Aminosäuresequenz-Varianten von HRG hängen die Stelle und die Beschaffenheit der Mutation von der bzw. den zu modifizierenden HRG-Eigenschaft(en) ab. Die Mutationsstellen können individuell oder in Reihe modifiziert sein, z. B. durch (1) Substituieren zunächst mit ausgewählten konservativen Aminosäuren und dann mit radikaleren Aminosäuren (je nach den erzielten Ergebnisen), (2) Deletieren des Zielrests oder (3) Insertieren von Resten anderer Liganden, die an die lokalisierte Stelle angrenzen.

[0101] Ein nützliches Verfahren zur Identifikation von HRG-Resten oder -Regionen für die Mutagenese wird als „Alanin-Scanning-Mutagenese" bezeichnet und wird von Cunningham und Wells (Science 244, 1081–1085 [1989]) beschrieben. Hier wird ein Rest oder eine Gruppe von Zielresten identifiziert (z. B. geladene Reste wie etwa arg, asp, his, lys und glu) und durch eine neutrale oder negativ geladene Aminosäure ersetzt (am häufigsten Alanin oder Polyalanin), um die Wechselwirkung der Aminosäuren mit der wässrigen Umgebung innerhalb und außerhalb der Zelle zu beeinflussen. Jene Domänen, die funktionale Sensitivität gegenüber den Substitutionen aufweisen, werden durch Einsetzen weiterer oder anderer Varianten an den oder für die Substitutionsstellen weiterentwickelt. Während somit die Stelle für das Einsetzen einer Aminosäuresequenz vorbestimmt ist, muss die Beschaffenheit der Mutation an sich nicht vorbestimmt sein. Um beispielsweise die Leistungsfähigkeit einer Mutation an einer bestimmten Stelle zu optimieren, kann ala-Scanning oder Zufallsmutagenese am Zielcodon oder in der Zielregion erfolgen, und die exprimierten HRG-Varianten werden hinsichtlich optimaler Kombination erwünschter Aktivität gescreent.

[0102] Es gibt zwei Hauptvariablen in der Konstruktion von Aminosäuresequenz-Varianten: die Position der Mutationsstelle und die Beschaffenheit der Mutation. Es handelt sich hier um Varianten der HRG-Sequenz, die natürlich vorkommende Allele (die keine Manipulation der HRG-DNA erfordern) oder vorbestimmte Mutantenformen (produziert durch Mutieren der DNA, entweder um ein Allel oder eine in der Natur nicht vorkommende Variante zu bilden) darstellen können. Im Allgemeinen hängt die Position und die Beschaffenheit der jeweiligen Mutation von der zu modifizierenden HRG-Eigenschaft ab. Natürlich sind Variationen, die z. B. HRG in einen bekannten Rezeptorliganden umwandeln, im Schutzbereich der Erfindung ebenso wenig enthalten wie andere HRG-Varianten oder Polypeptidsequenzen, die nicht neuartig sind.

[0103] Aminosäuresequenz-Varianten reichen im Allgemeinen von etwa 1 bis 30 Resten, noch bevorzugter von 1 bis 10 Resten, typischerweise von etwa 1 bis 5 zusammenhängenden Resten. Deletionen können in Regionen geringer Homologie mit anderen Vorläufern der EGF-Familie eingeführt werden, um die Aktivität von HRG zu modifizieren. Deletionen aus HRG in Bereichen substanzieller Homologie mit anderen Sequenzen der EGF-Familie können die biologische Aktivität von HRG wahrscheinlich deutlicher modifizieren. Die Anzahl hintereinander erfolgender Deletionen ist ausgewählt, um die Tertiärstruktur von HRG in der betroffenen Domäne zu bewahren, z. B. Cystein-Vernetzung, β-Faltblattstruktur oder α-Helix.

[0104] Aminosäuresequenz-Insertionen umfassen Amino- und/oder Carboxyl-terminale Fusionen, deren Länge von einem Rest bis zu Polypeptiden mit 100 oder mehr Resten reichen, sowie intrasequenzielle Insertionen einzelner oder mehrerer Aminosäurereste. Intrasequenzielle Insertionen (d. h. Insertionen innerhalb der HRG-Sequenz) können im Allgemeinen von etwa 1 bis 10 Resten, noch bevorzugter von 1 bis 5 Resten, am bevorzugtesten von 1 bis 3 Resten, reichen. Beispiele für terminale Insertionen sind HRG mit einem N-terminalen Methionylrest (einem Artefakt der direkten Expression von HRG in bakterieller rekombinanter Zellkultur) und die Fusion einer heterologen N-terminalen Signalsequenz mit dem N-Terminus von HRG, um die Sekretion von reifem HRG aus rekombinanten Wirtszellen zu vereinfachen. Solche Signalsequenzen werden im Allge-

meinen aus den beabsichtigten Wirtszellspezies erhalten und sind mit ihnen homolog. Geeignete Sequenzen sind STII oder Ipp für E. coli, α-Faktor für Hefe sowie virale Signale wie z. B. Herpes gD für Säugetierzellen.

[0105] Andere Insertionsvarianten von HRG umfassen die Fusion des N- oder C-Terminus von HRG mit einem immunogenen Polypeptid, z. B. mit bakteriellen Polypeptiden wie etwa β-Lactamase oder einem Enzym, für das E. coli trp-Locus kodiert, oder Hefeprotein, Rinderserumalbumin und mit chemotaktischen Polypeptiden. C-terminate Fusionen von HRG-NTD-GFD mit Proteinen mit langer Halbwertszeit wie z. B. Immunglobulin-Konstantregionen (oder anderen Immunglobulin-Regionen), Albumin oder Ferritin (siehe WO 89/02922 vom 6. April 1989) sind ebenfalls hierin enthalten.

[0106] Eine weitere Gruppe von Varianten von Aminsoäure-Substitutionsvarianten. In diesen ist zumindest ein Aminosäurerest im HRG-Molekül entfernt und ein anderer Rest an seine Stelle insertiert. Die Stellen von größtem Interesse für die Substitutionsmutagenese sind Stellen, die als aktive Stelle(n) von HRG identifiziert ist bzw. sind, und Stellen, worin sich die in HRG-Liganden aus verschiedenen Spezies vorgefundenen Aminosäuren hinsichtlich ihrer Seitenkettenmasse, Ladung und/oder Hydrophobie substanziell unterscheiden.

[0107] Der Arminoterminus der zytoplasmatischen Region von HRF kann mit dem Carboxyterminus heterologer Transmembran-Domänen und Rezeptoren fusioniert sein, um ein Fusionspolypeptid zu bilden, das sich für das intrazelluläre Signalisieren einer Ligandenbindung an den heterologen Rezeptor eignet.

[0108] Andere Stellen von Interesse sind jene, in denen bestimmte Reste von HRG-ähnlichen Liganden (aus verschiedenen Spezies erhalten) identisch sind. Diese Positionen können für die biologische Aktivität von HRG von Bedeutung sein. Diese Stellen, insbesondere jene, die in eine Sequenz von zumindest drei anderen identisch konservierten Stellen fallen, sind relativ konservativ substituiert. Solche konservativen Substitutionen sind aus Tabelle 1 unter der Überschrift „Bevorzugte Substitutionen" ersichtlich. Wenn solche Substitutionen zu einer Änderung der biologischen Aktivität führen, werden substanziellere Änderungen (unter „Beispiele für Substitutionen" in Tabelle 1 aufgelistet oder weiter unten im Zusammenhang mit Aminosäureklassen erläutert) eingesetzt und die Produkte gescreent.

Tabelle 1

Ursprünglicher Rest	Beispiele für Substitutionen	Bevorzugte Substitutionen
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
Gly (G)	pro	pro
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe; norleucine	leu
Leu (L)	norleucine; ile; val; met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala	leu
Pro (P)	gly	gly
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucine	leu

[0109] Substanzielle Veränderungen der Funktion oder immunologischen Identität von HRG erfolgen durch Auswahl von Substitutionen, die sich hinsichtlich ihres Einflusses auf (a) die Struktur des Polypeptid-Rückgrats im Bereich der Substitution, z. B. als Faltblatt- oder Helixkonformation, (b) die Ladung oder Hydrophobie des Moleküls an der Zielstelle oder (c) die Masse der Seitenkette signifikant unterscheiden. Natürlich vorkommende Reste sind in Gruppen unterteilt, die auf gemeinsamen Seitenketten-Eigenschaften beruhen:

1) hydrophob: Norleucin, met, ala, val, leu, ile;
2) neutral hydrophil: cys, ser, thr;
3) sauer: asp, glu;
4) basisch: asn, gln, his, lys, arg;
5) Kettenorientierung beeinflussende Reste: gly, pro; und
6) aromatisch; trp, tyr, phe.

[0110] Nicht-konservative Substitutionen bewirken das Austauschen eines Mitglieds einer dieser Klassen gegen ein anderes. Solche substituierten Reste können in HRG-Regionen, die homolog mit anderen Rezeptorliganden sind, oder – noch bevorzugter – in die nicht-homologen Regionen des Moleküls eingesetzt werden.

[0111] In einer Ausführungsform der Erfindung ist es wünschenswert, eine oder mehrere Protease-Spaltungsstellen, die im Molekül vorhanden sind, zu inaktivieren. Diese Stellen werden durch Kontrolle der kodierten Aminosäuresequenz identifiziert. Wenn potenzielle Protease-Spaltungsstellen identifiziert sind, z. B. an

K241 R242, werden sie gegenüber proteolytischer Spaltung inaktiviert, indem der Zielrest mit einem anderen Rest ersetzt wird, vorzugsweise einem basischen Rest wie etwa Glutamin oder einem hydrophilen Rest wie z. B. Serin; durch Deletieren des Rests; oder durch Insertieren eines Prolylrests unmittelbar nach dem Rest.

[0112] In einer weiteren Ausführungsform ist jeder beliebige Methionylrest mit Ausnahme des Anfangs-Methionylrests, oder jeder beliebige Rest, der sich innerhalb von etwa drei Resten N- oder C-terminal zu jedem solchen Methionylrest befindet, durch einen anderen Rest substituiert (vorzugsweise gemäß Tabelle 1) oder deletiert. Die Anmelder stellten fest, dass die Oxidation der 2 GFD M-Reste im Verlauf der E. coli-Expression die GFD-Aktivität stark zu reduzieren scheint. Somit sind diese M-Reste in Einklang mit Tabelle 1 mutiert. Alternativ dazu sind etwa 1 bis 3 Reste angrenzend an solche Stellen insertiert.

[0113] Jeder beliebige Cysteinrest, der nicht an der Beibehaltung der richtigen Konformation von HRG beteiligt ist, kann auch – im Allgemeinen mit Serin – substituiert sein, um die Oxidationsstabilität des Moleküls zu verbessern und abweichende Vernetzung zu verhindern.

[0114] Stellen, die sich für Substitutionen, Deletionen oder Insertionen oder zur Verwendung als Fragmente eignen, sind – nummeriert ab dem N-Terminus von HRG-α in **Fig.** 4 – die Folgenden:

1) potenzielle Glykosaminoglykan-Additionsstellen an den Serin-Glycin-Dipeptiden an 42–43, 64–65, 151–152;
2) potenzielle Asparagin-gebundene Glykosylierung an den Positionen 164, 170, 208 und 437, Stellen (NDS) 164–166, (NIT) 170–172, (NTS) 208–210 und NTS (609–611);
3) potenzielle O-Glykosylierung in einem Cluster von Serin und Threonin an 209-218;
4) Cysteine an 226, 234, 240, 254, 256 und 265;
5) Transmembran-Domäne an 287–309;
6) Schleife 1, begrenzt durch die Cysteine 226 und 240 ; 7) Schleife 2, begrenzt durch die Cysteine 234 und 254;
8) Schleife 3, begrenzt durch die Cysteine 256 und 265; und
9) potenzielle Protease-Verarbeitungsstellen an 2–3, 8–9, 23–24, 33–34, 36–37, 45-46; 48–49, 62–63, 66–67, 86–87, 110–11, 123–124, 134–135, 142–143, 272–273, 278–279 und 285–286.

[0115] Analoge Regionen in HRG-β_1 können durch Verweis auf **Fig.** 9 bestimmt sein, die analoge Aminosäuren in HRG-α und HRG-β_1 anführt. Die analogen HRG-β_1-Aminosäuren können – wie oben in Zusämmenhang mit HRG-α besprochen – mutiert oder modifiziert sein. Analoge Regionen in HRG-β_2 können durch Verweis auf **Fig.** 15 bestimmt werden, die analoge Aminosäuren in HRG-α, HRG-β_1 und HRG-β_2 anführt. Die analogen HRG-β_2-Aminosäuren können mutiert oder modifiziert sein, wie dies oben für HRGα oder HRG-β_1 besprochen ist. Die analogen Regionen in HRG-β_3 können durch Verweis auf **Fig.** 15 bestimmt werden, die analoge Aminosäuren in HRG-α, HRG-β_1 und HRG-β_2 anführt. Die analogen HRG-β_3-Aminosäuren können mutiert oder modifiziert sein, wie dies oben für HRG-α, HRG-β_1 oder HRG-β_2 besprochen ist.

[0116] DNA, die für Aminosäuresequenz-Varianten von HRG kodiert, kann durch eine Vielzahl von auf dem Gebiet bekannten Verfahren hergestellt werden. Zu diesen Verfahren zählen u. a. die Isolierung aus einer natürlichen Quelle (im Fall natürlich vorkommender Aminosäuresequenz-Varianten) oder die Herstellung durch Oligonucleotid-mediierte (oder stellengerichtete) Mutagenese, PCR-Mutagenese und Kassetten-Mutagenese einer früher gebildeten Variante oder Nicht-Varianten-Vesion von HRG. Diese Techniken können HRG-Nucleinsäure (DNA oder RNA) oder zu HRG-Nucleinsäure komplementäre Nucleinsäure verwenden.

[0117] Oligonucleotid-mediierte Mutagenese ist ein bevorzugtes Verfahren zur Erzeugung von Substitutions-, Deletions- und Insertionsvarianten von HRG-DNA. Diese Technik ist auf dem Gebiet der Erfindung allgemein bekannt und wird von Adelman et al., DNA 2, 183 (1983), beschrieben.

[0118] Im Allgemeinen werden Oligonucleotide einer Länge von zumindest 25 Nucleotiden verwendet. Ein optimales Oligonucleotid besitzt 12 bis 15 Nucleotide, die zum Templat auf beiden Seiten des bzw. der für die Mutation kodierenden Nucleotids bzw. Nucleotide vollständig komplementär sind. Dies stellt sicher, dass das Oligonucleotid richtig an das einzelstrangige DNA-Templatmolekül hybridisiert. Die Oligonucleotide werden unter Anwendung auf dem Gebiet bekannter Techniken problemlos synthetisiert; siehe z. B. Crea et al., Proc. Natl. Acad. Sci. USA 75, 5765 (1978).

[0119] Einzelstrangiges DNS-Templat kann durch Denatürieren doppelstrangiger Plasmid-DNA o. dgl. unter Anwendung herkömmlicher Verfahren erzeugt werden.

[0120] Zur Änderung der nativen DNA-Sequenz (um z. B. Aminosäuresequenz-Varianten zu erzeugen) wird das Oligonucleotid an das einzelstrangige Templat unter zweckgemäßen Hybridisierungsbedingungen hybridisiert. Ein DNA-Polymerisationsenzym, üblicherweise das Klenow-Fragment von DNA-Polymerase 1, wird dann zugesetzt, um den komplementären Strang des Templats mit dem Oligonucleotid als Synthese-Primer zu synthetisieren. Ein Heteroduplex-Molekül wird somit gebildet, sodass ein Strang der DNA für die mutierte Form von HRG kodiert und der andere Strang (das ursprüngliche Templat) für die native, unveränderte Sequenz von HRG kodiert. Das Heteroduplex-Molekül wird dann zu einer geeigneten Wirtszelle transformiert, üblicherweise einem Prokaryoten wie etwa E. coli JM101. Nachdem die Zellen gewachsen sind, werden sie auf Agarose-Plat-

ten aufgebracht und unter Einsatz des mit ^{32}P-Phosphat radiomarkierten Oligonucleotid-Primers gescreent, um die bakteriellen Kolonien zu identifizieren, die die mutierte DNA enthalten. Die mutierte Region wird dann entfernt und in einen geeigneten Vektor für die Proteinproduktion gesetzt, im Allgemeinen einen Expressionsvektor der Art, wie sie üblicherweise für die Transformation eines geeigneten Wirts verwendet wird.

[0121] Das soeben beschriebene Verfahren kann solcherart modifiziert werden, dass ein Homoduplex-Molekül geschaffen wird, in dem beide Plasmidstränge die Mutationen) aufweisen. Die Modifikationen sind wie folgt: Das einzelstrangige Oligonucleotid wird – wie oben beschrieben – an das einzelstrangige Templat anneliert. Ein Gemisch von 3 Desoxyribonucleotiden, Desoxyriboadenosin (dATP), Desoxyriboguanosin (dGTP) und Desoxyribothymidin (dTTP), wird mit einem modifizierten Thiodesoxyribocytosin mit Namen dCTP-(aS) (Amersham Corporation) kombiniert. Dieses Gemsch wird dem Templat-Oligonucleotid-Komplex zugesetzt. Nach Zugabe von DNA-Polymerase zu diesem Gemisch wird ein DNA-Strang erzeugt, der mit dem Templat bis auf die mutierten Basen identisch ist. Außerdem enthält dieser neue DNA-Strang dCTP-(as) anstelle von dCTP (dient dazu, ihn vor Restriktions-Endonuclease-Verdau zu schützen). Nach dem Schneiden des Templatstrangs des doppelstrangigen Heteroduplex mit einem geeigneten Restriktionsenzym kann der Templatstrang mit ExoIII-Nuclease oder einer anderen zweckgemäßen Nuclease nach der Region, die die zu mutagenisierende(n) Stelle(n) enthält, verdaut werden. Die Reaktion wird dann abgebrochen, um das Molekül, das nur teilweise einzelstrangig ist, zurückzulassen. Ein kompletter doppeltrangiger DNA-Homoduplex wird dann unter Verwendung von DNA-Polymerase in Gegenwart aller vier Desoxyribonucleotidtriphosphate, ATP und DNA-Ligase gebildet. Dieses Homoduplex-Molekül kann anschließend in ein geeignete Wirtszelle wie etwa E. coli JM101 (s. o.) transformiert werden.

[0122] Beispiele für Substitutionen, die für jedes HRG gelten, sind: S2T oder D; E3D oder K; R4 K oder E; K5R oder E; E6D oder K; G7P oder Y; R8K oder D; G9P oder Y; K10R oder E; G11 P oder Y; K12R oder E; G 19P oder Y; S20T oder F; G21 P oder Y; K22 oder E; K23R oder E; Q38D; S107N; G108P; N120K; D121K; 5122 T; N126S; 1126L; T127S; A163V; N 164K; T165–T174; jeder Rest von I, L, V, M, F, D, E, R oder K; G 175V oder P; T176S oder V; S177K oder T; H178K oder S; L179F oder I; V180L oder S; K181 R oder E; A 183N oder V; E184K oder D; K185R oder E; E186D oder Y; K187R oder D; T188S oder Q; F189Y oder S; V191 L oder D; N192Q oder N; G193P oder A; G194P oder A; E195D oder K; F197Y oder I; M198V oder Y; V199L oder T; K200V oder R; D201 E oder K; L202E oder K; S203A oder T; N204Δ; N204Q; P205Δ; P205G; S206T oder R; K207K oder A; Y208P oder F; L2091 I oder D; K211 I oder D; F216Y oder I; T217H oder S; G218A oder P; A/D219K oder R; R220K oder A; A235/240/232V oder F; E236/241/233D oder K; E237/242/234D oder K; L238/"43/2351 oder T; Y239/244/"36F oder T; Q240/245/237N oder K; K241/246/238H oder R; R242/247/238H oder K; V243/248/239L oder T; L244/249/240I oder S; T245/250/2415 oder I; I246/251/242V oder T und T247/252/243S oder I. Spezifisch in Bezug auf HRG-α, T222S, K oder V; E223D, R oder Q; N224Q, K oder F; V225A, R oder D; P226G, I, K oder F; M227V, T, R oder Y; K228R, H oder D; V229L, K oder D; Q230N, R oder Y; N231 Q, K oder Y; Q232N, R oder Y; E233D, K oder T und K 234R, N oder D (angrenzende K/R-Mutationen sind in alternativen Ausführungsform gepaart, um neue Proteolysestellen zu schaffen). Spezifisch in Bezug auf HRG-β (jedes Mitglied) sind Q222N, R oder Y; N223Q, K oder Y; Y224F, T oder R; V225A, K oder D; M226V, T oder R; A227V, K, Y oder D; S228T, Y oder R; F229Y, 1 oder K und Y230F, T oder R geeignete Varianten. Spezifisch in Bezug auf HRG-β_1 sind K231 R oder D, H232R oder D; L2331, K, F oder Y; G234P, R, A oder S; 12351, K, F oder Y; E236D, R oder A; F2371, Y, K oder A; M238V, T, R oder A und E239D, R oder A geeignete Varianten. Spezifisch in Bezug auf HRG-β_1 und HRG-β_2 sind K231 R oder D geeignete Varianten. Alternativ dazu kann jeder dieser Reste deletiert oder die angeführten Substituenten angrenzendend dazu insertiert sein. Außerdem werden etwa 1 bis 10 Varianten kombiniert, um Kombinationen zu bilden. Diese Veränderungen erfolgen in proHRG, NTD, GFD, NTD-GFD oder anderen Fragmenten oder Fusionen. Q213-G215, A219 und die etwa 11–21 Reste C-terminal zu C221 unterscheiden sich von den verschiedenen HRG-Klassen. Reste an diesen werden unter HRG-Klassen oder EGF-Familienmitgliedern wechselseitig ausgetauscht, deletiert oder ein Rest angrenzend dazu insertiert.

[0123] DNA, die für HRG-α-Mutanten mit mehr als einer zu substituierenden Aminosäure kodiert, kann auf unterschiedliche Weise erzeugt. Werden. Wenn die Aminosäuren in der Polypeptidkette nahe beisammen angeordnet sind, können sie gleichzeitig unter Einsatz eines Oligonucleotids, das für alle erwünschten Aminosäure-Substitutionen kodiert, mutiert werden. Wenn hingegen die Aminosäuren in einer Entfernung voneinander angeordnet sind (getrennt um mehr als etwa 10 Aminosäuren), ist es schwieriger, ein einzelnes Oligonucleotid zu erzeugen, das für alle erwünschten Änderungen kodiert. Stattdessen kann eines von zwei alternativen Verfahren herangezogen werden.

[0124] PCR-Mutagenese eignet sich zur Herstellung von Aminosäurevarianten von HRG-α. Die nachstehenden Ausführungen beziehen sich zwar auf DNA, doch ist zu beachten, dass die Technik auch auf RNA anwendbar ist. Die PCR-Technik bezieht sich im Allgemeinen auf die folgende Vorgangsweise (vgl. Erlich, s. o., Kapitel von R. Higuchi, S. 60–70). Wenn kleine Mengen an Templat-DNA als Ausgangsmaterial in einer PCR verwendet werden, können Primer, die sich hinsichtlich ihrer Sequenz nur geringfügig von der korrespondierenden Region in einer Templat-DNA unterscheiden, bis zu relativ großen Mengen eines spezifischen DNA-Fragments

verwendet werden, das sich von der Templatsequenz nur an jenen Positionen unterscheidet, an denen sich die Primer vom Templat unterscheiden. Für die Einsetzug einer Mutation in eine Plasmid-DNA ist einer der Primer ausgebildet, die Position der Mutation zu überlappen und die Mutation zu enthalten; die Sequenz des anderen Primers muss mit einem Sequenzabschnitt des entgegengesetzten Plasmidstrangs identisch sein, doch diese Sequenz kann sich an einer beliebigen Stelle entlang der Plasmid-DNA befinden. Es ist allerdings vorzuziehen, dass die Sequenz des zweiten Primers innerhalb von 200 Nucleotiden von jener des ersten angeordnet ist, sodass am Ende die gesamte von Primern begrenzte amplifizierte DNA-Region problemlos sequenziert werden kann. PCR-Amplifikation unter Verwendung eines Primerpaars wie die soeben beschriebene führt zu einer Population von DNA-Fragmenten, die sich an der durch den Primer spezifizierten Mutationsposition und möglicherweise auch an anderen Positionen unterscheiden, da das Templatkopieren etwas fehlerhaft ist.

[0125] Wenn das Verhältnis zwischen Templat und Produktmaterial extrem niedrig ist, inkorporiert die überwiegende Mehrheit von Produkt-DNA-Fragmenten die erwünschte(n) Mutation(en). Dieses Prouktmaterial dient dazu, die entsprechende Region im Plasmid, die als PCR-Templat diente, mittels Standard-DNA-Technologie zu ersetzen. Mutationen an getrennten Positionen können durch Verwendung einer zweiten Primermutante oder mittels Durchführung einer zweiten PCR mit anderen Primermutanten und durch gleichzeitiges Ligieren der zwei resultierenden PCR-Fragmente an das Vektorfragment in einer drei- oder mehrteiligen Ligation gleichzeitig eingeführt werden.

[0126] Ein weiteres Verfahren zur Herstellung von Varianten, die Kassetten-Mutagenese, basiert auf der von Wells et al., Gene 34, 315 (1985) beschriebenen Technik. Das Ausgangsmaterial ist das Plasmid (oder ein anderer Vektor), der die zu mutierende HRG-DNA umfasst. Das oder die Codon(s) in der zu mutierenden HRG-DNA wird/werden identifiziert. Es muss eine einzelne Restriktions-Endonuclease-Stelle auf jeder Seite der identifizierten Mutationsstelle(n) geben. Wenn keine solchen Restriktionsstellen vorliegen, können sie mittels des oben beschriebenen Oligonucleotid-mediierten Mutagenese-Verfahrens erzeugt werden, um sie an den geeigneten Positionen in der HRG-DNA einzusetzen. Nach der Einführung der Restriktionsstellen in das Plasmid wird das Plasmid an diesen Stellen geschnitten, um es zu linearisieren. Ein doppelstrangiges Oligonucleotid, das für die Sequenz der DNA zwischen den Restriktionsstellen kodiert, aber die erwünschte(n) Mutationen) enthält, wird durch Standardverfahren synthetisiert. Die zwei Stränge werden getrennt synthetisiert und dann zusammen unter Anwendung herkömmlicher Techniken hybridisiert. Dieses doppelstrangige Oligonucleotid wird als Kassette bezeichnet. Diese Kassette ist ausgebildet, 3'- und 5'-Enden aufzuweisen, die mit den Enden des linearisierten Plasmids kompatibel sind, sodass sie direkt mit dem Plasmid ligiert werden kann. Dieses Plasmid enthält nun die mutierte HRG-DNA-Sequenz.

C. Insertion von DNA in ein Klonierungs- oder Expressionsvehikel

[0127] Die cDNA oder genomische DNA; die für natives oder variantes HRG kodiert, wird in einen replizierbaren Vektor insertiert, um weiteres Klonieren (Amplifizieren der DNA) oder Expression vorzunehmen. Es stehen viele Vektoren zur Verfügung, wobei die Auswahl des geeigneten Vektors von folgenden Fragen abhängt: 1) Soll er für DNA-Amplifikation oder DNA-Expression verwendet werden? 2) Wie groß ist die in den Vektor zu insertierende DNA? 3) Wie ist die mit dem Vektor zu transformierende Wirtszelle? Jeder Vektor enthält je nach seiner Funktion (Amplifikation oer Expression von DNA) und der Wirtszelle, für die er kompatibel ist, verschiedene Komponenten. Die Vektorkomponenten enthalten u. a. eines oder mehrere der folgenden Elemente: eine Signalsequenz, einen Replikationsursprung, ein oder mehrere Marker-Gene, ein Enhancer-Element, einen Promotor und eine Transkriptions-Terminationssequenz.

(i) Signalsequenz-Komponente

[0128] Im allgemeinen kann die Signalsequenz eine Komponente des Vektors oder ein Teil der in den Vektor insertierten HRG-DNA sein. Man geht davon aus, dass die native HRG-DNA für eine Signalsequenz am Aminoterminus (5'-Ende der für HRG kodierenden DNA) des Polypeptids kodiert, das während post-translationaler Verarbeitung des Polypeptids gespalten wird, um den reifen HRG-Polypeptidliganden zu schaffen, der sich an den $p185^{HER2}$-Rezeptor bindet, obwohl eine konventionelle Signalstruktur nicht offensichtlich ist. Natives proHRG wird aus der Zelle sekretiert, kann aber in der Membran verbleiben, da es eine transmembrane Domäne und eine zytoplasmatische Region in der Carboxyl-terminalen Region des Polypeptids enthält. Somit ist in einer sekretierten löslichen Version von HRG die Carboxyl-terminale Domäne des Moleküls (einschließlich der Transmembran-Domäne) üblicherweise deletiert. Die verkürzte Variante des HRG-Polypeptids kann aus der Zelle sekretiert werden, sofern die für die verkürzte Variante kodierende DNA für eine vom Wirt erkannte Signalsequenz kodiert.

[0129] HRG der Erfindung kann nicht nur direkt, sondern auch als Fusion mit einem heterologen Polypeptid exprimiert werden, vorzugsweise mit einer Signalsequenz oder einem anderen Polypeptid mit einer spezifischen Spaltungsstelle an den N- und/oder C-Termini der reifen Proteins oder Polypeptids. Im Allgemeinen

kann die Signalsequenz eine Komponente des Vektors sein, sie kann aber auch Teil der in den Vektor insertierten HRG-DNA sein. Im Schutzumfang der Erfindung sind HRG enthalten, deren native Signalsequenz deletiert und mit einer heterologen Signalsequenz ersetzt ist. Die ausgewählte heterologe Signalsequenz sollte eine sein, die von der Wirtszelle erkannt und verarbeitet wird, d. h. von einer Signal-Peptidase gespalten wird. Für prokaryotische Wirtszellen, die die native HRG-Signalsequenz nicht erkennen und verarbeiten, ist die Signalsequenz mit einer prokaryotischen Signalsequenz substituiert, die z. B. aus der Gruppe der alkalischen Phosphatase-, Penicillinase-, lpp- oder wärmestabilen Enterotoxin II-Leadern ausgewählt ist. Für die Hefe-Sekretion kann die native HRG-Signalsequenz mit Hefe-Invertase-, α-Faktor- oder sauren Phosphatase-Leadern substituiert sein. In der Säugetierzellen-Expression ist die native Signalsequenz zweckmäßig, obwohl auch andere Säugetier-Signalsequenzen in Frage kommen.

(ii) Replikationsursprungs-Komponente

[0130] Sowohl Expressions- als auch Klonierungsvektoren enthalten im Allgemeinen eine Nucleinsäuresequenz, die es dem Vektor ermöglicht, sich in einer oder mehreren ausgewählten Wirtszellen zu replizieren. Im Allgemeinen ist diese Sequenz in Klonierungsvektoren eine, die es dem Vektor ermöglicht, sich unabhängig von der chromosomalen Wirts-DNA zu replizieren, und enthält Replikationsursprünge autonom replizierender Sequenzen. Solche Sequenzen sind für eine Vielzahl an Bakterien, Hefe und Viren allgemein bekannt. Der Replikationsursprung aus dem Plasmid pBR322 eignet sich für die meisten gramnegativen Bakterien, der 2 µ-Plasmidursprung eignet sich für Hefe, und verschiedene virale Ursprünge (SV40, Polyoma, Adenovirus, VSV oder BPV) kommen für Klonierungsvektoren in Säugetierzellen in Frage. Im Allgemeinen ist die Replikationsursprungs-Komponente für Säugetier-Expressionsvektorn nicht erforderlich (der SV40-Ursprung kann typischerweise nur verwendet werden, da er den frühen Promotor enthält).
[0131] Die meisten Expressionsvektoren sind „Shuttle"-Vektoren, d. h. sie sind zur Replikation in zumindest einer Klasse von Organismen fähig, können aber in einen anderen Expressionsorganismus transfiziert werden. Beispielsweise wird ein Vektor in E. coli kloniert und dann in Hefe oder Säugetierzellen zwecks Expression transfiziert, obwohl er nicht zur unabhängigen Replikation des Wirtszellen-Chromosoms fähig ist.
[0132] DNA kann auch durch Insertion in das Wirtsgenom amplifiziert werden. Dies erfolgt am besten unter Verwendung von Bacillus-Spezies als Wirte, z. B. durch Vorsehen einer DNA-Sequenz im Vektor, die komplementär zu einer in genomischer Bacillus-DNA vorgefundener Sequenz ist. Die Transfektion von Bacillus mit diesem Vektor führt zu homologer Kekombination mit dem Genom und Insertion von HRG-DNA. Die Gewinnung für HRG kodierender genomischer DNA ist jedoch komplexer als jene eines exogen replizierten Vektors, da Restriktionsenzym-Verdau zur Exzision von HRG-DNA erforderlich ist. DNA kann durch PCR ampliziert und direkt in die Wirtszellen (ohne Replikationskomponente) transfiziert werden.

(iii) Selektions-Gen-Komponente

[0133] Expressions- und Klonierungsvektoren sollten ein Selektions- bzw. Auswahl-Gen enthalten, das auch als auswählbarer Marker bezeichnet wird. Das Gen kodiert für ein Protein, das für das Überleben oder Wachstum transfomierter Wirtszellen notwendig ist, die in einem selektiven Kulturmedium gezüchtet werden. Wirtszellen, die nicht mit dem das Selektions-Gen enthaltenden Vektor transformiert sind, überleben im Kulturmedium nicht. Typische Auswahl-Gene kodieren für Proteine, die (a) Resistenz gegenüber Antibiotika oder anderen Toxinen verleihen, z. B. Ampicillin, Neomycin, Methotrexat oder Tetracyclin, (b) auxotrophe Defizienten komplementieren oder (c) wesentliche Nährstoffe zuführen, die aus Komplexmedien nicht erhältlich sind, z. B. das für D-Alanin-Racemase für Bacilli kodierende Gen.
[0134] Ein Beispiel für ein Selektionssystem verwendet ein Medikament, um das Wachstum einer Wirtszelle abzubrechen. Jene Zellen, die erfolgreich mit einem heterologen Gen transformiert sind, exprimieren ein Protein, das Medikamentenresistenz verleiht, und überleben somit das Selektionsregime. Beispiele für eine solch dominierende Auswahl verwenden die Medikamente Neomycin (Southern et al., J. Molec. Appl. Genet. 1, 327 [1982]), Mycophenolsäure (Mulligan et al., Science 209, 1422 [1980]), oder Hygromycin (Sugden et al., Mol. Cell. Biol. 5, 410–413 [1985]). Die drei oben angeführten Beispiele verwenden bakterielle Gene unter eukaryotischer Steuerung, um Resistenz gegenüber dem Arzneimittel G418 oder Neomycin (Geneticin), xgpt (Mycophenolsäure) bzw. Hygromycin zu verleihen.
[0135] Ein weiteres Beispiel für geeignete auswählbare Marker für Säugetierzellen sind jene, die Identifikation von Zellen ermöglichen, die zur Aufnahme von HRG-Nucleinsäuren fähig sind, z. B. Dihydrofolat-Reductase (DHFR) oder Thymidin-Kinase. Die Säugetierzellen-Transformanten werden Selektionsdruck ausgesetzt, und nur die Transformanten sind aufgrund ihrer Ausnahme des Markers ausgebildet, ihn zu überleben. Der Selektionsdruck wird ausgeübt, indem die Transformanten unter Bedingungen kultiviert werden, unter denen die Konzentration des Auswahlmittels im Medium nacheinander geändert wird, was zu Amplifikation sowohl des Selektions-Gens als auch der für HRG kodierenden DNA führt. Die Amplifikation ist der Prozess, durch den

Gene, nach denen größere Nachfrage bezüglich der Produktion eines für das Wachstum entscheidenden Proteins herrscht, in Tandem innerhalb der Chromosome aufeinander folgender Generationen rekombinanter Zellen wiederholt werden. Höhere Mengen an HRG werden aus der amplifizierten DNA synthetisiert.

[0136] Beispielsweise werden mit dem DHFR-Selektions-Gen transformierte Zellen zunächst durch Kultivieren der Transformanten in einem Kulturmedium identifiziert, das Methotrexat (Mtx), einen kompetitiven Antagonisten von DHFR, enthält. Eine geeignete Wirtszelle ist bei Verwendung von DHFR der Wildform die Chinahamster-Eierstock- (CHO-) Zelllinie mit DHFR-Aktivitäts-Defizienz, deren Herstellung und Vermehrung gemäß dem Verfahren von Urlaub und Chasin, Proc. Natl. Acad. Sci. USA 77, 4216 (1980 erfolgt. Die transformierten Zellen werden dann höheren Mengen an Methotrexat ausgesetzt. Dies führt zur Synthese mehrerer Kopien des DHFR-Gens und gleichzeitig zu mehreren Kopien anderer die Expressionsvektoren umfassender DNA, z. B. für HRG kodierender DNA. Dieses Amplifikationsverfahren kann mit jedem anderen geeigneten Wirt durchgeführt werden, z. B. ATCC Nr. CCL61 CHO-K1 (trotz der Gegenwart von endogenem DHFR), wenn z. B. eine DHFR-Genmutante, die gegenüber Mtx hochgradig resistent ist, verwendet wird (siehe EP 117.060). Alternativ dazu können Wirtszellen (insbesondere Wirte der Wildform, die endogenes DHFR enthalten), transformiert oder cotransformiert mit für HRG kodierenden DNA-Sequenzen, DHFR-Protein der Wildform und ein anderer auswählbarer Marker wie z. B. Aminoglykosid-3'-phosphotransferase (APH) durch Zellwachstum in Medium selektiert werden, das ein Auswahlmittel für den auswählbaren Marken wie z. B. ein Aminoglykosidin-Antibiotikum wie etwa Kanamycin, Neomycin oder G418 enthält (siehe US-Patent 4.965.199).

[0137] Ein geeignetes Auswahl-Gen zur Verwendung in Hefe ist das im Hefeplasmid Yrp7 vorhandene trpl-Gen (Stinchcomb et al., Nature 282, 39 [1979]; Kingsman et al., Gene 7, 141 [1979] oder Tschemper et al., Gene 10, 157 [1980]). Das trp1-Gen bietet einen Selektionsmarker für eine Stammmutante von Hefe, die sich nicht in Tryptophan vermehren kann, z. B. ATCC Nr. 44076 oder PEP4-1 (Jones, Genetics 85, 12 [1977]). Die Gegenwart der trpl-Läsion im Hefe-Wirtszellengenom sorgt dann für eine wirkungsvolle Umgebung zum Detektieren von Transformation durch Wachstum in Abwesenheit von Tryptophan. Ebenso werden Leu2-defiziente Wirtsstämme (ATCC 20.622 oder 38.626) durch das Leu2-Gen tragende bekannte Plasmide komplementiert.

(iv) Promotor-Komponente

[0138] Expressions- und Klonierungsvektoren enthalten üblicherweise einen Promotor, der durch den Wirtsorganismus erkannt wird und operabel mit HRG-Nucleinsäure verbunden ist. Promotoren sind untranslatierte Sequenzen, die sich stromauf (5') vom Startcodon eines strukturellen Gens (im allgemeinen innerhalb von etwa 100 bis 1000 bp) befinden und die Transkription und Translation einer bestimmten Nucleinsäuresequenz wie z. B. HRG, mit der sie operabel verbunden sind, steuern. Solche Promotoren fallen typischerweise in zwei Klassen: induzierbare und konstitutive. Induzierbare Promotoren sind jene, die ein höheres Ausmaß an Transkription aus DNA unter ihrer Steuerung als Reaktion auf eine Änderung der Kulturbedingungen – z. B. die Gegenwart oder Abwesenheit eines Nährstoffs oder ein Temperaturwechsel – initiieren. Es sind derzeit zahlreiche Promotoren bekannt, die von einer Vielzahl potenzieller Wirtszellen erkannt werden. Diese Promotoren werden operabel mit für HRG kodierender DNA verbunden, indem der Promotor aus der Quellen-DNA durch Restriktionsenzym-Verdau entfernt und die isolierte Promotorsequenz in den Vektor insertiert wird. Sowohl die native HRG-Promotorsequenz als auch zahlreiche heterologe Promotoren können dazu dienen, Amplifikation und/oder Expression von HRG-DNA zu lenken. Allerdings sind heterologe Promotoren vorzuziehen, da sie im Allgemeinen umfassendere Transkription und höhere Ausbeuten an exprimiertem HRG ermöglichen als der native HRG-Promotor.

[0139] Für prokaryotische Wirte geeignete Promotoren sind die β-Lactamase- und Lactose-Promotorsysteme (Chang et al., Nature 275, 615 [1978]; und Goeddel et al., Nature 281, 544 [1979]), alkalische Phosphatase, ein Tryptophan- (trp-) Promotorsystem (Goeddel, Nucleic Acids Res. 8, 4057 [1980] und EP 36.776) und Hybridpromotoren wie z. B. der tac-Promotor (deBoer et al., Proc. Natl. Acad. Sci. USA 80, 21–25 [1983]). Andere bekannte bakterielle Promotoren kommen allerdings auch in Frage. Ihre Nucleotidsequenzen wurden bereits veröffentlicht, wodurch Fachleute auf dem Gebiet der Erfindung in der Lage sind, sie mit für HRG kodierender DNA unter Verwendung von Linkern oder Adaptoren (zwecks Bereitstellung allenfalls erforderlicher Restriktionsstellen) operabel zu ligieren (Siebenlist et al., Cell 20, 269 [1980]). Promotoren zur Verwendung in bakteriellen Systemen enthalten im Allgemeinen auch eine Shine-Dalgarno- (S. D.-) Sequenz, die mit der für HRG kodierender DNA operabel verbunden ist.

[0140] Geeignete Promotorsequenzen zur Verwendung mit Hefewirten sind z. B. die Promotoren für 3-Phosphoglycerat-Kinase (Hitzeman et al., J. Biol. Chem. 255, 2073 [1980]) oder andere glykolytische Enzyme (Ness et al., J. Adv. Enzyme Reg. 7, 149 [1968]; und Holland, Biochemistry 17, 4900 [1978]) wie etwa Enolase, Glyceraldehyd-3-phosphat-Dehydrogenase, Hexokinase, Pyruvat-Decarboxylase, Phosphofructokinase, Glucose-6-phosphat-Isomerase, 3-Phosphoglycerat-Mutase, Pyruvat-Kinase, Triosephosphat-Isomerase, Phosphoglucose-Isomerase und Glucokinase.

[0141] Andere Hefepromotoren, die induzierbare Promotoren mit dem zusätzlichen Vorteil der durch Wachstumsbedingungen gesteuerten Transkription sind, sind die Promotorregionen für Alkohol-Dehydrogenase 2, Isocytochrom C, Saure Phosphatase, Abbauenzyme, die mit Stickstoff-Metabolismus assoziiert sind, Metallothionein, Glyceraldehyd-3-phosphat-Dehydrogenase und für Maltose- und Galactose-Verwendung verantwortliche Enzyme. Zweckmäßige Vektoren und Promotoren zur Verwendung in der Hefeexpression sind ausführlich in Hitzeman et al., EP 73.657A beschrieben. Hefe-Enhancer können mit Hefe-Promotoren auch vorteilhaft verwendet werden.

[0142] Promotorsequenzen sind für Eukaryoten bekannt. Praktisch alle eukaryotischen Gene besitzen eine AT-reiche Region, die sich etwa 25 bis 30 Basen stromauf von jener Stelle befindet, an der die Transkription initiiert wird. Eine weitere Sequenz, die sich 70 bis 80 Basen stromauf vom Transkriptionsstart zahlreicher Gene befindet, ist eine CXCAAT(SEQ.-ID Nr. 1) Region, worin X jedes beliebige Nucleotid sein kann. Am 3'-Ende der meisten eukaryotischen Gene ist eine AATAAA-Sequenz (SEQ.-ID Nr. 2) angeordnet, die das Signal für das Hinzufügen des polyA-Schwanzes zum 3'-Ende der Kodiersequenz sein kann. Alle diese Sequenzen können günstigerweise in Säugetier-Expressionsvektoren insertiert werden.

[0143] Die HRG-Gen-Transkription aus Vektoren in Säugetier-Wirtszellen wird durch Promotoren gesteuert, die aus den Genomen von Viren wie z. B. Polyoma-Virus, Geflügelpocken-Virus (GB-Patent 2.211.504 vom 5. Juli 1989), Adenovirus (z. B. Adenovirus 2), Rinderpapilloma-Virus, Vögel-Sarkomavirus, Zytomegalievirus, einem Retrovirus, Hepatitis-B-Virus und – am bevorzugtesten – Affenvirus 40 (SV40), aus heterologen Säugetier-Promotoren, z. B. dem Actin-Promotor oder einem Immunglobulin-Promotor, aus Hitzeschock-Promotoren und aus dem Promotor, der normalerweise mit HRG-Sequenz assoziiert wird, stammen, sofern solche Prootoren mit den Wirtszellensystemen verträglich sind.

