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

(10) **DE 693 31 832 T3** 2009.07.16

(12) Übersetzung der geänderten europäischen Patentschrift

(97) **EP 0 662 827 B2**
(21) Deutsches Aktenzeichen: **693 31 832.5**
(86) PCT-Aktenzeichen: **PCT/US93/09338**
(96) Europäisches Aktenzeichen: **93 923 752.5**
(87) PCT-Veröffentlichungs-Nr.: **WO 1994/007474**
(86) PCT-Anmeldetag: **30.09.1993**
(87) Veröffentlichungstag der PCT-Anmeldung: **14.04.1994**
(97) Erstveröffentlichung durch das EPA: **19.07.1995**
(97) Veröffentlichungstag der Patenterteilung beim EPA: **17.04.2002**
(97) Veröffentlichungstag des geänderten Patents beim EPA: **07.01.2009**
(47) Veröffentlichungstag im Patentblatt: **16.07.2009**

(51) Int Cl.[8]: ***A61K 31/00*** (2006.01)
A61K 31/70 (2006.01)
A61K 38/00 (2006.01)
A61K 39/44 (2006.01)
A61K 39/395 (2006.01)
C07H 21/04 (2006.01)
G01N 33/53 (2006.01)
G01N 33/68 (2006.01)
G01N 33/574 (2006.01)
C07K 14/47 (2006.01)

Patentschrift wurde im Einspruchsverfahren geändert

(30) Unionspriorität:

955012	**30.09.1992**	**US**
83590	**25.06.1993**	**US**

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

(73) Patentinhaber:
Yale University, New Haven, Conn., US

(74) Vertreter:
Andrae Flach Haug, 81541 München

(72) Erfinder:
ARTAVANIS-TSAKONAS, Spyridon, Hamden, Conn., US; FEHON, Richard Grant, Durham, N.C., US; ZAGOURAS, Panayiotis, New Haven, Conn., US; BLAUMUELLER, Christine Marie, New Haven, Conn., US

(54) Bezeichnung: **DIAGNOSTISCHE VERFAHREN UND PHARMAZEUTISCHE ZUSAMMENSETZUNGEN AUF DER BASIS VON NOTCH-PROTEINEN UND NUKLEINSÄUREN**

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

1. EINLEITUNG

[0001] Die Erfindung betrifft therapeutische Zusammensetzungen, die Notch-Proteine, Analoge und Derivate hiervon, Antikörper dagegen, Nukleinsäuren, die die Notch-Proteine, Derivate oder Analoge codieren, Notch-Antisinn-Nukleinsäuren enthalten. Therapeutische Anwendungen und Diagnoseverfahren sind ebenfalls bereitgestellt.

2. HINTERGRUND DER ERFINDUNG

2.1. DAS NOTCH-GEN UND -PROTEIN

[0002] Nullmutationen an einem der zygotischen neurogenen Loci – Notch (N), Delta (Dl), mastermind (mam), Enhancer of Split (E(spl)), neuralized (neu), und big brain (bib) – bewirken eine Hypertrophie des Nervensystems auf Kosten der ventralen und lateralen epidermalen Strukturen. Dieser Effekt beruht auf der Fehlleitung von epidermalen Vorläuferzellen zu einer neuronalen Route und legt nahe, daß die neurogene Genfunktion zur Diversifizierung der Zellen in der neurogenen Region von einem neuronalen zu einem epithelialen Werdegang notwendig ist. Studien, die die Auswirkungen der Laserablation spezifischer embryonaler Neuroblasten bei Heuschrecken bewerteten (Doe und Goodman 1985, Dev. Biol. 111, 206–219), haben gezeigt, daß zelluläre Wechselwirkungen zwischen Neuroblasten und den umgebenden akzessorischen Zellen den Zweck haben, diese akzessorischen Zellen davon abzuhalten, einen Neuroblasten-Werdegang einzuschlagen. Zusammengenommen haben diese entwicklungsgenetischen Beobachtungen zu der Hypothese geführt, daß die Proteinprodukte der neurogenen Loci als Komponenten eines zellulären Interaktionsmechanismus wirken, der zur ordnungsgemäßen Entwicklung der Epidermis notwendig ist (Artavanis-Tsakonas, 1988, Trends Genet. 4, 95–100).

[0003] Sequenzanalysen (Wharton et al., 1985, Cell 43, 567–581; Kidd et al., 1986, Mol. Cell. Biol. 6, 3094–3108; Vassin et al., 1987, EMBO J. 6, 3431–3440 ; Kopczynski et al., 1988, Genes Dev. 2, 1723–1735) haben gezeigt, daß zwei der neurogenen Loci, Notch und Delta, offensichtlich Transmembranproteine codieren, die sich einmal durch die Membran hindurch erstrecken. Das Drosophila-Notch-Gen codiert ein ca. 300 kd-Protein (wir verwenden "Notch" zur Bezeichnung dieses Proteins) mit einer großen N-terminalen extrazellulären Domäne, die 36 epidermale Wachstumsfaktor(EGF)-artige Sequenzwiederholungen in Tandemanordnung enthält, auf die drei weitere, cysteinreiche Sequenzwiederholungen folgen, die als Notch/lin-12-Sequenzwiederholungen bezeichnet werden (Wharton et al., 1985, Cell 43, 567–581; Kidd et al., 1986, Mol. Cell Biol. 6, 3094–3108; Yochem et al., 1988, Nature 335, 547–550). Die Sequenzen von Xenopus (Coffman et al., 1990, Science 249: 1438–1441) und einem menschlichen Notch-Homologen mit der Bezeichnung TAN-1 (Ellisen et al., 1991, Cell 66: 649–661) wurden ebenfalls beschrieben. Delta codiert ein ca. 100 kd-Protein (wir verwenden "Delta" zur Bezeichnung von DLZM), das Proteinprodukt der vorherrschenden zygotischen und materen Transkripte; Kopczynski et al., 1988, Genes Dev. 2, 1723–1735), das neun EGF-artige Sequenzwiederholungen in seiner extrazellulären Domäne aufweist (Vassin et al., 1987, EMBO J. 6, 3431–3440, Kopczynski et al., 1988, Genes Dev. 2, 1723–1735). Obwohl über die funktionelle Bedeutung dieser Sequenzwiederholungen wenig bekannt ist, wurde das EGF-artige Motiv bei einer Vielzahl von Proteinen gefunden, einschließlich derjenigen, die an der Blutgerinnungskaskade beteiligt sind (Furie und Furie, 1988, Cell 53, 505–518). Insbesondere wurde dieses Motiv in extrazellulären Proteinen nachgewiesen, wie in den Blutgerinnungsfaktoren IX und X (Rees et al., 1988, EMBO J. 7, 2053–2061; Furie und Furie, 1988, Cell 53, 505–518), in weiteren Drosophila-Genen (Knust et al., 1987, EMBO J. 761–766; Rothberg et al., 1988, Cell 55, 1047–1059) und in einigen Zelloberflächen-Rezeptorproteinen, wie Thrombomodulin (Suzuki et al., 1987, EMBO J. 6, 1891–1897), und im LDL-Rezeptor (Sudhof et al., 1985, Science 228, 815–822). In Thrombomodulin und Urokinase wurde eine Proteinbindungsstelle der EGF-Sequenzwiederholungsdomäne zugeordnet (Kurosawa et al., 1988, J. Biol. Chem. 263, 5993–5996; Appella et al., 1987, J. Biol. Chem. 262, 4437–4440).

[0004] Es wurde bereits eine interessante Reihe von Wechselwirkungen zwischen Notch- und Delta-Mutationen beschrieben (Vassin, et al., 1985, J. Neurogenet. 2, 291–308; Shepard et al., 1989, Genetics 122, 429–438; Xu et al., 1990, Genes Dev., 4, 464–475). Eine Reihe genetischer Studien (Zusammenfassung bei Alton et. al., 1989, Dev. Genet. 10, 261–272) hat gezeigt, daß die Gen-Dosierungen von Notch und Delta in Relation zueinander für die normale Entwicklung wesentlich sind. Eine 50%ige Verminderung in der Dosis von Delta bei einem Notch-Wild-Typ-Hintergrund verursacht eine Verbreiterung der Flügelvenen unter Erzeugung eines "Delta" an der Basis (Lindsley and Grell, 1968, Veröffentlichungsnummer 627, Washington, D. C., Carnegie Institute of Washington). Ein ähnlicher Phänotyp kommt durch eine 50%ige Erhöhung der Dosis von

Notch bei einem Delta-Wildtyp-Hintergrund zustande (ein "konfluenter" Phänotyp; Welshons, 1965, Science 150, 1122–1129). Dieser Delta-Phänotyp wird durch eine verminderte Notch-Dosis teilweise unterdrückt. Arbeiten haben gezeigt, daß letale Wechselwirkungen zwischen Allelen, die mit Änderungen in den EGF-artigen Sequenzwiederholungen in Notch korrelieren, durch Verminderung der Dosis von Delta verhindert werden können, (Xu et al., 1990, Genes Dev. 4, 464–475). Xu et al. (1990, Genes Dev. 4, 464–475) haben festgestellt, daß die als Abruptex-(Ax)-Mutationen bekannten Null-Mutationen entweder bei Delta oder bei mam die letalen Wechselwirkungen zwischen Heterozygoten-Kombinationen bestimmter Notch-Allele unterdrücken. Ax-Allele hängen mit Fehlsinn-Mutationen in den EGF-artigen Sequenzwiederholungen der extrazellulären Notch-Domäne zusammen (Kelley et al., 1987, Cell 51, 539–548; Hartley et al., 1987, EMBO J. 6, 3407–3417).

[0005] Neuere Studien haben gezeigt, daß Notch und Delta und Notch und Serrate direkt auf molekularem Niveau Wechselwirken (Fehon et al., 1990, Cell 61: 523–534; Rebay et al., 1991, Cell 67: 687–699).

[0006] Notch wird in Axon-Prozessen während des Herauswachsens embryonaler Neuronen exprimiert (Johansen et al., 1989. J. Cell. Biol. 109: 2427–2440; Kidd et al., 1989, Genes Dev. 3: 1113–1129; Fehon et al., 1991, J. Cell. Biol. 113: 657–669).

[0007] Eine Studie hat gezeigt, daß bestimmte Ax-Allele von Notch die Axon-Wegfindung während des sensorischen neuronalen Herauswachsens in den Imaginalscheiben gravierend verändern können, obwohl es bisher noch nicht bekannt ist, ob eine anormale Notch-Expression im Axon selbst oder im Epithel, an dem es entlang wächst, für diesen Defekt verantwortlich ist (Palka et al., 1990, Development 109, 167–175).

2.2. KREBS

[0008] Ein Neoplasma oder Tumor ist eine neoplastische Masse, die von einem abnormen unkontrollierten Zellwachstum herrührt, das ein Anschwellen auf der Körperoberfläche verursachen und gutartig oder bösartig sein kann. Gutartige Tumore bleiben in der Regel lokal. Bösartige Tumore werden kollektiv als Krebs bezeichnet. Der Begriff "bösartig" bedeutet allgemein, daß der Tumor in benachbarte Körperstrukturen eindringen und sie zerstören kann und sich zu weit entfernten Stellen ausbreiten und zum Tode führen kann (ein Überblick findet sich bei Robbins und Angell, 1976, Basic Pathology, 2. Ausg., W. B. Saunders Co., Philadelphia, Ss. 68–122).

[0009] Eine wirksame Behandlung und Prävention von Krebs bleibt immer noch ein lang verspürtes Bedürfnis und ein Hauptziel der biomedizinischen Forschung.

3. ZUSAMMENFASSUNG DER ERFINDUNG

[0010] Die vorliegende Erfindung wird in den Ansprüchen 1 bis 19 definiert, die am Ende der Beschreibung aufgeführt sind.

[0011] Die folgende Beschreibung beschreibt Einzelheiten, die sich auf die beiden Ausführungsformen beziehen, die von den Ansprüchen abgedeckt werden, sowie Gegenstände, die nicht von den Ansprüchen abgedeckt werden. Die Beschreibung von Gegenständen, die nicht von den Ansprüchen abgedeckt werden, erfolgt nur zur Information.

3.1. DEFINITIONEN

[0012] Wie hier verwendet, sollen die folgenden Begriffe die angegebenen Bedeutungen haben:

AA = Aminosäure
EGF = epidermaler Wachstumsfaktor
ELR = EGF-artige (homologe) Sequenzwiederholung
IC = intrazellulär
PCR = Polymerasekettenreaktion

[0013] Wie hier verwendet, bezeichnet die Unterstreichung des Namens eines Gens das Gen, im Gegensatz zu seinem codierten Protein-Produkt, das durch den Namen des Gens ohne Unterstreichung angegeben ist. Beispielsweise bedeutet "Notch" das Notch-Gen, wohin gegen "Notch" das Proteinprodukt des Notch-Gens bezeichnet.

4. BESCHREIBUNG DER FIGUREN

[0014] **Fig.** 1: Primäre Nukleotidsequenz der Delta-cDNA-D11 (SEQ ID NO: 1) und der Delta-Aminosäuresequenz (SEQ ID NO: 2). Gezeigt ist die DNA-Sequenz des 5'-3'-Strangs der D11-cDNA, die im Vergleich zu der bei Kopczynski et al. angegebenen (1988, Genes Dev. 2: 1723–1735) eine Reihe von Korrekturen enthält.

[0015] **Fig.** 2: Notch-Expressionskonstrukte und Deletionskartierung der Delta/Serrate-Bindungsdomäne. S2-Zellen im log-Phasenwachstum wurden vorübergehend mit der Reihe von gezeigten Expressionskonstrukten transfiziert; die Zeichnungen stellen die vorhergesagten Proteinprodukte der verschiedenen erzeugten Notch-Deletionsmutanten dar. Sämtliche Expressionskonstrukte stammten vom Konstrukt Nr. 1, pMtNMg. Die vorübergehend transfizierten Zellen wurden mit Delta-exprimierenden Zellen aus der stabil transformierten Linie L49-6-7 oder mit vorübergehend transfizierten Serrate-exprimierenden Zellen gemischt, mit $CuSO_4$ induziert, unter Aggregationsbedingungen inkubiert und sodann unter Verwendung spezieller Antiseren und unter Anwendung der Immunfluoreszenzmikroskopie auf ihre Aggregationsfähigkeit bewertet. Aggregate wurden als Cluster von vier oder mehr Zellen, die sowohl Notch- als auch Delta/Serrate-exprimierende Zellen enthalten, definiert. Die Werte, die für die Aggregation in % angegeben sind, beziehen sich auf den Prozentsatz sämtlicher Notch-exprimierender Zellen, die in solchen Clustern entweder mit Delta (D1) (linke Spalte) oder mit Serrate (Ser) (rechte Spalte) festgestellt wurden. Die verschiedenen Notch-Deletionskonstrukte sind diagrammartig dargestellt, wobei die geknickten Linien die Ligationsverbindungsstellen angeben. Jede EGF-Sequenzwiederholung ist als schraffierte, rechteckige Umrahmung angegeben, und die Anzahl der EGF-Sequenzwiederholungen auf jeder Seite einer Ligationsverbindungsstelle ist angegeben. An den Ligationsverbindungsstellen sind partielle EGF-Sequenzwiederholungen, die durch die verschiedenen Deletionen erzeugt werden, mit offenen Rechtecken und geschlossenen Klammern bezeichnet (siehe beispielsweise Nr. 23 ΔCla + EGF(10–12)). Die Konstrukte Nr. 3–13 stellen die ClaI-Deletionsreihe dar. Wie diagrammartig angegeben, brechen 4 der ClaI-Stellen in den Sequenzwiederholungen 7, 9, 17 und 26 die Sequenzwiederholung in der Mitte unmittelbar nach dem dritten Cystein (bezeichnet durch die offene-Rechteck-Sequenzwiederholungen; siehe **Fig. 3** zur weiteren Klärung), während die fünfte und am nächsten an 3' gelegene Stelle glatt zwischen der EGF-Sequenzwiederholung 30 und der EGF-Sequenzwiederholung 31 bricht (bezeichnet durch die geschlossene Rechteck-Sequenzwiederholung 31; siehe wiederum **Fig. 3**). Im gestückelten Konstrukt Nr. 15 ist die EGF-Sequenzwiederholung 14, die die gestückelte Punktmutation trägt, als gestreiftes Rechteck eingezeichnet. In Konstrukt Nr. 33 ΔCla + XEGF(10–13) sind die von Xenopus-Notch stammenden EGF-Sequenzwiederholungen von den Drosophila-Sequenzwiederholungen durch ein unterschiedliches Schraffurmuster unterschieden. SP, Signalpeptid; EGF, epidermale Wachstumsfaktor-Sequenzwiederholung; N, Notch/Lin-12-Sequenzwiederholung; TM, Transmembran-Domäne; cdc10, cdc10/Ankyrin-Sequenzwiederholung; PA, vermeintliche verwandte Nukleotidbindungssequenz; Opa, Polyglutamin-Verlängerung mit der Bezeichnung Opa; Dl, Delta; Ser, Serrate.

[0016] **Fig. 3**: Ausführliche Struktur der Notch-Deletionskonstrukte Nr. 19–24: Sowohl die EGF-Sequenzwiederholungen 11 als auch 12 sind für die Notch-Delta-Aggregation erforderlich. Die EGF-Sequenzwiederholungen 10–13 sind diagrammartig oben aufgeführt und zeigen den regulären Abstand der sechs Cysteinreste (C). Die für diese Konstrukte erzeugten PCR-Produkte (Name und Anzahl wie in **Fig.** 2 angegeben), sind durch die fetten schwarzen Balken dargestellt und die exakten Endpunkte sind relativ zu den verschiedenen EGF-Sequenzwiederholungen bezeichnet. Die Fähigkeit zur Aggregation mit Delta ist für jedes Konstrukt mit (+) oder (–) bezeichnet. Die PCR-Fragmente brechen die EGF-Sequenzwiederholungen entweder in der Mitte, unmittelbar nach dem dritten Cystein an der gleichen Stelle wie vier der fünf ClaI-Stellen oder genau zwischen zwei Sequenzwiederholungen an der gleichen Stelle wie die am meisten C-terminale ClaI-Stelle auf.

[0017] **Fig. 4**: Vergleich der Aminosäuresequenz der EGF-Sequenzwiederholungen 11 und 12 aus Drosophila- und Xenopus-Notch. Die Aminosäuresequenz der EGF-Sequenzwiederholungen 11 und 12 von Drosophila-Notch (SEQ ID NO: 14) (Wharton et al., 1985, Cell 43: 567–581; Kidd et al., 1986, Mol. Cell Biol. 6: 3094–3108) ist mit derjenigen der gleichen zwei EGF-Sequenzwiederholungen aus Xenopus-Notch (SEQ ID NO: 15) (Coffman et al., 1990, Science 249: 1438–1441) ausgerichtet. Identische Aminosäuren sind eingerahmt. Die sechs konservierten Cysteinreste jeder EGF-Sequenzwiederholung und die verwandten Ca-bindenden Reste (Rees et al., 1988, EMBO J. 7: 2053–2061) sind mit einem Stern (*) markiert. Die Änderung von Leucin zu Prolin, die im Xenopus-PCR-Klon, der nicht zu aggregieren vermochte, festgestellt wird, ist darunter gekennzeichnet.

[0018] **Fig.** 5: Nukleinsäuresequenzhomologien zwischen Serrate und Delta. Gezeigt ist ein Teil der Drosophila-Serrate-Nukleotidsequenz (SEQ ID NO: 3), wobei die codierte Serrate-Proteinsequenz (SEQ ID NO: 4) darunter aufgeführt ist (Fleming et al., 1990, Genes & Dev. 4: 2188–2201 und 2193–94). Die vier Bereiche, die

die hohe Sequenzhomologie mit der Drosophila-Delta-Sequenz zeigen, sind über der Zeile nummeriert und durch Klammern bezeichnet. Die homologe Gesamtregion umspannt die Nukleotidnummern 627 bis 1290 der Serrate-Nukleotidsequenz (Nummerierung wie in Fig. 4 von Fleming et al., 1990, Genes & Dev. 4: 2188–2201).

[0019] Fig. 6: Schematisches Diagramm menschlicher Notch-Klone. Gezeigt ist ein schematisches Diagramm von menschlichen Notch-Klonen. Die schwarzen Balkenlinien unter dem Diagramm zeigen den Teil der Notch-Sequenz, der in jedem der vier cDNA-Klone enthalten ist. Die Lokation der bei der PCR verwendeten Primer und ihre Orientierung sind durch Pfeile angegeben.

[0020] Fig. 7: Menschliche Notch-Sequenzen, die mit der Drosophila-Notch-Sequenz ausgerichtet sind. Die nummerierten vertikalen Linien entsprechen den Drosophila-Notch-Koordinaten. Die horizontalen Linien unter jeder Karte zeigen, wo relativ zu Sequenzverlängerungen (dicke horizontale Linien) Klone liegen.

[0021] **Fig.** 8: Nukleotidsequenzen des menschlichem Notch, das im Plasmid-cDNA-Klon hN2k enthalten ist. Fig. 8A: Gezeigt ist die DNA-Sequenz (SEQ ID NO: 5) eines Teils des menschlichen Notch-Insertionen, beginnend an der EcoRI-Stelle am 3'-Ende und weiter von 3' in Richtung 5'. Fig. 8B: Gezeigt ist die DNA-Sequenz (SEQ ID NO: 6) eines Teils des menschlichen Notch-Insertionen, beginnend an der EcoRI-Stelle am 5'-Ende und weiter von 5' in Richtung 3'. Fig. 8C: Gezeigt ist die DNA-Sequenz (SEQ ID NO: 7) eines Teils des menschlichen Notch-Insertionen, beginnend bei 3' der in Fig. 8B gezeigten Sequenz und weiter von 5' in Richtung 3'. Die gezeigten Sequenzen sind Versuche und bedürfen einer Bestätigung durch die Bestimmung der Überlappungssequenzen.

[0022] **Fig.** 9: Nukleotidsequenzen des menschlichen Notch, enthalten im Plasmid-cDNA-Klon hN4k. Fig. 9A: Gezeigt ist die DNA-Sequenz (SEQ ID NO: 8) eines Teils des menschlichen Notch-Insertionen, beginnend an der EcoRI-Stelle am 5'-Ende und weiter von 5' in Richtung 3'. Fig. 9B: Gezeigt ist die DNA-Sequenz (SEQ ID NO: 9) eines Teils der menschlichen Notch-Insertionen, beginnend nahe am 3'-Ende und weiter von 3' in Richtung 5'. Die gezeigten Sequenzen stellen Versuche dar und bedürfen einer Bestätigung durch die Bestimmung der Überlappungssequenzen.

[0023] **Fig.** 10: DNA (SEQ ID NO: 10) – und Aminosäure (SEQ ID NO: 11) – Sequenzen des menschlichen Notch, die im Plasmid-cDNA-Klon hN3k enthalten sind.

[0024] **Fig.** 11: DNA (SEQ ID NO: 12) – und Aminosäure (SEQ ID NO: 13) – Sequenzen des menschlichen Notch, die im Plasmid-cDNA-Klon hN5k enthalten sind.

[0025] **Fig.** 12: Vergleich von hN5k mit anderen Notch-Homologen. Fig. 12A: Schematische Darstellung von Drosophila-Notch. Angegeben sind die Signalsequenz (Signal), die 36 EGF-artigen Sequenzwiederholungen, die 3 Notch/lin-12-Sequenzwiederholungen, die Transmembrandomäne (TM), die 6 CDC10-Sequenzwiederholungen, die OPA-Sequenzwiederholung und die PEST-(Prolin, Glutaminsäure, Serin, Threonin)-reiche Region. Fig. 12B: Ausrichtung der abgeleiteten Aminosäuresequenz von hN5k mit Sequenzen anderer Notch-Homologe. Die Aminosäuren sind links nummeriert. Die cdc10- und PEST-reichen Regionen sind beide eingerahmt, und die einzelnen cdc10-Sequenzwiederholungen sind markiert. Aminosäuren, die in drei oder mehr Sequenzen identisch sind, sind gekennzeichnet. Die zur Klonierung von hN5k verwendeten Primer sind unter den Sequenzen, aus denen sie konstruiert wurden, angegeben. Die Kern-Lokalisierungssequenz (NLS), die Caseinkinase II(CKII)- und cdc2-Kinase(cdc2)-Stellen des vermeintlichen CcN-Motivs der Vertebraten-Notch-Homologe sind eingerahmt. Die mögliche zweiteilige Kern-Zielsequenz (BNTS) und die proximalen Phosphorylierungsstellen von Drosophila-Notch sind ebenfalls eingerahmt.

[0026] **Fig.** 13: Ausgerichtete Aminosäuresequenzen der Notch-Proteine verschiedener Spezies. humN: Das menschliche Notch-Protein, das von dem hN-Homologen (teilweise im Plasmid hN5k enthalten) (SEQ ID NO: 19) codiert wird. TAN-1: Das menschliche Notch-Protein, das von dem TAN-1-Homologen (SEQ ID NO: 20) (die gezeigte Sequenz entstammt zum Teil unserer Arbeit und zum Teil der TAN-1-Sequenz, wie von Ellisen et al., 1991, Cell 66: 649–661 veröffentlicht) codiert wird; Xen N: Xenopus-Notch-Protein (Coffman et al., 1990, Science 249: 1438–1441). Dros N: Drosophila-Notch-Protein (Wharton et al., 1985, Cell 43: 567–581). Die Strukturdomänen sind angegeben.

[0027] Fig. 14: Immunozytochemische Anfärbung von Brustkrebsgewebe eines menschlichen Patienten. Bösartiges Brustgewebe, das aus einer Probe eines menschlichen Patienten erhalten wurde, wurde in einen Paraffinschnitt eingebettet und einer immunozytochemischen Anfärbung mit dem monoklonalen Antihuman-Notch-Antikörper P4, der gegen das TAN-1-Protein gerichtet ist, unterzogen. Nicht bösartiges Brustgewe-

be zeigte eine viel geringere Färbung (nicht gezeigt).

[0028] **Fig.** 15: Immunozytochemische Anfärbung von Darmgewebe aus einem menschlichen Patienten mit Darmkrebs. Eine aus einem Patienten mit Darmkrebs erhaltene Darmgewebsprobe wurde in einen Paraffinschnitt eingebettet und einer immunozytochemischen Anfärbung mit dem monoklonalen Antihuman-Notch-Antikörper P1, der gegen das hN-codierte Protein gerichtet ist, unterzogen. Die Bereiche einer verstärkten Anfärbung sind die Bereiche, in denen die bösartigen Zellen vorhanden sind, wie bestimmt durch Zellmorphologie.

[0029] **Fig.** 16: Immunozytochemische Anfärbung von Zervikalgewebe. Menschliche Gewebeproben, die Zervikalkrebs (**Fig. 16A**) oder gesundes Zervikalepithel (**Fig. 16B**) enthielten, wurden vom gleichen Patienten erhalten, in einen Paraffinschnitt eingebettet und der immunozytochemischen Anfärbung mit monoklonalem Antihuman-Notch-Antikörper, der gegen das TAN-1-Protein gerichtet ist, unterzogen. Die Bereiche, die bösartige Zellen enthielten (morphologisch bestimmt), zeigten relativ zu nicht bösartigen Zellen eine verstärkte Anfärbung. Von den nicht bösartigen Zellen färben sich das Bindegewebe und die Basalschicht des Epithels (das Stammzellen enthält), mit dem Anti-Notch-Antikörper.

[0030] **Fig.** 17: Die DNA (SEQ ID NO: 21) und die codierte Aminosäuresequenz (enthalten in SEQ ID NO: 19) des menschlichen Notch-Homologen hN. Die gesamte DNA-codierende Sequenz (sowie eine nicht codierende Sequenz) ausschließlich derjenigen, die das Starter-Met codiert, ist dargestellt. Die letzten 8 Nukleotide, die gezeigt sind (Nummern 9716–9723), sind Vektor- und keine hN-Sequenzen.

5. AUSFÜHRLICHE BESCHREIBUNG DER ERFINDUNG

[0031] Die Erfindung betrifft therapeutische Anwendungen und diagnostische Verfahren und Zusammensetzungen auf der Basis von Notch-Proteinen und -Nukleinsäuren. Die Erfindung stellt die Behandlung von Störungen des zellulären Werdegangs oder der zellulären Differenzierung durch Verabreichung einer erfindungsgemäßen therapeutischen Verbindung bereit. Solche therapeutischen Verbindungen (hier als "Therapeutika" bezeichnet) umfassen: Notch-Proteine und Analoge und Derivate (einschließlich von Fragmenten) hiervon; Antikörper dagegen; Nukleinsäuren, die die Notch-Proteine, Annaloge oder Derivate codieren; und Notch-Antisinn-Nukleinsäuren. Ebenfalls eingeschlossen sind Proteine und Derivate und Analoge hiervon, die in der Lage sind, die Wechselwirkungen eines Notch-Proteins mit einem anderen toporhythmischen Protein zu hemmen. Bei einer bevorzugten Ausführungsform kann ein erfindungsgemäßes Therapeutikum zur Behandlung eines kanzerösen Zustands oder zur Verhinderung des Fortschreitens eines präneoplastischen oder nicht-bösartigen Zustands (z. B. metaplastischer Zustand) zu einem neoplastischen oder bösartigen Zustand verabreicht werden. Bei einer weiteren speziellen Ausführungsform kann ein erfindungsgemäßes Therapeutikum zur Behandlung einer nervösen systemischen Störung, wie eine Nervenverletzung, oder einer degenerativen Erkrankung verabreicht werden. Bei noch einer weiteren speziellen Ausführungsform kann ein erfindungsgemäßes Therapeutikum zur beschleunigten Geweberegenerierung und Reparatur, zur Behandlung verschiedener Zustände verabreicht werden.

[0032] Bei einer Ausführungsform können Therapeutika, die die Notch-Funktion antagonisieren oder hemmen (im folgenden "antagonistische Therapeutika"), zur therapeutischen Wirkung verabreicht werden; Störungen, die so behandelt werden können, können durch in vitro Tests, wie sie in Abschnitt 5.1, nachstehend, beschrieben sind, identifiziert werden. Solche antagonistischen Therapeutika umfassen, sind jedoch nicht beschränkt auf, Notch-Antisinn-Nukleinsäuren, Anti-Notch-Neutralisationsantikörper, kompetitive Inhibitoren der Notch-Protein-Protein-Wechselwirkungen (z. B. ein Protein, das Notch-ELR-11 und Notch-ELR-12 enthält) und Moleküle, die eine intrazelluläre Notch-Funktion, wie diejenige, die durch die cdc10-Sequenzwiederholungen vermittelt wird, wie nachstehend ausführlich beschrieben, beeinflussen.

[0033] Bei einer weiteren Ausführungsform können Therapeutika, die die Notch-Funktion beschleunigen (im folgenden "agonistische Therapeutika" genannt), zur therapeutischen Wirkung verabreicht werden; Störungen, die so behandelt werden können, können durch in vitro Tests, wie in Abschnitt 5.1, nachstehend, beschrieben, identifiziert werden. Solche agonistischen Therapeutika umfassen, sind jedoch nicht beschränkt auf, Notch-Proteine und Derivate hiervon, die die intrazelluläre Domäne enthalten, Notch-Nukleinsäuren, die die Vorgenannten codieren.

[0034] Störungen des zellulären Werdegangs, insbesondere präkanzeröse Zustände, wie Metaplasie und Dysplasie, und hyperproliferative (z. B. Krebs) oder hypoproliferative Störungen, an denen anormale oder unerwünschte Expressions- oder Aktivitätsniveaus des Notch-Proteins beteiligt sind, können, wie nachstehend

ausführlicher beschrieben, durch den Nachweis solcher Niveaus diagnostiziert werden.

[0035] Nach einem bevorzugten Aspekt ist ein erfindungsgemäßes Therapeutikum ein Protein, bestehend aus mindestens einem Fragment (im folgenden hier als "Adhäsionsfragment" bezeichnet) der Proteine, die von toporhythmischen Genen codiert werden, das die Bindung an Notch-Proteine oder Adhäsionsfragmente hiervon vermittelt. Toporhythmische Gene, wie hier verwendet, sollen Notch-Gene bedeuten.

[0036] Die Erfindung stellt ferner Anwendungen, pharmazeutische Zusammensetzungen oder die Herstellung eines Medikaments aus einem menschlichen Notch-Protein, das von dem hN-Homolog codiert wird, und Proteinen, die die extrazelluläre Domäne des Notch-Proteins und Untersequenzen hiervon enthalten, bereit. Nukleinsäuren, die die Vorgenannten codieren, und rekombinante Zellen sind ebenfalls bereitgestellt.

[0037] Zur Klarheit der Offenbarung und nicht zur Einschränkung ist die ausführliche Beschreibung der Erfindung in die folgenden Unterabschnitte eingeteilt:

(i) Therapeutische Anwendungen;
(ii) Prophylaktische Anwendungen;
(iii) Demonstration der therapeutischen oder prophylaktischen Brauchbarkeit;
(iv) Therapeutische/prophylaktische Verabreichung und Zusammensetzungen;
(v) Antisinn-Regulierung der Notch-Expression;
(vi) Diagnostische Brauchbarkeit;
(vii) Notch-Nukleinsäuren;
(viii) Rekombinationsproduktion von Protein-Therapeutika;
(ix) Derivate und Analoge von Notch- und anderen toporhythmischen Proteinen;
(x) Tests auf Notch-Proteine, -Derivate und -Analoge; und
(xi) Antikörper gegen Notch-Proteine, -Derivate und -Analoge.

5.1. THERAPEUTISCHE ANWENDUNGEN

[0038] Wie oben angegeben, sind die antagonistischen erfindungsgemäßen Therapeutika, diejenigen Therapeutika, die eine Notch-Funktion antagonisieren oder hemmen. Solche antagonistischen Therapeutika werden besonders bevorzugt unter Anwendung bekannter zweckmäßiger in vitro Tests identifiziert, z. B. auf der Basis ihrer Fähigkeit zur Hemmung der Bindung von Notch- oder anderen Proteinen (siehe Abschnitte 6–8 hier), oder zur Hemmung einer beliebigen bekannten Notch-Funktion, wie in vitro getestet; obwohl auch Gentests (z. B. mit Drosophila) eingesetzt werden können. Bei einer bevorzugten Ausführungsform ist das antagonistische Therapeutikum ein Protein oder ein Derivat hiervon, das ein funktionell aktives Fragment, wie ein Adhäsionsfragment von Notch, enthält. Bei speziellen Ausführungsformen kann ein solches antagonistisches Therapeutikum diejenigen Adhäsionsproteine, die von den entsprechenden, in den Abschnitten 6 und 7 nachstehend beschriebenen Konstrukten codiert werden, oder Proteine, die die extrazelluläre Notch-Region enthalten, insbesondere ELR-11 und ERL-12, oder ein Antikörper hierfür, oder ein analoger/kompetitiver Inhibitor einer intrazellulären Signal-transduzierenden Notch-Region, eine Nukleinsäure, die zur Expression eines Notch-Adhäsionsfragments in der Lage ist, oder eine Notch-Antisinn-Nukleinsäure (siehe Abschnitt 5.5 hier) sein. Es sollte angemerkt werden, daß in bestimmten Fällen ein Notch-Adhäsionsfragment (oder möglicherweise weitere vermutete antagonistische Therapeutika) alternativ als antagonistisches Therapeutikum, je nach Entwicklungsgeschichte des Gewebes, das dem Therapeutikum ausgesetzt ist, wirken kann; vorzugsweise sollten geeignete in vitro oder in vivo Tests, wie nachstehend beschrieben, angewandt werden, um die Wirkung eines bestimmten Therapeutikums und ob seine Verabreichung zur Behandlung des betroffenen Gewebes indiziert ist, zu bestimmen.

[0039] Bei einer weiteren Ausführungsform der Erfindung wird eine Nukleinsäure, die ein Teil eines Notch-Gens enthält, als antagonistisches Therapeutikum zur Beschleunigung der Notch-Inaktivierung durch homologe Rekombination verwendet (Kollen und Smithies, 1989, Proc. Natl. Acad. Sci. USA 86: 8932–8935; Zijlstra et al., 1989, Nature 342: 435–438).

[0040] Die erfindungsgemäßen antagonistischen Therapeutika, wie vorstehend beschrieben, beschleunigen die Notch-Funktion.

[0041] Weitere Beschreibungen und Quellen für die erfindungsgemäßen Therapeutika werden hier in den Abschnitten 5.4 bis 5.8 gefunden.

[0042] Die erfindungsgemäßen agonistischen und antagonistischen Therapeutika besitzen therapeutische

Anwendbarkeit gegen Störungen des zellulären Werdegangs. Die antagonistischen Therapeutika können therapeutisch (einschließlich prophylaktisch) verabreicht werden: (1) bei Krankheiten oder Störungen, die ein Fehlen oder eine Abnahme des Niveaus der Notch-Funktion (relativ zum normalen oder gewünschten Niveau) betreffen, beispielsweise bei Patienten, bei denen das Notch-Protein fehlt, genetisch defekt, biologisch inaktiv oder wenig aktiv oder unterexprimiert ist; und (2) bei Krankheiten oder Störungen, wobei die Brauchbarkeit der Notch-Agonisten-Verabreichung durch in vitro (oder in vivo) Tests (siehe nachstehend) angezeigt wird. Fehlende oder verminderte Niveaus der Notch-Funktion können leicht nachgewiesen werden, z. B. durch Erhalt einer Patienten-Gewebeprobe (z. B. aus Biopsiegewebe) und durch ihr Testen in vitro auf Proteinkonzentrationen, Struktur und/oder Aktivität des exprimierten Notch-Proteins. Es können somit viele technische Standardverfahren angewandt werden, einschließlich, jedoch nicht beschränkt auf, Immuntests zum Nachweis und/oder zur Visualisierung des Notch-Proteins (z. B. Western-Blots, Immunfällung und anschließende Natriumdodecylsulfatpolyacrylamidgelelektrophorese, Immunozytochemie etc.; siehe auch die in Abschnitt 5.6 nachstehend aufgeführten Tests) und/oder Hybridisierungstests zum Nachweise der Notch-Expression durch Nachweis und/oder Visualisierung von Notch-mRNA (z. B. Northern-Tests, Dotblots, in situ Hybridisierung etc.).

[0043] In vitro Tests, die verwendet werden können, um zu bestimmen, ob die Verabreichung eines speziellen agonistischen oder antagonistischen Therapeutikums angezeigt ist, umfassen in vitro Zellkulturtests, wobei eine Patienten-Gewebeprobe in Kultur gezüchtet und einem Therapeutikum ausgesetzt wird, oder an die ein Therapeutikum anderweitig verabreicht, und die Wirkung eines solchen Therapeutikums auf die Gewebeprobe festgestellt wird. Bei einer Ausführungsform, wobei der Patient eine Bösartigkeit aufweist, wird eine Probe von Zellen einer solchen Bösartigkeit in Kultur ausplattiert oder gezüchtet, und die Zellen werden anschließend einem Therapeutikum ausgesetzt. Ein Therapeutikum, das das Überleben oder Wachstum der bösartigen Zellen hemmt (z. B. durch Beschleunigung der terminalen Differenzierung), wird zur therapeutischen Verwendung in vivo gewählt. Viele Standardtests der Technik können zur Bewertung eines solchen Überlebens und/oder Wachstums verwendet werden; beispielsweise kann die Zellproliferation durch Messen des ^{3}H-Thymidineinbaus, durch direkte Zellzählung, durch Nachweis von Änderungen der transkriptionalen Aktivität bekannter Gene, wie Proto-Onkogene (z. B. fos, myc) oder durch Zellzyklus-Marker geprüft werden; die Zell-Überlebensfähigkeit kann durch Trypanblau-Anfärbung geprüft werden, die Differenzierung kann visuell auf der Basis von Änderungen in der Morphologie etc. geprüft werden. Bei einem bestimmten Aspekt werden die bösartigen Zellkulturen getrennt (1) einem agonistischen Therapeutikum und (2) einem antagonistischen Therapeutikum ausgesetzt; das Ergebnis des Tests kann anzeigen, welcher Typ von Therapeutikum therapeutisch wirksam ist.

[0044] Bei einer weiteren Ausführungsform ist ein Therapeutikum zur Verwendung indiziert, das die gewünschte Wirkung zeigt, die Hemmung oder Beschleunigung des Zellwachstums in einer Patienten-Zellprobe aus Gewebe mit einer hyper- bzw. hypoproliferativen Störung bzw. mit einer vermuteten hyper- bzw. hypoproliferativen Störung. Solche hyper- oder hypoproliferativen Störungen umfassen, sind jedoch nicht beschränkt auf, diejenigen, die in den Abschnitten 5.1.1 bis 5.1.3 nachstehend beschrieben sind.

[0045] Bei einer weiteren speziellen Ausführungsform ist ein Therapeutikum zur Verwendung bei der Behandlung von Nervenverletzungen oder einer Nervensystem-Degenerationsstörung indiziert (siehe Abschnitt 5.1.2), das in vitro eine beschleunigte Nerven-Regenerierung/Neuriten-Extension aus Nervenzellen des betroffenen Patiententyps zeigt.

[0046] Zusätzlich ist auch die Verabreichung eines erfindungsgemäßen antagonistischen Therapeutikums bei Krankheiten oder Störungen indiziert, von denen bestimmt wurde oder von denen bekannt ist, daß an ihnen ein aktiv-dominanter Notch-Phänotyp ("Funktionsgewinn"-Mutationen) beteiligt ist. Die Verabreichung eines agonistischen Therapeutikums ist bei Krankheiten oder Störungen indiziert, von denen bestimmt wurde oder von denen bekannt ist, daß an ihnen ein negativ-dominanter Notch-Phänotyp beteiligt ist ("Funktionsverlust"-Mutationen). Wir haben die Funktionen verschiedener struktureller Domänen des Notch-Proteins in vivo durch epitope Expression einer Reihe von Drosophila-Notch-Deletionsmutanten unter dem hsp70-Hitzeschock-Promotor sowie unter visuell-spezifische Promotoren erforscht. Es wurden zwei Klassen von dominanten Phänotypen festgestellt, wovon eine für die Notch-Funktionsverlust-Mutationen und die andere für die Notch-Funktionsgewinn-Mutationen naheliegend sind. "Aktiv"-dominante Phänotypen beruhen auf der Überexpression eines Proteins, dem die meisten extrazellulären Sequenzen fehlen, während "negativ"-dominante Phänotypen auf der Überexpression eines Proteins, dem die meisten intrazellulären Sequenzen fehlen, beruhen. Weitere Ergebnisse zeigen, daß Notch als Rezeptor fungiert, dessen extrazelluläre Domäne Ligandenbindung vermittelt, die die Weiterleitung von Entwicklungssignalen durch die cytoplasmatische Domäne bewirken. Die beobachteten Phänotypen legten auch nahe, daß die cdc10/Ankyrin-Sequenzwiederholungsregion in der intrazellulären Domäne bei den Notch-vermittelten Signaltransduktionsereignissen eine essentielle Rolle

spielt (intrazelluläre Funktion).

[0047] Bei verschieden speziellen Ausführungsformen können in vitro Tests mit repräsentativen Zellen von Zelltypen, die an einer Störung im Patienten beteiligt sind, durchgeführt werden, um zu bestimmen, ob ein Therapeutikum eine gewünschte Wirkung auf solche Zelltypen ausübt.

[0048] Bei einer weiteren Ausführungsform werden auch Zellen aus Patienten-Gewebeproben, die vermutlich präneoplastisch sind, ausplattiert oder in vitro wachsen gelassen und einem Therapeutikum ausgesetzt. Das Therapeutikum, das einen Zellphänotyp bewirkt, der normaler (das heißt, weniger repräsentativ für einen präneoplastischen Zustand, einen neoplastischen Zustand, einen bösartigen Zustand oder einen transformierten Phänotyp) ist, wird zur therapeutischen Verwendung gewählt. Es können viele Standardtests aus der Technik angewandt werden, um zu bewerten, ob ein präneoplastischer Zustand, ein neoplastischer Zustand oder eine transformierter oder bösartiger Phänotyp vorliegt (siehe Abschnitt 5.2.1). Beispielsweise umfassen die Merkmale, die mit einem transformierten Phänotyp (eine Reihe von in vitro Merkmalen, die in vivo mit einer Tumor-erzeugenden Fähigkeit zusammenhängen) zusammenhängen, eine eher rundliche Zellmorphologie, eine losere Substrat-Haftung, einen Verlust an Kontakthemmung, einen Verlust an Verankerungsabhängigkeit, eine Freisetzung von Proteasen, wie von Plasminogenaktivator, einen erhöhten Zuckertransport, eine verminderte Serum-Erfordernis, eine Expression von fetalen Antigenen, ein Verschwinden des 250000 Dalton-Oberflächenproteins etc. (sieh Luria et al., 1978, Genral Virology, 3. Ausg., John Wiley & Sons, New York, Ss. 436–446).

[0049] Bei weiteren speziellen Ausführungsformen können die vorstehend beschriebenen in vitro Tests unter Verwendung einer Zellinie statt einer Zellprobe durchgeführt werden, die sich von dem speziellen, zu behandelnden Patienten ableitet, wobei die Zellinie auf Merkmal(en) beruht oder aufweist, die mit der bösartigen neoplastischen oder präneoplastischen Störung zusammenhängen, deren Behandlung oder Prävention gewünscht ist, oder sie stammt vom neutralen Zelltyp oder einem anderen Zelltyp, auf den erfindungsgemäß eine Wirkung erwünscht ist.

[0050] Die antagonistischen Therapeutika können therapeutisch (einschließlich prophylaktisch) verabreicht werden: (1) bei Krankheiten oder Störungen, die erhöhte (relativ zu normalen oder erwünschten) Niveaus der Notch-Funktion aufweisen, beispielsweise wobei das Notch-Protein überexprimiert oder überaktiv ist; und (2) bei Krankheiten oder Störungen, wobei in vitro (oder in vivo) Tests die Anwendbarkeit einer Notch-Antagonisten-Verabreichung anzeigen. Die erhöhten Niveaus der Notch-Funktion können leicht durch Methoden, wie durch diejenigen; die vorstehend beschrieben sind, durch Quantifizierung von Protein und/oder RNA nachgewiesen werden. In vitro Tests mit Zellen aus Patienten-Gewebeproben oder mit der entsprechenden Zellinie oder mit dem entsprechenden Zelltyp zur Bestimmung der therapeutischen Brauchbarkeit können wie vorstehend beschrieben durchgeführt werden.

5.1.1 BÖSARTIGKEITEN

[0051] Bösartige und präneoplastische Zustände, die wie vorstehend beschrieben auf die Wirksamkeit eines Eingreifens mit Hilfe eines Antagonisten- oder Agonisten-Therapeutika getestet und beim Feststellen eines Anzeichens für therapeutische Anwendbarkeit behandelt werden können, umfassen, sind jedoch nicht beschränkt auf, diejenigen, die nachstehend in den Abschnitten 5.1.1 und 5.2.1 beschrieben sind.