[0144] Die frühen und späten Promotoren des SV40-Virus werden günstigerweise als SV40-Restriktionsfragment erhalten, das auch den viralen SV40-Replikationsursprung enthält (Fiers et al., Nature 273, 113 [1978]; Mulligan und Berg, Science 209, 1422–1427 [1980]; Pavlakis et al., Proc. Natl. Acad. Sci. USA 78, 7398–7402 [1981]). Der unmittelbar frühe Promotor des Human-Zytomegalievrus wird günstigerweise als HindIII E-Restriktionsframgnet erhalten (Greenaway et al., Gene 18, 355–360 [1982]). Ein System zur Expression von DNA in Säugetierwirten unter Einsatz des Ringer-Papillomavirus als Vektor ist in US-Patent 4.419.446 geoffenbart. Eine Modifikation dieses Systems ist in US-Patent 4.601.978 beschrieben. Siehe auch Gray et al., Nature 295, 503–508 (1982) betreffend die Expression von für Immuninterferon in Affenzellen kodierender cDNA; Reyes et al., Nature 297, 598–601 (1982) betreffend die Expression von Human-β-Interferon-cDNA in Mäusezellen unter der Steuerung eines Thymidin-Kinase-Promotors aus Herpes simplex-Virus; Canaani und Berg, Proc. Natl. Acad. Sci. USA 79, 5166–5170 (1982) betreffend die Expression des Human-Interferon-β1-Gens in kultivierten Mäuseund Kaninchenzellen; und Gorman et al., Proc. Natl. Acad. Sci. USA 79, 6777–6781 (1982) betreffend die Expression bakterieller CAT-Sequenzen in CV-1-Affennierenzellen, Hühnerembryo-Fibroblasten, Chinahamster-Eierstockzellen, HeLa-Zellen und Mäuse-NIH-3T3-Zellen unter Einsatz der langen terminalen Rous-Sarkomavirus-Wiederholung als Promotor.

(v) Enhancerelement-Komponente

[0145] Die Transkription eine für HRG kodierenden DNA der Erfindung durch höhere Eukaryoten wird oft durch Insertieren einer Enhancer-Sequenz in den Vektor gesteigert. Enhancer sind cis-wirkende Elemente der DNA (üblicherweise von etwa 10–300 bp), die auf einen Promotor einwirken, um dessen Transkription zu erhöhen. Enhancer sind relativ orientierungs- und positionsunabhängig und wurden 5' (Laimins et al., Proc. Natl. Acad. Sci. USA 78, 993 [1981]) und 3' (Lusky et al., Mol. Cell Bio. 3, 1108 [1983]) zur Transkriptionseinheit, innerhalb eines Introns (Banerji et al., Cell 33, 729 (1983]) sowie innerhalb der Kodiersequenz selbst (Osborne et al., Mol. Cell Bio. 4, 1293 [1984]) lokalisiert. Zahlreiche Enhancer-Sequenzen sind aus Säugetier-Genen bekannt (Globin, Elastase, Albumin, α-Fetoprotein und Insulin). Typischerweise verwendet man aber einen Enhancer aus einem eukaryotischen Zellvirus. Beispiele dafür sind der SV40-Enhancer auf der späten Seite des Replikationsursprungs (bp 100–270), der frühe Zytomegalievirus-Promotor-Enhancer, der Polyoma-Enhancer auf der späten Seite des Replikationsursprungs und Adenovirus-Enhancer (siehe auch Yaniv, Nature 297, 17–18 [1982]) betreffend die Verstärkung von Elementen zur Aktivierung eukaryotischer Promotoren. Der Enhancer kann auch an einer Position 5' oder 3' von der HRG-DNA in den Vektor gespleißt sein, befindet sich aber vorzugsweise an einer Stelle 5' vom Promotor.

(vi) Transkriptionsterminations-Komponente

[0146] In eukaryotischen Wirtszellen (Hefe-, Pilz-, Insekten-, Pflanzen-, Tier-, Human- oder kernhältigen Zellen aus anderem multizellulären Organismen) verwendete Expressionsvektoren enthalten auch Sequenzen, die für die Termination der Transkription oder für die Stabilisierung der mRNA erforderlich sind. Solche Sequenzen sind üblicherweise aus den untranslatierten 5'- und manchmal auch aus den untranslatierten 3'-Regionen

eukaryotischer oder viraler DNA oder cDNA erhältlich. Diese Regionen enthalten Nucleotidsegmente, die als polyadenylierte Fragmente im untranslatierten Abschnitt der für HRG kodierenden mRNA transkribiert sind. Die untranslatierten 3'-Regionen enthalten auch Transkriptionsterminations-Stellen.

[0147] Die Konstruktion geeigneter Vektoren, die eine oder mehrere der oben angeführten Komponenten der erwünschten Kodier- und Steuersequenzen enthalten, sieht die Anwendung von Standard-Ligationstechniken vor. Isolierte Plasmide oder DNA-Fragmente werden gespalten, zurechtgeschnitten und in der zur Erzeugung der erforderlichen Plasmide erwünschten Form religiert.

[0148] Für die Analyse zur Bestätigung der korrekten Sequenzen in den konstruierten Plasmiden werden die Ligationsgemische dazu verwendet, E. coli K12-Stamm 294 (ATCC 31.446) zu transformieren, und erfolgreiche Transformanten gegebenenfalls durch Ampicillin- oder Tretracyclin-Resistenz ausgewählt. Plasmide aus den Transformanten werden gebildet, durch Restriktions-Endonuclease-Verdau analysiert und/oder durch das Verfahren von Messing et al., Nucleic Acids. Res. 9, 309 (1981) oder durch das Verfahren von Maxam et al., Methods in Enzymology 65, 499 (1980) sequenziert.

[0149] Besonders nützlich sind hierin Expressionsvekeotren, die für transiente Expression von für HRG kodierender DNA in Säugetierzellen sorgen. Im Allgemeinen umfasst die transiente Expression die Verwendung eines Expressionsvektors, der zur effizienten Replikation in einer Wirtszelle fähig ist, sodass diese zahlreiche Kopien des Expressionsvektors akkumuliert und ihrerseits hohe Werte eines erwünschten Polypeptids synthetisiert, für das der Expressionsvektor kodiert. Transiente Expressionssysteme, die einen zweckmäßigen Expressionsvektor und eine Wirtszelle umfassen, ermöglichen die praktische und zuverlässige Identifikation von Polypeptiden, für die klonierte DNA kodieren, sowie das rasche Screenen solcher Polypeptide hinsichtlich erwünschter biologischer oder physiologischer Eigenschaften. Somit sind transiente Expressionssysteme hierin besonders geeignet, um Analoge und Varianten von HRG zu identifizieren, die HRG-ähnliche Aktivität aufweisen. Ein derartiges transientes Expressionssystem ist in der am 29. März 1989 veröffentlichten EP 309.237 beschrieben. Andere Verfahren, Vektoren und Wirtszellen, die sich zur Anpassung an die Synthese von HRG in rekombinanter Wirbeltier-Zellkultur eignen, sind in Gething et al., Nature 293, 620–625 (1981); Mantei et al., Nature 281, 40–46 (1979); Levinson et al., EP 117.060 und EP 117.058 beschrieben. Ein besonders geeignetes Expressionsplasmid für Säugetierzellenkultur-Expression von HRG ist pRKS (EP 307.247).

D. Selektion und Transformation von Wirtszellen

[0150] Zweckmäßige Wirtszellen für das Klonieren oder Exprimieren von Vektoren sind hierin die oben beschriebenen prokaryotischen, Hefe- oder höheren eukaryotischen Zellen. Geeignete Prokaryoten sind z. B. Eubakterien, z. B. gramnegative oder grampositive Organismen wie etwa E. coli, Bacilli wie etwa B. subtilis, Pseudomonas-Spezies wie etwa P. aeruginosa, Salmonella typhimurium oder Serratia marcescans. Ein bevorzugter E. coli-Klonierungswirt ist E. coli 294 (ATCC 31.446), obwohl auch andere Stämme wie z. B. E. coli B, E. coli $_x$1776 (ATCC 31.537) und E. coli W3110 (ATCC 27.325) in Frage kommen. Diese Beispiele sind lediglich veranschaulichend und keinesfalls einschränkend. Vorzugsweise sollte die Wirtszelle minimale Mengen proteolytischer Enzyme sekretieren. Alternativ dazu eignen sich In-vitro-Klönierungsverfahren wie etwa PCR oder andere Nucleinsäure-Polmyerase-Reaktionen.

[0151] Zusätzlich zu Prokaryoten, sind eukaryotische Mikroben wie etwa filamentöse Pilze oder Hefe zweckmäßige Wirte für HRG-kodierende Vektoren. Saccmaromyces cerevisiae, d. h. gemeine Backhefe, ist der am häufigsten verwendete der niederen eukaryotischen Wirts-Mikroorganismen. Doch einige andere Gattunge, Spezies und Stämme stehen ebenfalls zur Verfügung und sind hierin auch geeignet, z.. Schizosaccharomyces pombe (Beach und Nurse, Nature 290, 140 [1981]; EP 139.383 vom 2. Mai 1985), Kluyveromyces-Wirte (U.S.S.N. 4.943.529) wie z. B. K. lactis (Louvencourt et al., J. Bacteriol. 737 [1983]; K. fragilis, K. bulgaricus, K. thermotolerans und K. marxianus, Yarrowia (EP 402.226); Pichia pastoris (EP 183.070), Sreekrishna et al., J. Basic Microbiol. 28, 265-278 [1988]); Candida, Trichoderma reesia (EP 244.234); Neurospora crassa (Case et al., Proc. Natl. Acad. Sci. USA 76, 5259–5233 [1979]) und filamentöse Pilze wie z. B. Neurospora, Penicillium, Tolypocladium (WO 91/00357 vom 10. Januar 1991) sowie Aspergillus-Wirte wie etwa A. nidulans (Ballance et al., Biochem. Biophys. Res. Commun. 112, 284–289 [1983] ; Tilburn et al., Gene 26, 205–221 [1983] ; Yelton et al., Proc. Natl. Acad. Sci. USA 81, 1470–1474 [1984]); und A. niger (Kelly und Hynes, EMBO). 4, 475-479 [1985]).

[0152] Geeignete Wirtszellen für die Expression von glykosyliertem HRG-Polypeptid stammen aus multizellulären Organismen. Solche Wirtszellen sind zu komplexen Verarbeitungsund Glykosylierungsaktivitäten fähig. Im Prinzip eignet sich jede höhere eukaryotische Zellkultur, ob sie nun aus Wirbeltier- oder Wirbellosenkultur stammt. Beispiele für Wirbellosenzellen sind Pflanzen- und Insektenzellen. Zahlreiche bakulovirale Stämme und Varianten sowie entsprechende permissive Insektenwirtszellen aus Wirten wie z. B. Spodoptera frugiperda (Raupe), Aedes aegypti (Moskito), Aedes albopictus (Moskito), Drosophila melanogaster (Fruchtfliege) und Bombyx mori-Wirtszellen wurden bereits identifiziert (siehe z. B. Luckow et al, Bio/Technology 6, 47–55 [1988]; Miller et al., in Genetic Engineering , Setlow, J. K. et al., Hg. Bd. 8 [Plenum Publishing, 1986], S. 277–289; und

Maeda et al., Nature 315, 592–594 [1985]). Eine Vielzahl derartiger Stämme ist allgemein erhältlich, z. B. die L-1-Variante von Autographa californica-NPV und der Bm5-Stamm von Bomby mori-NPV, wobei solche Viren hierin als Virus gemäß der Erfindung verwendet werden können, insbesondere zur Transfektion von Spodoptera frugiperda-Zellen. Pflanzenzellkulturen von Baumwolle, Mais, Kartoffeln, Sojabohnen, Petunien, Tomaten und Tabak kommen als Wirte in Frage. Typischerweise werden Pflanzenzellen durch Inkubation mit bestimmten Stämmen des Bakteriums Agrobacterium tumefaciens transfiziert, das zuvor manipuliert wurde, um HRG-DNA zu enthalten. Während der Inkubation der Pflanzenzellkultur mit A. tumefaciens wird die für HRG kodierende DNA auf einen Pflanzenzellwirt übertragen, sodass er transfiziert wird und unter bestimmten Bedingungen HRG-DNA exprimiert. Außerdem stehen mit Pflanzenzellen verträgliche regulatorische und Signalsequenzen zur Verfügung, z. B. die Nopalin-Synthase-Promotor- und -Polyadenylationssignal-Sequenzen (Depicker et al., J. Mol. Appl. Gen. 1, 561 (1982]). Darüber hinaus können aus der stromaufwärtigen Region des T-DNA 780-Gens isolierte DNA-Segmente Transkriptionsmengen von pflanzlich exprimierbaren Genen in rekombinante DNA enthaltenden Pflanzengeweben aktivieren oder erhöhen (siehe EP 321.196 vom 21. Juni 1989).

[0153] Das meiste Interesse gilt jedoch den Wirbeltierzellen, wobei die Vermehrung von Wirbeltierzellen in Kultur (Gewebekultur) in den letzten Jahren zu einem Routinevorgang wurde (Tissue Culture, Academic Press, Kruse und Patterson, Hg. [1973]). Beispiele für nützliche Säugetier-Wirtszelllinien sind die Affennieren-CV1-Linie, transformiert durch SV40 (COS-7, ATCC CRL 1651); die Humanembryonieren-Linie (293 oder 293-Zellen, die für das Wachstum in Suspensionskultur subkloniert sind; siehe Graham et al., J. Gen. Virol. 36, 59 [1977]); Babyhamster-Nierenzellen (BHK, ATCC CCL 10); Chinahamster-Eierstockzellen/-DHFR (CHO, Urlaub und Chasin, Proc. Natl. Acad. Sci. USA 77, 4216 [1980]; Mäuse-Sertolizellen (TM4, Mather, Biol. Reprod. 23, 243–251 [1980]); Affennierenzellen (CV1 ATCC CCL 70); Nierenzellen aus afrikanischen Grünen Meerkatzen (VERO-76, ATCC CRL-1587); Humanzervikalkarzinom-Zellen (HELA, ATCC CCL 2); Hundenierenzellen (MDCK, ATCC CCL 34); Büffelratte-Leberzellen (BRL 3A, ATCC CRL 1442); Humanlungenzellen (W138, ATCC CCL 75); Humanleberzellen (Hep G2, HB 8065), Mäuse-Brusttumor (MMT 060562, ATCC CCL51); TRI-Zellen (Mather et al., Annals N.Y. Acad. Sci. 383, 44–68 [1982J); MRC 5-Zellen, FS4-Zellen; und eine Humanhepatom-Zelllinie (Hep G2). Bevorzugte Wirtszellen sind Numanembryonieren-293 und Chinahamster-Eierstockzellen.

[0154] Die meisten Wirtszellen werden mit den oben beschriebenen Expressions- oder Kloniervektoren der Erfindung transfiziert und vorzugsweise transformiert und in herkömmlichem Nährstoffmedium kultiviert, das gegebenenfalls zwecks Induktion von Promotoren, Selektion von Transformanten oder Amplifizieren der für die erwünschten Sequenzen kodierenden Gene modifiziert ist.

[0155] Transfektion bezieht sich auf die Aufnahme eines Expressionsvektors durch eine Wirtszelle, gleichgültig ob Kodiersequenzen exprimiert werden oder nicht. Es sind Fachleuten auf dem Gebiet der Erfindung zahlreiche Transfektionsverfahren bekannt, z. B. $CaPO_4$ und Elektroporation. Erfolgreiche Transfektion erkennt man im Allgemeinen dann, wenn Hinweise auf das Funktionieren dieses Vektors innerhalb der Wirtszelle auftreten.

[0156] Die Transformation bedeutet die Einsetzung von DNA in eine Organismus, sodass die DNA entweder als extrachromosomales Element oder durch chromosomale Integration replizierbar ist. Je nach der verwendeten Wirtszelle erfolgt die Transformation unter Anwendung von für solche Zellen geeigneten Standardtechniken. Die Calciumbehandlung unter Einsatz von Calciumchlorid, beschrieben in Kapitel 1.82 von Sambrook et al., s. o., wird im Allgemeinen für Prokaryoten oder andere Zellen verwendet, die substanzielle Zellwand-Barrieren aufweisen. Die Infektion mit Agrobacterium tumefaciens dient zur Transformation bestimmter Pflanzenzellen, wie dies von Shaw et al., Gene 23, 315 (1983) und WO 89/05859 vom 29. Juni 1989 beschrieben ist. Für Säugetierzellen ohne solche Zellwände ist das Calciumphosphat-Fällungsverfahren vorzuziehen, das in den Kapiteln 16.30–16.37 von Sambrook et al., s. o., erläutert ist. Allgemeine Aspekte von Transformationen mit Saugetierzellen-Wirtssystemen wurden von Axel im US-Patent 4.399.216 vom 16. August 1983 beschrieben. Transformationen in Hefe erfolgen typischerweise gemäß dem Verfahren von Van Solingen et al., J. Bact. 130, 946 (1977) und Hsiao et al., Proc. Natl. Acad. Sci. USA 76, 3829 (1979). Andere Verfahren zum Einsetzen von DNA in Zellen, z. B. durch Kerninjektion, Elektroporation oder Protoplasten-Fusion, kommen ebenfalls in Frage.

E. Kultivieren der Wirtszellen

[0157] Prokaryotische Zellen zur Produktion von HRG-Polypeptid der Erfindung werden in geeigneten Medien kultiviert, wie dies von Sambrook et al., s. o., beschrieben ist.

[0158] Die Säugetier-Wirtszellen zur Produktion von HRG der Erfindung können in einer Vielzahl an Medien kultiviert werden. Im Handel erhältliche Medien wie z. B. Ham's F10 (Sigma), Minimal Essential Medium (MEM, Sigma), RPMI-1640 (Sigma) und Dulbecco's Modified Eagle's Medium (DMEM, Sigma) eignen sich zum Kultivieren der Wirtszellen. Außerdem eignen sich die von Ham und Wallace, Meth. Enz. 44 (1979); Barnes und

Sato, Anal, Biochem. 102, 255 (1980); US-Patenten 4.767.866, 4.657.866, 4.927.762 oder 4.560.655; WO 90/03430; WO 87/00195 und US-Patent Re. 30.985 beschriebenen Medien als Kulturmedien für die Wirtszellen. Jedes dieser Medien kann gegebenenfalls mit Hormonen und/oder anderen Wachstumsfaktoren (z. B. Insulin, Transferrin oder Epidermis-Wachstumsfaktor), Salzen (z. B. Natriumchlorid, Calcium, Magnesium und Phosphat), Puffern (z. B. HEPES), Nucleosiden (z. B. Adenosin und Thymidin), Antibiotika (z. B. dem Medikament Gentamycin™), Spurenelementen (als anorganische Verbindungen definiert, die üblicherweise in Endkonzentrationen im mikromolekularen Bereich vorhanden sind) und Glucose oder einer äquivalenten Energiequelle ergänzt sein. Allfällige andere notwendige Ergänzungsstoffe können auch in geeigneten Konzentrationen enthalten sein, die für Fachleute auf dem Gebiet offenkundig sind. Die Kulturbedingungen, z. B. Temperatur, pH-Wert u. dgl., sind jene, die bereits für die zur Expression ausgewählten Wirtszelle herrschten und sind für Fachleute auf dem Gebiet ebenfalls offenkundig.

[0159] Die hierin angeführten Wirtszellen umfassen Zellen in In-vitro-Kultur ebenso wie Zellen, die sich in einem Wirtstier befinden.

[0160] Es ist weiters möglich, dass HRG der Erfindung durch homologe Rekombination oder mittels rekombinanter Produktionsverfahren unter Einsatz von Steuerelementen erzeugt wird, die in Zellen eingeführt werden, die für aktuell in der Praxis verwendetes HRG kodierende DNA enthalten. Beispielsweise wird ein wirkungsvolles Promotor/Enhancer-Element, ein Suppressor oder ein exogenes Transkriptions-Modulationselement in das Genom der intendierten Wirtszelle an einer nahen Position und in einer Orientierung insertiert, die ausreichen, um die Transkription von für das erwünschte HRG kodierender DNA zu beeinflussen. Das Steuerelement kodiert nicht für HRG der Erfindung, doch die DNA ist im Wirtszellengenom enthalten. Man screent als nächstes hinsichtlich Zellen, die HRG der Erfindung produzieren, oder erhöhter oder herabgesetzter Expressionswerte.

F. Detektieren von Gen-Amplifikation/Expression

[0161] Die Gen-Amplifikation und/oder -Expression kann in einer Probe direkt gemessen werden, z. B. durch herkömmliches Southern-Blotting, Northern-Blotting zur Quantifizierung der Transkription von mRNA (Thomas, Proc. Natl. Acad. Sci. USA 77, 5201–5205 [1980]), Dot-Blotting (DNA-Analyse) oder in situ-Hybridisierung; dabei kommt eine zweckmäßig markierte Sonde auf der Basis der hierin angeführten Sequenzen zum Einsatz. Es können verschiedene Marker verwendet werden, am häufigsten Radioisotope, insbesondere ^{32}P. Andere Techniken kommen allerdings auch in Frage, z. B. Biotin-modifizierte Nucleotide für die Einsetzung in ein Polynucleotid. Das Biotin dient dann als Stelle für die Bindung an Avidin oder Antikörper, die mit eine Vielzahl an Markern wie z. B. Radionucliden, Fluoreszenzmitteln, Enzymen o. dgl. markiert sein können. Alternativ dazu können Antikörper verwendet werden, die spezifische Doppelstränge, z. B. DNA-Doppelstränge, RNA-Doppelstränge und DNA-RNA-Hybrid-Doppelstränge oder DNA-Protein-Doppelstränge erkennen können. Die Antikörper können ihrerseits markiert und der Assay durchgeführt werden, wenn der Doppelstrang an eine Oberfläche gebunden ist, sodass nach Bildung des Doppelstrangs auf der Oberfläche die Gegenwart von an den Doppelstrang gebundenem Antikörper detektierbar ist.

[0162] Die Gen-Expression kann alternativ dazu durch immunologische Verfahren wie z. B. immunhistochemische Einfärbung von Gewebeabschnitten und Testen von Zellkultur oder Körperflüssigkeiten gemessen werden, um die Expression des Genprodukts direkt zu quantifizieren. Bei immunohistochemischen Einfärbungstechniken wird eine Probe typischerweise durch Dehydrierung und Fixierung gebildet, gefolgt von Reaktion mit markierten Antikörpern, die für das Genprodukt spezifisch sind, das dort gekoppelt ist, wo die Marker üblicherweise visuell detektierbar sind, z. B. enzymatische Marker, fluoreszierende Marker, lumneszierende Marker u. dgl. Eine besonders sensitive und hierin geeignete Einfärbungstechnik ist von Hsu et al., Am. J. Clin. Path. 75, 734–738 (1980) beschrieben.

[0163] Antikörper, die sich für immunhistochemisches Einfärben und/oder Testen von Probenflüssigkeiten eignen, können entweder monoklonal oder polyklonal sein und in jedem beliebigen Tier gebildet werden. Günstigeweise können Antikörper gegen ein natives HRG-Polypeptid oder ein synthetisches Peptid auf der Basis der hierin geoffenbarten DNA-Sequenzen gebildet werden, wie dies ausführlicher im nachstehenden Abschnitt 4 beschrieben ist.

G. Reinigung des HRG-Polypeptids

[0164] HRG wird aus einer zellulären Membranfraktion gewonnen. Alternativ dazu wird ein proteolytisch gespaltenes oder gekürztes lösliches HRG-Fragment oder eine Subdomäne aus dem Kulturmedium als lösliche Polypeptide gewonnen.

[0165] Wenn HRG in einer anderen Pekombinänten Zelle als aus menschlichem Ursprung exprimiert wird, ist HRG völlig frei von Proteinen oder Polypeptiden menschlichen Ursprungs. Es ist jedoch wünschenswert, HRG aus rekombinanten Zellproteinen oder Polypeptiden zu reinigen, um Präparate zu erhalten, die im Wesentli-

chen homogen in Bezug auf HRG sind. In einem ersten Schritt wird das Kulturmedium oder Lysat zentrifugiert, um Zellbruchstücke zu entfernen. Die Membran und lösliche Proteinfraktionen werden anschließend getrennt. HRG wird dann sowohl von der löslichen Proteinfraktion (erfordert die Gegenwart einer Protease) als auch von der Membranfraktion des Kulturlysats gereinigt (dies hängt davon ab, ob HRG membrangebunden ist). Die folgenden Vorgangsweisen sind Beispiele für geeignete Reinigungsverfahren: Fraktionierung auf Immunaffinitäts- oder Ionenaustauschsäulen; Ethanolfällung, RP-HPLC; Chromatographie auf Kieselsäure, Heparin-Sepharose oder einem Kationenaustauschharz wie z. B. DEAE; Chromatofokussieren; SDS-PAGE; Ammoniumsulfat-Fllung; und Gelfiltration unter Einsatz von z. B. Sephadex G-75.

[0166] HRG-Varianten, in denen Reste deletiert, insertiert oder substituiert wurden, werden in gleicher Weise wie natives HRG gewonnen, wobei allfällig substanzielle Änderungen der Eigenschaften aufgrund der Variation berücksichtigt werden. Beispielsweise erleichtert die Herstellung einer HRG-Fusion mit einem anderen Protein oder Polypeptid, z. B. einem bakteriellen oder viralen Antigen, die Reinigung; eine Immunaffinitätssäule mit einem Antikörper gegen das Antigen kann zur Absorption der Fusion dienen. Immunaffinitätssäulen wie z. B. eine polyklonale Kaninchen-Anti-HRG-Säule können dazu dienen, HRG-Variante durch deren Bindung an zumindest ein verbleibendes Immunepitop zu absorbieren. Ein Protease-Hemmer wie z. B. Phenylmethylsulfonylfluorid (PMSF) kann auch dazu verwendet werden, proteolytischen Abbau während der Reinigung zu hemmen, wobei Antibiotika enthalten sein können, um das Wachstum hinzukommender Schmutzstoffe zu verhindern. Es ist für Fachleute auf dem Gebiet der Erfindung offenkundig, dass Reinigungsverfahren für natives HRG möglicherweise modifiziert werden müssen, um Veränderungen der Eigenschaften von HRG-Varianten nach Expression in rekombinanter Zellkultur zu berücksichtigen.

H. Kovalente Modifikationen von HRG

[0167] Kovalente Modifikationen von HRG-Polypeptiden sind im Schutzbereich der Erfindung enthalten. Sowohl natives HRG als auch Aminosäuresequenz-Varianten von HRG sind gegebenenfalls kovalent modifiziert. Eine Art von kovalenter Modifikation, die ebenfalls im Schutzbereich der Erfindung enthalten ist, ist ein HRG-Polypeptidfragment. HRG-Fragmente wie z. B. HRG-GDF mit bis zu etwa 40 Aminosäureresten werden günstigerweise durch chemische Synthese oder durch enzymatische oder chemische Spaltung des HRG-Polypeptids voller Länge oder der HRG-Polypeptidvariante erzeugt. Andere Arten kovalenter Modifikationen von HRG oder seiner Fragmente werden in das Molekül inkorporiert, indem abgezielte Aminosäurereste von HRG oder Fragmenten davon mit einem organischen Derivatisierungsmittel umgesetzt werden, das zur Reaktion mit ausgewählten Seitenketten oder den N- oder C-terminalen Resten fähig ist.

[0168] Cysteinylreste werden am häufigsten mit α-Haloacetaten (und entsprechenden Aminen) wie z. B. Chloressigsäure oder Chloracetamid umgesetzt, um Carboxymethyl- oder Carboxyaminomethyl-Derivate zu ergeben. Cysteinylreste werden auch durch Reaktion mit Bromtrifluoraceton, α-Brom-β-(5-imidozoyl)propionsäure, Chloracetylphosphat, N-Alkylmaleimiden, 3-Nitro-2-pyridyldisulfid, Methyl-2-pyridyldisulfid, p-Chlormercuribenzoat, 2-Chlormercuri-4-nitrophenol oder Chlor-7-nitrobenzo-2-oxa-1,3-diazol derivatisiert.

[0169] Histidylreste werden durch Reaktion mit Diethylpyrocarbonat bei pH 5,5–7,0 derivatisiert, da dieses Mittel für die Histidyl-Seitenkette relativ stabil ist. p-Bromphenacylbromid ist auch geeignet; die Reaktion erfolgt vorzugsweise in 0,1 M Natriumcacodylat bei pH 6,0.

[0170] Lyisnyl- und aminoterminale Reste werden mit Bernstein- oder anderen Carbonsäureanhydriden umgesetzt. Die Derivatisierung dieser Mittel bewirkt die Umkehr der Ladung der Lysinylreste. Andere zweckmäßige Reagenzien für die Derivatisierung von α-Aminohältigen Resten sind Imidoester wie etwa Methylpicolinimidat; Pyridoxalphosphat; Pyridoxal; Chlorborhydrid; Trinitröbenzolsulfonsäure; O-Methylisoharnstoff; 2,4-Pentandion; und Transaminase-katalysierte Reaktion mit Glyoxylat.

[0171] Arginylreste werden durch Reaktion mit einem oder mehreren herkömmlichen Reagenzien modifiziert, z. B. Phenylglyoxal, 2,3-Butandion, 1,2-Cyclohexandion und Ninhydrin. Die Derivatisierung von Argininresten erfordert aufgrund des hohen pK_a der funktionellen Guanidingruppe , dass die Reaktion unter alkalischen Bedingungen stattfindet. Außerdem können diese Reagenzien mit den Lysingruppen sowie der Arginin-ε-Aminogruppe reagieren.

[0172] Die spezifische Modifikation von Tyrosylresten kann unter besonderer Berücksichtigung der Einführung spektraler Marker in Tyrosylreste durch Reaktion mit aromatischen Diazonium-Verbindungen oder Tetranitromethan erfolgen. Am häufigsten werden N-Acetylimidazol und Tetranitromethan dazu verwendet, O-Acetyltyrosyl-Spezies bzw. 3-Nitroderivate zu erzeugen. Tyrosylreste werden unter Einsatz von ^{125}I oder ^{131}I iodiert, um markierte Proteine zur Verwendung im Radioimmunassay zu herzustellen, wobei sich hier das oben beschriebene Chloramin T-Verfahren eignet.

[0173] Carboxyl-Seitengruppen (Aspartyl oder Glutamyl) werden selektiv durch Reaktion mit Carbodiimiden (R'-N=C=N-R') modifiziert, worin R und R' unterschiedliche Alkylgruppen sind, z. B. 1-Cyclohexyl-3-(2-morpholinyl-4-ethyl)carbodiimid oder 1-Ethyl-3-(4-azonia-4,4-dimethylpenyl)carbodiimid. Außerdem werden Aspartyl- und Glutamylreste durch Reaktion mit Ammoniumionen in Asparaginyl- und Glutaminylreste umgewandelt.

[0174] Die Derivatisierung mit bifunktionalen Mitteln eignet sich zur Vernetzung von HRG mit einer wasserunlöslichen Trägermatrix oder Oberfläche zur Verwendung im Verfahren zur Reinigung von Anti-HRG-Antikörpern. Üblicherweise verwendete Vernetzer sind z. B. 1,1-Bis(diazoacetyl)-2-phenylethan, Glutaraldehyd, N-Hydroxysuccinimidester, z. B. Ester mit 4-Azidosalicylsäure, homobifunktionale Imidoester, z. B. Disuccinimidylester wie etwa 3,3'-Dichiobis(succinimidylpropionat) und bifunktionale Maleimide wie etwa Bis-N-maleimido-l,8-octan. Derivatisierungsmittel wie z. B. Methyl-3-[(p-azidophenyl)dithio]propioimidat liefern photoaktivierbare Zwischenprodukte, die zur Bildung von Vernetzungen in Gegenwart von Licht fähig sind. Alternativ dazu werden reaktive wasserunlösliche Matrizen wie etwa Bromcyan-aktivierte Kohlenhydrate und die reaktiven Substrate der US-Patente 3.969.287, 3.691.0165, 4.195.128, 4.247.642, 4.229.537 und 4.330.440 zwecks Protein-Immobilisierung verwendet.

[0175] Glutaminyl- und Asparaginylreste werden häufig zu den entsprechenden Glutamylbzw. Aspartylresten deamidiert. Alternativ dazu werden diese Reste unter schwach sauren Bedingungen deamidiert. Beide Formen dieser Reste fallen in den Schutzumfang der Erfindung.

[0176] Andere Modifikationen sind die Hydroxylierung von Prolin und Lysin, die Phosphorylierung von Hydroxylgruppen oder Seryl- oder Threonylresten, die Methylierung der α-Aminogruppen von Lysin-, Arginin- und Histidin-Seitenketten (T. E. Creighton, Protein: Structure and Molecular Properties, W. H. Freeman & Co., San Francisco, S. 79–86 [1983]), die Acetylierung des N-terminalen Amins und die Amidierung beliebiger C-terminaler Carboxylgruppen.

[0177] HRG ist gegebenenfalls mit einem Polypeptide, das zu HRG heterolog ist, fusioniert. Das heterologe Polypeptid ist gegebenenfalls eine Ankersequenz, wie sie z. B. im Zerfalls-Beschleunigungssystem (DAF) vorkommt; ein Toxin wie z. B. Ricin, Pseudomonas-Exotoxin, Gelonin oder ein andres Polypeptid, das zum Tod der Zielzelle führt. Diese heterologen Polypeptide sind durch Seitenketten oder die endständigen Reste kovalent an HRG gebunden. In ähnlicher Weise ist HRG an andere Moleküle gebunden, die gegenüber einer Ziel-Säugetierzelle toxisch oder inhibitorisch sind, z. B. Tricothecene oder Antisense-DNA, die Expression von Ziel-Genen blockiert.

[0178] HRG wird auch durch Verandern seines nativen Glykosylierungsmusters kovalent modifiziert. Ein oder mehrere Kohlenhydrat-Substituenten werden modifiziert, indem die Monosaccharid-Komponenten an einer bestimmten Stelle hinzugefügt, entfernt oder variiert werden oder indem die Reste in HRG durch Zugabe oder Deletion von Glykosylierungsstellen modifiziert werden.

[0179] Die Glykosylierung von Polypeptiden ist typischerweise N- oder O-gebunden. N-gebunden bezieht sich auf die Befestigung der Kohlenhydratgruppe an der Seitenkette eines Asparaginrests. Die Tripeptidsequenzen Asparagin-X-Serin und Asparagin-X-Threonin, worin X jede beliebige Aminosäure mit Ausnahme von Prolin ist, sind die Erkennungssequenzen für enzymatische Bindung der Kohlenhydratgruppe an der Asparagin-Seitenkette. Somit schafft die Gegenwart einer dieser Tripeptidsequenzen in einem Polypeptid eine potenzielle Glykosylierungsstlle. O-gebundene Glykosylierung bezieht sich auf die Bindung eines der Zucker N-Acetylgalactosamin, Galactose oder Xylose an eine Hydroxyaminosäure, am häufigsten Serin oder Threonin, obwohl 5-Hydroxyprolin oder 5-Hydroxylysin auch in Frage kommen.

[0180] Glykosylierungsstellen werden HRG hinzugefügt, indem seine Aminosäuresequenz solcherart geändert wird, dass sie eine oder mehrere der oben angeführten Tripeptidsequenzen enthält (für N-gebundene Glykosylierungsstellen). Die Modifikation kann auch durch Addition von – oder Substitution mit – einem oder mehreren Serin- oder Threoninresten an HRG erfolgen (für O-gebundene Glykosylierungsstellen). Aus praktischen Gründen wird HRG vorzugsweise mittels Veränderungen auf DNA-Ebene modifiziert, insbesondere durch Mutieren der für HRG kodierenden DNA in vorausgewählten Basen, sodass Codons erzeugt werden, die in die erwünschten Aminosäuren translatieren.

[0181] Chemische oder enzymatische Kopplung von Glykosiden an HRG erhöht die Anzahl an Kohlenhydrat-Substituenten. Diese Vorgangsweisen sind insofern vorteilhaft, als sie keine Produktion des Polypeptids in einer Wirtszelle erfordern, die zur N- und O-gebundenen Glykosylierung fähig ist. Je nach der angewendeten Kopplungsmethode kann bzw. können der bzw. die Zucker an (a) Arginin und Histidin, (b) freie Carboxylgruppen, (c) freie Sulfhydrylgruppen wie z. B. jene von Cystein, (d) freie Hydroxylgruppen wie z. B. jene von Serin, Threonin oder Tryptophan oder (f) die Amidgruppe von Glutamin gebunden werden. Diese Verfahren sind in WO 87/05330 vom 11. September 1987 und in Aplin und Wriston, CRC Crit. Rev. Biochem, 259–306 (1981) beschrieben.

[0182] Auf einem HRG vorhandene Kohlenhydratgruppen werden ebenfalls chemisch oder enzymatisch entfernt. Die chemische Deglykosylierung erfordert, dass man das Polypeptid der Verbindung Trifluormethansulfonsäure oder einer äquivalenten Verbindung aussetzt. Diese Behandlung führt zur Spaltung der meisten oder aller Zucker mit der Ausnahme des Verbindungszuckers (N-Acetylglucosamin oder N-Acetylgalactosiamin), während das Polypeptid intakt bleibt. Die chemische Deglykosylierung wird von Hakimuddin et al., Arch. Biochem. Biophys. 259, 52 (1987) und von Edge et al., Anal. Biochem. 118, 131 (1981) beschrieben. Kohlenhydratgruppen werden mittels einer Vielzahl an Endo- und Exo-Glykosidasen aus HRG entfernt; siehe Thotakura et al., Meth. Enzymol. 138, 350 (1987).

[0183] Glykosylierung während der Expression in Zellen wird auch durch Tunicamycin unterdrückt, wie dies von Duskin et al., J. Biol. Chem. 257, 3105 (1982) erläutert wird. Tunicamycin blockiert die Bildung von Protein-N-Glykosid-Verknüpfungen.

[0184] HRG wird auch durch Binden von HRG an verschiedene nicht-proteinhältige Polymere modifiziert, z. B. Glykol, Polypropylenglykol oder Polyoxyalkylene, wie dies in den US-Patenten 4.640.835, 4.496.689, 4.301.144, 4.670.417, 4.791.192 oder 4.179.337 beschrieben ist.

[0185] Eine bevorzugte Möglichkeit der Verlängerung der In-vivo-Halbwertszeit von nichtmembrangebundenem HRG im Kreislauf ist seine Bindung an ein Polymer, das für verlängerte Halbwertszeit sorgt, z. B. Polyethylenglykol (PEG); siehe Maxfield et al., Polymer 16, 505–509 (1975; Bailey, F. E. et al., in Nonionic Surfactants (Schick, M. J., Hg.), S. 794–821 (1967); Abushowski, A. et al., J. Biol. Chem. 252, 3582–3586 (1977); Abuchowski, A. et al., Cancer Biochem. Biophys. 7, 175–186 (1984); Katre, N. V. et al., Proc. Natl. Acad. Sci. 84, 1487–1491 (1987); Goodson, R. et al., Bio/Technology 8, 343–346 (1990). Die Bindung von PEG konnte einer Publikation zufolge die Immunogenität und Toxizität reduzieren (Abuchowski, A. et al., J. Biol. Chem. 252, 3578–3581 [1977]).

[0186] HRG ist auch in Mikrokapseln, die z. B. durch Koazervationstechniken oder durch Grenzflächen-Polymerisation erzeugt werden (z. B. Hydroxymethylcellulose- oder Gelatine-Mikrokapseln bzw. Poly-(methylmethacrylat)-Mikrokapseln), in kolloidalen Medikamenten-Zufuhrsystemen (z. B. Liposomen, Alumin-Mikrokügelchen, Mikroemulsionen, Nano-Teilchen und Nano-Kapseln) oder in Makroemulsionen eingeschlossen. Solche Techniken sind in Remington's Pharamceutical Sciences, 16. Ausgabe, Osol, A., Hg., (1980) geoffenbart.

[0187] HRG eignet sich zur Erzeugung von Antikörpern als Standards in Tests auf HRG (z. B. durch Markieren von HRG zur Verwendung als Standard in einem Radioimmunassay, einem Enzym-gebundenen Immunassay oder einem Radiorezeptor-Assay), in Affinitäts-Reinigungsverfahren und in kompetitiven Rezeptor-Bindungsassay beim Markieren mit Radioiod, Enzymen, Fluorophoren, Spin-Markern u. gl.

[0188] Fachleute auf dem Gebiet können Varianten screenen, um die optimale Variante für den jeweiligen Zweck auszuwählen. Beispielsweise wird eine Veränderung der immunologischen Eigenschaft von HRG, z. B. eine Änderung der Affinität für ein bestimmtes Antigen im HER2-Rezeptor, durch einen kompetitiven Immunassay unter Verwendung eines Standards oder Vergleichs wie z. B. eines nativen HRG (insbesondere einer nativen HRG-GFD) gemessen. Andere potenzielle Modifikationen von Protein- oder Polypeptid-Eigenschaften wie z. B. Redox- oder thermische Stabilität, Hyrophobie, Anfälligkeit gegenüber proteolytischem Abbau, Stabilität in rekombinanter Zellkultur oder in Plasma oder die Neigung zur Aggregation mit Trägern oder in Multimere werden durch auf dem Gebiet bekannte Verfahren untersucht.

1. Therapeutische Verwendung von Heregulin-Liganden

[0189] Zwar ist die Rolle von $p185^{HER2}$ und seiner Liganden im normalem Zellwachstum und in der Differenzierung unbekannt, doch ist es ein Ziel der Erfindung, therapeutische Einsatzmöglichkeiten für die $p185^{HER2}$-Liganden der Erfindung zur Förderung von normalem Wachstum und normaler Entwicklung sowie zur Hemmung von abnormalem Wachstum, insbesondere in malignen oder neoplastischen Geweben, zu identifizieren.

2. Therapeutische Zusammensetzungen und Verabreichung von HRG

[0190] Therapeutische Formulierungen von HRG und HRG-Antikörper werden durch Vermischen des HRG-Proteins mit dem erwünschten Reinheitsgrad mit fakultativen physiologisch akzeptablen Trägern, Exzipienten oder Stabilisatoren (Remington's Pharmaceutical Sciences, s. o.) in Form von lyophilisiertem Kuchen oder wässrigen Lösungen auf die Lagerung vorbereitet. Annehmbare Träger, Exzipienten oder Stabilisatoren sind für die Rezipienten in den verwendeten Dosierungen und Konzentrationen nicht-toxisch; Beispiele sind Puffer wie etwa Phosphat, Citrat und andere organische Säuren; Antioxidanzien wie etwa Asorbinsäure; niedermolekulare (weniger als etwa 10 Reste) Polypeptide (zur Verhinderung von Methoxid-Bildung); Proteine wie etwa Serumalbumin, Gelatine oder Immunglobuline; hydrophile Polymere wie etwa Polyvnylpyrrolidon; Aminosäuren wie etwa Glycin, Glutamin, Asparagin, Arginin oder Lysin; Monosaccharide, Disaccharide oder andere Kohlenhydrate wie etwa Glucose, Mannose oder Dextrine; Chelatbildner wie etwa EDTA; Zuckeralkohole wie etwa Mannit oder Sorbit; salzbildenden Gegenionen wie etwa Natrium; und/oder nicht-ionische Tenside wie z. B. Tween, Pluronics oder Polyethylenglykol (PEG).

[0191] HRG oder HRG-Antikörper muss für die In-vivo-Verabreichung steril sein. Dies wird durch Filtration durch sterile Filtrationsmembranen vor oder während der Lyophilisierung und Rekonstitution erreicht. HRG oder HRG-Antikörper wird typischerweise in lyophilisierter Form oder in Lösung gelagert.

[0192] Therapeutisches HRG oder HRG-spezifische Antikörper-Zusammensetzungen werden im Allgemeinen in einem Behälter mit steriler Zugriffsöffnung, z. B. einem Beutel für intravenöse Lösungen oder einem Fläschchen mit einem Stopfen, der mit einer Nadel für subkutane Injektionen durchgestochen werden kann ist, aufbewahrt.

[0193] HRG, sein Antikörper oder eine HRG-Variante kann bei Verwendung als Antagonist gegebenenfalls mit anderen Mitteln, die zur Verwendung für die Behandlung von Malignität bekannt sind, kombiniert oder gemeinsam mit ihnen verabreicht werden. Wenn HRG als Agonist zur Stimulierung des HER2-Rezeptors verwendet wird (z. B. in Gewebekulturen), kann es mit anderen Wachstum stimulierenden Zusammensetzungen kombiniert oder gemeinsam mit ihnen verabreicht, z. B. PDGF, FGF, EGF, Wachstumshormon oder andere Proteinwachstumsfaktoren.

[0194] Die Verabreichung von HRG oder HRG-Antikörper erfolgt in Einklang mit bekannten Verfahren, z. B. Injektion oder Infusion auf intravenösem, intraperitonealem, intrazerebralem, intramuskulärem, intraokularem, intraarteriellem oder intraläsionalem Weg oder durch Retard-Freisetzungssysteme (siehe unten). HRG wird kontinuierlich durch Infusion oder durch Bolusinjektion verabreicht. HRG-Antikörper wird in gleicher Weise oder durch Verabreichung in den Blutstrom oder die Lymphe verabreicht.