[0052] Bösartigkeiten und damit zusammenhängende Störungen, Zellen, deren Typ in vitro (und/oder in vivo) getestet werden kann und die bei Feststellung des entsprechenden Testergebnisses erfindungsgemäß behandelt werden können, umfassen, sind jedoch nicht beschränkt auf, diejenigen, die in Tabelle 1 aufgeführt sind (eine Übersicht über solche Störungen findet sich bei Fishman et al., 1985, Medicine, 2. Ausg., J. B. Lippincott Co., Philadelphia):

TABELLE 1

BÖSARTIGKEITEN UND DAMIT ZUSAMMENHÄNGENDE STÖRUNGEN

- Leukämie
  - akute Leukämie
    - akute lymphozytäre Leukämie
    - akute myelozytäre Leukämie
      - Myeloblastenleulämie
      - Promyelozytenleukämie
      - myelomonozytäre Leukämie
      - Monozytenleukämie
      - Erythroleukämie
      - chronische Leukämie
        - chronische myelozytäre (granulozytäre) Leukämie
        - chronische lymphozytäre Leukämie
  - Polycythaemia vera
  - Lymphonom
    - Morbus Hodgkin
    - Non-Morbus Hodgkin
  - multiples Myelom
  - Waldenström Makroglobulinämie
  - Schwerkettenkrankheit
  - solide Tumore
    - Sarcome und Karzinome
      - Fibrosarcom
      - Myxosarcom
      - Liposarcom
      - Chondrosarcom

Osteogenes Sarcom

Chordom

Angiosarcom

Endothelsarcom

Lymphangiosarcom

Lymphangioendothelsarcom

Synoviom

Mesotheliom

Ewing Tumor

Leiomyosarcom

Rhabdomyosarcom

Darmkarzinom

Pankreaskrebs

Brustkrebs

Eierstockkrebs

Prostatakrebs

Plattenepithelkarzinom

Basalzellenkarzinom

Adenokarzinom

Schweißdrüsenkarzinom

Talgdrüsenkarzinom

Papillenkarzinom

Papillenadenokarzinome

Zystadenokarzinom

Medullokarzinom

bronchiogenes Karzinom

Nierenzellkarzinom

Hepatom

Gallengangkarzinom

Chorionkarzinom

Seminom

Embryonalkarzinom

Willms Tumor

Zervikalkrebs

Hodentumor

Lungenkarzinom

kleinzelliges Karzinom

Blasenkarzinom

Epithelialkarzinom

Gliom

Astrozytom

Medulloblastom

Kraniopharyngiom

Ependymom

Pinealom

Hämangioblastom

akustisches Neurom

Oligodendrogliom

Menangiom

Melanom

Neuroblastom

Retinoblastom

[0053] Bei speziellen Ausführungsformen werden die Bösartigkeit oder dysproliferative Änderungen (wie Metaplasie und Dysplasie) in Epithelgeweben, wie die Gewebe in Zervix, Ösophagus und in der Lunge, behandelt oder präventiv behandelt.

[0054] Wie in den Beispielen im Abschnitt 10.1 nachstehend ausgeführt, zeigten die Bösartigkeiten von Brust, Darm und Zervix eine erhöhte Expression des menschlichen Notch relativ zu einem solchen nicht bösartigen Gewebe. Somit werden bei speziellen Ausführungsformen Bösartigkeiten von Brust, Darm oder Zervix durch Verabreichung einer wirksamen Menge eines erfindungsgemäßen antagonistischen Therapeutikums behandelt oder vorgebeugt. Das Vorliegen einer erhöhten Notch-Expression bei Brust, Darm und Zervikalkrebs legt nahe, daß viel mehr kanzeröse Zustände überreguliertes Notch aufweisen. Somit sind wir davon ausgegangen, daß viel mehr Krebsarten, z. B. Seminom, Melanom und Lungenkrebs, durch Verabreichung eines antagonistischen Therapeutikums behandelt oder vorgebeugt werden können.

5.1.2 NERVENSYSTEMSTÖRUNGEN

[0055] Nervensystemstörungen, die Zelltypen betreffen, die wie vorstehend beschrieben auf ein wirksames Eingreifen antagonistischer oder agonistischer Therapeutika getestet werden können und die dann bei Beobachtung eines Anzeichens für eine therapeutischen Brauchbarkeit behandelt werden können, umfassen, sind jedoch nicht beschränkt auf, Nervensystem-Verletzungen und Krankheiten oder Störungen, die entweder zu einer Axon-Verbindungsunterbrechung, einer Verminderung oder Degenerierung von Neuronen oder zu einer Demyelierung führen. Nervensystem-Läsionen, die bei einem Patienten (einschließlich einem menschlichen und nicht menschlichen Säugerpatienten) erfindungsgemäß behandelt werden können, umfassen, sind jedoch nicht beschränkt auf, die folgenden Läsionen entweder des zentralen (einschließlich Rückenmark, Gehirn) oder peripheren Nervensystems:

(i) traumatische Läsionen, einschließlich Läsionen auf Grund physikalischer Verletzungen oder Läsionen in Verbindung mit einem chirurgischen Eingriff, beispielsweise Läsionen, die einen Teil des Nervensystems

durchtrennen, oder Kompressionsläsionen;
(ii) ischämische Läsionen, wobei ein Sauerstoffmangel in einem Teil des Nervensystems zu einer neuronalen Verletzung oder Absterben führt, einschließlich Zerebralinfarkt oder -ischämie oder Rückenmarksinfarkt oder -ischämie;
(iii) bösartige Läsionen, wobei ein Teil des Nervensystems durch bösartiges Gewebe zerstört oder verletzt ist, das entweder eine mit dem Nervensystem verbundene Bösartigkeit oder eine Bösartigkeit ist, die von einem Nichtnervensystemgewebe herrührt;
(iv) infektiöse Verletzungen, wobei ein Teil des Nervensystems als Ergebnis einer Infektion, beispielsweise durch einen Abszeß oder bedingt durch eine Infektion mit dem Humanimmundefizienz-Virus, Herpeszoster- oder Herpessimplex-Virus oder mit der Lyme-Krankheit, Tuberkulose, Syphilis zerstört oder verletzt wird;
(v) degenerative Läsionen, wobei ein Teil des Nervensystems als Ergebnis eines degenerativen Prozesses, einschließlich, jedoch nicht beschränkt auf, Degenerierung durch Parkinson Krankheit, Alzheimer Krankheit, Huntington Chorea oder amyotrophe Lateralsklerose, zerstört oder verletzt wird;
(vi) Läsionen auf Grund von Ernährungskrankheiten oder -störungen, wobei ein Teil des Nervensystems durch Ernährungsstörungen oder Störungen des Metabolismus, einschließlich, jedoch nicht beschränkt auf, Vitamin B12-Mangel, Folsäuremangel, Wernicke Krankheit, Tabak-Alkohol-Amblyopie, Marchiafava-Bignami-Krankheit (Primärdegeneration des Corpus callosum) und alkoholische Zerebellum-Degenerierung, zerstört oder verletzt wird;
(vii) Neurologische Läsionen auf Grund von systemischen Krankheiten, einschließlich Diabetes (diabetische Neuropathie, Bellsche Lähmung), systemischer Lupus erythematodes, Karzinom oder Sarkoidose, jedoch nicht darauf beschränkt;
(viii) Läsionen auf Grund von toxischen Substanzen, einschließlich Alkohol, Blei oder bestimmter Neurotoxine; und
(ix) Demyelinierungsläsionen, wobei ein Teil des Nervensystems durch eine demyelierende Krankheit zerstört oder verletzt wird, einschließlich Multipler Sklerose, Humanimmundefizienzvirus-bedingter Myelopathie, transverser Myelopathie oder verschiedener Ethiologien, progressiver multifokaler Leukoenzephalopathie und Zentralbrücken-Myelinolyse, jedoch nicht darauf beschränkt.

[0056] Therapeutika, die erfindungsgemäß zur Behandlung einer Nervensystemstörung geeignet sind, können durch Testen der biologischen Wirksamkeit im Hinblick auf die Unterstützung des Überlebens oder der Differenzierung von Neuronen ausgewählt werden (siehe auch Abschnitt 5.1), beispielsweise, jedoch nicht einschränkend, können Therapeutika, die eine der folgenden Wirkungen hervorrufen, erfindungsgemäß geeignet sein:
(i) erhöhte Überlebensdauer von Neuronen in Kultur;
(ii) erhöhtes Sprießen von Neuronen in Kultur oder in vivo;
(iii) erhöhte Produktion eines Neuronen-bedingten Moleküls in Kultur oder in vivo, z. B. Cholinacetyltransferase oder Acetylcholinesterase, bezüglich Motorneuronen; oder
(iv) verminderte Symptome einer neuronalen Fehlfunktion in vivo.

[0057] Solche Wirkungen können durch jedes beliebige aus der Technik bekannte Verfahren gemessen werden. Bei bevorzugten nicht einschränkenden Ausführungsformen können ein erhöhtes Überleben von Neuronen durch das bei Arakawa et al. (1990, J. Neurosci. 10: 3507–3515) ausgeführte Verfahren gemessen werden; ein vermehrtes Sprießen von Neuronen kann durch die bei Pestronk et al. (1980, Exp. Neurol. 70: 65–82) oder Brown et al. (1981, Ann. Rev. Neurosci. 4: 17–42) ausgeführten Verfahren nachgewiesen werden; eine vermehrte Produktion von Neuronenbedingten Molekülen kann durch Biotests, enzymatische Tests, Antikörperbindung, Northern-Blot-Tests etc., in Abhängigkeit von dem zu messenden Molekül, gemessen werden; und eine Motorneuron-Fehlfunktion kann durch Testen der physikalischen Manifestation von Motorneuronen-Störungen, z. B. Schwäche der Motorneuronen-Leitungsgeschwindigkeit oder durch Funktionsuntüchtigkeit gemessen werden.

[0058] Bei einer speziellen Ausführungsform umfassen Motorneuronen-Störungen, die erfindungsgemäß behandelt werden können, Störungen, die Infarkt, Infektion, Toxinexposition, Trauma, Schäden durch einen operativen Eingriff, degenerative Erkrankungen oder Bösartigkeit, die die motorischen Neuronen, sowie andere Komponenten des Nervensystems beeinflussen können, sowie Störungen, die Neuronen selektiv beeinflussen, wie amyotrophe Lateralsklerose, und einschließlich progressiver spinaler Muskelatrophie, progressiver Bulbärparalyse, primärer Lateralsklerose, infantiler und juveniler Muskelatrophie, progressiver Bulbärparalyse in der Kindheit (Fazio-Londe Syndrom), Poliomyelitis und des Postpoliosyndroms und der erblich bedingten motorsensorischen Neuropathie (Charot-Mariesche Krankheit), jedoch nicht darauf beschränkt.

5.1.3 GEWEBEREPARATUR UND -REGENERIERUNG

[0059] Bei einer weiteren Ausführungsform der Erfindung wird ein erfindungsgemäßes Therapeutikum zur Beschleunigung der Geweberegenerierung und -reparatur verwendet, einschließlich der Behandlung gutartiger dysproliferativer Störungen, jedoch nicht darauf beschränkt. Spezielle Ausführungsformen betreffen die Behandlung von Leberzirrhose (ein Zustand, wobei die Vernarbung über die normalen Leber-Regenerierungsprozesse überhand genommen hat), Behandlung von Keloidbildung (hypertrohe Grindbildung) (Fehlfiguration der Haut, wobei der Grindbildungsprozess in die normale Erneuerung eingreift), Psoriasis (ein häufiger Hautzustand, der durch eine übermäßige Proliferation der Haut und eine verzögerte Bestimmung des ordnungsgemäßen zellulären Werdegangs gekennzeichnet ist) und Kahlköpfigkeit (ein Zustand, wobei terminaldifferenzierte Haarfollikel (ein Gewebe, das reich ist an Notch) nicht ordnungsgemäß funktionieren).

5.2 PROPHYLAKTISCHE ANWENDUNGEN

5.2.1 BÖSARTIGKEITEN

[0060] Die erfindungsgemäßen Therapeutika können zur Verhinderung der Progression zu einem neoplastischen oder bösartigen Zustand, einschließlich der in Tabelle 1 aufgeführten Störungen, jedoch nicht darauf beschränkt, verabreicht werden. Eine solche Verabreichung ist indiziert, wenn sich erweist, daß das Therapeutikum in Tests, wie nachstehend beschrieben, zur Behandlung oder Prävention solcher Störungen brauchbar ist. Eine solche prophylaktische Anwendung ist bei Zuständen indiziert, die bekanntlich oder vermutlich der Progression zu Neoplasie oder Krebs vorausgehen, insbesondere bei denen, wobei nicht-neoplastisches Zellwachstum, bestehend aus Hyperplasie, Metaplasie oder insbesondere Dysplasie, aufgetreten ist (ein Überblick über solche abnormen Wachstumsbedingungen findet sich bei Robbins und Angell, 1976, Basic Pathology, 2. Ausg., W. B. Saunders Co., Philadelphia, Ss. 68–79). Die Hyperplasie ist eine Form der kontrollierten Zellproliferation, die eine Zunahme in der Zellanzahl in einem Gewebe oder einem Organ ohne signifikante Änderung in Struktur oder Funktion betrifft. Nur als Beispiel geht die endometriale Hyperplasie oft dem Endometrialkrebs voraus. Die Metaplasie ist eine Form des kontrollierten Zellwachstums, wobei ein Typ einer ausgewachsenen oder vollständig differenzierten Zelle einen anderen Typ von ausgewachsener Zelle ersetzt. Metaplasie kann in Epithel- oder Bindegewebszellen auftreten. An der atypischen Metaplasie ist ein etwas ungeordnetes metaplastisches Epithel beteiligt. Die Dysplasie ist häufig ein Vorläufer von Krebs und wird hauptsächlich im Epithel festgestellt. Es ist die besonders ungeordnete Form des neoplastischen Zellwachstums und umfaßt einen Verlust der individuellen Zellgleichmäßigkeit und architektonischen Ausrichtung der Zellen. Dysplastische Zellen besitzen oft abnorm große, stark gefärbte Kerne und zeigen Pleomorphismus. Die Dysplasie tritt charakteristischerweise dann ein, wenn chronische Irritation oder Entzündung vorliegt und wird oft im Zervix, in den Atemwegen, der Mundhöhle und in der Gallenblase festgestellt.

[0061] Alternativ oder zusätzlich zum Vorliegen eines abnormen Zellwachstums, das als Hyperplasie, Metaplasie oder Dysplasie gekennzeichnet ist, kann das Vorliegen eines oder mehrerer Merkmale eines transformierten Phänotyps oder eines bösartigen Phänotyps, der in vivo oder in vitro von einer Zellprobe eines Patienten gezeigt wird, anzeigen, daß eine prophylaktische/therapeutische Verabreichung eines erfindungsgemäßen Therapeutikums erwünscht ist. Wie vorstehend erwähnt, umfassen solche Merkmale eines transformierten Phänotyps morphologische Änderungen, losere Substrat-Haftung, Verlust von Kontakthemmung, Verlust von Verankerungsabhängigkeit, Protease-Freisetzung, verstärkter Zuckertransport, geringeres Serum-Erfordernis, Expression fetaler Antigene, Verschwinden des 250000 Dalton-Zelloberflächenproteins etc. (siehe auch id., Ss. 84–90 bezüglich Merkmalen im Zusammenhang mit einem transformierten oder bösartigen Phänotyp).

[0062] Bei einer speziellen Ausführungsform sind Leukoplakie, eine gutartig erscheinende hyperplastische oder dysplastische Läsion des Epithels oder die Bowen Krankheit, ein Karzinom in situ, präneoplastische Läsionen eine Indikation dafür sind, daß der prophylaktische Eingriff erwünscht ist.

[0063] Bei einer weiteren Ausführungsform ist die Zystofibrose (zystische Hyperplasie, Mammadysplasie, insbesondere Adenose (gutartige Epithel-Hyperplasie)) eine Indikation dafür, daß der prophylaktische Eingriff erwünscht ist.

[0064] Bei anderen Ausführungsformen wird ein Patient, der eine oder mehrere der folgenden Prädispositionsfaktoren für Bösartigkeit zeigt, durch Verabreichung einer wirksamen Menge eines Therapeutikums behandelt: chromosomale Translokation, verbunden mit einer Bösartigkeit (z. B. Philadelphia-Chromosom für chronische myelogene Leukämie, t(14; 18) für Follikel-Lymphom, etc.), Familiäre Polypose oder Gardner Syndrom

(mögliche Vorläufer von Darmkrebs), gutartige monoklonale Gammopathie (möglicher Vorläufer für multiples Myelom) und eine Verwandschaft ersten Grades mit Personen, die von Krebs oder einer präkanzerösen Erkrankung befallen sind und ein Mendelsches (genetisches) Erbmuster zeigen (z. B. familiäre Polypose des Darms, Gardner Syndrom, erbliche Exostose, polyendokrine Adenomatose, medulläres Thyroid-Karzinom mit Amyloid-Produktion und Phäochromozytom, Peutz-Jeghers Syndrom, Von Recklinghausen-Neurofibromatose, Retinoblastom, Karotidkörpertumor, kutanes Melanokarzinom, intraokulares Melanokarzinom, Xerodermpigmentosum, ataxia telangiectasia, Chediak-Higashi Syndrom, Albinismus, aplastische Fanconi-Anämie und Bloom Syndrom; 8 siehe Robbins und Angell, 1976, Basic Pathology, 2. Ausg., W. B. Saunders Co., Philadelphia, Ss. 112–113) etc.).

[0065] Bei einer weiteren speziellen Ausführungsform wird ein erfindungsgemäßes antagonistisches Therapeutikum einem menschlichen Patienten zur Verhinderung der Progression von Brust-, Darm- oder Zervikalkrebs verabreicht.

5.2.2 WEITERE STÖRUNGEN

[0066] Bei weiteren Ausführungsformen kann ein erfindungsgemäßes Therapeutikum zur Verhinderung einer Nervensystemstörung, die in Abschnitt 5.1.2 beschrieben ist, oder einer anderen Störung (z. B. Leberzirrhose, Psoriasis, Keloide, Kahlköpfigkeit) die in Abschnitt 5.1.3 beschrieben ist, verabreicht werden.

5.3 DEMONSTRATION DER THERAPEUTISCHEN ODER PROPHYLAKTISCHEN BRAUCHBARKEIT

[0067] Die erfindungsgemäßen Therapeutika können in vivo auf die gewünschte therapeutische oder prophylaktische Wirksamkeit getestet werden. Beispielsweise können solche Verbindungen vor dem Testen an Menschen in geeigneten Tiermodellsystemen getestet werden, die Ratten, Mäuse, Hühner, Kühe, Affen, Kaninchen etc. einschließen, jedoch nicht darauf beschränkt sind. Für den in vivo Test, vor der Verabreichung an Menschen, kann jedes aus der Technik bekannt Tiermodellsystem angewandt werden.

5.4 THERAPEUTISCHE/PROPHYLAKTISCHE VERABREICHUNG UND ZUSAMMENSETZUNGEN

[0068] Die Erfindung stellt pharmazeutische Zusammensetzungen zur Behandlung (und Prophylaxe) durch Verabreichung an ein Subjekt, einer wirksamen Menge eines erfindungsgemäßen Therapeutikums bereit. Bei einem bevorzugten Aspekt ist das Therapeutikum im wesentlichen rein. Das Subjekt ist vorzugsweise ein Tier, einschließlich von Tieren, wie Kühe, Schweine, Hühner etc., jedoch nicht darauf beschränkt, und ist vorzugsweise ein Säuger und besonders bevorzugt ein Mensch.

[0069] Verschiedene Abgabesysteme sind bekannt und können zur Verababreichung eines erfindungsgemäßen Therapeutikums verwendet werden, z. B. Einkapselung in Liposomen, Mikroteilchen, Mikrokapseln, Expression durch rekombinante Zellen, Rezeptor vermittelte Endozytose (siehe z. B. Wu und Wu, 1987, J. Biol. Chem. 262: 4429–4432), Konstruktion einer therapeutischen Nukleinsäure als Teil eines retroviralen oder eines anderen Vektors etc. Verfahren zum Einbringen umfassen, sind jedoch nicht beschränkt auf, intradermale, intramuskuläre, intraperitoneale, intravenöse, subkutane, intranasale und orale Wege. Die Verbindungen können durch jeden zweckmäßigen Weg verabreicht werden, beispielsweise durch Infusion oder Bolus-Injektion, durch Absorption über epitheliale oder mukokutane Auskleidungen (z. B. Mundschleimhaut, Rektal- und Intestinalschleimhaut etc.) und können zusammen mit anderen biologisch wirksamen Mitteln verabreicht werden. Die Verabreichung kann systemisch oder lokal erfolgen. Zusätzlich kann es erwünscht sein, die erfindungsgemäßen pharmazeutischen Zusammensetzungen durch einen beliebigen geeigneten Weg in das Zentralnervensystem einzuschleusen, einschließlich der intraventrikulären und intrathekalen Injektion; die intraventrikuläre Injektion kann durch einen Intraventrikulärkatheter erleichtert werden, der beispielsweise an ein Reservoir, wie ein Ommaya-Reservoir, angeschlossen ist.

[0070] Bei einer speziellen Ausführungsform kann die Verabreichung der erfindungsgemäßen pharmazeutischen Zusammensetzungen lokal auf den Bereich, der der Behandlung bedarf, erwünscht sein. Dies kann beispielsweise durch lokale Infusion während eines chirurgischen Eingriffs, topisches Aufbringen, z. B. in Verbindung mit einem Wundpflaster nach dem chirurgischen Eingriffs, durch Injektion über einen Katheter, ein Zäpfchen oder über ein Implantat, wobei das Implantat aus einem porösen, nicht porösen oder aus gelartigem Material ist, einschließlich Membranen, wie sialastische Membranen oder Fasern, erreicht werden, ist jedoch nicht darauf beschränkt. Bei einer Ausführungsform kann die Verabreichung durch direkte Injektion an die Stelle (oder ehemalige Stelle) eines bösartigen Tumors oder von neoplastischem oder präneoplastischem Gewebe erfolgen.

[0071] Bei einer speziellen Ausführungsform wird die Verabreichung eines Therapeutikums in eine Notch-Expressionszelle durch Binden des Therapeutikums an ein Delta (oder an ein anderes toporhythmisches) Protein oder an einen Teil hiervon, der in der Lage ist, die Bindung an Notch zu vermitteln, erreicht. Der Kontakt einer Notch-exprimierenden Zelle mit dem gebundenen Therapeutikum führt zum Binden des gebundenen Therapeutikums über seinen Deltateil an Notch an der Oberfläche der Zelle und zur anschießenden Aufnahme des gebundenen Therapeutikums in die Notch-exprimierende Zelle.

[0072] Bei einer speziellen Ausführungsform, wobei ein Analoges einer intrazellulären, signaltransduzierenden Notch-Domäne so als Therapeutikum eingesetzt wird, daß die Notch-Signaltransduktion gehemmt werden kann, wird das Analoge vorzugsweise intrazellulär abgegeben (z. B. durch Expression aus einem Nukleinsäurevektor oder durch Binden an ein Delta-Protein, das zum Binden an Notch und anschließendes Binden und Internalisieren in der Lage ist, oder durch Rezeptor-vermitteltem Mechanismen).

[0073] Bei einer speziellen Ausführungsform, wobei das Therapeutikum eine Nukleinsäure ist, die ein Protein-Therapeutikum codiert, kann die Nukleinsäure in vivo zur beschleunigten Expression ihres codierten Proteins verabreicht werden, durch ihre Konstruktion als Teil eines entsprechenden Nukleinsäure-Expressionsvektors und durch ihr Verabreichen, so daß sie intrazellulär wird, z. B. durch Verwendung eines retroviralen Vektors (siehe U.S. Patentschrift Nr. 4 980 286) oder direkte Injektion oder durch Verwendung von Mikroteilchen-Bombardierung (z. B. eine Genpistole, Biolistik, Dupont) oder durch Überziehen mit Lipiden oder Zelloberflächen-Rezeptoren oder Transfektionsmitteln oder durch ihr Verabreichen in Bindungen an ein Homöobox-artiges Peptid, das bekanntlich in den Zellkern eindringt (siehe z. B. Joliot et al., 1991, Proc. Natl. Acad. Sci. USA 88: 1864–1868), etc. verabreicht werden. Alternativ kann ein Nukleinsäure-Therapeutikum intrazellulär eingeschleust und zur Expression durch homologe Rekombination in die Wirtszellen DNA eingebaut werden.

[0074] Bei speziellen Ausführungsformen, die die Behandlung oder Prävention bestimmter Störungen betreffen, werden vorzugsweise die folgenden Verabreichungsformen angewandt:

Krankheit	Bevorzugte Verabreichungsformen
Zervikalkrebs	topisch
Gastrointestinalkrebs	oral; intravenös
Lungenkrebs	Inhalation; intravenös
Leukämie	intravenös; extrakorporal
Metastasenkarzinome	intravenös; oral
Hirntumor	gezielt; intravenös; intrathekal
Leberzirrhose	oral; intravenös
Psoriasis	topisch
Keloide	topisch
Kahlköpfigkeit	topisch
Rüchkenmarksverletzung	gezielt; intravenös; intrathekal
Parkinson Krankheit	gezielt; intravenös; intrathekal
Motorneuronenkrankheit	gezielt; intravenös; intrathekal
Alzheimer Krankheit	gezielt; intravenös; intrathekal

[0075] Die Erfindung stellt auch pharmazeutische Zusammensetzungen bereit. Solche Zusammensetzungen umfassen eine therapeutisch wirksame Menge eines Therapeutikums und einen pharmazeutisch verträglichen Träger oder Hilfsstoff. Ein solcher Träger umfaßt Salzlösung, gepufferte Salzlösung, Dextrose, Wasser, Glycerin, Ethanol und Kombinationen hiervon, ist jedoch nicht darauf beschränkt. Der Träger und die Zusammensetzung können steril sein. Die Formulierung sollte der Verabreichungsweise angepaßt sein.

[0076] Die Zusammensetzung, kann falls gewünscht auch geringe Mengen an Netz- oder Emulgationsmitteln oder pH-Puffermitteln enthalten. Die Zusammensetzung kann eine flüssige Lösung, Suspension, Emulsion, Tablette, Pille, Kapsel, eine Langzeit-Formulierung oder ein Pulver sein. Die Zusammensetzung kann als Zäpfchen mit traditionellen Bindemitteln und Trägern, wie Triglyceriden, formuliert sein. Die orale Formulierung kann Standardträger, wie Mannit, Laktose, Stärke, Magnesiumstaerat, Natriumsaccharin, Cellulose, Magnesiumcarbonat etc. in pharmazeutischer Qualität enthalten.

[0077] Bei einer bevorzugten Ausführungsform wird die Zusammensetzung in Übereinstimmung mit den routinemäßigen Verfahrensweisen als pharmazeutische Zusammensetzung, die zur intravenösen Verabreichung an Menschen geeignet ist, formuliert. Typischerweise sind die Zusammensetzungen zur intravenösen Verab-

reichung Lösungen in einem sterilen, isotonen, wäßrigen Puffer. Wo notwendig, kann die Zusammensetzung auch ein Solubilisierungsmittel und ein Lokalanästhetikum, wie Lidocain, zur Schmerzlinderung an der Injektionsstelle enthalten. In der Regel werden die Bestandteile entweder getrennt oder miteinander vermischt in einer Dosierungseinheitsform, beispielsweise als trockenes lyophilisiertes Pulver oder als wasserfreies Konzentrat in einem hermetisch verschlossenen Behälter, wie eine Ampulle oder ein Beutel, wobei die Menge des Wirkstoffs angegeben ist, geliefert. Wenn die Zusammensetzung durch Infusion verabreicht werden soll, kann sie mit einem Infusionsbeutel, der steriles Wasser oder Salzlösung von pharmazeutischer Qualität enthält, abgegeben werden. Wenn die Zusammensetzung durch Injektion verabreicht wird, kann eine Ampulle mit sterilem Wasser zur Injektion oder mit Salzlösung bereitgestellt werden, so daß die Bestandteile vor der Verabreichung gemischt werden können.

[0078] Die erfindungsgemäßen Therapeutika können als neutrale Formen oder als Salzformen formuliert sein. Pharmazeutisch verträgliche Salze umfassen diejenigen, die mit freien Aminogruppen gebildet sind, wie diejenigen, die sich von Chlorwasserstoffsäure, Phosphorsäure, Essigsäure, Oxalsäure, Weinsäure etc. ableiten, und diejenigen, die mit freien Carboxylgruppen gebildet sind, wie diejenigen, die sich von Natrium-, Kalium-, Ammonium-, Calcium-, Eisen(III)-hydroxiden, Isopropylamin, Triethylamin, 2-Ethylaminoethanol, Histidin, Procain etc. ableiten.

[0079] Die Menge des erfindungsgemäßen Therapeutikums, die bei der Behandlung einer bestimmten Störung oder eines bestimmten Zustandes wirksam ist, hängt von der Natur der Störung oder des Zustands ab und kann durch klinische Standardtechniken bestimmt werden. Außerdem können gegebenenfalls in vivo Tests zur Unterstützung der Identifizierung der optimalen Dosisbereiche angewandt werden. Die exakte, in der Formulierung zu verabreichende Dosis hängt auch vom Verabreichungsweg und von der Schwere der Krankheit oder Störung ab und sollte nach dem Urteil des Arztes in Abhängigkeit vom Befinden eines jeden Patienten entschieden werden. Allerdings betragen geeignete Dosisbereiche zur intravenösen Verabreichung in der Regel etwa 20–500 µg wirksame Verbindung/kg Körpergewicht. Geeignete Dosisbereiche zur intranasalen Verabreichung betragen in der Regel etwa 0,01 pg/kg Körpergewicht bis 1 mg/kg Körpergewicht. Die wirksamen Dosen können aus Dosis-Reaktionskurven, die sich aus in vitro Testmodellsystemen oder aus Tiermodelltestsystemen ergeben, extrapoliert werden.

[0080] Suppositorien enthalten in der Regel den Wirkstoff im Bereich von 0,5 bis 10 Gew.-%.; orale Formulierungen enthalten vorzugsweise 10 bis 95% Wirkstoff.

[0081] Die Erfindung stellt auch eine pharmazeutische Packung oder ein pharmazeutisches Kit bereit, das einen oder mehrere Behälter enthält, die mit einem oder mehreren der Bestandteile der erfindungsgemäßen pharmazeutischen Zusammensetzungen befällt sind. Gegebenenfalls kann zusammen mit solchen Behälter(n) eine Notiz in der Form vorhanden sein, wie sie von einer Regierungsbehörde vorgeschrieben ist, die die Herstellung, die Verwendung oder den Verkauf der Pharmazeutika oder der biologischen Produkte regelt, wobei die Notiz die behördlich zugelassene Herstellung, Verwendung oder den Verkauf zur menschlichen Verabreichung erläutert.

5.5 ANTISINN-REGULIERUNG DER NOTCH-EXPRESSION

[0082] Die Erfindung stellt die therapeutische oder prophylaktische Verwendung von Nukleinsäuren oder von mindestens sechs Nukleotiden bereit, die zu einem Gen oder einer cDNA, die das Notch oder einen Teil hiervon codiert, Antisinn sind. "Antisinn", wie hier verwendet, bezieht sich auf eine Nukleinsäure, die mit einem Teil einer Notch-RNA (vorzugsweise mRNA) aufgrund von einiger Sequenzkomplementarität zu hybridisieren vermag. Solche Antisinn-Nukleinsäuren sind als erfindungsgemäße antagonistische Therapeutika brauchbar und können bei der Behandlung oder Prävention von Störungen, wie supra in Abschnitt 5.1 und in seinen Unterabschnitten beschrieben, verwendet werden.

[0083] Die erfindungsgemäßen Antisinn-Nukleinsäuren können Oligonukleotide, die doppelt- oder einzelsträngig RNA oder DNA oder eine Modifikation oder Derivat hiervon, die direkt an eine Zelle verabreicht werden können oder intrazellulär durch Transkription exogener, eingebauter Sequenzen produziert werden kann, sein.

[0084] Bei einer speziellen Ausführungsform können die erfindungsgemäß bereitgestellten Notch-Antisinn-Nukleinsäuren zur Behandlung von Tumoren oder anderen Störungen verwendet werden, wobei (in vitro oder in vivo) gezeigt werden kann, daß die Zellen des Tumortyps oder der Störung das Notch-Gen exprimieren. Eine solcher Beweis kann durch den Nachweis von Notch-RNA oder von Notch-Protein erfolgen.

[0085] Die Erfindung stellt ferner pharmazeutische Zusammensetzungen bereit, die eine wirksame Menge der erfindungsgemäßen Notch-Antisinn-Nukleinsäuren in einem pharmazeutisch verträglichen Träger enthalten, wie vorstehend in Abschnitt 5.4 beschrieben.

[0086] Bei einer weiteren Ausführungsform betrifft die Erfindung die Verwendung einer Antisinn-Notch-Nukleinsäure zur Herstellung eines Medikaments zur Hemmung der Expression einer Notch-Nukleinsäuresequenz in einer prokaryotischen oder eukaryotischen Zelle, die die Versorgung der Zelle mit einer wirksamen Menge einer Zusammensetzung, die eine erfindungsgemäße Antisinn-Notch-Nukleinsäure enthält, umfaßt.

[0087] Bei einer weiteren Ausführungsform kann die Identifizierung von Zellen, die funktionelle Notch-Rezeptoren exprimieren, durch Beobachtung der Fähigkeit von Notch zur "Rettung" solcher Zellen vor den cytotoxischen Wirkungen einer Notch-Antisinn-Nukleinsäure durchgeführt werden.

[0088] Notch-Antisinn-Nukleinsäuren und ihre Verwendungen sind ausführlich nachstehend beschrieben.

5.5.1 NOTCH-ANTISINN-NUKLEINSÄUREN

[0089] Die Notch-Antisinn-Nukleinsäuren bestehen aus mindestens sechs Nukleotiden, und es sind vorzugsweise Oligonukleotide (im Bereich von 6 bis etwa 50 Oligonukletiden). Bei speziellen Aspekten weist das Nukleotid mindestens 10 Nukleotide, mindestens 15 Nukleotide, mindestens 100 Nukleotide oder mindestens 200 Nukleotide auf. Die Oligonukleotide können DNA oder RNA oder chimäre Gemische oder Derivate oder modifizierte Versionen hiervon, einzelsträngig oder doppelsträngig sein. Das Oligonukleotid kann an der Basengruppierung, Zuckergruppierung oder am Phosphatskelett modifiziert sein. Das Oligonukleotid kann weitere Seitengruppen, wie Peptide oder Mittel, die den Transport durch die Zellmembran (siehe z. B. Letsinger et al., 1989, Proc. Natl. Acad. Sci. U. S. A. 86: 6553–6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84: 648–652; PCT-Veröffentlichungs-Nr. WO 88/09810, veröffentlicht am 15. Dezember 1988), oder die Blut-Hirn-Schranke (siehe z. B. PCT-Veröffentlichungs-Nr. WO 89/10134, veröffentlicht am 25. April 1988) erleichtern, durch Hybridisierung ausgelöste Spaltungsmittel (siehe z. B. Krol et al., 1988, Bio Techniques 6: 958–976) oder Interkalationsmittel (siehe z. B. Zon, 1988, Pharm. Res. 5: 539–549) enthalten.

[0090] Bei einem bevorzugten Aspekt der Erfindung ist ein Notch-Antisinn-Oligonukleotid, vorzugsweise von einzelsträngiger DNA, bereitgestellt. Bei einem besonders bevorzugten Aspekt enthält ein solches Oligonukleotid eine Sequenz, die zur Sequenz, die ELR 11 und ELR 12 von Notch, besonders bevorzugt von menschlichem Notch codiert, Antisinn ist. Das Oligonukleotid kann an jeder Position seiner Struktur mit aus der Technik allgemein bekannten Substituenten modifiziert sein.

[0091] Das Notch-Antisinn-Oligonukleotid kann mindestens eine modifizierte Basengruppierung enthalten, die aus der Gruppe ausgewählt ist, einschließlich, jedoch nicht beschränkt auf, 5-Fluorouracil, 5-Bromuracil, 5-Chloruracil, 5-Ioduracil, Hypoxanthin, Xantin, 4-Acetylcytosin, 5-(Carboxyhydroxylmethyl)uracil, 5-Carboxymethylaminomethyl-2-thiouridin, 5-Carboxymethylaminomethyluracil, Dihydrouracil, β-D-Galactosylqueosin, Inosin, N6-Isopentenyladenin, 1-Methylguanin, 1-Methylinosin, 2,2-Dimethylguanin, 2-Methyladenin, 2-Methylguanin, 3-Methylcytosin, 5-Methylcytosin, N6-Adenin, 7-Methylguanin, 5-Methylaminomethyluracil, 5-Methoxyaminomethyl-2-thiouracil, β-D-Mannosylqueosin, 5'-Methoxycarboxymethyluracil, 5-Methoxyuracil, 2-Methylthio-N6-isopentenyladenin, Uracil-5-oxyessigsäure(v), Wybutoxosin, Pseudouracil, Queosin, 2-Thiocytosin, 5-Methyl-2-thiouracil, 2-Thiouracil, 4-Thiouracil, 5-Methyluracil, Uracil-5-oxyessigsäuremethylester, Uracil-5-oxyessigsäure(v), 5-Methyl-2-thiouracil, 3-(3-Amino-3-N-2-carboxyprophyl)uracil, (acp3)w und 2,6-Diaminopurin.

[0092] Bei einer anderen Ausführungsform enthält das Oligonukleotid mindestens eine modifizierte Zuckergruppierung, die aus der Gruppe ausgewählt ist, die Arabinose, 2-Fluorarabinose, Xylulose und Hexose umfaßt, jedoch nicht darauf beschränkt ist.

[0093] Bei noch einer weiteren Ausführungsform enthält das Oligonukleotid mindestens ein modifiziertes Phosphatskelett, das aus der Gruppe, bestehend aus einem Phosphorthioat, einem Phosphordithioat, einem Phosphoramidothioat, einem Phosphoramidat, einem Phosphordiamidat, einem Methylphosphonat, einem Alkylphosphortriester und einem Formacetal hiervon, oder einem Analogen hiervon ausgewählt ist.

[0094] Bei einer noch weiteren Ausführungsform ist das Oligonukleotid ein α-anomeres Oligonukleotid. Ein α-anomeres Oligonukleotid bildet mit komplementärer RNA spezielle doppelsträngige Hybride, wobei die Stränge im Gegensatz zu den üblichen β-Einheiten parallel zueinander verlaufen (Gautier et al., 1987, Nucl.

Acids Res. 15: 6625–6641).

[0095] Das Oligonukleotid kann mit einem anderen Molekül konjugiert sein, z. B. einem Peptid, einem durch Hybridisierung ausgelösten Vernetzungsmittel, Transportmittel, durch Hybridisierung ausgelösten Spaltungsmittel etc.

[0096] Die erfindungsgemäßen Oligonukleotide können durch aus der Technik bekannte Standardverfahren synthetisiert werden, z. B. unter Verwendung eines automatischen DNA-Synthesegerätes (wie diejenigen, die von der Firma Biosearch, Applied Biosystems, im Handel sind etc.). Als Beispiele können Phosphorthioat-Oligonukleotide durch das Verfahren von Stein et al., (1988, Nucl. Acids Res. 16: 3209) synthetisiert werden, Methylphosphonat-Oligonukleotide können unter Verwendung von kontrollierten porösen Glas-Polymerträgern hergestellt werden (Sarin et al., 1988, Proc. Natl. Acad. Sci. U. S. A. 85: 7448–7451), etc.

[0097] Bei einer speziellen Ausführungsform enthält das Notch-Antisinn-Oligonukleotid katalytische RNA oder ein Ribozym (siehe z. B. PCT Internationale Veröffentlichung WO 90/11364, veröffentlicht am 4. Oktober 1990; Sarver et al. 1990, Science 247: 1222–1225). Bei einer anderen Ausführungsform ist das Oligonukleotid ein 2'-O-Methylribonukleotid (Inoue et al., 1987, Nucl. Acids Res. 15: 6131–6148) oder ein chimäres RNA-DNA-Analoges (Inoue et al., 1987, FEBS Lett. 215: 327–330).

[0098] Bei einer alternativen Ausführungsform wird die erfindungsgemäße Notch-Antisinn-Nukleinsäure intrazellulär durch Transkription aus einer exogenen Sequenz hergestellt. Beispielsweise kann ein Vektor in vivo so eingebaut werden, daß er von einer Zelle aufgenommen wird, wobei der Vektor oder ein Teil hiervon unter Produktion einer erfindungsgemäßen Antisinn-Nukleinsäure (RNA) transkribiert wird. Ein solcher Vektor enthält eine Sequenz, die die Notch-Antisinn-Nukleinsäure codiert. Ein solcher Vektor kann episomal bleiben oder kann chromosomal eingebaut werden, solang er unter Produktion der gewünschten Antisinn-RNA transkribiert werden kann. Solche Vektoren können durch DNA-rekombinationstechnische Standardverfahren konstruiert werden. Die Vektoren können Plasmidvektoren, virale Vektoren oder andere aus der Technik bekannte Vektoren sein, die zur Replikation und Expression in Sängerzellen verwendet werden. Die Expression der Sequenz, die die Notch-Antisinn-RNA codiert, kann durch jeden beliebigen aus der Technik bekannten Promotor erfolgen, der in Sängerzellen, vorzugsweise in menschlichen Zellen, wirkt. Solche Promotoren können induzierbar oder konstitutiv sein. Solche Promotoren umfassen: die frühe SV40-Promotorregion (Bernoist und Chambon, 1981, Nature 290: 304–310), der in der langen, terminalen 3'-Sequenzwiederholung von Rous Sarcomavirus enthaltene Promotor (Yamamoto et al., 1980, Cell 22: 787–797), der Herpes-Thymidinkinase-Promotor (Wagner, et al., 1981, Proc. Natl. Acad. Sci. U. S. A. 78: 1441–1445), die regulatorischen Sequenzen des Metallothionein-Gens (Brinster et al., 1982, Nature 296: 39–42) etc., sind jedoch nicht darauf beschränkt.

[0099] Die erfindungsgemäßen Antisinn-Nukleinsäuren enthalten eine Sequenz, die zu mindestens einem Teil eines RNA-Transkripts eines Notch-Gens, vorzugsweise eines menschlichen Notch-Gens, komplementär ist. Allerdings ist eine absolute Komplementarität, obwohl bevorzugt, nicht erforderlich. Eine Sequenz, "die zu mindestens einem Teil einer RNA komplementär ist", wie hier beschrieben, bedeutet eine Sequenz mit ausreichender Komplementarität, die mit der RNA unter Bildung eines stabilen Doppelstrangs zu hybridisieren vermag; im Falle von doppelsträngigen Notch-Antisinn-Nukleinsäuren kann somit ein EinzelStrang der Doppelstrang-DNA getestet werden, oder es können Dreifachstrang-Formationen getestet werden. Die Fähigkeit der Hybridisierung hängt sowohl vom Komplementaritätsgrad als auch von der Länge der Antisinn-Nukleinsäure ab. Je länger die hybridisierende Nukleinsäure, desto mehr Basen-Fehlpaarungen mit einer Notch-RNA kann sie in der Regel enthalten und immer noch einen stabilen Doppelstrang (oder gegebenenfalls Dreifachstrang) bilden. Ein Fachmann kann unter Anwendung von Standardverfahren zur Bestimmung des Schmelzpunkts des hybridisierten Komplexes einen annehmbaren Fehlpaarungsgrad garantieren.

5.5.2 THERAPEUTISCHE BRAUCHBARKEIT VON NOTCH-ANTISINN-NUKLEINSÄUREN

[0100] Die Notch-Antisinn-Nukleinsäuren können zur Behandlung (oder Verhinderung) von Bösartigkeiten eines Zelltyps verwendet werden, von dem gezeigt wurde, daß er Notch-RNA exprimiert. Bösartige, neoplastische und präneoplastische Zellen, die auf eine solche Expression getestet werden können, umfassen, sind jedoch nicht beschränkt auf, diejenigen, die supra in den Abschnitten 5.1.1 und 5.2.1 beschrieben sind. Bei einer bevorzugten Ausführungsform wird ein einzelsträngiges DNA-Antisinn-Notch-Oligonukleotid verwendet.

[0101] Bösartige(insbesondere Tumor)-Zelltypen, die Notch-RNA exprimieren, können durch verschiedene aus der Technik bekannte Verfahren identifiziert werden. Solche Verfahren umfassen, sind jedoch nicht beschränkt auf, Hybridisierung mit einer Notch-spezifischen Nukleinsäure (z. B. durch Northern-Hybridisierung,

Dotblot-Hybridisierung, in situ Hybridisierung) unter Beobachtung der Fähigkeit der RNA aus dem Zelltyp in vitro in Notch translatiert zu werden etc. Bei einem bevorzugten Aspekt kann primäres Tumorgewebe von einem Patienten vor der Behandlung auf die Notch-Expression getestet werden.

[0102] Die pharmazeutischen erfindungsgemäßen Zusammensetzungen (siehe Abschnitt 5.1.4), die eine wirksame Menge einer Notch-Antisinn-Nukleinsäure in einem pharmazeutisch verträglichen Träger enthalten, können an einen Patienten mit einer Bösartigkeit, die von einem Notch-RNA exprimierenden Typ ist, verabreicht werden.

[0103] Die Menge an Antisinn-Nukleinsäure, die bei der Behandlung einer bestimmten Störung oder eines bestimmten Zustands wirksam ist, hängt von der Natur der Störung oder des Zustandes ab und kann durch klinische Standardtechniken bestimmt werden. Wo möglich, ist es erwünscht, die Antisinn-Zytotoxizität des zu behandelnden Tumortyps in vitro und anschließend vor den Testen und vor der Anwendung beim Menschen in geeigneten Tiermodellsystemen zu bestimmen.

[0104] Bei einer speziellen Ausführungsform werden pharmazeutische Zusammensetzungen, die Notch-Antisinn-Nukleinsäuren enthalten, über Liposomen, Mikroteilchen oder Mikrokapseln verabreicht. Bei verschiedenen Ausführungsformen der Erfindung kann die Verwendung solcher Zusammensetzungen zum Erreichen einer Langzeit-Freisetzung der Notch-Antisinn-Nukleinsäuren geeignet sein. Bei einer speziellen Ausführungsform kann es wünschenswert sein, Liposomen, die über Antikörper auf spezielle identifizierbare Tumor-Antigene gerichtet sind, zu verwenden (Leonetti et al., 1990, Proc. Natl. Acad. Sci. U. S. A. 87: 2448–2451; Renneisen et al., 1990, J. Biol. Chem. 265: 16337–16342).

5.6 DIAGNOSTISCHE BRAUCHBARKEIT

[0105] Notch-Proteine, Analoge und Derivate und Sequenzen hiervon, Notch-Nukleinsäuren (und hierzu komplementäre Sequenzen), Anti-Notch-Antikörper finden in der Diagnostik Anwendung. Solche Moleküle können bei Tests, wie Immuntests, zum Nachweis, zur Prognose, Diagnose oder Überwachung verschiedener, die Notch-Expression betreffender Zustände, Krankheiten und Störungen, oder zur Überwachung der Behandlung hiervon, eingesetzt werden. Insbesondere wird ein solcher Immuntest durch ein Verfahren durchgeführt, das den Kontakt einer Probe, die aus einem Patienten mit einem Anti-Notch-Antikörper stammt, unter solchen Bedingungen, daß eine immunspezifische Bindung auftreten kann, und den Nachweis oder das Messen der Menge einer immunspezifischen Bindung durch den Antikörper umfaßt. Bei einer speziellen Ausführungsform kann ein Antikörper gegen Notch zum Test einer Patienten-Gewebe- oder Serumprobe auf das Vorliegen von Notch verwendet werden, wo eine anormaler Notch-Konzentration ein Anzeichen für einen Krankheitszustand ist.

[0106] Die Immuntests, die angewandt werden können, umfassen, sind jedoch nicht beschränkt auf, kompetitive und nicht-kompetitive Testsysteme unter Verwendung von Techniken, wie Western-Blots, Radioimmuntests, ELISA (Enzyme linked immunosorbent assay). "Sandwich"-Immuntests, Immunfällungstests, Fällungsreaktionen, Geldiffusionsfällungsreaktionen, Immundiffusionstests, Agglutinationstests, Komplement-Fixierungstests, Immunradiometrietests, Fluoreszenzimmuntests, Protein A-Immuntests, um nur einige aufzuzählen.

[0107] Notch-Gene und verwandte Nukleinsäuresequenzen und Untersequenzen, einschließlich der Komplementärsequenzen und anderer toporhythmischer Gensequenzen, können ebenfalls bei Hybridisierungstests verwendet werden. Notch-Nukleinsäuresequenzen oder Untersequenzen hiervon, die mindestens etwa 8 Nukleotide enthalten, können als Hybridisierungssonden verwendet werden. Hybridisierungstests können zum Nachweis, zur Prognose, Diagnose oder zur Aufzeichnung von Zuständen, Störungen oder Krankheitszuständen verwendet werden, die mit anormalen Änderungen der Notch-Expression und/oder Aktivität, wie vorstehend beschrieben, einhergehen. Insbesondere wird ein solcher Hybridisierungstest durch ein Verfahren durchgeführt, das den Kontakt einer Probe, die Nukleinsäure enthält, mit einer Nukleinsäuresonde, die in der Lage ist, an Notch-DNA oder RNA zu hybridisieren, unter solchen Bedingungen, daß die Hybridisierung auftreten kann, und den Nachweis oder das Messen einer resultierenden Hybridisierung umfaßt.

[0108] Wie in den Beispielen in Abschnitt 10.1 nachstehend ausgeführt, tritt eine verstärkte Notch-Expression bei menschlichem Brustkrebs, Darmkrebs und Zervikalkrebs auf. Demnach wird bei speziellen Ausführungsformen menschlicher Brustkrebs, Darmkrebs oder Zervikalkrebs oder prämaligne Änderungen in solchen Geweben durch den Nachweis erhöhter Notch-Expression (oder Menge) in Patientenproben relativ zu dem Niveau der Notch-Expression (oder Menge) je nach Fall in einer analogen nicht malignen oder nicht prämalignen

Probe (vom Patienten oder von einer anderen Person, wie experimentell bestimmt, oder wie es als Standardniveau in solchen Proben bekannt ist) diagnostiziert.

[0109] Bei einer Ausführungsform befindet sich das Notch-Protein (oder ein Derivat mit Notch-Antigenität), das nachgewiesen oder gemessen wird, auf der Zelloberfläche. Bei einer anderen Ausführungsform ist das Notch-Protein (oder ein Derivat) ein zellfreies, lösliches Molekül (z. B. wird es im Blut oder in einer Serumprobe gemessen) oder ist intrazellulär. Ohne daß die Absicht besteht, mechanistisch festgelegt zu werden, glauben die Anwender, daß zellfreies Notch von der Sekretion oder dem Abstoßen von der Zelloberfläche herrührt. Bei noch einer anderen Ausführungsform werden lösliche, zelloberflächliche und intrazelluläre Mengen an Notch-Protein oder Derivat nachgewiesen oder gemessen.

5.7 NOTCH-NUKLEINSÄUREN

[0110] Die erfindungsgemäßen Therapeutika, die Notch-Nukleinsäuren oder Notch-Antisinn-Nukleinsäuren sind, sowie Nukleinsäuren, die Protein-Therapeutika codieren, umfassen diejenigen, die nachstehend beschrieben sind, die durch aus der Technik bekannte Verfahren und insbesondere, wie nachstehend beschrieben, erhalten werden können.

[0111] Bei bestimmten Aspekten stellt die Erfindung Aminosäurensequenzen von Notch, vorzugsweise von menschlichem Notch und Fragmente und Derivate hiervon, die eine Antigen-Determinante enthalten (d. h. Erkennung durch einen Antikörper ist möglich) oder die funktionell aktiv sind, sowie Nukleinsäuresequenzen die die Vorgenannten codieren, bereit. "Funktionell aktives" Material, wie hier verwendet, bezieht sich auf das Material, das eine oder mehrere bekannte funktionelle Aktivitäten aufweist, die mit dem Notch-Vollängen-Protein (Wildtyp) im Zusammenhang stehen, z. B. Bindung an Delta, Bindung an Serrate, Bindung an einen anderen Notch-Liganden, Antigeniät (Bindung an Anti-Notch-Antikörper) etc.

[0112] Bei speziellen Ausführungsformen stellt die Erfindung Fragmente eines Notch-Proteins bereit, bestehend aus mindestens 40 Aminosäuren oder aus mindestens 75 Aminosäuren. Bei anderen Ausführungsformen enthalten die Proteine die intrazellulären Domäne, Transmembranregion, extrazelluläre Domäne, cdc10-Region, Notch/lin-12-Sequenzwiederholungen oder die EGF-homologen Sequenzwiederholungen, oder einer Kombination der Vorgenannten, von einem Notch-Protein, oder bestehen im wesentlichen daraus. Fragmente oder Proteine, die Fragmente enthalten, denen einige oder sämtliche EGF-homologen Sequenzwiederholungen von Notch fehlen, sind ebenfalls bereitgestellt. Nukleinsäuren, die Vorgenanntes codieren, sind bereitgestellt.

[0113] Bei anderen spezifischen Ausführungsformen stellt die Erfindung Nukleotidsequenzen und Untersequenzen von Notch, vorzugsweise von menschlichem Notch, bereit, bestehend aus mindestens 25 Nukleotiden, mindestens 50 Nukleotiden oder mindestens 150 Nukleotiden. Nukleinsäuren, die die vorstehend beschriebenen Proteine und Proteinfragmente codieren, sowie Nukleinsäuren, die zu solchen Nukleinsäuren komplementär sind und in der Lage sind, mit ihnen zu hybridisieren, sind bereitgestellt. Bei einer Ausführungsform kann eine solche komplementäre Sequenz zu einer Notch-cDNA-Sequenz von mindestens 25 Nukleotiden oder von mindestens 100 Nukleotiden komplementär sein. Bei einem bevorzugten Aspekt verwendet die Erfindung cDNA-Sequenzen, die menschliches Notch oder einen Teil hiervon codieren. Bei einer speziellen Ausführungsform sind solche Sequenzen des menschlichen Notch-Gens oder der cDNA so, wie sie in den Plasmiden, hN3k, hN4k oder hN5k (siehe Abschnitt 9, nachstehend) oder in dem hierzu entsprechenden Gen enthalten sind; eine solche menschliche Notch-Proteinsequenz kann so sein, wie in den **Fig.** 10 (SEQ ID NO: 11) oder 11 (SEQ ID NO: 13) gezeigt. Bei anderen Ausführungsformen weist die Notch-Nukleinsäure und/oder ihr codiertes Protein mindestens einen Teil der in einer der folgenden Publikationen gezeigten Sequenz auf: Wharton et al., 1985, Cell 43: 567–581 (Drosophila-Notch); Kidd et al., 1986, Mol. Cell. Biol. 6: 3094–3108 (Drosophila-Notch); Coffman et al., 1990, Science 249: 1438–1441 (Xenopus Notch); Ellisen et al., 1991, Cell 66: 649–661 (ein menschliches Notch). Bei einem anderen Aspekt sind die Sequenzen von menschlichem Notch diejenigen, die die menschlichen Notch-Aminosäuresequenzen oder einen Teil hiervon, wie in **Fig.** 13 gezeigt, codieren. Bei einem bestimmten Aspekt sind die menschlichen Notch-Sequenzen diejenigen des hN-Homologen (zum Teil durch das Plasmid hN5k dargestellt) oder des TAN-1-Homologen.

[0114] Bei einer Ausführungsform der Erfindung betrifft die Erfindung das menschliche Vollängen-Notch-Protein, das von dem hN-Homologen, wie in **Fig.** 13 beschrieben, codiert wird, bei dem die Signalsequenz sowohl enthalten ist (d. h. das Vorläuferprotein, Aminosäuren 1–2169) als auch fehlt (d. h. das reife Protein, Aminosäuren ca. 26–2169), sowie Teile des Vorgenannten (z. B. extrazelluläre Domäne, die EGF-homologe Sequenzwiederholungsregion, die EGF-artigen Sequenzwiederholungen 11 und 12, die cdc-10/Ankyrin-Sequenzwie-

derholungen, etc.) und Proteine, die das Vorgenannten enthalten, sowie Nukleinsäuren, die das Vorgenannten codieren.

[0115] Wie leicht offensichtlich ist, soll, wie hier verwendet, eine "Nukleinsäure, die ein Fragment oder einen Teil eines Notch-Proteins codiert", so aufgebaut sein, daß eine Nukleinsäure gemeint ist, die nur das aufgeführte Fragment oder ein Teil des Notch-Proteins und keine anderen Teile des Notch-Proteins codiert.

[0116] Bei einem bevorzugten, allerdings nicht einschränkenden Aspekt der Erfindung kann eine menschliche Notch-DNA-Sequenz kloniert und durch das in Abschnitt 9, nachstehend, beschriebene Verfahren sequenziert werden.

[0117] Bei einem weiteren bevorzugten Aspekt wird die PCR zur Amplifikation der gewünschten Sequenz in der Bibliothek vor der Selektion angewandt. Beispielsweise können Oligonukleotid-Primer, die einen Teil der Adhäsionsdomänen darstellen, die von einem Homologen des gewünschten Gens codiert werden, als Primer bei der PCR eingesetzt werden.

[0118] Die obigen Verfahren sollen für die folgende allgemeine Beschreibung der Verfahren, durch die Notch-Klone erhalten werden können, nicht einschränkend sein.

[0119] Jede eukaryotische Zelle kann potentiell als Nukleinsäurequelle für das molekulare Klonieren des Notch-Gens dienen. Die DNA kann durch aus der Technik bekannte Standardverfahrensweisen aus klonierter DNA (z. B. einer DNA-"Bibliothek"), durch chemische Synthese, cDNA-Klonierung oder durch Klonieren von Genom-DNA oder von Fragmenten hiervon, die aus der gewünschten menschlichen Zelle gereinigt wurden (siehe beispielsweise Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, 2. Ausg. Cold Spring Harbor, New York; Glover, D. M. (Hrsg.), 1985, DNA Cloning: A Practical Approach, MRL Press Ltd., Oxford, U. K. Vol. I, II.) erhalten werden. Klone, die aus Genom-DNA stammen, können zusätzlich zu den codierenden Regionen regulatorische und Intron-DNA-Regionen enthalten; Klone, die von cDNA stammen, enthalten nur Exon-Sequenzen. Ohne Rücksicht auf die Quelle sollte das Gen molekular in einen geeigneten Vektor zur Propagation des Gens kloniert werden.

[0120] Beim molekularen Klonieren des Gens aus Genom-DNA werden DNA-Fragmente erzeugt, von denen einige das gewünschte Gen codieren. Die DNA kann unter Verwendung verschiedener Restriktionsenzyme an bestimmten Stellen gespalten werden. Alternativ kann DNAse in Gegenwart von Mangan zur Fragmentierung der DNA verwendet werden, oder die DNA kann physikalisch geschert werden, wie beispielsweise durch Ultraschall. Anschließend können die linearen DNA-Fragmente durch Standardtechniken, einschließlich, jedoch nicht beschränkt auf, Agarose- und Polyacrylamidgelelektrophorese und Säulenchromatographie, der Größe nach aufgetrennt werden.

[0121] Nach Erzeugung der DNA-Fragmente kann die Identifizierung des speziellen DNA-Fragments, das das gewünschte Gen enthält, auf einer Reihe von Wegen erreicht werden. Wenn beispielsweise ein Teil eines Notch(von einer beliebigen Spezies)-Gens oder seine spezielle RNA oder ein Fragment hiervon, z. B. die Adhäsionsdomäne, verfügbar ist und gereinigt und markiert werden kann, können die erzeugten DNA-Fragmente durch Nukleinsäure-Hybridisierung mit der markierten Sonde durchmustert werden (Benton, W. und Davis, R., 1977, Science 196, 180; Grunstein, M. und Hogness, D., 1975, Proc. Natl. Acad. Sci. U. S. A. 72, 3961). Diejenigen DNA-Fragmente mit nennenswerter Homologie zur Sonde hybridisieren. Es ist auch möglich, das entsprechende Fragment durch Restriktionsenzymverdauung(en) zu identifizieren und die Fragmentgrößen mit denjenigen zu vergleichen, die nach einer bekannten Restriktionskarte, sofern verfügbar, erwartet werden. Eine weitere Selektion kann auf der Grundlage der Eigenschaften des Gens erfolgen. Alternativ kann das Vorliegen des Gens durch Tests auf der Basis der physikalischen, chemischen oder immunologischen Eigenschaften seines exprimierten Produkts nachgewiesen werden. Beispielsweise können cDNA-Klone oder DNA-Klone gewählt werden, die durch Hybridisierung die richtigen mRNAs selektieren, die ein Protein produzieren, das z. B. ein ähnliches oder identische elektrophoretische Migrations-, isoelektrisches Fokussierungsverhalten, proteolytische Verdau-Muster, in vitro Aggregationsaktivität ("Adhäsionsvermögen") oder Antigen-Eigenschaften, wie sie für Notch bekannt sind, aufweist. Wenn ein Antikörper auf Notch verfügbar ist, kann das Notch-Protein durch Binden des markierten Antikörpers an die vermeintlichen Notch-synthetisierenden Klone durch ein Verfahren vom Typ ELISA (enzyme-linked immunosorbent assay) identifiziert werden.

[0122] Das Notch-Gen kann auch durch mRNA-Selektion durch Nukleinsäure-Hybridisierung und anschließende in vivo Translation identifiziert werden. Bei diesem Verfahren werden Fragmente zur Isolierung komplementärer mRNAs durch Hybridisierung verwendet. Solche DNA-Fragmente können verfügbare, gereinigte

Notch-DNA einer anderen Spezies (z. B. Drosophila) darstellen. Immunfällungsanalyse oder Funktionstests (z. B. Aggregationsvermögen in vitro; siehe Beispiele nachstehend) mit den in vitro Translationsprodukten der isolierten Produkte der isolierten mRNAs identifizieren die mRNA und darum die komplementären DNA-Fragmente, die die gewünschten Sequenzen enthalten. Zusätzlich können spezielle mRNAs durch Adsorption von aus Zellen isolierten Polysomen unter Immobilisierung von Antikörpern, die spezifisch gegen Notch- oder Deltaprotein gerichtet sind, selektiert werden. Eine radioaktiv markierte Notch-cDNA kann unter Verwendung der gewählten mRNA (aus den adsorbierten Polysomen) als Templat synthetisiert werden. Sodann kann die radioaktiv markierte mRNA oder cDNA als Sonde zur Identifizierung der Notch-DNA-Fragmente unter anderen genomischen DNA-Fragmenten verwendet werden.

[0123] Alternativen zur Isolierung der Notch-Genom-DNA umfassen, sind jedoch nicht beschränkt auf, chemische Synthese der Gensequenz selbst aus einer bekannten Sequenz oder Überführen von cDNA in die mRNA, die das Notch-Gen codiert. Beispielsweise kann RNA zum cDNA-Klonieren des Notch-Gens aus Zellen, die Notch exprimieren, isoliert werden. Weitere Methoden sind möglich und liegen im Umfang der Erfindung.

[0124] Das identifizierte und isolierte Gen kann dann in einen geeigneten Klonierungsvektor inseriert werden. Eine große Anzahl aus der Technik bekannter Vektor-Wirt-Systeme kann verwendet werden. Mögliche Vektoren umfassen, sind jedoch nicht beschränkt auf, Plasmide oder modifizierte Viren, allerdings muß das Vektorsystem mit der verwendeten Wirtszelle kompatibel sein. Solche Vektoren umfassen, sind jedoch nicht beschränkt auf, Bakteriophagen, wie Lambda-Derivate oder Plasmide, wie PBR322, oder pUC-Plasmidderivate. Die Insertion in einen Klonierungsvektor kann beispielsweise durch Ligation des DNA-Fragments in einem Klonierungsvektor, der komplementäre, klebrige Enden aufweist, erreicht werden. Wenn allerdings die komplementären, zur Fragmentierung der DNA verwendeten Restriktionsstellen nicht in dem Klonierungsvektor vorhanden sind, können die Enden der DNA-Moleküle enzymatisch modifiziert werden. Alternativ kann eine beliebige gewünschte Stelle durch Ligation von Nukleotidsequenzen (Linker) an die DNA-Termini produziert werden; diese ligierten Linker können spezielle chemisch synthetisierte Oligonukleotide enthalten, die Restriktionsendonuklease-Erkennungssequenzen codieren. Bei einem alternativen Verfahren, können der gespaltene Vektor und das Notch-Gen durch homopolymere Schwanzbildung modifiziert werden. Rekombinante Moleküle können über Transformation, Transfektion, Infektion, Elektroporation, etc. in Wirtszellen eingeschleust werden, so daß viele Kopien der Gensequenz erzeugt werden.

[0125] Bei einem alternativen Verfahren kann das gewünschte Gen identifiziert und nach Insertion in einen geeigneten Klonierungsvektor "shot gun"-artig isoliert werden. Die Anreicherung mit dem gewünschten Gen, beispielsweise durch Größenfraktionierung, kann vor der Insertion in den Klonierungsvektor erfolgen.

[0126] Bei speziellen Ausführungsformen ermöglicht die Transformation von Wirtszellen mit rekombinanten DNA-Molekülen, in die das isolierte Notch-Gen, die cDNA oder die synthetisierte DNA-Sequenz eingeschleust ist, die Erzeugung von Mehrfachkopien des Gens. Somit kann das Gen durch Züchten von Transformanten, Isolieren der rekombinanten DNA-Moleküle aus den Transformanten und wenn notwendig, Gewinnung des inserierten Gens aus der isolierten rekombinanten DNA in großen Mengen erhalten werden.

[0127] Die erfindungsgemäß bereitgestellten Notch-Sequenzen umfassen diejenigen Nukleotidsequenzen, die praktisch die gleichen Aminosäuresequenzen codieren, wie sie in nativem Notch-Protein vorkommen, und die codierten Aminosäuresequenzen mit funktionell equivalenten Aminosäuren, alles wie in Abschnitt 5.6, nachstehend für Notch-Derivate beschrieben.

5.8 REKOMBINANTE HERSTELLUNG VON PROTEINTHERAPEUTIKA

[0128] Die Nukleinsäure, die ein erfindungsgemäßes Protein-Therapeutikum codiert, kann in einen geeigneten Expressionsvektor, das heißt, einen Vektor, der die zur Transkription und Translation der inserierten Protein-Kodierungssequenz notwendigen Elemente enthält, inseriert werden. Die notwendigen Transkriptions- und Translationssignale können auch von dem nativen toporhythmischen Gen und/oder seinen Flankierungsregionen bereitgestellt werden. Eine Vielzahl von Wirt-Vektor-Systemen kann zur Expression der Protein-codierenden Sequenz verwendet werden. Diese umfassen, sind jedoch nicht beschränkt auf, Säugerzellsysteme, die mit einem Virus infiziert worden sind (z. B. Vacciniavirus, Adenovirus, etc.); Insektenzellen, die mit einem Virus infiziert worden sind (z. B. Baculovirus); Mikroorganismen, wie Hefe, die Hefevektoren enthalten, oder Bakterien, die mit Bakteriophagen-DNA, Plasmid-DNA oder Cosmid-DNA transformiert worden sind. Die Expressionselemente von Vektoren variieren in ihrer Stärke und Spezifität. In Abhängigkeit von dem verwendeten Wirt-Vektor-System kann eine beliebige Anzahl von geeigneten Transkriptions- und Translationselementen

verwendet werden. Bei einer speziellen Ausführungsform wird der klebrige Teil des Notch-Gens, z. B. die codierenden EGF-artigen Sequenzwiederholungen (ELR)-11 und -12, exprimiert. Bei weiteren speziellen Ausführungsformen wird das menschliche Notch-Gen oder eine Sequenz, die einen funktionell aktiven Teil des menschlichen Notch codiert, exprimiert.

[0129] Jedes der zuvor zur Insertion von DNA-Fragmenten in einen Vektor beschriebenen Verfahren kann zur Konstruktion von Expressionsvektoren angewandt werden, die ein chimäres Gen enthalten, bestehend aus geeigneten Transkriptions/Translations-Kontrollsignalen und den Protein-codierenden Sequenzen. Diese Verfahren können in vitro DNA-Rekombinationstechniken und Synthesetechniken und in vivo Rekombinationen (genetische Rekombination) einschließen. Die Expression einer Nukleinsäuresequenz, die ein Notch-Protein oder ein Peptidfragment codiert, kann durch eine zweite Nukleinsäuresequenz reguliert werden, so daß das Notch-Protein oder -Peptid in einem mit dem rekombinanten DNA-Molekül transformierten Wirt exprimiert wird. Beispielsweise kann die Expression eines Notch-Proteins durch ein beliebiges, aus der Technik bekanntes Promotor/Enhancer-Element kontrolliert werden. Promotoren, die zur Kontrolle der togorhythmischen Genexpression verwendet werden, umfassen, sind jedoch nicht beschränkt auf, die frühe SV40-Promotorregion (Bernoist und Chambon, 1981, Nature 290, 304–310), den Promotor, der in der langen 3'-terminalen Sequenzwiederholung des Rous-Sarcom-Virus enthalten ist (Yamamoto, et al., 1980, Cell 22, 787–797), den Herpes-Thymidinkinase-Promotor (Wagner et al., 1981, Proc. Natl. Acad. Sci. U. S. A. 78, 1441–1445), die regulatorischen Sequenzen des Metallothionein-Gens (Brinster et al., 1982, Nature 296, 39–42); die prokaryotischen Expressionsvektoren, wie der β-Lactamase-Promotor (Villa-Kamaroff, et al., 1978, Proc. Natl. Acad. Sci. U. S. A. 75, 3727–3731), oder den tac-Promotor (DeBoer, et al., 1983, Proc. Natl. Acad. Sci. U. S. A. 80, 21–25); siehe auch "Useful Proteins from recombinant bacteria" in Scientific American, 1980, 242, 74–94; pflanzliche Expressionsvektoren, die die Nopalinsynthetase-Promotorregion enthalten (Herrera-Estrella et al., Nature 303, 209–213) oder den Blumenkohlmosaikvirus-35S-RNA-Promotor (Gardner, et al., 1981, Nucl. Acids Res. 9, 2871) und den Promotor des Photosynthese-Enzyms Ribulosebiphosphatcarboxylase (Herrera-Estrella et al., 1984, Nature, 310, 115–120); Promotorelemente aus Hefe oder anderen Pilzen, wie der Gal-4-Promotor, der ADC(Alkoholdehydrogenase)-Promotor, der PGK(Phosphoglycerinkinase)-Promotor, der alkalische Phosphatasepromotor, und die folgenden tierischen Transkriptionskontrollregionen, die Gewebespezifität aufweisen und in transgenen Tieren eingesetzt werden: die Elastase-1-Genkontrollregion, die in pankreatischen Azinuszellen aktiv ist (Swift et al., 1984, Cell 38, 639–646; Ornitz et al., 1986, Cold Spring Harbor Symp. Quant. Biol. 50, 399–409; MacDonald, 1987, Hepatology 7, 425–515); Insulin-Gen-Kontrollregion, die in pankreatischen β-Zellen aktiv ist (Hanahan, 1985, Nature 315, 115–122), Immunglobulin-Gen-Kontrollregion, die in Lymphoidzellen aktiv ist (Grosschedl et al., 1984, Cell 38, 647–658; Adames et al., 1985, Nature 318, 533–538; Alexander et al., 1987, Mol. Cell. Biol. 7, 1436–1444), Maus-Mammatumorvirus-Kontrollregion, die in Hoden, Brust, Lymphoid und Mastzellen aktiv ist (Leder et al., 1986, Cell 45, 485–495), Albumin-Gen-Kontrollregion, die in der Leber aktiv ist (Pinkert et al., 1987, Genes und Devel. 1, 268–276), Alpha-Fetoprotein-Gen-Kontrollregion, die in der Leber aktiv ist (Krumlauf et al. 1985, Mol. Cell. Biol. 5, 1639–1648; Hammer et al., 1987, Science 235, 53–58); Alpha-1-Antitrypsin-Gen-Kontrollregion, die in der Leber aktiv ist (Kelsey et al., 1987, Genes und Devel. 1, 161–171), β-Globin-Gen-Kontrollregion, die in Myeloidzellen aktiv ist (Mogram et al., 1985, Nature 315, 338–340; Kollias et al., 1986, Cell 46, 89–94); Myelinbasisprotein-Gen-Kontrollregion, die in Oligodendrozytzellen im Gehirn aktiv ist (Readhead et al., 1987, Cell 48, 703–712); Myosin-leichte Kette 2-Gen-Kontrollregion, die im Skelettmuskel aktiv ist (Scani, 1985, Nature 314, 283–286) und Gonadotropinfreisetzungshormon-Gen-Kontrollregion, die im Hypothalamus aktiv ist (Mason et al., 1986, Science 234, 1372–1378).

[0130] Expressionsvektoren, die Notch-Gen-Insertionen enthalten, können durch drei Möglichkeiten identifiziert werden: (a) Nukleinsäure-Hybridisierung, (b) Vorliegen oder Abwesenheit von "Marker"-Genfunktionen und (c) Expression von inserierten Sequenzen. Bei der ersten Möglichkeit kann das Vorliegen eines in einen Expressionsvektor eingebauten Fremdgens durch Nukleinsäure-Hybridisierung unter Verwendung von Sonden nachgewiesen werden, die Sequenzen enthalten, die zu einem inserierten toporhythmischen Gen homolog sind. Bei der zweiten Möglichkeit kann das rekombinante Vektor/Wirtssystem auf der Basis des Vorliegens und Fehlens bestimmter "Marker"-Genfunktionen (z. B. Thymidinkinase-Aktivität, Antibiotika-Resistenz, Transformations-Phänotyp, Occlusionskörperbildung in Baculovirus, etc.), die durch die Insertion der Fremdgene in den Vektor verursacht werden, identifiziert und selektiert werden. Wenn beispielsweise das Notch-Gen in die Markergensequenz des Vektors inseriert wird, können Rekombinanten, die die Notch-Insertion enthalten, durch das Fehlen der Markergenfunktion identifiziert werden. Bei der dritten Möglichkeit können rekombinante Expressionsvektoren durch Testen des durch die Rekombinante exprimierten Fremdgenprodukts identifiziert werden. Solche Tests können beispielsweise auf den physikalischen oder funktionellen Eigenschaften des Notch-Gen-Produkts in in vitro Testsystemen, z. B. Aggregations(Adhäsions)-Vermögen (siehe Abschnitt 6–7, nachstehend) beruhen.

[0131] Nach der Identifizierung und Isolierung eines bestimmten rekombinanten DNA-Moleküls können mehrere aus der Technik bekannte Verfahren zu seiner Propagation angewandt werden. Nach Einrichten eines geeigneten Wirtssystems und der Wachstumsbedingungen können rekombinante Expressionsvektoren propagiert und quantitativ hergestellt werden. Wie bereits erklärt, umfassen die Expressionsvektoren, die verwendet werden können, die folgenden Vektoren oder ihre Derivate, sind jedoch nicht darauf beschränkt: menschliche oder tierische Viren, wie Vacciniavirus oder Adenovirus; Insektenviren, wie Baculovirus; Hefevektoren; Bakteriophagenvektoren (z. B. Lambda) und Plasmid- und Cosmid-DNA-Vektoren, um nur einige zu nennen.

[0132] Zusätzlich kann ein Wirtszellenstamm gewählt werden, der die Expression der inserierten Sequenzen moduliert oder das Genprodukt in der speziellen gewünschten Weise modifiziert und weiter verarbeitet. Die Expression aus bestimmten Promotoren kann in Gegenwart bestimmter Induktoren gesteigert werden; somit kann die Expression von gentechnisch hergestelltem Notch-Protein kontrolliert werden. Außerdem haben verschiedene Wirtszellen charakteristische und spezielle Mechanismen für die Translations- und Posttranslationsverarbeitung und Modifikation (z. B. Glycosylierung, Spaltung) der Proteine. Geeignete Zell-Linien oder Wirtssysteme können zur Sicherstellung der gewünschten Modifikation und Weiterverarbeitung des exprimierten Fremdproteins gewählt werden. Beispielsweise kann die Expression in einem Bakteriensystem zur Produktion eines nicht glycosylierten Kernproteinprodukts angewandt werden. Die Expression in Hefe erzeugt ein glycosyliertes Produkt. Die Expression in Säugerzellen kann angewandt werden, um eine "native" Glycosylierung eines heterologen togorhythmischen Säugerproteins sicherzustellen. Außerdem können verschiedene Vektor/Wirtsexpressionssysteme Verarbeitungsreaktionen bewirken, wie proteolytische Spaltungen in unterschiedlichem Ausmaß.

[0133] Bei weiteren speziellen Ausführungsformen kann das Notch-Protein, das Fragment-Analoge oder Derivat als Fusionsprodukt oder chimäres Proteinprodukt, das das Protein, Fragment, Analoge oder Derivat enthält, das über eine Peptidbindung an eine heterologe Proteinsequenz (von einem unterschiedlichen Protein) gebunden worden ist, exprimiert werden. Ein solches chimäres Produkt kann durch Ligation der geeigneten Nukleinsäuresequenzen, die die gewünschten Aminosäuresequenzen miteinander codieren, durch aus der Technik bekannte Verfahren im richtigen Codierungsrahmen und Expression des chimären Produkts durch üblicherweise aus der Technik bekannte Verfahren hergestellt werden. Alternativ kann ein solches chimäres Produkt durch Proteinsynthesetechniken hergestellt werden, z. B. unter Verwendung eines Peptidsynthesegerätes.

[0134] Sowohl cDNA- als Genomsequenzen können kloniert und exprimiert werden. Bei anderen Ausführungsformen kann eine Notch-cDNA-Sequenz chromosomal integriert und exprimiert werden. Aus der Technik bekannte homologe Rekombinationsverfahren können verwendet werden.

5.8.1 IDENTIFIZIERUNG UND REINIGUNG DES EXPRIMIERTEN GENPRODUKTS

[0135] Nach Identifizierung einer Rekombinante, die die Notch-Gen-Sequenz exprimiert, kann das Genprodukt analysiert werden. Dies kann durch Tests auf der Basis der physikalischen oder funktionellen Eigenschaften des Produkts, einschließlich radioaktiver Markierung des Produkts und anschließende Analyse durch Genelektrophorese, erfolgen.

[0136] Nach der Identifizierung des Notch-Proteins kann es isoliert und durch Standardverfahren, einschließlich Chromatographie (z. B. Innenaustausch-, Affinitätschromatographie und Größenausschlußsäulenchromatographie), Zentrifugation, differenzielle Löslichkeit, oder durch jedes andere Standardverfahren zur Reinigung von Proteinen, gereinigt werden. Die funktionellen Eigenschaften können unter Verwendung eines beliebigen geeigneten Tests, einschließlich, jedoch nicht beschränkt auf, Aggregationstests (siehe Abschnitte 6–7) bewertet werden.

5.9 DERIVATE UND ANALOGE VON NOTCH- UND ANDEREN TOPORHYTHMISCHEN PROTEINEN

[0137] Die Erfindung stellt weiterhin als Therapeutika Derivate (einschließlich, jedoch nicht beschränkt auf, Fragmente) und Analoge von Notch-Proteinen bereit.

[0138] Die Produktion und Verwendung von mit Notch verwandten Derivaten und Analogen liegen im Umfang der Erfindung. Bei einer speziellen Ausführungsform ist das Derivat oder Analoge funktionell aktiv, das heißt, es kann eine oder mehrere funktionellen Aktivitäten aufweisen, die mit einem Vollängen-Wildtyp-Notch-Protein zusammenhängen. Als Beispiel können solche Derivate oder Analoge verwendet werden, die die gewünschte Antigenität, beispielsweise in diagnostischen Immuntests, wie in Abschnitt 5.3 beschrieben, aufweisen. Mole-

küle, die eine gewünschte Notch-Eigenschaft bewahren oder alternativ hemmen, z. B. die Bindung an Delta- oder an andere toporhythmische Proteine, Bindung an einen intrazellulären Liganden, können therapeutisch als Induktoren bzw. als Inhibitoren einer solchen Eigenschaft und seiner physiologischen Korrelierungen verwendet werden. Derivate oder Analoge von Notch können durch aus der Technik bekannte Verfahren, einschließlich der nachstehend beschriebenen Tests, jedoch nicht darauf beschränkt, auf die gewünschte Aktivität getestet werden. Bei einer speziellen Ausführungsform können Peptid-Bibliotheken zur Selektion eines Peptids mit der gewünschten Aktivität durchmustert werden. Eine solche Durchmusterung kann durch Testen, z. B. auf die Bindung an Notch, durchgeführt werden.

[0139] Insbesondere können Notch-Derivate durch Änderung der Notch-Sequenzen durch Substitutionen, Additionen oder Deletionen, die funktionell equivalente Moleküle bereitstellen, hergestellt werden. Aufgrund der Degeneration Nukleotid-codierender Sequenzen können andere DNA-Sequenzen, die praktisch die gleiche Aminosäuresequenz wie das Notch-Gen codieren, in der Praxis der Erfindung verwendet werden. Diese umfassen, sind jedoch nicht beschränkt auf, Nukleotidsequenzen, die alles oder Teile von Notch-Genen enthalten, die durch den Ersatz verschiedener Codons geändert worden sind, die einen funktionell equivalenten Aminosäurerest innerhalb der Sequenz codieren und so eine stumme Änderung erzeugen. Gleichermaßen umfassen die erfindungsgemäßen Notch-Derivate, sind jedoch nicht beschränkt auf, diejenigen, die als primäre Aminosäuresequenz alles oder ein Teil der Aminosäuresequenz eines Notch-Proteins enthalten, einschließlich geänderter Sequenzen, in denen funktionell equivalente Aminosäurereste durch Reste innerhalb der Sequenz, die eine stumme Änderung herbeiführen, ersetzt worden sind. Beispielsweise können eine oder mehrere Aminosäurereste in der Sequenz durch eine andere Aminosäure ähnlicher Polarität, die als funktionelles Äquivalent wirkt, ersetzt werden, was zu einer stummen Änderung führt. Ersatz-Aminosäuren für eine Aminosäure in der Sequenz können aus anderen Elementen der Klasse ausgewählt werden, der die Aminosäure angehört. Beispielsweise umfassen die nichtpolaren (hydrophoben) Aminosäuren Alanin, Leucin, Isoleucin, Valin, Prolin, Phenylalanin, Tryptophan und Methionin. Die polaren neutralen Aminosäuren umfassen Glycin, Serin, Threonin, Cystein, Tyrosin, Asparagin und Glutamin. Die positiv geladenen (basischen) Aminosäuren umfassen Arginin, Lysin und Histidin. Die negativ geladenen (sauren) Aminosäuren umfassen Aspartamsäure und Glutaminsäure.

[0140] Derivate oder Analoge von Notch umfassen, sind jedoch nicht beschränkt auf, diejenigen Peptide, die praktisch zu Notch oder zu Fragmenten hiervon homolog sind, oder deren codierende Nukleinsäure zur Hybridisierung an eine Notch-Nukleinsäuresequenz in der Lage ist.

[0141] Die erfindungsgemäßen Notch-Derivate und -Analoge können durch verschiedene aus der Technik bekannte Verfahren hergestellt werden. Die Manipulationen, die zu ihrer Herstellung führen, können auf Gen- oder Proteinniveau erfolgen. Beispielsweise kann die klonierte Notch-Gen-Sequenz durch eine der zahlreichen, aus der Technik bekannten Strategien modifiziert werden (Maniatis. T., 1989, Molecular Cloning, A Laboratory Manual, 2. Ausg., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). Die Sequenz kann an geeigneten Stellen mit Restriktionsendonuklease(n) gespalten und anschließend, wenn gewünscht, enzymatisch modifiziert werden und isoliert und in vitro ligiert werden. Bei der Herstellung des Gens, das ein Derivat oder Analoges von Notch codiert, ist Vorsicht geboten, um zu sicherzustellen, daß das modifizierte Gen im gleichen Translationsleserahmen wie Notch, nicht von Translationsstoppsignalen unterbrochen, in der Genregion, in der die gewünschte Notch-Aktivität codiert wird, verbleibt.

[0142] Zusätzlich kann die Notch-codierende Nukleinsäuresequenz in vitro oder in vivo unter Erzeugung und/oder Zerstörung von Translations-, Start- und/oder Terminationssequenzen oder Erzeugung von Variationen in den codierenden Regionen und/oder Bildung neuer Restriktionsendonukleasestellen oder Zerstörung der bereits existierenden zur leichteren weiteren in vitro Modifikation, mutiert werden. Jede Technik, die aus der Technik zur Mutagenese bekannt ist, kann angewandt werden, einschließlich ortsgerichteter in vitro Mutagenese (Hutchinson, C., et al., 1978, J. Biol. Chem. 253: 6551), Verwendung von TAB®-Linkern (Pharmacia), etc., jedoch nicht darauf beschränkt.

[0143] Manipulationen der Notch-Sequenz können auch auf Protein-Niveau vorgenommen werden. Eingeschlossen im Umfang der Erfindung sind Notch-Proteinfragmente oder andere Derivate oder Analoge, die während oder nach der Translation, z. B. durch Glycosylierung, Acetylierung, Phosphorylierung, Amidierung, Derivatisierung durch bekannte Schutz/Blockierungsgruppen, proteolytsche Spaltung, Verknüpfung mit einem Antikörpermolekül oder einem anderen Zelliganden etc. differentiell modifiziert werden. Jede der zahlreichen chemischen Modifikationen kann durch bekannte Techniken, einschließlich spezifischer chemischer Spaltung durch Cyanogenbromid, Trypsin, Chymotrypsin, Papain, V8-Protease, $NaBH_4$; Acetylierung, Formulierung, Oxidation, Reduktion; metabolischer Synthese in Gegenwart von Tunicamycin; etc., jedoch nicht darauf be-

schränkt, durchgeführt werden.

[0144] Zusätzlich können Analoge und Derivate von Notch chemisch synthetisiert werden. Beispielsweise kann ein Peptid, entsprechend einem Teil eines Notch-Proteins, das die gewünschte Domäne enthält oder das die gewünschte Aggregationsaktivität in vitro vermittelt, oder das an einen Rezeptor bindet, durch Verwendung eines Peptidsynthesegeräts synthetisiert werden. Wenn gewünscht, können außerdem nicht klassische Aminosäuren oder chemische Aminosäure-Analoge als Substitution oder Addition in die Notch-Sequenz eingebaut werden. Nicht klassische Aminosäuren umfassen, sind jedoch nicht beschränkt auf, die D-Isomere der üblichen Aminosäuren, α-Aminoisobuttersäuren, 4-Aminobuttersäure, Hydroxyprolin, Sarcosin, Citrullin, Cysteinsäure, t-Butylglycin, t-Butylalanin, Phenylglycin, Cyclohexylalanin, β-Alanin, Designer-Aminosäuren, wie β-Methylaminosäuren, Cα-Methylaminosäuren und Nα-Methylaminosäuren.

[0145] Bei einer speziellen Ausführungsform ist das Notch-Derivat ein chimäres Protein oder ein Fusionsprotein, das ein Notch-Protein oder ein Fragment hiervon enthält, das über eine Peptidbindung an seinem Amino- und/oder Carboxyterminus an eine Nicht-Notch-Aminosäuresequenz fusioniert ist. Bei einer Ausführungsform wird ein solches chimäres Protein durch rekombinante Expression einer Nukleinsäure, die das Protein (umfassend eine Notch-codierende Sequenz, die im Raster mit einer Nicht-Notch-codierender Sequenz verbunden ist) codiert, hergestellt. Ein solches chimäres Produkt kann durch Ligation der geeigneten Nukleinsäuresequenzen hergestellt werden, die die gewünschte Aminosäuresequenzen füreinander codieren, durch aus der Technik bekannte Verfahren im richtigen Codierungsraster und Expression des chimären Produkts durch aus der Technik allgemein bekannte Verfahren hergestellt werden. Alternativ kann ein solches chimäres Produkt durch Proteinsynthesetechniken hergestellt werden, z. B. Verwendung eines Peptidssynthesegerätes. Bei einer speziellen Ausführungsform wird eine chimäre Nukleinsäure, die ein reifes Notch-Protein mit einer heterologen Signalsequenz codiert, so exprimiert, daß das chimäre Protein exprimiert und durch die Zelle zum reifen Notch-Protein weiter verarbeitet wird. Als anderes Beispiel, jedoch nicht als Einschränkung, kann erfindungsgemäß ein rekombinantes Molekül konstruiert werden, das codierende Portionen von Notch und einem anderen togorhythmischen Genen, z. B. Delta, enthält. Das codierte Protein eines solchen rekombinanten Moleküls kann Eigenschaften aufweisen, die sowohl mit Notch als auch Delta in Verbindung stehen, und ein neues Profil von biologischen Aktivitäten porträtieren, einschließlich Agonisten und Antagonisten. Die Primärsequenz von Notch und Delta kann auch zur Vorhersage der Tertiärstruktur der Moleküle unter Anwendung von Computersimulation eingesetzt werden (Hopp und Woods, 1981, Proc. Natl. Acad. Sci. U. S. A. 78: 3824–3828); chimäre Notch/Delta-Rekombinationsgene können im Hinblick auf die Korrelationen zwischen Tertiärstruktur und biologischer Funktion entworfen werden. Gleichermaßen können chimäre Gene, die Teile von Notch enthalten, die mit einer heterologen (Nicht-Notch)-Protein-codierenden Sequenz fusioniert sind, konstruiert werden. Eine spezielle Ausführungsform betrifft ein chimäres Protein, das ein Fragment von Notch mit mindestens sechs Aminosäuren enthält.

[0146] Bei einer speziellen Ausführungsform ist das Notch-Derivat ein Notch-Fragment, das mit einem anderen toporhythmischen Protein eine homologe Region aufweist. Wie hier verwendet, soll eine Region eines ersten Proteins als "homolog" mit einem zweiten Protein angesehen werden, wenn die Aminosäuresequenz der Region zu mindestens 30% identisch ist oder zu mindestens 75% entweder identisch ist oder im Vergleich zu einer Sequenz im zweiten Protein mit einer gleichen Anzahl von Aminosäuren, wie die Anzahl, die in der Region enthalten ist, konservative Änderungen aufweist.

5.9.1 DERIVATE VON NOTCH, DIE EINE ODER MEHRERE DOMÄNEN DES PROTEINS ENTHALTEN

[0147] Bei einer speziellen Ausführungsform stellt die Erfindung Therapeutika bereit, die Notch-Derivate und -Analoge sind, insbesondere Notch-Fragmente und -Derivate von solchen Fragmenten, die eine oder mehrere Domänen des Notch-Proteins enthalten, einschließlich der extrazellulären Domäne, Transmembrandomäne, intrazellulären Domäne, Membran-assoziierten Region, einer oder mehrerer der EGF-artigen Sequenzwiederholungen (ELR) des Notch-Proteins, der cdc10-Sequenzwiederholungen und der Notch/lin-12-Sequenzwiederholungen, jedoch nicht darauf beschränkt. Bei speziellen Ausführungsformen können dem Notch-Derivat alle oder ein Teil der ELRs oder eine oder mehrere andere Regionen des Proteins fehlen.

[0148] Bei einer speziellen Ausführungsform, bezüglich eines Notch-Proteins, einer Spezies, die anders als D. Melanogaster und vorzugsweise menschlich ist, sind Fragmente, die spezielle Teile von Notch enthalten, diejenigen Fragmente, die in dem entsprechenden Notch-Protein Teile enthalten, die zu speziellen Fragmenten des Drosophila-Notch-Proteins (z. B. ELR-11 und ELR-12) besonders homolog sind.

5.9.2 DERIVATE VON NOTCH ODER ANDERE TOPORHYTHMISCHE PROTEINE, DIE DIE BINDUNG AN TOPORHYTHMISCHE PROTEINDOMÄNEN VERMITTELN, UND INHIBITOREN HIERFÜR

[0149] Die Erfindung stellt auch Notch-Fragmente und Analoge oder Derivate solcher Fragmente bereit, die die Bindung an toporhythmische Proteine vermitteln (und somit hier als "klebend" bezeichnet werden), und Nukleinsäuresequenzen die die Vorgenannten codieren.

[0150] Ebenfalls als erfindungsgemäße Therapeutika eingeschlossen, sind Inhibitoren (z. B. Peptid-Inhibitoren) der vorgenannten toporhythmischen Protein-Wechselwirkungen mit Notch.

[0151] Die Fähigkeit zur Bindung an ein toporhythmisches Protein kann durch in vitro Aggregationstests mit Zellen gezeigt werden, die ein solches toporhythmisches Protein exprimieren, sowie Zellen die Notch oder ein Notch-Derivat exprimieren (siehe Abschnitt 6). Das heißt, die Fähigkeit eines Proteinfragments zur Bindung an ein Notch-Protein kann durch Nachweis der Fähigkeit des Fragments gezeigt werden, bei Expresseion auf der Oberfläche einer ersten Zelle an ein Notch-Protein zu binden, das auf der Oberfläche einer zweiten Zelle exprimiert wird. Inhibitoren der vorgenannten Wechselwirkungen können durch ihre Fähigkeit zur Hemmung einer solchen Aggregation in vitro nachgewiesen werden.