[0195] Geeignete Beispiee für Retard-Freisetzungspräparate sind halbdurchlässige Matrizen fester hydrophober Polymere mit dem Protein, welche Matrizen als Formteile wie z. B. Filme oder Mikrokapseln vorliegen. Beispiele für Retard-Freisetzungsmatrizen sind Polyester, Hydrogele (z. B. Poly(2-hydroxyethyelmethacrylat, z. B. beschrieben von Langer et al., J. Bomed. Mater. Res. 15, 167–277 [1981] und Langer, Chem. Tech. 12, 98–105 [1982] oder Poly(vinylalkohol), Polylactide (US-Patent 3.773.919, EP 58.481), Copolymere von L-Glutaminsäure und γ-Ethyl-L-glutamat (Sidman et al., Biopolymers 22, 547-556 [1983]), nicht-abbaubares Ethylenvinylacetat (Langer et al., s. o.), abbaubare Milchsäure-Glykolsäure-Copolymere wie z. B. Lupron Depot™ (injizierbare Mikrokügelchen, die aus Milchsäure-Glykolsäure-Copolymer und Leuprolidacetat bestehen) und Poly-D(–)-3-hydroxybuttersäure (EP 133,988). Während Polymere wie z. B. Ethylenvinylacetat und Milchsäure-Glykolsäure die Freisetzung von Molekülen über einen Zeitraum von mehr als 100 Tagen ermöglichen, setzen bestimmte Hydrogele Proteine über kürzere Zeiträume frei. Wenn eingekapselte Proteine über lange Zeit im Körper verbleiben, können sie infolge des Kontakts mit Feuchtigkeit bei 37°C denaturieren und aggregieren, was zu einem Verlust biologischer Aktivität und möglicherweise zu Veränderungen der Immunogenität führt. Es können je nach dem betreffenden Mechanismus vernünftige Strategien für die Proteinstabilisierung entwickelt werden. Wenn man z. B. entdeckt, dass der Aggregationsmechanismus die Entstehung einer intermolekularen S-S-Bindung durch Thio-Disulfid-Austausch ist, kann Stabilisierung durch Modifizieren von Sulfhydrylresten, Lyophilisieren aus sauren Lösungen, Regulieren des Feuchtigkeitsgehalts, Einsatz geeigneter Additive und Entwicklung spezifischer Polymermatrix-Zusammensetzungen erzielt werden.

[0196] Retard-Freisetzungs-HRG oder Antikörper-Zusammensetzungen enthalten auch liposomal eingeschlossenes HRG oder Antikörper. Liposome mit HRG oder Antikörper werden durch an sich bekannte Verfahren erzeugt; siehe DE 3.218.121; Epstein et al., Proc. Natl. Acad. Sci. USA 82, 3688–3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77, 4030–4034 (1980); EP 52.322; EP 36.676; EP 143.949; EP 142.641; JP-Patent 83-118008; US-Patente 4.485.045 und 4.544.545; und EP 102.324. Üblicherweise sind die Liposome klein (etwa 200–800 Angstrom) und unilamellar, wobei der Lipidanteil mehr als etwa 30 Mol.-% Cholesterin beträgt und der ausgewählte Anteil auf die optimale HRG-Therapie abgestimmt ist. Liposome mit verbesserter Zirkulationszeit sind in US-Patent 5.013.556 geoffenbart.

[0197] Eine weitere Verwendungsmöglichkeit der Erfindung umfasst das Inkorporieren von HRG-Polypeptid oder Antikörper in Formteile. Solche Teile können zum Modulieren von Zellwachstum und Zellentwicklung verwendet werden. Außerdem können Zellwachstum, Zellteilung und das Eindringen von Tumor mittels dieser Formteile moduliert werden.

[0198] Eine wirkungsvolle therapeutisch zu verwendende Menge von HRG oder Antikörper hängt z. B. von den therapeutischen Zielen, der Verabreichungsweise und dem Zustand des Patienten ab. Demzufolge ist es notwendig, dass der Therapeut die Dosierung titert und die Verabreichungsmethode gegebenenfalls ändert, um den optimalen therapeutischen Nutzen zu erzielen. Eine typische Tagesdosis reicht z. B. von etwa 1 μg/kg bis zu 100 mg/kg oder mehr (in Abhängigkeit von den oben angeführten Faktoren). Typischerweise verabreicht der Kliniker HRG oder Antikörper, bis eine Dosierung erreicht ist, die den erwünschten Effekt hat. Der Therapieverlauf wird mittels herkömmlicher Tests problemlos überwacht.

3. Heregulin-Antikörper-Herstellung und therapeutische Nutzung

[0199] Die Antikörper der Erfindung werden durch routinemäßiges Screenen erhalten. Polyklonale Antikörper gegen HRG werden im Allgemeinen durch mehrere subkutane (sc) oder intraperitoneale (ip) Injektionen von HRG und eines Adjuvans gebildet. Es kann nützlich sein, HRG oder ein HRg-Fragment mit der Ziel-Aminosäuresequenz an ein Protein zu binden, das in der zu immunisierenden Spezies immunogen ist, z. B. Napfschnecken-Hämocyanin, Serumalbumin, Rinder-Thyroglobulin oder Sojabohnen-Trypsin-Inhibitor; dies erfolgt unter Einsatz eins bifunktionalen oder derivatisierenden Mittels, z. B. Maleimidobenzoylsulfosuccinimidester (Konjugation durch Cysteinreste), N-Hydroxysuccinimid (durch Lysinreste), Glutaraldehyd, Bernsteinsäuranhydrid, $SOCl_2$ oder $R^1N=C=NR$, worin R und R^1 unterschiedlice Alkylgruppen sind.

[0200] Die Methode und der Zeitplan der Immunisierung eines Tiers oder der Entnahme und Kultivierung von

Antikörper-produzierenden Zellen entsprechen im Allgemeinen bekannten und herkömmlichen Techniken der Antikörper-Stimulierung und -Produktion. Während Mäuse häufig immunisiert werden, ist davon auszugehen, dass beliebige Säugetierprobanden (einschließlich menschlicher Probanden oder aus ihnen erhaltene Antikörperproduzierende Zellen) immunisiert werden können, um Antikörperproduzierende Zellen zu erzeugen.

[0201] Die Probanden werden typischerweise gegen HRG oder seine immunogenen Konjugate oder Derivate immunisiert, indem 1 mg oder 1 μg HRG-Immunogen (für Kaninchen bzw. Mäuse) mit 3 Volumina komplettem Freund'schem Adjuvans kombiniert werden und die Lösung intradermal an mehreren Stellen injiziert wird. Einen Monat später erhalten die Versuchstiere erneut eine Injektion mit 1/5 bis 1/10 der ursprünglichen Menge an Immunogen in komplettem Freund'schem Adjuvans durch subkutane Injektion an mehreren Stellen. 7 bis 14 Tage später wird den Tieren Blut entnommen und das Serum auf Anti-HRG-Antikörper-Titer getestet. Die Tiere erhalten Auffrischungsimpfung, bis der Titer ein Plateau erreicht. Vorzugsweise erhalten die Versuchstiere mit einem Konjugat des gleichen HRG ihre Auffrischungsimpfung, jedoch an ein unterschiedliches Protein und/oder durch einen unterschiedlichen Vernetzer gebunden. Konjugate können in rekombinanter Zellkultur als Proteinfusionen erzeugt werden. Aggregationsmittel wie z. B. Alaun dienen dazu, die Immunreaktion zu steigern.

[0202] Nach der Immunisierung werden monoklonale Antikörper durch Gewinnung von Lymphzellen – typischerweise Milzzellen oder Lymphozyten aus Lymphknotengewebe – aus immunisierten Tieren und Immortalisieren der Zellen in herkömmlicher Weise, z. B. durch Fusion mit Myelomzellen oder durch Epstein-Barr-(EB-) Virus-Transformation und Screenen auf Klone, die den erwünschten Antikörper exprimieren, produziert. Das Hybridom-Verfahren wurde ursprünglich von Kohler und Milstein, Eur. J. Immunol. 6, 511 (1976) beschrieben und wird sehr häufig angewendet, um Hybridzelllinien zu er- zeugen, die hohe Werte monoklonaler Antikörper gegen zahlreiche spezifische Antigene sekretieren.

[0203] Es ist ferner möglich, Zellen einer Spezies mit einer anderen zu fusionieren. Es ist aber vorzuziehen, dass die Quelle des Zellen produzierenden immunisierten Antikörpers und des Myelom die gleiche Spezies ist.

[0204] Hybridom-Zelllinien, die Anti-HRG produzieren, werden durch Screenen der Kulturüberstände auf HRG bindenden Antikörper identifiziert. Dies erfolgt typischerweise durch herkömmliche Immunassays mittels löslicher HRG-Präparate oder durch FACS unter Einsatz von zellgebundenem HRG und markiertem Antikörperkandidaten.

[0205] Die Hybridzelllinien können in vitro in Zellkulturmedium in Kultur gehalten werden. Die Zelllinien der Erfindung können in einer Zusammensetzung, die die kontinuierliche Zelllinie in Hypoxanthin-Aminopterin-Thymidin- (HAT-) Medium umfasst, ausgewählt und/oder gehalten werden. Sobald die Hybridom-Zelllinie festgelegt ist, kann sie in einer Vielzahl nährstoffmäßig adäquater Medien gehalten werden. Außerdem können die Hybrid-Zelllinien in unterschiedlicher Weise konserviert werden, z. B. mittels Gefrieren und Lagern unter flüssigem Stickstoff. Gefrorene Zelllinien können durch erneute Aufnahme der Synthese und Sekretion von monoklonalem Antikörper ad infinitum wiederbelebt und kultiviert werden. Der sekretierte Antikörper wird aus Gewebekulturüberstand durch herkömmliche Verfahren gewonnen, z. B. Fällung, Ionenaustausch-Chromatographie, Affinitätschromatographie o. dgl. Die hierin beschriebenen Antikörper werden aus Hybridom-Zellkulturen durch herkömmliche Verfahren zur Reinigung von IgG oder IgM gewonnen, die bislang dazu dienten, diese Immunglobuline aus gepooltem Plasma zu reinigen, z. B. Ethanol- oder Polyethylenglykol-Fällungsverfahren. Die gereinigten Antikörper werden sterilfiltriert und gegebenenfalls an einen detektierbaren Marker wie z. B. ein Enzym oder einen Spinmarker zur Verwendung in diagnostischen Untersuchungen von HRG in Testproben gebunden.

[0206] Zwar werden routinemäßig monoklonale Mäuse-Antikörper verwendet, doch ist die Erfindung nicht darauf beschränkt; Human-Antikörper können ebenfalls verwendet werden und erweisen sich möglicherweise als bevorzugt. Solche Antikörper können unter Einsatz von Human-Hybridomen erhalten werden (Cote et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, 77 [1985]). Chimäre Antikörper (Cabilly et al., Morrison et al., Proc. Natl. Acad. Sci. USA 81, 6851 [1984]; Neuberger et al., Nature 312, 604 [1984]; und Takeda et al., Nature 314, 452 [(1985]) mit einer variablem Mäuse-Anti-HRG-Region und einer konstanten Humanregion geeigneter biologischer Aktivität (z. B. mit der Fähigkeit, Humankomplement zu aktivieren und ADCC zu mediieren) liegen ebenso im Schutzbereich der Erfindung wie humanisierte Anti-HRG-Antikörper, die durch herkömmliche CRD-Graftingverfahren erzeugt werden.

[0207] Techniken zur Bildung rekombinanter DNA-Versionen der Antigen-Bindungsregionen von Antikörpermolekülen (als Fab- oder variable Regionsfragmente bekannt), die ohne Erzeugung monoklonaler Antikörper auskommen, liegen ebenfalls im Schutzbereich der Erfindung. Man extrahiert Antikörper-spezifische Messenger-RNA-Moleküle aus Immunsystemzellen, die einem immunisierten Versuchstier entnommen wurden, transkribiert diese in komplementäre DNA (cDNA), kloniert die cDNA in ein bakterielles Expressionssystem und trifft hinsichtlich der erwünschten Bindungseigenschaft eine Auswahl. Das Scripps/Stratagene-Verfahren verwendet ein Bakteriophagen-λ-Vektorsystem mit einer Leader-Sequenz, die bewirkt, dass das exprimierte Fab-Protein in den periplasmatischen Raum migriert (zwischen der bakteriellen Zellmembran und der Zellwand) oder sekretiert wird. Man kann eine hohe Anzahl funktionaler Fab-Fragmente rasch produzieren und

screenen, um jene zu identifizieren, die HRG mit den erwünschten Eigenschaften binden.

[0208] Antikörper, die für HRG-α, HRG-β_1, HRG-β_2 und HRG-β_3 spezifisch sind, können in der oben beschriebenen Weise produziert und verwendet werden. HRG-α-, HRG-β_1-, HRGβ_2- und HRG-β_3-spezifische Antikörper der Erfindung unterliegen vorzugsweise keiner Kreuzreaktion mit anderen Mitgliedern der EGF-Familie (**Fig.** 6) oder miteinander.

[0209] Antikörper, die zur spezifischen Bindung an HRG-NTD, NRG-GFD oder HRG-CTP fähig sind, sind von besonderem Interesse. Ebenfalls von Interesse sind Antikörper, die zur spezifischen Bindung an proteolytische Verarbeitungsstellen zwischen der GFD und Transmembran-Domänen fähig sind. Diese Antikörper werden durch an sich herkömmliche Verfahren identifiziert. Beispielsweise erhält man durch eines der obigen Verfahren mittels Immunisierung mit Voll-proHRG eine Bank von Antikörperkandidaten, die zur Bindung an HRG-ECD oder proHRG fähig sind. Sie können dann anhand ihrer Fähigkeit untergliedert werden, sich an die verschiedenen HRG-Domänen zu binden; dies erfolgt unter Heranziehung herkömmlicher Kartierungstechniken. Es ist weniger vorzuziehen, wenn für eine vorbestimmte Domäne spezifische Antikörper anfänglich durch Immunisieren des Versuchstiers mit einem Polypeptid gebildet werden, das im Wesentlichen nur die betreffende Domäne umfasst, z. B. HRg-GFD frei von NTD- oder CTP-Polypeptiden. Diese Antikörper erfordern keine Kartierung, es sei denn die Bindung an ein bestimmtes Epitop ist erwünscht.

[0210] Antikörper, die zur Bindung an proteolytische Verarbeitungsstellen fähig sind, sind von besonderem Interesse. Sie werden folgendermaßen produziert: Durch Immunisieren eines HRG-Fragments, das die CTP-Verarbeitungsstelle enthält, mit intaktem HRG oder mit HRG-NTD-GFD und anschließend durch Screenen hinsichtlich der Fähigkeit, proteolytische Verarbeitung von HRG zum NTD-GFD-Fragment durch rekombinante Wirtszellen oder isolierte Zelllinien, die sonst in der Lage sind, HRG zum Fragment zu verarbeiten, zu blockieren oder zu hemmen. Diese Antikörper eignen sich zur Unterdrückung der Freisetzung von NTD-GFD und daher auch dazu, die Freisetzung von HTD-GFD und die Stimulierung des HER-2-Rezeptors zu verhindern. Sie kommen auch für die Steuerung von Zellwachstum und -Replikation in Frage. Anti-GFD-Antikörper sind aus den gleichen Gründen geeignet, doch sie sind möglicherweise biologisch nicht so wirkungsvoll wie Antikörper gegen eine Verarbeitungsstelle.

[0211] Es werden Antikörper ausgewählt, die zur Bindung an nur eines der Mitglieder der HRG-Familie fähig sind, z. B. HRG-α oder eine der HRG-β-Isoformen. Da jedes der HRG-Familienmitglieder eine eigene GFD-Transmenbran-Domänen-Spaltungsstelle besitzt, können Antikörper, die spezifisch gegen diese nur einmalig vorhandenen Sequenzen gerichtet sind, die hochspezifische Hemmung jeder der GFD oder Verarbeitungsstellen ermöglichen und dadurch die erwünschten biologischen Reaktionen verfeinern. Beispielsweise können Brustkarzinomzellen, die HER-2-abhängig sind, möglicherweise nur durch einen einzelnen GFD-Isotyp aktiviert werden; wenn nicht, kann die aktivierende GFD möglicherweise nur aus einer bestimmten Verarbeitungssequenz stammen – entweder auf der HER-2-tragenden Zelle selbst oder auf einer GFD-erzeugenden Zelle. Diese Identifikation der Ziel-Aktivierungs-GFD oder Verarbeitungsstelle ist eine unkomplizierte Angelegenheit und umfasst die Analyse von HER-2-abhängigen Karzinomen, z. B. durch Analysieren der Gewebe hinsichtlich der Expression eines HRG-Familienmitglieds (würde dann als therapeutisches Ziel dienen). Diese selektiven Antikörper werden in gleicher Weise wie oben produziert – entweder durch Immunisierung mit der Zielsequenz oder -domäne oder durch Selektieren aus einer Bank von Antikörpern mit breiterer Spezifität.

[0212] Wie oben erläutert, sollten die Antikörper hohe Spezifität und Affinität für die Zielsequenz aufweisen. Beispielsweise sollten die gegen GFD-Sequenzen gebildeten Antikörper höhere Affinität für GFD aufweisen, als GFD für den HER-2-Rezeptor besitzt. Solche Antikörper werden durch herkömmliche Screening-Methoden ausgewählt.

4. Nicht-therapeutische Verwendungszwecke von Heregulin und seiner Antikörper

[0213] Die für HRG kodierende Nucleinsäure kann als Diagnostikum für gewebespezifische Typisierung verwendet werden. Beispielsweise können Verfahren wie In-situ-Hybridisierung, Northern und Southern Blotting sowie PCR-Analyse zur Bestimmung dienen, ob für HRG kodierende DNA und/oder RNA in dem bzw. den untersuchten Zelltypus bzw. Zelltypen vorhanden ist bzw. sind. insbesondere kann die Nucleinsäure als spezifische Sonde für bestimmte Arten von Tumorzellen nützlich sein, z. B. für Brustdrüsen, Magenund Darm-Adenokarzinome, Speicheldrüsen und andere Gewebe, die $p185^{HER2}$ enthalten.

[0214] Isoliertes HRG kann in quantitativen diagnostischen Tests als Standard oder Kontrolle verwendet werden, womit Proben verglichen werden, die unbekannte Mengen an HRG enthalten.

[0215] Isoliertes HRG kann als Wachstumsfaktor für In-vitro-Zelkultur und in vivo verwendet werden, um das Wachstum von Zellen zu unterstützen, die $p185^{HER2}$ oder andere analoge Rezeptoren enthalten.

[0216] HRG-Antikörper eignen sich für diagnostische Tests betreffend die HRG-Expression in spezifischen Zellen oder Geweben. Die Antikörper sind in gleicher Weise wie oben beschriebene HRG markiert und/oder auf einer unlöslichen Matrix immobilisiert.

[0217] HRG-Antikörper eignen sich auch zur Affinitätsreinigung von HRG aus rekombinanter Zellkultur oder

natürlichen Quellen. HRG-Antikörper, die mit anderen HRG nicht detektierbar kreuzreagieren, können dazu dienen, HRG von anderen bekannten Liganden oder kontaminierendem Protein zu reinigen.

[0218] Geeignete Diagnosetests für HRG und seine Antikörper sind auf dem Gebiet an sich allgemein bekannt. Solche Assays sind kompetitive und Sandwich-Assays sowie sterische Hinderungsassays. Kompetitive und Sandwich-Methoden sehen einen Phasentrennschritt als integralen Bestandteil des Verfahrens vor, während Tests mit sterischer Hinderung in einem einzelnen Reaktionsgemisch stattfinden. Es werden im Grunde genommen die gleichen Vorgangsweisen wie für den HRG-Test und für die HRG bindenden Substanzen gewählt, obwohl bestimmte Verfahren je nach dem Molekulargewicht der untersuchten Substanz gegenüber anderen bevorzugt sind. Daher wird die zu untersuchende Substanz hierin als Analyt bezeichnet – ungeachtet ihres sonstigen Status als Antigen oder Antikörper. Proteine, die sich an den Analyte binden, werden hierin als Bindungspartner bezeichnet – ob sie nun Antikörper, Zelloberflächen-Rezeptoren oder Antigene sind.

[0219] Analytische Verfahren für HRG oder seine Antikörper verwenden alle eines oder mehrere der folgenden Reagenzien: markierten Analyten-Analog, immobilisierten Analyten-Analog, markierten Bindungspartner, immobilisierten Bindungspartner und sterische Konjugate. Die markierten Reagenzien sind auch als „Tracer" bekannt.

[0220] Der verwendete Marker (eignet sich auch zum Markieren von für HRG kodierender Nucleinsäure als Sonde) ist jede beliebige Funktionalität, die die Bindung von Analyt und seinem Bindungspartner nicht beeinträchtigt. Es sind zur Verwendung im Immunassay zahlreiche Marker bekannt, z. B. Gruppen, die direkt detektierbar sind wie etwa fluorochrome, chemolumineszierende und radioaktive Marker, sowie Gruppen wie etwa Enzyme, die zur Detektion umgesetzt oder derivatisiert werden müssen. Beispiele für solche Marker sind die Radioisotope ^{32}P, ^{14}C, ^{125}I, ^{3}H und ^{131}I, Fluorophore wie etwa Seltenerd-Chelate oder Fluorescein und seine Derivate, Rhodamin und seine Derivate, Dansyl, Umbelliferon, Luciferase, z. B. Glühwürmchen-Luciferase und bakterielle Luciferase (US-Patent 4.737.456), Luciferin, 2,3-Dihydrophthalazindione, Meerrettich-Peroxidase (HRP), alkalische Phosphatase, β-Galactosidase, Glucoamylase, Lysozym, Saccharidoxidasen, z. B. Glucose-Oxidase, Galactose-Oxidase und Glucose-6-phosphat-Dehydrogenase, Heterozyklen-Oxidasen wie etwa Uricase und Xanthin-Oxidase, gekoppelt mit einem Enzym, das Wasserstoffperoxid verwendet, um einen Farbstoffvorläufer zu peroxidieren, z. B. HRP, Lactoperoxidase oder Microperoxidase, Biotin/Avidin, Spinmarker, Bakteriophagenmarker, stabile freie Radikale u. dgl.

[0221] Es stehen herkömmliche Verfahren zur Verfügung, um diese Marker kovalent an Proteine oder Polypeptide zu binden. Beispielsweise können Kopplungsmittel wie etwa Dialdehyde, Carbodiimide, Dirnaleimide, Bisimidate, bis-diazotiertes Benzidin u. dgl. dazu dienen, die Antikörper mit den oben beschriebenen fluoreszierenden, chemolumineszierenden und Enzymmarkern zu markieren. Siehe z. B. US-Patente 3.940.475 (Fluorimetrie) und 3.645.090 (Enzyme); Hunter et al., Nature 144, 945 (1962); David et al., Biochemistry 13, 1014–1021 (1974); Pain et al., J. Immunol. Methods 40, 219–230 (1981); und Nygren, J. Histochem. and Cytochem. 30, 407–412 (1982). Bevorzugte Marken sind hierin Enzyme wie z. B. Meerrettich-Peroxidase und alkalische Phosphatase. Die Bindung eines solchen Markers (z. B. der Enzyme) an den Antikörper ist ein herkömmliches Manipulationsverfahren für Fachleute auf dem Gebiet von Immuntests. Siehe z. B. O'Sullivan et al., „Methods for the Preparation of Enzyme-antibody Conjugates for Use in Enzyme Immunoassay", in Methods in Enzymology, J. J. Langone und H. Van Vunakis Hg., Bd. 73 (Academic Press, New York, NY, 1981), S. 147–166. Solche Bindungsverfahren eignen sich zur Verwendung mit HRG oder seinen Antikörpern, von denen alle proteinhältig sind.

[0222] Andere Tests, die als kompetitive oder Sandwich-Assays bekannt sind, wurden hinlänglich beschrieben und sind im Bereich der kommerziellen Diagnostik allgemein vorhanden.

[0223] Kompetitive Assays beruhen auf der Fähigkeit eines Tracer-Analogs, mit dem Testprobenanalyten um eine begrenzte Anzahl an Bindungsstellen auf einem gemeinsamen Bindungspartner zu konkurrieren. Der Bindungspartner wird im Allgemeinen vor oder nach der Kometition insolubilisiert; der Tracer und der an den Bindungspartner gebundene Analyt werden dann vom ungebundenen Tracer und Analyten getrennt. Diese Trennung erfolgt durch Dekantieren (sofern der Bindungspartner vorinsolubilisiert wurde) oder durch Zentrifugieren (sofern der Bindungspartner nach der kompetitiven Reaktion gefällt wurde). Die Menge des Testprobenanalyten ist umgekehrt proportional zur Menge an gebundenem Tracer (gemessen anhand der Menge an Markersubstanz). Dosis-Reaktions-Kurven mit bekannten Mengen an Analyt werden erstellt und mit den Testergebnissen verglichen, um die Menge an in der Testprobe vorhandenem Analyten quantitativ zu bestimmen. Diese Assays werden als ELISA-Systeme bezeichnet, wenn Enzyme als detektierbare Marker verwendet werden.

[0224] Eine weitere Kategorie von kompetitivem Assay sind „homogene" Tests, die keine Phasentrennung erfordern. Hier wird ein Konjugat eines Enzyms mit dem Analyten gebildet und solcherart verwendet, dass bei Bindung des Anti-Analyten am Analyten die Gegenwart des Anti-Analyten die Enzymaktivität modifiziert. In diesem Fall sind HRG oder seine immunologisch aktiven Fragmente mit einer bifunktionalen organischen Brücke zu einem Enzym wie z. B. Peroxidase verbunden. Konjugate werden zur Verwendung mit Anti-HRG solcherart ausgewählt, dass die Bindung des Anti-HRG-Antikörpers die Enzymaktivität des Markers inhibiert oder potenziert. Dieses Verfahren wird sehr häufig unter der Bezeichnung EMIT durchgeführt.

[0225] Sterische Konjugate werden in Verfahren auf Basis sterischer Hinderung für homogene Tests verwendet. Diese Konjugate werden durch kovalente Bindung eines niedermolekularen Haptens an einen kleinen Analyten synthetisiert, sodass Antikörper gegen Hapten im Wesentlichen nicht in der Lage ist, das Konjugat gleichzeitig wie Anti-Analyt zu binden. Unter den Testbedingungen bindet der in der Testprobe vorhandene Analyt Anti-Analyt, wodurch Anti-Hapten das Konjugat binden kann, was zu einer Veränderung der Beschaffenheit des Konjugathaptens führt, z. B. zu einer Veränderung der Fluoreszenz, wenn das Hapten ein Fluorophor ist.

[0226] Sandwich-Assays eignen sich besonders zur Bestimmung von HRG oder HRG-Antikörpern. In sequenziellen Sandwich-Assays wird ein immobilisierter Bindungspartner dazu verwendet, Testprobenanalyt zu adsorbieren, die Testprobe wird z. B. durch Waschen entfernt, der gebundene Analyt dient zum Adsorbieren von markiertem Bindungspartner, und gebundenes Material wird dann vom restlichen Tracer getrennt. Die Menge an gebundenem Tracer ist direkt proportional zum Testprobenanalyten. In „simultanen" Sandwich-Assays wird die Testprobe vor der Zugabe des markierten Bindungspartners nicht getrennt. Ein sequenzieller Sandwich-Assay unter Einsatz eines monoklonalen Anti-HRG-Antikörpers als ein Antikörper und eines polyklonalen Anti-HRG-Antikörpers als anderer Antikörper eignet sich zur Untersuchung von Proben auf HRG-Aktivität.

[0227] Die obigen Ausführungen betreffen Beispiele für diagnostische Untersuchungen hinsichtlich HRG und Antikörpern. Andere Verfahren, die für die Bestimmung dieser Analyten heute oder in Zukunft entwickelt werden, sind im Schutzbereich der Erfindung enthalten, z. B. die oben beschriebenen Bioassays.

[0228] HRG-Polypeptide können für Affinitätsreinigung von Rezeptoren wie z. B. $p185^{HER2}$ und ähnlichen Rezeptoren verwendet werden, die Bindungsaffinität für HRG besitzen, insbesondere HRG-α, HRG-β_1, HRG-β_2 und HRG-β_3. HRG-α, HRG-β_1, HRG-β_2 und HRG-β_3 können dazu dienen, Fusionspolypeptide zu bilden, worin der HRG-Abschnitt für Affinitätsbindung an Nucleinsäuren und Heparin geeignet ist.

[0229] HRG-Polypeptide können als Liganden für das kompetitive Screening potenzieller Agonisten oder Antagoisten für die Bindung an $p185^{HER2}$ verwendet werden. HRG-Varianten kommen als Standards oder Vergleiche in HRG-Tests in Frage, sofern sie durch das verwendete analytische System erkannt werden, z. B. Anti-HRG-Antikörper. Antikörper, der zur Bindung an denaturiertes HRG oder ein Fragment davon in der Lage ist, wird in Tests verwendet, in denen HRG vor dem Assay denaturiert wird; in diesem Test wird das denaturierte HRG oder Fragment als Standard oder Vergleich verwendet. Vorzugsweise werden HRG-α, HRG-β_1, HRG-β_2 und HRG-β_3 detektierbar markiert, und es erfolgt ein Kompetitionstest für gebundenes $p185^{HER2}$ unter Anwendung von herkömmlichen Assayverfahren.

[0230] Die hierin beschriebenen Verfahren und Vorgangsweisen unter Verwendung von HRGα können ebenso auf HRG-β_1, HRG-β_2 und HRG-β_3 sowie andere neuartige HRG-Liganden und ihre Varianten angewendet werden. Die folgenden Beispiele dienen lediglich der Veranschaulichung und sind keinesfalls einschränkend.

BEISPIELE

Beispiel 1

Herstellung von Brustkrebs-Überständen

[0231] HRG-α wurde aus dem Überstand des Human-Brustkarzinom-MDA-MB-231 isoliert. HRG wurde in Zellkultur freigesetzt und von diesem isoliert.

a. Zellkultur

[0232] MDA-MB-231, Human-Brustkarzinomzellen, erhältlich bei der ATCC (ATCC HTB 25), wurden zunächst von 25 cm^2 großen Gewebekulturkolben auf 890 cm^2 große Kunststoff-Rollflaschen (Corning, Corning, NY, USA) durch Reihenpassagieren maßstabsvergrößert und der Seed Train im Maßstab der Rollflaschen gehalten. Um die Zellen zu passagieren und den seed train zu halten, wurden die Kolben und die Rollflaschen zunächst mit Phosphat-gepufferter Salzlösung (PBS) gespült und dann mit Trypsin/EDTA (Sigma, St. Louis, MO, USA) 1–3 Minuten lang bei 37°C inkubiert. Die gelösten Zellen wurden dann mehrmals in frischem Kulturmedium, das Rinderfötenserum (FBS; Gibco, Grand Island, NY, USA) enthielt, pipettiert, um Zellklumpen aufzubrechen und das Trypsin zu inaktivieren. Die Zellen wurden schließlich in einem Verhältnis von 1 : 10 in frisches Medium aufgeteilt, in neue Kolben oder Flaschen gefüllt, bei 37°C inkubiert und fast bis zur Konfluenz gezüchtet. Das Wachstumsmedium, in dem die Zellen gehalten wurden, war eine kombinierte DME/Ham'sF-12-Mediumformulierung, die hinsichtlich der Konzentrationen einiger Aminosäuren, Vitamine, Zucker und Salze modifiziert und mit 5% FBS ergänzt war. Das gleiche Hauptkulturmedium wird für die serumfreie Ligandenproduktion verwendet und mit 0,5% Primatone RL (Sheffield, Norwich, NY, USA) ergänzt.

b. Produktion im großen Maßstab

[0233] MDA-MB-231-Zellwachstum im großen Maßstab wurde durch Verwendung von Percell Biolytica-Mikroträgern (Hyclone Laboratories, Logan, UT; USA) aus gewichteter vernetzter Gelatine erzielt. Die Mikroträger wurden zunächst hydriert, autoklaviert und gespült (gemäß den Empfehlungen des Herstellers). Zellen aus 10 Rollflaschen wurden mit Trypsin behandelt und einem Impfungs-Zentrifugengefäß zugegeben, das 3 l Wachstumsmedium und 10–20 g hydrierte Mikroträger enthielt. Die Zellen wurden etwa 1 Stunde lang vorsichtig geführt und in einen 10 l-Fermentor mit Instrumenten übertragen, der 7 l Wachstumsmedium enthielt. Die Kultur wurde bei 65–75 U/min geschüttelt, um die Mikroträger in Suspension zu halten. Der Fermentor wurde bei 37°C und der pH-Wert bei 7,0–7,2 gehalten (dies erfolgte durch Zugabe von Natriumcarbonat und CO). Man ließ Luft und Sauerstoffgase durchperlen, um die Kultur bei etwa 40% Luftsättigung zu halten. Die Zellpopulation wurde mikroskopisch mit einem fluoreszierenden Vitalfarbstoff (Fluoresceindiacetat) überwacht und mit Trypanblau-Färbung verglichen, um die relative Lebensfähigkeit der Zellen und den Grad an Mikroträger-Invasion durch die Zellen zu beurteilen. Die Veränderung der Zell-Mikroträger-Aggregatgröße wurden durch mikroskopische Fotografie dokumentiert.
[0234] Sobald die Mikroträger zu 90–100% konfluent waren, wurde die Kultur mit serumfreiem Medium gewaschen, um das Serum zu entfernen. Dies erfolgte durch Abbrechen des Schüttelns und anderer Steuerungsmöglichkeiten, sodass sich die Träger am Gefäßboden setzen konnten. Etwa 9 l des Kulturüberstands wurden aus dem Gefäß gepumpt und mit einem gleichen Volumen serumfreiem Medium ersetzt (dem gleichen Hauptkulturmedium wie oben, ergänzt mit Primatone RL oder nicht). Die Mikroträger wurden kurz resuspendiert und der Vorgang wiederholt, bis eine 1000fache Entfernung von FBS erzielt wurde. Die Zellen wurden anschließend 3–5 Tage lang in serumfreiem Medium inkubiert. Die Glucose-Konzentration in der Kultur wurde täglich überwacht und gegebenenfalls mit Zugaben von Glucose ergänzt, um die Konzentration im Fermentor bei oder über 1 g/l zu halten. Zum Zeitpunkt der Ernte waren die Mikroträger gesetzt (siehe oben), und der Überstand wurde aseptisch entfernt und zwecks Reinigung bei 2–8°C gelagert. Frisches serumfreies Medium wurde in den Fermentor gefüllt, die Mikroträger wurden resuspendiert, und die Kultur wurde wie oben inkubiert und geerntet. Diese Vorgangsweise könnte vier Mal wiederholt werden.

Beispiel 2

Reinigung von Wachstumsfaktor-Aktivität

[0235] Konditioniertes Medium (10–20 l) aus MDA-MB-231-Zellen wurden durch Zentrifugieren bei 10.000 U/min in einer Sorvall-Zentrifuge geklärt, durch einen 0,22 µm-Filter filtriert und dann 10– bis 50fach (etwa 25fach) mit einer Minitan Tangential Flow Unit (Millipor Corp.) mit einer 10 kDa-Cutoff-Polysulfon-Membran bei Raumtemperatur konzentriert. Alternativ dazu wurde das Medium mit einer 2,5 1-Amicon Stirred Cell bei 4 °C mit einer YM3-Membran konzentriert. Nach der Konzentration wurde das Medium nochmals bei 10.000 U/min zentrifugiert und der Überstand in 35–50 ml-Aliquoten bei –80°C gefroren.
[0236] Heparin-Sepharose stammte von Pharmacia (Piscatawy, NJ, USA) und wurde gemäß den Anweisungen des Herstellers erzeugt. 5 ml des Harzes wurden in eine Säule gepackt, ausgiebig gewaschen (100 Säulenvolumina) und mit PBS äquilibriert. Das konzentrierte konditionierte Medium wurde aufgetaut, durch einen 0,22 µm-Filter filtriert, um Teilchenmaterial zu entfernen, und mit einer Flussrate von 1 ml/min auf die Heparin-Sepharose-Säule geladen. Die normale Ladung bestand aus 30–50 ml 40fach konzentriertem Medium. Nach dem Beladen wurde die Säule mit PBS gewaschen, bis die Extinktion bei 280 nm zur Grundlage zurückkehrte (vor Beginn der Elution des Proteins). Die Säule wurde bei 1 ml/min mit aufeinander folgenden Salzschritten von 0,3 M, 0,6 M, 0,9 M und (gegebenenfalls) 2,0 M NaCl (gebildet in PBS) eluiert. Jeder Schritt wurde fortgesetzt, bis die Extinktion zur Grundlinie zurückkehrte (üblicherwise 6–10 Säulenvolumina). Es wurden Fraktionen mit einem Volumen von 1 ml gesammelt. Alle jedem Waschoder Salzschritt entsprechenden Fraktionen wurden gesammelt und für den nachfolgenden Test im MDA-MB-453-Zellassay gelagert.
[0237] Ein Großteil der Tyrosinphosphorylierungs-Stimulierungsaktivität wurde im 0,6 M NaCl-Pool festgestellt, der für den nächsten Reinigungsschritt verwendet wurde. Aktive Fraktionen aus der Heparin-Sepharose-Chromatographie wurden getaut, dreifach mit entionisiertem Wasser (MilliQ) verdünnt, um die Salzkonzentration zu reduzieren, und auf eine Polyasparaginsäure-Säule geladen (PolyCAT A, 4,6 × 100 mm, PolyLC, Columbia, MD, SA), die in 17 mM Na-Phosphat, pH 6,8, äquilibriert war. Alle Puffer aus diesem Reinigungsschritt enthielten 30% Ethanol, um die Auflösung von Protein auf dieser Säule zu verbessern. Nach dem Laden wurde die Säule mit Äquilibrierungspuffer gewaschen und mit einem linearen Salzgradienten von 0,3 M bis 0,6 M NaCl in 17 mM Na-Phosphatpuffer, pH 6,8, eluiert. Die Säule wurde mit 1 ml/min beladen und entwickelt, und 1 ml-Fraktionen wurden mittels Gradientenelution gesammelt. Fraktionen wurden bei 4°C gelagert. Mehrere Heparin-Sepharose- und PolyCat-Säulen wurden verarbeitet, um ausreichend Material für den nächsten Reinigungsschritt zu erhalten. Ein typisches Extinktionsprofil aus der PolyCat A-Säule ist in **Fig.** 1 zu sehen. Ali-

quoten von 10–25 µl wurden jeder Fraktion zwecks Untersuchung und SDS-Gelanalyse entnommen.

[0238] Tyrosinphosphorylierungs-Stimulierungsaktivität wurde in allen eluierten Fraktionen der PolyCAT A-Säule festgestellt – die meiste Aktivität in den Fraktionen, die Peak C des Chromatogramms entsprachen (Salzkonzentration von etwa 0,45 M NaCl). Diese Fraktionen wurden gesammelt und mittels Zugabe von 0,1 Vol.-% TFA auf 0,1% Trifluoressigsäure (TFA) eingestellt. 2 Volumina entionisiertes Wasser wurden zugesetzt, um das Ethanol und Salz vom vorherigen Schritt zu verdünnen, und die Probe weiterer Reinigung auf HPLC unter Verwendung einer C4-Reverse-phase-Säule (SynChropak RP-4, 4,6 × 100 mm), äquilibriert in einem Puffer aus 0,1% TFA in Wasser mit 15% Acetonitril, unterzogen. Der HPLC-Prozess erfolgte bei Raumtemperatur mit einer Durchflussrate von 1 ml/min. Nach dem Laden der Probe wurde die Säule in 0,1% TFA/15% Acetonitril reäquilibriert. Ein Gradient von Acetonitril wurde solcherart gebildet, dass über einen Zeitraum von 10 Minuten die Acetonitril-Konzentration von 15 auf 25% anstieg (1%/min). Anschließend wurde die Säule mit einem Gradienten von 25 bis 40% Acetonitril während eines Zeitraums von 60 min (0,25%/min) entwickelt. Es wurden 1 ml-Fraktionen abgenommen, zwecks Verhinderung von Verdampfung abgedeckt und bei 4 °C gelagert. Aliquoten von 10–50 µl wurden entnommen, unter Vakuum bis zur Trockene eingeengt (SpeedVac) und mit Assaypuffer (PBS mit 0,1% Rinderserumalbumin) für den Tyrosinphosphorylierungs-Test rekonstituiert. Außerdem wurden Aliquoten von 10-50 µl entnommen und wie oben für die Analyse durch SDS-Gelelektrophorese getrocknet. Ein typisches HPLC-Profil ist aus **Fig.** 2 ersichtlich.

[0239] Ein Haupt-Aktivitätspeak wurde in Fraktion 17 (**Fig.** 2B) festgestellt. Durch SDS-Gelanalyse stellte man fest, dass Fraktion 17 eine einzige Hauptproteinspezies enthielt, die mit dem Standard eines Molekulargewichts von 45.000 co-migrierte (**Fig.** 2C, 3). In anderen Präparaten co-migrierte die Gegenwart des 45 kD-Proteins mit der Stimulierung von Tyrosin-Phosphorylierungsaktivität im MDA-MB-453-Zellassay. Die chromatographischen Eigenschaften des 45 kD-Proteins waren atypisch; im Gegensatz zu vielen anderen Proteinen im Präparat eluierte das 45 kD-Protein aus der Reversed-Phase-Säule nicht innerhalb von 2 oder 3 Fraktionen. Stattdessen eluierte es über 5–10 Fraktionen. Dies ist möglicherweise auf extensive posttranslationale Modifikationen zurückzuführen.

a. Proteinsequenz-Bestimmung

[0240] Fraktionen mit dem 45 kD-Protein wurden unter Vakuum für das Aminosäuren-Sequenzieren getrocknet. Proben wurden in 70% Ameisensäure erneut gelöst und für die Nterminale Sequenzierung in einen Dampfphasen-Sequenzierer von Applied Biosystems, Inc., Modell 470 A geladen. Man erhielt keine erkennbare N-terminate Sequenz, und dies legt den Schluss nahe, dass der N-terminale Rest blockiert war. Ähnliche Ergebnisse wurden erzielt, als das Protein zunächst auf einem SDS-Gel laufen gelassen und auf ProBlott-Membran transgeblottet und die 45 kD-Bande nach der Lokalisierung durch rasche Einfärbung mit Coomassie-Brilliantblau exzidiert wurde.

[0241] Die interne Aminosäuresequenz wurde erhalten, indem die das 45 kD-Protein enthaltenden Fraktionen partiellem Verdau unter Einsatz von Bromcyan (zwecks Spalten an Methioninresten), Lysin-C (zwecks Spalten an der C-terminalen Seite von Lysinresten) oder Asp-N (zwecks Spalten an der N-terminalen Seite von Asparatinsäureresten) unterzogen wurden. Proben wurden nach dem Verdau direkt sequenziert oder die Peptide zunächst durch HPLC-Chromatographie auf einer Synchrom C4-Säule (4000 A, 2 × 100 mm), äquilibriert in 0,1% TFA und eluiert mit einem 1-Propanol-Gradienten in 0,1 TFA, gelöst. Peaks aus dem chromatographen Lauf wurden vor dem Sequenzieren unter Vakuum getrocknet.

[0242] Nach dem Sequenzieren des Peptids im Peak Nr. 15 (Lysin C-15) wurden mehrere Aminosäuren auf jedem Zyklus des Laufs identifiziert. Nach einer sorgfältigen Analyse war klar, dass die Fraktion das gleiche Grundpeptid mit verschiedenen N-Termini enthielt, was zu der Vielzahl an Aminosäuren in jedem Zyklus führt. Nach der Dekonvolution wurde die folgende Sequenz bestimmt (SEQ.-ID Nr. 3):

[A] A E K E K T F [C] V N G G E X F M V K D L X N P

1 5 10 15 20

[0243] (Die Reste in Klammern waren unsicher, während ein X einen Zyklus darstellt, in dem es nicht möglich war, die Aminosäure zu identifizieren.) Die Anfangsausbeute betrug 8,5 pMol. Diese Sequenz mit 24 Aminosäuren stimmte mit keinem zuvor bekannten Protein überein. Rest 1 wurde später anhand der cDNA-Sequenz als Cys und Rest 9 als korrekt identifiziert. Die unbekannten Aminosäuren an den Positionen 15 und 22 stellten sich als Cys bzw. Cys heraus.

[0244] Das Sequenzieren auf den Proben nach Bromcyan- und Asp-N-Verdau, jedoch ohne Trennung durch HPLC, erfolgte, um die cDNA-Sequenz zu bestätigen. Die erhaltenen Sequenzen sind aus Tabelle I ersichtlich, und sie bestätigen die Sequenz für das 45 kD-Protein (aus der cDNA-Sequenz abgeleitet). Der N-Terminal des Proteins scheint mit einer unbekannten Blockiergruppe blockiert zu sein. Einmal zeigte direktes Sequenzieren der 45 kD-Bande aus einem PVDF-Blot die folgende Sequenz mit sehr geringer anfänglicher Ausbeute (0,2

pMol; SEQ.-ID Nr. 4):

X E X K E (G) (R) G K (G) K (G) K K K E X G X G (K)

[0245] (Unbestimmbare Reste sind durch „X" gekennzeichnet, während vorläufige Reste in Klammern angeführt sind.) Diese Sequenz stimmt mit einer Sequenz überein, die am Serin an Position 46 in der Nähe des vorliegenden N-Terminus der HRG-cDNA-Sequenz beginnt; dies legt nahe, dass der N-Terminus des 45 kD-Proteins an diesem oder vor diesem Punkt in der Sequenz angeordnet ist.