[0152] Die Nukleinsäuresequenzen, die toporhythmische Proteine oder die klebenden Domänen hiervon codieren, können aus Menschen, Schweinen, Rindern, Katzen, Vögeln, Pferden, Hunden oder Insekten, sowie Primatenquellen und aus jeder anderen Spezies isoliert werden, bei denen Homologe bekannter toporhythmischer Gene identifiziert werden können.

[0153] Bei einer speziellen Ausführungsform ist das klebende Fragment von Notch dasjenige, das den Teil von Notch enthält, der zu ELR-11 und ELR-12 besonders homolog ist, das heißt zu den Aminosäuren mit den Nummern 447 bis 527 (SEQ ID NO: 14) der Drosophila-Notch-Sequenz (siehe **Fig. 4**).

[0154] Auf Grund der Degeneration der Nukleotid-codierenden Sequenz können andere DNA-Sequenzen, die im wesentlichen die gleiche Aminosäuresequenz codieren wie die Adhäsionssequenzen, in der Praxis der Erfindung eingesetzt werden. Diese umfassen, sind jedoch nicht beschränkt auf, Nukleotidsequenzen, die alles oder Teile des Notch enthalten, die durch Ersatz verschiedener Codons, die einen funktionell äquivalenten Aminosäurerest innerhalb der Sequenz codieren und somit eine stumme Änderung erzeugen, geändert wurden. Gleichermaßen umfassen die erfindungsgemäßen klebenden Proteinfragmente oder Derivate hiervon, sind jedoch nicht beschränkt auf, diejenigen, die als primäre Aminosäuresequenz alles oder einen Teil der Aminosäuresequenz der klebenden Domänen enthalten, einschließlich geänderter Sequenzen, in denen die Reste innerhalb der Sequenz durch funktionell äquivalente Aminosäurereste ersetzt worden sind, was zu einer stummen Änderung führt.

[0155] Klebende Fragmente toporhythmischer Proteine und potentielle Derivate, Analoge oder Peptide, die mit klebenden toporhythmischen Proteinsequenzen verwandt sind, können auf die gewünschte Bindungsaktivität getestet werden, z. B. durch die in den folgenden Beispielen beschriebenen in vitro Aggregationstests. Klebende Derivate oder klebende Analoge von klebenden Fragmenten toporhythmischer Proteine umfassen, sind jedoch nicht beschränkt auf, diejenigen Peptide, die im wesentlichen zu den klebenden Fragmenten homolog sind oder deren codierende Nukleinsäure in der Lage ist, mit der Nukleinsäuresequenz zu hybridisieren, die die klebenden Fragmente codiert, und wobei Peptide und Peptidanaloge positive Bindungsaktivität aufweisen, z. B. wie in vitro durch einen Aggregationstest, wie nachstehend in den Abschnitten für die Beispiele beschrieben, getestet. Solche Derivate und Analoge werden als Therapeutika in Betracht gezogen und liegen im Umfang der Erfindung.

[0156] Die mit dem Adhäsionsprotein verwandten erfindungsgemäßen Derivate, Analoge und Peptide können durch verschiedene, aus der Technik bekannte Verfahren hergestellt werden. Die Manipulationen, die zu ihrer Herstellung führen, können auf dem Gen- oder Proteinniveau erfolgen (siehe Abschnitt 5.6).

[0157] Zusätzlich können die Adhäsionsprotein-codierenden Nukleinsäuresequenzen in vitro oder in vivo mutiert werden; und Manipulationen an der klebenden Sequenz können ebenfalls auf Proteinniveau vorgenommen werden (siehe Abschnitt 5.6).

[0158] Zusätzlich können Analoge und Peptide, die mit den klebenden Fragmenten verwandt sind, chemisch synthetisiert werden.

5.10 TESTS VON NOTCH-PROTEINEN, -DERIVATEN UND -ANALOGEN

[0159] Die in vitro Aktivität von Notch-Proteinen, -Derivaten und -Analogen und anderen togorhythmischen Proteinen, die an Notch binden, kann durch verschiedene Verfahren getestet werden.

[0160] Bei einer Ausführungsform, wobei die Fähigkeit zur Bindung oder zur Konkurrenz mit Wild-Typ-Notch und die Bindung an Anti-Notch-Antikörper getestet wird, können verschiedene, aus der Technik bekannte Immuntests eingesetzt werden, einschließlich kompetitiver und nicht kompetitiver Testsysteme, unter Anwendung von Techniken, wie Radioimmuntests, ELISA (enzyme linked Immunosorbent assay), "Sandwich"-Immuntests, Immunradiometrie-Tests, Geldiffusionfällungsreaktionen, Immundiffusionstests, in situ Immuntests (unter Verwendung beispielsweise von kolloidalem Gold, enzymatischer oder Radioisotopen-Markierung), Western-Blots, Fällungsreaktionen, Agglutinationstests (z. B. Gelagglutinationstests, Hämagglutinationstests), Komplementfixierungstests, Immunfluoreszenztests, Protein-A-Tests und Immunelektrophoresetests, etc., jedoch nicht darauf beschränkt. Bei einer Ausführungsform wird die Antikörperbindung durch Nachweis einer Markierung auf dem primären Antikörper nachgewiesen. Bei einer anderen Ausführungsform wird der primäre Antikörper durch Nachweis der Bindung eines sekundären Antikörpers oder Reagenz an den primären Antikörper nachgewiesen. Bei einer weiteren Ausführungsform wird der sekundäre Antikörper markiert. Aus der Technik sind viele Mittel zum Nachweis der Bindung bei einem Immuntest bekannt und liegen im Umfang der Erfindung.

[0161] Bei einer anderen Ausführungsform, wobei die Fähigkeit zur Vermittlung der Bindung an Notch getestet wird, kann ein in vitro Aggregationstest durchgeführt werden, wie infra in Abschnitt 6 oder 7 beschrieben (siehe auch Fehon et al., 1990, Cell 61: 523–534; Rebay et al., 1991, Cell 67: 687–699).

[0162] Bei einer anderen Ausführungsform, wobei ein anderer Ligand für Notch identifiziert wird, kann die Ligandenbindung getestet werden, z. B. durch aus der Technik gut bekannte Bindungstests. Bei einer anderen Ausführungsform können physiologische Korrelate von Ligandenbindung an Zellen, die einen Notch-Rezeptor exprimieren, (Signaltransduktion) getestet werden.

[0163] Bei einer anderen Ausführungsform in einem Insekten- oder einem anderen Modellsystem können genetische Studien zum Studium des phänotypischen Effekts einer Notch-Mutante, die ein Derivat oder Analoges von Wild-Typ-Notch ist, vorgenommen werden.

[0164] Weitere Verfahren sind dem Fachmann bekannt und liegen im Umfang der Erfindung.

5.11 ANTIKÖRPER AUF NOTCH-PROTEINE, -DERIVATE UND -ANALOGE

[0165] Bei einer Ausführungsform der Erfindung sind Antikörper und Fragmente, die die gegen Notch gerichtete Bindungsdomäne hiervon enthalten, Therapeutika. Demnach können Notch-Proteine, Fragmente oder Analoge oder Derivate hiervon, insbesondere menschliche Notch-Proteine oder Fragmente hiervon, als immunogene Erzeugung von Anti-Notch-Proteinantikörpern verwendet werden. Solche Antikörper können polyklonal, monoklonal, chimär, einzelkettig, Fab-Fragmente sein oder aus einer Fab-Expressionsbibliothek stammen. Bei einer speziellen Ausführungsform können Antikörper, die auf die EGF-artigen Sequenzwiederholungen 11 und 12, von Notch spezifisch sind, produziert werden. Bei anderen Ausführungsformen können Antiköper, die mit der extrazellulären Domäne von Notch reaktiv sind, produziert werden. Ein Beispiel für solche Antikörper kann die Aggregation bei einem in vitro Test verhindern. Bei einer anderen Ausführungsform werden Antikörper, die auf menschliches Notch spezifisch sind, produziert.

[0166] Verschiedene, aus der Technik bekannte Verfahrensweisen können zur Produktion von polyklonalen Antikörpern auf ein Notch-Protein oder -Peptid angewandt werden. Bei einer besonderen Ausführungsform können polyklonale Kaninchen-Antikörper auf ein Epitop des menschlichen Notch-Proteins, das von einer in **Fig.** 10 oder 11 abgebildeten Sequenz oder einer Untersequenz hiervon codiert wird, erhalten werden. Zur Produktion von Antikörpern können verschiedene Wirtstiere durch Injektion mit dem nativen Notch-Protein oder einer synthetischen Version oder einem Fragment hiervon immunisiert werden, einschließlich, jedoch nicht beschränkt auf, Kaninchen, Mäuse, Ratten, etc. Verschiedene Hilfsstoffe können zur Steigerung der Immunantwort in Abhängigkeit von der Wirtsspezies verwendet werden und umfassen, sind jedoch nicht beschränkt auf, Freund Medium (komplett und inkomplett), mineralische Gele, wie Aluminiumhydroxid, oberflächenaktive Substanzen, wie Lysolecithin, Pluronic-Polyole, Polyanionen, Peptide, Ölemulsionen, Schloß-und-Schlüssel-Hämocyanine, Dinitrophenol und potentiell geeignete menschliche Hilfsstoffe, wie BCG(Bacillus Calmette-Guerin) und Corynebacterium parvum.

[0167] Zur Herstellung von monoklonalen Antikörpern, die gegen eine Notch-Proteinsequenz gerichtet sind, kann jede Technik eingesetzt werden, die die Produktion von Antikörpermolekülen durch kontinuierliche Zell-Linien in Kultur bereitstellt; beispielsweise die Hybridomtechnik, die ursprünglich von Kohler und Milstein entwickelt wurde (1975, Nature 256, 495–497), sowie die Triomtechnik, die menschliche B-Zellen-Hybridomtechnik (Kozbor et al., 1983, Immunology Today 4, 72) und die EBV-Hybridomtechnik zur Produktion von menschlichen monoklonalen Antikörpern (Cole et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., Ss. 77–96).

[0168] Antikörperfragmente, die den Idiotyp (Bindungsdomäne) des Moleküls enthalten, können durch bekannte Techniken erzeugt werden. Beispielsweise umfassen solche Fragmente, sind jedoch nicht beschränkt auf, das $F(ab')_2$-Fragment, das durch Pepsinverdau des Antikörpermoleküls erzeugt werden kann; die Fab'-Fragmente, die durch eine Verminderung der Disulfidbrücken des $F(ab')_2$-Fragments erzeugt werden können, und die Fab-Fragmente, die durch Behandeln des Antikörpermoleküls mit Papain und einem Reduktionsmittel erzeugt werden können.

[0169] Bei der Produktion von Antikörpern kann die Durchmusterung auf den gewünschten Antikörper, durch aus der Technik bekannte Verfahren erreicht werden, z. B. ELISA (enzyme linked immunosorbent assay). Beispielsweise können zur Selektion von Antikörpern, die die klebende Domäne eines Notch-Proteins erkennen, erzeugte Hybridome auf ein Produkt getestet werden, das an ein Proteinfragment bindet, das eine solche Domäne enthält. Zur Selektion eines auf menschliches Notch spezifischen Antikörpers kann ein Antikörper auf der Grundlage der positiven Bindung an menschliches Notch und einer fehlenden Bindung an Drosophila-Notch gewählt werden.

[0170] Zusätzlich zur therapeutischen Brauchbarkeit besitzen die vorgenannten Antikörper Brauchbarkeit bei diagnostischen Immuntests, wie in Abschnitt 5.6 vorstehend beschrieben.

[0171] Ähnliche Verfahren wie diejenigen, die vorstehend beschrieben worden sind, können zur Herstellung von Therapeutika angewandt werden, die Antiköper auf Domänen anderer Proteine (insbesondere toporhythmischer Proteine) sind, die an Notch binden oder anderweitig mit Notch Wechselwirken (z. B. klebende Fragmente von Delta oder Serrate).

6. DOMÄNEN VON NOTCH VERMITTELN DIE BINDUNG MIT DELTA

[0172] Die intermolekulare Assoziation zwischen den Produkten der Notch- und Delta-Gene wurde durch das Studium der Wirkungen ihrer Expression auf die Aggregation in Drosophila-Schneider-2(S2)-Zellen (Fehon et al., 1990, Cell 61, 523–534) nachgewiesen. Ein direkter Nachweis der intermolekularen Wechselwirkungen zwischen Notch und Delta sowie ein Testsystem, das zur Zerlegung der erfindungsgemäßen Komponenten angewandt werden kann, ist hier beschrieben. Normalerweise binden nicht klebende Drosophila-S2-Kulturzellen, die Notch exprimieren, spezifisch Calcium-abhängig an Zellen, die Delta exprimieren.

[0173] Während Zellen, die Notch exprimieren, nicht aneinander binden, binden Zellen, die Delta exprimieren, doch aneinander, was zudem nahe legt, daß Notch und Delta um die Bindung an Delta an der Zelloberfläche konkurrieren können. Notch und Delta bilden sowohl in Kulturzellen als auch embryonalen Zellen Detergenz-lösliche Komplexe, was nahe legt, daß Notch und Delta direkt auf molekularem Niveau in vitro und in vivo Wechselwirken. Die Analysen legen nahe, daß Notch- und Delta-Proteine an der Zelloberfläche über ihre extrazellulären Domänen Wechselwirken.

6.1 EXMPERIMENTELLE VERFAHREN

6.1.1 EXPRESSIONSKONSTRUKTE

[0174] Expressionskonstrukte sind bei Fehon et al., 1990, Cell 61: 523–534, beschrieben. Kurz gesagt wurde Notch, das durch das MgIIa-Miniaturgen, ein cDNA/Genom-Chimärkonstrukt, (Ramos et al., 1989, Genetics 123, 337–348) codiert wurde, exprimiert und anschließend in pRmHa-3 inseriert (Bunch, et al., 1988, Nucl. Acids Res. 16, 1043–1061). In dem resultierenden Konstrukt, das mit pMtNMg bezeichnet wurde, ist der Metallothionein-Promotor in pRmHa-3 an die Notch-Sequenzen, ausgehend 20 Nukleotide, stromaufwärts der Translationsstartstelle fusioniert.

[0175] Das extrazelluläre Notch-Konstrukt (ECN1) wurde von einem Notch-Cosmid abgeleitet (Ramos et al., 1989, Genetics 123, 337–348) und weist eine interne Deletion der Notch-codierenden Sequenzen von Amino-

säure 1790 bis einschließlich 2625 (Wharton et al., 1985, Cell 43, 567–581) und eine vorhergesagte Rahmenverschiebung, die einen neuen 59 Aminosäurecarboxyl-Terminus erzeugt, auf.

[0176] Für das Delta-Expressionskonstrukt, wurde die D11-cDNA (Kopczynski et al., 1988, Genes Dev. 2, 1723–1735; **Fig.** 1; SEQ ID NO: 1), die die vollständige Kodierungskapazität für Delta enthält, in die EcoRI-Stelle von pRmHa-3 inseriert. Dieses Konstrukt wurde pMTDI1 genannt.

6.1.2 ANTIKÖRPER HERSTELLUNG

[0177] Die Hybridomzellinie C17.9C6 wurde aus Mäusen erhalten, die mit einem Fusionsprotein auf der Basis eines 2,1 kb-SalI-HindIII-Fragments immunisiert wurden, das die Kodierungssequenzen für den größten Teils der intrazellulären Domäne von Notch (Aminosäuren 1791–2504; Wharton et al., 1985, Cell 43, 567–581) enthält. Das Fragment wurde in pUR289 subkloniert (Ruther und Muller-Hill, 1983, EMBO J. 2, 1791–1794) und anschließend in den pATH1-Expressionsvektor (Dieckmann und Tzagoloff, 1985, J. Biol. Chem. 260, 1513–1520) als BglII-HindIII-Fragment transferiert. Lösliches Fusionsprotein wurde exprimiert mit 25% $(NH_4)_2SO_4$ gefällt, in 6 M Harnstoff resuspendiert und durch präparative isoelektrische Fokussierung unter Verwendung eines Rotofor (Bio-Rad) (ausführliche Angaben, siehe Fehon, 1989, Rotofor Review No. 7, Bulletin 1518, Richmond, California; Bio-Rad Laboratories) gereinigt.

[0178] Polyklonale Maus-Antiseren wurden gegen die extrazelluläre Domäne von Notch unter Verwendung von vier BstY1-Fragmenten einer Länge von 08 kb (Aminosäuren 237–501; Wharton et al., 1985, Cell 43, 567–581), 1,1 kb (Aminosäuren 501–868), 0,99 kb (Aminosäuren 868–1200) und 1,4 kb (Aminosäuren 1465–1935), die sich von der fünften EGF-artigen Sequenzwiederholung über die Transmembrandomäne erstreckten, und einzeln rahmenkonform in den geeigneten pGEX-Expressionsvektor inseriert wurden (Smith und Johnson, 1988, Gene 67, 31–40), hergestellt. Die Fusionsproteine wurden über Glutathionagaroseperlen (SIGMA) gereinigt. Die Maus- und Ratten-Antiseren wurden mit 50% $(NH_4)_2SO_4$ gefällt und in PBS (150 mM NaCl, 14 mM Na_2HPO_4, 6 mM NaH_2PO_4) mit 0,02% NaN_3 resuspendiert.

[0179] Die Hybridomzellinie 201 wurde aus einer Maus erhalten, die mit einem Fusionsprotein immunisiert war, das Codierungssequenzen aus der extrazellulären Domäne von Delta (Kopczynski et al., 1988, Genes Dev. 2, 1723–1735), einschließlich Sequenzen, die von der vierten bis zur neunten EGF-artigen Sequenzwiederholung in Delta verlaufen (Aminosäuren 350–529), enthält.

[0180] Die polyklonalen Ratten-Antiseren wurden nach Immunisierung mit Antigen erhalten, das aus dem gleichen Fusionsproteinkonstrukt stammte. In diesem Fall wurde das Fusionsprotein durch Lyse IPTG-induzierte Zellen in SDS-Laemmli-Puffer hergestellt (Carroll und Laughon, 1987, in DNA Cloning, Band III, D. M. Glover, Hrsg. (Oxford: IRL Press), Ss. 89–111), die Proteine wurden durch SDS-PAGE aufgetrennt, die entsprechenden Banden wurden aus dem Gel ausgeschnitten und das Antigen aus der Gelscheibe zur Verwendung bei der Immunisierung elektroeluiert (Harlow und Lane, 1988, Antibodies: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory))

6.1.3 ZELLKULTUR UND TRANSFEKTION

[0181] Die S2-Zell-Linie (Schneider, 1972, J. Embryol. Exp. Morph. 27, 353–365) wurde in M3-Medium (hergestellt von Hazleton Co.), angereichert mit 2,5 mg/ml Bacto-Pepton (Difco), 1 mg/ml TC Yeastolate (Difco), 11% wärmeinaktiviertem fetalem Kälberserum (FCS) (Hyklone) und 100 U/ml Penizillin, 100 µg/ml Streptomycin-0,25 µg/ml Fungizone (Hazelton), gezüchtet. Zellen, die in der Log-Phase bei ca. 2×10^6-Zellen/ml wachsen, wurden mit 20 µg DNA-Calciumphosphat-Kopräzipitat in 1 ml/5 ml Kultur, wie bereits beschrieben (Wigler et al., 1979, Proc. Natl. Acad. Sci. USA 78, 1373–1376), mit der Ausnahme, daß BES-Puffer (SIGMA) anstelle von HEPES-Puffer (Chen und Okayama, 1987, Mol. Cell. Biol. 7, 2745–2752) verwendet wurde, transfiziert. Nach 16–18 h wurden die Zellen in konische Zentrifugenröhrchen übergeführt, in einer klinischen Zentrifuge bei höchster Geschwindigkeit 30 s pelletiert, einmal mit 1/4 Volumen frischem Komplettmedium gespült in ihrem Originalvolumen an Komplettmedium resuspendiert und wieder in den Originalkolben übergeführt. Die transfizierten Zellen wurden dann vor der Inkubation 24 h ruhen gelassen.

6.1.4 AGGREGATIONSTESTS

[0182] Die Expression der Notch- und Delta-Metallothionein-Konstrukte wurde durch die Zugabe von $CuSO_4$ bis auf 0,7 mM induziert. Zellen, die mit dem ECN1-Konstrukt transfiziert worden sind, wurden ähnlich behandelt. Die Zellen wurden anschließend gemischt, unter Aggregationsbedingungen inkubiert und unter Verwen-

dung spezifischer Antiseren und Immunfluoreszenzmikroskopie zur Visualisierung exprimierender Zellen auf ihre Fähigkeit zur Aggregation bewertet.

[0183] Zwei Typen von Aggregationstests wurden angewandt. Beim ersten Test wurden insgesamt 3 ml Zellen (5–10 × 10^{6-} Zellen/ml) in einen 25-ml-Erlenmeyerkolben vorgelegt und bei 40 bis 50 u/min auf einem Rotationsschüttler 24–48 h bei Raumtemperatur rotiert. Für diese Experimente wurden die Zellen 1–4 h nach Induktionsbeginn gemischt, und die Induktion wurde während der gesamten Aggregationsdauer fortgesetzt. Im zweiten Test wurden ca. 0,6 ml Zellen nach einer Über-Nacht-Induktion (12–16 h) bei Raumtemperatur in ein 0,6-ml-Eppendorfröhrchen (wobei eine kleine Luftblase zurückblieb) gegeben und vorsichtig 1–2 h bei 4°C hin und her geschüttelt. Die Antikörper-Hemmungs- und Ca^{2+}-Abhängigkeitsexperimente wurden unter Anwendung von letzterem Test durchgeführt. Für die Ca^{2+}-Abhängigkeitsexperimente wurden die Zellen zuerst gesammelt und in ausgewogener Salzlösung (BSS) mit 11% FCS (BSS-FCS; FCS wurde gegen 0,9% NaCl, 5 mM-Tris [pH 7,5] dialysiert) oder in Ca^{2+}-freiem BSS-FCS, das 10 mM EGTA enthielt (Snow et al., 1989, Cell 59, 313–323), gespült und dann im gleichen Medium bis zum ursprünglichen Volumen resuspendiert. Für die Antiköper-Hemmungsexperimente wurden die Notchtransfizierten Zellen gesammelt und in M3-Medium gespült und so dann vor der Aggregation in M3-Medium 1 h bei 4°C mit einer 1:250-Verdünnung von Immun- oder Präimmunserum aus jedem der vier Mäuse, die mit Fusionsproteinen, die Segmente aus der extrazellulären Domäne von Notch enthielten (siehe Antikörperherstellung vorstehend), immunisiert worden sind, behandelt.

6.1.5 IMMUNFLUORESZENZ

[0184] Die Zellen wurden durch Zentrifugation (3000 U/min für 20 s in einer Eppendorf-Mikrozentrifuge) gesammelt und in 0,6-ml-Eppendorfröhrchen mit 0,5 ml frisch hergestelltem 2% Paraformaldehyd in PBS 10 min bei Raumtemperatur fixiert. Nach dem Fixieren wurden die Zellen durch Zentrifugation gesammelt, zweimal in PBS gespült und 1 h mit primärem Antikörper in PBS mit 0,1% Saponin (SIGMA) und 1% normalem Ziegenserum (Pocono Rabbit Farm. Canadensis. PA) angefärbt. Monoklonale Antikörper-Überstände wurden 1:10 verdünnt, und Maus- oder Rattenserum wurde für diesen Schritt 1:1000 verdünnt. Die Zellen wurden dann einmal in PBS gespült und 1 h mit speziellen sekundären Antikörpern (Ziegen-Anti-Maus- und Ziegen-Anti-Ratten-Antikörper von Doppelmarkierungsqualität, Jackson Immunoresearch) in normalem PBS-Saponin-Ziegenserum angefärbt. Nach dieser Inkubation wurden die Zellen zweimal in PBS gespült und in 90% Glycerin, 10% 1M-Tris (pH 8,0) und 0,5% n-Propylgallat auf Deckgläser aufgebracht. Die Zellen wurden unter Epifluoreszenz auf einem Leitz-Orthoplan-2-Mikroskop betrachtet.

[0185] Konfokale Mikrographien wurden unter Verwendung des Bio-Rad MRC 500-Systems aufgenommen, das mit einem Zeiss Axiovert-Verbindungsmikroskop gekoppelt war. Die Bilder wurden unter Verwendung der BHS und GHS-Filterserien gesammelt, die unter Verwendung des ALIGN-Programms ausgerichtet und unter Verwendung von MERGE zusammengeführt wurden. Fluoreszenz-Ausbluten aus dem grünen in den roten Kanal wurde unter Verwendung des BLEED-Programms (sämtliche Software bereitgestellt von Bio-Rad) vermindert. Die. Fotografien wurden direkt auf dem Computermonitor unter Verwendung eines Kodak Ektar 125-Films erhalten.

6.1.6 ZELLLYSATE, IMMUNFÄLLUNGEN UND WESTERN-BLOTS

[0186] Nicht denaturierende Detergens-Lysate von Gewebekulturen und Wild-Typ-Canton-S-Embryonen wurden auf Eis in ca. 10 Zellvolumen Lysepuffer (300 mM NaCl, 50 mM Tris [pH 8,0], 0,5% NP-40, 0,5% Deoxycholat, 1 mM $CaCl_2$, 1 mM $MgCl_2$) mit 1 mM Phenylmethysulfonylfluorid (PMSF) und Diisopropylfluorphosphat, verdünnt 1:2500, als Proteaseinhibitoren, hergestellt. Die Lysate wurden unter Verwendung von 18G, 21G und 25G Nadeln, die an 1 cc-Tuberkulinspritzen angeschlossen waren, nach und nach verrieben und anschließend bei voller Geschwindigkeit in einer Mikrozentrifuge 10 min bei 4°C zur Entfernung von unlöslichem Material zentrifugiert. Die Immunfällung wurde durch Zugabe von ca. 1 µg Antikörper (1–2 µl polyklonales Antiserum) zu 250–500 µl Zelllysat und Inkubation für 1 h bei 4°C unter Rühren durchgeführt. Diesem Gemisch wurden 15 µg Ziegen-Antimaus-Antikörper (Jackson Immunoresearch; diese Antikörper erkennen sowohl Maus- als auch Ratten-IgG) zugesetzt und man ließ 1 h bei 4°C unter Rühren inkubieren. Hierauf folgt die Zugabe von 100 µl fixierten Staphylococcus aureus (Staph A)-Bakterien (Zysorbin, Zymed; resuspendiert nach den Anleitungen des Herstellers), die gesammelt, fünfmal in Lysepuffer gewaschen und noch 1 h inkubiert worden waren. Die Staph A-Antikörperkomplexe wurden dann durch Zentrifugation pelletiert und dreimal in Lysepuffer gewaschen, gefolgt von zwei 15 min-Waschvorgängen in Lysepuffer. Nach der Überführung in ein neues Röhrchen wurde das ausgefällte Material in 50 µl SDS-PAGE-Probenpuffer suspendiert, sofort 10 min gekocht, auf 3- bis 15%-igen Gradientengelen laufen gelassen, auf Nitrocellulose geblottet und unter Verwendung monoklonaler Antikörper und HRP-konjugierter, sekundärer Ziegen-Anti-Maus-Antikörper, wie bereits beschrie-

ben (Johansen et al., 1989, J. Cell Biol. 109, 2427–2440), nachgewiesen. Für die Gesamtzellproteinproben, die aus den Western-Blots verwendet wurden, wurden die Zellen durch Zentrifugation gesammelt, in 10 Zellvolumina Probenpuffer, der 1 mM PMSF enthielt, lysiert und sofort gekocht.

6.2 ERGEBNISSE

6.2.1 DIE EXPRESSION VON NOTCH UND DELTA IN KULTIVIERTEN ZELLEN

[0187] Zum Nachweis von Wechselwirkungen zwischen Notch und Delta überprüften wir das Verhalten der Zellen, die diese Proteine auf ihren Oberflächen exprimieren, unter Anwendung eines Aggregationstests. Für diese Studien wählten wir die S2-Zelllinie (Schneider, 1972, J. Embryol. Exp. Morph. 27, 353–365). S2-Zellen exprimieren eine anormale Notch-Botschaft und kein nachweisbares Notch, aufgrund einer Umlagerung des 5'-Endes der Notch-codierenden Sequenz. Diese Zellen exprimieren auch kein nachweisbares Delta.

[0188] Die Ergebnisse des Western-Blut und der Immunfluoreszenzanalyse zeigten eindeutig, daß die Notch- und Delta-Konstrukte die Expression von Proteinen der erwarteten Größen und subzellulären Lokalisation unterstützen.

6.2.2 ZELLEN, DIE NOTCH UND DELTA EXPRIMIEREN AGGREGIEREN

[0189] Ein einfacher Aggregationstest wurde zum Nachweis von Wechselwirkungen zwischen Notch und Delta, die auf der Oberfläche von S2-Zellen exprimiert wurden, angewandt.

[0190] S2-Zellen im Log-Phasenwachstum wurden entweder mit dem Notch- oder Delta-Metallothionein-Promotorkonstrukt getrennt transfiziert. Nach Induktion mit $CuSO_4$ wurden die transfizierten Zellen in gleicher Anzahl gemischt, und man ließ sie über Nacht bei Raumtemperatur aggregieren (ausführliche Angaben, siehe experimentelle Verfahren, Abschnitt 6.1). Alternativ wurden bei einigen Experimenten, die die Verminderung der metabolischen Aktivität beabsichtigten, die Zellen vorsichtig bei 4°C 1–2 h gemischt. Um zu bestimmen, ob sich Aggregate sich gebildet hatten, wurden die Zellen für die Immunfluoreszenzmikroskopie unter Verwendung von Antikörpern, die für jedes Genprodukt spezifisch waren, und verschieden markierten fluoreszierenden sekundären Antikörpern aufgearbeitet. Die exprimierenden Zellen machten in der Regel weniger als 5% der gesamten Zellpopulation aus, da transiente statt stabile Transformanten verwendet wurden. Die verbleibenden Zellen exprimierten entweder ein gegebenes Protein nicht oder exprimierten auf Niveaus, die zum Nachweis durch Immunfluoreszenzmikroskopie zu gering waren. Als Kontrollen führten wir Aggregationen mit nur einem einzigen transfizierten Zelltyp durch.

[0191] Die Ergebnisse (Fehon et al., 1990, Cell 61: 523–534) zeigten, daß, obgleich bei dem Test Notch-exprimierende ($Notch^+$)-Zellen allein keine Aggregate bildeten, dies mit Deltaexprimierenden($Delta^+$)-Zellen erfolgte. Das Bestreben der $Delta^+$-Zellen zur Aggregation war auch in den nicht aggregierten Kontrollproben offensichtlich, wo üblicherweise Zellcluster von 4–8 Zellen, die wahrscheinlich von der Adhäsion zwischen mitotischen Schwesterzellen herrührten, auftraten. Allerdings befanden sich die Cluster nach der Inkubation in der Regel unter Aggregationsbedingungen (z. B. 19% $Delta^+$-Zellen in Aggregaten vor der Inkubation, gegenüber 37% $Delta^+$-Zellen in Aggregaten nach Inkubation), was anzeigt, daß die $Delta^+$-Zellen in der Lage sind, bei diesen Tests untereinander stabile Kontakte zu bilden.

[0192] In bemerkenswertem Kontrast zu Kontrollexperimenten mit $Notch^+$-Zellen allein ergab die Aggregation von Gemischen von $Notch^+$- und $Delta^+$-Zellen die Bildung von Clustern von bis zu 20 oder mehr Zellen. Die Fraktion exprimierender Zellen, die nach 24 h Aggregation in Clustern von vier oder mehr angefärbten Zellen vorgefunden wurde, reichte in Gemischen von $Notch^+$- und $Delta^+$-Zellen von 32% bis 54%. Dieser Bereich entsprach demjenigen, der für $Delta^+$-Zellen allein (37–40%) festgestellt wurde, unterschied sich jedoch sehr von demjenigen für $Notch^+$-Zellen allein (nur 0%–5%). Obwohl einige Cluster, die nur aus $Delta^+$-Zellen bestanden, gefunden wurden, wurden in Clustern von mehr als vier bis fünf Zellen niemals $Notch^+$-Zellen gefunden, wenn nicht auch Delta+-Zellen vorhanden waren. Wiederum exprimierten sämtliche Zellen in diesen Clustern entweder Notch oder Delta, auch wenn transfizierte Zellen nur eine kleine Fraktion der Gesamtzellpopulation ausmachten. Nach 48 h war der Aggregationsgrad offenbar höher (63–71%), was nahe legt, daß die Aggregation nach 24 h unter diesen Bedingungen noch kein Maximum erreicht hatte. Ferner aggregierten Zellen, die mit Notch- und Delta-Konstrukten kotransfiziert worden waren (so daß sämtliche transfizierten Zellen beide Proteine exprimieren), unter den gleichen experimentellen Bedingungen auf ähnliche Weise.

[0193] Die Notch-Beteiligung am Aggregationsprozeß wurde direkt durch Überprüfen der Wirkung eines Ge-

misches von polyklonalen Antiseren getestet, die gegen Fusionsproteine gerichtet waren, die fast die gesamte extrazelluläre Domäne von Notch bei der Aggregation umspannten (siehe experimentellen Verfahrensweisen-abschnitt 6.1). Zur Minimierung von Artefakten, die auf Grund einer metabolischen Reaktion das Patching von Oberflächen-Antigene entstehen könnten, wurden bei diesen Experimenten eine Antikörperbehandlung und der Aggregationstest bei 4°C durchgeführt. Notch$^+$-Zellen wurden entweder mit Maus-Vorimmun- oder -Immunserum 1 h inkubiert, Delta$^+$-Zellen wurden zugesetzt, und die Aggregation wurde 1 bis 2 h durchgeführt. Obgleich Notch$^+$-Zellen, die mit Präimmunseren vorbehandelt worden waren, mit Delta$^+$-Zellen aggregierten (bei einem von drei Experimenten, befanden sich 23% der Notch$^+$-Zellen in Notch$^+$-Delta$^+$-Zellaggregaten), erfolgte dies bei denjenigen, die mit Immunserum behandelt worden waren, nicht (nur 2% der Notch$^+$-Zellen befanden sich in Aggregaten). Dieses Ergebnis legt nahe, daß die extrazelluläre Domäne von Notch für die Notch$^+$-Delta$^+$-Zellaggregation erforderliche war.

6.2.3 DIE NOTCH-DELTA-VERMITTELTE AGGREGATION IST CALCIUMABHÄNGIG

[0194] Die Fähigkeit von Zellen zur Aggregation in Gegenwart oder Abwesenheit von Ca^{2+}-Ionen wurde getestet, um zu bestimmen, ob eine Calcium^{2+}-Ionen-Anforderung für die Notch-Delta-Aggregation besteht. Zur Minimierung möglicher nichtspezifischer Effekte aufgrund metabolischer Reaktionen auf die Entfernung von Ca^{2+} wurden diese Experimente bei 4°C durchgeführt. Die Ergebnisse zeigten eindeutig eine Abhängigkeit der Notch-Delta-vermittelten Aggregation von exogenem Ca^{2+}.

6.2.4 NOTCH UND DELTA WECHSELWIRKEN MIT EINER EINZELNEN ZELLE

[0195] Die Frage, ob Notch und Delta innerhalb der Membran von einer Zelle, die beide Proteine exprimiert, assoziiert sind, wurde durch Überprüfen der Verteilungen von Notch und Delta in kotransfizierten Zellen gestellt. Um zu testen, ob die beobachtete Kolokalisation zufällig war oder eine stabile Wechselwirkung zwischen Notch und Delta darstellte, wurden lebende Zellen mit einem Überschuß an polyklonalem Anti-Notch-Antiserum behandelt. Diese Behandlung führte zu einem "Patching" von Notch auf der Oberfläche exprimierender Zellen in diskreten Flicken, wie durch Immunfluoreszenz nachgewiesen. Es bestand eine distinkte Korrelation zwischen den Verteilungen von Notch und Delta auf der Oberfläche dieser Zellen nach dieser Behandlung, was anzeigte, daß diese Proteine in der Membran assoziiert sind.

6.2.5 WECHSELWIRKUNGEN MIT DELTA ERFORDERN DIE INTRAZELLULÄRE DOMÄNE VON NOTCH NICHT

[0196] Zusätzlich zu der großen extrazellulären Domäne, die EGF-artige Sequenzwiederholungen enthält, besitzt Notch eine anpaßbare intrazelluläre (IC)-Domäne von ca. 940 Aminosäuren. Die IC-Domäne umfaßt eine Phosphorylierungsstelle (Kidd et al., 1989, Genes Dev. 3, 1113–1129), eine vermeintliche Nukleotidbindungsdomäne, eine Polyglutaminverlängerung (Wharton et al., 1985, Cell 43, 567–581; Kidd, et al., 1986, Mol. Cell. Biol. 6, 3094–3108) und Sequenzen, die zu dem Hefe-cdc10-Gen homolog sind, das an der Zellzykluskontrolle in Hefe beteiligt ist (Breeden und Nasmyth, 1987, Nature 329, 651–654). Eine Notch-Konstrukt-Variante wurde verwendet, bei der die codierenden Sequenzen für die ca. 835 Aminosäuren der IC-Domäne, einschließlich sämtlicher struktureller Merkmale, die vorstehend vermerkt sind, deletiert worden sind (wobei 25 membranproximale Aminosäuren und ein neuer 59 Aminosäure-Carboxyl-Terminus zurückblieben; siehe experimentielle Verfahrensweisen).

[0197] In Aggregationstests bildeten die Zellen, die das ECN1-Konstrukt exprimierten, mit Delta$^+$-Zellen konsistente Aggregate, jedoch nicht mit sich selbst, ebenso wie es für die Zellen festgestellt wurde, die intaktes Notch exprimierten. Scharfe Banden von angefärbtem ECN1 wurden in Regionen des Kontakts mit Delta$^+$-Zellen festgestellt, was wiederum eine Lokalisierung von ECN1 innerhalb von Regionen des Kontakts zwischen den Zellen anzeigte. Zum Testen auf Wechselwirkungen innerhalb der Membran wurden Oberflächenantigen-Copatching-Experimente unter Verwendung von Zellen durchgeführt, die mit den ECN1- und Delta-Konstrukten kotransfiziert worden sind. Wie für intaktes Notch festgestellt, wenn ECN1 unter Verwendung polyklonaler Antiseren gegen die extrazelluläre Domäne von Notch gepatcht wurde, kolokalisierten ECN1 und Delta an der Zelloberfläche. Diese Ergebnisse zeigen, daß die beobachteten Wechselwirkungen zwischen Notch und Delta in der Membran nicht den deletierten Teil der IC-Domäne von Notch erfordern und darum wahrscheinlich durch die extrazelluläre Domäne vermittelt werden.

6.2.6 NOTCH UND DELTA BILDEN DETERGENSLÖSLICHE INTERMOLEKULARE KOMPLEXE

[0198] Die vorgenannten Ergebnisse zeigten molekulare Wechselwirkungen zwischen Notch und Delta, die

innerhalb der gleichen Membran und zwischen diesen Proteinen vorhanden sind, die auf verschiedenen Zellen exprimiert wurden. Ein weiterer Test solcher Wechselwirkungen besteht in der Kopräzipitation dieser Proteine aus nicht denaturierenden Detergensextrakten von Zellen, die Notch und Delta exprimieren. Wenn Notch und Delta einen stabilen intermolekularen Komplex, entweder zwischen oder innerhalb von Zellen, bilden, sollte es möglich sein, beide Proteine aus Zellextrakten unter Verwendung spezifischer Antiseren, die gegen eines dieser Proteine gerichtet sind, auszufällen. Diese Analyse wurde durch Immunfällung von Delta mit polyklonalem Antiserum aus NP-40/Deoxycholat-Lysaten (siehe experimentelle Verfahren) aus Zellen, die mit den Notch- und Delta-Konstrukten cotransfiziert worden waren, die man über Nacht aggregieren ließ, oder von 0–24 h Wild-Typ-Embryos durchgeführt.

[0199] Die Kopräzipitation von Notch wurde in Delta-Immunpräzipitaten aus kotransfizierten Zellen und Embryonen nachgewiesen. Allerdings ist kopräzipitierendes Notch anscheinend in viel geringeren Mengen vorhanden als Delta, und es war darum schwer nachzuweisen. Die Tatsache, daß die Immunfällung von Delta zur Kopräzipitation von Notch führt, stellt einen direkten Beweis dafür dar, daß diese beiden Proteine in transfizierten S2-Zellen und in Embryonenzellen stabile intermolekulare Komplexe bilden.

6.3 DISKUSSION

[0200] Die Anwendung eines in vitro Aggregationstests, der normalerweise nicht klebende S2-Zellen einsetzt, zeigte, daß Zellen, die Notch und Delta exprimieren, spezifisch aneinander haften.

7. EGF-SEQUENZWIEDERHOLUNGEN 11 UND 12 VON NOTCH SIND ZUR NOTCH-DELTA-VERMITTELTEN AGGREGATION ERFORDERLICH UND AUSREICHEND

[0201] Der gleiche Aggregationstest, wie in Abschnitt 6 beschrieben, wurde zusammen mit Deletionsmutanten von Notch zur Identifizierung von Regionen in der extrazellulären Domäne von Notch, die für die Wechselwirkungen mit Delta notwendig sind, verwendet. Der Beweis zeigt, daß die EGF-Sequenzwiederholungen von Notch direkt an dieser Wechselwirkung beteiligt sind und daß nur zwei (ELR-11 und 12) der 36 EGF-Sequenzwiederholungen notwendig erscheinen. Diese zwei EGF-Sequenzwiederholungen sind zum Binden an Delta und dafür, daß die Calcium-Abhängigkeit der Notch-Delta-vermittelten Aggregation ebenfalls mit diesen zwei Sequenzwiederholungen assoziiert ist. Schließlich vermitteln die zwei entsprechenden EGF-Sequenzwiederholungen aus dem Xenopus-Homolog von Notch auch die Aggregation mit Delta, was impliziert, das nicht nur die Struktur von Notch, sondern auch ihre Funktion evolutionär konserviert worden ist. Diese Ergebnisse legen nahe, daß die extrazelluläre Domäne von Notch überraschenderweise modular ist und zusätzlich zu Delta potentiell eine Vielzahl von Proteinen binden könnte. (See Rebay et al., 1991, Cell 67: 687–699).

7.1 EXPERIMENTELLE VERFAHREN

7.1.1 EXPRESSIONSKONSTRUKTE

[0202] Die beschriebenen Konstrukte sind alles Derivate des Notch-Vollängen-Expressionsprodukts Nr. 1, pMtNMg (siehe Abschnitt 6, vorstehend), und wurden wie beschrieben (Rebay et al., 1991, Cell 67: 687–699) hergestellt.

7.1.2 ZELLKULTUR UND TRANSFEKTION

[0203] Die Drosophila-S2-Zell-Linie wurde gezüchtet und wie in Abschnitt 6, vorstehend beschrieben, transfiziert. Die Deltaexprimierende stabil transformierte S2-Zelllinie L-49-6-7 (freundlicherweise zu Verfügung gestellt von L. Cherbas), wurde in M3-Medium (hergestellt von Hazleton Co.), angereichert mit 11% Hefe-inaktiviertem fetalem Kälberserum (FCS) (Hyclone), 100 U/ml Penizillin-100 µg/ml Streptomycin-0,25 µg/ml Fungizone (Hazelton), 2×10^{-7} M Methotrexat, 0,1 mM Hypoxanthin und 0,016 mM Thymidin, gezüchtet.

7.1.3 AGGREGATIONSTESTS UND IMMUNFLUORESZENZ

[0204] Aggregationsstest und Ca^{++}-Abhängigkeitsexperimente wurden, wie vorstehend in Abschnitt 6 beschrieben, durchgeführt. Die Zellen wurden mit dem monoklonalen Anti-Notch-Antikörper 9C6.C17 und polyklonalem Anti-Delta-Ratten-Antiserum (ausführlich in Abschnitt 6, vorstehend beschrieben), angefärbt. Die Oberflächenexpression von Notch-Konstrukten in nicht permeabilisierten Zellen wurde unter Verwendung von polyklonalem Ratten-Antiserum, das gegen das 0,8 kb (Aminosäuren 237–501; Wharton et al., 1985, Cell 43, 567–581)-BstY1-Fragment aus der extrazellulären Domäne von Notch gezüchtet worden war, durchgeführt.

Die Zellen wurden unter Epifluoreszenz auf einem Leitz-Orthoplan-2-Mikroskop betrachtet.

7.2 ERGEBNISSE

7.2.1 EGF-SEQUENZWIEDERHOLUNGEN 11 UND 12 VON NOTCH SIND ZUR NOTCH-DELTA-VERMITTELTEN AGGREGATION ERFORDERLICH

[0205] Eine breit angelegte Deletionsanalyse wurde für die extrazelluläre Domäne des Notch-Proteins, von der sich erwiesen hatte, daß sie an den Notch-Delta-Wechselwirkungen beteiligt war (vorstehend Abschnitt 6; Fehon et al., 1990, Cell 61: 523–534), zur Identifizierung der exakten Domäne von Notch, die diese Wechselwirkungen vermittelt, durchgeführt. Die Fähigkeit von Zellen, die mit den verschiedenen Deletionskonstrukten transfiziert worden waren, zur Wechselwirkung mit Delta, wurde unter Anwendung des Aggregationstest, der in Abschnitt 6 beschrieben ist, getestet. Kurz gesagt, wurden Notch-Deletionskonstrukte vorübergehend in Drosophila-S2-Zellen transfiziert, mit $CuSO_4$ induziert und anschließend über Nacht bei Raumtemperatur mit einer kleinen Menge an Zellen aus der stabil transformierten Delta-exprimierenden Zell-Linie L49-6-7 (Cherbas) aggregiert, was eine Population ergab, die typischerweise aus ca. 1% Notch-exprimierenden Zellen und ca. 5% Delta-exprimierenden Zellen bestand, wobei die restlichen Zellen keines der beiden Proteine exprimierten.

[0206] Schematische Zeichnungen der getesteten Konstrukte und die Ergebnisse der Aggregationsexperimente sind in **Fig.** 2 gezeigt. Zum Testen des Aggregationsgrads wurden die Zellen mit Antiseren, die auf jedes Genprodukt spezifisch waren, angefärbt und mit Immunfluoreszenzmikroskopie geprüft. Die Aggregate wurden als Cluster von vier oder mehr Zellen definiert, die sowohl Notch- als auch Delta-exprimierende Zellen enthielten, und die in **Fig.** 2 gezeigten Werte stellen den Prozentsatz von sämtlichen Notch-exprimierenden Zellen dar, der in solchen Clustern festgestellt wurde. Sämtliche Zahlen spiegeln das Durchschnittsergebnis von mindestens zwei getrennten Transfektionsexperimenten wider, wobei mindestens 100 Notch-exprimierende Zelleinheiten (entweder einzelne Zellen oder Cluster) bewertet wurden.

[0207] Die Ursprungskonstrukte (Nr. 2 DSph und Nr. 3 ΔCla) deletierten große Teile der EGF-Sequenzwiederholungen. Ihre Fähigkeit zur Beschleunigung der Notch-Delta-Aggregation legte nahe, daß die EGF-Sequenzwiederholungen von Notch an der Wechselwirkung mit Delta beteiligt waren. Eine Reihe von sechs rahmenkonformen ClaI-Restriktionsstellen wurde zur weiteren Aufspaltung des Bereiches der Region zwischen den EGF-Sequenzwiederholungen 7 und 30 verwendet. Aufgrund einer Sequenzhomologie zwischen den Sequenzwiederholungen traten fünf der ClaI-Stellen an der gleichen relativen Stelle in der EGF Sequenzwiederholung, unmittelbar nach dem dritten Cystein auf, während die sechste Stelle unmittelbar vor dem ersten Cystein der EGF-Sequenzwiederholung 31 auftrat (Fig. 3). Somit wurden durch Durchführung eines ClaI-Teilverdaus und durch anschließende Religation Deletionen erhalten, die nicht nur den offenen Leserahmen des Notch-Proteins beibehielten, sondern zusätzlich häufig die Strukturintegrität aufrecht erhielten und den Abstand der drei Disulfidbindungen in den chimären EGF-Sequenzwiederholungen, die durch die Religation produziert worden waren (**Fig.** 2, Konstrukte-Nr. 4–14), mindestens theoretisch konservierten. Leider war die am meisten 3' gelegene ClaI-Stelle gegenüber dem Abbau resistent, während die nächste, 3'-nahe ClaI-Stelle die EGF-Sequenzwiederholungen 30 und 31 aufbrach. Wenn darum verschiedene ClaI-Verdauungsfragmente wieder in das Rahmenwerk des vollständigen ClaI-Verdaus (Konstrukt Nr. 3, Δ-Cla) inseriert wurden, war die Gesamtstruktur der EGF-Sequenzwiederholungen offensichtlich an der 3'-Verbindungsstelle unterbrochen.

[0208] Mehrere Punkte hinsichtlich dieser Serie von Konstrukten sind bemerkenswert. Zuerst hob die Entfernung des ClaI-Restriktionsfragments, das die EGF-Sequenzwiederholungen 9 und 17 aufbricht (Konstrukt Nr. 8, ΔEGF9–17), die Aggregation mit Delta auf, während die Reinsertion dieses Stücks in Konstrukt Nr. 3 Δ-Cla, dem die EGF-Sequenzwiederholungen 7–30 fehlen, die Aggregation etwa auf Wild-Typ-Niveau wieder herstellte (Konstrukt Nr. 13, Δ-Cla + EGF9–17), was nahe legt, daß die EGF-Sequenzwiederholungen 9–17 Sequenzen enthalten, die zur Bindung an Delta wichtig sind. Zweitens standen sämtliche Konstrukte dieser Serie (Nr. 4–14) mit der Bindungsstellen-Zuordnung zu den EGF-Sequenzwiederholungen 9–17 im Einklang. Expressionskonstrukte, die diese Sequenzwiederholungen enthalten (Nr. 6, 7, 9, 10, 13), beschleunigten die Notch-Delta-Wechselwirkungen, während bei Konstrukten, denen diese Sequenzwiederholungen fehlten (Nr. 4, 5, 8, 11, 12, 14), dies nicht auftrat. Um zu bestätigen, daß das Unvermögen zur Aggregation mit Delta-Zellen nicht einfach auf dem Unvermögen des mutagenierten Notch-Proteins zum Erreichen der Zelloberfläche beruhte, sondern tatsächlich die Deletion der notwendigen Bindungsstelle widerspiegelte, wurde die Zelloberflächenexpression sämtlicher Konstrukte durch Immunfluoreszenzanfärben lebender transfizierter Zellen mit Antikörpern, die spezifisch auf die extrazelluläre Domäne von Notch sind, getestet. Sämtlich Konstrukte versagten bei der Vermittlung von Notch-Delta-Wechselwirkungen, die ein Protein produzierte, das anscheinend normaler-

weise an der Zelloberfläche exprimiert wird. Obwohl der Aggregationstest nicht quantitativ ist, aggregierten drittens zwei Konstrukte, die die EGF-Sequenzwiederholungen 9–17, Nr. 9 ΔEGF17–26, oder am bemerkenswertesten Nr. 10 ΔEGF26–30 enthielten, auf einem offenbar niedrigeren Niveau. Die Zellen, die mit den Konstrukten-Nr. 9 ΔEGF17–26 und 10 ΔEGF26–30 transfiziert worden waren, zeigten eine beträchtlich geringere Oberflächenanfärbung als normale Zellen, obwohl fixierte und permeabilisierte Zellen, die mit dem gleichen Antikörper reagierten, sich normalerweise färbten, was zeigt, daß die Epitope, die von den Antiseren erkannt wurden, nicht einfach deletiert worden sind. Durch Vergleich des Prozentgehalts transfizierter Zellen in entweder permeabilisierten oder lebenden Zellpopulationen, wurde festgestellt, daß für das Konstrukt Nr. 9 ΔEGF17–26 etwa 50% der transfizierten Zellen und für das Konstrukt Nr. 10 ΔEGF26–30 10% nachweisbares Protein an der Zelloberfläche produzierte. Somit produzierten diese beiden Konstrukte Proteine, die oft nicht die Zelloberfläche zu erreichen vermochten, vielleicht aufgrund von Fehlfaltung, wodurch die Fähigkeit von transfizierten Zellen zur Aggregation mit Delta-exprimierenden Zellen vermindert, jedoch nicht aufgehoben war.

[0209] Nach dem Einordnen der Bindungsstelle in die EGF-Sequenzwiederholungen 9 bis 17, zeigten weitere Experimente (Rebay et al., 1991, Cell 67: 687–699), daß die EGF-Sequenzwiederholung 14 von Notch nicht an den Wechselwirkungen mit Delta, die im Modellversuch mit dem Gewebekulturtest dargestellt wurden, beteiligt war.

[0210] Um weiterhin die Delta-Bindungsdomäne in die EGF-Sequenzwiederholungen 9–17 einzuordnen, wurden spezielle Oligonukleotid-Primer und die PCR-Technik zur Erzeugung mehrerer Unterfragmente dieser Region eingesetzt. Drei überlappende Konstrukte Nr. 16, 17 und 18 wurden hergestellt, wovon nur eines, Nr. 16 ΔCla + EGF9–13, bei Transfektion in S2-Zellen die Aggregation mit Delta-Zellen zuließ. Konstrukt Nr. 19, ΔCla + EGF(10–13), dem die EGF-Sequenzwiederholung 9 fehlte, definierte weiterhin die EGF-Sequenzwiederholungen 10–13 als die für die Notch-Delta-Wechselwirkungen notwendige Region.

[0211] Die Konstrukte Nr. 20–24 stellten Versuche dar, diese Domäne unter Anwendung der gleichen PCR-Strategien noch weiter abzubauen (siehe **Fig. 3**). Das Konstrukt Nr. 20 ΔCla + EGF(11–13), wobei die EGF-Sequenzwiederholung 12 die einzige vollständige Sequenzwiederholung ist, die hinzugefügt wurde, und das Konstrukt Nr. 21 ΔCla + EGF(10–12), wobei die EGF-Sequenzwiederholung 11 die einzige vollständige Sequenzwiederholung ist, die hinzugefügt wurde, vermittelten die Aggregation nicht, was nahe legt, daß das Vorliegen entweder der EGF-Sequenzwiederholung 11 oder 12 allein für die Notch-Delta-Wechselwirkungen nicht ausreichte. Da allerdings die 3'-Ligationsverbidnungsstelle dieser Konstrukte die Gesamtstruktur der EGF-Sequenzwiederholungen unterbrach, war es möglich, daß eine kurze "Pufferzone" benötigt wurde, um ein normales Funktionieren der essentiellen Sequenzwiederholung zu ermöglichen. Somit ist beispielsweise in Konstrukt Nr. 19 ΔCla + EGF(10–13) die EGF-Sequenzwiederholung 12 vielleicht nicht direkt an der Bindung mit Delta beteiligt, sondern könnte statt dessen zu der minimale Menge an Puffersequenz beitragen, die zum Schutze der Struktur der EGF-Sequenzwiederholung 11 benötigt wird, wodurch Wechselwirkungen mit Delta möglich sind. Die Konstrukte Nr. 22–24 sprechen dieses Thema an. Die Konstrukte Nr. 22 ΔCla + EGF(10–11), die keine Aggregation vermittelten, und Nr. 23 ΔCla + EGF(10–12), die diese vermittelten, legten wiederum nahe, daß beide Sequenzwiederholungen 11 und 12 erforderlich sind, während die flankierende Sequenz der Sequenzwiederholung 13 eindeutig nicht erforderlich ist. Schließlich zeigte das Konstrukt Nr. 24 ΔCla + EGF(11–12), obwohl nun an der 5'-Verbindungsstelle potentiell strukturell unterbrochen, überzeugend, daß die Sequenzen aus der EGF-Sequenzwiederholung 10 nicht essentiell sind. Basierend auf den vollkommen konsistenten Daten aus 24 Konstrukten definieren somit die EGF-Sequenzwiederholungen 11 und 12 von Notch zusammen die aus dieser Analyse erhältliche kleinste funktionelle Einheit, die die notwendigen Stellen zur Bindung an Delta in transfizierten S2-Zellen enthält.

7.2.2 EGF-SEQUENZWIEDERHOLUNGEN 11 UND 12 VON NOTCH SIND FÜR DIE NOTCH-DELTA VERMITTELTE AGGREGATION AUSREICHEND

[0212] Die große ClaI-Deletion in die PCR-Fragmente inseriert wurden (Nr. 3 ΔCla), behält grob gesagt ein Drittel der ursprünglichen 36 EGF-Sequenzwiederholungen, sowie drei Notch/lin-12-Sequenzwiederholungen bei. Während diese eindeutig zur Unterstützung der Aggregation nicht ausreichen, ist es möglich, daß sie ein notwendiges Netzwerk bilden in dem spezifische EGF-Sequenzwiederholungen mit Delta Wechselwirken. Um zu testen, ob tatsächlich nur einige EGF-Sequenzwiederholungen zur Unterstützung der Aggregation ausreichend waren, wurden zwei Konstrukte, Nr. 25 ΔEGF, bei dem sämtliche 36 EGF-Sequenzwiederholungen deletiert waren, mit der Ausnahme der ersten zwei Drittel von Sequenzwiederholung 1, und Nr. 30 ΔECN, bei dem der gesamte extrazellulären Teil von Notch deletiert war, mit der Ausnahme des ersten Drittels der EGF-Sequenzwiederholung 1 und der ca. 35 Aminosäuren unmittelbar vor der Transmembran-Domäne, konstruiert.

Fragmente, die die Notch-Delta-Aggregation im Hintergrund von Konstrukt Nr. 3 ΔCla bei der Insertion in Konstrukt Nr. 25 ΔEGF vermittelten hatten, waren wiederum in der Lage, die Interaktionen mit Delta zu fördern (Konstrukte Nr. 26–30). Analoge Konstrukte (Nr. 31, 32), in denen die Notch/lin-12-Sequenzwiederholungen ebenfalls fehlten, vermittelten die Notch-Delta-Aggregation wiederum erfolgreich. Somit funktionieren die EGF-Sequenzwiederholungen 11 und 12 anscheinend als unabhängige molekulare Einheiten, die zur Vermittlung von Notch-Delta-Wechselirkungen in S2-Zellen ausreichen, auch in Abwesenheit des größten Teils der extrazellulären Domäne von Notch.

7.2.3 Die EGF-SEQUENZWIEDERHOLUNGEN 11 und 12 VON NOTCH HALTEN DIE CALCIUMABHÄNGIGKEIT DER NOTCH-DELTA VERMITTELTEN AGGREGATION AUFRECHT

[0213] Die Fähigkeit von Zellen, die bestimmte Deletionskonstrukte exprimieren, zur Aggregation mit Delta-exprimierenden Zellen, wurde in Gegenwart oder Abwesenheit von Ca^{++}-Ionen überprüft. Die Calciumabhängigkeit der Wechselwirkung wurde sogar im kleinsten Konstrukt bewahrt, was mit der Feststellung Einklang steht, daß die minimalen Konstrukte, die die EGF-Sequenzwiederholungen 11 und 12 enthalten, an Delta auf eine Weise binden, die derjenigen des Vollängen-Notch entspricht.

7.2.4 DIE DELTA-BINDUNGSFUNKTION DER EGF-SEQUENZWIEDERHOLUNGEN 11 UND 12 VON NOTCH IST IM XENOPUS-HOMOLOGEN VON NOTCH KONSERVIERT

[0214] PCR-Primer auf der Basis der Xenopus-Notch-Sequenz (Coffman et al., 1990, Science 249, 1438–1441) wurde zum Erhalt eines ca. 350 bp-Fragments aus einer Xenopus-Stadium 17-cDNA-Bibliothek verwendet, die die EGF-Sequenzwiederholungen 11 und 12, flankiert jeweils von der Hälfte der Sequenzwiederholungen 10 und 13 auf einer Seite, enthielt. Diese Fragmente wurden in Konstrukt Nr. 3 ΔCla kloniert, und drei unabhängige Klone wurden auf die Fähigkeit zur Wechselwirkung mit Delta im Zellkulturaggregationstest getestet. Zwei der Klone Nr. 33a & b ΔCla + XEGF(10–13) waren bei Transfektion in S2-Zellen in der Lage, Notch-Delta-Wechselwirkungen auf einem Niveau, das in etwa demjenigen des analogen Drosophila-Notch-Konstrukts Nr. 19 ΔCla + XEGF(10–13) entsprach, wiederum Calcium-abhängig, zu vermitteln (Tabelle III). Allerdings vermocht der dritte Klone Nr. 33c ΔCla + XEGF(10–13) nicht, die Notch-Delta-Wechselwirkungen zu vermitteln, obwohl das Protein normalerweise an der Zelloberfläche exprimiert wurde, wie durch Anfärben von lebenden unpermeabilisierten Zellen bewertet. Der Sequenzvergleich des Xenopus-PCR-Produktes in den Konstrukten Nr. 33a und 33c ergab eine Fehlsinn-Mutation, die zu einer Leucin-zu-Prolin-Änderung (Aminosäure Nr. 453, Coffman, et al., 1990, Science 249, 1438–1441) in der EGF-Sequenzwiederholung 11 des Konstrukts Nr. 33c führte. Obwohl dieser Rest nicht zwischen Drosophila- und Xenopus-Notch konserviert ist (**Fig.** 8), könnte der Einbau eines Prolinrestes leicht die Struktur der EGF-Sequenzwiederholung auseinanderbrechen und sie somit an der richtigen Wechselwirkung mit Delta hindern.

[0215] Der Vergleich der Aminosäuresequenz der EGF-Sequenzwiederholungen 11 und 12 von Drosophila- und Xenopus-Notch ergibt einen hohen Aminosäure-Identitätsgrad, einschließlich der verwandten, calciumbindenden Sequenz (**Fig. 4**, SEQ ID NO: 1 und NR. 2). Allerdings ist das Homologieniveau von demjenigen, der den meisten anderen EGF-Sequenzwiederholungen gemeinsam ist, die insgesamt etwa 50%ige Identität auf Aminosäureniveau aufweisen, nicht auffallend verschieden. Diese eins-zu-eins Entsprechung zwischen den einzelnen EGF-Sequenzwiederholungen von Drosophila- und Xenopus-Notch, zusammen mit der funktionellen Konservierung von ELR-11 und 12, legt nahe, daß die 36 EGF-Sequenzwiederholungen von Notch einen Bereich mit Tandem-Anordnung von konservierten funktionellen Einheiten enthalten.

7.3 DISKUSSION

[0216] Eine umfassende Deletionsanalyse der extrazellulären Domäne von Notch wurde angewandt, um zu zeigen, daß die Regionen von Notch, die die EGF-homologen Sequenzwiederholungen 11 und 12 enthalten, für die Notch-Delta-vermittelte Aggregation sowohl notwendig als auch ausreichend sind und daß dieses Delta-Bindungsvermögen in den gleichen beiden EGF-Sequenzwiederholungen von Xenopus-Notch bewahrt wurde. Die Feststellung, daß die Aggregation, die bei den EGF-Sequenzwiederholungen 11 und 12 von Notch zugeordnet wurde, zeigt, daß die EGF-Sequenzwiederholungen von Notch auch als spezifische Proteinbindungsdomänen fungieren. Die EGF-Sequenzwiederholungen 11 und 12 allein (Nr. 32 ΔECN + EGF(11–12)) waren zur Aufrechterhaltung der Ca^{++}-Abhängigkeit der Notch-Delta-Wechselwirkungen ausreichend.

[0217] Die verschiedenen Deletionskonstrukte legen nahe, daß ELR-11 und ELR-12 als modulare Einheit funktionieren, unabhängig vom unmittelbaren Zusammenhang, in dem sie angeordnet sind. Somit sind weder die restlichen 34 EGF-Sequenzwiederholungen noch die drei Notch/lin-12-Sequenzwiederholungen offen-

sichtlich zum Erstellen eines strukturellen Rahmens notwendig, der für die EGF-SequenzSequenzwiederholungen 11 und 12 zum Funktionieren erforderlich ist. Interessanterweise wurde fast der gegenteilige Effekt festgestellt: obwohl der Aggregationstest nicht die Stärke der Wechselwirkung mißt, da sich die Bindungsstelle zu immer kleiner werdenden Fragmenten verschmälerte, wurde eine Zunahme in der Fähigkeit der transfizierten Zellen zur Aggregation mit Delta-exprimierenden Zellen festgestellt, was nahe legt, daß die normalen flankierenden EGF-Sequenzen tatsächlich die Assoziation zwischen den Proteinen behindern. Die verbleibenden 34 EGF-Sequenzwiederholungen können auch modulare Bindungsdomänen für andere Proteine bilden, die mit Notch zu verschiedenen Zeiten während der Entwicklung Wechselwirken.

[0218] Die Feststellung, daß die EGF-Sequenzwiederholungen 11 und 12 von Notch eine diskrete Delta-Bindungseinheit bilden, stellt den ersten konkreten Beweis dar, der die Idee unterstützt, daß jede EGF-Sequenzwiederholung oder kleine Untereinheit von Sequenzwiederholungen eine einzigartige Rolle während der Entwicklung spielen kann, möglicherweise durch direkte Wechselwirkungen mit anderen Proteinen. Die zwischen der klebenden Domäne von Delta und Serrate festgestellten Homologien (**Fig.** 5), legen nahe, daß der homologe Teile von Serrate insofern "klebend" ist, als er die Bindung an andere toporhythmische Proteine vermittelt (siehe Abschnitt 8, nachstehend). Zusätzlich kann das Gen Scabrous, das ein sezerniertes Protein mit Ähnlichkeit zu Fibrinogen codiert, mit Notch Wechselwirken.

[0219] Zusätzlich zu der EGF-Sequenzwiederholung treten in der Regel mehrere Kopien anderer Strukturmotive in einer Vielzahl von Proteinen auf. Ein relevantes Beispiel ist das cdc10/Ankyrin-Motiv, wovon sechs Kopien in der intrazellulären Domäne von Notch gefunden werden. Ankyrin enthält 22 dieser Sequenzwiederholungen. Vielleicht können wiederholte Anordnungen von strukturellen Motiven in der Regel einen linearen Aufbau einer Serie von modularen Proteinbindungseinheiten darstellen. Angesichts dieser Ergebnisse zusammen mit der bekannten strukturellen genetischen und entwicklungsbedingten Komplexität von Notch, kann Notch während der Entwicklung mit einer Anzahl von verschiedenen Liganden in einem exakt regulierten zeitlichen und räumlichen Muster Wechselwirken. Solche Kontext-spezifische Wechselwirkungen mit extrazellulären Proteinen könnten durch die EGF- und Notch/lin-12-Sequenzwiederholungen vermittelt werden, während die Wechselwirkungen mit Cytoskelett- und Cytoplasma-Proteinen durch die intrazellulären cdc10/Ankyrin-Motive vermittelt werden könnten.

8. SEQUENZEN DIE NOTCH-SERRATE-WECHSELWIRKUNGEN VERMITTELN

[0220] Wie hier beschrieben, bauen die zwei EGF-Sequenzwiederholungen von Notch, die die Wechselwirkungen mit Delta vermitteln, nämlich die EGF-Sequenzwiederholungen 11 und 12, auch eine Serrate-Bindungsdomäne auf (siehe Rebay et al., 1991, Cell 67: 687–699).

[0221] Um zu testen, ob Notch und Serrate direkt Wechselwirken, wurden S2-Zellen mit einem Serrate-Expressionskonstrukt transfiziert und mit Notch-exprimierenden Zellen in einem Aggregationstest gemischt. Für das Serrate-Expressionskonstrukt wurde ein synthetischer Primer, der eine künstliche BamHI-Stelle unmittelbar 5' zum Initiator AUG an Position 442 enthielt (sämtliche Sequenznummern sind nach Fleming et al., 1990, Genes & Dev. 4: 2188–2201), und bis zu Position 464 homolog ist in Verbindung mit einem zweiten Primer von Position 681–698 unter Erzeugung eines DNA-Fragments von ca. 260 Basenpaaren verwendet. Dieses Fragment wurde mit BamHI und KpnI (Position 571) geschnitten und in Bluescript-KS+(Stratagene) ligiert. Dieses Konstrukt, BTSer5'PCR, wurde durch Sequenzieren geprüft, anschließend mit KpnI geschnitten. Das Serrate-KpnI-Fragment (571–2981) wurde inseriert und die richtige Orientierung selektiert, um BTSer5'PCR-Kpn zu erzeugen. Das 5'-SacII-Fragment von BTSer5'PCR-Kpn (SacII-Stellen im Bluescript-Polylinker und in Serrate (1199)) wurde isoliert und zum Ersatz des 5'SacII-Fragments von cDNA-C1 verwendet (Fleming et al., 1990, Genes & Dev. 4: 2188–2201), wodurch die Vollängen-Serrate cDNA abzüglich der 5'-untranslatierten Regionen erzeugt wurde. Diese Insertion wurde durch einen SalI- und teilweisen BamHI-Verdau isoliert und zwischen den BamHI- und SalI-Stellen von pRmHa-3 unter Erzeugung des fertigen Expressionskonstrukts Ser-mtn eingebaut.

[0222] Serrate-exprimierende Zellen hafteten Calcium-abhängig an Notch-exprimierenden Zellen (**Fig.** 2 und Rebay et al., 1991, supra). Im Gegensatz zu Delta wechselwirkte Serrate allerdings unter den getesteten experimentellen Bedingungen anscheinend nicht homotypisch. Zusätzlich wurden keine Wechselwirkungen zwischen Serrate und Delta nachgewiesen.

[0223] Eine Unterserie von Notch-Deletionskonstrukten wurde getestet und zeigte, daß die EGF-Sequenzwiederholungen 11 und 12, zusätzlich zum Binden an Delta, auch Wechselwirkungen mit Serrate vermitteln (**Fig.** 2; Konstrukte Nr. 1, 7–10, 13, 16, 17, 19, 28 und 32). Zusätzlich wurde die Serrate-Bindungsfunktion die-

ser Sequenzwiederholungen offensichtlich auch in den entsprechenden zwei EGF-Sequenzwiederholungen von Xenopus-Notch (Nr. 33 ΔCla + XEGF(10–13)) konserviert. Diese Ergebnisse zeigen unzweideutig, daß Notch sowohl mit Delta als auch Serrate wechselwirkt und daß die gleichen zwei EGF-Sequenzwiederholungen von Notch beide Wechselwirkungen vermitteln. Die Serrate-Region, die für die Notch/Serrate-Aggregation essentiell ist, wurde ebenfalls definiert. Die Deletion der Nukleotide 676–1287 (d. h. die Aminosäuren 79–282) (siehe **Fig.** 5; SEQ ID NO: 3 und NR. 4) beseitigte die Fähigkeit des Serrate-Proteins zur Aggregation mit Notch.

[0224] Notch und Serrate aggregieren anscheinend weniger effizient als Notch und Delta, vielleicht aufgrund der schwächeren Notch-Serrate-Wechselwirkung. Eine triviale Erklärung für diese verminderte Aggregationsmenge könnte sein, daß das Serrate-Konstrukt einfach nicht so viel Protein wie das Delta-Konstrukt an der Zelloberfläche exprimierte, wodurch die Stärke der Wechselwirkung vermindert wird. Alternativ kann der Unterschied in der Stärke der Wechselwirkung eine fundamentalen funktionellen Differenz zwischen Notch-Delta- und Notch-Serrate-Wechselwirkungen anzeigen, die in vivo signifikant sein kann.

9. KLONIERUNG, SEQUENZIERUNG UND EXPRESSION VON MENSCHLICHEM NOTCH

9.1 ISOLIERUNG UND SEQUENZIERUNG VON MENSCHLICHEM NOTCH

[0225] Zunächst wurden Klone für die menschliche Notch-Sequenz unter Anwendung der Polymerasekettenreaktion (PCR) zur Amplifikation von DNA aus einer 17–18 Wochen alten menschlichen, fetalen Gehirn-cDNA-Bibliothek im Lambda-ZapII-Vektor (Stratagene) erhalten.

[0226] Das auf diese Weise erhaltene 400 bp-Fragment wurde anschließend als Sonde verwendet, mit der die gleiche Bibliothek nach menschlichen Notch-Klonen durchmustert wurde. Das erste Durchmustern ergab drei einzigartige Klons, hN3k, hN2k und hN5k, die alle durch anschließende Sequenzanalyse ergaben, daß sie in daß 3'-Ende von menschlichem Notch (Fig. 6) fallen. Ein zweites Durchmustern unter Verwendung des 5'-Endes von hN3k als Sonde wurde durchgeführt, um nach Klonen zu suchen, die das 5'-Ende von menschlichem Notch enthalten. Ein einziger Klon, hN4k, wurde durch diese Durchmusterung erhalten, und die vorläufigen Sequenzierungsdaten zeigen, daß er den größten Teil des 5'-Endes des Gens enthält (Fig. 6). Zusammen enthalten die Klone hN4k, hN3k und hN5k etwa 10 kb des menschlichen Notch-Homologen (der Homologen), beginnend frühzeitig in der EGF-Sequenzwiederholung und bis in die untranslatierte 3'-Region des Gens. Alle drei Klone sind cDNA-Insertionen an der EcoRI-Stelle von pBluescript-SK$^-$ (Stratagene). Der E. coli-Wirtsstamm ist XL1-Blue (siehe Maniatis, T., 1990, Molecular Cloning. A Laboratory Manual, 2. Ausg., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, S. A12). Eine Aneinanderreihung der menschlichen Notch-Sequenzen mit Drosophila-Notch ist in Fig. 7 gezeigt.

[0227] Die Sequenz von verschiedenen Teilen von Notch, die in den cDNA-Klonen erhalten war, wurde bestimmt (unter Verwendung von Sequenase®, U. S. Biochemical Corp.) und ist für hN2k bzw. hN4k in den **Fig.** 8 (SEQ ID NO: 5–7) und 9 (SEQ ID NO: 8, 9) gezeigt. Die weitere Sequenzanalyse von hN2k ergab, daß es eine menschliche Notch-Sequenz codierte, die mit derjenigen überlappt, die in hN5k enthalten ist.

[0228] Die vollständigen Nukleotidsequenzen der menschlichen Notch- cDNA, die in hN3k und hN5k enthalten war, wurde durch das Disdesoxykettenterminationsverfahren unter Verwendung des Sequenase®-Kit (U. S. Biochemical Corp.) bestimmt. Diejenigen Nukleotidsequenzen, die menschliches Notch in dem entsprechenden Leserahmen codieren, wurden leicht identifiziert, da keine Introns vorhanden sind und die Translation in nur einem der drei möglichen Leserahmen eine Sequenz ergibt, die bei Vergleich mit der veröffentlichten von Drosophila-Notch abgeleiteten Aminosäuresequenz eine Sequenz mit einem wesentlichen Homologiegrad mit der Drosophila-Notch-Sequenz ergibt. Die DNA und die abgeleiteten Proteinsequenzen der menschlichen Notch-cDNA in hN3k und hN5k sind in den **Fig.** 10 (SEQ ID NO: 10, 11) und 11 (SEQ ID NO: 12, 13) dargestellt. Der Klon hN3k codiert einen Teil eines Notch-Polypeptids, beginnend bei ungefähr der dritten Notch/lin-12-Sequenzwiederholung bis mehrere Aminosäuren kurz vor der Carboxyl-terminalen Aminosäure. Klon hN5k codiert einen Teil eines Notch-Polypeptids, beginnend ungefähr vor der cdc10-Region bis zum Ende des Polypeptids und enthält auch eine untranslatierte 3'-Region.

[0229] Der Vergleich der cDNA und Proteinsequenzen, die in **Fig.** 10 dargestellt sind (SEQ ID NO: 10, 11), mit denjenigen in **Fig.** 11 (SEQ ID NO: 12, 13), ergibt wesentliche Unterschiede zwischen den Sequenzen, was nahe legt, daß hN3k und hN5k einen Teil von zwei distinkten Notch-homologen Genen darstellen. Die Daten legen somit nahe, daß das menschliche Genom mehr als ein Notch-homologes Gen beherbergt. Dies steht im Gegensatz zu Drosophila, wo Notch offenbar ein Einzelkopie-Gen ist.

[0230] Der Vergleich der DNA- und Aminosäuresequenzen der menschlichen Notch-Homologen, die in hN3k und hN5k enthalten sind, mit den entsprechenden Drosophila-Notch-Sequenzen (wie veröffentlicht bei Wharton et al., 1985, Cell 43: 567–581) und mit den entsprechenden Xenopus-Notch-Sequenzen (wie veröffentlicht bei Coffman et al., 1990, Science 249: 1438–1441 oder erhältlich von Genbank® (Zugangsnummer M33874)) ergab ebenfalls Unterschiede.

[0231] Die in **Fig.** 10 gezeigte Aminosäuresequenz (hN3k) wurde mit der von Ellisen et al., August 1991, Cell 66: 649–661, vorhergesagten Sequenz des TAN-1-Polypeptids verglichen, die in **Fig.** 2 gezeigt ist. Zwischen den abgeleiteten Aminosäuresequenzen wurden einige Unterschiede festgestellt, allerdings ist die hN3k-Notch-Polypeptidsequenz überall zu 99% mit der entsprechenden TAN-1-Region (TAN-1, Aminosäuren 1455 bis 2506) identisch. Vier Unterschiede wurden festgestellt: (1) in der Region zwischen der dritten Notch/lin-12-Sequenzwiederholung und dem ersten cdc10-Motiv befindet sich ein Arginin (hN3k) anstelle eines X (TAN-1, Aminosäure 1763); (2) es existiert ein Prolin (hN3k) anstelle eines X (TAN-1, Aminosäure 1787); (3) es existiert eine konservative Änderung eines Aspartamsäurerestes (hN3k) anstelle eines Glutaminsäurerestes (TAN-1, Aminosäure 2495); und (4) die Carboxylterminale Region unterscheidet sich wesentlich zwischen den TAN-1-Aminosäuren 2507 und 2535.

[0232] Die in **Fig.** 11 gezeigte Aminosäuresequenz (hN5k) wurde mit der von Ellisen et al., August 1991, Cell 66: 649–661, vorhergesagten Sequenz des TAN-1-Polypeptids, die in **Fig.** 2 gezeigt ist, verglichen. Zwischen den abgeleiteten Aminosäuresequenzen wurden Unterschiede festgestellt. Das abgeleitete Notch-Polypeptid von hN5k ist mit dem TAN-1-Polypeptid (64% identisch zu Drosophila-Notch) in der cdc10-Region, die sowohl das cdc10-Motiv (TAN-1, Aminosäuren 1860 bis 2217) als auch die gut konservierten Flankierungsregionen (**Fig.** 12) enthält, zu 79% identisch. Die cdc10-Region deckt die Aminosäuren 1860 bis 2217 der TAN-1-Sequenz ab. Zusätzlich ist das hN5k-codierte Polypeptid mit dem TAN-1-Polypeptid (44% identisch zu Drosophila-Notch) am Carboxyl-terminalen Ende des Moleküls, das eine PEST(Prolin, Glutaminsäure, Serin, Threonin)-reiche Region (TAN-1, Aminosäuren 2482 bis 2551) (Fig. 12B) enthält, zu 65% identisch. Die Verlängerung von 215 Aminosäuren, die zwischen den zuvor genannten Regionen liegt, ist bei keinem der durch hN3k, hN5k und TAN-1 dargestellten Notch-homologen Klone gut konserviert. Weder das hN5k-Polypeptid noch Drosophila-Notch zeigt nennenswerte Niveaus von Aminosäureidentität gegenüber den anderen Proteinen in dieser Region (z. B. hN5k/TAN-1 = 24% Identität; hN5k/Drosophila-Notch = 11% Identität; TAN-1/Drosophila-Notch = 17% Identität). Im Gegensatz dazu teilen sich Xenopus-Notch (Xotch) (SEQ ID NO: 16), Ratten-Notch (SEQ ID NO: 17) und TAN-1 (SEQ ID NO: 18) weiterhin signifikante Niveaus der Sequenzidentität (z. B. TAN-1/Ratten-Notch = 75% Identität, TAN-1/Xenopus-Notch = 45% Identität, Ratten-Notch/Xenopus-Notch = 50% Identität).

[0233] Die Überprüfung der Sequenz der intrazellulären Domänen der Vertebraten-Notch-Homologen, die in Fig. 12B gezeigt sind, ergab eine unerwartete Feststellung: alle diese Proteine, einschließlich hN5k, enthalten ein vermeintliches CcN-Motiv, das mit einer Kern-gerichteten Funktion in der konservierten Region verbunden ist, die auf die letzte der sechs cdc10-Sequenzwiederholungen folgt (Fig. 12B). Obwohl bei Drosophila-Notch ein solches definiertes Motiv fehlt, ergab eine nähere Überprüfung seiner Sequenz das Vorliegen einer möglichen zweiteiligen Kern-Lokalisierungssequenz (Robbins et al., 1991, Cell 64: 615–623) sowie möglicher CKII- und cdc2-Phosphorylierungsstellen, alle in relativer Nähe zueinander, wodurch möglicherweise ein alternativer Typ von CcN-Motiv definiert ist (Fig. 12B).

[0234] Zur Isolierung von Klonen, die das 5'-Ende von hN (dem menschlichen Notch-Homologen, das zum Teil in hN5k enthalten ist) abdecken, wurde der Klon hN2k als Sonde zur Durchmusterung von 260000 Plaques aus der Phagen-Bibliothek des menschlichen fetalen Gehirns, die im Handel von Stratagene erhältlich ist, auf Kreuzhybridisierungsklone verwendet. Vier Klone wurden identifiziert und unter Anwendung von Standardverfahren isoliert (Maniatis et al., 1982, Molecular Cloning: A Laboratory Manual., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). Vier Klone wurden auch durch Hybridisierung mit der Notch-homologen Sequenz von Adams et al., 1992, Nature 355: 632–655 isoliert, die von der ATCC erhalten wurde.

[0235] Zur Isolierung von Klonen, die das 5'-Ende von TAN-1 abdeckten, wurde die Bibliothek des menschlichen fetale Gehirns, die käuflich von Stratagene erhältlich ist, auf Klone durchmustert, die die Sequenz bis zum 5'-Ende verlängern würden. 880000 Plaques wurden durchmustert, und vier Klone, die mit den hN3k-Sequenzen kreuzhybridisieren, wurden identifiziert. Die Sequenzierung bestätigte die relative Position dieser Sequenzen innerhalb des Notch-Proteins, das von TAN-1 codiert wird.

[0236] Die 5'-Sequenz unseres isolierten TAN-1-Homologen wurde durch die Nukleotidnummer 972 (wobei Nukleotidnummer 1 das A im ATG-Startcodon ist) bestimmt und mit der Sequenz (wie veröffentlicht von Ellisen

et al., 1991, Cell 66: 649–661) verglichen. Bei Nukleotid 559 besitzt unser TAN-1-Homolog ein G, wohin gegen Ellisen et al. ein A offenbaren, wobei die Änderung zu einer unterschiedlich codierten Aminosäure führt. Somit unterscheidet sich unser TAN-1-codiertes Protein innerhalb den ersten 324 Aminosäuren von demjenigen, das von Ellisen et al. gelehrt wird, da unser Protein ein Gly in Position 187 aufweist, wohin gegen Ellisen et al. an dieser Position ein Arg offenbaren (wie in **Fig.** 13 dargestellt).

[0237] Die Vollängen-Aminosäuresequenzen sowohl der hN- (SEQ ID NO: 19) als auch TAN-1-codierten (SEQ ID NO: 20) Proteine, sowie der Xenopus- und Drosophila-Notch-Proteine sind in **Fig.** 13 gezeigt. Die Vollängen-DNA-Codierungssequenz (mit der Ausnahme derjenigen, die den Initiator Met codiert) (enthalten in der SEQ ID NO: 21), und die codierte Aminosäuresequenz (mit der Ausnahme, daß der Initiator Met nicht gezeigt ist) (enthalten in der SEQ ID NO: 19) von hN sind in **Fig.** 17 gezeigt.

9.2 EXPRESSION VON MENSCHLICHEM NOTCH

[0238] Expressionskonstrukte wurden unter Verwendung der menschlichen Notch-cDNA-Klone hergestellt, die in Abschnitt 9.1 vorstehend diskutiert wurden. In den Fällen von hN3k und hN2k wurde der gesamte Klon als EcoRI-Restriktionsfragment aus seinem Vektor herausgeschnitten und in die EcoRI-Restriktionsstelle von jedem der drei pGEX-Vektoren (Glutathion-S-Transferaseexpressionsvektoren; Smith und Johnson, 1988, Gene 7, 31–40) subkloniert. Dies gestattet die Expression des Notch-Proteinprodukts aus dem Subklon in dem korrekten Leserahmen. Im Falle von hN5k enthält der Klon zwei interne Eco-RI-Restriktionsstellen, die 2,6, 1,5 und 0,6 kb-Fragmente produzieren. Sowohl das 2,6- als auch das 1,5 kb-Fragment wurden ebenfalls in jeden der pGEX-Vektoren subkloniert.

[0239] Das pGEX-Vektorsystem wurde zum Erhalt der Expression von menschlichen Notch-Fusions(chimären)-Proteinen aus den nachstehend beschriebenen Konstrukten verwendet. Die klonierte Notch-DNA wurde in jedem Fall in Phase in den entsprechenden pGEX-Vektor inseriert. Anschließend wurde jedes Konstrukt durch Elektroporation in Bakterien (E. coli) eingeschleust und als die Notch-Protein-Sequenzen-enthaltendes Fusionsprotein, das an den Carboxyl-Terminus des Glutathion-S-Transferase-Proteins fusioniert war, exprimiert. Die Expression der Fusionsproteine wurde durch Analyse von bakteriellen Proteinextrakten, durch Polyacrylamidgelelektrophorese unter Vergleich von Proteinextrakten, die aus Bakterien erhalten wurden, die die pGEX-Plasmide mit und ohne die inserierte Notch-DNA enthielten, bestätigt. Die Fusionsproteine waren in wässriger Lösung löslich und wurden durch Affinitätschromatographie unter Verwendung von Glutathion-beschichteter Agarose (da der Carboxylterminus von Glutathion-S-Transferase an Glutathion bindet) aus Bakterienlysaten gereinigt. Die exprimierten Fusionsproteine wurden über einen Antikörper an Drosophila-Notch gebunden, wie durch Western-Blots geprüft.

[0240] Die zur Herstellung von menschlichen Notch-Glutathion-S-Transferasefusionsproteinen verwendeter Konstrukte waren wie folgt:
hNFP Nr. 2 – PCR wurde zum Erhalt eines Fragments, beginnend unmittelbar vor den cdc10-Sequenzwiederholungen bei Nukleotid 192 der hN5k-Insertion bis unmittelbar vor die PEST-reiche Region bei Nukleotid 1694, verwendet. Die DNA wurde anschließend mit BamBI und SmaI verdaut, und das resultierende Fragment wurde in pGEX-3 ligiert. Nach der Expression wurde das Fusionsprotein durch Binden an Glutathionagarose gereinigt. Das gereinigte Polypeptid wurde über ein 4- bis 15%iges Gradienten-Polyacrylamidgel quantifiziert. Das resultierende Fusionsprotein besaß ein ungefähres Molekulargewicht von 83 kD.
hN3FP-Nr. 1 – Die gesamte hN3k-DNA-Insertion (Nukleotid 1 bis 3235) wurde aus dem Bluescript(SK)-Vektor durch Verdau mit EcoRI herausgeschnitten. Die DNA wurde in pGEX-3 ligiert.
hNFP-Nr. 2 – Ein 3'-Segment von hN3k-DNA (Nukleotid 1847 bis 3235) plus etwas Polylinker wurde aus dem Bluescript (SK)-Vektor durch Verdau mit XmaI herausgeschnitten. Das Fragment wurde in pGEX-1 ligiert.

[0241] Nach der Reinigung werden diese Fusionsproteine zur Herstellung von entweder polyklonalen und/oder monoklonalen Antikörpern gegen menschliches Notch verwendet.

10. NOTCH-EXPRESSION IN GESUNDEN UND BÖSARTIGEN ZELLEN

[0242] Verschiedene menschliche Patientengewebe-Proben und Zell-Linien, die sowohl gesundes als auch eine breite Vielzahl bösartiger Zellen darstellten, werden zum Nachweis und/oder zur Quantifizierung der Expression von Notch getestet. Patientengewebe-Proben werden vom Department für Pathologie der Yale University School of Medicine erhalten.

[0243] Die folgenden Test werden zum Messen der Notch-Expression in Patientengewebe-Proben verwen-

det: (a) Northern-Hybridisierung; (b) Western-Blots; (c) in situ Hybridisierung; und (d) Immunozytochemie. Die Tests werden unter Anwendung von Standardtechniken durchgeführt. Die Northern-Hybridisierung und in situ Hybridisierung werden (i) unter Verwendung einer DNA-Sonde, die auf die Notch-Sequenz des Klons hN3k spezifisch ist; und (ii) unter Verwendung einer DNA-Sonde, die auf die Notch-Sequenz von Klon hN5k spezifisch ist, durchgeführt. Western-Blots und Immunozytochemie werden unter Verwendung eines Antikörpers auf Drosophila-Notch-Protein (der auch menschliche Notch-Proteine erkennt) durchgeführt.

[0244] Northern-Hybridisierung und Western-Blots, wie vorstehend beschrieben, werden ebenfalls zur Analyse zahlreicher menschlicher Zell-Linien verwendet, die verschiedene gesunde oder kanzeröse Gewebe darstellen. Die getesteten Zell-Linien sind in Tabelle 2 aufgeführt.

TABELLE 2

MENSCHLICHE ZELLINIEN

Gewebe/Tumor	Zellinie
Knochenmark	IM-9
	KG-1
Gehirn	A-172
	HS 683
	U-87MG
	TE 671
Brust	BT-20
	Hs 578Bs
	MDA-MB-330
Darm	Caco-2
	SW 48
	T84
	WiDr
Embryo	FHs 173We
Nieren	A-498
	A-704
	Caki-2
Leukämie	ARH-77
	KG-1
Leber	Hep G2
	WRL 68
Lunge	Calu-1
	HLF-a
	SK-Lu-I
Lymphoblasten	CCRF-CEM
	HuT 78
Lymphome	Hs 445
	MS116
	U-937

Melanome	A-375
	G-361
	Hs 294T
	SK-MEL-1
Myelome	IM-9
	RPMI 8226
Neuroblastome	IMR-32
	SK-N-SH
	SK-N-MC
Eierstock	Caov-3
	Caov-4
	PA-1
Plasmazellen	ARH-77
Sarkom	A-204
	A673
	HOS
Haut	Amdur II
	BUD-8
Hoden	Tera-1
	Tera-2
Thymus	Hs67
Uterus	AN3 Ca
	HEC-1-A

[0245] Die Bösartigkeit bösartiger Zellgewebstypen, bei denen sich gezeigt hat, daß sie Notch spezifisch exprimieren, kann wie in Abschnitt 5.1 et seq. beschrieben behandelt werden.

10.1 DIE EXPRESSION VON MENSCHLICHEM NOTCH-PROTEIN IST BEI VERSCHIEDENEN BÖSARTIGKEITEN ERHÖHT

[0246] Wie nachstehend beschrieben, haben wir festgestellt, daß die menschliche Notch-Proteinexpression bei mindestens drei menschlichen Krebsarten, nämlich Zervikal-, Brust- und Darmkrebs, erhöht ist. Die immunozytochemische Anfärbung von Gewebeproben aus Zervikal-, Brust- und Darmkrebs von menschlichen Patienten zeigte eindeutig, daß das bösartige Gewebe hohe Konzentrationen an Notch, auf im Vergleich zu nicht bösartigen Gewebeschnitten erhöhten Niveaus, exprimieren. Dieses breite Spektrum an diversen Neoplasien mit erhöhter Notch-Expression, legt nahe, daß viel mehr kanzeröse Zustände mit Notch-Überregulierung vorkommen.

[0247] Scheiben von menschlichen Tumorproben (aus Brust-, Darm- und Zervikaltumoren) wurden von der Gewebebank der pathologischen Abteilung der Yale Medical School erhalten. Die Anfärbungen erfolgten unter Verwendung monoklonaler Antikörper, die gegen die P1- und P4-Fusionsproteine gezüchtet wurden, die aus Sequenzen von hN bzw. TAN-1 erzeugt worden sind.

[0248] Die P1- und P4-Fusionsproteine wurden durch Insertion der gewünschten menschlichen Notch-Sequenz in den entsprechenden pGEX-Expressionsvektor erhalten (Smith und Johnson, 1988, Gene 7: 31–40; AMRAD Corp., Melbourne, Australia) und wurden nach den Anweisungen des Herstellers (AMRAD Corp.) affinitätschromatographisch gereinigt. Zur Herstellung des P1-Fusionsproteins wurde pGEX-2 mit BamHI geschnitten und an ein Concatamer ligiert, das aus drei Kopien eines 518 bp-BamHI-BgIII-Fragment von hN besteht. Ratten wurden mit dem exprimierten Gen immunisiert, und monoklonale Antikörper wurden durch Standardverfahren produziert. Zur Produktion des P4-Fusionsproteins wurden pGEX-2 mit BamHI geschnitten und mit einem Concatamer ligiert, das aus drei Kopien eines 473 bp-BamHI-BgIII-Fragments von TAN-1 besteht. Mit dem exprimierten Gen wurden Ratten immunisiert, und monoklonale Antiköper wurden durch Standardverfahren produziert.