Beispiel 3

Klonieren und Sequenzieren von Human-Heregulin

[0246] Die cDNA-Klonierung des $p185^{HER2}$-Liganden fand wie folgt statt. Ein Abschnitt der Lysin C-15-Peptid-Aminosäuresequenz wurde dekodiert, um eine Sonde für cDNA zu konstruieren, die für den 45 kD-HRG-α-Liganden kodieren. Das folgende aus 39 Resten bestehende achtfach degenerierte Desoxyoligonucleotid, das der Aminosäuresequenz (SEQ.-ID Nr. 5) NH-2-...AEKEKTFXVNGGE entsprach, wurde chemisch synthetisiert (SEQ.-ID Nr. 6):

3' GCTGAGAAGGAGAAGACCTTCTGT/CGTGAAT/CGGA/CGGCGAG 5'.

[0247] Der unbekannte Aminosäurerest (gekennzeichnet durch X in der Aminosäuresequenz) erhielt ein Cystein zugewiesen, um die Sonde zu konstruieren. Diese wurde radioaktiv phosphoryliert und dazu verwendet, die unter wenig strengen Bedingungen erfolgende Hybridisierung einer Oligo dT-geprimten cDNA-Bibliothek, konstruiert aus der Human-MDA-MB-231-Zellen-mRNA in ?gt10, zu screenen (Huyung et al., 1984, in DNA Cloning, Bd. 1: A Practical Approach, D. Glover, Hg., S. 49–78 (IRL Press Oxford). DNA-Sequenzanalyse zeigte, dass diese zwei Klone identisch waren.

[0248] Die 2010 bp-cDNA-Nucleotidsequenz von λgt10her16 (**Fig.** 4) enthält einen einzigen langen offenen Leserahmen von 669 Aminosäuren, beginnend mit Alanin an Nucleotidpositionen 3–5 und endend mit Glutamin an Nucleotidpositionen 2007–2009. Es wurde kein Stopcodon in der translatierten Sequenz vorgefunden; eine spätere Analyse von HRG-β-Klonen zeigte allerdings, dass das an Nucleotidpositionen 135–137 kodierte Methionin das Anfangs-Methionin war. Nucleotid-Sequenzhomologie mit der Sonde existiert zwischen und einschließlich der Basen 681–719. Homologie zwischen jenen Aminosäuren, für die die Sonde kodiert, und den die Sonde flankierenden mit der Aminosäuresequenz, die für das Lysin C-15-Fragment bestimmt ist, bestätigt, dass der isolierte Klon zumindest für das Lysin C-15-Fragment des 45 kD-Proteins kodiert.

[0249] Hydropathie-Analyse zeigt die Gegenwart einer stark hydrophoben Aminosäureregion einschließlich der Reste 287–309 (**Fig.** 4) – ein Indikator dafür, dass dieses Protein eine Transmembran- oder interne Signalsequenzdomäne enthält und somit an der Zellmembran verankert ist.

[0250] Die aus 669 Aminosäuren bestehende Sequenz, für die die 2010 bp-cDNA-Sequenz kodiert, enthält potenzielle Stellen für Asparagin-gebundene Glykosylierung (Winzler, R. in Hormonal Proteins and Peptids, Li, C. H., Hg., S. 1–15, Academic Press, New York [1973]) an Positionen Asparagin 164, 170, 208, 437 und 609. Eine potenzielel O-Glykosylierungsstelle (Marshall, R. D. Biochem. Soc. Symp. 40, 17–26 [1974]) befindet sich in der Region, die ein Cluster von Serin- und Threoninresten an Aminosäurepositionen 209–218 enthält. 3 Stellen potenzieller Glykosaminoglykan-Addition (Goldstein, L. A. et al. Cell 56, 1063–1072 [1989]) sind an den Serin-Glycin-Dipeptiden an den Aminosäuren 42–43, 64–65 und 151–152 positioniert. Die Glykosylierung ist wahrscheinlich für die Diskrepanz zwischen dem errechneten NW von etwa 26 kD für die (extrazelluläre) NTD-GFD-Region von HRG und dem beobachteten NW von etwa 45 kD für gereinigtes HRG verantwortlich.

[0251] Diese Aminosäuresequenz weist einige gemeinsame Merkmale mit der EGF-Familie transmembran-gebundener Wachstumsfaktoren auf (Carpenter, G. und Cohen, S., Ann. Rev. Biochem, 48, 193–216 [1979]; Massenque, J., J. Biol. Chem. 265, 21393–21396 [1990]): 1) das Vorhandensein einer Vorform jedes Wachstumsfaktors, aus der die reife Form proteolytisch freigesetzt wird (Gray, A. Dull, T. J. und Ullrich, A., Nature 303, 722-725 [1983]; Bell, G. I. et al., Nuc. Acid Res. 14, 8427–8477 [1986]; Derynck, R. et al., Cell 287–297 [1984]); 2) die Konservierung von 6 Cysteinresten, die sich typischerweise über einem Abschnitt von ewta 40 Aminosäuren erstrecken (das EGF-ähnliche strukturelle Motiv) (Savage, R. C. et al., J. Biol. Chem. 248, 7669–7672 [1973]; HRG-α-Cysteine 226, 234, 240, 254, 256 und 265); und 3) das Vorhandensein einer Transmembran-Domäne, die proximal zur Carboxy-terminalen Seite der EGF-homologen Region auftritt (**Fig.** 4 und 6).

[0252] Die Ausrichtung der Aminosäuresequenzen in der Region des EGF-Motivs und flankierender Transmembran-Domänen mehrerer Human-EGF-ähnlicher Proteine (**Fig.** 6) zeigt, dass zwischen dem 1. und 6. Cystein des EGF-Motivs HRG dem Heparin-Bindungs-EGFähnlichen Wachstumsfaktor (HB-EGF) am meisten

ähnelte (50%) (Higashiyama, S. et al., Science 251, 936–939 [1991]). In der gleichen Region ist HRG – 35% homolog mit Amphiregulin (AR) (Plowman, G.D. et al., Mol. Cell. Biol. 10, 1969–1981 [1990]), – 32 homolog mit transformierendem Wachstumsfaktor a (TGF-α) (8), 27% homolog mit EGF (Bell, G. I. et al., Nuc Acid. Res. 14, 8427–8446 [1986]) und 39% homolog mit dem aus Schwanoma abgeleiteten Wachstumsfaktor (Kimura, N. et al., Nature 348, 257-260 [1980]). Disulfidbindungen zwischen Cysteinresten im EGF-Motiv wurden für EGF bestimmt (Savage, R. C. et al., J. Biol. Chem. 248, 7669–7672 [1973]). Diese Disulfide definieren die sekundäre Struktur dieser Region und begrenzten drei Schleifen. Durch Nummerieren der Cysteine, beginnend mit 1 auf dem aminoterminalen Ende, ist Schleife 1 durch Cysteine 1 und 3 begrenzt; Schleife 2 durch Cysteine 2 und 4; und Schleife 3 durch Cysteine 5 und 6. Obwohl die exakte Disulfid-Konfiguration in der Region für die anderen Familienmitglieder nicht bestimmt wurde, zeigt die strikte Beibehaltung der 6 Cysteine sowie mehrerer anderer Reste, d. h. Glycin 238 und 262 und Arginin an Position 264, dass sie auch höchstwahrscheinlich die gleiche Anordnung aufweisen. HRG-α und EGF besitzen beide 13 Aminosäuren in Schleife 1. HB-EGF-Amphregulin (AR) und TGF-α besitzen 12 Aminosäuren in Schleife 1. Jedes Mitglied besitzt 10 Reste in Schleife 2, außer HRG-α, das 13 aufweist. Alle 5 Mitglieder besitzen in der 3. Schleife 8 Reste.

[0253] EGF, AR, HB-EGF und TGF-α werden alle aufgrund ihrer Transmembran-Domäne als Membran-verankerte Proteine neu synthetisiert. Die Proproteine werden anschließend verarbeitet, um reife aktive Moleküle zu liefern. im Fall von TGF-α gibt es Hinweise darauf, dass die Membran-assoziierten Proformen der Moleküle auch biologisch aktiv sind (Brachmann, R. et al., Cell 56, 691–700 [1989]), ein Charakteristikum, das auch für HRG-α gelten könnte. EGF wird als ein transmembraner 1 168-Aminosäuren-gebundener proEGF synthetisiert, der auf dem aminoterminalen Ende zwischen Arginin 970 und Asparagin 971 sowie am Carboxy-terminalen Ende zwischen Arginin 1023 und Histidin 1024 gespalten ist (Carpenter, G. und Cohen, S., Ann Rev. Biochem. 48, 193–216 [1979]), um das aus 53 Aminosäuren bestehende reife EGF-Molekül zu erhalten, das die aus 3 Schleifen und 3 Disulfidbindungen bestehende typische Struktur enthält. Das aus 252 Aminosäuren bestehende proAR wird zwischen Asparaginsäure 100 und Serin 101 sowie zwischen Lysin 184 und Serin 185 gespalten, um eine aus 84 Aminosäuren bestehende Form von reifem AR zu liefern; eine aus 78 Aminosäuren bestehende Form wird durch NN_2-terminate Spaltung zwischen Glutamin 106 und Valin 107 erzeugt (Plowman, G. D. et al., Mol. Cell. Biol. 10, 1969–1981 [1990]). HB-EGF wird mittels Spaltung zwischen Arginin 73 und Valin 74 sowie an einer zweiten Stelle etwa 84 Aminosäuren weg in Carboxy-terminaler Richtung von seinem primären 208 Aminosäuren-Translationsprodukt in seine vorgeschlagene 84 Aminosäuren-Form verarbeitet (Higashiyama, S. et al. und Klagsburn, M, Science 251, 936–939 (1991]). Die 160 Aminosäuren umfassende Proform von TGF-α wird mittels Spaltungen zwischen Alanin 39 und Valin 40 auf einer Seite und stromabwärtige Spaltung zwischen Alanin 89 und Valin 90 in ein reifes 50 Aminosäuren umfassendes Protein umgewandelt (Derynck et al., Cell 38, 287–297 [1984]). Für jedes der oben beschriebenen Moleküle findet COOH-terminate Verarbeitung im Bereich statt, der vom 6. Cystein des EGF-Motivs und dem Beginn der Transmembran-Domäne begrenzt wird.

[0254] Die Reste zwischen dem 1. und dem 6. Cystein von HRG weisen die größte Ähnlichkeit (45%) mit NP-EGF auf. In dieser gleichen Region sind sie zu 35% identisch mit AR, zu 32% identisch mit TGF-α und zu 27% identisch mit EGF Außerhalb des EGF-Motivs besteht wenig Ähnlichkeit zwischen HRG und anderen Mitgliedern der EGF-Familie, EGF, AR HB-EGF und TGF-α sind alle aus Membran-verankerten Proproteinen abgeleitet, die auf beiden Seiten der EGF-Struktureinheit verarbeitet werden, um 50–84 reife Aminosäureproteine (16–19) zu liefern. Wie andere EGF-Familienmitglieder scheinen die HRG aus einer Membran-gebundenen Proform abgeleitet zu sein, doch sie erfordern nur eine einzige Spaltung (C-terminal zum Cystein-Cluster), um reifes Protein zu produzieren.

[0255] HRG kann seine biologische Funktion zur Geltung bringen, indem es sich an Rezeptor bindet und die Transduktion eines Wachstums-modulierenden Signals auslöst. Dies kann als lösliches Molekül oder vielleicht als Membran-verankerte Form erfolgen, wie dies manchmal bei TGF-α der Fall ist (Brachmann, R. et al., Cell 56, 691–700 [1989]). Im Gegenzug oder zusätzlich zur Stimulierung von Signaltransduktion kann HRG durch eine Zielzelle internalisiert werden, d. h. es kann mit den regulierenden Regionen anderer Regulator-Gene in Wechselwirkung treten und somit direkt die Botschaft zum Zellkern übermitteln. Die Möglicheit, dass HRG einige seiner Wirkungen durch einen Mechanismus wie diesen mediiert, wird durch die Tatsache nahe gelegt, dass ein potenzielles Kernpositionssignal (Roberts, Biochem-Biophys. Acta 1008, 263–280 [1989]) in der Region um die 3 Lysinreste an Positionen 58–60 besteht (**Fig.** 4).

[0256] Die Isolierung von HRG-α-cDNA voller Länge erfolgt unter Verwendung der DNA-Sequenz aus **Fig.** 4, um zusätzliche cDNA-Sequenzen aus der cDNA-Bibliothek auszuwählen, die aus Human-Mda-MB-231 besteht. cDNA-Klone voller Länge, die für HRG-α kodieren, werden durch Identifizieren von für HRG-α kodierender cDNA, die sowohl in 3'- als auch in 5'-Richtung länger sind, und durch anschließendes Spleißen eines Verbunds der unterschiedlichen cDNA erhalten. Zusätzliche cDNA-Bibliotheken werden dann je nach Bedarf konstruiert. Die 3 Typen möglicher cDNA-Bibliotheken: 1) OligodT-geprimt, worin vor allem Abschnitte von Polyadenosinresten geprimt sind, 2) zufallsgeprimt mittels kurzer synthetischer Desoxyoligonucleotide, die für keine bestimmte mRNA-Region spezifisch sind, und 3) spezifisch geprimt unter Verwendung kurzer synthetischer

Desoxyoligonucleotide, die für eine erwünschte Region der mRNA spezifisch sind. Verfahren zur Isolierung solcher cDNA-Bibliotheken sind oben angeführt.

Beispiel 4

Detektion von HRG-α-mRNA-Expression durch Northern-Analysen

[0257] Northern Blot-Analyse von MDA-MB-231 und SK-BR-3-Zelen-mRNA unter sehr strengen Bedingungen zeigt zumindest 5 Hybridisierungsbanden in MDA-MB-231-mRNA, worin eine 6,4 Kb-Bande vorherrscht; andere schwächere Banden befinden sich an 9,4, 6,9, 2,8 und 1,8 Kb (**Fig.** 5). Keine Hybridisierungsbande liegt in SK-BR-3-mRNA vor (p185^{HER2} wird von dieser Zelllinie überexprimiert. Das Vorhandensein dieser Vielzahl an Messages in MDA-MB-231-Zellen zeigt entweder alternatives Spleißen des Gens, verschiedene Verarbeitungen des primären Transkripts der Gene oder die Gegenwart eines Transkripts einer weiteren homologen Message auf. Eine dieser Messages kann für eine lösliche, nicht-transmembran gebundene Form von HRG-α kodieren. Solche Messages (**Fig.** 5) können zur Produktion von cDNA herangezogen werden, die für lösliche nicht-transmembran-gebundene Formen von HRG-α kodieren.

Beispiel 5

Zellwachstums-Stimulierung durch HRG-α

[0258] Mehrere unterschiedliche Brustkrebs-Zelllinien, die den EGF-Rezeptor oder den p185^{HER2}-Rezeptor exprimieren, wurden auf ihre Sensitivität gegenüber Wachstumshemmung oder -Stimulierung durch Ligandenpräparate untersucht. Die getesteten Zelllinien waren SK-BR-3 (ATCC NTB 30), eine Zelllinie, die p185^{HER2} überexprimiert; MDA-MB-468 (ATCC NTB 132), eine Linie, die den EGF-Rezeptor übeexprimiert; und MCF-7-Zellen (ATCC NTB 22), die einen moderaten Wert an p185^{HER2}-Expression aufweisen. Diese Zellen wurden gemäß bekannter Zellkulturtechniken in Kultur gehalten und passagiert. Die Zellen wurden in einem 1 : 1-Gemisch von DMEM und F-12-Medium mit 10 Rinderfötenserum gezüchtet. Für diesen Test wurden die Stammkulturen mit Trypsin behandelt, um die Zellen von der Kulturschale zu lösen, und in einer Menge von etwa 20.000 Zellen/Napf auf eine 96-Napf-Mikrotiterplatte plattiert. Während des Wachstumstests wurden sie mit 1% Rinderfötenserum in Medium gehalten. Die Testproben wurden durch Filtration durch 0,22 µm-Filter sterilisiert und jeweils 4 Näpfen zugegeben und die Zellen bei 37°C 3–5 Tage inkubiert. Am Ende der Wachstumsperiode wurde das Medium aus jedem Napf abgesaugt, und die Zellen wurden mit Kristallviolett behandelt (Lewis, G. et al., Cancer Research 347, 5382–5385 [1987]). Die Menge an Kristallviolett-Extinktion, die proportional zur Anzahl an Zellen in jedem Napf ist, wurde auf einem Durchfluss-Plattenlesegerät gemessen. Die Werte von mehreren Näpfen für jede Testprobe wurden gemittelt. Unbehandelte Näpfe dienten auf jeder Schale als Vergleiche. Die Ergebnisse sind als prozentuelles Wachstum im Verhältnis zu den Vergleichszellen ausgedrückt.

[0259] Der gerinigte HRG-α-Ligand wurde auf Aktivität im Zellwachstumstest untersucht; Die Ergebnisse sind in **Fig.** 7 veranschaulicht. Bei einer Konzentration von etwa 1 nM Ligand zeigten beide Zelllinien, die den p185^{HER2}-Rezeptor exprimieren (SK-BR-2 und MCF-7) Wachstumsstimulierung im Verhältnis zu den Vergleichen, während der Zelltyp (MDA-MB-468), der nur den EGF-Rezeptor exprimiert, keine deutliche Reaktion aufwies. Diese Ergebnisse stimmten mit jenen überein, die man aus den Autophosphorylierungs-Experimenten mit den diversen Zelllinien erzielte. Die Ergebnisse belegen, dass HRG-α-Ligand für den p185^{HER2}-Rezeptor spezifisch ist und keine deutliche Wechselwirkung mit dem EGF-Rezeptor bei diesen Konzentrationen aufweist.

[0260] HRG konkurriert nicht mit Antikörpern gegen die extrazelluläre Domäne von p185^{HER2} doch die monoklonalen Antikörper Anti-p185^{HER2} 2C4 und 7F3 (die selbst antiproliferativ sind) antagonisieren HRG.

Beispiel 6

Klonieren und Sequenzieren von HRG-β_1

[0261] Die Isolierung von HRG-β_1-cDNA erfolgte unter Einsatz eines Hybridisierungsfragments der für HRG-α kodierenden DNA-Sequenz, um zusätzliche cDNA-Sequenzen aus der aus MDA-MB-231-Zellen bestehenden cDNA-Bibliothek auszuwählen. Klon λher11.1-dbI (HRG-β_1) wurde in einer λgt$_{10}$ Oligo-dT-geprimten cDNA-Bibliothek aus MDA-MB-231-poly A$^+$-mRNA identifiziert. Radioaktiv markierte synthetische DNA-Sonden, die mit den 5'- und 3'-Enden von λher16 (HRG-α) übereinstimmten, wurden in einer Hybridisierungsreaktion unter sehr strengen Bedingungen verwendet, um den λher11.1 dbI-Klon zu isolieren. Die DNA-Nucleotidsequenz des ?her11.1dbI-Klons ist in **Fig.** 8 (SEQ.-ID Nr. 9) dargestellt. HRG-β-Aminosäuresequenz ist mit HRG-a homolog

– vom aminoterminalen Ende an Position Asp 15 von HRG-α bis zum 3'-Ende von HRG-α (mit Ausnahme der unten angeführten Positionen). Außerdem erstreckt sich für HRG-β_1 kodierende DNA über 189 Basenpaare länger als λher16 in 3'-Richtung und liefert ein Stopcodon nach Val 675. An Nucleotidposition 247 von λher11.1dbI befindet sich ein A ersetzendes G, was zur Substitution von Gln(Q) anstelle von Arg(R)in HRG-β_1 führt, wie dies aus der 2. Zeile von **Fig.** 9 ersichtlich ist (SEQ.-ID Nr. 8 und SEQ 9) führt.

[0262] Im Bereich des EGF-Motivs gibt es eine weitere Differenz zwischen HRG-α und HRGβ_1. Diese Unterschiede sind unten in einer vergrößerten Ansicht der Homologie zwischen HRG-α und HRG-β_1, in der Region des EGF-Motivs oder der GFD veranschaulicht. Die spezifische Sequenz entspricht HRG-α-Aminosäuren 221–286 (siehe **Fig.** 9). Die Sterne zeigen identische Reste im nachstehenden Vergleich an (SEQ.-ID Nr. 10 und SEQ.-ID Nr. 11).

```
HEREGULIN-α    S H L V K C A E K E K T F C V N G G E C
HEREGULIN-β1   * * * * * * * * * * * * * * * * * * *

HEREGULIN-α    F M V K D L S N P S R Y L C K C Q P G F
HEREGULIN-β1   * * * * * * * * * * * * * * * P N E *

HEREGULIN-α    T G A R C T E N V P M K V Q N Q E K - -
HEREGULIN-β1   * * D * * Q N Y * M A S F Y K H L G I E

HEREGULIN-α    - - - A E E L Y Q K R (-Transmembran)
HEREGULIN-β1   F M E * * * * * * * * (-Transmembran)
```

Beispiel 7

Expression von HRG in E. coli

[0263] HRG-α und HRG-β_1 wurden in E. coli unter Verwendung der DNA-Sequenzen von **Fig.** 4 und 8 exprimiert, die für HRG kodieren (unter der Steuerung des alkalischen Phosphatase-Promotors und der STII-Leadersequenz). In der anfänglichen Charakterisierung von HRG-Aktivität waren die natürlichen Amino- und Carboxytermini des HRG-Moleküls nicht präzise definiert. Nach dem Vergleich von HRG mit EGF und TGF-α-Sequenzen jedoch konnten die Anmelder erwarten, dass verkürzte Formen von HRG, die etwa bei Ser 221 beginnen und etwa bei Glu 277 von **Fig.** 4 enden, möglicherweise biologische Aktivität aufweisen. Analoge Regionen aller Hereguline können identifiziert und exprimiert werden. Eine andere verkürzte Form wurde konstruiert, um einen N-terminalen Asp-Rest, gefolgt von den Resten 221–277 von HRG-α, aufzuweisen. Aufgrund einer zufälligen Rahmenverlagerungs-Mutation nach Glu 277 war die HRG-α-Sequenz um 13 Aminosäuren auf dem Carboxy-terminalen Ende verlängert. Somit war das Carboxy-terminale Ende Glu 277 von HRG-α, gefolgt von der 13 Aminosäuren umfassenden Sequenz RPNARLPPGVFYC (SEQ.-ID Nr. 20).

[0264] Die Expression dieses Konstrukts wurde durch Wachstum der Zellen in Phosphat-abgereichertem Medium etwa 20 Stunden lang induziert. Rekombinantes Protein wurde durch Ernten der Zellpaste und Resuspendieren in 10 mM Tris (pH 8), Homogenisieren, Inkubieren bei 4°C über den Zeitraum von 40 Minuten und anschließendes Zentrifugieren bei 15.000 U/min (Sorvall) gereinigt. Der Überstand wurde auf einer 30 K-Ultrafiltrationsmembran (Amicon) konzentriert und das Filtrat auf eine in 10 mM Tris, pH 8, äquilibrierte MonoQ-Säule aufgebracht. Die Durchflussfraktionen aus der MonoQ-Säule wurden auf 0,05% T' FA eingestellt und C4-Reverse-Phase-HPLC unterzogen. Die Elution erfolgte mit einem Gradienten von 10–25 Acenotril in 0,1 TFA/H_2O. Das Lösungsmittel wurde durch Lyophilisierung entfernt und gereinigtes Protein in 0,1% Rinderserumalbumin in PBS resuspendiert. **Fig.** 10 zeigt HER-2-Rezeptor-Autophosphorylierungs-Daten mit MCF-7-Zellen als Reaktion auf das gereinigte E. coli-abgeleitete Protein. Dieses Material zeigt volle biologische Aktivität mit einem EC_{50} von 0,8 nM. Das gereinigte Material wurde auch in Zellwachstumstests (Beispiel 5) untersucht und stellte sich als hochwirksamer Stimulator von Zellwachstum heraus.

[0265] Der rekombinante Expressionsvektor für die Synthese von HRG-β_1 wurde in ähnlicher Weise wie HRG-α konstruiert. Der Expressionsvektor enthielt DNA, die für HRG-β_1-Aminosäuren kodierte (von Ser 207 bis Leu 273; **Fig.** 4), Diese für HRG-β_1, kodierende DNA wurde in den Expressionsvektor stromab vom alkalischen Phosphatase-Promotor und die STII-Leadersequenz rekombinant gespleißt. Ein zusätzlicher Serinrest

wurde infolge des rekombinanten Konstruktionsprozesses auf dem Carboxy-Terminus gespleißt. Der für HRG-β_1 kodierende Expressionsvektor diente zum Transformieren von E. coli und exprimierte in Phosphat-abgereichertem Medium. Induzierte E. coli wurden pelletiert, in 10 mM Tris (pH 7,5) resuspendiert und ultraschallbehandelt. Zellbruchstücke wurden durch Zentrifugieen pelletiert und der Überstand vor dem Test durch einen sterilen Filter filtriert. Die Expression von HRG-β_1 wurde durch die Detektion von Protein mit der Fähigkeit, Autophosphorylierung des HER-2-Rezeptors in MCF-7-Zellen zu stimulieren, bestätigt.

[0266] Es wurde ein ähnlicher Expressionsvektor wie oben für HRG-β_1 beschrieben mit einem C-terminalen Tyrosinrest anstelle des Serinrests konstruiert. Dieser Vektor wurde in E. coli transformiert und wie oben exprimiert. Die Reinigung dieses rekombinanten Proteins erfolgte wie für rekombinantes HRG-α beschrieben. Die massenspektrometrische Analyse zeigte, dass das gereinigte Protein aus Formen bestand, die kürzer als erwartet waren. Aminosäure-Sequenzieren belegte, dass das Protein den erwünschten N-terminalen Rest (Ser) besaß, aber am C-Terminus gekürzt war (wurde durch Massenspektrometrie ermittelt). Die Großteil (> 80%) des Proteins bestand aus einer Form, die 51 Aminosäuren lang war und ein C-terminales Methionin (Met 271) besaß (SEQ.-ID Nr. 9). Es wurde auch ein kleiner Abschnitt einer kürzeren Form (49 Reste) detektiert, der an Val 269 gekürzt war. Beide verkürzten Formen zeigten allerdings volle biologische Aktivität im HER-2-Rezeptor-Autophosphorylierungs-Test.

Beispiel 8

Isolierung von HRG-β_2- und -β_3-Varianten

[0267] Es wurden HRG-β_2- und -β_3-Varianten isoliert, um cDNA-Klone zu erhalten, die sich in 5'-Richtung weiter erstrecken. Eine spezifisch geprimte cDNA-Bibliothek wurde in λgt10 konstruiert, indem der chemisch synthetisierte Antisense-Primer 3' CCTTCCCGTTCT-TCTTCCTCGCTCC (SEQ.-ID Nr. 21) verwendet wurde. Dieser Primer befindet sich zwischen den Nucleotiden 167–190 in der Sequenz von λher16 (**Fig.** 4). Die Isolierung von Klon λ5'her13 (nicht mit λher13 zu verwechseln) erfolgte durch Hybridisieren einer synthetischen DNA-Sonde, die mit dem 5'-Ende von λherl6 übereinstimmte, unter sehr strengen Bedingungen mit der spezifisch geprimten cDNA-Bibliothek. Die Nucleotidsequenz von λ5'her13 ist aus **Fig.** 11 ersichtlich (SEQ.-ID Nr. 22). Die aus 496 Basenpaaren bestehende Nucleotidsequenz von λ5'her13 ist homolog mit der Sequenz von λher-16 zwischen den Nucleotiden 309–496 von λ5'her13 und 3–190 von λher16. λ5'her13 erstreckt sich um 102 Aminosäuren über den offenen Leserahmen von λherl6.

[0268] Die Isolierung von HRG-β-Varianten erfolgte durch Sondieren einer neu produzierten cDNA-Bibliothek (aus oligo dT-geprimter λgt10 MDA-MB-231-mRNA abgeleitet) mit synthetischen Sonden, die mit dem 5'-Ende von λ5'her13 und der Cystein-reichen EGFähnlichen Region von λher16 übereinstimmen. 3 Varianten von HRG-β wurden identifiziert, isoliert und sequenziert. Die Aminosäure-Homologien zwischen allen Heregulinen sind aus **Fig.** 15 ersichtlich (SEQ.-ID NR. 26–30).

[0269] HRG-Polypeptide λher76 (HRG-β_2) (SEQ.-ID Nr. 23), λher78 (HRG-β_3) (SEQ.-ID Nr. 24) und λher84 (HRG-β_2-ähnlich) (SEQ.-ID Nr. 25) gelten als Varianten von λher11.1.dbl (HRG-β_1), da zwar die abgeleitete Aminosäuresequenz zwischen Cystein 1 und Cystein 6 des EGF-ähnlichen Motivs identisch ist, aber ihre Sequenzen vor der vorhergesagten Transmembran-Domäne abweichen, die wahrscheinlich mit Aminosäure 248 in λher-11.1 dbl beginnt. Die Nucleotidsequenzen und abgeleiteten Aminosäuresequenzen von λher76, λher78 und λher84 sind in **Fig.** 12, 13 bzw. 14 zu sehen.

[0270] Die Varianten enthalten jeweils ein TGA-Stopcodon 148 Basen 5' vom ersten Methioninrest in ihren Sequenzen. Daher kann das ATG-Codon an Nucleotidposition 135–137 von λherl6 und das entsprechende ATG in den anderen HRG-Klonen als Anfangs-Methionin (Aminosäure 1) definiert werden. Die Klone λher11.1 dbl, λher76, λher84 und λher78 kodieren alle für Glutamm an Aminosaure 38 (**Fig.** 15), während Klon her16 für Arginin kodiert (**Fig.** 4, Position 82).

[0271] Die abgeleitete Aminosäuresequenz von λher76 (HRG-β_1) weist einen Klon voller Länge auf, der für 637 Aminosäuren kodiert. Es teilt sich ein identische abgeleitete Aminosäuresequenz als λher11.1 dbl, außer dass die Reste, die Aminosäuren 232–239 von λher11.1 dbl entsprechen, deletiert wurden. Die abgeleitete Aminosäure von λher84 besitzt die gleiche Aminosäuresequenz wie λher76 – beginnend mit dem Anfangs-Methionin (Aminosäure 1, **Fig.** 15) durch den EGF-ähnlichen Bereich und die Transmembran-Domäne. λher84 gelangt jedoch an Arginin 421 (λher84-Nummerierung) zu einem frühen Stopcodon. Anschließend zweigt die untranslatierte 3'-Sequenz ab. Die abgeleitete Aminosäuresequenz von λher78 (HRG-$_{\beta 3}$) ist mit HRG-β_1 und HRG-β_2 durch Aminosäure 230 homolog, wo die Sequenz für 11 Aminosäuren abzweigt und dann aufhört. Somit besitzt HRG-β_3 keine Transmembran-Region. Die untranslatierte 3'-Sequenz ist mit den anderen Klonen nicht homolog.

Beispiel 9

Expression von HRG-β-Formen

[0272] Um HRG-β-Formen in Säugetierzellen zu exprimieren, wurden cDNA-Nucleotidsequenzen voller Länge aus λher76 (HRG-β_2) oder λher84 in den Säugetier-Expressionsvektor pRK5.1 subkloniert. Dieser Vektor ist ein Derivat von pRK5, das einen Zytomegalie-Promotor, gefolgt von einem 5'-Intron, einem Klonierungs-Polylinker und einem frühen SV-40-Polyadenylationssignal, enthält. COS7-, Affen- oder Humannieren-293-Zellen wurden transfiziert und konditioniertes Medium im MCF-7-Zellen-p185/her2-Autophosphorylierungs-Assay untersucht. Eine positive Response bestätigte die Expression der cDNA aus λher76 (HRG-β_2) und λher84 (HRG-β_3).

[0273] Überstände aus einem im großem Maßstab durchgetührten transienten Expressionsexperiments wurden auf YM10-Membran (Amicon) konzentriert und auf eine Sepharose-Säule aufgebracht (siehe Beispiel 1). Aktivität (Tyrosin-Phosphorylierungsassay) wurde im 0,6 M NaCl-Elutionspool detektiert und auf einer Polyasparaginsäuren-Säule weiter gereinigt (siehe oben). Durch SDS-Gelanalyse und Aktivitätstests wurde festgestellt, dass die aktiven Fraktionen dieser Säule hochgereinigt waren und eine einzelne Bande von Protein mit einem scheinbaren Molekulargewicht von 45 kD enthielten. Somit besitzt das exprimierte Protein chromatographische und strukturelle Eigenschaften, die jenen der nativen Form von Heregulin (ursprünglich aus den MDA-231-Zellen isoliert) stark ähneln. Im kleinen Maßstab durchgeführte transiente Expressionsexperimete mit Konstrukten aus λher84-cDNA zeigten auch vergleichbare Werte an Aktivität in den Zellüberständen aus dieser Variantenform. Die Expression der Transmembran-minus-Variante, HRG-β_3, wird derzeit untersucht.

Beispiel 10

[0274] proHRG-α- und proHRG-β_1-cDNA wurden in Epstein-Barr-Virus gespleißt, das aus Zytomegalievirus enthaltenden Expressionsvektoren abgeleitet war. rHRG wurden (im Wesentlichen wie in Beispiel 2) von serumfeiem Medium stabil transfizierter CEN4-Zellen (Humannieren-293-Zellen; ATCC Nr. 1573) gereinigt, die den EB-Virus-EBNA-1-Transaktivator exprimieren. In anderen Experimenten lieferten transiente proHRG-α-, -β_1- und -β_2-Expressionskonstrukte voller Länge $p185^{HER2}$-Phosphorylierungsaktivität im konditionierten Medium transfizierter COS7-Affennierenzellen. Ähnliche Konstrukte von pro-HRG-β_3 voller Länge ergaben allerdings keine Aktivität, was den Schluss nahe legt, dass die hydrophobe Domäne, die in proHRG-β_3 fehlt, aber in den anderen proHRG vorhanden ist, für die Sekretion von reifem Protein notwendig ist. Gekürzte Versionen von pro-HRG-α (63 Aminosäuren, Serin 177 bis Tyrosin 239) und proHRG-β_1 (68 Aminosäuren, Serin 177 bis Tyrosin 241), die jeweils für die GFD-Struktureinheit und dazwischen liegende Flankierungsregionen kodieren, wurden auch in E. coli exprimiert. Man kann davon ausgehen, dass homologe gekürzte Versionen von NRG-β_3 als aktive Moleküle exprimiert werden. Diese gekürzten Proteine wurden vom periplasmatischen Raum und Kulturbrühe von E. coli gereinigt, die mit Expressionsvektoren transformiert waren; diese waren ausgebildet, rekombinante Proteine zu sekretieren (C. N. Change, M. Rey, B. Bochener, N. Heyneker, G. Gray, Gene 55, 189 [1987]). Diese Proteine stimulierten auch Tyrosin-Phosphorylierung von $p185^{HER2}$, jedoch nicht $p107^{HER1}$ – ein Indikator dafür, dass die biologische Aktivität von HRG in der EGF-ähnlichen Domäne des Proteins liegt und dass Kohlenhydratgruppen für die Aktivität in diesem Test nicht wesentlich sin. Die NTD hemmt oder unterdrückt diese Aktivität nicht.

Beispiel 11

[0275] Verschiedene Humangewebe wurden auf Gegenwart von HRG-mRNA untersucht. Transkripte wurden in Brust, Eierstöcken, Hoden, Prostata, Herz, Skelettmuskeln, Lunge, Leber, Niere, Speicheldrüsen, Dünndarm und Milz, nicht jedoch in Bauch, Bauchspeicheldrüse, Uterus oder Plazenta festgestellt. Während die meisten dieser Gewebe die gleichen 3 Klassen von Transkripten aufweisen wie die MDA-MB-231-Zellen (6,6 kb, 2,5 kb und 1,8 kb), wurde nur die 6,6 kb-Message für das Herz und die Skelettmuskeln beobachtet. Im Gehirn wird ein einzelnes Transkript von 2,2 kb beobachtet, und in den Hoden tritt das Transkript von 6,6 kb gemeinsam mit den anderen von 2,2 kb, 1,9 kb und 1,5 kb in Erscheinung. Das für HRG identifizierte gewebespezifische Expressionsmuster unterscheidet sich von jenem von $p185^{HER2}$; Beispielsweise enthalten die Leber, die Milz und das Gehirn von Erwachsenen HRG, aber keine $p185^{HER2}$-Transkripte, während Bauch, Bauchspeicheldrüse, Uterus und Plazenta $p185^{HER2}$-Transkripte, aber keine HRG-mRNA aufweisen.

SEQUENZPROTOKOLL

(1) ALLGEMEINE INFORMATIONEN

(i) ANMELDER: Genentech, Inc.

(ii) TITEL DER ERFINDUNG: Struktur, Herstellung und Verwendung von Heregulin

(iii) ANZAHL DER SEQUENZEN: 30

(iv) GESCHÄFTSADRESSE:

(A) ADRESSAT: Genentech, Inc.

(B) STRASSE: 460 Point San Bruno Blvd

(C) STADT: South San Francisco

(D) BUNDESSTAAT: Kalifornien

(E) STAAT: USA

(F) POSTLEITZAHL: 94080

(v) COMPUTERLESBARE FORM:

(A) TYP DES MEDIUMS: 5,25 Zoll, 360 kB Diskette

(B) COMPUTER: IBM PC-kompatibel

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

(D) SOFTWARE: Patin (Genentech)

(vi) AKTUELLE ANMELDUNGSDATEN:

(A) ANMELDUNGSNUMMER:

(B) EINREICHUNGSDATUM: 21. Mai 1992

(C) KLASSIFIZIERUNG:

(vii) PRIORITÄTSANMELDUNGSDATEN:

(A) ANMELDUNGSNUMMER:

(B) EINREICHUNGSDATUM: 11. Mai 1992

(vii) PRIORITÄTSANMELDUNGSDATEN:

(A) ANMELDUNGSNUMMER: 07/847743

(B) EINREICHUNGSDATUM: 06. März 1992

(vii) PRIORITÄTSANMELDUNGSDATEN:

(A) ANMELDUNGSNUMMER: 07/847743

(B) EINREICHUNGSDATUM: 24. Mai 1991

(vii) PRIORITÄTSANMELDUNGSDATEN:

(A) ANMELDUNGSNUMMER: 07/765212

(B) EINREICHUNGSDATUM: 25. September 1991

(vii) PRIORITÄTSANMELDUNGSDATEN:

(A) ANMELDUNGSNUMMER: 07/765212

(B) EINREICHUNGSDATUM: 25. September 1991

(vii) PRIORITÄTSANMELDUNGSDATEN:

(A) ANMELDUNGSNUMMER: 07/790801

(B) EINREICHUNGSDATUM: 08. November 1991

(viii) INFORMATIONEN ÜBER ANWALT/VERTRETER:

(A) NAME: Hensley, Max D.

(B) REGISTRIERUNGSNUMMER: 27,043

(C) KENNZEICHEN/AKTENZAHL: 712P4

(ix) TELEKOMMUNIKATIONSINFORMATIONEN:

(A) TELEFON: 415/266-1994

(B) TELEFAX: 415/952-9881

(C) TELEX: 910/371-7168

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 1:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 6 Basen

(B) TYP: Nucleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

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

CNCAAT 6

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 2:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 6 Basen

(B) TYP: Nucleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 2:

AATAAA 6

(2) INFORMATIONEN UBER SEQ.-ID NR. 3:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 24 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 3:

```
Ala Ala Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Xaa
 1               5                  10                  15

Phe Met Val Lys Asp Leu Xaa Asn Pro
            20                  24
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 4:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 21 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 4:

```
Xaa Glu Xaa Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys Lys Lys
 1               5                  10                  15

Glu Xaa Gly Xaa Gly Lys
            20  21
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 5:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 13 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 5:

```
Ala Glu Lys Glu Lys Thr Phe Xaa Val Asn Gly Gly Glu
  1               5                  10          13
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 6:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 42 Basen

(B) TYP: Nucleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 6:

```
GCTGAGAAGG AGAAGACCTT CTGTCGTGAA TCGGACGGCG AG 42
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 7:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 2199 Basen

(B) TYP: Nucleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 7:

```
GG  GAC AAA CTT TTC CCA AAC CCG ATC CGA GCC CTT GGA 38
    Asp Lys Leu Phe Pro Asn Pro Ile Arg Ala Leu Gly
      1               5                  10

CCA AAC TCG CCT GCG CCG AGA GCC GTC CGC GTA GAG CGC 77
Pro Asn Ser Pro Ala Pro Arg Ala Val Arg Val Glu Arg
         15                  20                  25

TCC GTC TCC GGC GAG ATG TCC GAG CGC AAA GAA GGC AGA 116
Ser Val Ser Gly Glu Met Ser Glu Arg Lys Glu Gly Arg
                 30                  35

GGC AAA GGG AAG GGC AAG AAG AAG GAG CGA GGC TCC GGC 155
Gly Lys Gly Lys Gly Lys Lys Lys Glu Arg Gly Ser Gly
     40                  45                  50

AAG AAG CCG GAG TCC GCG GCG GGC AGC CAG AGC CCA GCC 194
Lys Lys Pro Glu Ser Ala Ala Gly Ser Gln Ser Pro Ala
             55                  60

TTG CCT CCC CAA TTG AAA GAG ATG AAA AGC CAG GAA TCG 233
Leu Pro Pro Gln Leu Lys Glu Met Lys Ser Gln Glu Ser
 65                  70                  75
```

```
GCT GCA GGT TCC AAA CTA GTC CTT CGG TGT GAA ACC AGT 272
Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser
         80                  85                  90

TCT GAA TAC TCC TCT CTC AGA TTC AAG TGG TTC AAG AAT 311
Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn
                 95                 100

GGG AAT GAA TTG AAT CGA AAA AAC AAA CCA CAA AAT ATC 350
Gly Asn Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile
    105                 110                 115

AAG ATA CAA AAA AAG CCA GGG AAG TCA GAA CTT CGC ATT 389
Lys Ile Gln Lys Lys Pro Gly Lys Ser Glu Leu Arg Ile
            120                 125

AAC AAA GCA TCA CTG GCT GAT TCT GGA GAG TAT ATG TGC 428
Asn Lys Ala Ser Leu Ala Asp Ser Gly Glu Tyr Met Cys
130                 135                 140

AAA GTG ATC AGC AAA TTA GGA AAT GAC AGT GCC TCT GCC 467
Lys Val Ile Ser Lys Leu Gly Asn Asp Ser Ala Ser Ala
        145                 150                 155

AAT ATC ACC ATC GTG GAA TCA AAC GAG ATC ATC ACT GGT 506
Asn Ile Thr Ile Val Glu Ser Asn Glu Ile Ile Thr Gly
                160                 165

ATG CCA GCC TCA ACT GAA GGA GCA TAT GTG TCT TCA GAG 545
Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu
    170                 175                 180

TCT CCC ATT AGA ATA TCA GTA TCC ACA GAA GGA GCA AAT 584
Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn
            185                 190

ACT TCT TCA TCT ACA TCT ACA TCC ACC ACT GGG ACA AGC 623
Thr Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser
195                 200                 205

CAT CTT GTA AAA TGT GCG GAG AAG GAG AAA ACT TTC TGT 662
His Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys
        210                 215                 220

GTG AAT GGA GGG GAG TGC TTC ATG GTG AAA GAC CTT TCA 701
Val Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser
                225                 230

AAC CCC TCG AGA TAC TTG TGC AAG TGC CCA AAT GAG TTT 740
Asn Pro Ser Arg Tyr Leu Cys Lys Cys Pro Asn Glu Phe
    235                 240                 245

ACT GGT GAT CGC TGC CAA AAC TAC GTA ATG GCC AGC TTC 779
Thr Gly Asp Arg Cys Gln Asn Tyr Val Met Ala Ser Phe
            250                 255

TAC AAG CAT CTT GGG ATT GAA TTT ATG GAG GCG GAG GAG 818
Tyr Lys His Leu Gly Ile Glu Phe Met Glu Ala Glu Glu
260                 265                 270

      CTG TAC CAG AAG AGA GTG CTG ACC ATA ACC GGC ATC TGC 857
Leu Tyr Gln Lys Arg Val Leu Thr Ile Thr Gly Ile Cys
        275                 280                 285
```

```
ATC GCC CTC CTT GTG GTC GGC ATC ATG TGT GTG GTG GCC 896
Ile Ala Leu Leu Val Val Gly Ile Met Cys Val Val Ala
                290                 295

TAC TGC AAA ACC AAG AAA CAG CGG AAA AAG CTG CAT GAC 935
Tyr Cys Lys Thr Lys Lys Gln Arg Lys Lys Leu His Asp
    300                 305                 310

CGT CTT CGG CAG AGC CTT CGG TCT GAA CGA AAC AAT ATG 974
Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg Asn Asn Met
            315                 320

ATG AAC ATT GCC AAT GGG CCT CAC CAT CCT AAC CCA CCC 1013
Met Asn Ile Ala Asn Gly Pro His His Pro Asn Pro Pro
325                 330                 335

CCC GAG AAT GTC CAG CTG GTG AAT CAA TAC GTA TCT AAA 1052
Pro Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys
        340                 345                 350

AAC GTC ATC TCC AGT GAG CAT ATT GTT GAG AGA GAA GCA 1091
Asn Val Ile Ser Ser Glu His Ile Val Glu Arg Glu Ala
                355                 360

GAG ACA TCC TTT TCC ACC AGT CAC TAT ACT TCC ACA GCC 1130
Glu Thr Ser Phe Ser Thr Ser His Tyr Thr Ser Thr Ala
    365                 370                 375

CAT CAC TCC ACT ACT GTC ACC CAG ACT CCT AGC CAC AGC 1169
His His Ser Thr Thr Val Thr Gln Thr Pro Ser His Ser
            380                 385

TGG AGC AAC GGA CAC ACT GAA AGC ATC CTT TCC GAA AGC 1208
Trp Ser Asn Gly His Thr Glu Ser Ile Leu Ser Glu Ser
390                 395                 400

CAC TCT GTA ATC GTG ATG TCA TCC GTA GAA AAC AGT AGG 1247
His Ser Val Ile Val Met Ser Ser Val Glu Asn Ser Arg
        405                 410                 415

CAC AGC AGC CCA ACT GGG GGC CCA AGA GGA CGT CTT AAT 1286
His Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg Leu Asn
                420                 425

GGC ACA GGA GGC CCT CGT GAA TGT AAC AGC TTC CTC AGG 1325
Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg
    430                 435                 440

CAT GCC AGA GAA ACC CCT GAT TCC TAC CGA GAC TCT CCT 1364
His Ala Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro
            445                 450

CAT AGT GAA AGG TAT GTG TCA GCC ATG ACC ACC CCG GCT 1403
His Ser Glu Arg Tyr Val Ser Ala Met Thr Thr Pro Ala
455                 460                 465

CGT ATG TCA CCT GTA GAT TTC CAC ACG CCA AGC TCC CCC 1442
Arg Met Ser Pro Val Asp Phe His Thr Pro Ser Ser Pro
        470                 475                 480

AAA TCG CCC CCT TCG GAA ATG TCT CCA CCC GTG TCC AGC 1481
Lys Ser Pro Pro Ser Glu Met Ser Pro Pro Val Ser Ser
                485                 490
```

```
ATG ACG GTG TCC ATG CCT TCC ATG GCG GTC AGC CCC TTC 1520
Met Thr Val Ser Met Pro Ser Met Ala Val Ser Pro Phe
    495                 500                 505

ATG GAA GAA GAG AGA CCT CTA CTT CTC GTG ACA CCA CCA 1559
Met Glu Glu Glu Arg Pro Leu Leu Leu Val Thr Pro Pro
            510                 515

AGG CTG CGG GAG AAG AAG TTT GAC CAT CAC CCT CAG CAG 1598
Arg Leu Arg Glu Lys Lys Phe Asp His His Pro Gln Gln
520                 525                 530

TTC AGC TCC TTC CAC CAC AAC CCC GCG CAT GAC AGT AAC 1637
Phe Ser Ser Phe His His Asn Pro Ala His Asp Ser Asn
        535                 540                 545

AGC CTC CCT GCT AGC CCC TTG AGG ATA GTG GAG GAT GAG 1676
Ser Leu Pro Ala Ser Pro Leu Arg Ile Val Glu Asp Glu
                550                 555

GAG TAT GAA ACG ACC CAA GAG TAC GAG CCA GCC CAA GAG 1715
Glu Tyr Glu Thr Thr Gln Glu Tyr Glu Pro Ala Gln Glu
    560                 565                 570

CCT GTT AAG AAA CTC GCC AAT AGC CGG CGG GCC AAA AGA 1754
Pro Val Lys Lys Leu Ala Asn Ser Arg Arg Ala Lys Arg
          575                 580

ACC AAG CCC AAT GGC CAC ATT GCT AAC AGA TTG GAA GTG 1793
Thr Lys Pro Asn Gly His Ile Ala Asn Arg Leu Glu Val
585                 590                 595

GAC AGC AAC ACA AGC TCC CAG AGC AGT AAC TCA GAG AGT 1832
Asp Ser Asn Thr Ser Ser Gln Ser Ser Asn Ser Glu Ser
        600                 605                 610

GAA ACA GAA GAT GAA AGA GTA GGT GAA GAT ACG CCT TTC 1871
Glu Thr Glu Asp Glu Arg Val Gly Glu Asp Thr Pro Phe
                615                 620

CTG GGC ATA CAG AAC CCC CTG GCA GCC AGT CTT GAG GCA 1910
Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser Leu Glu Ala
    625                 630                 635

ACA CCT GCC TTC CGC CTG GCT GAC AGC AGG ACT AAC CCA 1949
Thr Pro Ala Phe Arg Leu Ala Asp Ser Arg Thr Asn Pro
            640                 645

GCA GGC CGC TTC TCG ACA CAG GAA GAA ATC CAG GCC AGG 1988
Ala Gly Arg Phe Ser Thr Gln Glu Glu Ile Gln Ala Arg
650                 655                 660

CTG TCT AGT GTA ATT GCT AAC CAA GAC CCT ATT GCT GTA TA 2029
Leu Ser Ser Val Ile Ala Asn Gln Asp Pro Ile Ala Val
        665                 670                 675

A AACCTAAATA AACACATAGA TTCACCTGTA AAACTTTATT 2070

TTATATAATA AAGTATTCCA CCTTAAATTA AACAATTTAT TTTATTTTAG 2120

CAGTTCTGCA AATAGAAAAC AGGAAAAAAA CTTTTATAAA TTAAATATAT 2170

GTATGTAAAA ATGAAAAAAA AAAAAAAAA 2199
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 8:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 669 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 8:

```
Ala Arg Ala Pro Gln Arg Gly Arg Ser Leu Ser Pro Ser Arg Asp
 1               5                  10                  15

Lys Leu Phe Pro Asn Pro Ile Arg Ala Leu Gly Pro Asn Ser Pro
            20                  25                  30

Ala Pro Arg Ala Val Arg Val Glu Arg Ser Val Ser Gly Glu Met
        35                  40                  45

Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys Lys Lys
    50                  55                  60

Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala Ala Gly Ser Gln
65                  70                  75

Ser Pro Ala Leu Pro Pro Arg Leu Lys Glu Met Lys Ser Gln Glu
80                  85                  90

Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser Ser
            95                  100                 105

Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn Glu
        110                 115                 120

Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys Lys
    125                 130                 135

Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala Asp
140                 145                 150

Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu Gly Asn Asp
            155                 160                 165

Ser Ala Ser Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile Ile
        170                 175                 180

Thr Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu
    185                 190                 195

Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr Ser
200                 205                 210

Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys
            215                 220                 225
```

Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys
230 235 240

Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu Cys Lys
245 250 255

Cys Gln Pro Gly Phe Thr Gly Ala Arg Cys Thr Glu Asn Val Pro
260 265 270

Met Lys Val Gln Asn Gln Glu Lys Ala Glu Glu Leu Tyr Gln Lys
275 280 285

Arg Val Leu Thr Ile Thr Gly Ile Cys Ile Ala Leu Leu Val Val
290 295 300

Gly Ile Met Cys Val Val Ala Tyr Cys Lys Thr Lys Lys Gln Arg
305 310 315

Lys Lys Leu His Asp Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg
320 325 330

Asn Asn Met Met Asn Ile Ala Asn Gly Pro His His Pro Asn Pro
335 340 345

Pro Pro Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys Asn
350 355 360

Val Ile Ser Ser Glu His Ile Val Glu Arg Glu Ala Glu Thr Ser
365 370 375

Phe Ser Thr Ser His Tyr Thr Ser Thr Ala His His Ser Thr Thr
380 385 390

Val Thr Gln Thr Pro Ser His Ser Trp Ser Asn Gly His Thr Glu
395 400 405

Ser Ile Leu Ser Glu Ser His Ser Val Ile Val Met Ser Ser Val
410 415 420

Glu Asn Ser Arg His Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg
425 430 435

Leu Asn Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg
440 445 450

His Ala Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro His Ser
455 460 465

Glu Arg Tyr Val Ser Ala Met Thr Thr Pro Ala Arg Met Ser Pro
470 475 480

Val Asp Phe His Thr Pro Ser Ser Pro Lys Ser Pro Pro Ser Glu
485 490 495

Met Ser Pro Pro Val Ser Ser Met Thr Val Ser Met Pro Ser Met
500 505 510

Ala Val Ser Pro Phe Met Glu Glu Glu Arg Pro Leu Leu Leu Val
515 520 525

Thr Pro Pro Arg Leu Arg Glu Lys Lys Phe Asp His His Pro Gln
530 535 540

```
Gln Phe Ser Ser Phe His His Asn Pro Ala His Asp Ser Asn Ser
                545                 550                 555

Leu Pro Ala Ser Pro Leu Arg Ile Val Glu Asp Glu Glu Tyr Glu
            560                 565                 570
Thr Thr Gln Glu Tyr Glu Pro Ala Gln Glu Pro Val Lys Lys Leu
        575                 580                 585

Ala Asn Ser Arg Arg Ala Lys Arg Thr Lys Pro Asn Gly His Ile
    590                 595                 600

Ala Asn Arg Leu Glu Val Asp Ser Asn Thr Ser Ser Gln Ser Ser
605                 610                 615

Asn Ser Glu Ser Glu Thr Glu Asp Glu Arg Val Gly Glu Asp Thr
620                 625                 630

Pro Phe Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser Leu Glu Ala
635                 640                 645

Thr Pro Ala Phe Arg Leu Ala Asp Ser Arg Thr Asn Pro Ala Gly
650                 655                 660

Arg Phe Ser Thr Gln Glu Glu Ile Gln
665                 669
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 9:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 732 Aminosäure

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

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

```
Asp Lys Leu Phe Pro Asn Pro Ile Arg Ala Leu Gly Pro Asn Ser
  1               5                  10                  15

Pro Ala Pro Arg Ala Val Arg Val Glu Arg Ser Val Ser Gly Glu
            20                  25                  30

Met Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys Lys
        35                  40                  45

Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala Ala Gly Ser
    50                  55                  60

Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys Glu Met Lys Ser Gln
65                  70                  75

Glu Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser
            80                  85                  90

Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn
        95                  100                 105

Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys
    110                 115                 120
```

Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala
125 130 135

Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu Gly Asn
140 145 150

Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile
155 160 165

Ile Thr Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser
170 175 180

Glu Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr
185 190 195

Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu Val
200 205 210

Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu
215 220 225

Cys Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu Cys
230 235 240

Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln Asn Tyr Val
245 250 255

Met Ala Ser Phe Tyr Lys His Leu Gly Ile Glu Phe Met Glu Ala
260 265 270

Glu Glu Leu Tyr Gln Lys Arg Val Leu Thr Ile Thr Gly Ile Cys
275 280 285

Ile Ala Leu Leu Val Val Gly Ile Met Cys Val Val Ala Tyr Cys
290 295 300

Lys Thr Lys Lys Gln Arg Lys Lys Leu His Asp Arg Leu Arg Gln
305 310 315

Ser Leu Arg Ser Glu Arg Asn Asn Met Met Asn Ile Ala Asn Gly
320 325 330

Pro His His Pro Asn Pro Pro Pro Glu Asn Val Gln Leu Val Asn
335 340 345

Gln Tyr Val Ser Lys Asn Val Ile Ser Ser Glu His Ile Val Glu
350 355 360

Arg Glu Ala Glu Thr Ser Phe Ser Thr Ser His Tyr Thr Ser Thr
365 370 375

Ala His His Ser Thr Thr Val Thr Gln Thr Pro Ser His Ser Trp
380 385 390

Ser Asn Gly His Thr Glu Ser Ile Leu Ser Glu Ser His Ser Val
395 400 405

Ile Val Met Ser Ser Val Glu Asn Ser Arg His Ser Ser Pro Thr
410 415 420

Gly Gly Pro Arg Gly Arg Leu Asn Gly Thr Gly Gly Pro Arg Glu
425 430 435

Cys Asn Ser Phe Leu Arg His Ala Arg Glu Thr Pro Asp Ser Tyr
440 445 450

Arg Asp Ser Pro His Ser Glu Arg Tyr Val Ser Ala Met Thr Thr
455 460 465

Pro Ala Arg Met Ser Pro Val Asp Phe His Thr Pro Ser Ser Pro
470 475 480

Lys Ser Pro Pro Ser Glu Met Ser Pro Pro Val Ser Ser Met Thr
485 490 495

Val Ser Met Pro Ser Met Ala Val Ser Pro Phe Met Glu Glu Glu
500 505 510

Arg Pro Leu Leu Leu Val Thr Pro Pro Arg Leu Arg Glu Lys Lys
515 520 525

Phe Asp His His Pro Gln Gln Phe Ser Ser Phe His His Asn Pro
530 535 540

Ala His Asp Ser Asn Ser Leu Pro Ala Ser Pro Leu Arg Ile Val
545 550 555

Glu Asp Glu Glu Tyr Glu Thr Thr Gln Glu Tyr Glu Pro Ala Gln
560 565 570

Glu Pro Val Lys Lys Leu Ala Asn Ser Arg Arg Ala Lys Arg Thr
575 580 585

Lys Pro Asn Gly His Ile Ala Asn Arg Leu Glu Val Asp Ser Asn
590 595 600

Thr Ser Ser Gln Ser Ser Asn Ser Glu Ser Glu Thr Glu Asp Glu
605 610 615

Arg Val Gly Glu Asp Thr Pro Phe Leu Gly Ile Gln Asn Pro Leu
620 625 630

Ala Ala Ser Leu Glu Ala Thr Pro Ala Phe Arg Leu Ala Asp Ser
635 640 645

Arg Thr Asn Pro Ala Gly Arg Phe Ser Thr Gln Glu Glu Ile Gln
650 655 660

Ala Arg Leu Ser Ser Val Ile Ala Asn Gln Asp Pro Ile Ala Val
665 670 675

Xaa Asn Leu Asn Lys His Ile Asp Ser Pro Val Lys Leu Tyr Phe
680 685 690

Ile Xaa Xaa Ser Ile Pro Pro Xaa Ile Lys Gln Phe Ile Leu Phe
695 700 705

```
Xaa Gln Phe Cys Lys Xaa Lys Thr Gly Lys Lys Leu Leu Xaa Ile
                710                 715                 720

Lys Tyr Met Tyr Val Lys Met Lys Lys Lys Lys Lys
                725                 730     732
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 10:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 66 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 10:

```
Ser His Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val
  1                 5                  10                  15

Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser Asn Pro Ser
                 20                  25                  30

Arg Tyr Leu Cys Lys Cys Gln Pro Gly Phe Thr Gly Ala Arg Cys
                 35                  40                  45

Thr Glu Asn Val Pro Met Lys Val Gln Asn Gln Glu Lys Ala Glu
                 50                  55                  60

Glu Leu Tyr Gln Lys Arg
                 65  66
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 11:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 71 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 11:

```
Ser His Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val
 1               5                  10                  15

Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser Asn Pro Ser
            20                  25                  30

Arg Tyr Leu Cys Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys
            35                  40                  45

Gln Asn Tyr Val Met Ala Ser Phe Tyr Lys His Leu Gly Ile Glu
            50                  55                  60

Phe Met Glu Ala Glu Glu Leu Tyr Gln Lys Arg
            65                  70  71
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 12:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 2010 Basen

(B) TYP: Nucleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 12:

GGGCGCGAGC GCCTCAGCGC GGCCGCTCGC TCTCCCCCTC GAGGGACAAA 50

CTTTTCCCAA ACCCGATCCG AGCCCTTGGA CCAAACTCGC CTGCGCCGAG 100

AGCCGTCCGC GTAGAGCGCT CCGTCTCCGG CGAGATGTCC GAGCGCAAAG 150

AAGGCAGAGG CAAAGGGAAG GGCAAGAAGA AGGAGCGAGG CTCCGGCAAG 200

AAGCCGGAGT CCGCGGCGGG CAGCCAGAGC CCAGCCTTGC CTCCCCGATT 250

GAAAGAGATG AAAAGCCAGG AATCGGCTGC AGGTTCCAAA CTAGTCCTTC 300

GGTGTGAAAC CAGTTCTGAA TACTCCTCTC TCAGATTCAA GTGGTTCAAG 350

AATGGGAATG AATTGAATCG AAAAAACAAA CCACAAAATA TCAAGATACA 400

AAAAAAGCCA GGGAAGTCAG AACTTCGCAT TAACAAAGCA TCACTGGCTG 450

ATTCTGGAGA GTATATGTGC AAAGTGATCA GCAAATTAGG AAATGACAGT 500

GCCTCTGCCA ATATCACCAT CGTGGAATCA AACGAGATCA TCACTGGTAT 550

GCCAGCCTCA ACTGAAGGAG CATATGTGTC TTCAGAGTCT CCCATTAGAA 600

TATCAGTATC CACAGAAGGA GCAAATACTT CTTCATCTAC ATCTACATCC 650

ACCACTGGGA CAAGCCATCT TGTAAAATGT GCGGAGAAGG AGAAAACTTT 700

CTGTGTGAAT GGAGGGGAGT GCTTCATGGT GAAAGACCTT TCAAACCCCT 750

CGAGATACTT GTGCAAGTGC CAACCTGGAT TCACTGGAGC AAGATGTACT 800

GAGAATGTGC CCATGAAAGT CCAAAACCAA GAAAAGGCGG AGGAGCTGTA 850

CCAGAAGAGA GTGCTGACCA TAACCGGCAT CTGCATCGCC CTCCTTGTGG 900

TCGGCATCAT GTGTGTGGTG GCCTACTGCA AAACCAAGAA ACAGCGGAAA 950

AAGCTGCATG ACCGTCTTCG GCAGAGCCTT CGGTCTGAAC GAAACAATAT 1000

GATGAACATT GCCAATGGGC CTCACCATCC TAACCCACCC CCCGAGAATG 1050

TCCAGCTGGT GAATCAATAC GTATCTAAAA ACGTCATCTC CAGTGAGCAT 1100

ATTGTTGAGA GAGAAGCAGA GACATCCTTT TCCACCAGTC ACTATACTTC 1150

CACAGCCCAT CACTCCACTA CTGTCACCCA GACTCCTAGC CACAGCTGGA 1200

GCAACGGACA CACTGAAAGC ATCCTTTCCG AAAGCCACTC TGTAATCGTG 1250

ATGTCATCCG TAGAAAACAG TAGGCACAGC AGCCCAACTG GGGGCCCAAG 1300

AGGACGTCTT AATGGCACAG GAGGCCCTCG TGAATGTAAC AGCTTCCTCA 1350

GGCATGCCAG AGAAACCCCT GATTCCTACC GAGACTCTCC TCATAGTGAA 1400

AGGTATGTGT CAGCCATGAC CACCCCGGCT CGTATGTCAC CTGTAGATTT 1450

CCACACGCCA AGCTCCCCCA AATCGCCCCC TTCGGAAATG TCTCCACCCG 1500

TGTCCAGCAT GACGGTGTCC ATGCCTTCCA TGGCGGTCAG CCCCTTCATG 1550

GAAGAAGAGA GACCTCTACT TCTCGTGACA CCACCAAGGC TGCGGGAGAA 1600

GAAGTTTGAC CATCACCCTC AGCAGTTCAG CTCCTTCCAC CACAACCCCG 1650

CGCATGACAG TAACAGCCTC CCTGCTAGCC CCTTGAGGAT AGTGGAGGAT 1700

GAGGAGTATG AAACGACCCA AGAGTACGAG CCAGCCCAAG AGCCTGTTAA 1750

GAAACTCGCC AATAGCCGGC GGGCCAAAAG AACCAAGCCC AATGGCCACA 1800

TTGCTAACAG ATTGGAAGTG GACAGCAACA CAAGCTCCCA GAGCAGTAAC 1850

TCAGAGAGTG AAACAGAAGA TGAAAGAGTA GGTGAAGATA CGCCTTTCCT 1900

GGGCATACAG AACCCCCTGG CAGCCAGTCT TGAGGCAACA CCTGCCTTCC 1950

GCCTGGCTGA CAGCAGGACT AACCCAGCAG GCCGCTTCTC GACACAGGAA 2000

GAAATCCAGG 2010

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 13:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 669 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 13:

```
Ala Arg Ala Pro Gln Arg Gly Arg Ser Leu Ser Pro Ser Arg Asp
 1               5                  10                  15

Lys Leu Phe Pro Asn Pro Ile Arg Ala Leu Gly Pro Asn Ser Pro
            20                  25                  30

Ala Pro Arg Ala Val Arg Val Glu Arg Ser Val Ser Gly Glu Met
            35                  40                  45

Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys Lys Lys
            50                  55                  60

Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala Ala Gly Ser Gln
            65                  70                  75

  Ser Pro Ala Leu Pro Pro Arg Leu Lys Glu Met Lys Ser Gln Glu
              80                  85                  90

  Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser Ser
              95                  100                 105

  Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn Glu
             110                 115                 120

  Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys Lys
             125                 130                 135

  Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala Asp
             140                 145                 150

  Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu Gly Asn Asp
             155                 160                 165

  Ser Ala Ser Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile Ile
             170                 175                 180

  Thr Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu
             185                 190                 195

  Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr Ser
             200                 205                 210

  Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys
             215                 220                 225
```

```
Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys
230                 235                 240

Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu Cys Lys
245                 250                 255

Cys Gln Pro Gly Phe Thr Gly Ala Arg Cys Thr Glu Asn Val Pro
260                 265                 270

Met Lys Val Gln Asn Gln Glu Lys Ala Glu Glu Leu Tyr Gln Lys
275                 280                 285

Arg Val Leu Thr Ile Thr Gly Ile Cys Ile Ala Leu Leu Val Val
290                 295                 300

Gly Ile Met Cys Val Val Ala Tyr Cys Lys Thr Lys Lys Gln Arg
305                 310                 315

Lys Lys Leu His Asp Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg
320                 325                 330

Asn Asn Met Met Asn Ile Ala Asn Gly Pro His His Pro Asn Pro
335                 340                 345

Pro Pro Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys Asn
350                 355                 360

Val Ile Ser Ser Glu His Ile Val Glu Arg Glu Ala Glu Thr Ser
365                 370                 375

Phe Ser Thr Ser His Tyr Thr Ser Thr Ala His His Ser Thr Thr
380                 385                 390

Val Thr Gln Thr Pro Ser His Ser Trp Ser Asn Gly His Thr Glu

395                 400                 405

Ser Ile Leu Ser Glu Ser His Ser Val Ile Val Met Ser Ser Val
410                 415                 420

Glu Asn Ser Arg His Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg
425                 430                 435

Leu Asn Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg
440                 445                 450

His Ala Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro His Ser
455                 460                 465

Glu Arg Tyr Val Ser Ala Met Thr Thr Pro Ala Arg Met Ser Pro
470                 475                 480

Val Asp Phe His Thr Pro Ser Ser Pro Lys Ser Pro Pro Ser Glu
485                 490                 495

Met Ser Pro Pro Val Ser Ser Met Thr Val Ser Met Pro Ser Met
500                 505                 510

Ala Val Ser Pro Phe Met Glu Glu Glu Arg Pro Leu Leu Leu Val
515                 520                 525

Thr Pro Pro Arg Leu Arg Glu Lys Lys Phe Asp His His Pro Gln
530                 535                 540
```

```
Gln Phe Ser Ser Phe His His Asn Pro Ala His Asp Ser Asn Ser
545                 550                 555

Leu Pro Ala Ser Pro Leu Arg Ile Val Glu Asp Glu Glu Tyr Glu
560                 565                 570

Thr Thr Gln Glu Tyr Glu Pro Ala Gln Glu Pro Val Lys Lys Leu
575                 580                 585

Ala Asn Ser Arg Arg Ala Lys Arg Thr Lys Pro Asn Gly His Ile
590                 595                 600

Ala Asn Arg Leu Glu Val Asp Ser Asn Thr Ser Ser Gln Ser Ser
605                 610                 615

Asn Ser Glu Ser Glu Thr Glu Asp Glu Arg Val Gly Glu Asp Thr
620                 625                 630

Pro Phe Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser Leu Glu Ala
635                 640                 645

Thr Pro Ala Phe Arg Leu Ala Asp Ser Arg Thr Asn Pro Ala Gly
650                 655                 660

Arg Phe Ser Thr Gln Glu Glu Ile Gln
665                 669
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 14:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 95 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 14:

```
Ser His Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val
 1                5                  10                  15

Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser Asn Pro Ser
                20                  25                  30

Arg Tyr Leu Cys Lys Cys Gln Pro Gly Phe Thr Gly Ala Arg Cys
                35                  40                  45

Thr Glu Asn Val Pro Met Lys Val Gln Asn Gln Glu Lys Ala Glu
                50                  55                  60

Glu Leu Tyr Gln Lys Arg Val Leu Thr Ile Thr Gly Ile Cys Ile
                65                  70                  75

Ala Leu Leu Val Val Gly Ile Met Cys Val Val Ala Tyr Cys Lys
                80                  85                  90

Thr Lys Lys Gln Arg
                95
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 15:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 91 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 15:

```
Asn Ser Asp Ser Glu Cys Pro Leu Ser His Asp Gly Tyr Cys Leu
 1               5                  10                  15

His Asp Gly Val Cys Met Tyr Ile Glu Ala Leu Asp Lys Tyr Ala
            20                  25                  30

Cys Asn Cys Val Val Gly Tyr Ile Gly Glu Arg Cys Gln Tyr Arg
        35                  40                  45

Asp Leu Lys Trp Trp Glu Leu Arg His Ala Gly His Gly Gln Gln
    50                  55                  60

Gln Lys Val Ile Val Val Ala Val Cys Val Val Val Leu Val Met
65                  70                  75

Leu Leu Leu Leu Ser Leu Trp Gly Ala His Tyr Tyr Arg Thr Gln
            80                  85                  90

Lys
 91
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 16:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 82 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 16:

```
Asn Asp Cys Pro Asp Ser His Thr Gln Phe Cys Phe His Gly Thr
 1               5                  10                  15

Cys Arg Phe Leu Val Gln Glu Asp Lys Pro Ala Cys Val Cys His
            20                  25                  30

Ser Gly Tyr Val Gly Ala Arg Cys Glu His Ala Asp Leu Leu Ala
            35                  40                  45

Val Val Ala Ala Ser Gln Lys Lys Gln Ala Ile Thr Ala Leu Val
            50                  55                  60

Val Val Ser Ile Val Ala Leu Ala Val Leu Ile Ile Thr Cys Val
            65                  70                  75

Leu Ile His Cys Cys Gln Val
            80      82
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 17:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 87 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 17:

```
Lys Lys Lys Asn Pro Cys Asn Ala Glu Phe Gln Asn Phe Cys Ile
 1               5                  10                  15

His Gly Glu Cys Lys Tyr Ile Glu His Leu Glu Ala Val Thr Cys
            20                  25                  30

Lys Cys Gln Gln Glu Tyr Phe Gly Glu Arg Cys Gly Glu Lys Ser
            35                  40                  45

Met Lys Thr His Ser Met Ile Asp Ser Ser Leu Ser Lys Ile Ala
            50                  55                  60

Leu Ala Ala Ile Ala Ala Phe Met Ser Ala Val Ile Leu Thr Ala
            65                  70                  75

Val Ala Val Ile Thr Val Gln Leu Arg Arg Gln Tyr
            80                  85      87
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 18:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 87 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 18:

```
Lys Lys Lys Asn Pro Cys Ala Ala Lys Phe Gln Asn Phe Cys Ile
 1               5                  10                  15

His Gly Glu Cys Arg Tyr Ile Glu Asn Leu Glu Val Val Thr Cys
            20                  25                  30

His Cys His Gln Asp Tyr Phe Gly Glu Arg Cys Gly Glu Lys Thr
            35                  40                  45

Met Lys Thr Gln Lys Lys Asp Asp Ser Asp Leu Ser Lys Ile Ala
            50                  55                  60

Leu Ala Ala Ile Ile Val Phe Val Ser Ala Val Ser Val Ala Ala
            65                  70                  75

Ile Gly Ile Ile Thr Ala Val Leu Leu Arg Lys Arg
            80                  85      87
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 19:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 86 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 19:

```
Lys Lys Arg Asp Pro Cys Leu Arg Lys Tyr Lys Asp Phe Cys Ile
 1               5                  10                  15

His Gly Glu Cys Lys Tyr Val Lys Glu Leu Arg Ala Pro Ser Cys
            20                  25                  30

Ile Cys His Pro Gly Tyr His Gly Glu Arg Cys His Gly Leu Ser
        35                  40                  45

Leu Pro Val Glu Asn Arg Leu Tyr Thr Tyr Asp His Thr Thr Ile
    50                  55                  60

Leu Ala Val Val Ala Val Val Leu Ser Ser Val Cys Leu Leu Val
65                  70                  75

Ile Val Gly Leu Leu Met Phe Arg Tyr His Arg
            80                  85  86
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 20:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 13 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 20:

```
Arg Pro Asn Ala Arg Leu Pro Pro Gly Val Phe Tyr Cys
 1               5                  10          13
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 21:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 25 Basen

(B) TYP: Nucleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 21:

```
CCTCGCTCCT TCTTCTTGCC CTTCC 25
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 22:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 496 Basen

(B) TYP: Nucleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 22:

```
AA  AGA GCC GGC GAG GAG TTC CCC GAA ACT TGT TGG AAC 38
    Arg Ala Gly Glu Glu Phe Pro Glu Thr Cys Trp Asn
      1               5                  10

TCC GGG CTC GCG CGG AGG CCA GGA GCT GAG CGG CGG CGG 77
Ser Gly Leu Ala Arg Arg Pro Gly Ala Glu Arg Arg Arg
         15                  20                  25

CTG CCG GAC GAT GGG AGC GTG AGC AGG ACG GTG ATA ACC 116
Leu Pro Asp Asp Gly Ser Val Ser Arg Thr Val Ile Thr
                30                  35

TCT CCC CGA TCG GGT TGC GAG GGC GCC GGG CAG AGG CCA 155
Ser Pro Arg Ser Gly Cys Glu Gly Ala Gly Gln Arg Pro
     40                  45                  50

GGA CGC GAG CCG CCA GCG GTG GGA CCC ATC GAC GAC TTC 194
Gly Arg Glu Pro Pro Ala Val Gly Pro Ile Asp Asp Phe
            55                  60
```

```
CCG GGG CGA CAG GAG CAG CCC CGA GAG CCA GGG CGA GCG 233
Pro Gly Arg Gln Glu Gln Pro Arg Glu Pro Gly Arg Ala
 65                  70                  75

CCC GTT CCA GGT GGC CGG ACC GCC CGC CGC GTC CGC GCC 272
Pro Val Pro Gly Gly Arg Thr Ala Arg Arg Val Arg Ala
         80                  85                  90

GCG CTC CCT GCA GGC AAC GGG AGA CGC CCC CGC GCA GCG 311
Ala Leu Pro Ala Gly Asn Gly Arg Arg Pro Arg Ala Ala
                95                 100

CGA GCG CCT CAG CGC GGC CGC TCG CTC TCC CCC TCG AGG 350
Arg Ala Pro Gln Arg Gly Arg Ser Leu Ser Pro Ser Arg
    105                 110                 115

GAC AAA CTT TTC CCA AAC CCG ATC CGA GCC CTT GGA CCA 389
Asp Lys Leu Phe Pro Asn Pro Ile Arg Ala Leu Gly Pro
            120                 125

AAC TCG CCT GCG CCG AGA GCC GTC CGC GTA GAG CGC TCC 428
Asn Ser Pro Ala Pro Arg Ala Val Arg Val Glu Arg Ser
130                 135                 140

GTC TCC GGC GAG ATG TCC GAG CGC AAA GAA GGC AGA GGC 467
Val Ser Gly Glu Met Ser Glu Arg Lys Glu Gly Arg Gly
        145                 150                 155

AAA GGG AAG GGC AAG AAG AAG GAG CGA GG 496
Lys Gly Lys Gly Lys Lys Lys Glu Arg
               160                 164
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 23:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 2490 Basen

(B) TYP: Nucleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 23:

```
GTGGCTGCGG GGCAATTGAA AAAGAGCCGG CGAGGAGTTC CCCGAAACTT 50

GTTGGAACTC CGGGCTCGCG CGGAGGCCAG GAGCTGAGCG GCGGCGGCTG 100

CCGGACGATG GGAGCGTGAG CAGGACGGTG ATAACCTCTC CCCGATCGGG 150

TTGCGAGGGC GCCGGGCAGA GGCCAGGACG CGAGCCGCCA GCGGCGGGAC 200

CCATCGACGA CTTCCCGGGG CGACAGGAGC AGCCCCGAGA GCCAGGGCGA 250

GCGCCCGTTC CAGGTGGCCG GACCGCCCGC CGCGTCCGCG CCGCGCTCCC 300

TGCAGGCAAC GGGAGACGCC CCCGCGCAGC GCGAGCGCCT CAGCGCGGCC 350

GCTCGCTCTC CCCATCGAGG GACAAACTTT TCCCAAACCC GATCCGAGCC 400

CTTGGACCAA ACTCGCCTGC GCCGAGAGCC GTCCGCGTAG AGCGCTCCGT 450

CTCCGGCGAG  ATG TCC GAG CGC AAA GAA GGC AGA GGC AAA 490
            Met Ser Glu Arg Lys Glu Gly Arg Gly Lys
              1               5                  10

GGG AAG GGC AAG AAG AAG GAG CGA GGC TCC GGC AAG AAG 529
Gly Lys Gly Lys Lys Lys Glu Arg Gly Ser Gly Lys Lys
                 15                  20

CCG GAG TCC GCG GCG GGC AGC CAG AGC CCA GCC TTG CCT 568
Pro Glu Ser Ala Ala Gly Ser Gln Ser Pro Ala Leu Pro
     25                  30                  35

CCC CAA TTG AAA GAG ATG AAA AGC CAG GAA TCG GCT GCA 607
Pro Gln Leu Lys Glu Met Lys Ser Gln Glu Ser Ala Ala
            40                  45

GGT TCC AAA CTA GTC CTT CGG TGT GAA ACC AGT TCT GAA 646
Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser Ser Glu
 50                  55                  60

TAC TCC TCT CTC AGA TTC AAG TGG TTC AAG AAT GGG AAT 685
Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn
         65                  70                  75
```

```
GAA TTG AAT CGA AAA AAC AAA CCA CAA AAT ATC AAG ATA 724
Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile
                  80                  85

CAA AAA AAG CCA GGG AAG TCA GAA CTT CGC ATT AAC AAA 763
Gln Lys Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys
     90                  95                 100

GCA TCA CTG GCT GAT TCT GGA GAG TAT ATG TGC AAA GTG 802
Ala Ser Leu Ala Asp Ser Gly Glu Tyr Met Cys Lys Val
            105                 110

ATC AGC AAA TTA GGA AAT GAC AGT GCC TCT GCC AAT ATC 841
Ile Ser Lys Leu Gly Asn Asp Ser Ala Ser Ala Asn Ile
115                 120                 125

ACC ATC GTG GAA TCA AAC GAG ATC ATC ACT GGT ATG CCA 880
Thr Ile Val Glu Ser Asn Glu Ile Ile Thr Gly Met Pro
        130                 135                 140

GCC TCA ACT GAA GGA GCA TAT GTG TCT TCA GAG TCT CCC 919
Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu Ser Pro
                145                 150

ATT AGA ATA TCA GTA TCC ACA GAA GGA GCA AAT ACT TCT 958
Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr Ser
    155                 160                 165

TCA TCT ACA TCT ACA TCC ACC ACT GGG ACA AGC CAT CTT 997
Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu
            170                 175

GTA AAA TGT GCG GAG AAG GAG AAA ACT TTC TGT GTG AAT 1036
Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn
180                 185                 190

GGA GGG GAG TGC TTC ATG GTG AAA GAC CTT TCA AAC CCC 1075
Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser Asn Pro
        195                 200                 205

TCG AGA TAC TTG TGC AAG TGC CCA AAT GAG TTT ACT GGT 1114
Ser Arg Tyr Leu Cys Lys Cys Pro Asn Glu Phe Thr Gly
                210                 215

GAT CGC TGC CAA AAC TAC GTA ATG GCC AGC TTC TAC AAG 1153
Asp Arg Cys Gln Asn Tyr Val Met Ala Ser Phe Tyr Lys
    220                 225                 230

GCG GAG GAG CTG TAC CAG AAG AGA GTG CTG ACC ATA ACC 1192
Ala Glu Glu Leu Tyr Gln Lys Arg Val Leu Thr Ile Thr
            235                 240

GGC ATC TGC ATC GCC CTC CTT GTG GTC GGC ATC ATG TGT 1231
Gly Ile Cys Ile Ala Leu Leu Val Val Gly Ile Met Cys
245                 250                 255

GTG GTG GCC TAC TGC AAA ACC AAG AAA CAG CGG AAA AAG 1270
Val Val Ala Tyr Cys Lys Thr Lys Lys Gln Arg Lys Lys
        260                 265                 270

CTG CAT GAC CGT CTT CGG CAG AGC CTT CGG TCT GAA CGA 1309
Leu His Asp Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg
                275                 280
```

```
AAC AAT ATG ATG AAC ATT GCC AAT GGG CCT CAC CAT CCT 1348
Asn Asn Met Met Asn Ile Ala Asn Gly Pro His His Pro
    285                 290                 295

AAC CCA CCC CCC GAG AAT GTC CAG CTG GTG AAT CAA TAC 1387
Asn Pro Pro Pro Glu Asn Val Gln Leu Val Asn Gln Tyr
            300                 305

GTA TCT AAA AAC GTC ATC TCC AGT GAG CAT ATT GTT GAG 1426
Val Ser Lys Asn Val Ile Ser Ser Glu His Ile Val Glu
310                 315                 320

AGA GAA GCA GAG ACA TCC TTT TCC ACC AGT CAC TAT ACT 1465
Arg Glu Ala Glu Thr Ser Phe Ser Thr Ser His Tyr Thr
        325                 330                 335

TCC ACA GCC CAT CAC TCC ACT ACT GTC ACC CAG ACT CCT 1504
Ser Thr Ala His His Ser Thr Thr Val Thr Gln Thr Pro
                340                 345

AGC CAC AGC TGG AGC AAC GGA CAC ACT GAA AGC ATC CTT 1543
Ser His Ser Trp Ser Asn Gly His Thr Glu Ser Ile Leu
    350                 355                 360

TCC GAA AGC CAC TCT GTA ATC GTG ATG TCA TCC GTA GAA 1582
Ser Glu Ser His Ser Val Ile Val Met Ser Ser Val Glu
            365                 370

AAC AGT AGG CAC AGC AGC CCA ACT GGG GGC CCA AGA GGA 1621
Asn Ser Arg His Ser Ser Pro Thr Gly Gly Pro Arg Gly
375                 380                 385

CGT CTT AAT GGC ACA GGA GGC CCT CGT GAA TGT AAC AGC 1660
Arg Leu Asn Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser
        390                 395                 400

TTC CTC AGG CAT GCC AGA GAA ACC CCT GAT TCC TAC CGA 1699
Phe Leu Arg His Ala Arg Glu Thr Pro Asp Ser Tyr Arg
                405                 410

GAC TCT CCT CAT AGT GAA AGG TAT GTG TCA GCC ATG ACC 1738
Asp Ser Pro His Ser Glu Arg Tyr Val Ser Ala Met Thr
    415                 420                 425

ACC CCG GCT CGT ATG TCA CCT GTA GAT TTC CAC ACG CCA 1777
Thr Pro Ala Arg Met Ser Pro Val Asp Phe His Thr Pro
            430                 435

AGC TCC CCC AAA TCG CCC CCT TCG GAA ATG TCT CCA CCC 1816
Ser Ser Pro Lys Ser Pro Pro Ser Glu Met Ser Pro Pro
440                 445                 450

GTG TCC AGC ATG ACG GTG TCC AAG CCT TCC ATG GCG GTC 1855
Val Ser Ser Met Thr Val Ser Lys Pro Ser Met Ala Val
        455                 460                 465

AGC CCC TTC ATG GAA GAA GAG AGA CCT CTA CTT CTC GTG 1894
Ser Pro Phe Met Glu Glu Glu Arg Pro Leu Leu Leu Val
                470                 475

ACA CCA CCA AGG CTG CGG GAG AAG AAG TTT GAC CAT CAC 1933
Thr Pro Pro Arg Leu Arg Glu Lys Lys Phe Asp His His
    480                 485                 490
```

```
CCT CAG CAG TTC AGC TCC TTC CAC CAC AAC CCC GCG CAT 1972
Pro Gln Gln Phe Ser Ser Phe His His Asn Pro Ala His
            495                 500

GAC AGT AAC AGC CTC CCT GCT AGC CCC TTG AGG ATA GTG 2011
Asp Ser Asn Ser Leu Pro Ala Ser Pro Leu Arg Ile Val
505                 510                 515

GAG GAT GAG GAG TAT GAA ACG ACC CAA GAG TAC GAG CCA 2050
Glu Asp Glu Glu Tyr Glu Thr Thr Gln Glu Tyr Glu Pro
        520                 525                 530

GCC CAA GAG CCT GTT AAG AAA CTC GCC AAT AGC CGG CGG 2089
Ala Gln Glu Pro Val Lys Lys Leu Ala Asn Ser Arg Arg
                535                 540

GCC AAA AGA ACC AAG CCC AAT GGC CAC ATT GCT AAC AGA 2128
Ala Lys Arg Thr Lys Pro Asn Gly His Ile Ala Asn Arg
    545                 550                 555

TTG GAA GTG GAC AGC AAC ACA AGC TCC CAG AGC AGT AAC 2167
Leu Glu Val Asp Ser Asn Thr Ser Ser Gln Ser Ser Asn
            560                 565

TCA GAG AGT GAA ACA GAA GAT GAA AGA GTA GGT GAA GAT 2206
Ser Glu Ser Glu Thr Glu Asp Glu Arg Val Gly Glu Asp
570                 575                 580

ACG CCT TTC CTG GGC ATA CAG AAC CCC CTG GCA GCC AGT 2245
Thr Pro Phe Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser
        585                 590                 595

CTT GAG GCA ACA CCT GCC TTC CGC CTG GCT GAC AGC AGG 2284
Leu Glu Ala Thr Pro Ala Phe Arg Leu Ala Asp Ser Arg
                600                 605

ACT AAC CCA GCA GGC CGC TTC TCG ACA CAG GAA GAA ATC 2323
Thr Asn Pro Ala Gly Arg Phe Ser Thr Gln Glu Glu Ile
    610                 615                 620

CAG GCC AGG CTG TCT AGT GTA ATT GCT AAC CAA GAC CCT 2362
Gln Ala Arg Leu Ser Ser Val Ile Ala Asn Gln Asp Pro
            625                 630

ATT GCT GTA TAAAACCTA AATAAACACA TAGATTCACC TGTAAAACTT 2410
Ile Ala Val
635     637

TATTTTATAT AATAAAGTAT TCCACCTTAA ATTAAACAAT TTATTTTATT 2460


TTAGCAGTTC TGCAAATAAA AAAAAAAAAA 2490
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 24:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 1715 Basen

(B) TYP: Nucleinsäure

(C) ANZAHL DER STRÄNGE: einzeln

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 24:

```
GCGCCTGCCT CCAACCTGCG GGCGGGAGGT GGGTGGCTGC GGGGCAATTG 50

AAAAAGAGCC GGCGAGGAGT TCCCCGAAAC TTGTTGGAAC TCCGGGCTCG 100

CGCGGAGGCC AGGAGCTGAG CGGCGGCGGC TGCCGGACGA TGGGAGCGTG 150

AGCAGGACGG TGATAACCTC TCCCCGATCG GGTTGCGAGG GCGCCGGGCA 200

GAGGCCAGGA CGCGAGCCGC CAGCGGCGGG ACCCATCGAC GACTTCCCGG 250

GGCGACAGGA GCAGCCCCGA GAGCCAGGGC GAGCGCCCGT TCCAGGTGGC 300

CGGACCGCCC GCCGCGTCCG CGCCGCGCTC CCTGCAGGCA ACGGGAGACG 350

CCCCCGCGCA GCGCGAGCGC CTCAGCGCGG CCGCTCGCTC TCCCCATCGA 400

GGGACAAACT TTTCCCAAAC CCGATCCGAG CCCTTGGACC AAACTCGCCT 450

GCGCCGAGAG CCGTCCGCGT AGAGCGCTCC GTCTCCGGCG AG  ATG 495
                                                Met
                                                 1

TCC GAG CGC AAA GAA GGC AGA GGC AAA GGG AAG GGC AAG 534
Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys
            5                  10

AAG AAG GAG CGA GGC TCC GGC AAG AAG CCG GAG TCC GCG 573
Lys Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala
 15                 20                  25

GCG GGC AGC CAG AGC CCA GCC TTG CCT CCC CAA TTG AAA 612
Ala Gly Ser Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys
        30                  35                  40
```

```
GAG ATG AAA AGC CAG GAA TCC GCT GCA GGT TCC AAA CTA 651
Glu Met Lys Ser Gln Glu Ser Ala Ala Gly Ser Lys Leu
                  45                  50

GTC CTT CGG TGT GAA ACC AGT TCT GAA TAC TCC TCT CTC 690
Val Leu Arg Cys Glu Thr Ser Ser Glu Tyr Ser Ser Leu
     55                  60                  65

AGA TTC AAG TGG TTC AAG AAT GGG AAT GAA TTG AAT CGA 729
Arg Phe Lys Trp Phe Lys Asn Gly Asn Glu Leu Asn Arg
               70                  75

AAA AAC AAA CCA CAA AAT ATC AAG ATA CAA AAA AAG CCA 768
Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys Lys Pro
 80                  85                  90

GGG AAG TCA GAA CTT CGC ATT AAC AAA GCA TCA CTG GCT 807
Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala
         95                 100                 105

GAT TCT GGA GAG TAT ATG TGC AAA GTG ATC AGC AAA TTA 846
Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu
               110                 115

GGA AAT GAC AGT GCC TCT GCC AAT ATC ACC ATC GTG GAA 885
Gly Asn Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu
    120                 125                 130

TCA AAC GAG ATC ATC ACT GGT ATG CCA GCC TCA ACT GAA 924
Ser Asn Glu Ile Ile Thr Gly Met Pro Ala Ser Thr Glu
            135                 140

GGA GCA TAT GTG TCT TCA GAG TCT CCC ATT AGA ATA TCA 963
Gly Ala Tyr Val Ser Ser Glu Ser Pro Ile Arg Ile Ser
145                 150                 155

GTA TCC ACA GAA GGA GCA AAT ACT TCT TCA TCT ACA TCT 1002
Val Ser Thr Glu Gly Ala Asn Thr Ser Ser Ser Thr Ser
        160                 165                 170

ACA TCC ACC ACT GGG ACA AGC CAT CTT GTA AAA TGT GCG 1041
Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys Cys Ala
               175                 180

GAG AAG GAG AAA ACT TTC TGT GTG AAT GGA GGG GAG TGC 1080
Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys
    185                 190                 195

TTC ATG GTG AAA GAC CTT TCA AAC CCC TCG AGA TAC TTG 1119
Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu
            200                 205

TGC AAG TGC CCA AAT GAG TTT ACT GGT GAT CGC TGC CAA 1158
Cys Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln
210                 215                 220

AAC TAC GTA ATG GCC AGC TTC TAC AGT ACG TCC ACT CCC 1197
Asn Tyr Val Met Ala Ser Phe Tyr Ser Thr Ser Thr Pro
        225                 230                 235

TTT CTG TCT CTG CCT GAA TAGGA GCATGCTCAG TTGGTGCTGC 1240
Phe Leu Ser Leu Pro Glu
               240 241

TTTCTTGTTG CTGCATCTCC CCTCAGATTC CACCTAGAGC TAGATGTGTC 1290
```

TTACCAGATC TAATATTGAC TGCCTCTGCC TGTCGCATGA GAACATTAAC 1340

AAAAGCAATT GTATTACTTC CTCTGTTCGC GACTAGTTGG CTCTGAGATA 1390

CTAATAGGTG TGTGAGGCTC CGGATGTTTC TGGAATTGAT ATTGAATGAT 1440

GTGATACAAA TTGATAGTCA ATATCAAGCA GTGAAATATG ATAATAAAGG 1490

CATTTCAAAG TCTCACTTTT ATTGATAAAA TAAAAATCAT TCTACTGAAC 1540

AGTCCATCTT CTTTATACAA TGACCACATC CTGAAAAGGG TGTTGCTAAG 1590

CTGTAACCGA TATGCACTTG AAATGATGGT AAGTTAATTT TGATTCAGAA 1640

TGTGTTATTT GTCACAAATA AACATAATAA AAGGAGTTCA GATGTTTTTC 1690

TTCATTAACC AAAAAAAAAA AAAAA 1715

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 25:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 2431 Basen

(B) TYP: Nucleinsäure

(C) ANZAHL DER STRÄNGE: N.A.