[0249] In sämtlichen getesteten Tumoren waren die Notch-Proteine, die sowohl von dem menschlichen Notch-Homologen TAN-1 als auch von hN codiert wurden, im bösartigen Teil des Gewebes im Vergleich zum gesunden Teil in erhöhten Konzentrationen vorhanden. Repräsentative Anfärbungen sind in den bereitgestellten Bildern gezeigt (Fig. 14–16).

[0250] Das Anfärbeverfahren war wie folgt: Die Gewebe wurden in Paraformaldehyd fixiert, in Paraffin eingebettet, in 5 µm dicke Schnitte geschnitten und auf Glasträger gelegt. Anschließend wurden die folgenden Schritte durchgeführt:

1. Paraffinentfernung durch viermaligen Xylolaustausch jeweils 4 min.
2. Entfernung von Xylol durch dreimaligen Austausch mit absolutem Ethanol, jeweils 4 min.
3. Stufenweise Rehydratation der Gewebe durch Eintauchen der Scheiben in 95%, 90%, 80%, 60% und 30% Ethanol, jeweils 4 min lang. Am Ende wurden die Scheiben 5 min in destilliertem Wasser gespült.
4. Stoppen von endogener Peroxidase durch Inkubation für 30 min in 0,3% Wasserstoffperoxid in Methanol.
5. Waschen in PBS (10 mM Natriumphosphat, pH 7,5, 0,9% NaCl) für 20 min.
6. Inkubation für 1 h in Blockierungslösung. (Blockierungslösung: PBS, enthaltend 4% normales Kaninchenserum und 0,1% Triton X-100).
7. Inkubation über Nacht bei 4°C mit primärem Antikörper, verdünnt in Blockierungslösung. Endkonzentration von primärem Antikörper 20–50 µg/ml.
8. Waschen für 20 min mit PBS + 0,1% Triton X-100 (dreimal ausgetauscht).
9. Inkubation für 30 min mit biotinyliertem Kaninchen-Anti-Ratten-Antikörper: 50 µl biotinylierter Antiköper (VECTOR) in 10 ml Blockierungslösung.
10. Waschen für 20 min in PBS + 0,1% Triton X-100 (dreimal ausgetauscht).
11. Inkubation mit ABC-Reagenz (VECTOR) für 30 min (das Reagenz wird in PBS + 0,1% Triton-X-100 hergestellt).
12. Waschen für 20 min in PBS + 0,1% Triton-X-100. Anschließend Inkubation für 2 min in PBS + 0,5% Triton-X-100.
13. Inkubation für 2–5 min in Peroxidasesubstratlösung. Peroxidasesubstratlösung: Gleiche Volumina von 0,02% Wasserstoffperoxid in destilliertem Wasser und 0,1% Diaminobenzidintetrahydrochlorid (DAB) in 0,1 M Trispuffer, pH 7,5 werden unmittelbar vor der Inkubation mit den Geweben gemischt. Triton X-100 wird der Endlösung in einer Konzentration von 0,5% zugesetzt.
14. Waschen für 15 min in Leitungswasser.
15. Gegenfärben für 10 min mit Mayer's Hämatoxylin.
16. Waschen für 15 min in Leitungswasser.
17. Dehydratation durch Austausch mit 30%, 60%, 80% , 90%, 95% und absoluten Ethanol (jeweils 4 min).
18. Eintauchen in Xylol (2maliger Austausch, jeweils 4 min).
19. Fixieren, Lichtmikroskopie.

11. HINTERLEGUNG VON MIKROORGANISMEN

[0251] Die folgenden rekombinanten Bakterien, die jeweils ein Plasmid tragen, das einen Teil des menschlichen Notch codiert, wurden am 2. Mai 1991 bei der American Type Culture Collection, 1201 Parklawn Drive, Rockville, Maryland 20852, unter den Auflagen des Budapester Vertrags über die Internationale Anerkennung der Hinterlegung von Mikroorganismen zu Patentierungszwecken hinterlegt.

Bakterien	mit:	Plasmid	ATCC-Zugangs-Nr.
E. coli XL1-Blue		hN4k	68610
E. coli XL1-Blue		hN3k	68609
E. coli XL1-Blue		hNk5	68611

[0252] Es sind hier verschiedene Veröffentlichungen zitiert, die der Fachmann wegen weiterer Informationen über das besprochene Problem heranziehen könnte.

SEQUENZPROTOKOLL

(1) ALLGEMEINE ANGABEN:

(i) ANMELDER : Artavanis-Tsakonas, S. et al.

(ii) BEZEICHNUNG DER ERFINDUNG: THERAPEUTISCHE UND DIAGNOSTISCHE VERFAHREN UND ZUSAMMENSETZUNGEN AUF DER BASIS VON NOTCH-PROTEINEN UND NUKLEINSÄUREN

(iii) ANZAHL DER SEQUENZEN: 21

(iv) KORRESPONDENZADRESSE:
(A) ADRESSE: Pennie & Edmonds
(B) STRASSE: 1155 AVENUE OF THE AMERICAS
(C) STADT: New York
(D) STAAT: New York
(E) LAND: U.S.A.
(F) ZIP: 10036

(v) COMPUTERLESBARE FORM:
(A) DATENTRÄGER: Diskette
(B) COMPUTER: IBM PC compatible
(C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) VORLIEGENDE PATENTANMELDUNG:
(A) ANMELDENUMMER: neu zu vergeben
(B) ANMELDEDATUM: gleiches Datum
(C) KLASSIFIKATION:

(viii) PATENTANWALT/INFORMATION:
(A) NAME: Misrock, S. Leslie
(B) REGISTRIERNUMMER: 18,872
(C) REFERENZ/AKTENNUMMER: 7326-019

(ix) TELEKOMMUNIKATIONSINFORMATION:
(A) TELEPHON: 212 790-9090
(B) TELEFAX: 212 8698864/9741
(C) TELEX: 66141 PENNIE

(2) ANGABEN ZU SEQ ID NO:1:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 2892 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Doppelstrang
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: cDNA

(ix) MERKMAL:
(A) NAME/SCHLÜSSEL: CDS
(B) LAGE: 142..2640

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:1:

```
GAATTCGGAG GAATTATTCA AAACATAAAC ACAATAAACA ATTTGAGTAG TTGCCGCACA      60

CACACACACA CACAGCCCGT GGATTATTAC ACTAAAAGCG ACACTCAATC CAAAAAATCA     120

GCAACAAAAA CATCAATAAA C ATG CAT TGG ATT AAA TGT TTA TTA ACA GCA        171
                        Met His Trp Ile Lys Cys Leu Leu Thr Ala
                          1               5                  10
```

```
TTC ATT TGC TTC ACA GTC ATC GTG CAG GTT CAC AGT TCC GGC AGC TTT      219
Phe Ile Cys Phe Thr Val Ile Val Gln Val His Ser Ser Gly Ser Phe
                15                  20                  25

GAG TTG CGC CTG AAG TAC TTC AGC AAC GAT CAC GGG CGG GAC AAC GAG      267
Glu Leu Arg Leu Lys Tyr Phe Ser Asn Asp His Gly Arg Asp Asn Glu
            30                  35                  40

GGT CGC TGC TGC AGC GGG GAG TCG GAC GGA GCG ACG GGC AAG TGC CTG      315
Gly Arg Cys Cys Ser Gly Glu Ser Asp Gly Ala Thr Gly Lys Cys Leu
        45                  50                  55

GGC AGC TGC AAG ACG CGG TTT CGC GTC TGC CTA AAG CAC TAC CAG GCC      363
Gly Ser Cys Lys Thr Arg Phe Arg Val Cys Leu Lys His Tyr Gln Ala
    60                  65                  70

ACC ATC GAC ACC ACC TCC CAG TGC ACC TAC GGG GAC GTG ATC ACG CCC      411
Thr Ile Asp Thr Thr Ser Gln Cys Thr Tyr Gly Asp Val Ile Thr Pro
75                  80                  85                  90

ATT CTC GGC GAG AAC TCG GTC AAT CTG ACC GAC GCC CAG CGC TTC CAG      459
Ile Leu Gly Glu Asn Ser Val Asn Leu Thr Asp Ala Gln Arg Phe Gln
                95                  100                 105

AAC AAG GGC TTC ACG AAT CCC ATC CAG TTC CCC TTC TCG TTC TCA TGG      507
Asn Lys Gly Phe Thr Asn Pro Ile Gln Phe Pro Phe Ser Phe Ser Trp
            110                 115                 120

CCG GGT ACC TTC TCG CTG ATC GTC GAG GCC TGG CAT GAT ACG AAC AAT      555
Pro Gly Thr Phe Ser Leu Ile Val Glu Ala Trp His Asp Thr Asn Asn
        125                 130                 135

AGC GGC AAT GCG CGA ACC AAC AAG CTC CTC ATC CAG CGA CTC TTG GTG      603
Ser Gly Asn Ala Arg Thr Asn Lys Leu Leu Ile Gln Arg Leu Leu Val
    140                 145                 150

CAG CAG GTA CTG GAG GTG TCC TCC GAA TGG AAG ACG AAC AAG TCG GAA      651
Gln Gln Val Leu Glu Val Ser Ser Glu Trp Lys Thr Asn Lys Ser Glu
155                 160                 165                 170

TCG CAG TAC ACG TCG CTG GAG TAC GAT TTC CGT GTC ACC TGC GAT CTC      699
Ser Gln Tyr Thr Ser Leu Glu Tyr Asp Phe Arg Val Thr Cys Asp Leu
                175                 180                 185

AAC TAC TAC GGA TCC GGC TGT GCC AAG TTC TGC CGG CCC CGC GAC GAT      747
Asn Tyr Tyr Gly Ser Gly Cys Ala Lys Phe Cys Arg Pro Arg Asp Asp
            190                 195                 200

TCA TTT GGA CAC TCG ACT TGC TCG GAG ACG GGC GAA ATT ATC TGT TTG      795
Ser Phe Gly His Ser Thr Cys Ser Glu Thr Gly Glu Ile Ile Cys Leu
        205                 210                 215

ACC GGA TGG CAG GGC GAT TAC TGT CAC ATA CCC AAA TGC GCC AAA GGC      843
Thr Gly Trp Gln Gly Asp Tyr Cys His Ile Pro Lys Cys Ala Lys Gly
    220                 225                 230

TGT GAA CAT GGA CAT TGC GAC AAA CCC AAT CAA TGC GTT TGC CAA CTG      891
Cys Glu His Gly His Cys Asp Lys Pro Asn Gln Cys Val Cys Gln Leu
235                 240                 245                 250

GGC TGG AAG GGA GCC TTG TGC AAC GAG TGC GTT CTG GAA CCG AAC TGC      939
Gly Trp Lys Gly Ala Leu Cys Asn Glu Cys Val Leu Glu Pro Asn Cys
                255                 260                 265

ATC CAT GGC ACC TGC AAC AAA CCC TGG ACT TGC ATC TGC AAC GAG GGT      987
Ile His Gly Thr Cys Asn Lys Pro Trp Thr Cys Ile Cys Asn Glu Gly
            270                 275                 280
```

```
TGG GGA GGC TTG TAC TGC AAC CAG GAT CTG AAC TAC TGC ACC AAC CAC      1035
Trp Gly Gly Leu Tyr Cys Asn Gln Asp Leu Asn Tyr Cys Thr Asn His
        285                 290                 295

AGA CCC TGC AAG AAT GGC GGA ACC TGC TTC AAC ACC GGC GAG GGA TTG      1083
Arg Pro Cys Lys Asn Gly Gly Thr Cys Phe Asn Thr Gly Glu Gly Leu
    300                 305                 310

TAC ACA TGC AAA TGC GCT CCA GGA TAC AGT GGT GAT GAT TGC GAA AAT      1131
Tyr Thr Cys Lys Cys Ala Pro Gly Tyr Ser Gly Asp Asp Cys Glu Asn
315                 320                 325                 330

GAG ATC TAC TCC TGC GAT GCC GAT GTC AAT CCC TGC CAG AAT GGT GGT      1179
Glu Ile Tyr Ser Cys Asp Ala Asp Val Asn Pro Cys Gln Asn Gly Gly
                335                 340                 345

ACC TGC ATC GAT GAG CCG CAC ACA AAA ACC GGC TAC AAG TGT CAT TGC      1227
Thr Cys Ile Asp Glu Pro His Thr Lys Thr Gly Tyr Lys Cys His Cys
            350                 355                 360

GCC AAC GGC TGG AGC GGA AAG ATG TGC GAG GAG AAA GTG CTC ACG TGT      1275
Ala Asn Gly Trp Ser Gly Lys Met Cys Glu Glu Lys Val Leu Thr Cys
        365                 370                 375

TCG GAC AAA CCC TGT CAT CAG GGA ATC TGC CGC AAC GTT CGT CCT GGC      1323
Ser Asp Lys Pro Cys His Gln Gly Ile Cys Arg Asn Val Arg Pro Gly
    380                 385                 390

TTG GGA AGC AAG GGT CAG GGC TAC CAG TGC GAA TGT CCC ATT GGC TAC      1371
Leu Gly Ser Lys Gly Gln Gly Tyr Gln Cys Glu Cys Pro Ile Gly Tyr
395                 400                 405                 410

AGC GGA CCC AAC TGC GAT CTC CAG CTG GAC AAC TGC AGT CCG AAT CCA      1419
Ser Gly Pro Asn Cys Asp Leu Gln Leu Asp Asn Cys Ser Pro Asn Pro
                415                 420                 425

TGC ATA AAC GGT GGA AGC TGT CAG CCG AGC GGA AAG TGT ATT TGC CCA      1467
Cys Ile Asn Gly Gly Ser Cys Gln Pro Ser Gly Lys Cys Ile Cys Pro
            430                 435                 440

GCG GGA TTT TCG GGA ACG AGA TGC GAG ACC AAC ATT GAC GAT TGT CTT      1515
Ala Gly Phe Ser Gly Thr Arg Cys Glu Thr Asn Ile Asp Asp Cys Leu
        445                 450                 455

GGC CAC CAG TGC GAG AAC GGA GGC ACC TGC ATA GAT ATG GTC AAC CAA      1563
Gly His Gln Cys Glu Asn Gly Gly Thr Cys Ile Asp Met Val Asn Gln
    460                 465                 470

TAT CGC TGC CAA TGC GTT CCC GGT TTC CAT GGC ACC CAC TGT AGT AGC      1611
Tyr Arg Cys Gln Cys Val Pro Gly Phe His Gly Thr His Cys Ser Ser
475                 480                 485                 490

AAA GTT GAC TTG TGC CTC ATC AGA CCG TGT GCC AAT GGA GGA ACC TGC      1659
Lys Val Asp Leu Cys Leu Ile Arg Pro Cys Ala Asn Gly Gly Thr Cys
                495                 500                 505

TTG AAT CTC AAC AAC GAT TAC CAG TGC ACC TGT CGT GCG GGA TTT ACT      1707
Leu Asn Leu Asn Asn Asp Tyr Gln Cys Thr Cys Arg Ala Gly Phe Thr
            510                 515                 520

GGC AAG GAT TGC TCT GTG GAC ATC GAT GAG TGC AGC AGT GGA CCC TGT      1755
Gly Lys Asp Cys Ser Val Asp Ile Asp Glu Cys Ser Ser Gly Pro Cys
        525                 530                 535

CAT AAC GGC GGC ACT TGC ATG AAC CGC GTC AAT TCG TTC GAA TGC GTG      1803
His Asn Gly Gly Thr Cys Met Asn Arg Val Asn Ser Phe Glu Cys Val
    540                 545                 550
```

```
TGT GCC AAT GGT TTC AGG GGC AAG CAG TGC GAT GAG GAG TCC TAC GAT      1851
Cys Ala Asn Gly Phe Arg Gly Lys Gln Cys Asp Glu Glu Ser Tyr Asp
555                 560                 565                 570

TCG GTG ACC TTC GAT GCC CAC CAA TAT GGA GCG ACC ACA CAA GCG AGA      1899
Ser Val Thr Phe Asp Ala His Gln Tyr Gly Ala Thr Thr Gln Ala Arg
                575                 580                 585

GCC GAT GGT TTG ACC AAT GCC CAG GTA GTC CTA ATT GCT GTT TTC TCC      1947
Ala Asp Gly Leu Thr Asn Ala Gln Val Val Leu Ile Ala Val Phe Ser
            590                 595                 600

GTT GCG ATG CCT TTG GTG GCG GTT ATT GCG GCG TGC GTG GTC TTC TGC      1995
Val Ala Met Pro Leu Val Ala Val Ile Ala Ala Cys Val Val Phe Cys
        605                 610                 615

ATG AAG CGC AAG CGT AAG CGT GCT CAG GAA AAG GAC GAC GCG GAG GCC      2043
Met Lys Arg Lys Arg Lys Arg Ala Gln Glu Lys Asp Asp Ala Glu Ala
    620                 625                 630

AGG AAG CAG AAC GAA CAG AAT GCG GTG GCC ACA ATG CAT CAC AAT GGC      2091
Arg Lys Gln Asn Glu Gln Asn Ala Val Ala Thr Met His His Asn Gly
635                 640                 645                 650

AGT GGG GTG GGT GTA GCT TTG GCT TCA GCC TCT CTG GGC GGC AAA ACT      2139
Ser Gly Val Gly Val Ala Leu Ala Ser Ala Ser Leu Gly Gly Lys Thr
                655                 660                 665

GGC AGC AAC AGC GGT CTC ACC TTC GAT GGC GGC AAC CCG AAT ATC ATC      2187
Gly Ser Asn Ser Gly Leu Thr Phe Asp Gly Gly Asn Pro Asn Ile Ile
            670                 675                 680

AAA AAC ACC TGG GAC AAG TCG GTC AAC AAC ATT TGT GCC TCA GCA GCA      2235
Lys Asn Thr Trp Asp Lys Ser Val Asn Asn Ile Cys Ala Ser Ala Ala
        685                 690                 695

GCA GCG GCG GCG GCG GCA GCA GCG GCG GAC GAG TGT CTC ATG TAC GGC      2283
Ala Ala Ala Ala Ala Ala Ala Ala Ala Asp Glu Cys Leu Met Tyr Gly
    700                 705                 710

GGA TAT GTG GCC TCG GTG GCG GAT AAC AAC AAT GCC AAC TCA GAC TTT      2331
Gly Tyr Val Ala Ser Val Ala Asp Asn Asn Asn Ala Asn Ser Asp Phe
715                 720                 725                 730

TGT GTG GCT CCG CTA CAA AGA GCC AAG TCG CAA AAG CAA CTC AAC ACC      2379
Cys Val Ala Pro Leu Gln Arg Ala Lys Ser Gln Lys Gln Leu Asn Thr
                735                 740                 745

GAT CCC ACG CTC ATG CAC CGC GGT TCG CCG GCA GGC AGC TCA GCC AAG      2427
Asp Pro Thr Leu Met His Arg Gly Ser Pro Ala Gly Ser Ser Ala Lys
            750                 755                 760

GGA GCG TCT GGC GGA GGA CCG GGA GCG GCG GAG GGC AAG AGG ATC TCT      2475
Gly Ala Ser Gly Gly Gly Pro Gly Ala Ala Glu Gly Lys Arg Ile Ser
        765                 770                 775

GTT TTA GGC GAG GGT TCC TAC TGT AGC CAG CGT TGG CCC TCG TTG GCG      2523
Val Leu Gly Glu Gly Ser Tyr Cys Ser Gln Arg Trp Pro Ser Leu Ala
    780                 785                 790

GCG GCG GGA GTG GCC GGA GCC TGT TCA TCC CAG CTA ATG GCT GCA GCT      2571
Ala Ala Gly Val Ala Gly Ala Cys Ser Ser Gln Leu Met Ala Ala Ala
795                 800                 805                 810

TCG GCA GCG GGC AGC GGA GCG GGG ACG GCG CAA CAG CAG CGA TCC GTG      2619
Ser Ala Ala Gly Ser Gly Ala Gly Thr Ala Gln Gln Gln Arg Ser Val
                815                 820                 825
```

```
GTC TGC GGC ACT CCG CAT ATG TAACTCCAAA AATCCGGAAG GGCTCCTGGT          2670
Val Cys Gly Thr Pro His Met
            830

AAATCCGGAG AAATCCGCAT GGAGGAGCTG ACAGCACATA CACAAAGAAA AGACTGGGTT    2730

GGGTTCAAAA TGTGAGAGAG ACGCCAAAAT GTTGTTGTTG ATTGAAGCAG TTTAGTCGTC    2790

ACGAAAAATG AAAAATCTGT AACAGGCATA ACTCGTAAAC TCCCTAAAAA ATTTGTATAG    2850

TAATTAGCAA AGCTGTGACC CAGCCGTTTC GATCCCGAAT TC                       2892
```

(2) ANGABEN ZU SEQ ID NO:2:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 833 Aminosäuren
(B) ART: Aminosäure
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:2:

```
Met His Trp Ile Lys Cys Leu Leu Thr Ala Phe Ile Cys Phe Thr Val
  1               5                  10                  15

Ile Val Gln Val His Ser Ser Gly Ser Phe Glu Leu Arg Leu Lys Tyr
            20                  25                  30

Phe Ser Asn Asp His Gly Arg Asp Asn Glu Gly Arg Cys Cys Ser Gly
        35                  40                  45

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

Phe Arg Val Cys Leu Lys His Tyr Gln Ala Thr Ile Asp Thr Thr Ser
65                  70                  75                  80

Gln Cys Thr Tyr Gly Asp Val Ile Thr Pro Ile Leu Gly Glu Asn Ser
                85                  90                  95

Val Asn Leu Thr Asp Ala Gln Arg Phe Gln Asn Lys Gly Phe Thr Asn
            100                 105                 110

Pro Ile Gln Phe Pro Phe Ser Phe Ser Trp Pro Gly Thr Phe Ser Leu
        115                 120                 125

Ile Val Glu Ala Trp His Asp Thr Asn Asn Ser Gly Asn Ala Arg Thr
    130                 135                 140

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

Ser Ser Glu Trp Lys Thr Asn Lys Ser Glu Ser Gln Tyr Thr Ser Leu
                165                 170                 175

Glu Tyr Asp Phe Arg Val Thr Cys Asp Leu Asn Tyr Tyr Gly Ser Gly
            180                 185                 190

Cys Ala Lys Phe Cys Arg Pro Arg Asp Asp Ser Phe Gly His Ser Thr
        195                 200                 205

Cys Ser Glu Thr Gly Glu Ile Ile Cys Leu Thr Gly Trp Gln Gly Asp
    210                 215                 220

Tyr Cys His Ile Pro Lys Cys Ala Lys Gly Cys Glu His Gly His Cys
```

225 230 235 240

Asp Lys Pro Asn Gln Cys Val Cys Gln Leu Gly Trp Lys Gly Ala Leu
245 250 255

Cys Asn Glu Cys Val Leu Glu Pro Asn Cys Ile His Gly Thr Cys Asn
260 265 270

Lys Pro Trp Thr Cys Ile Cys Asn Glu Gly Trp Gly Gly Leu Tyr Cys
275 280 285

Asn Gln Asp Leu Asn Tyr Cys Thr Asn His Arg Pro Cys Lys Asn Gly
290 295 300

Gly Thr Cys Phe Asn Thr Gly Glu Gly Leu Tyr Thr Cys Lys Cys Ala
305 310 315 320

Pro Gly Tyr Ser Gly Asp Asp Cys Glu Asn Glu Ile Tyr Ser Cys Asp
325 330 335

Ala Asp Val Asn Pro Cys Gln Asn Gly Gly Thr Cys Ile Asp Glu Pro
340 345 350

His Thr Lys Thr Gly Tyr Lys Cys His Cys Ala Asn Gly Trp Ser Gly
355 360 365

Lys Met Cys Glu Glu Lys Val Leu Thr Cys Ser Asp Lys Pro Cys His
370 375 380

Gln Gly Ile Cys Arg Asn Val Arg Pro Gly Leu Gly Ser Lys Gly Gln
385 390 395 400

Gly Tyr Gln Cys Glu Cys Pro Ile Gly Tyr Ser Gly Pro Asn Cys Asp
405 410 415

Leu Gln Leu Asp Asn Cys Ser Pro Asn Pro Cys Ile Asn Gly Gly Ser
420 425 430

Cys Gln Pro Ser Gly Lys Cys Ile Cys Pro Ala Gly Phe Ser Gly Thr
435 440 445

Arg Cys Glu Thr Asn Ile Asp Asp Cys Leu Gly His Gln Cys Glu Asn
450 455 460

Gly Gly Thr Cys Ile Asp Met Val Asn Gln Tyr Arg Cys Gln Cys Val
465 470 475 480

Pro Gly Phe His Gly Thr His Cys Ser Ser Lys Val Asp Leu Cys Leu
485 490 495

Ile Arg Pro Cys Ala Asn Gly Gly Thr Cys Leu Asn Leu Asn Asn Asp
500 505 510

Tyr Gln Cys Thr Cys Arg Ala Gly Phe Thr Gly Lys Asp Cys Ser Val
515 520 525

Asp Ile Asp Glu Cys Ser Ser Gly Pro Cys His Asn Gly Gly Thr Cys
530 535 540

Met Asn Arg Val Asn Ser Phe Glu Cys Val Cys Ala Asn Gly Phe Arg
545 550 555 560

Gly Lys Gln Cys Asp Glu Glu Ser Tyr Asp Ser Val Thr Phe Asp Ala
565 570 575

His Gln Tyr Gly Ala Thr Thr Gln Ala Arg Ala Asp Gly Leu Thr Asn
580 585 590

```
Ala Gln Val Val Leu Ile Ala Val Phe Ser Val Ala Met Pro Leu Val
        595                 600                 605

Ala Val Ile Ala Ala Cys Val Val Phe Cys Met Lys Arg Lys Arg Lys
    610                 615                 620

Arg Ala Gln Glu Lys Asp Asp Ala Glu Ala Arg Lys Gln Asn Glu Gln
625                 630                 635                 640

Asn Ala Val Ala Thr Met His His Asn Gly Ser Gly Val Gly Val Ala
                645                 650                 655

Leu Ala Ser Ala Ser Leu Gly Gly Lys Thr Gly Ser Asn Ser Gly Leu
            660                 665                 670

Thr Phe Asp Gly Gly Asn Pro Asn Ile Ile Lys Asn Thr Trp Asp Lys
        675                 680                 685

Ser Val Asn Asn Ile Cys Ala Ser Ala Ala Ala Ala Ala Ala Ala Ala
    690                 695                 700

Ala Ala Ala Asp Glu Cys Leu Met Tyr Gly Gly Tyr Val Ala Ser Val
705                 710                 715                 720

Ala Asp Asn Asn Asn Ala Asn Ser Asp Phe Cys Val Ala Pro Leu Gln
                725                 730                 735

Arg Ala Lys Ser Gln Lys Gln Leu Asn Thr Asp Pro Thr Leu Met His
            740                 745                 750

Arg Gly Ser Pro Ala Gly Ser Ser Ala Lys Gly Ala Ser Gly Gly Gly
        755                 760                 765

Pro Gly Ala Ala Glu Gly Lys Arg Ile Ser Val Leu Gly Glu Gly Ser
    770                 775                 780

Tyr Cys Ser Gln Arg Trp Pro Ser Leu Ala Ala Ala Gly Val Ala Gly
785                 790                 795                 800

Ala Cys Ser Ser Gln Leu Met Ala Ala Ala Ser Ala Ala Gly Ser Gly
                805                 810                 815

Ala Gly Thr Ala Gln Gln Gln Arg Ser Val Val Cys Gly Thr Pro His
            820                 825                 830

Met
```

(2) ANGABEN ZU SEQ ID NO:3:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 1320 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORMN: Doppelstrang
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: cDNA

(ix) MERKMAL:
(A) NAME/SCHLÜSSEL. cds
(B) LAGE: 442..1320

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:3:

```
CCGAGTCGAG CGCCGTGCTT CGAGCGGTGA TGAGCCCCTT TTCTGTCAAC GCTAAAGATC      60

TACAAAACAT CAGCGCCTAT CAAGTGGAAG TGTCAAGTGT GAACAAAACA AAAACGAGAG     120

AAGCACATAC TAAGGTCCAT ATAAATAATA AATAATAATT GTGTGTGATA ACAACATTAT     180

CCAAACAAAA CCAAACAAAA CGAAGGCAAA GTGGAGAAAA TGATACAGCA TCCAGAGTAC     240

GGCCGTTATT CAGCTATCCA GAGCAAGTGT AGTGTGGCAA AATAGAAACA AACAAAGGCA     300

CCAAAATCTG CATACATGGG CTAATTAAGG CTGCCCAGCG AATTTACATT TGTGTGGTGC     360

CAATCCAGAG TGAATCCGAA ACAAACTCCA TCTAGATCGC CAACCAGCAT CACGCTCGCA     420

AACGCCCCCA GAATGTACAA A ATG TTT AGG AAA CAT TTT CGG CGA AAA CCA        471
                         Met Phe Arg Lys His Phe Arg Arg Lys Pro
                           1               5                  10

GCT ACG TCG TCG TCG TTG GAG TCA ACA ATA GAA TCA GCA GAC AGC CTG        519
Ala Thr Ser Ser Ser Leu Glu Ser Thr Ile Glu Ser Ala Asp Ser Leu
                  15                  20                  25

GGA ATG TCC AAG AAG ACG GCG ACA AAA AGG CAG CGT CCG AGG CAT CGG        567
Gly Met Ser Lys Lys Thr Ala Thr Lys Arg Gln Arg Pro Arg His Arg
              30                  35                  40

GTA CCC AAA ATC GCG ACC CTG CCA TCG ACG ATC CGC GAT TGT CGA TCA        615
Val Pro Lys Ile Ala Thr Leu Pro Ser Thr Ile Arg Asp Cys Arg Ser
          45                  50                  55

TTA AAG TCT GCC TGC AAC TTA ATT GCT TTA ATT TTA ATA CTG TTA GTC        663
Leu Lys Ser Ala Cys Asn Leu Ile Ala Leu Ile Leu Ile Leu Leu Val
      60                  65                  70

CAT AAG ATA TCC GCA GCT GGT AAC TTC GAG CTG GAA ATA TTA GAA ATC        711
His Lys Ile Ser Ala Ala Gly Asn Phe Glu Leu Glu Ile Leu Glu Ile
  75                  80                  85                  90

TCA AAT ACC AAC AGC CAT CTA CTC AAC GGC TAT TGC TGC GGC ATG CCA        759
Ser Asn Thr Asn Ser His Leu Leu Asn Gly Tyr Cys Cys Gly Met Pro
                  95                 100                 105

GCG GAA CTT AGG GCC ACC AAG ACG ATA GGC TGC TCG CCA TGC ACG ACG        807
Ala Glu Leu Arg Ala Thr Lys Thr Ile Gly Cys Ser Pro Cys Thr Thr
             110                 115                 120

GCA TTC CGG CTG TGC CTG AAG GAG TAC CAG ACC ACG GAG CAG GGT GCC        855
Ala Phe Arg Leu Cys Leu Lys Glu Tyr Gln Thr Thr Glu Gln Gly Ala
         125                 130                 135

AGC ATA TCC ACG GGC TGT TCG TTT GGC AAC GCC ACC ACC AAG ATA CTG        903
Ser Ile Ser Thr Gly Cys Ser Phe Gly Asn Ala Thr Thr Lys Ile Leu
     140                 145                 150

GGT GGC TCC AGC TTT GTG CTC AGC GAT CCG GGT GTG GGA GCC ATT GTG        951
Gly Gly Ser Ser Phe Val Leu Ser Asp Pro Gly Val Gly Ala Ile Val
155                 160                 165                 170

CTG CCC TTT ACG TTT CGT TGG ACG AAG TCG TTT ACG CTG ATA CTG CAG        999
Leu Pro Phe Thr Phe Arg Trp Thr Lys Ser Phe Thr Leu Ile Leu Gln
                 175                 180                 185

GCG TTG GAT ATG TAC AAC ACA TCC TAT CCA GAT GCG GAG AGG TTA ATT       1047
Ala Leu Asp Met Tyr Asn Thr Ser Tyr Pro Asp Ala Glu Arg Leu Ile
             190                 195                 200

GAG GAA ACA TCA TAC TCG GGC GTG ATA CTG CCG TCG CCG GAG TGG AAG       1095
```

```
Glu Glu Thr Ser Tyr Ser Gly Val Ile Leu Pro Ser Pro Glu Trp Lys
        205                 210                 215

ACG CTG GAC CAC ATC GGG CGG AAC GCG CGG ATC ACC TAC CGT GTC CGG      1143
Thr Leu Asp His Ile Gly Arg Asn Ala Arg Ile Thr Tyr Arg Val Arg
    220                 225                 230

GTG CAA TGC GCC GTT ACC TAC TAC AAC ACG ACC TGC ACG ACC TTC TGC      1191
Val Gln Cys Ala Val Thr Tyr Tyr Asn Thr Thr Cys Thr Thr Phe Cys
235                 240                 245                 250

CGT CCG CGG GAC GAT CAG TTC GGT CAC TAC GCC TGC GGC TCC GAG GGT      1239
Arg Pro Arg Asp Asp Gln Phe Gly His Tyr Ala Cys Gly Ser Glu Gly
                255                 260                 265

CAG AAG CTC TGC CTG AAT GGC TGG CAG GGC GTC AAC TGC GAG GAG GCC      1287
Gln Lys Leu Cys Leu Asn Gly Trp Gln Gly Val Asn Cys Glu Glu Ala
            270                 275                 280

ATA TGC AAG GCG GGC TGC GAC CCC GTC CAC GGC                          1320
Ile Cys Lys Ala Gly Cys Asp Pro Val His Gly
        285                 290
```

(2) ANGABEN ZU SEQ ID NO:4:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 293 Aminosäuren
(B) ART: Aminosäure
(D) TROPOLOGIE: unbekannt

(ii) MOLEKULARTYP: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:4:

```
Met Phe Arg Lys His Phe Arg Arg Lys Pro Ala Thr Ser Ser Ser Leu
  1               5                  10                  15

Glu Ser Thr Ile Glu Ser Ala Asp Ser Leu Gly Met Ser Lys Lys Thr
            20                  25                  30

Ala Thr Lys Arg Gln Arg Pro Arg His Arg Val Pro Lys Ile Ala Thr
        35                  40                  45

Leu Pro Ser Thr Ile Arg Asp Cys Arg Ser Leu Lys Ser Ala Cys Asn
    50                  55                  60

Leu Ile Ala Leu Ile Leu Ile Leu Leu Val His Lys Ile Ser Ala Ala
65                  70                  75                  80

Gly Asn Phe Glu Leu Glu Ile Leu Glu Ile Ser Asn Thr Asn Ser His
                85                  90                  95

Leu Leu Asn Gly Tyr Cys Cys Gly Met Pro Ala Glu Leu Arg Ala Thr
            100                 105                 110

Lys Thr Ile Gly Cys Ser Pro Cys Thr Thr Ala Phe Arg Leu Cys Leu
        115                 120                 125

Lys Glu Tyr Gln Thr Thr Glu Gln Gly Ala Ser Ile Ser Thr Gly Cys
    130                 135                 140

Ser Phe Gly Asn Ala Thr Thr Lys Ile Leu Gly Gly Ser Ser Phe Val
145                 150                 155                 160

Leu Ser Asp Pro Gly Val Gly Ala Ile Val Leu Pro Phe Thr Phe Arg
                165                 170                 175
```

```
Trp Thr Lys Ser Phe Thr Leu Ile Leu Gln Ala Leu Asp Met Tyr Asn
            180                 185                 190

Thr Ser Tyr Pro Asp Ala Glu Arg Leu Ile Glu Glu Thr Ser Tyr Ser
        195                 200                 205

Gly Val Ile Leu Pro Ser Pro Glu Trp Lys Thr Leu Asp His Ile Gly
    210                 215                 220

Arg Asn Ala Arg Ile Thr Tyr Arg Val Arg Val Gln Cys Ala Val Thr
225                 230                 235                 240

Tyr Tyr Asn Thr Thr Cys Thr Thr Phe Cys Arg Pro Arg Asp Asp Gln
                245                 250                 255

Phe Gly His Tyr Ala Cys Gly Ser Glu Gly Gln Lys Leu Cys Leu Asn
            260                 265                 270

Gly Trp Gln Gly Val Asn Cys Glu Glu Ala Ile Cys Lys Ala Gly Cys
        275                 280                 285

Asp Pro Val His Gly
    290
```

(2) ANGABEN ZU SEQ ID NO:5:

(i) SEQUENZKENNZEICHEN:
- (A) LÄNGE: 267 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: cDNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:5:

```
CGGTGGACTT CCTTCGTGTA TTGGTGGGAG CCCTCGGGAA CGGGGGGTAA CACTGAAAGG      60

TCGAGTACCC ATTTCCGTCA TAACGGGTTG GTCGCCCCCT AGGGGTCGGA GTCAGGTGGA     120

CGGGAGGTCG ACAACGCCCG GGGGACGGGT GGTACATGGT GTAAGGTCTT TACCGGACCG     180

GGCAAACGGG TCACACCGAA AGGGGTGAAC GGTAACTACG GGGTCGTCCT GCCCGTCCAT     240

CGAGTCTGGT AAGAGGGTCG CCTTAAG                                         267
```

(2) ANGABEN ZU SEQ ID NO:6:

(i) SEQUENZKENNZEICHEN:
- (A) LÄNGE: 574 Baenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Doppelstrang
- (D) TOPLOGIE: unbekannt

(ii) MOLEKULARTYP: cDNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:6:

```
GAATTCCTTC CATTATACGT GACTTTTCTG AAACTGTAGC CACCCTAGTG TCTCTAACTC      60

CCTCTGGAGT TTGTCAGCTT TGGTCTTTTC AAAGAGCAGG CTCTCTTCAA GCTCCTTAAT     120
```

GCGGGCATGC TCCAGTTTGG TCTGCGTCTC AAGATCACCT TTGGTAATTG ATTCTTCTTC 180

AACCCGGAAC TGAAGGCTGG CTCTCACCCT CTAGGCAGAG CAGGAATTCC GAGGTGGATG 240

TGTTAGATGT GAATGTCCGT GGCCCAGATG GCTGCACCCC ATTGATGTTG GCTTCTCTCC 300

GAGGAGGCAG CTCAGATTTG AGTGATGAAG ATGAAGATGC AGAGGACTGT TCTGCTAACA 360

TCATCACAGA CTTGGTCTAC CAGGGTGCCA GCCTCCAGNC CAGACAGACC GGACTGGTGA 420

GATGGCCCTG CACCTTGCAG CCCGCTACTC ACGGGCTGAT GCTGCCAAGC GTCTCCTGGA 480

TGCAGGTGCA GATGCCAATG CCCAGGACAA CATGGGCCGC TGTCCACTCC ATGCTGCAGT 540

GGCACGTGAT GCCAAGGTGT ATTCAGATCT GTTA 574

(2) ANGABEN ZU SEQ ID NO:7:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 295 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Doppelstrang
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: cDNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:7:

TCCAGATTCT GATTCGCAAC CGAGTAACTG ATCTAGATGC CAGGATGAAT GATGGTACTA 60

CACCCCTGAT CCTGGCTGCC CGCCTGGCTG TGGAGGGAAT GGTGGCAGAA CTGATCAACT 120

GCCAAGCGGA TGTGAATGCA GTGGATGACC ATGGAAAATC TGCTCTTCAC TGGGCAGCTG 180

CTGTCAATAA TGTGGAGGCA ACTCTTTTGT TGTTGAAAAA TGGGGCCAAC CGAGACATGC 240

AGGACAACAA GGAAGAGACA CCTCTGTTTC TTGCTGCCCG GGAGGAGCTA TAAGC 295

(2) ANGABEN ZU SEQ ID NO:8:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 248 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Doppelstrang
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: cDNA

(xi) ANGABEN ZU SEQ ID NO:8:

GAATTCCATT CAGGAGGAAA GGGTGGGGAG AGAAGCAGGC ACCCACTTTC CCGTGGCTGG 60

ACTCGTTCCC AGGTGGCTCC ACCGGCAGCT GTGACCGCCG CAGGTGGGGG CGGAGTGCCA 120

TTCAGAAAAT TCCAGAAAAG CCCTACCCCA ACTCGGACGG CAACGTCACA CCCGTGGGTA 180

GCAACTGGCA CACAAACAGC CAGCGTGTCT GGGGCACGGG GGGATGGCAC CCCCTGCAGG 240

CAGAGCTG 248

(2) ANGABEN ZU SEQ ID NO:9:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE:323 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Doppelstrang
(D) TOPOLOGIE: cDNA

(ii) MOLEKULARTYP: cDNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:9:

```
TACGTATCTC GAGCACAGAC AGCTGACGTA CACTTTTNNA GTGCGAGGGA CATTCGTCCG      60

ACCAGTACGA ACATTTAGGC TCAGTACGGT AGGTCCATGG CCAAGACTAG GAGACGTAGG     120

GAGCTACAGG TCCCGCTCGC TAAACTCGGA CCACTGAAAC CTCCGGTCGA CAGTCGGTAA     180

GCGAACAAGA GGGCCAGATC TTAGAGAAGG TGTCGCGGCG AGACTCGGGC TCGGGTCAGG     240

CGGCCTTAAG GACGTCGGGC CCNNNAGGTG ATCAAGATCT CGNCNCGGCG GGCGCCACCT     300

CGAGGNCGAA AACAAGGGAA ATC                                              323
```

(2) ANGABEN ZU SEQ ID NO:10:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 3234 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Doppelstrang
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: cDNA