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 25:

GAGGCGCCTG CCTCCAACCT GCGGGCGGGA GGTGGGTGGC TGCGGGGCAA 50

TTGAAAAAGA GCCGGCGAGG AGTTCCCCGA AACTTGTTGG AACTCCGGGC 100

TCGCGCGGAG GCCAGGAGCT GAGCGGCGGC GGCTGCCGGA CGATGGGAGC 150

GTGAGCAGGA CGGTGATAAC CTCTCCCCGA TCGGGTTGCG AGGGCGCCGG 200

GCAGAGGCCA GGACGCGAGC CGCCAGCGGC GGGACCCATC GACGACTTCC 250

```
CGGGGCGACA GGAGCAGCCC CGAGAGGCAG GGCGAGCGCC CGTTCCAGGT 300

GGCCGGACCG CCCGCCGCGT CCGCGCCGCG CTCCCTGCAG GCAACGGGAG 350

ACGCCCCCGC GCAGCGCGAG CGCCTCAGCG CGGCCGCTCG CTCTCCCCAT 400

CGAGGGACAA ACTTTTCCCA AACCCGATCC GAGCCCTTGG ACCAAACTCG 450

CCTGCGCCGA GAGCCGTCCG CGTAGAGCGC TCCGTCTCCG GCGAG   AT 497
                                                      Met
                                                       1

G TCC GAG CGC AAA GAA GGC AGA GGC AAA GGG AAG GGC AAG 537
  Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys
              5                  10

AAG AAG GAG CGA GGC TCC GGC AAG AAG CCG GAG TCC GCG 576
Lys Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala
 15                  20                  25

 GCG GGC AGC CAG AGC CCA GCC TTG CCT CCC CAA TTG AAA 615
 Ala Gly Ser Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys
          30                  35                  40

 GAG ATG AAA AGC CAG GAA TCG GCT GCA GGT TCC AAA CTA 654
 Glu Met Lys Ser Gln Glu Ser Ala Ala Gly Ser Lys Leu
                  45                  50

 GTC CTT CGG TGT GAA ACC AGT TCT GAA TAC TCC TCT CTC 693
 Val Leu Arg Cys Glu Thr Ser Ser Glu Tyr Ser Ser Leu
      55                  60                  65

 AGA TTC AAG TGG TTC AAG AAT GGG AAT GAA TTG AAT CGA 732
 Arg Phe Lys Trp Phe Lys Asn Gly Asn Glu Leu Asn Arg
              70                  75

 AAA AAC AAA CCA CAA AAT ATC AAG ATA CAA AAA AAG CCA 771
 Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys Lys Pro
  80                  85                  90

 GGG AAG TCA GAA CTT CGC ATT AAC AAA GCA TCA CTG GCT 810
 Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala
          95                  100                 105

 GAT TCT GGA GAG TAT ATG TGC AAA GTG ATC AGC AAA TTA 849
 Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu
                  110                 115

 GGA AAT GAC AGT GCC TCT GCC AAT ATC ACC ATC GTG GAA 888
 Gly Asn Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu
     120                 125                 130

 TCA AAC GAG ATC ATC ACT GGT ATG CCA GCC TCA ACT GAA 927
 Ser Asn Glu Ile Ile Thr Gly Met Pro Ala Ser Thr Glu
              135                 140

 GGA GCA TAT GTG TCT TCA GAG TCT CCC ATT AGA ATA TCA 966
 Gly Ala Tyr Val Ser Ser Glu Ser Pro Ile Arg Ile Ser
 145                 150                 155
```

```
GTA TCC ACA GAA GGA GCA AAT ACT TCT TCA TCT ACA TCT 1005
Val Ser Thr Glu Gly Ala Asn Thr Ser Ser Ser Thr Ser
        160                 165                 170

ACA TCC ACC ACT GGG ACA AGC CAT CTT GTA AAA TGT GCG 1044
Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys Cys Ala
                175                 180

GAG AAG GAG AAA ACT TTC TGT GTG AAT GGA GGG GAG TGC 1083
Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys
    185                 190                 195

TTC ATG GTG AAA GAC CTT TCA AAC CCC TCG AGA TAC TTG 1122
Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu
            200                 205

TGC AAG TGC CCA AAT GAG TTT ACT GGT GAT CGC TGC CAA 1161
Cys Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln
210                 215                 220

AAC TAC GTA ATG GCC AGC TTC TAC AAG GCG GAG GAG CTG 1200
Asn Tyr Val Met Ala Ser Phe Tyr Lys Ala Glu Glu Leu
        225                 230                 235

 TAC CAG AAG AGA GTG CTG ACC ATA ACC GGC ATC TGC ATC 1239
 Tyr Gln Lys Arg Val Leu Thr Ile Thr Gly Ile Cys Ile
                240                 245

 GCC CTC CTT GTG GTC GGC ATC ATG TGT GTG GTG GCC TAC 1278
 Ala Leu Leu Val Val Gly Ile Met Cys Val Val Ala Tyr
    250                 255                 260

 TGC AAA ACC AAG AAA CAG CGG AAA AAG CTG CAT GAC CGT 1317
 Cys Lys Thr Lys Lys Gln Arg Lys Lys Leu His Asp Arg
            265                 270

 CTT CGG CAG AGC CTT CGG TCT GAA CGA AAC AAT ATG ATG 1356
 Leu Arg Gln Ser Leu Arg Ser Glu Arg Asn Asn Met Met
 275                 280                 285

 AAC ATT GCC AAT GGG CCT CAC CAT CCT AAC CCA CCC CCC 1395
 Asn Ile Ala Asn Gly Pro His His Pro Asn Pro Pro Pro
        290                 295                 300

 GAG AAT GTC CAG CTG GTG AAT CAA TAC GTA TCT AAA AAC 1434
 Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys Asn
                305                 310

 GTC ATC TCC AGT GAG CAT ATT GTT GAG AGA GAA GCA GAG 1473
 Val Ile Ser Ser Glu His Ile Val Glu Arg Glu Ala Glu
    315                 320                 325

 ACA TCC TTT TCC ACC AGT CAC TAT ACT TCC ACA GCC CAT 1512
 Thr Ser Phe Ser Thr Ser His Tyr Thr Ser Thr Ala His
            330                 335

 CAC TCC ACT ACT GTC ACC CAG ACT CCT AGC CAC AGC TGG 1551
 His Ser Thr Thr Val Thr Gln Thr Pro Ser His Ser Trp
 340                 345                 350

 AGC AAC GGA CAC ACT GAA AGC ATC CTT TCC GAA AGC CAC 1590
 Ser Asn Gly His Thr Glu Ser Ile Leu Ser Glu Ser His
        355                 360                 365
```

```
TCT GTA ATC GTG ATG TCA TCC CTA GAA AAC AGT AGG CAC 1629
Ser Val Ile Val Met Ser Ser Val Glu Asn Ser Arg His
                370                 375

AGC AGC CCA ACT GGG GGC CCA AGA GGA CGT CTT AAT GGC 1668
Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg Leu Asn Gly
    380                 385                 390

ACA GGA GGC CCT CGT GAA TGT AAC AGC TTC CTC AGG CAT 1707
Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg His
            395                 400

GCC AGA GAA ACC CCT GAT TCC TAC CGA GAC TCT CCT CAT 1746
Ala Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro His
405                 410                 415

AGT GAA AGG TAAAA CCGAAGGCAA AGCTACTGCA GAGGAGAAAC 1790
Ser Glu Arg
        420

TCAGTCAGAG AATCCCTGTG AGCACCTGCG GTCTCACCTC AGGAAATCTA 1840

CTCTAATCAG AATAAGGGGC GGCAGTTACC TGTTCTAGGA GTGCTCCTAG 1890

TTGATGAAGT CATCTCTTTG TTTGACGGAA CTTATTTCTT CTGAGCTTCT 1940

CTCGTCGTCC CAGTGACTGA CAGGCAACAG ACTCTTAAAG AGCTGGGATG 1990

CTTTGATGCG GAAGGTGCAG CACATGGAGT TTCCAGCTCT GGCCATGGGC 2040

TCAGACCCAC TCGGGGTCTC AGTGTCCTCA GTTGTAACAT TAGAGAGATG 2090

GCATCAATGC TTGATAAGGA CCCTTCTATA ATTCCAATTG CCAGTTATCC 2140

AAACTCTGAT TCGGTGGTCG AGCTGGCCTC GTGTTCTTAT CTGCTATCC 2190

TGTCTTACCT TCCAGCCTCA GTTAAGTCAA ATCAAGGGCT ATGTCATTGC 2240

TGAATGTCAT GGGGGGCAAC TGCTTGCCCT CCACCCTATA GTATCTATTT 2290

TATGAAATTC CAAGAAGGGA TGAATAAATA AATCTCTTGG ATGCTGCGTC 2340

TGGCAGTCTT CACGGGTGGT TTTCAAAGCA GAAAAAAAAA AAAAAAAAAA 2390

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA A 2431
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 26:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 625 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 26:

```
Met Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys Lys
 1               5                  10                  15

Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala Ala Gly Ser
            20                  25                  30

Gln Ser Pro Ala Leu Pro Pro Arg Leu Lys Glu Met Lys Ser Gln
            35                  40                  45

Glu Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser
            50                  55                  60

Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn
            65                  70                  75

Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys
            80                  85                  90
```

Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala
95 100 105

Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu Gly Asn
110 115 120

Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile
125 130 135

Ile Thr Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser
140 145 150

Glu Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr
155 160 165

Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu Val
170 175 180

Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu
185 190 195

Cys Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu Cys
200 205 210

Lys Cys Gln Pro Gly Phe Thr Gly Ala Arg Cys Thr Glu Asn Val
215 220 225

Pro Met Lys Val Gln Asn Gln Glu Lys Ala Glu Glu Leu Tyr Gln
230 235 240

Lys Arg Val Leu Thr Ile Thr Gly Ile Cys Ile Ala Leu Leu Val
245 250 255

Val Gly Ile Met Cys Val Val Ala Tyr Cys Lys Thr Lys Lys Gln
260 265 270

Arg Lys Lys Leu His Asp Arg Leu Arg Gln Ser Leu Arg Ser Glu
275 280 285

Arg Asn Asn Met Met Asn Ile Ala Asn Gly Pro His His Pro Asn
290 295 300

Pro Pro Pro Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys
305 310 315

Asn Val Ile Ser Ser Glu His Ile Val Glu Arg Glu Ala Glu Thr
320 325 330

Ser Phe Ser Thr Ser His Tyr Thr Ser Thr Ala His His Ser Thr
335 340 345

Thr Val Thr Gln Thr Pro Ser His Ser Trp Ser Asn Gly His Thr
350 355 360

Glu Ser Ile Leu Ser Glu Ser His Ser Val Ile Val Met Ser Ser
365 370 375

Val Glu Asn Ser Arg His Ser Ser Pro Thr Gly Gly Pro Arg Gly
380 385 390

Arg Leu Asn Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu
395 400 405

```
Arg His Ala Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro His
                410                 415                 420

Ser Glu Arg Tyr Val Ser Ala Met Thr Thr Pro Ala Arg Met Ser
                425                 430                 435

Pro Val Asp Phe His Thr Pro Ser Ser Pro Lys Ser Pro Pro Ser
                440                 445                 450

Glu Met Ser Pro Pro Val Ser Ser Met Thr Val Ser Met Pro Ser
                455                 460                 465

Met Ala Val Ser Pro Phe Met Glu Glu Glu Arg Pro Leu Leu Leu
                470                 475                 480

Val Thr Pro Pro Arg Leu Arg Glu Lys Lys Phe Asp His His Pro
                485                 490                 495

Gln Gln Phe Ser Ser Phe His His Asn Pro Ala His Asp Ser Asn
                500                 505                 510

Ser Leu Pro Ala Ser Pro Leu Arg Ile Val Glu Asp Glu Glu Tyr
                515                 520                 525

Glu Thr Thr Gln Glu Tyr Glu Pro Ala Gln Glu Pro Val Lys Lys
                530                 535                 540

Leu Ala Asn Ser Arg Arg Ala Lys Arg Thr Lys Pro Asn Gly His
                545                 550                 555

Ile Ala Asn Arg Leu Glu Val Asp Ser Asn Thr Ser Ser Gln Ser
                560                 565                 570

Ser Asn Ser Glu Ser Glu Thr Glu Asp Glu Arg Val Gly Glu Asp
                575                 580                 585

Thr Pro Phe Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser Leu Glu
                590                 595                 600

Ala Thr Pro Ala Phe Arg Leu Ala Asp Ser Arg Thr Asn Pro Ala
                605                 610                 615

Gly Arg Phe Ser Thr Gln Glu Glu Ile Gln
                620                 625
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 27:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 645 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 27:

```
Met Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys Lys
 1               5                  10                  15

Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala Ala Gly Ser
            20                  25                  30

Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys Glu Met Lys Ser Gln
            35                  40                  45
```

```
Glu Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser
                50                  55                  60

Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn
                65                  70                  75

Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys
                80                  85                  90

Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala
                95                 100                 105

Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu Gly Asn
               110                 115                 120

Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile
               125                 130                 135

Ile Thr Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser
               140                 145                 150

Glu Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr
               155                 160                 165

Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu Val
               170                 175                 180

Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu
               185                 190                 195

Cys Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu Cys
               200                 205                 210

Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln Asn Tyr Val
               215                 220                 225

Met Ala Ser Phe Tyr Lys His Leu Gly Ile Glu Phe Met Glu Ala
               230                 235                 240

Glu Glu Leu Tyr Gln Lys Arg Val Leu Thr Ile Thr Gly Ile Cys
               245                 250                 255

Ile Ala Leu Leu Val Val Gly Ile Met Cys Val Val Ala Tyr Cys
               260                 265                 270

Lys Thr Lys Lys Gln Arg Lys Lys Leu His Asp Arg Leu Arg Gln
               275                 280                 285

Ser Leu Arg Ser Glu Arg Asn Asn Met Met Asn Ile Ala Asn Gly
               290                 295                 300

Pro His His Pro Asn Pro Pro Pro Glu Asn Val Gln Leu Val Asn
               305                 310                 315

Gln Tyr Val Ser Lys Asn Val Ile Ser Ser Glu His Ile Val Glu
               320                 325                 330

Arg Glu Ala Glu Thr Ser Phe Ser Thr Ser His Tyr Thr Ser Thr
               335                 340                 345

Ala His His Ser Thr Thr Val Thr Gln Thr Pro Ser His Ser Trp
               350                 355                 360
```

Ser Asn Gly His Thr Glu Ser Ile Leu Ser Glu Ser His Ser Val
365 370 375

Ile Val Met Ser Ser Val Glu Asn Ser Arg His Ser Ser Pro Thr
380 385 390

Gly Gly Pro Arg Gly Arg Leu Asn Gly Thr Gly Gly Pro Arg Glu
395 400 405

Cys Asn Ser Phe Leu Arg His Ala Arg Glu Thr Pro Asp Ser Tyr
410 415 420

Arg Asp Ser Pro His Ser Glu Arg Tyr Val Ser Ala Met Thr Thr
425 430 435

Pro Ala Arg Met Ser Pro Val Asp Phe His Thr Pro Ser Ser Pro
440 445 450

Lys Ser Pro Pro Ser Glu Met Ser Pro Pro Val Ser Ser Met Thr
455 460 465

Val Ser Met Pro Ser Met Ala Val Ser Pro Phe Met Glu Glu Glu
470 475 480

Arg Pro Leu Leu Leu Val Thr Pro Pro Arg Leu Arg Glu Lys Lys
485 490 495

Phe Asp His His Pro Gln Gln Phe Ser Ser Phe His His Asn Pro
500 505 510

Ala His Asp Ser Asn Ser Leu Pro Ala Ser Pro Leu Arg Ile Val
515 520 525

Glu Asp Glu Glu Tyr Glu Thr Thr Gln Glu Tyr Glu Pro Ala Gln
530 535 540

Glu Pro Val Lys Lys Leu Ala Asn Ser Arg Arg Ala Lys Arg Thr
545 550 555

Lys Pro Asn Gly His Ile Ala Asn Arg Leu Glu Val Asp Ser Asn
560 565 570

Thr Ser Ser Gln Ser Ser Asn Ser Glu Ser Glu Thr Glu Asp Glu
575 580 585

Arg Val Gly Glu Asp Thr Pro Phe Leu Gly Ile Gln Asn Pro Leu
590 595 600

Ala Ala Ser Leu Glu Ala Thr Pro Ala Phe Arg Leu Ala Asp Ser
605 610 615

Arg Thr Asn Pro Ala Gly Arg Phe Ser Thr Gln Glu Glu Ile Gln
620 625 630

Ala Arg Leu Ser Ser Val Ile Ala Asn Gln Asp Pro Ile Ala Val
635 640 645

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 28:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 637 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 28:

```
Met Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys Lys
 1               5                  10                  15

Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala Ala Gly Ser
            20                  25                  30

Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys Glu Met Lys Ser Gln
            35                  40                  45

Glu Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser
            50                  55                  60

Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn
            65                  70                  75

Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys
            80                  85                  90

Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala
            95                 100                 105

Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu Gly Asn
           110                 115                 120

Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile
           125                 130                 135

Ile Thr Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser
           140                 145                 150

Glu Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr
           155                 160                 165

Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu Val
           170                 175                 180

Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu
           185                 190                 195

Cys Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu Cys
           200                 205                 210

Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln Asn Tyr Val
           215                 220                 225

Met Ala Ser Phe Tyr Lys Ala Glu Glu Leu Tyr Gln Lys Arg Val
           230                 235                 240
```

Leu Thr Ile Thr Gly Ile Cys Ile Ala Leu Leu Val Val Gly Ile
245 250 255

Met Cys Val Val Ala Tyr Cys Lys Thr Lys Lys Gln Arg Lys Lys
260 265 270

Leu His Asp Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg Asn Asn
275 280 285

Met Met Asn Ile Ala Asn Gly Pro His His Pro Asn Pro Pro Pro
290 295 300

Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys Asn Val Ile
305 310 315

Ser Ser Glu His Ile Val Glu Arg Glu Ala Glu Thr Ser Phe Ser
320 325 330

Thr Ser His Tyr Thr Ser Thr Ala His His Ser Thr Thr Val Thr
335 340 345

Gln Thr Pro Ser His Ser Trp Ser Asn Gly His Thr Glu Ser Ile
350 355 360

Leu Ser Glu Ser His Ser Val Ile Val Met Ser Ser Val Glu Asn
365 370 375

Ser Arg His Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg Leu Asn
380 385 390

Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg His Ala
395 400 405

Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro His Ser Glu Arg
410 415 420

Tyr Val Ser Ala Met Thr Thr Pro Ala Arg Met Ser Pro Val Asp
425 430 435

Phe His Thr Pro Ser Ser Pro Lys Ser Pro Pro Ser Glu Met Ser
440 445 450

Pro Pro Val Ser Ser Met Thr Val Ser Lys Pro Ser Met Ala Val
455 460 465

Ser Pro Phe Met Glu Glu Glu Arg Pro Leu Leu Leu Val Thr Pro
470 475 480

Pro Arg Leu Arg Glu Lys Lys Phe Asp His His Pro Gln Gln Phe
485 490 495

Ser Ser Phe His His Asn Pro Ala His Asp Ser Asn Ser Leu Pro
500 505 510

Ala Ser Pro Leu Arg Ile Val Glu Asp Glu Glu Tyr Glu Thr Thr
515 520 525

Gln Glu Tyr Glu Pro Ala Gln Glu Pro Val Lys Lys Leu Ala Asn
530 535 540

Ser Arg Arg Ala Lys Arg Thr Lys Pro Asn Gly His Ile Ala Asn
545 550 555

```
Arg Leu Glu Val Asp Ser Asn Thr Ser Ser Gln Ser Ser Asn Ser
                560                 565                 570

Glu Ser Glu Thr Glu Asp Glu Arg Val Gly Glu Asp Thr Pro Phe
                575                 580                 585

Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser Leu Glu Ala Thr Pro
                590                 595                 600

Ala Phe Arg Leu Ala Asp Ser Arg Thr Asn Pro Ala Gly Arg Phe
                605                 610                 615


Ser Thr Gln Glu Glu Ile Gln Ala Arg Leu Ser Ser Val Ile Ala
                620                 625                 630

Asn Gln Asp Pro Ile Ala Val
                635     637
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 29:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 420 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 29:

```
Met Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys Lys
  1                 5                  10                  15

Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala Ala Gly Ser
                20                  25                  30

Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys Glu Met Lys Ser Gln
                35                  40                  45

Glu Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser
                50                  55                  60

Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn
                65                  70                  75

Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys
                80                  85                  90
```

Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala
95 100 105

Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu Gly Asn
110 115 120

Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile
125 130 135

Ile Thr Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser
140 145 150

Glu Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr
155 160 165

Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu Val
170 175 180

Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu
185 190 195

Cys Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu Cys
200 205 210

Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln Asn Tyr Val
215 220 225

Met Ala Ser Phe Tyr Lys Ala Glu Glu Leu Tyr Gln Lys Arg Val
230 235 240

Leu Thr Ile Thr Gly Ile Cys Ile Ala Leu Leu Val Val Gly Ile
245 250 255

Met Cys Val Val Ala Tyr Cys Lys Thr Lys Lys Gln Arg Lys Lys
260 265 270

Leu His Asp Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg Asn Asn
275 280 285

Met Met Asn Ile Ala Asn Gly Pro His His Pro Asn Pro Pro Pro
290 295 300

Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys Asn Val Ile
305 310 315

Ser Ser Glu His Ile Val Glu Arg Glu Ala Glu Thr Ser Phe Ser
320 325 330

Thr Ser His Tyr Thr Ser Thr Ala His His Ser Thr Thr Val Thr
335 340 345

Gln Thr Pro Ser His Ser Trp Ser Asn Gly His Thr Glu Ser Ile
350 355 360

Leu Ser Glu Ser His Ser Val Ile Val Met Ser Ser Val Glu Asn
365 370 375

```
Ser Arg His Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg Leu Asn
                380                 385                 390

Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg His Ala
                395                 400                 405

Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro His Ser Glu Arg
                410                 415                 420
```

(2) INFORMATIONEN ÜBER SEQ.-ID NR. 30:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 241 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ.-ID NR. 30:

```
Met Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys Lys
  1               5                  10                  15

Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala Ala Gly Ser
                20                  25                  30

Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys Glu Met Lys Ser Gln
                35                  40                  45

Glu Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser
                50                  55                  60

Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn
                65                  70                  75

Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys
                80                  85                  90
```

```
Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala
                 95                  100                 105

Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu Gly Asn
                110                 115                 120

Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile
                125                 130                 135

Ile Thr Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser
                140                 145                 150

Glu Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr
                155                 160                 165

Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu Val
                170                 175                 180

Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu
                185                 190                 195

Cys Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu Cys
                200                 205                 210

Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln Asn Tyr Val
                215                 220                 225

Met Ala Ser Phe Tyr Ser Thr Ser Thr Pro Phe Leu Ser Leu Pro
                230                 235                 240

Glu
241
```

Patentansprüche

1. Zusammensetzung, die ein isoliertes Polypeptid umfasst, das eine Sequenz aus zumindest 15 Aminosäureresten umfasst, die in vivo Effektor- oder Antigen-Aktivität eines Heregulins (HRG) der **Fig.** 4, 8, 12, 13 oder 15 und zumindest 75%ige Aminosäuresequenzidentität damit aufweist.

2. Zusammensetzung nach Anspruch 1, worin das isolierte Polypeptid-Heregulin Antigen-Aktivität aufweist.

3. Zusammensetzung nach Anspruch 1, worin das isolierte Polypeptid eine biologische Wirkung von Heregulin aufweist.

4. Zusammensetzung nach Anspruch 3, worin das Heregulin HRG-GFD (-Wachstumsfaktor-Domäne) ist.

5. Zusammensetzung nach Anspruch 1, worin das Heregulin Heregulin-α, -β1, -β2 oder -β3, wie in **Fig.** 15 gezeigt, oder ein Allel oder ein Tieranalog davon ist.

6. Zusammensetzung nach Anspruch 3, worin das Heregulin Human-Heregulin-α-Wachstumsfaktor-Domäne (-GFD) ist.

7. Zusammensetzung nach Anspruch 3, worin das Heregulin Human-Heregulin-β-GFD, Heregulin-β2-GF oder Heregulin-β3-GFD ist.

8. Zusammensetzung nach Anspruch 1, die weiters einen pharmazeutisch annehmbaren Träger umfasst.

9. Zusammensetzung nach Anspruch 8, worin das Heregulin eine Heregulin-GFD (-Wachstumsfaktor-Domäne) ist.

10. Zusammensetzung nach Anspruch 28, die weiters ein Immunadjuvans umfasst.

11. Zusammensetzung nach Anspruch 10, worin die Heregulin-GFD ein immunogenes Nicht-Heregulin-Polypeptid umfasst.

12. Zusammensetzung nach Anspruch 1, worin das Heregulin aus N-terminaler Domäne (NTD) und Wachstumsfaktor-Domäne (GFD) besteht.

13. Zusammensetzung nach Anspruch 1, worin das Heregulin NTD-GFD-Transmembran-Polypeptid ist.

14. Zusammensetzung nach Anspruch 1, worin das Heregulin HRG-GFD ist.

15. Zusammensetzung nach Anspruch 1, worin das Heregulin eine Zytoplasma-Domäne umfasst.

16. Zusammensetzung nach Anspruch 1, worin das Heregulin NTD-GFD ist und eine Aminosäuresequenz aufweist, die zu zumindest 85% homolog mit der nativen Heregulin-α, -β1, -β3 NTD-GFD-Sequenz, wie in **Fig.** 15 gezeigt, oder ein Allel oder ein Tieranalog davon ist.

17. Zusammensetzung nach Anspruch 1, worin das Heregulin-Polypeptid ein Enzym umfasst.

18. Zusammensetzung nach Anspruch 16, worin das Heregulin HRG-α ist.

19. Zusammensetzung nach Anspruch 18, worin das Heregulin-α benachbart zu einem der Reste 1–23, 107–108, 121–123, 128–130 und 163–247, wie in **Fig.** 15 gezeigt, eine substituierte, deletierte oder insertierte Aminosäure aufweist.

20. Zusammensetzung nach Anspruch 16, worin das Neregulin HRG-β1 ist.

21. Zusammensetzung nach Anspruch 20, worin das Heregulin β1 benachbart zu den Resten 1–23, 107–108, 121–123, 128–130 und 163–252, wie in **Fig.** 15 gezeigt, eine substituierte, deletierte oder insertierte Aminosäure aufweist.

22. Zusammensetzung nach Anspruch 16, worin das Heregulin HRG-β2 ist.

23. Zusammensetzung nach Anspruch 22, worin das Heregulin β2 benachbart zu einem der Reste 1–23, 107–108, 121–123, 128–130 und 163–244, wie in **Fig.** 15 gezeigt, eine substituierte, deletierte oder insertierte Aminosäure aufweist.

24. Zusammensetzung nach Anspruch 16, worin das Heregulin HRG-β3 ist.

25. Zusammensetzung nach Anspruch 24, worin das Heregulin β3 benachbart zu einem der Reste 1–23, 107–108, 121–123, 128–130 und 163–241, wie in **Fig.** 15 gezeigt, eine substituierte, deletierte oder insertierte Aminosäure aufweist.

26. Isolierter Antikörper, der fähig ist, sich spezifisch an ein Heregulin-Polypeptid nach Anspruch 1 zu binden.

27. Isolierter Antikörper nach Anspruch 26, der fähig ist, sich spezifisch an Heregulin-α, Heregulin-β1, Heregulin-β2 oder Heregulin-β3 zu binden.

28. Nucleinsäure, die für ein Polypeptid nach Anspruch 1 kodiert.

29. Nucleinsäure nach Anspruch 28, die für Heregulin-α-, Heregulin-β1-, Heregulinβ2- oder Heregulin-β3-Polypeptid kodiert.

30. Nucleinsäure nach Anspruch 28, die für eine Heregulin-GFD kodiert.

31. Expressionsvektor, der eine nucleinsäure nach Anspruch 28 umfasst.

32. Expression nach Anspruch 31, worin die Nucleinsäure für eine Heregulin-GFD kodiert.

33. Wirtszelle, die mit einem Vektor nach Anspruch 31 transformiert ist.

34. Verfahren, umfassend das Kultivieren einer Wirtszelle nach Anspruch 33 zum Exprimieren des Heregulins sowie das Gewinnen des Heregulins aus der Wirtszelle.

35. Verfahren nach Anspruch 34, worin das Heregulin Heregulin-α, Heregulin-β1, Heregulin-β2 oder Heregulin-β3, wie in **Fig.** 15 gezeigt, oder ein Allel oder ein Tieranalog davon ist.

36. Verfahren nach Anspruch 34, worin das Heregulin Heregulin-NTD-GFD ist.

37. Verfahren nach Anspruch 34, worin das Heregulin Heregulin-GFD ist.

38. Verfahren zum Nachweis der Gegenwart einer Heregulin-Nucleinsäure, umfassend das Kontaktieren von Nucleinsäure nach Anspruch 28 mit einer Testproben-Nucleinsäure und die Bestimmung, ob Hybridisierung stattgefunden hat.

39. Verfahren zum Amplifizieren einer Nucleinsäure-Testprobe, umfassend das Primen einer Nucleinsäure-Polymerasekettenreaktion mit Nucleinsäure nach Anspruch 28.

40. Verfahren zur Reinigung eines Heregulins nach Anspruch 1, umfassend das Adsorbieren von Heregulin aus einer kontaminierten Lösung davon an Heparin-Sepharose oder ein Kationen-Austauschharz.

41. Verfahren, umfassend die Herstellung einer Zusammensetzung, die ein isoliertes Polypeptid umfasst, das eine Sequenz aus zumindest 15 Aminosäureresten umfasst, die in vivo Effektor- oder Antigen-Aktivität eines Heregulins (HRG) der **Fig.** 4, 8, 12, 13 oder 15 und zumindest 75%ige Aminosäuresequenzidentität damit aufweist.

42. Verfahren nach Anspruch 1, worin das isolierte Polypeptid-Heregulin Antigen-Aktivität aufweist.

43. Verfahren nach Anspruch 1, worin das isolierte Polypeptid eine biologische Wirkung von Heregulin aufweist.

44. Verfahren nach Anspruch 3, worin das Heregulin HRG-GFD (-Wachstumsfaktor-Domäne) ist.

45. Verfahren nach Anspruch 1, worin das Heregulin Heregulin-α, -β1, -β2 oder -β3, wie in **Fig.** 15 gezeigt, oder ein Allel oder ein Tieranalog davon ist.

46. Verfahren nach Anspruch 3, worin das Heregulin Human-Heregulin-α-Wachstumsfaktor-Domäne (-GFD) ist.

47. Verfahren nach Anspruch 3, worin das Heregulin Human-Heregulin-β-GFD, Heregulin-2β-GFD oder Heregulin-β3-GFD ist.

48. Verfahren nach Anspruch 1, das weiters einen pharmazeutisch annehmbaren Trägerumfasst.

49. Verfahren nach Anspruch 8, worin das Heregulin eine Heregulin-GFD (-Wachstumsfaktor-Domäne) ist.

50. Verfahren nach Anspruch 9; das weiters ein Immunadjuvans umfasst.

51. Verfahren nach Anspruch 10, worin die Heregulin-GFD ein immunogenes Nicht-Heregulin-Polypeptid umfasst.

52. Verfahren nach Anspruch 1, worin das Heregulin aus N-terminaler Domäne (NTD) und Wachstumsfaktor-Domäne (GFD) besteht.

53. Verfahren nach Anspruch 1, worin das Heregulin NTD-GFD-Transmembran-Polypeptid ist.

54. Verfahren nach Anspruch 1, worin das Heregulin HRG-GFD ist.

55. Verfahren nach Anspruch 1, worin das Heregulin eine Zytoplasma-Domäne umfasst.

56. Verfahren nach Anspruch 1, worin das Heregulin NTD-GFD ist und eine Aminosäuresequenz aufweist, die zu zumindest 85% homolog mit einer nativen Heregulin-α, -β1, -β3 NTD-GFD-Sequenz, wie in **Fig.** 15 gezeigt, oder ein Allel oder ein Tieranalog davon ist.

57. Verfahren nach Anspruch 1, worin das Heregulin-Polypeptid ein Enzym umfasst.

58. Verfahren nach Anspruch 16, worin das Heregulin HRG-α ist.

59. Verfahren nach Anspruch 18, worin das Heregulin-α benachbart zu einem der Reste 1–23, 107–108, 121–123, 128–130 und 163–247, wie in **Fig.** 15 gezeigt, eine substituierte, deletierte oder insertierte Aminosäure aufweist.

60. Verfahren nach Anspruch 16, worin das Heregulin HRG-β1 ist.

61. Verfahren nach Anspruch 20, worin das Heregulin β1 benachbart zu den Resten 1–23, 107–108, 121–123, 128–130 und 163–252, wie in **Fig.** 15 gezeigt, eine substituierte, deletierte oder insertierte Aminosäure aufweist.

62. Verfahren nach Anspruch 16, worin das Heregulin HRG-β2 ist.

63. Verfahren nach Anspruch 22, worin das Heregulin β2 benachbart zu einem der Reste 1–23, 107–108, 121–123, 128–130 und 163–244, wie in **Fig.** 15 gezeigt, eine substituierte, deletierte oder insertierte Aminosäure aufweist.

64. Verfahren nach Anspruch 16, worin das Heregulin HRG-β3 ist.

65. Verfahren nach Anspruch 24, worin das Heregulin β3 benachbart zu einem der Reste 1–23, 107–108, 121–123, 128–130 und 163–241, wie in **Fig.** 15 gezeigt, eine substituierte, deletierte oder insertierte Aminosäure aufweist.

66. Verfahren, das die Herstellung eines isolierten Antikörpers umfasst, der fähig ist, sich spezifisch an ein Heregulin-Polypeptid nach Anspruch 1 zu binden.

67. Verfahren nach Anspruch 26, das zur spezifischen Bindung an Heregulin-α, Heregulin-β1, Heregulin-β2 oder Heregulin-β3 fähig ist.

68. Verfahren, das die Herstellung von Nucleinsäure umfasst, die für ein Polypeptid nach Anspruch 1 kodiert.

69. Verfahren nach Anspruch 28, worin die Nucleinsäure für Heregulin-α-, Heregulin-β1-, Heregulin-β2- oder Heregulin-β3-Polypeptid kodiert.

70. Verfahren nach Anspruch 28, worin die Nucleinsäure für eine Heregulin-GFD kodiert.

71. Verfahren, das die Herstellüng eines Expressionsvektors umfasst, der Nucleinsäure nach Anspruch 28 umfasst.

72. Verfahren nach Anspruch 31, worin die Nucleinsäure für eine Heregulin-GFD kodiert.

73. Verfahren, das die Herstellung einer Wirtszelle umfasst, die mit einem Vektor nach Anspruch 31 transformiert ist.

74. Verfahren, umfassend das Kultivieren einer Wirtszelle nach Anspruch 33 zum Exprimieren des Heregulins sowie das Gewinnen des Heregulins aus der Wirtszelle.

75. Verfahren nach Anspruch 34, worin das Heregulin Heregulin-α, Heregulin-β1, Heregulin-β2 oder Here-

gulin-β3, wie in **Fig.** 15 gezeigt, oder ein Allel oder ein Tieranalog davon ist.

76. Verfahren nach Anspruch 34, worin das Heregulin Heregulin-NTD-GFD ist.

77. Verfahren nach Anspruch 34, worin das Heregulin Heregulin-GFD ist.

78. Verfahren zum Nachweis der Gegenwart einer Heregulin-Nucleinsäure, umfassend das Kontaktieren von Nucleinsäure nach Anspruch 28 mit einer Testproben-Nucleinsäure und die Bestimmung, ob Hybridisierung stattgefunden hat.

79. Verfahren zum Amplifizieren einer Nucleinsäure-Testprobe, umfassend das Primen einer Nucleinsäure-Polymerasekettenreaktion mit Nucleinsäure nach Anspruch 28.

80. Verfahren zur Reinigung eines Heregulins nach Anspruch 1, umfassend das Adsorbieren von Heregulin aus einer kontaminierten Lösung davon an Heparin-Sepharose oder ein Kationen-Austauschharz.

Es folgen 33 Blatt Zeichnungen

Anhängende Zeichnungen

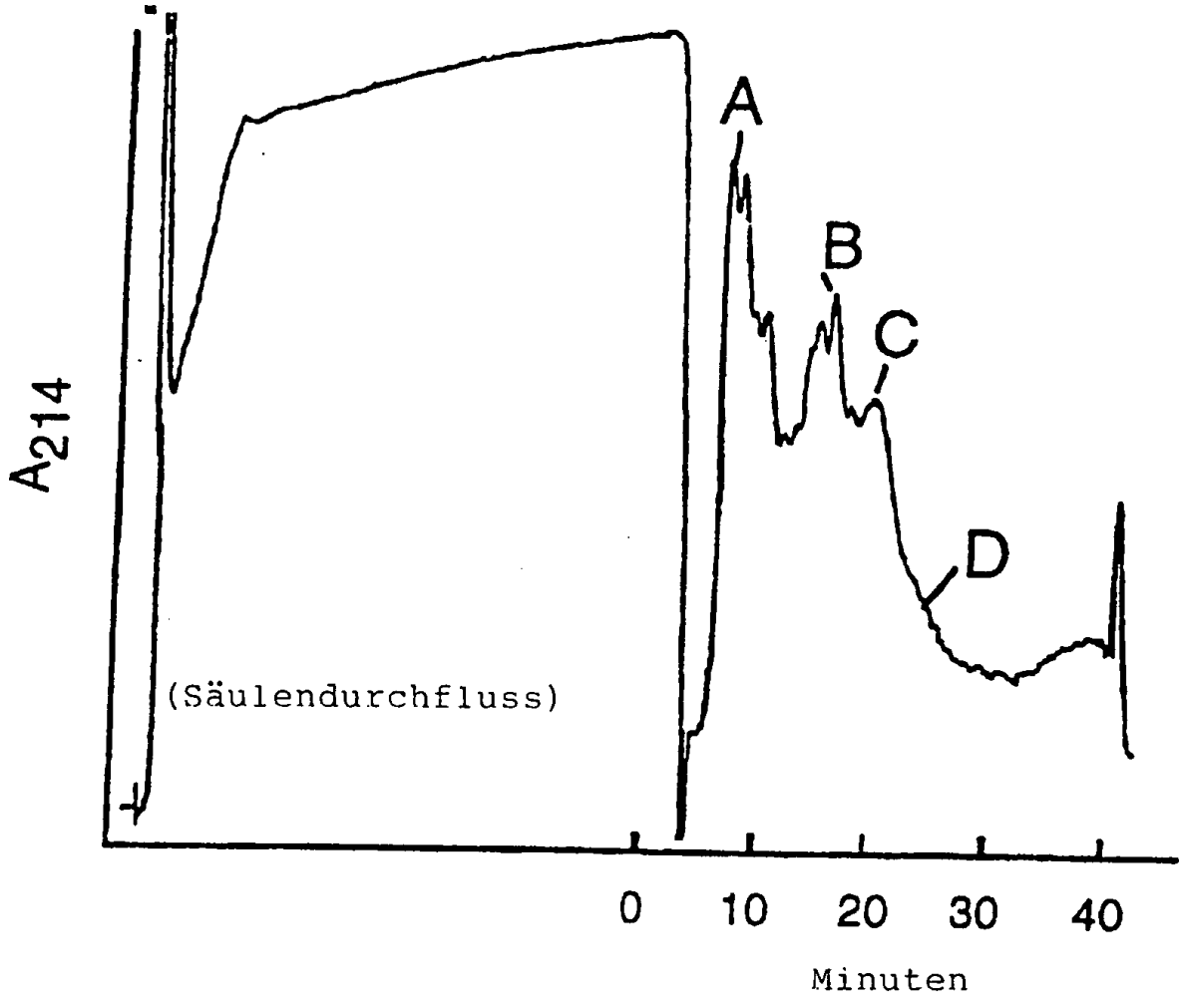


FIG. 1

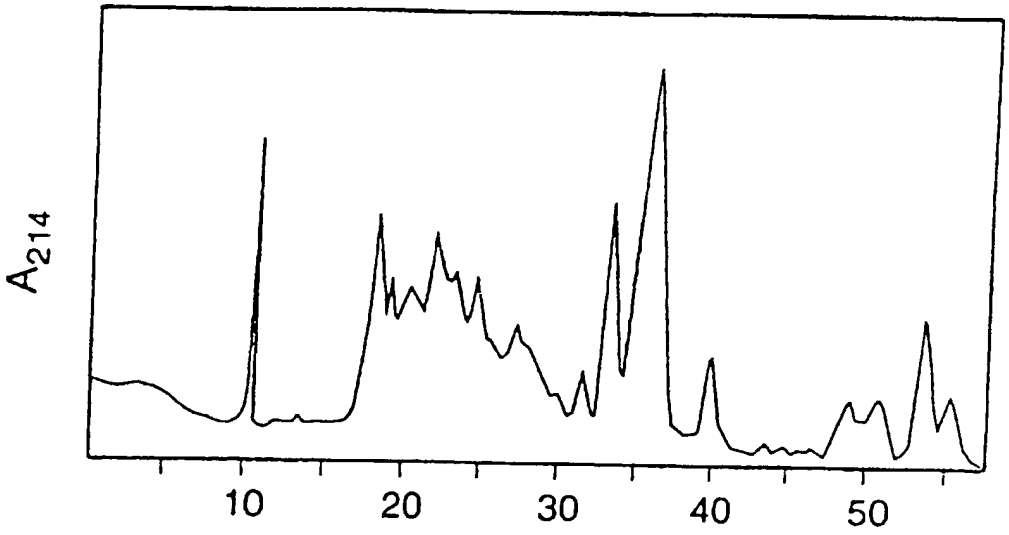


FIG. 2A

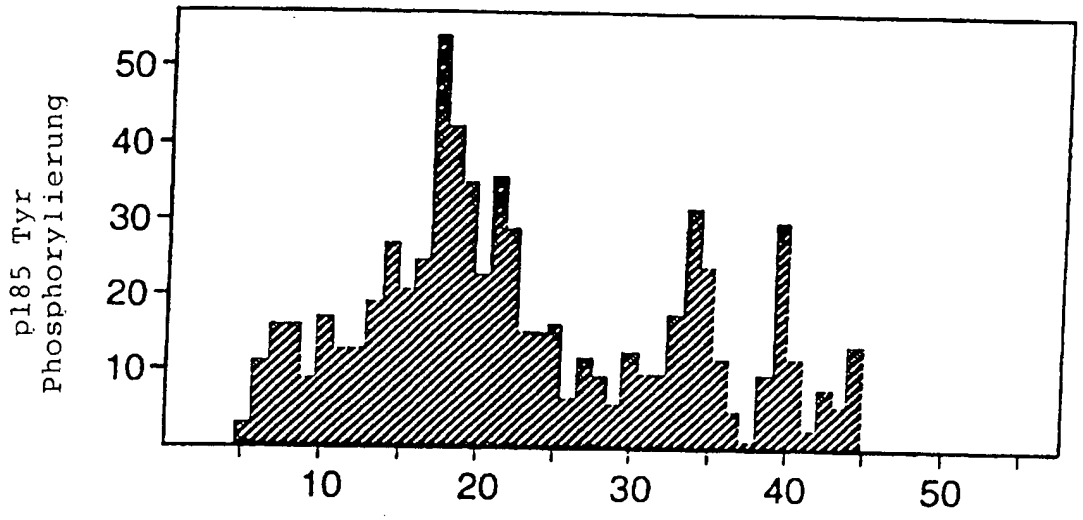


FIG. 2B

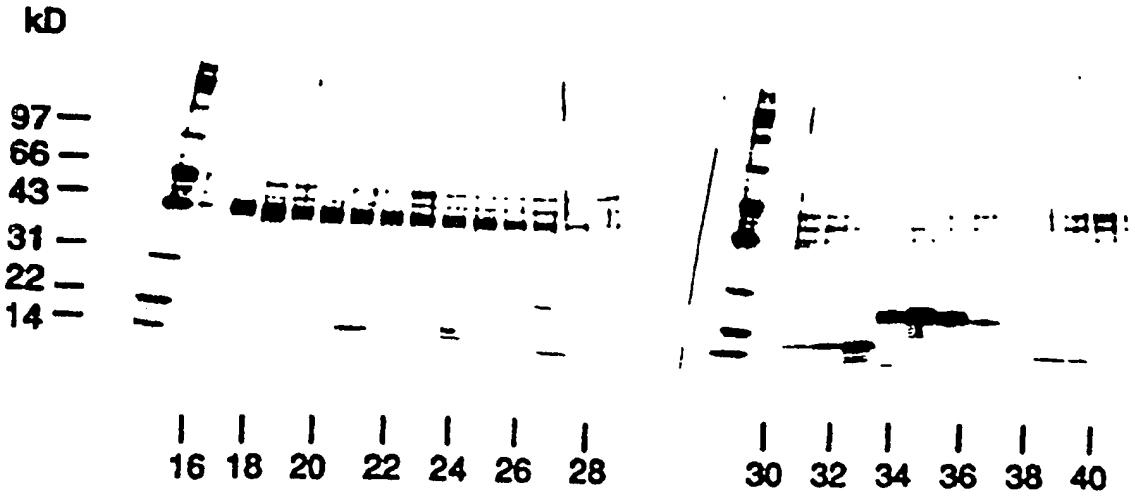


FIG. 2C

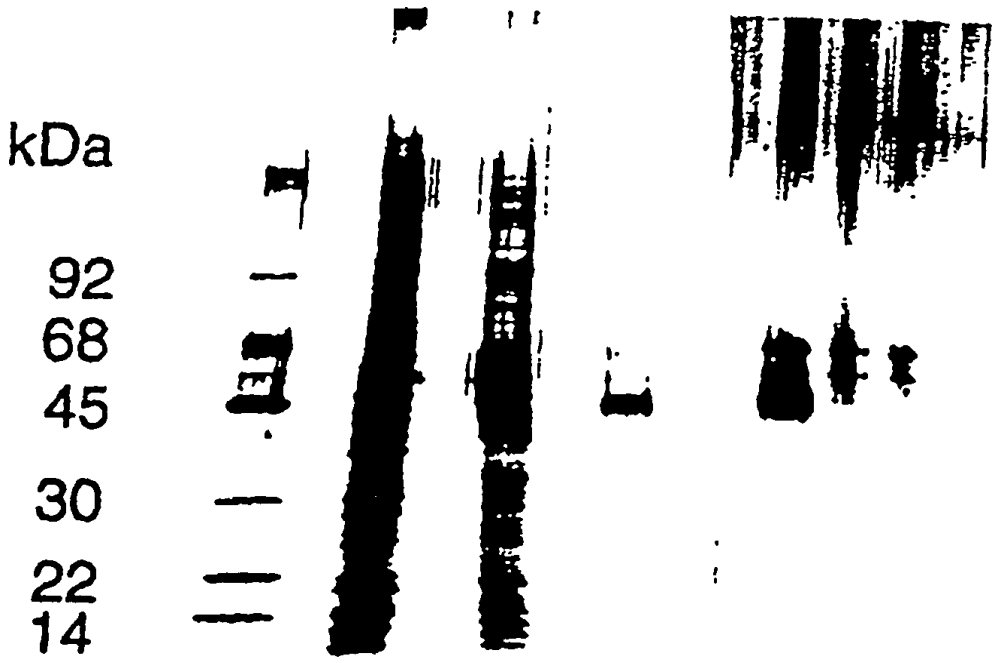


FIG. 3

```
GG  GCG CGA GCG CCT CAG CGC GGC CGC TCG CTC TCC CCC 38
    Ala Arg Ala Pro Gln Arg Gly Arg Ser Leu Ser Pro
      1               5                  10

TCG AGG GAC AAA CTT TTC CCA AAC CCG ATC CGA GCC CTT 77
Ser Arg Asp Lys Leu Phe Pro Asn Pro Ile Arg Ala Leu
         15                  20                  25

GGA CCA AAC TCG CCT GCG CCG AGA GCC GTC CGC GTA GAG 116
Gly Pro Asn Ser Pro Ala Pro Arg Ala Val Arg Val Glu
                30                  35

CGC TCC GTC TCC GGC GAG ATG TCC GAG CGC AAA GAA GGC 155
Arg Ser Val Ser Gly Glu Met Ser Glu Arg Lys Glu Gly
     40                  45                  50

AGA GGC AAA GGG AAG GGC AAG AAG AAG GAG CGA GGC TCC 194
Arg Gly Lys Gly Lys Gly Lys Lys Lys Glu Arg Gly Ser
            55                  60

GGC AAG AAG CCG GAG TCC GCG GCG GGC AGC CAG AGC CCA 233
Gly Lys Lys Pro Glu Ser Ala Ala Gly Ser Gln Ser Pro
 65                  70                  75

GCC TTG CCT CCC CGA TTG AAA GAG ATG AAA AGC CAG GAA 272
Ala Leu Pro Pro Arg Leu Lys Glu Met Lys Ser Gln Glu
         80                  85                  90

TCG GCT GCA GGT TCC AAA CTA GTC CTT CGG TGT GAA ACC 311
Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr
                95                 100

AGT TCT GAA TAC TCC TCT CTC AGA TTC AAG TGG TTC AAG 350
Ser Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys
    105                 110                 115

AAT GGG AAT GAA TTG AAT CGA AAA AAC AAA CCA CAA AAT 389
Asn Gly Asn Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn
            120                 125

ATC AAG ATA CAA AAA AAG CCA GGG AAG TCA GAA CTT CGC 428
Ile Lys Ile Gln Lys Lys Pro Gly Lys Ser Glu Leu Arg
130                 135                 140

ATT AAC AAA GCA TCA CTG GCT GAT TCT GGA GAG TAT ATG 467
Ile Asn Lys Ala Ser Leu Ala Asp Ser Gly Glu Tyr Met
        145                 150                 155

TGC AAA GTG ATC AGC AAA TTA GGA AAT GAC AGT GCC TCT 506
Cys Lys Val Ile Ser Lys Leu Gly Asn Asp Ser Ala Ser
            160                 165
```

FIG. 4A

```
GCC AAT ATC ACC ATC GTG GAA TCA AAC GAG ATC ATC ACT 545
Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile Ile Thr
    170                 175                 180

GGT ATG CCA GCC TCA ACT GAA GGA GCA TAT GTG TCT TCA 584
Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser
            185                 190

GAG TCT CCC ATT AGA ATA TCA GTA TCC ACA GAA GGA GCA 623
Glu Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala
195                 200                 205

AAT ACT TCT TCA TCT ACA TCT ACA TCC ACC ACT GGG ACA 662
Asn Thr Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr
        210                 215                 220

AGC CAT CTT GTA AAA TGT GCG GAG AAG GAG AAA ACT TTC 701
Ser His Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe
                225                 230

TGT GTG AAT GGA GGG GAG TGC TTC ATG GTG AAA GAC CTT 740
Cys Val Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu
    235                 240                 245

TCA AAC CCC TCG AGA TAC TTG TGC AAG TGC CAA CCT GGA 779
Ser Asn Pro Ser Arg Tyr Leu Cys Lys Cys Gln Pro Gly
            250                 255

TTC ACT GGA GCA AGA TGT ACT GAG AAT GTG CCC ATG AAA 818
Phe Thr Gly Ala Arg Cys Thr Glu Asn Val Pro Met Lys
260                 265                 270

GTC CAA AAC CAA GAA AAG GCG GAG GAG CTG TAC CAG AAG 857
Val Gln Asn Gln Glu Lys Ala Glu Glu Leu Tyr Gln Lys
        275                 280                 285

AGA GTG CTG ACC ATA ACC GGC ATC TGC ATC GCC CTC CTT 896
Arg Val Leu Thr Ile Thr Gly Ile Cys Ile Ala Leu Leu
                290                 295

GTG GTC GGC ATC ATG TGT GTG GTG GCC TAC TGC AAA ACC 935
Val Val Gly Ile Met Cys Val Val Ala Tyr Cys Lys Thr
    300                 305                 310

AAG AAA CAG CGG AAA AAG CTG CAT GAC CGT CTT CGG CAG 974
Lys Lys Gln Arg Lys Lys Leu His Asp Arg Leu Arg Gln
            315                 320

AGC CTT CGG TCT GAA CGA AAC AAT ATG ATG AAC ATT GCC 1013
Ser Leu Arg Ser Glu Arg Asn Asn Met Met Asn Ile Ala
325                 330                 335
```

FIG. 4B

```
AAT GGG CCT CAC CAT CCT AAC CCA CCC CCC GAG AAT GTC 1052
Asn Gly Pro His His Pro Asn Pro Pro Pro Glu Asn Val
        340                 345                 350

CAG CTG GTG AAT CAA TAC GTA TCT AAA AAC GTC ATC TCC 1091
Gln Leu Val Asn Gln Tyr Val Ser Lys Asn Val Ile Ser
                355                 360

AGT GAG CAT ATT GTT GAG AGA GAA GCA GAG ACA TCC TTT 1130
Ser Glu His Ile Val Glu Arg Glu Ala Glu Thr Ser Phe
    365                 370                 375

TCC ACC AGT CAC TAT ACT TCC ACA GCC CAT CAC TCC ACT 1169
Ser Thr Ser His Tyr Thr Ser Thr Ala His His Ser Thr
            380                 385

ACT GTC ACC CAG ACT CCT AGC CAC AGC TGG AGC AAC GGA 1208
Thr Val Thr Gln Thr Pro Ser His Ser Trp Ser Asn Gly
390                 395                 400

CAC ACT GAA AGC ATC CTT TCC GAA AGC CAC TCT GTA ATC 1247
His Thr Glu Ser Ile Leu Ser Glu Ser His Ser Val Ile
        405                 410                 415

GTG ATG TCA TCC GTA GAA AAC AGT AGG CAC AGC AGC CCA 1286
Val Met Ser Ser Val Glu Asn Ser Arg His Ser Ser Pro
                420                 425

ACT GGG GGC CCA AGA GGA CGT CTT AAT GGC ACA GGA GGC 1325
Thr Gly Gly Pro Arg Gly Arg Leu Asn Gly Thr Gly Gly
    430                 435                 440

CCT CGT GAA TGT AAC AGC TTC CTC AGG CAT GCC AGA GAA 1364
Pro Arg Glu Cys Asn Ser Phe Leu Arg His Ala Arg Glu
            445                 450

ACC CCT GAT TCC TAC CGA GAC TCT CCT CAT AGT GAA AGG 1403
Thr Pro Asp Ser Tyr Arg Asp Ser Pro His Ser Glu Arg
455                 460                 465

TAT GTG TCA GCC ATG ACC ACC CCG GCT CGT ATG TCA CCT 1442
Tyr Val Ser Ala Met Thr Thr Pro Ala Arg Met Ser Pro
        470                 475                 480

GTA GAT TTC CAC ACG CCA AGC TCC CCC AAA TCG CCC CCT 1481
Val Asp Phe His Thr Pro Ser Ser Pro Lys Ser Pro Pro
                485                 490

TCG GAA ATG TCT CCA CCC GTG TCC AGC ATG ACG GTG TCC 1520
Ser Glu Met Ser Pro Pro Val Ser Ser Met Thr Val Ser
    495                 500                 505
```

FIG. 4C

```
ATG CCT TCC ATG GCG GTC AGC CCC TTC ATG GAA GAA GAG 1559
Met Pro Ser Met Ala Val Ser Pro Phe Met Glu Glu Glu
            510                 515

AGA CCT CTA CTT CTC GTG ACA CCA CCA AGG CTG CGG GAG 1598
Arg Pro Leu Leu Leu Val Thr Pro Pro Arg Leu Arg Glu
520                 525                 530

AAG AAG TTT GAC CAT CAC CCT CAG CAG TTC AGC TCC TTC 1637
Lys Lys Phe Asp His His Pro Gln Gln Phe Ser Ser Phe
        535                 540                 545

CAC CAC AAC CCC GCG CAT GAC AGT AAC AGC CTC CCT GCT 1676
His His Asn Pro Ala His Asp Ser Asn Ser Leu Pro Ala
                550                 555

AGC CCC TTG AGG ATA GTG GAG GAT GAG GAG TAT GAA ACG 1715
Ser Pro Leu Arg Ile Val Glu Asp Glu Glu Tyr Glu Thr
    560                 565                 570

ACC CAA GAG TAC GAG CCA GCC CAA GAG CCT GTT AAG AAA 1754
Thr Gln Glu Tyr Glu Pro Ala Gln Glu Pro Val Lys Lys
            575                 580

CTC GCC AAT AGC CGG CGG GCC AAA AGA ACC AAG CCC AAT 1793
Leu Ala Asn Ser Arg Arg Ala Lys Arg Thr Lys Pro Asn
585                 590                 595

GGC CAC ATT GCT AAC AGA TTG GAA GTG GAC AGC AAC ACA 1832
Gly His Ile Ala Asn Arg Leu Glu Val Asp Ser Asn Thr
        600                 605                 610

AGC TCC CAG AGC AGT AAC TCA GAG AGT GAA ACA GAA GAT 1871
Ser Ser Gln Ser Ser Asn Ser Glu Ser Glu Thr Glu Asp
                615                 620

GAA AGA GTA GGT GAA GAT ACG CCT TTC CTG GGC ATA CAG 1910
Glu Arg Val Gly Glu Asp Thr Pro Phe Leu Gly Ile Gln
    625                 630                 635

AAC CCC CTG GCA GCC AGT CTT GAG GCA ACA CCT GCC TTC 1949
Asn Pro Leu Ala Ala Ser Leu Glu Ala Thr Pro Ala Phe
            640                 645

CGC CTG GCT GAC AGC AGG ACT AAC CCA GCA GGC CGC TTC 1988
Arg Leu Ala Asp Ser Arg Thr Asn Pro Ala Gly Arg Phe
650                 655                 660

TCG ACA CAG GAA GAA ATC CAG G 2010
Ser Thr Gln Glu Glu Ile Gln
        665                 669
```

FIG. 4D

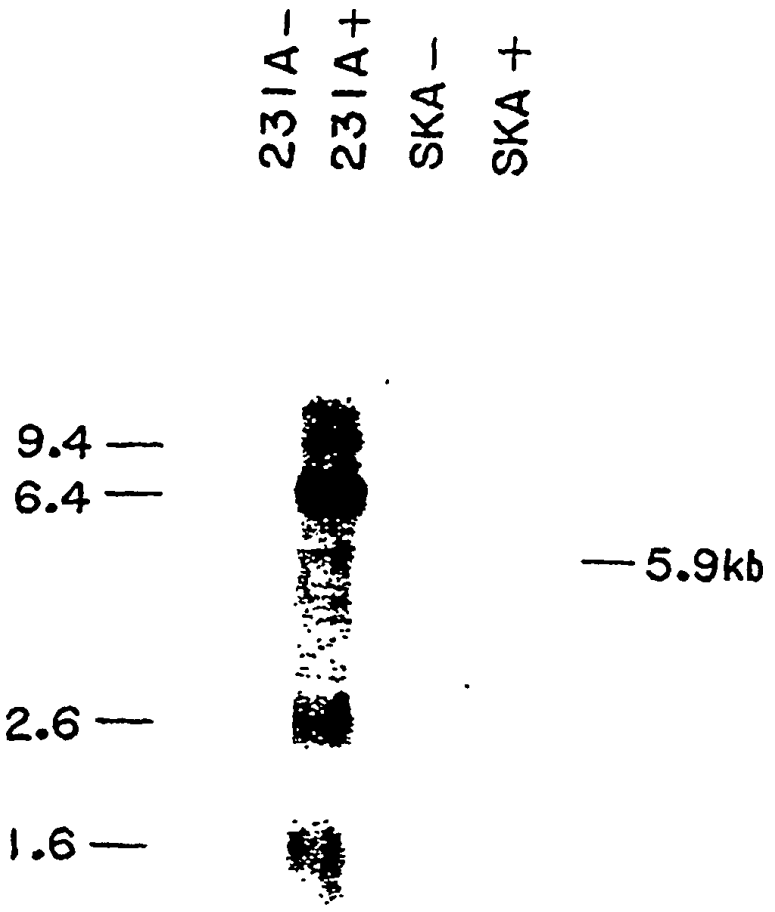


FIG. 5

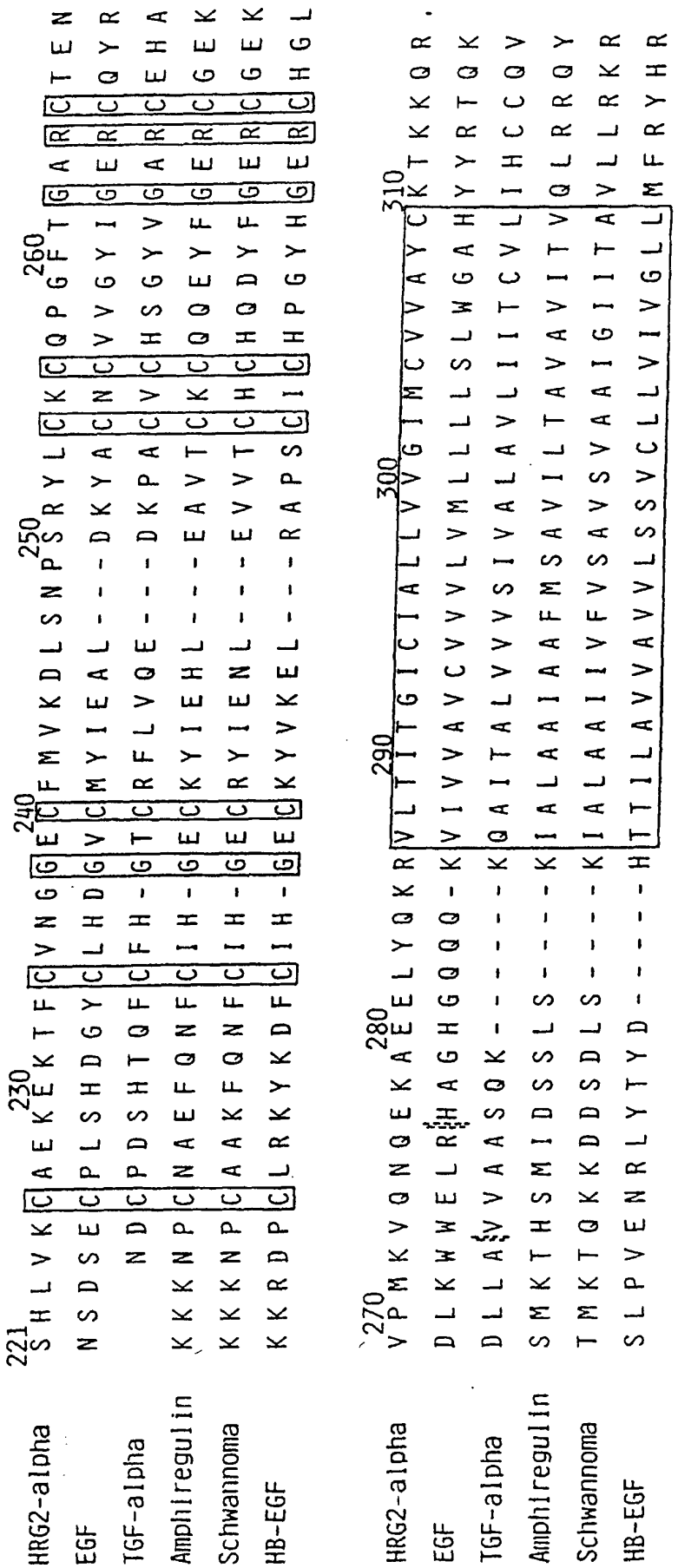

HRG2-alpha
EGF
TGF-alpha
Amphiregulin
Schwannoma
HB-EGF
221
230
240
250
260
270
280
290
300
310

FIG. 6

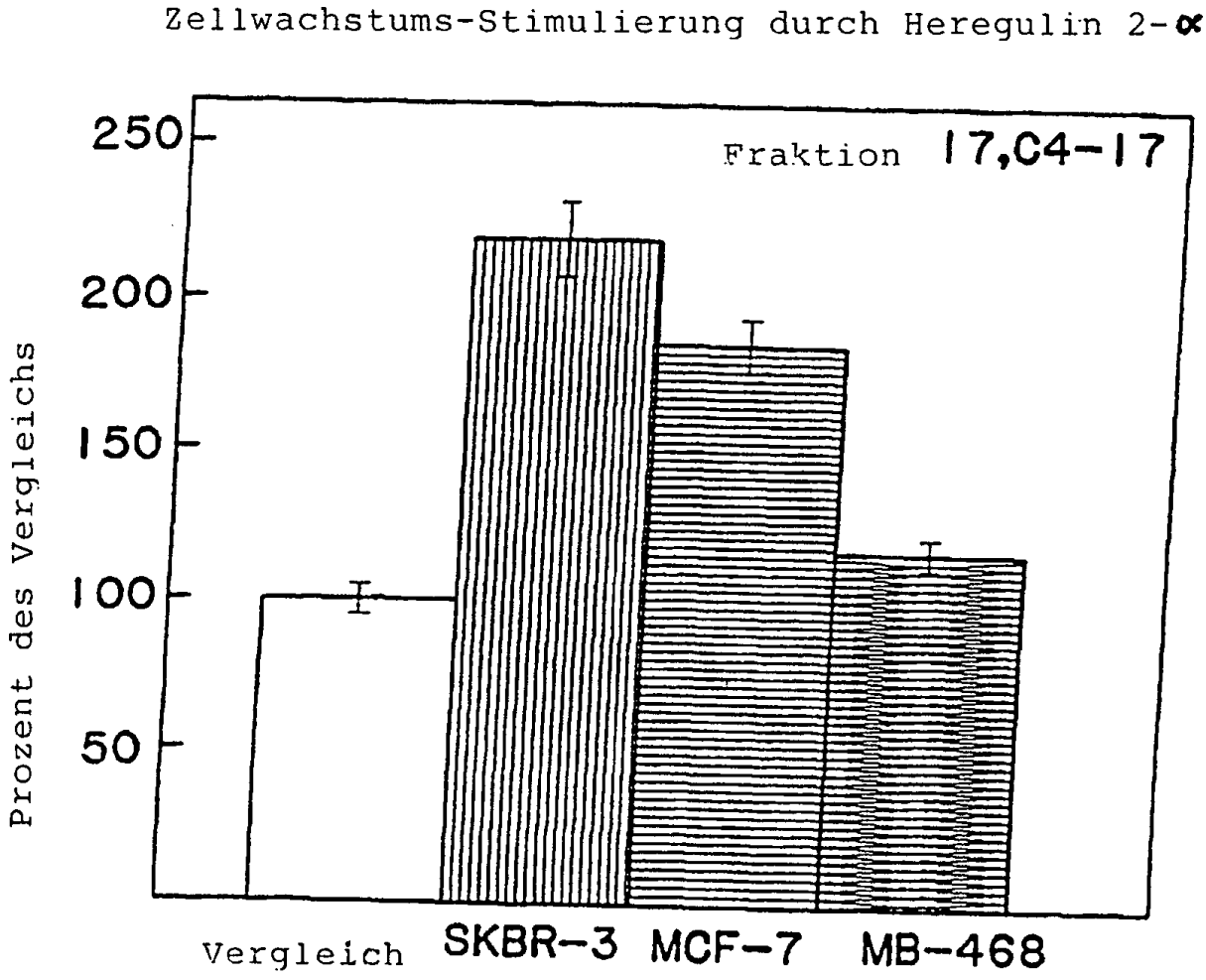

Zellwachstums-Stimulierung durch Heregulin 2-α
Fraktion 17,C4-17
Prozent des Vergleichs
250
200
150
100
50
Vergleich
SKBR-3
MCF-7
MB-468

FIG. 7

```
GG  GAC AAA CTT TTC CCA AAC CCG ATC CGA GCC CTT GGA 38
    Asp Lys Leu Phe Pro Asn Pro Ile Arg Ala Leu Gly
      1               5                  10

CCA AAC TCG CCT GCG CCG AGA GCC GTC CGC GTA GAG CGC 77
Pro Asn Ser Pro Ala Pro Arg Ala Val Arg Val Glu Arg
         15                  20                  25

TCC GTC TCC GGC GAG ATG TCC GAG CGC AAA GAA GGC AGA 116
Ser Val Ser Gly Glu Met Ser Glu Arg Lys Glu Gly Arg
                30                  35

GGC AAA GGG AAG GGC AAG AAG AAG GAG CGA GGC TCC GGC 155
Gly Lys Gly Lys Gly Lys Lys Lys Glu Arg Gly Ser Gly
     40                  45                  50

AAG AAG CCG GAG TCC GCG GCG GGC AGC CAG AGC CCA GCC 194
Lys Lys Pro Glu Ser Ala Ala Gly Ser Gln Ser Pro Ala
            55                  60

TTG CCT CCC CAA TTG AAA GAG ATG AAA AGC CAG GAA TCG 233
Leu Pro Pro Gln Leu Lys Glu Met Lys Ser Gln Glu Ser
 65                  70                  75

GCT GCA GGT TCC AAA CTA GTC CTT CGG TGT GAA ACC AGT 272
Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser
         80                  85                  90

TCT GAA TAC TCC TCT CTC AGA TTC AAG TGG TTC AAG AAT 311
Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn
                95                  100

GGG AAT GAA TTG AAT CGA AAA AAC AAA CCA CAA AAT ATC 350
Gly Asn Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile
    105                 110                 115

AAG ATA CAA AAA AAG CCA GGG AAG TCA GAA CTT CGC ATT 389
Lys Ile Gln Lys Lys Pro Gly Lys Ser Glu Leu Arg Ile
            120                 125

AAC AAA GCA TCA CTG GCT GAT TCT GGA GAG TAT ATG TGC 428
Asn Lys Ala Ser Leu Ala Asp Ser Gly Glu Tyr Met Cys
130                 135                 140

AAA GTG ATC AGC AAA TTA GGA AAT GAC AGT GCC TCT GCC 467
Lys Val Ile Ser Lys Leu Gly Asn Asp Ser Ala Ser Ala
        145                 150                 155

AAT ATC ACC ATC GTG GAA TCA AAC GAG ATC ATC ACT GGT 506
Asn Ile Thr Ile Val Glu Ser Asn Glu Ile Ile Thr Gly
                160                 165
```

FIG. 8A

```
ATG CCA GCC TCA ACT GAA GGA GCA TAT GTG TCT TCA GAG 545
Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu
    170                 175                 180

TCT CCC ATT AGA ATA TCA GTA TCC ACA GAA GGA GCA AAT 584
Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn
            185                 190

ACT TCT TCA TCT ACA TCT ACA TCC ACC ACT GGG ACA AGC 623
Thr Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser
195                 200                 205

CAT CTT GTA AAA TGT GCG GAG AAG GAG AAA ACT TTC TGT 662
His Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys
        210                 215                 220

GTG AAT GGA GGG GAG TGC TTC ATG GTG AAA GAC CTT TCA 701
Val Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser
                225                 230

AAC CCC TCG AGA TAC TTG TGC AAG TGC CCA AAT GAG TTT 740
Asn Pro Ser Arg Tyr Leu Cys Lys Cys Pro Asn Glu Phe
    235                 240                 245

ACT GGT GAT CGC TGC CAA AAC TAC GTA ATG GCC AGC TTC 779
Thr Gly Asp Arg Cys Gln Asn Tyr Val Met Ala Ser Phe
            250                 255

TAC AAG CAT CTT GGG ATT GAA TTT ATG GAG GCG GAG GAG 818
Tyr Lys His Leu Gly Ile Glu Phe Met Glu Ala Glu Glu
260                 265                 270

CTG TAC CAG AAG AGA GTG CTG ACC ATA ACC GGC ATC TGC 857
Leu Tyr Gln Lys Arg Val Leu Thr Ile Thr Gly Ile Cys
        275                 280                 285

ATC GCC CTC CTT GTG GTC GGC ATC ATG TGT GTG GTG GCC 896
Ile Ala Leu Leu Val Val Gly Ile Met Cys Val Val Ala
                290                 295

TAC TGC AAA ACC AAG AAA CAG CGG AAA AAG CTG CAT GAC 935
Tyr Cys Lys Thr Lys Lys Gln Arg Lys Lys Leu His Asp
    300                 305                 310

CGT CTT CGG CAG AGC CTT CGG TCT GAA CGA AAC AAT ATG 974
Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg Asn Asn Met
            315                 320

ATG AAC ATT GCC AAT GGG CCT CAC CAT CCT AAC CCA CCC 1013
Met Asn Ile Ala Asn Gly Pro His His Pro Asn Pro Pro
-325                330                 335
```

FIG. 8B

```
CCC GAG AAT GTC CAG CTG GTG AAT CAA TAC GTA TCT AAA 1052
Pro Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys
        340                 345                 350

AAC GTC ATC TCC AGT GAG CAT ATT GTT GAG AGA GAA GCA 1091
Asn Val Ile Ser Ser Glu His Ile Val Glu Arg Glu Ala
                355                 360

GAG ACA TCC TTT TCC ACC AGT CAC TAT ACT TCC ACA GCC 1130
Glu Thr Ser Phe Ser Thr Ser His Tyr Thr Ser Thr Ala
    365                 370                 375

CAT CAC TCC ACT ACT GTC ACC CAG ACT CCT AGC CAC AGC 1169
His His Ser Thr Thr Val Thr Gln Thr Pro Ser His Ser
            380                 385

TGG AGC AAC GGA CAC ACT GAA AGC ATC CTT TCC GAA AGC 1208
Trp Ser Asn Gly His Thr Glu Ser Ile Leu Ser Glu Ser
390                 395                 400

CAC TCT GTA ATC GTG ATG TCA TCC GTA GAA AAC AGT AGG 1247
His Ser Val Ile Val Met Ser Ser Val Glu Asn Ser Arg
        405                 410                 415

CAC AGC AGC CCA ACT GGG GGC CCA AGA GGA CGT CTT AAT 1286
His Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg Leu Asn
                420                 425

GGC ACA GGA GGC CCT CGT GAA TGT AAC AGC TTC CTC AGG 1325
Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg
    430                 435                 440

CAT GCC AGA GAA ACC CCT GAT TCC TAC CGA GAC TCT CCT 1364
His Ala Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro
            445                 450

CAT AGT GAA AGG TAT GTG TCA GCC ATG ACC ACC CCG GCT 1403
His Ser Glu Arg Tyr Val Ser Ala Met Thr Thr Pro Ala
455                 460                 465

CGT ATG TCA CCT GTA GAT TTC CAC ACG CCA AGC TCC CCC 1442
Arg Met Ser Pro Val Asp Phe His Thr Pro Ser Ser Pro
        470                 475                 480

AAA TCG CCC CCT TCG GAA ATG TCT CCA CCC GTG TCC AGC 1481
Lys Ser Pro Pro Ser Glu Met Ser Pro Pro Val Ser Ser
                485                 490

ATG ACG GTG TCC ATG CCT TCC ATG GCG GTC AGC CCC TTC 1520
Met Thr Val Ser Met Pro Ser Met Ala Val Ser Pro Phe
    495                 500                 505
```

→ FIG. 8D

FIG. 8C

```
ATG GAA GAA GAG AGA CCT CTA CTT CTC GTG ACA CCA CCA 1559
Met Glu Glu Glu Arg Pro Leu Leu Leu Val Thr Pro Pro
               510                 515

AGG CTG CGG GAG AAG AAG TTT GAC CAT CAC CCT CAG CAG 1598
Arg Leu Arg Glu Lys Lys Phe Asp His His Pro Gln Gln
520                 525                 530

TTC AGC TCC TTC CAC CAC AAC CCC GCG CAT GAC AGT AAC 1637
Phe Ser Ser Phe His His Asn Pro Ala His Asp Ser Asn
        535                 540                 545

AGC CTC CCT GCT AGC CCC TTG AGG ATA GTG GAG GAT GAG 1676
Ser Leu Pro Ala Ser Pro Leu Arg Ile Val Glu Asp Glu
                550                 555

GAG TAT GAA ACG ACC CAA GAG TAC GAG CCA GCC CAA GAG 1715
Glu Tyr Glu Thr Thr Gln Glu Tyr Glu Pro Ala Gln Glu
    560                 565                 570

CCT GTT AAG AAA CTC GCC AAT AGC CGG CGG GCC AAA AGA 1754
Pro Val Lys Lys Leu Ala Asn Ser Arg Arg Ala Lys Arg
            575                 580

ACC AAG CCC AAT GGC CAC ATT GCT AAC AGA TTG GAA GTG 1793
Thr Lys Pro Asn Gly His Ile Ala Asn Arg Leu Glu Val
585                 590                 595

GAC AGC AAC ACA AGC TCC CAG AGC AGT AAC TCA GAG AGT 1832
Asp Ser Asn Thr Ser Ser Gln Ser Ser Asn Ser Glu Ser
        600                 605                 610
```

FIG. 8D

DE 692 33 137 T2 2004.04.29

```
GAA ACA GAA GAT GAA AGA GTA GGT GAA GAT ACG CCT TTC 1871
Glu Thr Glu Asp Glu Arg Val Gly Glu Asp Thr Pro Phe
                615                 620

CTG GGC ATA CAG AAC CCC CTG GCA GCC AGT CTT GAG GCA 1910
Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser Leu Glu Ala
    625                 630                 635

ACA CCT GCC TTC CGC CTG GCT GAC AGC AGG ACT AAC CCA 1949
Thr Pro Ala Phe Arg Leu Ala Asp Ser Arg Thr Asn Pro
            640                 645

GCA GGC CGC TTC TCG ACA CAG GAA GAA ATC CAG GCC AGG 1988
Ala Gly Arg Phe Ser Thr Gln Glu Glu Ile Gln Ala Arg
650                 655                 660

CTG TCT AGT GTA ATT GCT AAC CAA GAC CCT ATT GCT GTA TA 2029
Leu Ser Ser Val Ile Ala Asn Gln Asp Pro Ile Ala Val
        665                 670                 675

A AACCTAAATA AACACATAGA TTCACCTGTA AAACTTTATT              2070

TTATATAATA AAGTATTCCA CCTTAAATTA AACAATTTAT TTTATTTTAG     2120

CAGTTCTGCA AATAGAAAAC AGGAAAAAAA CTTTTATAAA TTAAATATAT     2170

GTATGTAAAA ATGAAAAAAA AAAAAAAAA                            2199
```

FIG. 8D-1

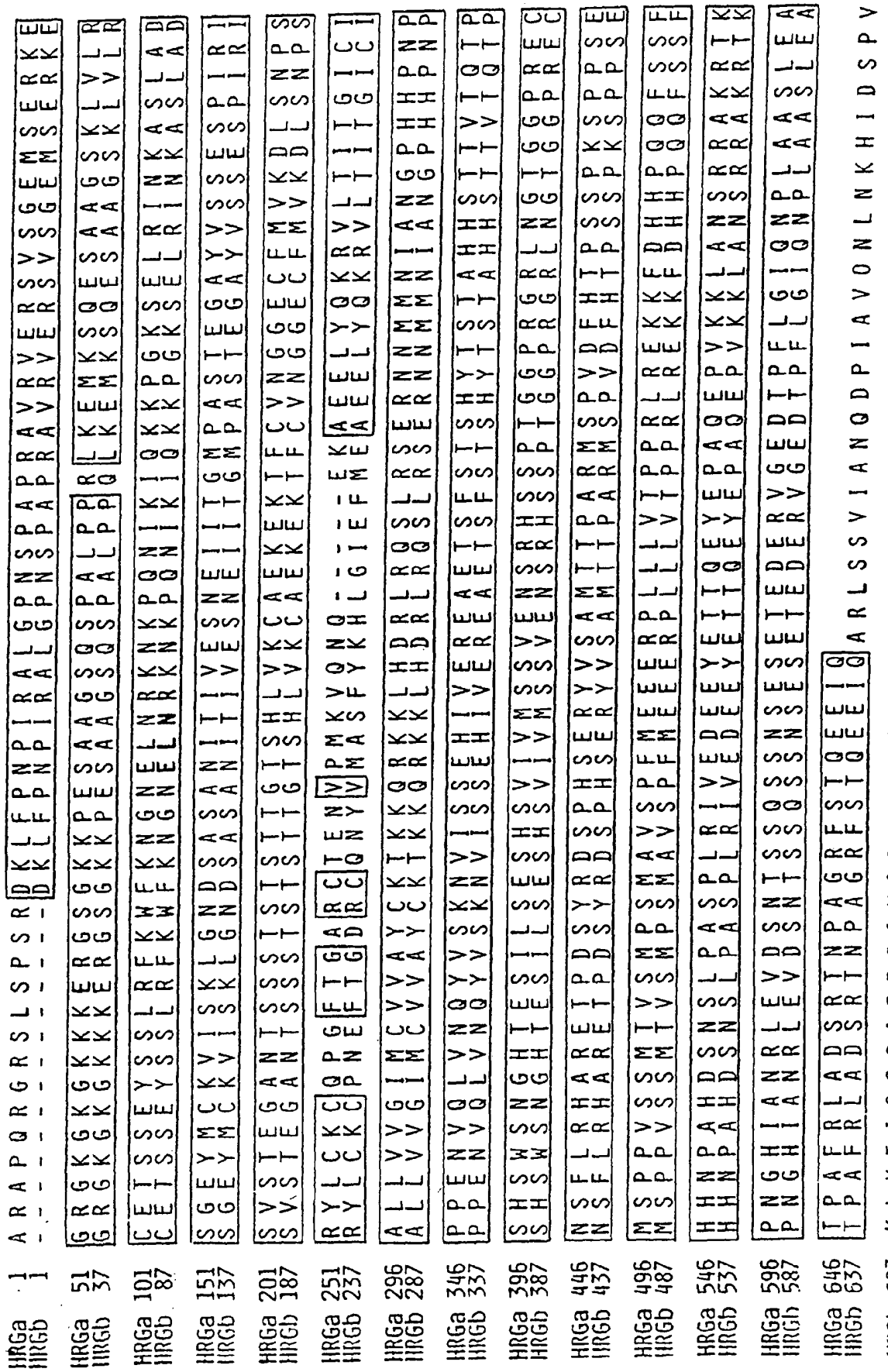


FIG. 9

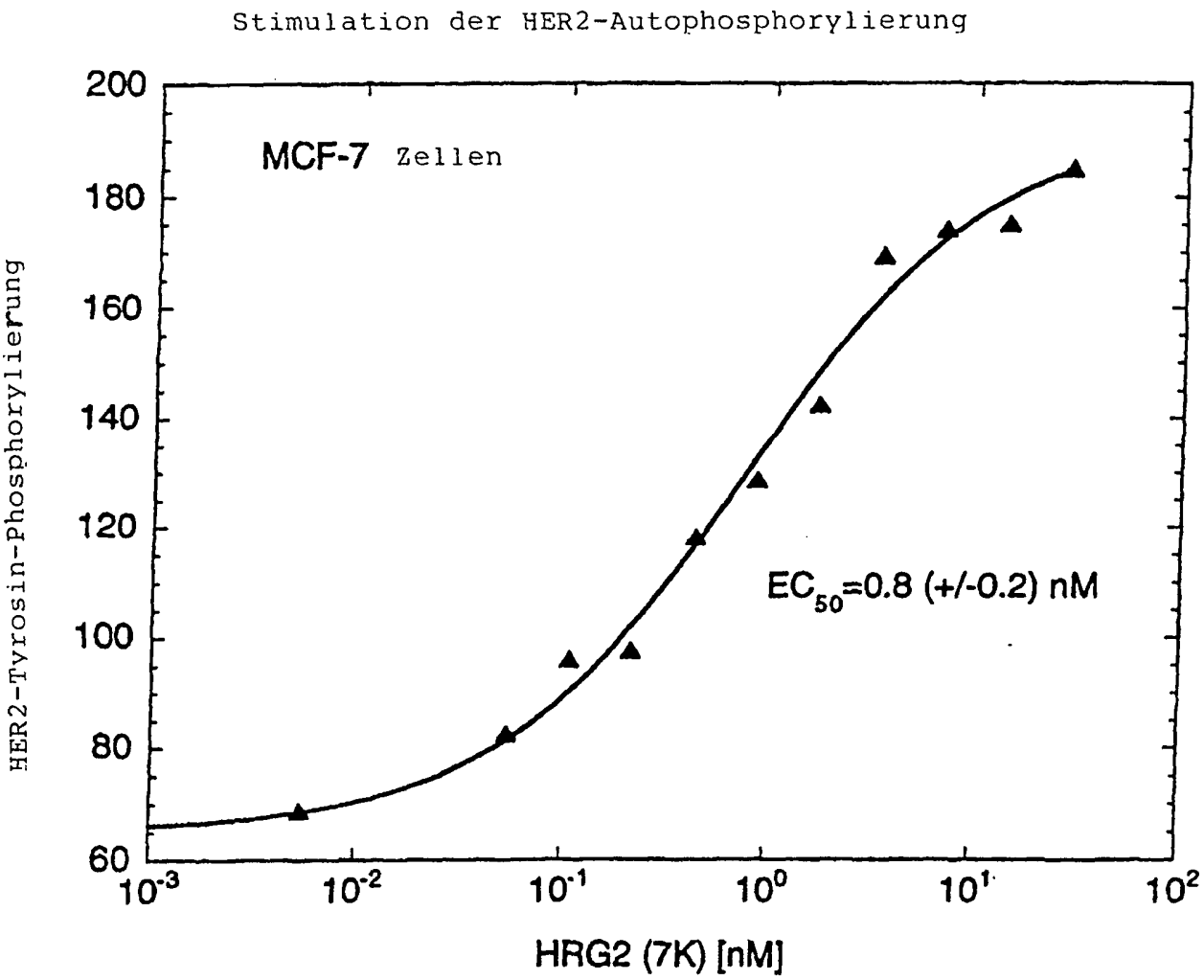

Stimulation der HER2-Autophosphorylierung
MCF-7 Zellen
HER2-Tyrosin-Phosphorylierung
200
180
160
140
120
100
80
60
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
EC_{50}=0.8 (+/-0.2) nM
HRG2 (7K) [nM]

FIG. 10

```
AA  AGA GCC GGC GAG GAG TTC CCC GAA ACT TGT TGG AAC 38
    Arg Ala Gly Glu Glu Phe Pro Glu Thr Cys Trp Asn
      1               5                  10

TCC GGG CTC GCG CGG AGG CCA GGA GCT GAG CGG CGG CGG 77
Ser Gly Leu Ala Arg Arg Pro Gly Ala Glu Arg Arg Arg
          15                  20                  25