(ix) MERKMAL:
(A) NAME/SCHLÜSSEL: CDS
(B) LAGE: 1..3234

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:10:

```
TGC CAG GAG GAC GCG GGC AAC AAG GTC TGC AGC CTG CAG TGC AAC AAC        48
Cys Gln Glu Asp Ala Gly Asn Lys Val Cys Ser Leu Gln Cys Asn Asn
 1               5                  10                  15

CAC GCG TGC GGC TGG GAC GGC GGT GAC TGC TCC CTC AAC TTC AAT GAC        96
His Ala Cys Gly Trp Asp Gly Gly Asp Cys Ser Leu Asn Phe Asn Asp
            20                  25                  30

CCC TGG AAG AAC TGC ACG CAG TCT CTG CAG TGC TGG AAG TAC TTC AGT       144
Pro Trp Lys Asn Cys Thr Gln Ser Leu Gln Cys Trp Lys Tyr Phe Ser
        35                  40                  45

GAC GGC CAC TGT GAC AGC CAG TGC AAC TCA GCC GGC TGC CTC TTC GAC       192
Asp Gly His Cys Asp Ser Gln Cys Asn Ser Ala Gly Cys Leu Phe Asp
    50                  55                  60

GGC TTT GAC TGC CAG CGT GCG GAA GGC CAG TGC AAC CCC CTG TAC GAC       240
Gly Phe Asp Cys Gln Arg Ala Glu Gly Gln Cys Asn Pro Leu Tyr Asp
65                  70                  75                  80

CAG TAC TGC AAG GAC CAC TTC AGC GAC GGG CAC TGC GAC CAG GGC TGC       288
Gln Tyr Cys Lys Asp His Phe Ser Asp Gly His Cys Asp Gln Gly Cys
                85                  90                  95
```

```
AAC AGC GCG GAG TGC GAG TGG GAC GGG CTG GAC TGT GCG GAG CAT GTA        336
Asn Ser Ala Glu Cys Glu Trp Asp Gly Leu Asp Cys Ala Glu His Val
            100                 105                 110

CCC GAG AGG CTG GCG GCC GGC ACG CTG GTG GTG GTG GTG CTG ATG CCG        384
Pro Glu Arg Leu Ala Ala Gly Thr Leu Val Val Val Val Leu Met Pro
        115                 120                 125

CCG GAG CAG CTG CGC AAC AGC TCC TTC CAC TTC CTG CGG GAG CTC AGC        432
Pro Glu Gln Leu Arg Asn Ser Ser Phe His Phe Leu Arg Glu Leu Ser
    130                 135                 140

CGC GTG CTG CAC ACC AAC GTG GTC TTC AAG CGT GAC GCA CAC GGC CAG        480
Arg Val Leu His Thr Asn Val Val Phe Lys Arg Asp Ala His Gly Gln
145                 150                 155                 160

CAG ATG ATC TTC CCC TAC TAC GGC CGC GAG GAG GAG CTG CGC AAG CAC        528
Gln Met Ile Phe Pro Tyr Tyr Gly Arg Glu Glu Glu Leu Arg Lys His
                165                 170                 175

CCC ATC AAG CGT GCC GCC GAG GGC TGG GCC GCA CCT GAC GCC CTG CTG        576
Pro Ile Lys Arg Ala Ala Glu Gly Trp Ala Ala Pro Asp Ala Leu Leu
            180                 185                 190

GGC CAG GTG AAG GCC TCG CTG CTC CCT GGT GGC AGC GAG GGT GGG CGG        624
Gly Gln Val Lys Ala Ser Leu Leu Pro Gly Gly Ser Glu Gly Gly Arg
        195                 200                 205

CGG CGG AGG GAG CTG GAC CCC ATG GAC GTC CGC GGC TCC ATC GTC TAC        672
Arg Arg Arg Glu Leu Asp Pro Met Asp Val Arg Gly Ser Ile Val Tyr
    210                 215                 220

CTG GAG ATT GAC AAC CGG CAG TGT GTG CAG GCC TCC TCG CAG TGC TTC        720
Leu Glu Ile Asp Asn Arg Gln Cys Val Gln Ala Ser Ser Gln Cys Phe
225                 230                 235                 240

CAG AGT GCC ACC GAC GTG GCC GCA TTC CTG GGA GCG CTC GCC TCG CTG        768
Gln Ser Ala Thr Asp Val Ala Ala Phe Leu Gly Ala Leu Ala Ser Leu
                245                 250                 255

GGC AGC CTC AAC ATC CCC TAC AAG ATC GAG GCC GTG CAG AGT GAG ACC        816
Gly Ser Leu Asn Ile Pro Tyr Lys Ile Glu Ala Val Gln Ser Glu Thr
            260                 265                 270

GTG GAG CCG CCC CCG CCG GCG CAG CTG CAC TTC ATG TAC GTG GCG GCG        864
Val Glu Pro Pro Pro Pro Ala Gln Leu His Phe Met Tyr Val Ala Ala
        275                 280                 285

GCC GCC TTT GTG CTT CTG TTC TTC GTG GGC TGC GGG GTG CTG CTG TCC        912
Ala Ala Phe Val Leu Leu Phe Phe Val Gly Cys Gly Val Leu Leu Ser
    290                 295                 300

CGC AAG CGC CGG CGG CAG CAT GGC CAG CTC TGG TTC CCT GAG GGC TTC        960
Arg Lys Arg Arg Arg Gln His Gly Gln Leu Trp Phe Pro Glu Gly Phe
305                 310                 315                 320

AAA GTG TCT GAG GCC AGC AAG AAG AAG CGG CGG GAG CCC CTC GGC GAG       1008
Lys Val Ser Glu Ala Ser Lys Lys Lys Arg Arg Glu Pro Leu Gly Glu
                325                 330                 335

GAC TCC GTG GGC CTC AAG CCC CTG AAG AAC GCT TCA GAC GGT GCC CTC       1056
Asp Ser Val Gly Leu Lys Pro Leu Lys Asn Ala Ser Asp Gly Ala Leu
            340                 345                 350

ATG GAC GAC AAC CAG AAT GAG TGG GGG GAC GAG GAC CTG GAG ACC AAG       1104
Met Asp Asp Asn Gln Asn Glu Trp Gly Asp Glu Asp Leu Glu Thr Lys
        355                 360                 365
```

```
AAG TTC CGG TTC GAG GAG CCC GTG GTT CTG CCT GAC CTG GAC GAC CAG      1152
Lys Phe Arg Phe Glu Glu Pro Val Val Leu Pro Asp Leu Asp Asp Gln
    370                 375                 380

ACA GAC CAC CGG CAG TGG ACT CAG CAG CAC CTG GAT GCC GCT GAC CTG      1200
Thr Asp His Arg Gln Trp Thr Gln Gln His Leu Asp Ala Ala Asp Leu
385                 390                 395                 400

CGC ATG TCT GCC ATG GCC CCC ACA CCG CCC CAG GGT GAG GTT GAC GCC      1248
Arg Met Ser Ala Met Ala Pro Thr Pro Pro Gln Gly Glu Val Asp Ala
                405                 410                 415

GAC TGC ATG GAC GTC AAT GTC CGC GGG CCT GAT GGC TTC ACC CCG CTC      1296
Asp Cys Met Asp Val Asn Val Arg Gly Pro Asp Gly Phe Thr Pro Leu
            420                 425                 430

ATG ATC GCC TCC TGC AGC GGG GGC GGC CTG GAG ACG GGC AAC AGC GAG      1344
Met Ile Ala Ser Cys Ser Gly Gly Gly Leu Glu Thr Gly Asn Ser Glu
        435                 440                 445

GAA GAG GAG GAC GCG CCG GCC GTC ATC TCC GAC TTC ATC TAC CAG GGC      1392
Glu Glu Glu Asp Ala Pro Ala Val Ile Ser Asp Phe Ile Tyr Gln Gly
    450                 455                 460

GCC AGC CTG CAC AAC CAG ACA GAC CGC ACG GGC GAG ACC GCC TTG CAC      1440
Ala Ser Leu His Asn Gln Thr Asp Arg Thr Gly Glu Thr Ala Leu His
465                 470                 475                 480

CTG GCC GCC CGC TAC TCA CGC TCT GAT GCC GCC AAG CGC CTG CTG GAG      1488
Leu Ala Ala Arg Tyr Ser Arg Ser Asp Ala Ala Lys Arg Leu Leu Glu
                485                 490                 495

GCC AGC GCA GAT GCC AAC ATC CAG GAC AAC ATG GGC CGC ACC CCG CTG      1536
Ala Ser Ala Asp Ala Asn Ile Gln Asp Asn Met Gly Arg Thr Pro Leu
            500                 505                 510

CAT GCG GCT GTG TCT GCC GAC GCA CAA GGT GTC TTC CAG ATC CTG ATC      1584
His Ala Ala Val Ser Ala Asp Ala Gln Gly Val Phe Gln Ile Leu Ile
        515                 520                 525

CGG AAC CGA GCC ACA GAC CTG GAT GCC CGC ATG CAT GAT GGC ACG ACG      1632
Arg Asn Arg Ala Thr Asp Leu Asp Ala Arg Met His Asp Gly Thr Thr
    530                 535                 540

CCA CTG ATC CTG GCT GCC CGC CTG GCC GTG GAG GGC ATG CTG GAG GAC      1680
Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu Gly Met Leu Glu Asp
545                 550                 555                 560

CTC ATC AAC TCA CAC GCC GAC GTC AAC GCC GTA GAT GAC CTG GGC AAG      1728
Leu Ile Asn Ser His Ala Asp Val Asn Ala Val Asp Asp Leu Gly Lys
                565                 570                 575

TCC GCC CTG CAC TGG GCC GCC GCC GTG AAC AAT GTG GAT GCC GCA GTT      1776
Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn Val Asp Ala Ala Val
            580                 585                 590

GTG CTC CTG AAG AAC GGG GCT AAC AAA GAT ATG CAG AAC AAC AGG GAG      1824
Val Leu Leu Lys Asn Gly Ala Asn Lys Asp Met Gln Asn Asn Arg Glu
        595                 600                 605

GAG ACA CCC CTG TTT CTG GCC GCC CGG GAG GGC AGC TAC GAG ACC GCC      1872
Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly Ser Tyr Glu Thr Ala
    610                 615                 620

AAG GTG CTG CTG GAC CAC TTT GCC AAC CGG GAC ATC ACG GAT CAT ATG      1920
Lys Val Leu Leu Asp His Phe Ala Asn Arg Asp Ile Thr Asp His Met
625                 630                 635                 640
```

```
GAC CGC CTG CCG CGC GAC ATC GCA CAG GAG CGC ATG CAT CAC GAC ATC      1968
Asp Arg Leu Pro Arg Asp Ile Ala Gln Glu Arg Met His His Asp Ile
                645                 650                 655

GTG AGG CTG CTG GAC GAG TAC AAC CTG GTG CGC AGC CCG CAG CTG CAC      2016
Val Arg Leu Leu Asp Glu Tyr Asn Leu Val Arg Ser Pro Gln Leu His
            660                 665                 670

GGA GCC CCG CTG GGG GGC ACG CCC ACC CTG TCG CCC CCG CTC TGC TCG      2064
Gly Ala Pro Leu Gly Gly Thr Pro Thr Leu Ser Pro Pro Leu Cys Ser
        675                 680                 685

CCC AAC GGC TAC CTG GGC AGC CTC AAG CCC GGC GTG CAG GGC AAG AAG      2112
Pro Asn Gly Tyr Leu Gly Ser Leu Lys Pro Gly Val Gln Gly Lys Lys
    690                 695                 700

GTC CGC AAG CCC AGC AGC AAA GGC CTG GCC TGT GGA AGC AAG GAG GCC      2160
Val Arg Lys Pro Ser Ser Lys Gly Leu Ala Cys Gly Ser Lys Glu Ala
705                 710                 715                 720

AAG GAC CTC AAG GCA CGG AGG AAG AAG TCC CAG GAT GGC AAG GGC TGC      2208
Lys Asp Leu Lys Ala Arg Arg Lys Lys Ser Gln Asp Gly Lys Gly Cys
                725                 730                 735

CTG CTG GAC AGC TCC GGC ATG CTC TCG CCC GTG GAC TCC CTG GAG TCA      2256
Leu Leu Asp Ser Ser Gly Met Leu Ser Pro Val Asp Ser Leu Glu Ser
            740                 745                 750

CCC CAT GGC TAC CTG TCA GAC GTG GCC TCG CCG CCA CTG CTG CCC TCC      2304
Pro His Gly Tyr Leu Ser Asp Val Ala Ser Pro Pro Leu Leu Pro Ser
        755                 760                 765

CCG TTC CAG CAG TCT CCG TCC GTG CCC CTC AAC CAC CTG CCT GGG ATG      2352
Pro Phe Gln Gln Ser Pro Ser Val Pro Leu Asn His Leu Pro Gly Met
    770                 775                 780

CCC GAC ACC CAC CTG GGC ATC GGG CAC CTG AAC GTG GCG GCC AAG CCC      2400
Pro Asp Thr His Leu Gly Ile Gly His Leu Asn Val Ala Ala Lys Pro
785                 790                 795                 800

GAG ATG GCG GCG CTG GGT GGG GGC GGC CGG CTG GCC TTT GAG ACT GGC      2448
Glu Met Ala Ala Leu Gly Gly Gly Gly Arg Leu Ala Phe Glu Thr Gly
                805                 810                 815

CCA CCT CGT CTC TCC CAC CTG CCT GTG GCC TCT GGC ACC AGC ACC GTC      2496
Pro Pro Arg Leu Ser His Leu Pro Val Ala Ser Gly Thr Ser Thr Val
            820                 825                 830

CTG GGC TCC AGC AGC GGA GGG GCC CTG AAT TTC ACT GTG GGC GGG TCC      2544
Leu Gly Ser Ser Ser Gly Gly Ala Leu Asn Phe Thr Val Gly Gly Ser
        835                 840                 845

ACC AGT TTG AAT GGT CAA TGC GAG TGG CTG TCC CGG CTG CAG AGC GGC      2592
Thr Ser Leu Asn Gly Gln Cys Glu Trp Leu Ser Arg Leu Gln Ser Gly
    850                 855                 860

ATG GTG CCG AAC CAA TAC AAC CCT CTG CGG GGG AGT GTG GCA CCA GGC      2640
Met Val Pro Asn Gln Tyr Asn Pro Leu Arg Gly Ser Val Ala Pro Gly
865                 870                 875                 880

CCC CTG AGC ACA CAG GCC CCC TCC CTG CAG CAT GGC ATG GTA GGC CCG      2688
Pro Leu Ser Thr Gln Ala Pro Ser Leu Gln His Gly Met Val Gly Pro
                885                 890                 895

CTG CAC AGT AGC CTT GCT GCC AGC GCC CTG TCC CAG ATG ATG AGC TAC      2736
Leu His Ser Ser Leu Ala Ala Ser Ala Leu Ser Gln Met Met Ser Tyr
            900                 905                 910
```

```
CAG GGC CTG CCC AGC ACC CGG CTG GCC ACC CAG CCT CAC CTG GTG CAG      2784
Gln Gly Leu Pro Ser Thr Arg Leu Ala Thr Gln Pro His Leu Val Gln
        915                 920                 925

ACC CAG CAG GTG CAG CCA CAA AAC TTA CAG ATG CAG CAG CAG AAC CTG      2832
Thr Gln Gln Val Gln Pro Gln Asn Leu Gln Met Gln Gln Gln Asn Leu
    930                 935                 940

CAG CCA GCA AAC ATC CAG CAG CAG CAA AGC CTG CAG CCG CCA CCA CCA      2880
Gln Pro Ala Asn Ile Gln Gln Gln Gln Ser Leu Gln Pro Pro Pro Pro
945                 950                 955                 960

CCA CCA CAG CCG CAC CTT GGC GTG AGC TCA GCA GCC AGC GGC CAC CTG      2928
Pro Pro Gln Pro His Leu Gly Val Ser Ser Ala Ala Ser Gly His Leu
                965                 970                 975

GGC CGG AGC TTC CTG AGT GGA GAG CCG AGC CAG GCA GAC GTG CAG CCA      2976
Gly Arg Ser Phe Leu Ser Gly Glu Pro Ser Gln Ala Asp Val Gln Pro
            980                 985                 990

CTG GGC CCC AGC AGC CTG GCG GTG CAC ACT ATT CTG CCC CAG GAG AGC      3024
Leu Gly Pro Ser Ser Leu Ala Val His Thr Ile Leu Pro Gln Glu Ser
        995                 1000                1005

CCC GCC CTG CCC ACG TCG CTG CCA TCC TCG CTG GTC CCA CCC GTG ACC      3072
Pro Ala Leu Pro Thr Ser Leu Pro Ser Ser Leu Val Pro Pro Val Thr
    1010                1015                1020

GCA GCC CAG TTC CTG ACG CCC CCC TCG CAG CAC AGC TAC TCC TCG CCT      3120
Ala Ala Gln Phe Leu Thr Pro Pro Ser Gln His Ser Tyr Ser Ser Pro
1025                1030                1035                1040

GTG GAC AAC ACC CCC AGC CAC CAG CTA CAG GTG CCT GTT CCT GTA ATG      3168
Val Asp Asn Thr Pro Ser His Gln Leu Gln Val Pro Val Pro Val Met
                1045                1050                1055

GTA ATG ATC CGA TCT TCG GAT CCT TCT AAA GGC TCA TCA ATT TTG ATC      3216
Val Met Ile Arg Ser Ser Asp Pro Ser Lys Gly Ser Ser Ile Leu Ile
            1060                1065                1070

GAA GCT CCC GAC TCA TGG                                             3234
Glu Ala Pro Asp Ser Trp
        1075
```

(2) ANGABEN ZU SEQ ID NO:11:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 1078 Aminosäuren
(B) ART: Aminosäure
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:11:

```
Cys Gln Glu Asp Ala Gly Asn Lys Val Cys Ser Leu Gln Cys Asn Asn
  1               5                  10                  15

His Ala Cys Gly Trp Asp Gly Gly Asp Cys Ser Leu Asn Phe Asn Asp
             20                  25                  30

Pro Trp Lys Asn Cys Thr Gln Ser Leu Gln Cys Trp Lys Tyr Phe Ser
         35                  40                  45

Asp Gly His Cys Asp Ser Gln Cys Asn Ser Ala Gly Cys Leu Phe Asp
     50                  55                  60
```

Gly Phe Asp Cys Gln Arg Ala Glu Gly Gln Cys Asn Pro Leu Tyr Asp
65 70 75 80

Gln Tyr Cys Lys Asp His Phe Ser Asp Gly His Cys Asp Gln Gly Cys
85 90 95

Asn Ser Ala Glu Cys Glu Trp Asp Gly Leu Asp Cys Ala Glu His Val
100 105 110

Pro Glu Arg Leu Ala Ala Gly Thr Leu Val Val Val Val Leu Met Pro
115 120 125

Pro Glu Gln Leu Arg Asn Ser Ser Phe His Phe Leu Arg Glu Leu Ser
130 135 140

Arg Val Leu His Thr Asn Val Val Phe Lys Arg Asp Ala His Gly Gln
145 150 155 160

Gln Met Ile Phe Pro Tyr Tyr Gly Arg Glu Glu Glu Leu Arg Lys His
165 170 175

Pro Ile Lys Arg Ala Ala Glu Gly Trp Ala Ala Pro Asp Ala Leu Leu
180 185 190

Gly Gln Val Lys Ala Ser Leu Leu Pro Gly Gly Ser Glu Gly Gly Arg
195 200 205

Arg Arg Arg Glu Leu Asp Pro Met Asp Val Arg Gly Ser Ile Val Tyr
210 215 220

Leu Glu Ile Asp Asn Arg Gln Cys Val Gln Ala Ser Ser Gln Cys Phe
225 230 235 240

Gln Ser Ala Thr Asp Val Ala Ala Phe Leu Gly Ala Leu Ala Ser Leu
245 250 255

Gly Ser Leu Asn Ile Pro Tyr Lys Ile Glu Ala Val Gln Ser Glu Thr
260 265 270

Val Glu Pro Pro Pro Pro Ala Gln Leu His Phe Met Tyr Val Ala Ala
275 280 285

Ala Ala Phe Val Leu Leu Phe Phe Val Gly Cys Gly Val Leu Leu Ser
290 295 300

Arg Lys Arg Arg Arg Gln His Gly Gln Leu Trp Phe Pro Glu Gly Phe
305 310 315 320

Lys Val Ser Glu Ala Ser Lys Lys Lys Arg Arg Glu Pro Leu Gly Glu
325 330 335

Asp Ser Val Gly Leu Lys Pro Leu Lys Asn Ala Ser Asp Gly Ala Leu
340 345 350

Met Asp Asp Asn Gln Asn Glu Trp Gly Asp Glu Asp Leu Glu Thr Lys
355 360 365

Lys Phe Arg Phe Glu Glu Pro Val Val Leu Pro Asp Leu Asp Asp Gln
370 375 380

Thr Asp His Arg Gln Trp Thr Gln Gln His Leu Asp Ala Ala Asp Leu
385 390 395 400

Arg Met Ser Ala Met Ala Pro Thr Pro Pro Gln Gly Glu Val Asp Ala
405 410 415

Asp Cys Met Asp Val Asn Val Arg Gly Pro Asp Gly Phe Thr Pro Leu

420 425 430

Met Ile Ala Ser Cys Ser Gly Gly Gly Leu Glu Thr Gly Asn Ser Glu
435 440 445

Glu Glu Glu Asp Ala Pro Ala Val Ile Ser Asp Phe Ile Tyr Gln Gly
450 455 460

Ala Ser Leu His Asn Gln Thr Asp Arg Thr Gly Glu Thr Ala Leu His
465 470 475 480

Leu Ala Ala Arg Tyr Ser Arg Ser Asp Ala Ala Lys Arg Leu Leu Glu
485 490 495

Ala Ser Ala Asp Ala Asn Ile Gln Asp Asn Met Gly Arg Thr Pro Leu
500 505 510

His Ala Ala Val Ser Ala Asp Ala Gln Gly Val Phe Gln Ile Leu Ile
515 520 525

Arg Asn Arg Ala Thr Asp Leu Asp Ala Arg Met His Asp Gly Thr Thr
530 535 540

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

Leu Ile Asn Ser His Ala Asp Val Asn Ala Val Asp Asp Leu Gly Lys
565 570 575

Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn Val Asp Ala Ala Val
580 585 590

Val Leu Leu Lys Asn Gly Ala Asn Lys Asp Met Gln Asn Asn Arg Glu
595 600 605

Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly Ser Tyr Glu Thr Ala
610 615 620

Lys Val Leu Leu Asp His Phe Ala Asn Arg Asp Ile Thr Asp His Met
625 630 635 640

Asp Arg Leu Pro Arg Asp Ile Ala Gln Glu Arg Met His His Asp Ile
645 650 655

Val Arg Leu Leu Asp Glu Tyr Asn Leu Val Arg Ser Pro Gln Leu His
660 665 670

Gly Ala Pro Leu Gly Gly Thr Pro Thr Leu Ser Pro Pro Leu Cys Ser
675 680 685

Pro Asn Gly Tyr Leu Gly Ser Leu Lys Pro Gly Val Gln Gly Lys Lys
690 695 700

Val Arg Lys Pro Ser Ser Lys Gly Leu Ala Cys Gly Ser Lys Glu Ala
705 710 715 720

Lys Asp Leu Lys Ala Arg Arg Lys Lys Ser Gln Asp Gly Lys Gly Cys
725 730 735

Leu Leu Asp Ser Ser Gly Met Leu Ser Pro Val Asp Ser Leu Glu Ser
740 745 750

Pro His Gly Tyr Leu Ser Asp Val Ala Ser Pro Pro Leu Leu Pro Ser
755 760 765

Pro Phe Gln Gln Ser Pro Ser Val Pro Leu Asn His Leu Pro Gly Met
770 775 780

```
Pro Asp Thr His Leu Gly Ile Gly His Leu Asn Val Ala Ala Lys Pro
785                 790                 795                 800

Glu Met Ala Ala Leu Gly Gly Gly Gly Arg Leu Ala Phe Glu Thr Gly
                805                 810                 815

Pro Pro Arg Leu Ser His Leu Pro Val Ala Ser Gly Thr Ser Thr Val
            820                 825                 830

Leu Gly Ser Ser Ser Gly Gly Ala Leu Asn Phe Thr Val Gly Gly Ser
        835                 840                 845

Thr Ser Leu Asn Gly Gln Cys Glu Trp Leu Ser Arg Leu Gln Ser Gly
    850                 855                 860

Met Val Pro Asn Gln Tyr Asn Pro Leu Arg Gly Ser Val Ala Pro Gly
865                 870                 875                 880

Pro Leu Ser Thr Gln Ala Pro Ser Leu Gln His Gly Met Val Gly Pro
                885                 890                 895

Leu His Ser Ser Leu Ala Ala Ser Ala Leu Ser Gln Met Met Ser Tyr
            900                 905                 910

Gln Gly Leu Pro Ser Thr Arg Leu Ala Thr Gln Pro His Leu Val Gln
        915                 920                 925

Thr Gln Gln Val Gln Pro Gln Asn Leu Gln Met Gln Gln Gln Asn Leu
    930                 935                 940

Gln Pro Ala Asn Ile Gln Gln Gln Gln Ser Leu Gln Pro Pro Pro Pro
945                 950                 955                 960

Pro Pro Gln Pro His Leu Gly Val Ser Ser Ala Ala Ser Gly His Leu
                965                 970                 975

Gly Arg Ser Phe Leu Ser Gly Glu Pro Ser Gln Ala Asp Val Gln Pro
            980                 985                 990

Leu Gly Pro Ser Ser Leu Ala Val His Thr Ile Leu Pro Gln Glu Ser
        995                 1000                1005

Pro Ala Leu Pro Thr Ser Leu Pro Ser Ser Leu Val Pro Pro Val Thr
    1010                1015                1020

Ala Ala Gln Phe Leu Thr Pro Pro Ser Gln His Ser Tyr Ser Ser Pro
1025                1030                1035                1040

Val Asp Asn Thr Pro Ser His Gln Leu Gln Val Pro Val Pro Val Met
                1045                1050                1055

Val Met Ile Arg Ser Ser Asp Pro Ser Lys Gly Ser Ser Ile Leu Ile
            1060                1065                1070

Glu Ala Pro Asp Ser Trp
        1075
```

(2) ANGABEN ZU SEQ ID NO:12:

(i) SEQUENZKENNZEICHEN:
- (A) LÄNGE: 4268 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: cDNA

(ix) MERKMAL:
(A) NAME/SCHLÜSSEL: CDS
(B) LAGE: 2..1972

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
G GAG GTG GAT GTG TTA GAT GTG AAT GTC CGT GGC CCA GAT GGC TGC          46
  Glu Val Asp Val Leu Asp Val Asn Val Arg Gly Pro Asp Gly Cys
    1               5                  10                  15

ACC CCA TTG ATG TTG GCT TCT CTC CGA GGA GGC AGC TCA GAT TTG AGT        94
Thr Pro Leu Met Leu Ala Ser Leu Arg Gly Gly Ser Ser Asp Leu Ser
                20                  25                  30

GAT GAA GAT GAA GAT GCA GAG GAC TCT TCT GCT AAC ATC ATC ACA GAC       142
Asp Glu Asp Glu Asp Ala Glu Asp Ser Ser Ala Asn Ile Ile Thr Asp
            35                  40                  45

TTG GTC TAC CAG GGT GCC AGC CTC CAG GCC CAG ACA GAC CGG ACT GGT       190
Leu Val Tyr Gln Gly Ala Ser Leu Gln Ala Gln Thr Asp Arg Thr Gly
        50                  55                  60

GAG ATG GCC CTG CAC CTT GCA GCC CGC TAC TCA CGG GCT GAT GCT GCC       238
Glu Met Ala Leu His Leu Ala Ala Arg Tyr Ser Arg Ala Asp Ala Ala
    65                  70                  75

AAG CGT CTC CTG GAT GCA GGT GCA GAT GCC AAT GCC CAG GAC AAC ATG       286
Lys Arg Leu Leu Asp Ala Gly Ala Asp Ala Asn Ala Gln Asp Asn Met
80                  85                  90                  95

GGC CGC TGT CCA CTC CAT GCT GCA GTG GCA GCT GAT GCC CAA GGT GTC       334
Gly Arg Cys Pro Leu His Ala Ala Val Ala Ala Asp Ala Gln Gly Val
                100                 105                 110

TTC CAG ATT CTG ATT CGC AAC CGA GTA ACT GAT CTA GAT GCC AGG ATG       382
Phe Gln Ile Leu Ile Arg Asn Arg Val Thr Asp Leu Asp Ala Arg Met
            115                 120                 125

AAT GAT GGT ACT ACA CCC CTG ATC CTG GCT GCC CGC CTG GCT GTG GAG       430
Asn Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu
        130                 135                 140

GGA ATG GTG GCA GAA CTG ATC AAC TGC CAA GCG GAT GTG AAT GCA GTG       478
Gly Met Val Ala Glu Leu Ile Asn Cys Gln Ala Asp Val Asn Ala Val
    145                 150                 155

GAT GAC CAT GGA AAA TCT GCT CTT CAC TGG GCA GCT GCT GTC AAT AAT       526
Asp Asp His Gly Lys Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn
160                 165                 170                 175

GTG GAG GCA ACT CTT TTG TTG TTG AAA AAT GGG GCC AAC CGA GAC ATG       574
Val Glu Ala Thr Leu Leu Leu Leu Lys Asn Gly Ala Asn Arg Asp Met
                180                 185                 190

CAG GAC AAC AAG GAA GAG ACA CCT CTG TTT CTT GCT GCC CGG GAG GGG       622
Gln Asp Asn Lys Glu Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly
            195                 200                 205

AGC TAT GAA GCA GCC AAG ATC CTG TTA GAC CAT TTT GCC AAT CGA GAC       670
Ser Tyr Glu Ala Ala Lys Ile Leu Leu Asp His Phe Ala Asn Arg Asp
        210                 215                 220

ATC ACA GAC CAT ATG GAT CGT CTT CCC CGG GAT GTG GCT CGG GAT CGC       718
Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp Val Ala Arg Asp Arg
    225                 230                 235
```

```
ATG CAC CAT GAC ATT GTG CGC CTT CTG GAT GAA TAC AAT GTG ACC CCA      766
Met His His Asp Ile Val Arg Leu Leu Asp Glu Tyr Asn Val Thr Pro
240                 245                 250                 255

AGC CCT CCA GGC ACC GTG TTG ACT TCT GCT CTC TCA CCT GTC ATC TGT      814
Ser Pro Pro Gly Thr Val Leu Thr Ser Ala Leu Ser Pro Val Ile Cys
                260                 265                 270

GGG CCC AAC AGA TCT TTC CTC AGC CTG AAG CAC ACC CCA ATG GGC AAG      862
Gly Pro Asn Arg Ser Phe Leu Ser Leu Lys His Thr Pro Met Gly Lys
            275                 280                 285

AAG TCT AGA CGG CCC AGT GCC AAG AGT ACC ATG CCT ACT AGC CTC CCT      910
Lys Ser Arg Arg Pro Ser Ala Lys Ser Thr Met Pro Thr Ser Leu Pro
        290                 295                 300

AAC CTT GCC AAG GAG GCA AAG GAT GCC AAG GGT AGT AGG AGG AAG AAG      958
Asn Leu Ala Lys Glu Ala Lys Asp Ala Lys Gly Ser Arg Arg Lys Lys
    305                 310                 315

TCT CTG AGT GAG AAG GTC CAA CTG TCT GAG AGT TCA GTA ACT TTA TCC     1006
Ser Leu Ser Glu Lys Val Gln Leu Ser Glu Ser Ser Val Thr Leu Ser
320                 325                 330                 335

CCT GTT GAT TCC CTA GAA TCT CCT CAC ACG TAT GTT TCC GAC ACC ACA     1054
Pro Val Asp Ser Leu Glu Ser Pro His Thr Tyr Val Ser Asp Thr Thr
                340                 345                 350

TCC TCT CCA ATG ATT ACA TCC CCT GGG ATC TTA CAG GCC TCA CCC AAC     1102
Ser Ser Pro Met Ile Thr Ser Pro Gly Ile Leu Gln Ala Ser Pro Asn
            355                 360                 365

CCT ATG TTG GCC ACT GCC GCC CCT CCT GCC CCA GTC CAT GCC CAG CAT     1150
Pro Met Leu Ala Thr Ala Ala Pro Pro Ala Pro Val His Ala Gln His
        370                 375                 380

GCA CTA TCT TTT TCT AAC CTT CAT GAA ATG CAG CCT TTG GCA CAT GGG     1198
Ala Leu Ser Phe Ser Asn Leu His Glu Met Gln Pro Leu Ala His Gly
    385                 390                 395

GCC AGC ACT GTG CTT CCC TCA GTG AGC CAG TTG CTA TCC CAC CAC CAC     1246
Ala Ser Thr Val Leu Pro Ser Val Ser Gln Leu Leu Ser His His His
400                 405                 410                 415

ATT GTG TCT CCA GGC AGT GGC AGT GCT GGA AGC TTG AGT AGG CTC CAT     1294
Ile Val Ser Pro Gly Ser Gly Ser Ala Gly Ser Leu Ser Arg Leu His
                420                 425                 430

CCA GTC CCA GTC CCA GCA GAT TGG ATG AAC CGC ATG GAG GTG AAT GAG     1342
Pro Val Pro Val Pro Ala Asp Trp Met Asn Arg Met Glu Val Asn Glu
            435                 440                 445

ACC CAG TAC AAT GAG ATG TTT GGT ATG GTC CTG GCT CCA GCT GAG GGC     1390
Thr Gln Tyr Asn Glu Met Phe Gly Met Val Leu Ala Pro Ala Glu Gly
        450                 455                 460

ACC CAT CCT GGC ATA GCT CCC CAG AGC AGG CCA CCT GAA GGG AAG CAC     1438
Thr His Pro Gly Ile Ala Pro Gln Ser Arg Pro Pro Glu Gly Lys His
    465                 470                 475

ATA ACC ACC CCT CGG GAG CCC TTG CCC CCC ATT GTG ACT TTC CAG CTC     1486
Ile Thr Thr Pro Arg Glu Pro Leu Pro Pro Ile Val Thr Phe Gln Leu
480                 485                 490                 495

ATC CCT AAA GGC AGT ATT GCC CAA CCA GCG GGG GCT CCC CAG CCT CAG     1534
Ile Pro Lys Gly Ser Ile Ala Gln Pro Ala Gly Ala Pro Gln Pro Gln
                500                 505                 510
```

```
TCC ACC TGC CCT CCA GCT GTT GCG GGC CCC CTG CCC ACC ATG TAC CAG      1582
Ser Thr Cys Pro Pro Ala Val Ala Gly Pro Leu Pro Thr Met Tyr Gln
            515                 520                 525

ATT CCA GAA ATG GCC CGT TTG CCC AGT GTG GCT TTC CCC ACT GCC ATG      1630
Ile Pro Glu Met Ala Arg Leu Pro Ser Val Ala Phe Pro Thr Ala Met
        530                 535                 540

ATG CCC CAG CAG GAC GGG CAG GTA GCT CAG ACC ATT CTC CCA GCC TAT      1678
Met Pro Gln Gln Asp Gly Gln Val Ala Gln Thr Ile Leu Pro Ala Tyr
    545                 550                 555

CAT CCT TTC CCA GCC TCT GTG GGC AAG TAC CCC ACA CCC CCT TCA CAG      1726
His Pro Phe Pro Ala Ser Val Gly Lys Tyr Pro Thr Pro Pro Ser Gln
560                 565                 570                 575

CAC AGT TAT GCT TCC TCA AAT GCT GCT GAG CGA ACA CCC AGT CAC AGT      1774
His Ser Tyr Ala Ser Ser Asn Ala Ala Glu Arg Thr Pro Ser His Ser
                580                 585                 590

GGT CAC CTC CAG GGT GAG CAT CCC TAC CTG ACA CCA TCC CCA GAG TCT      1822
Gly His Leu Gln Gly Glu His Pro Tyr Leu Thr Pro Ser Pro Glu Ser
            595                 600                 605

CCT GAC CAG TGG TCA AGT TCA TCA CCC CAC TCT GCT TCT GAC TGG TCA      1870
Pro Asp Gln Trp Ser Ser Ser Ser Pro His Ser Ala Ser Asp Trp Ser
        610                 615                 620

GAT GTG ACC ACC AGC CCT ACC CCT GGG GGT GCT GGA GGA GGT CAG CGG      1918
Asp Val Thr Thr Ser Pro Thr Pro Gly Gly Ala Gly Gly Gly Gln Arg
    625                 630                 635

GGA CCT GGG ACA CAC ATG TCT GAG CCA CCA CAC AAC AAC ATG CAG GTT      1966
Gly Pro Gly Thr His Met Ser Glu Pro Pro His Asn Asn Met Gln Val
640                 645                 650                 655

TAT GCG TGAGAGAGTC CACCTCCAGT GTAGAGACAT AACTGACTTT TGTAAATGCT       2022
Tyr Ala
```

```
GCTGAGGAAC AAATGAAGGT CATCCGGGAG AGAAATGAAG AAATCTCTGG AGCCAGCTTC    2082

TAGAGGTAGG AAAGAGAAGA TGTTCTTATT CAGATAATGC AAGAGAAGCA ATTCGTCAGT    2142

TTCACTGGGT ATCTGCAAGG CTTATTGATT ATTCTAATCT AATAAGACAA GTTTGTGGAA    2202

ATGCAAGATG AATACAAGCC TTGGGTCCAT GTTTACTCTC TTCTATTTGG AGAATAAGAT    2262

GGATGCTTAT TGAAGCCCAG ACATTCTTGC AGCTTGGACT GCATTTTAAG CCCTGCAGGC    2322

TTCTGCCATA TCCATGAGAA GATTCTACAC TAGCGTCCTG TTGGGAATTA TGCCCTGGAA    2382

TTCTGCCTGA ATTGACCTAC GCATCTCCTC CTCCTTGGAC ATTCTTTTGT CTTCATTTGG    2442

TGCTTTTGGT TTTGCACCTC TCCGTGATTG TAGCCCTACC AGCATGTTAT AGGGCAAGAC    2502

CTTTGTGCTT TTGATCATTC TGGCCCATGA AAGCAACTTT GGTCTCCTTT CCCCTCCTGT    2562

CTTCCCGGTA TCCCTTGGAG TCTCACAAGG TTTACTTTGG TATGGTTCTC AGCACAAACC    2622

TTTCAAGTAT GTTGTTTCTT TGGAAAATGG ACATACTGTA TTGTGTTCTC CTGCATATAT    2682

CATTCCTGGA GAGAGAAGGG GAGAAGAATA CTTTTCTTCA ACAAATTTTG GGGGCAGGAG    2742

ATCCCTTCAA GAGGCTGCAC CTTAATTTTT CTTGTCTGTG TGCAGGTCTT CATATAAACT    2802

TTACCAGGAA GAAGGGTGTG AGTTTGTTGT TTTTCTGTGT ATGGGCCTGG TCAGTGTAAA    2862
```

GTTTTATCCT TGATAGTCTA GTTACTATGA CCCTCCCCAC TTTTTTAAAA CCAGAAAAAG 2922

GTTTGGAATG TTGGAATGAC CAAGAGACAA GTTAACTCGT GCAAGAGCCA GTTACCCACC 2982

CACAGGTCCC CCTACTTCCT GCCAAGCATT CCATTGACTG CCTGTATGGA ACACATTTGT 3042

CCCAGATCTG AGCATTCTAG GCCTGTTTCA CTCACTCACC CAGCATATGA AACTAGTCTT 3102

AACTGTTGAG CCTTTCCTTT CATATCCACA GAAGACACTG TCTCAAATGT TGTACCCTTG 3162

CCATTTAGGA CTGAACTTTC CTTAGCCCAA GGGACCCAGT GACAGTTGTC TTCCGTTTGT 3222

CAGATGATCA GTCTCTACTG ATTATCTTGC TGCTTAAAGG CCTGCTCACC AATCTTTCTT 3282

TCACACCGTG TGGTCCGTGT TACTGGTATA CCCAGTATGT TCTCACTGAA GACATGGACT 3342

TTATATGTTC AAGTGCAGGA ATTGGAAAGT TGGACTTGTT TTCTATGATC CAAAACAGCC 3402

CTATAAGAAG GTTGGAAAAG GAGGAACTAT ATAGCAGCCT TTGCTATTTT CTGCTACCAT 3462

TTCTTTTCCT CTGAAGCGGC CATGACATTC CCTTTGGCAA CTAACGTAGA AACTCAACAG 3522

AACATTTTCC TTTCCTAGAG TCACCTTTTA GATGATAATG GACAACTATA GACTTGCTCA 3582

TTGTTCAGAC TGATTGCCCC TCACCTGAAT CCACTCTCTG TATTCATGCT CTTGGCAATT 3642

TCTTTGACTT TCTTTTAAGG GCAGAAGCAT TTTAGTTAAT TGTAGATAAA GAATAGTTTT 3702

CTTCCTCTTC TCCTTGGGCC AGTTAATAAT TGGTCCATGG CTACACTGCA ACTTCCGTCC 3762

AGTGCTGTGA TGCCCATGAC ACCTGCAAAA TAAGTTCTGC CTGGGCATTT TGTAGATATT 3822

AACAGGTGAA TTCCCGACTC TTTTGGTTTG AATGACAGTT CTCATTCCTT CTATGGCTGC 3882

AAGTATGCAT CAGTGCTTCC CACTTACCTG ATTTGTCTGT CGGTGGCCCC ATATGGAAAC 3942

CCTGCGTGTC TGTTGGCATA ATAGTTTACA AATGGTTTTT TCAGTCCTAT CCAAATTTAT 4002

TGAACCAACA AAAATAATTA CTTCTGCCCT GAGATAAGCA GATTAAGTTT GTTCATTCTC 4062

TGCTTTATTC TCTCCATGTG GCAACATTCT GTCAGCCTCT TTCATAGTGT GCAAACATTT 4122

TATCATTCTA AATGGTGACT CTCTGCCCTT GGACCCATTT ATTATTCACA GATGGGGAGA 4182

ACCTATCTGC ATGGACCCTC ACCATCCTCT GTGCAGCACA CACAGTGCAG GGAGCCAGTG 4242

GCGATGGCGA TGACTTTCTT CCCCTG 4268

(2) ANGABEN ZU SEQ ID NO:13:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 657 Aminosäuren
(B) ART: Aminosäure
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:13:

Glu Val Asp Val Leu Asp Val Asn Val Arg Gly Pro Asp Gly Cys Thr
1 5 10 15

Pro Leu Met Leu Ala Ser Leu Arg Gly Gly Ser Ser Asp Leu Ser Asp
20 25 30

Glu Asp Glu Asp Ala Glu Asp Ser Ser Ala Asn Ile Ile Thr Asp Leu

```
        35                  40                  45

Val Tyr Gln Gly Ala Ser Leu Gln Ala Gln Thr Asp Arg Thr Gly Glu
     50                  55                  60

Met Ala Leu His Leu Ala Ala Arg Tyr Ser Arg Ala Asp Ala Ala Lys
 65                  70                  75                  80

Arg Leu Leu Asp Ala Gly Ala Asp Ala Asn Ala Gln Asp Asn Met Gly
                 85                  90                  95

Arg Cys Pro Leu His Ala Ala Val Ala Ala Asp Ala Gln Gly Val Phe
            100                 105                 110

Gln Ile Leu Ile Arg Asn Arg Val Thr Asp Leu Asp Ala Arg Met Asn
        115                 120                 125

Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu Gly
    130                 135                 140

Met Val Ala Glu Leu Ile Asn Cys Gln Ala Asp Val Asn Ala Val Asp
145                 150                 155                 160

Asp His Gly Lys Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn Val
                165                 170                 175

Glu Ala Thr Leu Leu Leu Leu Lys Asn Gly Ala Asn Arg Asp Met Gln
            180                 185                 190

Asp Asn Lys Glu Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly Ser
        195                 200                 205

Tyr Glu Ala Ala Lys Ile Leu Leu Asp His Phe Ala Asn Arg Asp Ile
    210                 215                 220

Thr Asp His Met Asp Arg Leu Pro Arg Asp Val Ala Arg Asp Arg Met
225                 230                 235                 240

His His Asp Ile Val Arg Leu Leu Asp Glu Tyr Asn Val Thr Pro Ser
                245                 250                 255

Pro Pro Gly Thr Val Leu Thr Ser Ala Leu Ser Pro Val Ile Cys Gly
            260                 265                 270

Pro Asn Arg Ser Phe Leu Ser Leu Lys His Thr Pro Met Gly Lys Lys
        275                 280                 285

Ser Arg Arg Pro Ser Ala Lys Ser Thr Met Pro Thr Ser Leu Pro Asn
    290                 295                 300

Leu Ala Lys Glu Ala Lys Asp Ala Lys Gly Ser Arg Arg Lys Lys Ser
305                 310                 315                 320

Leu Ser Glu Lys Val Gln Leu Ser Glu Ser Ser Val Thr Leu Ser Pro
                325                 330                 335

Val Asp Ser Leu Glu Ser Pro His Thr Tyr Val Ser Asp Thr Thr Ser
            340                 345                 350

Ser Pro Met Ile Thr Ser Pro Gly Ile Leu Gln Ala Ser Pro Asn Pro
        355                 360                 365

Met Leu Ala Thr Ala Ala Pro Pro Ala Pro Val His Ala Gln His Ala
    370                 375                 380

Leu Ser Phe Ser Asn Leu His Glu Met Gln Pro Leu Ala His Gly Ala
385                 390                 395                 400
```

```
Ser Thr Val Leu Pro Ser Val Ser Gln Leu Leu Ser His His His Ile
                405                 410                 415

Val Ser Pro Gly Ser Gly Ser Ala Gly Ser Leu Ser Arg Leu His Pro
            420                 425                 430

Val Pro Val Pro Ala Asp Trp Met Asn Arg Met Glu Val Asn Glu Thr
        435                 440                 445

Gln Tyr Asn Glu Met Phe Gly Met Val Leu Ala Pro Ala Glu Gly Thr
    450                 455                 460

His Pro Gly Ile Ala Pro Gln Ser Arg Pro Pro Glu Gly Lys His Ile
465                 470                 475                 480

Thr Thr Pro Arg Glu Pro Leu Pro Pro Ile Val Thr Phe Gln Leu Ile
                485                 490                 495

Pro Lys Gly Ser Ile Ala Gln Pro Ala Gly Ala Pro Gln Pro Gln Ser
            500                 505                 510

Thr Cys Pro Pro Ala Val Ala Gly Pro Leu Pro Thr Met Tyr Gln Ile
        515                 520                 525

Pro Glu Met Ala Arg Leu Pro Ser Val Ala Phe Pro Thr Ala Met Met
    530                 535                 540

Pro Gln Gln Asp Gly Gln Val Ala Gln Thr Ile Leu Pro Ala Tyr His
545                 550                 555                 560

Pro Phe Pro Ala Ser Val Gly Lys Tyr Pro Thr Pro Pro Ser Gln His
                565                 570                 575

Ser Tyr Ala Ser Ser Asn Ala Ala Glu Arg Thr Pro Ser His Ser Gly
            580                 585                 590

His Leu Gln Gly Glu His Pro Tyr Leu Thr Pro Ser Pro Glu Ser Pro
        595                 600                 605

Asp Gln Trp Ser Ser Ser Ser Pro His Ser Ala Ser Asp Trp Ser Asp
    610                 615                 620

Val Thr Thr Ser Pro Thr Pro Gly Gly Ala Gly Gly Gly Gln Arg Gly
625                 630                 635                 640

Pro Gly Thr His Met Ser Glu Pro Pro His Asn Asn Met Gln Val Tyr
                645                 650                 655

Ala
```

(2) ANGABEN ZU SEQ ID NO:14:

(i) SEQUENZKENNZEICHEN:
- (A) LÄNGE: 77 AMINOSÄUREN
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:14:

```
Glu Asp Ile Asp Glu Cys Asp Gln Gly Ser Pro Cys Glu His Asn Gly
1               5                   10                  15
```

Ile Cys Val Asn Thr Pro Gly Ser Tyr Arg Cys Asn Cys Ser Gln Gly
20 25 30

Phe Thr Gly Pro Arg Cys Glu Thr Asn Ile Asn Glu Cys Glu Ser His
35 40 45

Pro Cys Gln Asn Glu Gly Ser Cys Leu Asp Asp Pro Gly Thr Phe Arg
50 55 60

Cys Val Cys Met Pro Gly Phe Thr Gly Thr Gln Cys Glu
65 70 75

(2) ANGABEN ZU SEQ ID NO:15:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 78 Aminosäuren
(B) ART: Aminosäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:15:

Asn Asp Val Asp Glu Cys Ser Leu Gly Ala Asn Pro Cys Glu His Gly
1 5 10 15

Gly Arg Cys Thr Asn Thr Leu Gly Ser Phe Gln Cys Asn Cys Pro Gln
20 25 30

Gly Tyr Ala Gly Pro Arg Cys Glu Ile Asp Val Asn Glu Cys Leu Ser
35 40 45

Asn Pro Cys Gln Asn Asp Ser Thr Cys Leu Asp Gln Ile Gly Glu Phe
50 55 60

Gln Cys Ile Cys Met Pro Gly Tyr Glu Gly Leu Tyr Cys Glu
65 70 75

(2) ANGABEN ZU SEQ ID NO:16:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 654 Aminosäuren
(B) ART: Aminosäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:16:

Thr Pro Pro Gln Gly Glu Ile Glu Ala Asp Cys Met Asp Val Asn Val
1 5 10 15

Arg Gly Pro Asp Gly Phe Thr Pro Leu Met Ile Ala Ser Cys Ser Gly
20 25 30

Gly Gly Leu Glu Thr Gly Asn Ser Glu Glu Glu Glu Asp Ala Ser Ala
35 40 45

Asn Met Ile Ser Asp Phe Ile Gly Gln Gly Ala Gln Leu His Asn Gln

50 55 60

Thr Asp Arg Thr Gly Glu Thr Ala Leu His Leu Ala Ala Arg Tyr Ala
65 70 75 80

Arg Ala Asp Ala Ala Lys Arg Leu Leu Glu Ser Ser Ala Asp Ala Asn
85 90 95

Val Gln Asp Asn Met Gly Arg Thr Pro Leu His Ala Ala Val Ala Ala
100 105 110

Asp Ala Gln Gly Val Phe Gln Ile Leu Ile Arg Asn Arg Ala Thr Asp
115 120 125

Leu Asp Ala Arg Met Phe Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala
130 135 140

Arg Leu Ala Val Glu Gly Met Val Glu Glu Leu Ile Asn Ala His Ala
145 150 155 160

Asp Val Asn Ala Val Asp Glu Phe Gly Lys Ser Ala Leu His Trp Ala
165 170 175

Ala Ala Val Asn Asn Val Asp Ala Ala Ala Val Leu Leu Lys Asn Ser
180 185 190

Ala Asn Lys Asp Met Gln Asn Asn Lys Glu Glu Thr Ser Leu Phe Leu
195 200 205

Ala Ala Arg Glu Gly Ser Tyr Glu Thr Ala Lys Val Leu Leu Asp His
210 215 220

Tyr Ala Asn Arg Asp Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp
225 230 235 240

Ile Ala Gln Glu Arg Met His His Asp Ile Val His Leu Leu Asp Glu
245 250 255

Tyr Asn Leu Val Lys Ser Pro Thr Leu His Asn Gly Pro Leu Gly Ala
260 265 270

Thr Thr Leu Ser Pro Pro Ile Cys Ser Pro Asn Gly Tyr Met Gly Asn
275 280 285

Met Lys Pro Ser Val Gln Ser Lys Lys Ala Arg Lys Pro Ser Ile Lys
290 295 300

Gly Asn Gly Cys Lys Glu Ala Lys Glu Leu Lys Ala Arg Arg Lys Lys
305 310 315 320

Ser Gln Asp Gly Lys Thr Thr Leu Leu Asp Ser Gly Ser Ser Gly Val
325 330 335

Leu Ser Pro Val Asp Ser Leu Glu Ser Thr His Gly Tyr Leu Ser Asp
340 345 350

Val Ser Ser Pro Pro Leu Met Thr Ser Pro Phe Gln Gln Ser Pro Ser
355 360 365

Met Pro Leu Asn His Leu Thr Ser Met Pro Glu Ser Gln Leu Gly Met
370 375 380

Asn His Ile Asn Met Ala Thr Lys Gln Glu Met Ala Ala Gly Ser Asn
385 390 395 400

Arg Met Ala Phe Asp Ala Met Val Pro Arg Leu Thr His Leu Asn Ala
405 410 415