CTG CCG GAC GAT GGG AGC GTG AGC AGG ACG GTG ATA ACC 116
Leu Pro Asp Asp Gly Ser Val Ser Arg Thr Val Ile Thr
                  30                  35

TCT CCC CGA TCG GGT TGC GAG GGC GCC GGG CAG AGG CCA 155
Ser Pro Arg Ser Gly Cys Glu Gly Ala Gly Gln Arg Pro
     40                  45                  50

GGA CGC GAG CCG CCA GCG GTG GGA CCC ATC GAC GAC TTC 194
Gly Arg Glu Pro Pro Ala Val Gly Pro Ile Asp Asp Phe
              55                  60

CCG GGG CGA CAG GAG CAG CCC CGA GAG CCA GGG CGA GCG 233
Pro Gly Arg Gln Glu Gln Pro Arg Glu Pro Gly Arg Ala
 65                  70                  75

CCC GTT CCA GGT GGC CGG ACC GCC CGC CGC GTC CGC GCC 272
Pro Val Pro Gly Gly Arg Thr Ala Arg Arg Val Arg Ala
          80                  85                  90

GCG CTC CCT GCA GGC AAC GGG AGA CGC CCC CGC GCA GCG 311
Ala Leu Pro Ala Gly Asn Gly Arg Arg Pro Arg Ala Ala
                  95                  100

CGA GCG CCT CAG CGC GGC CGC TCG CTC TCC CCC TCG AGG 350
Arg Ala Pro Gln Arg Gly Arg Ser Leu Ser Pro Ser Arg
    105                 110                 115

GAC AAA CTT TTC CCA AAC CCG ATC CGA GCC CTT GGA CCA 389
Asp Lys Leu Phe Pro Asn Pro Ile Arg Ala Leu Gly Pro
            120                 125

AAC TCG CCT GCG CCG AGA GCC GTC CGC GTA GAG CGC TCC 428
Asn Ser Pro Ala Pro Arg Ala Val Arg Val Glu Arg Ser
130                 135                 140

GTC TCC GGC GAG ATG TCC GAG CGC AAA GAA GGC AGA GGC 467
Val Ser Gly Glu Met Ser Glu Arg Lys Glu Gly Arg Gly
        145                 150                 155

AAA GGG AAG GGC AAG AAG AAG GAG CGA GG 496
Lys Gly Lys Gly Lys Lys Lys Glu Arg
                160             164
```

FIG. 11

```
GTGGCTGCGG GGCAATTGAA AAAGAGCCGG CGAGGAGTTC CCCGAAACTT 50

GTTGGAACTC CGGGCTCGCG CGGAGGCCAG GAGCTGAGCG GCGGCGGCTG 100

CCGGACGATG GGAGCGTGAG CAGGACGGTG ATAACCTCTC CCCGATCGGG 150

TTGCGAGGGC GCCGGGCAGA GGCCAGGACG CGAGCCGCCA GCGGCGGGAC 200

CCATCGACGA CTTCCCGGGG CGACAGGAGC AGCCCCGAGA GCCAGGGCGA 250

GCGCCCGTTC CAGGTGGCCG GACCGCCCGC CGCGTCCGCG CCGCGCTCCC 300

TGCAGGCAAC GGGAGACGCC CCCGCGCAGC GCGAGCGCCT CAGCGCGGCC 350

GCTCGCTCTC CCCATCGAGG GACAAACTTT TCCCAAACCC GATCCGAGCC 400

CTTGGACCAA ACTCGCCTGC GCCGAGAGCC GTCCGCGTAG AGCGCTCCGT 450

CTCCGGCGAG  ATG TCC GAG CGC AAA GAA GGC AGA GGC AAA 490
            Met Ser Glu Arg Lys Glu Gly Arg Gly Lys
              1               5                  10

GGG AAG GGC AAG AAG AAG GAG CGA GGC TCC GGC AAG AAG 529
Gly Lys Gly Lys Lys Lys Glu Arg Gly Ser Gly Lys Lys
                  15                  20

CCG GAG TCC GCG GCG GGC AGC CAG AGC CCA GCC TTG CCT 568
Pro Glu Ser Ala Ala Gly Ser Gln Ser Pro Ala Leu Pro
     25                  30                  35

CCC CAA TTG AAA GAG ATG AAA AGC CAG GAA TCG GCT GCA 607
Pro Gln Leu Lys Glu Met Lys Ser Gln Glu Ser Ala Ala
             40                  45

GGT TCC AAA CTA GTC CTT CGG TGT GAA ACC AGT TCT GAA 646
Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser Ser Glu
 50                  55                  60

TAC TCC TCT CTC AGA TTC AAG TGG TTC AAG AAT GGG AAT 685
Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn
         65                  70                  75

GAA TTG AAT CGA AAA AAC AAA CCA CAA AAT ATC AAG ATA 724
Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile
             80                  85
```

FIG. 12A

```
CAA AAA AAG CCA GGG AAG TCA GAA CTT CGC ATT AAC AAA 763
Gln Lys Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys
 90                  95                 100

GCA TCA CTG GCT GAT TCT GGA GAG TAT ATG TGC AAA GTG 802
Ala Ser Leu Ala Asp Ser Gly Glu Tyr Met Cys Lys Val
        105                 110

ATC AGC AAA TTA GGA AAT GAC AGT GCC TCT GCC AAT ATC 841
Ile Ser Lys Leu Gly Asn Asp Ser Ala Ser Ala Asn Ile
115                 120                 125

ACC ATC GTG GAA TCA AAC GAG ATC ATC ACT GGT ATG CCA 880
Thr Ile Val Glu Ser Asn Glu Ile Ile Thr Gly Met Pro
        130                 135                 140

GCC TCA ACT GAA GGA GCA TAT GTG TCT TCA GAG TCT CCC 919
Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu Ser Pro
                145                 150

ATT AGA ATA TCA GTA TCC ACA GAA GGA GCA AAT ACT TCT 958
Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr Ser
    155                 160                 165

TCA TCT ACA TCT ACA TCC ACC ACT GGG ACA AGC CAT CTT 997
Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu
            170                 175

GTA AAA TGT GCG GAG AAG GAG AAA ACT TTC TGT GTG AAT 1036
Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn
180                 185                 190

GGA GGG GAG TGC TTC ATG GTG AAA GAC CTT TCA AAC CCC 1075
Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser Asn Pro
        195                 200                 205

TCG AGA TAC TTG TGC AAG TGC CCA AAT GAG TTT ACT GGT 1114
Ser Arg Tyr Leu Cys Lys Cys Pro Asn Glu Phe Thr Gly
                210                 215

GAT CGC TGC CAA AAC TAC GTA ATG GCC AGC TTC TAC AAG 1153
Asp Arg Cys Gln Asn Tyr Val Met Ala Ser Phe Tyr Lys
    220                 225                 230

GCG GAG GAG CTG TAC CAG AAG AGA GTG CTG ACC ATA ACC 1192
Ala Glu Glu Leu Tyr Gln Lys Arg Val Leu Thr Ile Thr
            235                 240

GGC ATC TGC ATC GCC CTC CTT GTG GTC GGC ATC ATG TGT 1231
Gly Ile Cys Ile Ala Leu Leu Val Val Gly Ile Met Cys
245                 250                 255

GTG GTG GCC TAC TGC AAA ACC AAG AAA CAG CGG AAA AAG 1270
Val Val Ala Tyr Cys Lys Thr Lys Lys Gln Arg Lys Lys
260                 265                 270
```

FIG. 12B

```
CTG CAT GAC CGT CTT CGG CAG AGC CTT CGG TCT GAA CGA 1309
Leu His Asp Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg
                275                 280

AAC AAT ATG ATG AAC ATT GCC AAT GGG CCT CAC CAT CCT 1348
Asn Asn Met Met Asn Ile Ala Asn Gly Pro His His Pro
    285                 290                 295

AAC CCA CCC CCC GAG AAT GTC CAG CTG GTG AAT CAA TAC 1387
Asn Pro Pro Pro Glu Asn Val Gln Leu Val Asn Gln Tyr
            300                 305

GTA TCT AAA AAC GTC ATC TCC AGT GAG CAT ATT GTT GAG 1426
Val Ser Lys Asn Val Ile Ser Ser Glu His Ile Val Glu
310                 315                 320

AGA GAA GCA GAG ACA TCC TTT TCC ACC AGT CAC TAT ACT 1465
Arg Glu Ala Glu Thr Ser Phe Ser Thr Ser His Tyr Thr
        325                 330                 335

TCC ACA GCC CAT CAC TCC ACT ACT GTC ACC CAG ACT CCT 1504
Ser Thr Ala His His Ser Thr Thr Val Thr Gln Thr Pro
                340                 345

AGC CAC AGC TGG AGC AAC GGA CAC ACT GAA AGC ATC CTT 1543
Ser His Ser Trp Ser Asn Gly His Thr Glu Ser Ile Leu
    350                 355                 360

TCC GAA AGC CAC TCT GTA ATC GTG ATG TCA TCC GTA GAA 1582
Ser Glu Ser His Ser Val Ile Val Met Ser Ser Val Glu
            365                 370

AAC AGT AGG CAC AGC AGC CCA ACT GGG GGC CCA AGA GGA 1621
Asn Ser Arg His Ser Ser Pro Thr Gly Gly Pro Arg Gly
375                 380                 385

CGT CTT AAT GGC ACA GGA GGC CCT CGT GAA TGT AAC AGC 1660
Arg Leu Asn Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser
        390                 395                 400

TTC CTC AGG CAT GCC AGA GAA ACC CCT GAT TCC TAC CGA 1699
Phe Leu Arg His Ala Arg Glu Thr Pro Asp Ser Tyr Arg
                405                 410

GAC TCT CCT CAT AGT GAA AGG TAT GTG TCA GCC ATG ACC 1738
Asp Ser Pro His Ser Glu Arg Tyr Val Ser Ala Met Thr
    415                 420                 425

ACC CCG GCT CGT ATG TCA CCT GTA GAT TTC CAC ACG CCA 1777
Thr Pro Ala Arg Met Ser Pro Val Asp Phe His Thr Pro
            430                 435

AGC TCC CCC AAA TCG CCC CCT TCG GAA ATG TCT CCA CCC 1816
Ser Ser Pro Lys Ser Pro Pro Ser Glu Met Ser Pro Pro
440                 445                 450
```

FIG. 12C

```
GTG TCC AGC ATG ACG GTG TCC AAG CCT TCC ATG GCG GTC 1855
Val Ser Ser Met Thr Val Ser Lys Pro Ser Met Ala Val
        455                 460                 465

AGC CCC TTC ATG GAA GAA GAG AGA CCT CTA CTT CTC GTG 1894
Ser Pro Phe Met Glu Glu Glu Arg Pro Leu Leu Leu Val
                470                 475

ACA CCA CCA AGG CTG CGG GAG AAG AAG TTT GAC CAT CAC 1933
Thr Pro Pro Arg Leu Arg Glu Lys Lys Phe Asp His His
    480                 485                 490

CCT CAG CAG TTC AGC TCC TTC CAC CAC AAC CCC GCG CAT 1972
Pro Gln Gln Phe Ser Ser Phe His His Asn Pro Ala His
            495                 500

GAC AGT AAC AGC CTC CCT GCT AGC CCC TTG AGG ATA GTG 2011
Asp Ser Asn Ser Leu Pro Ala Ser Pro Leu Arg Ile Val
505                 510                 515

GAG GAT GAG GAG TAT GAA ACG ACC CAA GAG TAC GAG CCA 2050
Glu Asp Glu Glu Tyr Glu Thr Thr Gln Glu Tyr Glu Pro
        520                 525                 530

GCC CAA GAG CCT GTT AAG AAA CTC GCC AAT AGC CGG CGG 2089
Ala Gln Glu Pro Val Lys Lys Leu Ala Asn Ser Arg Arg
                535                 540

GCC AAA AGA ACC AAG CCC AAT GGC CAC ATT GCT AAC AGA 2128
Ala Lys Arg Thr Lys Pro Asn Gly His Ile Ala Asn Arg
    545                 550                 555

TTG GAA GTG GAC AGC AAC ACA AGC TCC CAG AGC AGT AAC 2167
Leu Glu Val Asp Ser Asn Thr Ser Ser Gln Ser Ser Asn
            560                 565

TCA GAG AGT GAA ACA GAA GAT GAA AGA GTA GGT GAA GAT 2206
Ser Glu Ser Glu Thr Glu Asp Glu Arg Val Gly Glu Asp
570                 575                 580

ACG CCT TTC CTG GGC ATA CAG AAC CCC CTG GCA GCC AGT 2245
Thr Pro Phe Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser
        585                 590                 595

CTT GAG GCA ACA CCT GCC TTC CGC CTG GCT GAC AGC AGG 2284
Leu Glu Ala Thr Pro Ala Phe Arg Leu Ala Asp Ser Arg
                600                 605

ACT AAC CCA GCA GGC CGC TTC TCG ACA CAG GAA GAA ATC 2323
Thr Asn Pro Ala Gly Arg Phe Ser Thr Gln Glu Glu Ile
    610                 615                 620

CAG GCC AGG CTG TCT AGT GTA ATT GCT AAC CAA GAC CCT 2362
Gln Ala Arg Leu Ser Ser Val Ile Ala Asn Gln Asp Pro
625                 630
```

FIG. 12D

```
ATT GCT GTA TAAAACCTA AATAAACACA TAGATTCACC TGTAAAACTT 2410
Ile Ala Val
635     637

TATTTTATAT AATAAAGTAT TCCACCTTAA ATTAAACAAT TTATTTTATT 2460

TTAGCAGTTC TGCAAATAAA AAAAAAAAAA 2490
```

FIG. 12E

```
GCGCCTGCCT CCAACCTGCG GGCGGGAGGT GGGTGGCTGC GGGGCAATTG 50

AAAAAGAGCC GGCGAGGAGT TCCCCGAAAC TTGTTGGAAC TCCGGGCTCG 100

CGCGGAGGCC AGGAGCTGAG CGGCGGCGGC TGCCGGACGA TGGGAGCGTG 150

AGCAGGACGG TGATAACCTC TCCCCGATCG GGTTGCGAGG GCGCCGGGCA 200

GAGGCCAGGA CGCGAGCCGC CAGCGGCGGG ACCCATCGAC GACTTCCCGG 250

GGCGACAGGA GCAGCCCCGA GAGCCAGGGC GAGCGCCCGT TCCAGGTGGC 300

CGGACCGCCC GCCGCGTCCG CGCCGCGCTC CCTGCAGGCA ACGGGAGACG 350

CCCCCGCGCA GCGCGAGCGC CTCAGCGCGG CCGCTCGCTC TCCCCATCGA 400

GGGACAAACT TTTCCCAAAC CCGATCCGAG CCCTTGGACC AAACTCGCCT 450

GCGCCGAGAG CCGTCCGCGT AGAGCGCTCC GTCTCCGGCG AG  ATG 495
                                                 Met
                                                   1

TCC GAG CGC AAA GAA GGC AGA GGC AAA GGG AAG GGC AAG 534
Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys
               5                  10

AAG AAG GAG CGA GGC TCC GGC AAG AAG CCG GAG TCC GCG 573
Lys Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala
 15                  20                  25

GCG GGC AGC CAG AGC CCA GCC TTG CCT CCC CAA TTG AAA 612
Ala Gly Ser Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys
         30                  35                  40

GAG ATG AAA AGC CAG GAA TCG GCT GCA GGT TCC AAA CTA 651
Glu Met Lys Ser Gln Glu Ser Ala Ala Gly Ser Lys Leu
                45                  50

GTC CTT CGG TGT GAA ACC AGT TCT GAA TAC TCC TCT CTC 690
Val Leu Arg Cys Glu Thr Ser Ser Glu Tyr Ser Ser Leu
    55                  60                  65

AGA TTC AAG TGG TTC AAG AAT GGG AAT GAA TTG AAT CGA 729
Arg Phe Lys Trp Phe Lys Asn Gly Asn Glu Leu Asn Arg
            70                  75
```

FIG. 13A

```
AAA AAC AAA CCA CAA AAT ATC AAG ATA CAA AAA AAG CCA 768
Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys Lys Pro
 80                  85                  90

GGG AAG TCA GAA CTT CGC ATT AAC AAA GCA TCA CTG GCT 807
Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala
         95                 100                 105

GAT TCT GGA GAG TAT ATG TGC AAA GTG ATC AGC AAA TTA 846
Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu
                110                 115

GGA AAT GAC AGT GCC TCT GCC AAT ATC ACC ATC GTG GAA 885
Gly Asn Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu
    120                 125                 130

TCA AAC GAG ATC ATC ACT GGT ATG CCA GCC TCA ACT GAA 924
Ser Asn Glu Ile Ile Thr Gly Met Pro Ala Ser Thr Glu
            135                 140

GGA GCA TAT GTG TCT TCA GAG TCT CCC ATT AGA ATA TCA 963
Gly Ala Tyr Val Ser Ser Glu Ser Pro Ile Arg Ile Ser
145                 150                 155

GTA TCC ACA GAA GGA GCA AAT ACT TCT TCA TCT ACA TCT 1002
Val Ser Thr Glu Gly Ala Asn Thr Ser Ser Ser Thr Ser
        160                 165                 170

ACA TCC ACC ACT GGG ACA AGC CAT CTT GTA AAA TGT GCG 1041
Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys Cys Ala
                175                 180

GAG AAG GAG AAA ACT TTC TGT GTG AAT GGA GGG GAG TGC 1080
Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys
    185                 190                 195

TTC ATG GTG AAA GAC CTT TCA AAC CCC TCG AGA TAC TTG 1119
Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu
            200                 205

TGC AAG TGC CCA AAT GAG TTT ACT GGT GAT CGC TGC CAA 1158
Cys Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln
210                 215                 220

AAC TAC GTA ATG GCC AGC TTC TAC AGT ACG TCC ACT CCC 1197
Asn Tyr Val Met Ala Ser Phe Tyr Ser Thr Ser Thr Pro
        225                 230                 235

TTT CTG TCT CTG CCT GAA TAGGA GCATGCTCAG TTGGTGCTGC 1240
Phe Leu Ser Leu Pro Glu
                240 241

TTTCTTGTTG CTGCATCTCC CCTCAGATTC CACCTAGAGC TAGATGTGTC 1290
```

FIG. 13B

TTACCAGATC TAATATTGAC TGCCTCTGCC TGTCGCATGA GAACATTAAC 1340

AAAAGCAATT GTATTACTTC CTCTGTTCGC GACTAGTTGG CTCTGAGATA 1390

CTAATAGGTG TGTGAGGCTC CGGATGTTTC TGGAATTGAT ATTGAATGAT 1440

GTGATACAAA TTGATAGTCA ATATCAAGCA GTGAAATATG ATAATAAAGG 1490

CATTTCAAAG TCTCACTTTT ATTGATAAAA TAAAAATCAT TCTACTGAAC 1540

AGTCCATCTT CTTTATACAA TGACCACATC CTGAAAAGGG TGTTGCTAAG 1590

CTGTAACCGA TATGCACTTG AAATGATGGT AAGTTAATTT TGATTCAGAA 1640

TGTGTTATTT GTCACAAATA AACATAATAA AAGGAGTTCA GATGTTTTTC 1690

TTCATTAACC AAAAAAAAAA AAAAA 1715

FIG. 13C

```
GAGGCGCCTG CCTCCAACCT GCGGGCGGGA GGTGGGTGGC TGCGGGGCAA 50

TTGAAAAAGA GCCGGCGAGG AGTTCCCCGA AACTTGTTGG AACTCCGGGC 100

TCGCGCGGAG GCCAGGAGCT GAGCGGCGGC GGCTGCCGGA CGATGGGAGC 150

GTGAGCAGGA CGGTGATAAC CTCTCCCCGA TCGGGTTGCG AGGGCGCCGG 200

GCAGAGGCCA GGACGCGAGC CGCCAGCGGC GGGACCCATC GACGACTTCC 250

CGGGGCGACA GGAGCAGCCC CGAGAGCCAG GGCGAGCGCC CGTTCCAGGT 300

GGCCGGACCG CCCGCCGCGT CCGCGCCGCG CTCCCTGCAG GCAACGGGAG 350

ACGCCCCCGC GCAGCGCGAG CGCCTCAGCG CGGCCGCTCG CTCTCCCCAT 400

CGAGGGACAA ACTTTTCCCA AACCCGATCC GAGCCCTTGG ACCAAACTCG 450

CCTGCGCCGA GAGCCGTCCG CGTAGAGCGC TCCGTCTCCG GCGAG   AT 497
                                                    Met
                                                     1

G TCC GAG CGC AAA GAA GGC AGA GGC AAA GGG AAG GGC AAG 537
  Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys
                5                   10

AAG AAG GAG CGA GGC TCC GGC AAG AAG CCG GAG TCC GCG 576
Lys Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala
 15                  20                  25

GCG GGC AGC CAG AGC CCA GCC TTG CCT CCC CAA TTG AAA 615
Ala Gly Ser Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys
         30                  35                  40

GAG ATG AAA AGC CAG GAA TCG GCT GCA GGT TCC AAA CTA 654
Glu Met Lys Ser Gln Glu Ser Ala Ala Gly Ser Lys Leu
                45                  50

GTC CTT CGG TGT GAA ACC AGT TCT GAA TAC TCC TCT CTC 693
Val Leu Arg Cys Glu Thr Ser Ser Glu Tyr Ser Ser Leu
     55                  60                  65

AGA TTC AAG TGG TTC AAG AAT GGG AAT GAA TTG AAT CGA 732
Arg Phe Lys Trp Phe Lys Asn Gly Asn Glu Leu Asn Arg
             70                  75
```

FIG. 14A

```
AAA AAC AAA CCA CAA AAT ATC AAG ATA CAA AAA AAG CCA 771
Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys Lys Pro
 80                  85                  90

GGG AAG TCA GAA CTT CGC ATT AAC AAA GCA TCA CTG GCT 810
Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala
         95                 100                 105

GAT TCT GGA GAG TAT ATG TGC AAA GTG ATC AGC AAA TTA 849
Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu
                110                 115

GGA AAT GAC AGT GCC TCT GCC AAT ATC ACC ATC GTG GAA 888
Gly Asn Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu
    120                 125                 130

TCA AAC GAG ATC ATC ACT GGT ATG CCA GCC TCA ACT GAA 927
Ser Asn Glu Ile Ile Thr Gly Met Pro Ala Ser Thr Glu
            135                 140

GGA GCA TAT GTG TCT TCA GAG TCT CCC ATT AGA ATA TCA 966
Gly Ala Tyr Val Ser Ser Glu Ser Pro Ile Arg Ile Ser
145                 150                 155

GTA TCC ACA GAA GGA GCA AAT ACT TCT TCA TCT ACA TCT 1005
Val Ser Thr Glu Gly Ala Asn Thr Ser Ser Ser Thr Ser
        160                 165                 170

ACA TCC ACC ACT GGG ACA AGC CAT CTT GTA AAA TGT GCG 1044
Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys Cys Ala
                175                 180

GAG AAG GAG AAA ACT TTC TGT GTG AAT GGA GGG GAG TGC 1083
Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys
    185                 190                 195

TTC ATG GTG AAA GAC CTT TCA AAC CCC TCG AGA TAC TTG 1122
Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu
            200                 205

TGC AAG TGC CCA AAT GAG TTT ACT GGT GAT CGC TGC CAA 1161
Cys Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln
210                 215                 220

AAC TAC GTA ATG GCC AGC TTC TAC AAG GCG GAG GAG CTG 1200
Asn Tyr Val Met Ala Ser Phe Tyr Lys Ala Glu Glu Leu
        225                 230                 235

TAC CAG AAG AGA GTG CTG ACC ATA ACC GGC ATC TGC ATC 1239
Tyr Gln Lys Arg Val Leu Thr Ile Thr Gly Ile Cys Ile
                240                 245

GCC CTC CTT GTG GTC GGC ATC ATG TGT GTG GTG GCC TAC 1278
Ala Leu Leu Val Val Gly Ile Met Cys Val Val Ala Tyr
    250                 255                 260
```

FIG. 14B

```
TGC AAA ACC AAG AAA CAG CGG AAA AAG CTG CAT GAC CGT 1317
Cys Lys Thr Lys Lys Gln Arg Lys Lys Leu His Asp Arg
            265                 270

CTT CGG CAG AGC CTT CGG TCT GAA CGA AAC AAT ATG ATG 1356
Leu Arg Gln Ser Leu Arg Ser Glu Arg Asn Asn Met Met
275                 280                 285

AAC ATT GCC AAT GGG CCT CAC CAT CCT AAC CCA CCC CCC 1395
Asn Ile Ala Asn Gly Pro His His Pro Asn Pro Pro Pro
        290                 295                 300

GAG AAT GTC CAG CTG GTG AAT CAA TAC GTA TCT AAA AAC 1434
Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys Asn
                305                 310

GTC ATC TCC AGT GAG CAT ATT GTT GAG AGA GAA GCA GAG 1473
Val Ile Ser Ser Glu His Ile Val Glu Arg Glu Ala Glu
    315                 320                 325

ACA TCC TTT TCC ACC AGT CAC TAT ACT TCC ACA GCC CAT 1512
Thr Ser Phe Ser Thr Ser His Tyr Thr Ser Thr Ala His
            330                 335

CAC TCC ACT ACT GTC ACC CAG ACT CCT AGC CAC AGC TGG 1551
His Ser Thr Thr Val Thr Gln Thr Pro Ser His Ser Trp
340                 345                 350

AGC AAC GGA CAC ACT GAA AGC ATC CTT TCC GAA AGC CAC 1590
Ser Asn Gly His Thr Glu Ser Ile Leu Ser Glu Ser His
        355                 360                 365

TCT GTA ATC GTG ATG TCA TCC GTA GAA AAC AGT AGG CAC 1629
Ser Val Ile Val Met Ser Ser Val Glu Asn Ser Arg His
                370                 375

AGC AGC CCA ACT GGG GGC CCA AGA GGA CGT CTT AAT GGC 1668
Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg Leu Asn Gly
    380                 385                 390

ACA GGA GGC CCT CGT GAA TGT AAC AGC TTC CTC AGG CAT 1707
Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg His
            395                 400

GCC AGA GAA ACC CCT GAT TCC TAC CGA GAC TCT CCT CAT 1746
Ala Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro His
405                 410                 415

AGT GAA AGG TAAAA CCGAAGGCAA AGCTACTGCA GAGGAGAAAC 1790
Ser Glu Arg
        420
```

FIG. 14C

TCAGTCAGAG AATCCCTGTG AGCACCTGCG GTCTCACCTC AGGAAATCTA 1840

CTCTAATCAG AATAAGGGGC GGCAGTTACC TGTTCTAGGA GTGCTCCTAG 1890

TTGATGAAGT CATCTCTTTG TTTGACGGAA CTTATTTCTT CTGAGCTTCT 1940

CTCGTCGTCC CAGTGACTGA CAGGCAACAG ACTCTTAAAG AGCTGGGATG 1990

CTTTGATGCG GAAGGTGCAG CACATGGAGT TTCCAGCTCT GGCCATGGGC 2040

TCAGACCCAC TCGGGGTCTC AGTGTCCTCA GTTGTAACAT TAGAGAGATG 2090

GCATCAATGC TTGATAAGGA CCCTTCTATA ATTCCAATTG CCAGTTATCC 2140

AAACTCTGAT TCGGTGGTCG AGCTGGCCTC GTGTTCTTAT CTGCTAACCC 2190

TGTCTTACCT TCCAGCCTCA GTTAAGTCAA ATCAAGGGCT ATGTCATTGC 2240

TGAATGTCAT GGGGGGCAAC TGCTTGCCCT CCACCCTATA GTATCTATTT 2290

TATGAAATTC CAAGAAGGGA TGAATAAATA AATCTCTTGG ATGCTGCGTC 2340

TGGCAGTCTT CACGGGTGGT TTTCAAAGCA GAAAAAAAAA AAAAAAAAAA 2390

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA A 2431

FIG. 14D

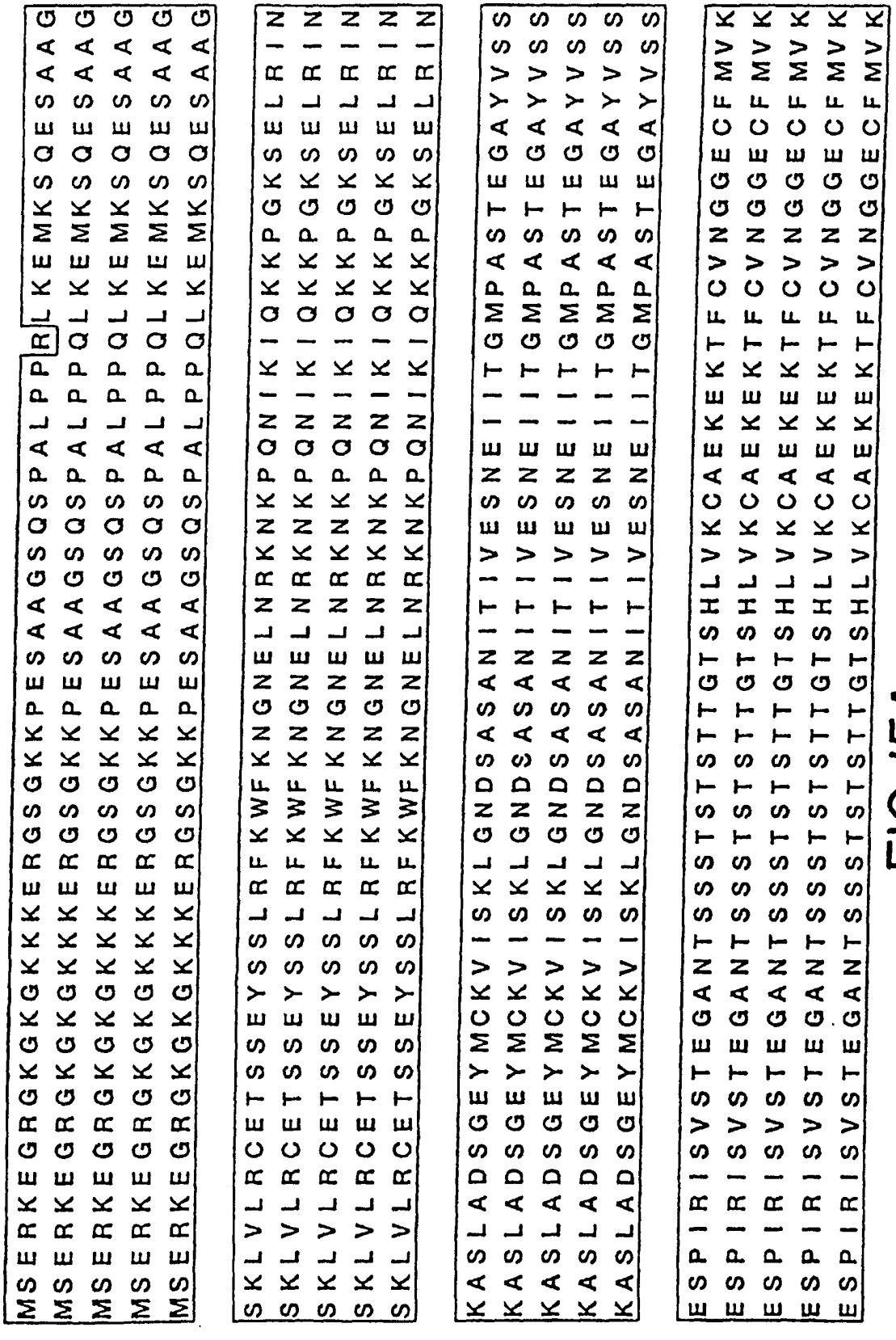

1 MSERKEGRGKGKKKERGSGKKPESAAGSQSPALPPRLKEMKSQESAAG
16
1 MSERKEGRGKGKKKERGSGKKPESAAGSQSPALPPQLKEMKSQESAAG
11
1 MSERKEGRGKGKKKERGSGKKPESAAGSQSPALPPQLKEMKSQESAAG
76
1 MSERKEGRGKGKKKERGSGKKPESAAGSQSPALPPQLKEMKSQESAAG
84
1 MSERKEGRGKGKKKERGSGKKPESAAGSQSPALPPQLKEMKSQESAAG
78
51 SKLVLRCETSSEYSSLRFKWFKNGNELNRKNKPQNIKIQKKKPGKSELRIN
16
51 SKLVLRCETSSEYSSLRFKWFKNGNELNRKNKPQNIKIQKKKPGKSELRIN
11
51 SKLVLRCETSSEYSSLRFKWFKNGNELNRKNKPQNIKIQKKKPGKSELRIN
76
51 SKLVLRCETSSEYSSLRFKWFKNGNELNRKNKPQNIKIQKKKPGKSELRIN
84
51 SKLVLRCETSSEYSSLRFKWFKNGNELNRKNKPQNIKIQKKKPGKSELRIN
78
101 KASLADSGEYMCKVISKLGNDSASANITIVESNEIITGMPASTEGAYVSS
16
101 KASLADSGEYMCKVISKLGNDSASANITIVESNEIITGMPASTEGAYVSS
11
101 KASLADSGEYMCKVISKLGNDSASANITIVESNEIITGMPASTEGAYVSS
76
101 KASLADSGEYMCKVISKLGNDSASANITIVESNEIITGMPASTEGAYVSS
84
101 KASLADSGEYMCKVISKLGNDSASANITIVESNEIITGMPASTEGAYVSS
78
151 ESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTFCVNGGECFMVK
16
151 ESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTFCVNGGECFMVK
11
151 ESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTFCVNGGECFMVK
76
151 ESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTFCVNGGECFMVK
84
151 ESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTFCVNGGECFMVK
78

FIG. 15A

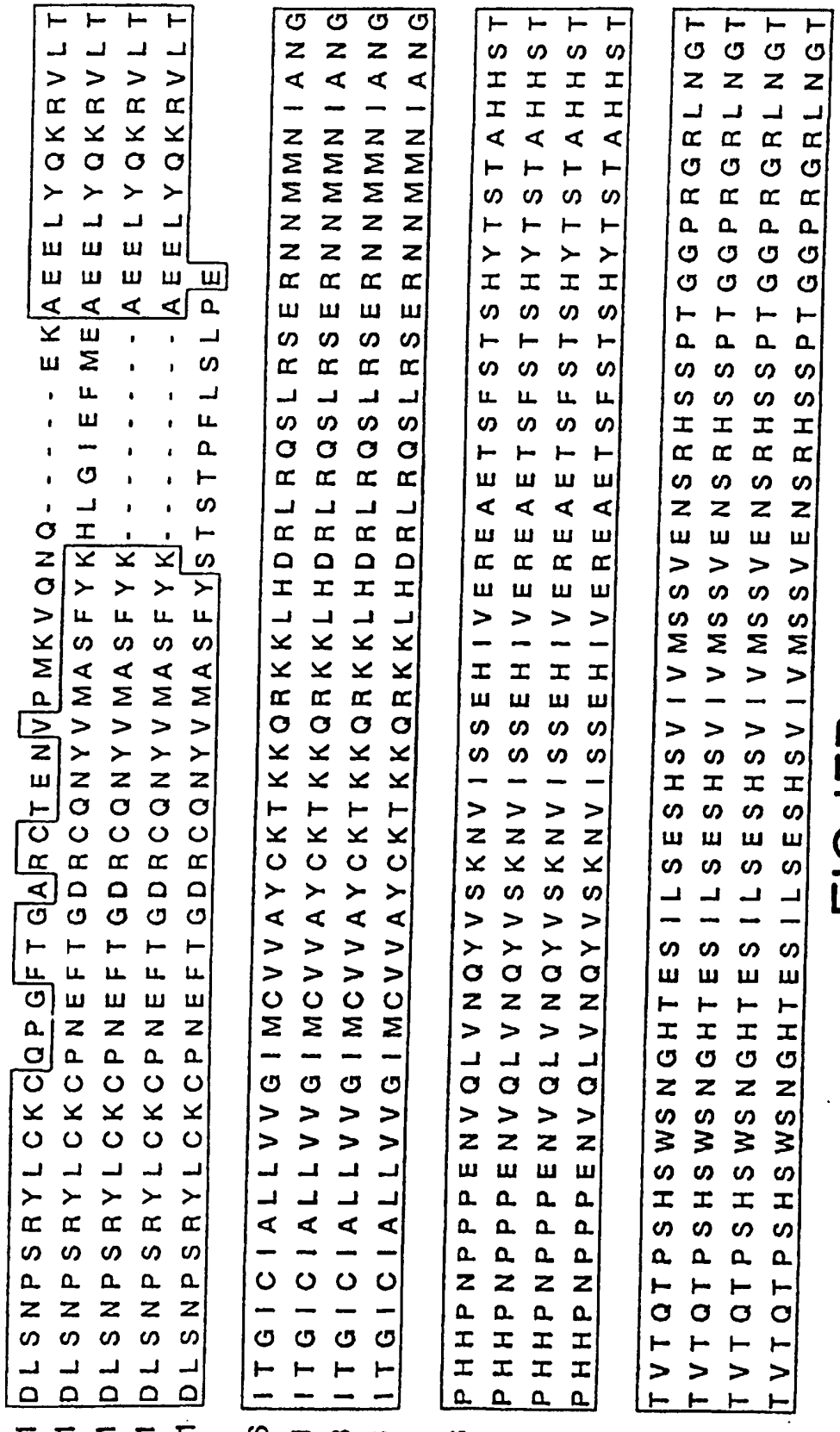

16 201 DLSNPSRYLCKCQPGFTGARCTENVPMKVQNQ- - - - - EK AEELYQKRVLT
11 201 DLSNPSRYLCKCPNEFTGDRCQNYVMASFYKHLGIEFME AEELYQKRVLT
76 201 DLSNPSRYLCKCPNEFTGDRCQNYVMASFYK- - - - - - - AEELYQKRVLT
84 201 DLSNPSRYLCKCPNEFTGDRCQNYVMASFYK- - - - - - - AEELYQKRVLT
78 201 DLSNPSRYLCKCPNEFTGDRCQNYVMASFYSTSTPFLSLPE AEELYQKRVLT
16 246 ITGICIALLVVGIMCVVAYCKTKKQRKKLHDRLRQSLRSERNNMMNIANG
11 251 ITGICIALLVVGIMCVVAYCKTKKQRKKLHDRLRQSLRSERNNMMNIANG
76 243 ITGICIALLVVGIMCVVAYCKTKKQRKKLHDRLRQSLRSERNNMMNIANG
84 243 ITGICIALLVVGIMCVVAYCKTKKQRKKLHDRLRQSLRSERNNMMNIANG
16 296 PHHPNPPPENVQLVNQYVSKNVISSEHIVEREAETSFSTSHYTSTAHHST
11 301 PHHPNPPPENVQLVNQYVSKNVISSEHIVEREAETSFSTSHYTSTAHHST
76 293 PHHPNPPPENVQLVNQYVSKNVISSEHIVEREAETSFSTSHYTSTAHHST
84 293 PHHPNPPPENVQLVNQYVSKNVISSEHIVEREAETSFSTSHYTSTAHHST
16 346 TVTQTPSHSWSNGHTESILSESHSVIVMSSVENSRHSSPTGGPRGRLNGT
11 351 TVTQTPSHSWSNGHTESILSESHSVIVMSSVENSRHSSPTGGPRGRLNGT
76 343 TVTQTPSHSWSNGHTESILSESHSVIVMSSVENSRHSSPTGGPRGRLNGT
84 343 TVTQTPSHSWSNGHTESILSESHSVIVMSSVENSRHSSPTGGPRGRLNGT

FIG. 15B

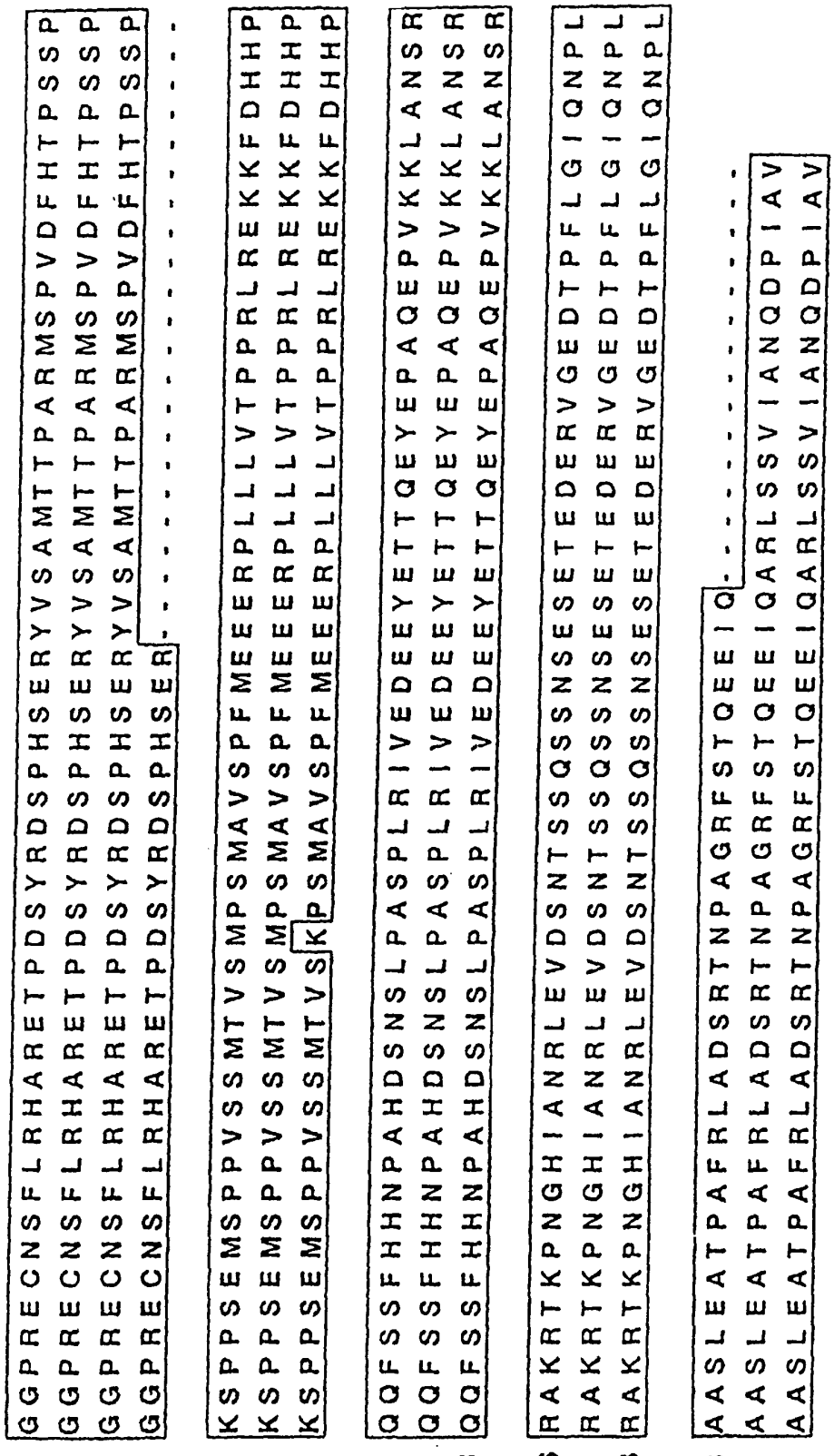

16 396 GGPRECNSFLRHARETPDSYRDSPHSERYVSAMTTPARMSPVDFHTPSSP
11 401 GGPRECNSFLRHARETPDSYRDSPHSERYVSAMTTPARMSPVDFHTPSSP
76 393 GGPRECNSFLRHARETPDSYRDSPHSERYVSAMTTPARMSPVDFHTPSSP
84 393 GGPRECNSFLRHARETPDSYRDSPHSER----------
16 446 KSPPSEMSPPVSSMTVSMPSMAVSPFMEEERPLLLVTPPRLREKKFDHHP
11 451 KSPPSEMSPPVSSMTVSMPSMAVSPFMEEERPLLLVTPPRLREKKFDHHP
76 443 KSPPSEMSPPVSSMTVSKPSMAVSPFMEEERPLLLVTPPRLREKKFDHHP
16 496 QQFSSFHHNPAHDSNSLPASPLRIVEDEEYETTQEYEPAQEPVKKLANSR
11 501 QQFSSFHHNPAHDSNSLPASPLRIVEDEEYETTQEYEPAQEPVKKLANSR
76 493 QQFSSFHHNPAHDSNSLPASPLRIVEDEEYETTQEYEPAQEPVKKLANSR
16 546 RAKRTKPNGHIANRLEVDSNTSSQSSNSESETEDERVGEDTPFLGIQNPL
11 551 RAKRTKPNGHIANRLEVDSNTSSQSSNSESETEDERVGEDTPFLGIQNPL
76 543 RAKRTKPNGHIANRLEVDSNTSSQSSNSESETEDERVGEDTPFLGIQNPL
16 596 AASLEATPAFRLADSRTNPAGRFSTQEEIQ----------
11 601 AASLEATPAFRLADSRTNPAGRFSTQEEIQARLSSVIANQDPIAV
76 593 AASLEATPAFRLADSRTNPAGRFSTQEEIQARLSSVIANQDPIAV

FIG. 15C