```
Ser Ser Pro Asn Thr Ile Met Ser Asn Gly Ser Met His Phe Thr Val
            420                 425                 430

Gly Gly Ala Pro Thr Met Asn Ser Gln Cys Asp Trp Leu Ala Arg Leu
        435                 440                 445

Gln Asn Gly Met Val Gln Asn Gln Tyr Asp Pro Ile Arg Asn Gly Ile
    450                 455                 460

Gln Gln Gly Asn Ala Gln Gln Ala Gln Ala Leu Gln His Gly Leu Met
465                 470                 475                 480

Thr Ser Leu His Asn Gly Leu Pro Ala Thr Thr Leu Ser Gln Met Met
                485                 490                 495

Thr Tyr Gln Ala Met Pro Asn Thr Arg Leu Ala Asn Gln Pro His Leu
            500                 505                 510

Met Gln Ala Gln Gln Met Gln Gln Gln Gln Asn Leu Gln Leu His Gln
        515                 520                 525

Ser Met Gln Gln Gln His His Asn Ser Ser Thr Thr Ser Thr His Ile
    530                 535                 540

Asn Ser Pro Phe Cys Ser Ser Asp Ile Ser Gln Thr Asp Leu Gln Gln
545                 550                 555                 560

Met Ser Ser Asn Asn Ile His Ser Val Met Pro Gln Asp Thr Gln Ile
                565                 570                 575

Phe Ala Ala Ser Leu Pro Ser Asn Leu Thr Gln Ser Met Thr Thr Ala
            580                 585                 590

Gln Phe Leu Thr Pro Pro Ser Gln His Ser Tyr Ser Ser Pro Met Asp
        595                 600                 605

Asn Thr Pro Ser His Gln Leu Gln Val Pro Asp His Pro Phe Leu Thr
    610                 615                 620

Pro Ser Pro Glu Ser Pro Asp Gln Trp Ser Ser Ser Ser Pro His Ser
625                 630                 635                 640

Asn Met Ser Asp Trp Ser Glu Gly Ile Ser Ser Pro Pro Thr
                645                 650
```

(2) ANGABEN ZU SEQ ID NO:17:

(i) SEQUENZKENNZEICHEN:
- (A) LÄNGE: 666 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:17:

```
Thr Pro Pro Gln Gly Glu Val Asp Ala Asp Cys Met Asp Val Asn Val
1               5                   10                  15

Arg Gly Pro Asp Gly Phe Thr Pro Leu Met Ile Ala Ser Cys Ser Gly
            20                  25                  30

Gly Gly Leu Glu Thr Gly Asn Ser Glu Glu Glu Glu Asp Ala Pro Ala
        35                  40                  45
```

Val Ile Ser Asp Phe Ile Tyr Gln Gly Ala Ser Leu His Asn Gln Thr
50 55 60

Asp Arg Thr Gly Glu Thr Ala Leu His Leu Ala Ala Arg Tyr Ser Arg
65 70 75 80

Ser Asp Ala Ala Lys Arg Leu Leu Glu Ala Ser Ala Asp Ala Asn Ile
85 90 95

Gln Asp Asn Met Gly Arg Thr Pro Leu His Ala Ala Val Ser Ala Asp
100 105 110

Ala Gln Gly Val Phe Gln Ile Leu Leu Arg Asn Arg Ala Thr Asp Leu
115 120 125

Asp Ala Arg Met His Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg
130 135 140

Leu Ala Val Glu Gly Met Leu Glu Asp Leu Ile Asn Ser His Ala Asp
145 150 155 160

Val Asn Ala Val Asp Asp Leu Gly Lys Ser Ala Leu His Trp Ala Ala
165 170 175

Ala Val Asn Asn Val Asp Ala Ala Val Val Leu Leu Lys Asn Gly Ala
180 185 190

Asn Lys Asp Met Gln Asn Asn Lys Glu Glu Thr Pro Leu Phe Leu Ala
195 200 205

Ala Arg Glu Gly Ser Tyr Glu Thr Ala Lys Val Leu Leu Asp His Phe
210 215 220

Ala Asn Arg Asp Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp Ile
225 230 235 240

Ala Gln Glu Arg Met His His Asp Ile Val Arg Leu Leu Asp Glu Tyr
245 250 255

Asn Leu Val Arg Ser Pro Gln Leu His Gly Thr Ala Leu Gly Gly Thr
260 265 270

Pro Thr Leu Ser Pro Thr Leu Cys Ser Pro Asn Gly Tyr Leu Gly Asn
275 280 285

Leu Lys Ser Ala Thr Gln Gly Lys Lys Ala Arg Lys Pro Ser Thr Lys
290 295 300

Gly Leu Ala Cys Ser Ser Lys Glu Ala Lys Asp Leu Lys Ala Arg Arg
305 310 315 320

Lys Lys Ser Gln Asp Gly Lys Gly Cys Leu Leu Asp Ser Ser Ser Met
325 330 335

Leu Ser Pro Val Asp Ser Leu Glu Ser Pro His Gly Tyr Leu Ser Asp
340 345 350

Val Ala Ser Pro Pro Leu Pro Ser Pro Phe Gln Gln Ser Pro Ser Met
355 360 365

Pro Leu Ser His Leu Pro Gly Met Pro Asp Thr His Leu Gly Ile Ser
370 375 380

His Leu Asn Val Ala Ala Lys Pro Glu Met Ala Ala Leu Ala Gly Gly
385 390 395 400

Ser Arg Leu Ala Phe Glu Pro Pro Pro Pro Arg Leu Ser His Leu Pro

405 410 415

Val Ala Ser Ser Ala Ser Thr Val Leu Ser Thr Asn Gly Thr Gly Ala
420 425 430

Met Asn Phe Thr Val Gly Ala Pro Ala Ser Leu Asn Gly Gln Cys Glu
435 440 445

Trp Leu Pro Arg Leu Gln Asn Gly Met Val Pro Ser Gln Tyr Asn Pro
450 455 460

Leu Arg Pro Gly Val Thr Pro Gly Thr Leu Ser Thr Gln Ala Ala Gly
465 470 475 480

Leu Gln His Gly Met Met Ser Pro Ile His Ser Ser Leu Ser Thr Asn
485 490 495

Thr Leu Ser Pro Ile Ile Tyr Gln Gly Leu Pro Asn Thr Arg Leu Ala
500 505 510

Thr Gln Pro His Leu Val Gln Thr Gln Gln Val Gln Pro Gln Asn Leu
515 520 525

Gln Ile Gln Pro Gln Asn Leu Gln Pro Pro Ser Gln Pro His Leu Ser
530 535 540

Val Ser Ser Ala Ala Asn Gly His Leu Gly Arg Ser Phe Leu Ser Gly
545 550 555 560

Glu Pro Ser Gln Ala Asp Val Gln Pro Leu Gly Pro Ser Ser Leu Pro
565 570 575

Val His Thr Ile Leu Pro Gln Glu Ser Gln Ala Leu Pro Thr Ser Leu
580 585 590

Pro Ser Ser Met Val Pro Pro Met Thr Thr Thr Gln Phe Leu Thr Pro
595 600 605

Pro Ser Gln His Ser Tyr Ser Ser Ser Pro Val Asp Asn Thr Pro Ser
610 615 620

His Gln Leu Gln Val Pro Glu His Pro Phe Leu Thr Pro Ser Pro Glu
625 630 635 640

Ser Pro Asp Gln Trp Ser Ser Ser Ser Arg His Ser Asn Ile Ser Asp
645 650 655

Trp Ser Glu Gly Ile Ser Ser Pro Pro Thr
660 665

(2) ANGABEN ZU SEQ ID NO:18:

(i) SEQUENZKENNZEICHEN:
(A) LANGE: 681 Aminosäuren
(B) ART: Aminosäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:18:

Thr Pro Pro Gln Gly Glu Val Asp Ala Asp Cys Met Asp Val Asn Val
1 5 10 15

Arg Gly Pro Asp Gly Phe Thr Pro Leu Met Ile Ala Ser Cys Ser Gly
20 25 30

Gly Gly Leu Glu Thr Gly Asn Ser Glu Glu Glu Glu Asp Ala Pro Ala
35 40 45

Val Ile Ser Asp Phe Ile Tyr Gln Gly Ala Ser Leu His Asn Gln Thr
50 55 60

Asp Arg Thr Gly Glu Thr Ala Leu His Leu Ala Ala Arg Tyr Ser Arg
65 70 75 80

Ser Asp Ala Ala Lys Arg Leu Leu Glu Ala Ser Ala Asp Ala Asn Ile
85 90 95

Gln Asp Asn Met Gly Arg Thr Pro Leu His Ala Ala Val Ser Ala Asp
100 105 110

Ala Gln Gly Val Phe Gln Ile Leu Ile Arg Asn Arg Ala Thr Asp Leu
115 120 125

Asp Ala Arg Met His Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg
130 135 140

Leu Ala Val Glu Gly Met Leu Glu Asp Leu Ile Asn Ser His Ala Asp
145 150 155 160

Val Asn Ala Val Asp Asp Leu Gly Lys Ser Ala Leu His Trp Ala Ala
165 170 175

Ala Val Asn Asn Val Asp Ala Ala Val Val Leu Leu Lys Asn Gly Ala
180 185 190

Asn Lys Asp Met Gln Asn Asn Arg Glu Glu Thr Pro Leu Phe Leu Ala
195 200 205

Ala Arg Glu Gly Ser Tyr Glu Thr Ala Lys Val Leu Leu Asp His Phe
210 215 220

Ala Asn Arg Asp Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp Ile
225 230 235 240

Ala Gln Glu Arg Met His His Asp Ile Val Arg Leu Leu Asp Glu Tyr
245 250 255

Asn Leu Val Arg Ser Pro Gln Leu His Gly Ala Pro Leu Gly Gly Thr
260 265 270

Pro Thr Leu Ser Pro Pro Leu Cys Ser Pro Asn Gly Tyr Leu Gly Ser
275 280 285

Leu Lys Pro Gly Val Gln Gly Lys Lys Val Arg Lys Pro Ser Ser Lys
290 295 300

Gly Leu Ala Cys Gly Ser Lys Glu Ala Lys Asp Leu Lys Ala Arg Arg
305 310 315 320

Lys Lys Ser Gln Asp Gly Lys Gly Cys Leu Leu Asp Ser Ser Gly Met
325 330 335

Leu Ser Pro Val Asp Ser Leu Glu Ser Pro His Gly Tyr Leu Ser Asp
340 345 350

Val Ala Ser Pro Pro Leu Leu Pro Ser Pro Phe Gln Gln Ser Pro Ser
355 360 365

Val Pro Leu Asn His Leu Pro Gly Met Pro Asp Thr His Leu Gly Ile

```
    370                 375                 380

Gly His Leu Asn Val Ala Ala Lys Pro Glu Met Ala Ala Leu Gly Gly
385                 390                 395                 400

Gly Gly Arg Leu Ala Phe Glu Thr Gly Pro Pro Arg Leu Ser His Leu
                405                 410                 415

Pro Val Ala Ser Gly Thr Ser Thr Val Leu Gly Ser Ser Ser Gly Gly
            420                 425                 430

Ala Leu Asn Phe Thr Val Gly Gly Ser Thr Ser Leu Asn Gly Gln Cys
        435                 440                 445

Glu Trp Leu Ser Arg Leu Gln Ser Gly Met Val Pro Asn Gln Tyr Asn
    450                 455                 460

Pro Leu Arg Gly Ser Val Ala Pro Gly Pro Leu Ser Thr Gln Ala Pro
465                 470                 475                 480

Ser Leu Gln His Gly Met Val Gly Pro Leu His Ser Ser Leu Ala Ala
                485                 490                 495

Ser Ala Leu Ser Gln Met Met Ser Tyr Gln Gly Leu Pro Ser Thr Arg
            500                 505                 510

Leu Ala Thr Gln Pro His Leu Val Gln Thr Gln Gln Val Gln Pro Gln
        515                 520                 525

Asn Leu Gln Met Gln Gln Gln Asn Leu Gln Pro Ala Asn Ile Gln Gln
    530                 535                 540

Gln Gln Ser Leu Gln Pro Pro Pro Pro Pro Pro Gln Pro His Leu Gly
545                 550                 555                 560

Val Ser Ser Ala Ala Ser Gly His Leu Gly Arg Ser Phe Leu Ser Gly
                565                 570                 575

Glu Pro Ser Gln Ala Asp Val Gln Pro Leu Gly Pro Ser Ser Leu Ala
            580                 585                 590

Val His Thr Ile Leu Pro Gln Glu Ser Pro Ala Leu Pro Thr Ser Leu
        595                 600                 605

Pro Ser Ser Leu Val Pro Pro Val Thr Ala Ala Gln Phe Leu Thr Pro
    610                 615                 620

Pro Ser Gln His Ser Tyr Ser Ser Pro Val Glu Asn Thr Pro Ser His
625                 630                 635                 640

Gln Leu Gln Val Pro Glu His Pro Phe Leu Thr Pro Ser Pro Glu Ser
                645                 650                 655

Pro Asp Gln Trp Ser Ser Ser Ser Pro His Ser Asn Val Ser Asp Trp
            660                 665                 670

Ser Glu Gly Val Ser Ser Pro Pro Thr
        675                 680
```

(2) ANGABEN ZU SEQ ID NO:19:

(i) SEQUENZKENNZEICHEN:
- (A) LÄNGE: 2471 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:19:

```
Met Pro Ala Leu Arg Pro Ala Leu Leu Trp Ala Leu Leu Ala Leu Trp
1               5                  10                  15

Leu Cys Cys Ala Ala Pro Ala His Ala Leu Gln Cys Arg Asp Gly Tyr
            20                  25                  30

Glu Pro Cys Val Asn Glu Gly Met Cys Val Thr Tyr His Asn Gly Thr
        35                  40                  45

Gly Tyr Cys Lys Cys Pro Glu Gly Phe Leu Gly Glu Tyr Cys Gln His
    50                  55                  60

Arg Asp Pro Cys Glu Lys Asn Arg Cys Gln Asn Gly Gly Thr Cys Val
65                  70                  75                  80

Ala Gln Ala Met Leu Gly Lys Ala Thr Cys Arg Cys Ala Ser Gly Phe
                85                  90                  95

Thr Gly Glu Asp Cys Gln Tyr Ser Thr Ser His Pro Cys Phe Val Ser
            100                 105                 110

Arg Pro Cys Leu Asn Gly Gly Thr Cys His Met Leu Ser Arg Asp Thr
        115                 120                 125

Tyr Glu Cys Thr Cys Gln Val Gly Phe Thr Gly Lys Glu Cys Gln Trp
    130                 135                 140

Thr Asp Ala Cys Leu Ser His Pro Cys Ala Asn Gly Ser Thr Cys Thr
145                 150                 155                 160

Thr Val Ala Asn Gln Phe Ser Cys Lys Cys Leu Thr Gly Phe Thr Gly
                165                 170                 175

Gln Lys Cys Glu Thr Asp Val Asn Glu Cys Asp Ile Pro Gly His Cys
            180                 185                 190

Gln His Gly Gly Thr Cys Leu Asn Leu Pro Gly Ser Tyr Gln Cys Gln
        195                 200                 205

Cys Pro Gln Gly Phe Thr Gly Gln Tyr Cys Asp Ser Leu Tyr Val Pro
    210                 215                 220

Cys Ala Pro Ser Pro Cys Val Asn Gly Gly Thr Cys Arg Gln Thr Gly
225                 230                 235                 240

Asp Phe Thr Phe Glu Cys Asn Cys Leu Pro Gly Phe Glu Gly Ser Thr
                245                 250                 255

Cys Glu Arg Asn Ile Asp Asp Cys Pro Asn His Arg Cys Gln Asn Gly
            260                 265                 270

Gly Val Cys Val Asp Gly Val Asn Thr Tyr Asn Cys Arg Cys Pro Pro
        275                 280                 285

Gln Trp Thr Gly Gln Phe Cys Thr Glu Asp Val Asp Glu Cys Leu Leu
    290                 295                 300

Gln Pro Asn Ala Cys Gln Asn Gly Gly Thr Cys Ala Asn Arg Asn Gly
305                 310                 315                 320

Gly Tyr Gly Cys Val Cys Val Asn Gly Trp Ser Gly Asp Asp Cys Ser
```

```
                325                 330                 335

Glu Asn Ile Asp Asp Cys Ala Phe Ala Ser Cys Thr Pro Gly Ser Thr
            340                 345                 350

Cys Ile Asp Arg Val Ala Ser Phe Ser Cys Met Cys Pro Glu Gly Lys
        355                 360                 365

Ala Gly Leu Leu Cys His Leu Asp Asp Ala Cys Ile Ser Asn Pro Cys
    370                 375                 380

His Lys Gly Ala Leu Cys Asp Thr Asn Pro Leu Asn Gly Gln Tyr Ile
385                 390                 395                 400

Cys Thr Cys Pro Gln Gly Tyr Lys Gly Ala Asp Cys Thr Glu Asp Val
                405                 410                 415

Asp Glu Cys Ala Met Ala Asn Ser Asn Pro Cys Glu His Ala Gly Lys
            420                 425                 430

Cys Val Asn Thr Asp Gly Ala Phe His Cys Glu Cys Leu Lys Gly Tyr
        435                 440                 445

Ala Gly Pro Arg Cys Glu Met Asp Ile Asn Glu Cys His Ser Asp Pro
    450                 455                 460

Cys Gln Asn Asp Ala Thr Cys Leu Asp Lys Ile Gly Gly Phe Thr Cys
465                 470                 475                 480

Leu Cys Met Pro Gly Phe Lys Gly Val His Cys Glu Leu Glu Ile Asn
                485                 490                 495

Glu Cys Gln Ser Asn Pro Cys Val Asn Asn Gly Gln Cys Val Asp Lys
            500                 505                 510

Val Asn Arg Phe Gln Cys Leu Cys Pro Pro Gly Phe Thr Gly Pro Val
        515                 520                 525

Cys Gln Ile Asp Ile Asp Asp Cys Ser Ser Thr Pro Cys Leu Asn Gly
    530                 535                 540

Ala Lys Cys Ile Asp His Pro Asn Gly Tyr Glu Cys Gln Cys Ala Thr
545                 550                 555                 560

Gly Phe Thr Gly Val Leu Cys Glu Glu Asn Ile Asp Asn Cys Asp Pro
                565                 570                 575

Asp Pro Cys His His Gly Gln Cys Gln Asp Gly Ile Asp Ser Tyr Thr
            580                 585                 590

Cys Ile Cys Asn Pro Gly Tyr Met Gly Ala Ile Cys Ser Asp Gln Ile
        595                 600                 605

Asp Glu Cys Tyr Ser Ser Pro Cys Leu Asn Asp Gly Arg Cys Ile Asp
    610                 615                 620

Leu Val Asn Gly Tyr Gln Cys Asn Cys Gln Pro Gly Thr Ser Gly Val
625                 630                 635                 640

Asn Cys Glu Ile Asn Phe Asp Asp Cys Ala Ser Asn Pro Cys Ile His
                645                 650                 655

Gly Ile Cys Met Asp Gly Ile Asn Arg Tyr Ser Cys Val Cys Ser Pro
            660                 665                 670

Gly Phe Thr Gly Gln Arg Cys Asn Ile Asp Ile Asp Glu Cys Ala Ser
        675                 680                 685
```

Asn Pro Cys Arg Lys Gly Ala Thr Cys Ile Asn Gly Val Asn Gly Phe
690 695 700

Arg Cys Ile Cys Pro Glu Gly Pro His His Pro Ser Cys Tyr Ser Gln
705 710 715 720

Val Asn Glu Cys Leu Ser Asn Pro Cys Ile His Gly Asn Cys Thr Gly
725 730 735

Gly Leu Ser Gly Tyr Lys Cys Leu Cys Asp Ala Gly Trp Val Gly Ile
740 745 750

Asn Cys Glu Val Asp Lys Asn Glu Cys Leu Ser Asn Pro Cys Gln Asn
755 760 765

Gly Gly Thr Cys Asp Asn Leu Val Asn Gly Tyr Arg Cys Thr Cys Lys
770 775 780

Lys Gly Phe Lys Gly Tyr Asn Cys Gln Val Asn Ile Asp Glu Cys Ala
785 790 795 800

Ser Asn Pro Cys Leu Asn Gln Gly Thr Cys Phe Asp Asp Ile Ser Gly
805 810 815

Tyr Thr Cys His Cys Val Leu Pro Tyr Thr Gly Lys Asn Cys Gln Thr
820 825 830

Val Leu Ala Pro Cys Ser Pro Asn Pro Cys Glu Asn Ala Ala Val Cys
835 840 845

Lys Glu Ser Pro Asn Phe Glu Ser Tyr Thr Cys Leu Cys Ala Pro Gly
850 855 860

Trp Gln Gly Gln Arg Cys Thr Ile Asp Ile Asp Glu Cys Ile Ser Lys
865 870 875 880

Pro Cys Met Asn His Gly Leu Cys His Asn Thr Gln Gly Ser Tyr Met
885 890 895

Cys Glu Cys Pro Pro Gly Phe Ser Gly Met Asp Cys Glu Glu Asp Ile
900 905 910

Asp Asp Cys Leu Ala Asn Pro Cys Gln Asn Gly Gly Ser Cys Met Asp
915 920 925

Gly Val Asn Thr Phe Ser Cys Leu Cys Leu Pro Gly Phe Thr Gly Asp
930 935 940

Lys Cys Gln Thr Asp Met Asn Glu Cys Leu Ser Glu Pro Cys Lys Asn
945 950 955 960

Gly Gly Thr Cys Ser Asp Tyr Val Asn Ser Tyr Thr Cys Lys Cys Gln
965 970 975

Ala Gly Phe Asp Gly Val His Cys Glu Asn Asn Ile Asn Glu Cys Thr
980 985 990

Glu Ser Ser Cys Phe Asn Gly Gly Thr Cys Val Asp Gly Ile Asn Ser
995 1000 1005

Phe Ser Cys Leu Cys Pro Val Gly Phe Thr Gly Ser Phe Cys Leu His
1010 1015 1020

Glu Ile Asn Glu Cys Ser Ser His Pro Cys Leu Asn Glu Gly Thr Cys
1025 1030 1035 1040

Val Asp Gly Leu Gly Thr Tyr Arg Cys Ser Cys Pro Leu Gly Tyr Thr

1045 1050 1055

Gly Lys Asn Cys Gln Thr Leu Val Asn Leu Cys Ser Arg Ser Pro Cys
1060 1065 1070

Lys Asn Lys Gly Thr Cys Val Gln Lys Lys Ala Glu Ser Gln Cys Leu
1075 1080 1085

Cys Pro Ser Gly Trp Ala Gly Ala Tyr Cys Asp Val Pro Asn Val Ser
1090 1095 1100

Cys Asp Ile Ala Ala Ser Arg Arg Gly Val Leu Val Glu His Leu Cys
1105 1110 1115 1120

Gln His Ser Gly Val Cys Ile Asn Ala Gly Asn Thr His Tyr Cys Gln
1125 1130 1135

Cys Pro Leu Gly Tyr Thr Gly Ser Tyr Cys Glu Glu Gln Leu Asp Glu
1140 1145 1150

Cys Ala Ser Asn Pro Cys Gln His Gly Ala Thr Cys Ser Asp Phe Ile
1155 1160 1165

Gly Gly Tyr Arg Cys Glu Cys Val Pro Gly Tyr Gln Gly Val Asn Cys
1170 1175 1180

Glu Tyr Glu Val Asp Glu Cys Gln Asn Gln Pro Cys Gln Asn Gly Gly
1185 1190 1195 1200

Thr Cys Ile Asp Leu Val Asn His Phe Lys Cys Ser Cys Pro Pro Gly
1205 1210 1215

Thr Arg Gly Leu Leu Cys Glu Glu Asn Ile Asp Asp Cys Ala Arg Gly
1220 1225 1230

Pro His Cys Leu Asn Gly Gly Gln Cys Met Asp Arg Ile Gly Gly Tyr
1235 1240 1245

Ser Cys Arg Cys Leu Pro Gly Phe Ala Gly Glu Arg Cys Glu Gly Asp
1250 1255 1260

Ile Asn Glu Cys Leu Ser Asn Pro Cys Ser Ser Glu Gly Ser Leu Asp
1265 1270 1275 1280

Cys Ile Gln Leu Thr Asn Asp Tyr Leu Cys Val Cys Arg Ser Ala Phe
1285 1290 1295

Thr Gly Arg His Cys Glu Thr Phe Val Asp Val Cys Pro Gln Met Pro
1300 1305 1310

Cys Leu Asn Gly Gly Thr Cys Ala Val Ala Ser Asn Met Pro Asp Gly
1315 1320 1325

Phe Ile Cys Arg Cys Pro Pro Gly Phe Ser Gly Ala Arg Cys Gln Ser
1330 1335 1340

Ser Cys Gly Gln Val Lys Cys Arg Lys Gly Glu Gln Cys Val His Thr
1345 1350 1355 1360

Ala Ser Gly Pro Arg Cys Phe Cys Pro Ser Pro Arg Asp Cys Glu Ser
1365 1370 1375

Gly Cys Ala Ser Ser Pro Cys Gln His Gly Gly Ser Cys His Pro Gln
1380 1385 1390

Arg Gln Pro Pro Tyr Tyr Ser Cys Gln Cys Ala Pro Pro Phe Ser Gly
1395 1400 1405

Ser Arg Cys Glu Leu Tyr Thr Ala Pro Pro Ser Thr Pro Pro Ala Thr
1410 1415 1420

Cys Leu Ser Gln Tyr Cys Ala Asp Lys Ala Arg Asp Gly Val Cys Asp
1425 1430 1435 1440

Glu Ala Cys Asn Ser His Ala Cys Gln Trp Asp Gly Gly Asp Cys Ser
1445 1450 1455

Leu Thr Met Glu Asn Pro Trp Ala Asn Cys Ser Ser Pro Leu Pro Cys
1460 1465 1470

Trp Asp Tyr Ile Asn Asn Gln Cys Asp Glu Leu Cys Asn Thr Val Glu
1475 1480 1485

Cys Leu Phe Asp Asn Phe Glu Cys Gln Gly Asn Ser Lys Thr Cys Lys
1490 1495 1500

Tyr Asp Lys Tyr Cys Ala Asp His Phe Lys Asp Asn His Cys Asn Gln
1505 1510 1515 1520

Gly Cys Asn Ser Glu Glu Cys Gly Trp Asp Gly Leu Asp Cys Ala Ala
1525 1530 1535

Asp Gln Pro Glu Asn Leu Ala Glu Gly Thr Leu Val Ile Val Val Leu
1540 1545 1550

Met Pro Pro Glu Gln Leu Leu Gln Asp Ala Arg Ser Phe Leu Arg Ala
1555 1560 1565

Leu Gly Thr Leu Leu His Thr Asn Leu Arg Ile Lys Arg Asp Ser Gln
1570 1575 1580

Gly Glu Leu Met Val Tyr Pro Tyr Tyr Gly Glu Lys Ser Ala Ala Met
1585 1590 1595 1600

Lys Lys Gln Arg Met Thr Arg Arg Ser Leu Pro Gly Glu Gln Glu Gln
1605 1610 1615

Glu Val Ala Gly Ser Lys Val Phe Leu Glu Ile Asp Asn Arg Gln Cys
1620 1625 1630

Val Gln Asp Ser Asp His Cys Phe Lys Asn Thr Asp Ala Ala Ala Ala
1635 1640 1645

Leu Leu Ala Ser His Ala Ile Gln Gly Thr Leu Ser Tyr Pro Leu Val
1650 1655 1660

Ser Val Val Ser Glu Ser Leu Thr Pro Glu Arg Thr Gln Leu Leu Tyr
1665 1670 1675 1680

Leu Leu Ala Val Ala Val Val Ile Ile Leu Phe Ile Ile Leu Leu Gly
1685 1690 1695

Val Ile Met Ala Lys Arg Lys Arg Lys His Gly Ser Leu Trp Leu Pro
1700 1705 1710

Glu Gly Phe Thr Leu Arg Arg Asp Ala Ser Asn His Lys Arg Arg Glu
1715 1720 1725

Pro Val Gly Gln Asp Ala Val Gly Leu Lys Asn Leu Ser Val Gln Val
1730 1735 1740

Ser Glu Ala Asn Leu Ile Gly Thr Gly Thr Ser Glu His Trp Val Asp
1745 1750 1755 1760

Asp Glu Gly Pro Gln Pro Lys Lys Val Lys Ala Glu Asp Glu Ala Leu

1765 1770 1775

Leu Ser Glu Glu Asp Asp Pro Ile Asp Arg Arg Pro Trp Thr Gln Gln
1780 1785 1790

His Leu Glu Ala Ala Asp Ile Arg Arg Thr Pro Ser Leu Ala Leu Thr
1795 1800 1805

Pro Pro Gln Ala Glu Gln Glu Val Asp Val Leu Asp Val Asn Val Arg
1810 1815 1820

Gly Pro Asp Gly Cys Thr Pro Leu Met Leu Ala Ser Leu Arg Gly Gly
1825 1830 1835 1840

Ser Ser Asp Leu Ser Asp Glu Asp Glu Asp Ala Glu Asp Ser Ser Ala
1845 1850 1855

Asn Ile Ile Thr Asp Leu Val Tyr Gln Gly Ala Ser Leu Gln Ala Gln
1860 1865 1870

Thr Asp Arg Thr Gly Glu Met Ala Leu His Leu Ala Ala Arg Tyr Ser
1875 1880 1885

Arg Ala Asp Ala Ala Lys Arg Leu Leu Asp Ala Gly Ala Asp Ala Asn
1890 1895 1900

Ala Gln Asp Asn Met Gly Arg Cys Pro Leu His Ala Ala Val Ala Ala
1905 1910 1915 1920

Asp Ala Gln Gly Val Phe Gln Ile Leu Ile Arg Asn Arg Val Thr Asp
1925 1930 1935

Leu Asp Ala Arg Met Asn Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala
1940 1945 1950

Arg Leu Ala Val Glu Gly Met Val Ala Glu Leu Ile Asn Cys Gln Ala
1955 1960 1965

Asp Val Asn Ala Val Asp Asp His Gly Lys Ser Ala Leu His Trp Ala
1970 1975 1980

Ala Ala Val Asn Asn Val Glu Ala Thr Leu Leu Leu Leu Lys Asn Gly
1985 1990 1995 2000

Ala Asn Arg Asp Met Gln Asp Asn Lys Glu Glu Thr Pro Leu Phe Leu
2005 2010 2015

Ala Ala Arg Glu Gly Ser Tyr Glu Ala Ala Lys Ile Leu Leu Asp His
2020 2025 2030

Phe Ala Asn Arg Asp Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp
2035 2040 2045

Val Ala Arg Asp Arg Met His His Asp Ile Val Arg Leu Leu Asp Glu
2050 2055 2060

Tyr Asn Val Thr Pro Ser Pro Pro Gly Thr Val Leu Thr Ser Ala Leu
2065 2070 2075 2080

Ser Pro Val Ile Cys Gly Pro Asn Arg Ser Phe Leu Ser Leu Lys His
2085 2090 2095

Thr Pro Met Gly Lys Lys Ser Arg Arg Pro Ser Ala Lys Ser Thr Met
2100 2105 2110

Pro Thr Ser Leu Pro Asn Leu Ala Lys Glu Ala Lys Asp Ala Lys Gly
2115 2120 2125

Ser Arg Arg Lys Lys Ser Leu Ser Glu Lys Val Gln Leu Ser Glu Ser
2130 2135 2140

Ser Val Thr Leu Ser Pro Val Asp Ser Leu Glu Ser Pro His Thr Tyr
2145 2150 2155 2160

Val Ser Asp Thr Thr Ser Ser Pro Met Ile Thr Ser Pro Gly Ile Leu
2165 2170 2175

Gln Ala Ser Pro Asn Pro Met Leu Ala Thr Ala Ala Pro Pro Ala Pro
2180 2185 2190

Val His Ala Gln His Ala Leu Ser Phe Ser Asn Leu His Glu Met Gln
2195 2200 2205

Pro Leu Ala His Gly Ala Ser Thr Val Leu Pro Ser Val Ser Gln Leu
2210 2215 2220

Leu Ser His His His Ile Val Ser Pro Gly Ser Gly Ser Ala Gly Ser
2225 2230 2235 2240

Leu Ser Arg Leu His Pro Val Pro Val Pro Ala Asp Trp Met Asn Arg
2245 2250 2255

Met Glu Val Asn Glu Thr Gln Tyr Asn Glu Met Phe Gly Met Val Leu
2260 2265 2270

Ala Pro Ala Glu Gly Thr His Pro Gly Ile Ala Pro Gln Ser Arg Pro
2275 2280 2285

Pro Glu Gly Lys His Ile Thr Thr Pro Arg Glu Pro Leu Pro Pro Ile
2290 2295 2300

Val Thr Phe Gln Leu Ile Pro Lys Gly Ser Ile Ala Gln Pro Ala Gly
2305 2310 2315 2320

Ala Pro Gln Pro Gln Ser Thr Cys Pro Pro Ala Val Ala Gly Pro Leu
2325 2330 2335

Pro Thr Met Tyr Gln Ile Pro Glu Met Ala Arg Leu Pro Ser Val Ala
2340 2345 2350

Phe Pro Thr Ala Met Met Pro Gln Gln Asp Gly Gln Val Ala Gln Thr
2355 2360 2365

Ile Leu Pro Ala Tyr His Pro Phe Pro Ala Ser Val Gly Lys Tyr Pro
2370 2375 2380

Thr Pro Pro Ser Gln His Ser Tyr Ala Ser Ser Asn Ala Ala Glu Arg
2385 2390 2395 2400

Thr Pro Ser His Ser Gly His Leu Gln Gly Glu His Pro Tyr Leu Thr
2405 2410 2415

Pro Ser Pro Glu Ser Pro Asp Gln Trp Ser Ser Ser Ser Pro His Ser
2420 2425 2430

Ala Ser Asp Trp Ser Asp Val Thr Thr Ser Pro Thr Pro Gly Gly Ala
2435 2440 2445

Gly Gly Gly Gln Arg Gly Pro Gly Thr His Met Ser Glu Pro Pro His
2450 2455 2460

Asn Asn Met Gln Val Tyr Ala
2465 2470

(2) ANGABEN ZU SEQ ID NO:20:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 2556 Aminosäuren
(B) ART: Aminosäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:20:

```
Met Pro Pro Leu Leu Ala Pro Leu Leu Cys Leu Ala Leu Leu Pro Ala
1               5                  10                  15

Leu Ala Ala Arg Gly Pro Arg Cys Ser Gln Pro Gly Glu Thr Cys Leu
            20                  25                  30

Asn Gly Gly Lys Cys Glu Ala Ala Asn Gly Thr Glu Ala Cys Val Cys
        35                  40                  45

Gly Gly Ala Phe Val Gly Pro Arg Cys Gln Asp Pro Asn Pro Cys Leu
    50                  55                  60

Ser Thr Pro Cys Lys Asn Ala Gly Thr Cys His Val Val Asp Arg Arg
65                  70                  75                  80

Gly Val Ala Asp Tyr Ala Cys Ser Cys Ala Leu Gly Phe Ser Gly Pro
                85                  90                  95

Leu Cys Leu Thr Pro Leu Asp Asn Ala Cys Leu Thr Asn Pro Cys Arg
            100                 105                 110

Asn Gly Gly Thr Cys Asp Leu Leu Thr Leu Thr Glu Tyr Lys Cys Arg
        115                 120                 125

Cys Pro Pro Gly Trp Ser Gly Lys Ser Cys Gln Gln Ala Asp Pro Cys
    130                 135                 140

Ala Ser Asn Pro Cys Ala Asn Gly Gly Gln Cys Leu Pro Phe Glu Ala
145                 150                 155                 160

Ser Tyr Ile Cys His Cys Pro Pro Ser Phe His Gly Pro Thr Cys Arg
                165                 170                 175

Gln Asp Val Asn Glu Cys Gly Gln Lys Pro Arg Leu Cys Arg His Gly
            180                 185                 190

Gly Thr Cys His Asn Glu Val Gly Ser Tyr Arg Cys Val Cys Arg Ala
        195                 200                 205

Thr His Thr Gly Pro Asn Cys Glu Arg Pro Tyr Val Pro Cys Ser Pro
    210                 215                 220

Ser Pro Cys Gln Asn Gly Gly Thr Cys Arg Pro Thr Gly Asp Val Thr
225                 230                 235                 240

His Glu Cys Ala Cys Leu Pro Gly Phe Thr Gly Gln Asn Cys Glu Glu
                245                 250                 255

Asn Ile Asp Asp Cys Pro Gly Asn Asn Cys Lys Asn Gly Gly Ala Cys
            260                 265                 270

Val Asp Gly Val Asn Thr Tyr Asn Cys Pro Cys Pro Pro Glu Trp Thr
        275                 280                 285

Gly Gln Tyr Cys Thr Glu Asp Val Asp Glu Cys Gln Leu Met Pro Asn
```

```
     290                 295                 300

 Ala Cys Gln Asn Gly Gly Thr Cys His Asn Thr His Gly Gly Tyr Asn
 305                 310                 315                 320

 Cys Val Cys Val Asn Gly Trp Thr Gly Glu Asp Cys Ser Glu Asn Ile
                 325                 330                 335

 Asp Asp Cys Ala Ser Ala Ala Cys Phe His Gly Ala Thr Cys His Asp
             340                 345                 350

 Arg Val Ala Ser Phe Tyr Cys Glu Cys Pro His Gly Arg Thr Gly Leu
         355                 360                 365

 Leu Cys His Leu Asn Asp Ala Cys Ile Ser Asn Pro Cys Asn Glu Gly
     370                 375                 380

 Ser Asn Cys Asp Thr Asn Pro Val Asn Gly Lys Ala Ile Cys Thr Cys
 385                 390                 395                 400

 Pro Ser Gly Tyr Thr Gly Pro Ala Cys Ser Gln Asp Val Asp Glu Cys
                 405                 410                 415

 Ser Leu Gly Ala Asn Pro Cys Glu His Ala Gly Lys Cys Ile Asn Thr
             420                 425                 430

 Leu Gly Ser Phe Glu Cys Gln Cys Leu Gln Gly Tyr Thr Gly Pro Arg
         435                 440                 445

 Cys Glu Ile Asp Val Asn Glu Cys Val Ser Asn Pro Cys Gln Asn Asp
     450                 455                 460

 Ala Thr Cys Leu Asp Gln Ile Gly Glu Phe Gln Cys Met Cys Met Pro
 465                 470                 475                 480

 Gly Tyr Glu Gly Val His Cys Glu Val Asn Thr Asp Glu Cys Ala Ser
                 485                 490                 495

 Ser Pro Cys Leu His Asn Gly Arg Cys Leu Asp Lys Ile Asn Glu Phe
             500                 505                 510

 Gln Cys Glu Cys Pro Thr Gly Phe Thr Gly His Leu Cys Gln Tyr Asp
         515                 520                 525

 Val Asp Glu Cys Ala Ser Thr Pro Cys Lys Asn Gly Ala Lys Cys Leu
     530                 535                 540

 Asp Gly Pro Asn Thr Tyr Thr Cys Val Cys Thr Glu Gly Tyr Thr Gly
 545                 550                 555                 560

 Thr His Cys Glu Val Asp Ile Asp Glu Cys Asp Pro Asp Pro Cys His
                 565                 570                 575

 Tyr Gly Ser Cys Lys Asp Gly Val Ala Thr Phe Thr Cys Leu Cys Arg
             580                 585                 590

 Pro Gly Tyr Thr Gly His His Cys Glu Thr Asn Ile Asn Glu Cys Ser
         595                 600                 605

 Ser Gln Pro Cys Arg Leu Arg Gly Thr Cys Gln Asp Pro Asp Asn Ala
     610                 615                 620

 Tyr Leu Cys Phe Cys Leu Lys Gly Thr Thr Gly Pro Asn Cys Glu Ile
 625                 630                 635                 640

 Asn Leu Asp Asp Cys Ala Ser Ser Pro Cys Asp Ser Gly Thr Cys Leu
                 645                 650                 655
```

Asp Lys Ile Asp Gly Tyr Glu Cys Ala Cys Glu Pro Gly Tyr Thr Gly
660 665 670

Ser Met Cys Asn Ser Asn Ile Asp Glu Cys Ala Gly Asn Pro Cys His
675 680 685

Asn Gly Gly Thr Cys Glu Asp Gly Ile Asn Gly Phe Thr Cys Arg Cys
690 695 700

Pro Glu Gly Tyr His Asp Pro Thr Cys Leu Ser Glu Val Asn Glu Cys
705 710 715 720

Asn Ser Asn Pro Cys Val His Gly Ala Cys Arg Asp Ser Leu Asn Gly
725 730 735

Tyr Lys Cys Asp Cys Asp Pro Gly Trp Ser Gly Thr Asn Cys Asp Ile
740 745 750

Asn Asn Asn Glu Cys Glu Ser Asn Pro Cys Val Asn Gly Gly Thr Cys
755 760 765

Lys Asp Met Thr Ser Gly Ile Val Cys Thr Cys Arg Glu Gly Phe Ser
770 775 780

Gly Pro Asn Cys Gln Thr Asn Ile Asn Glu Cys Ala Ser Asn Pro Cys
785 790 795 800

Leu Asn Lys Gly Thr Cys Ile Asp Asp Val Ala Gly Tyr Lys Cys Asn
805 810 815

Cys Leu Leu Pro Tyr Thr Gly Ala Thr Cys Glu Val Val Leu Ala Pro
820 825 830

Cys Ala Pro Ser Pro Cys Arg Asn Gly Gly Glu Cys Arg Gln Ser Glu
835 840 845

Asp Tyr Glu Ser Phe Ser Cys Val Cys Pro Thr Ala Gly Ala Lys Gly
850 855 860

Gln Thr Cys Glu Val Asp Ile Asn Glu Cys Val Leu Ser Pro Cys Arg
865 870 875 880

His Gly Ala Ser Cys Gln Asn Thr His Gly Gly Tyr Arg Cys His Cys
885 890 895

Gln Ala Gly Tyr Ser Gly Arg Asn Cys Glu Thr Asp Ile Asp Asp Cys
900 905 910

Arg Pro Asn Pro Cys His Asn Gly Gly Ser Cys Thr Asp Gly Ile Asn
915 920 925

Thr Ala Phe Cys Asp Cys Leu Pro Gly Phe Arg Gly Thr Phe Cys Glu
930 935 940

Glu Asp Ile Asn Glu Cys Ala Ser Asp Pro Cys Arg Asn Gly Ala Asn
945 950 955 960

Cys Thr Asp Cys Val Asp Ser Tyr Thr Cys Thr Cys Pro Ala Gly Phe
965 970 975

Ser Gly Ile His Cys Glu Asn Asn Thr Pro Asp Cys Thr Glu Ser Ser
980 985 990

Cys Phe Asn Gly Gly Thr Cys Val Asp Gly Ile Asn Ser Phe Thr Cys
995 1000 1005

Leu Cys Pro Pro Gly Phe Thr Gly Ser Tyr Cys Gln His Val Val Asn

```
1010                1015                1020

Glu Cys Asp Ser Arg Pro Cys Leu Leu Gly Gly Thr Cys Gln Asp Gly
1025                1030                1035                1040

Arg Gly Leu His Arg Cys Thr Cys Pro Gln Gly Tyr Thr Gly Pro Asn
                1045                1050                1055

Cys Gln Asn Leu Val His Trp Cys Asp Ser Ser Pro Cys Lys Asn Gly
            1060                1065                1070

Gly Lys Cys Trp Gln Thr His Thr Gln Tyr Arg Cys Glu Cys Pro Ser
        1075                1080                1085

Gly Trp Thr Gly Leu Tyr Cys Asp Val Pro Ser Val Ser Cys Glu Val
    1090                1095                1100

Ala Ala Gln Arg Gln Gly Val Asp Val Ala Arg Leu Cys Gln His Gly
1105                1110                1115                1120

Gly Leu Cys Val Asp Ala Gly Asn Thr His His Cys Arg Cys Gln Ala
                1125                1130                1135

Gly Tyr Thr Gly Ser Tyr Cys Glu Asp Leu Val Asp Glu Cys Ser Pro
            1140                1145                1150

Ser Pro Cys Gln Asn Gly Ala Thr Cys Thr Asp Tyr Leu Gly Gly Tyr
        1155                1160                1165

Ser Cys Lys Cys Val Ala Gly Tyr His Gly Val Asn Cys Ser Glu Glu
    1170                1175                1180

Ile Asp Glu Cys Leu Ser His Pro Cys Gln Asn Gly Gly Thr Cys Leu
1185                1190                1195                1200

Asp Leu Pro Asn Thr Tyr Lys Cys Ser Cys Pro Arg Gly Thr Gln Gly
                1205                1210                1215

Val His Cys Glu Ile Asn Val Asp Asp Cys Asn Pro Pro Val Asp Pro
            1220                1225                1230

Val Ser Arg Ser Pro Lys Cys Phe Asn Asn Gly Thr Cys Val Asp Gln
        1235                1240                1245

Val Gly Gly Tyr Ser Cys Thr Cys Pro Pro Gly Phe Val Gly Glu Arg
    1250                1255                1260

Cys Glu Gly Asp Val Asn Glu Cys Leu Ser Asn Pro Cys Asp Ala Arg
1265                1270                1275                1280

Gly Thr Gln Asn Cys Val Gln Arg Val Asn Asp Phe His Cys Glu Cys
                1285                1290                1295

Arg Ala Gly His Thr Gly Arg Arg Cys Glu Ser Val Ile Asn Gly Cys
            1300                1305                1310

Lys Gly Lys Pro Cys Lys Asn Gly Gly Thr Cys Ala Val Ala Ser Asn
        1315                1320                1325

Thr Ala Arg Gly Phe Ile Cys Lys Cys Pro Ala Gly Phe Glu Gly Ala
    1330                1335                1340

Thr Cys Glu Asn Asp Ala Arg Thr Cys Gly Ser Leu Arg Cys Leu Asn
1345                1350                1355                1360

Gly Gly Thr Cys Ile Ser Gly Pro Arg Ser Pro Thr Cys Leu Cys Leu
                1365                1370                1375
```

Gly Pro Phe Thr Gly Pro Glu Cys Gln Phe Pro Ala Ser Ser Pro Cys
1380 1385 1390

Leu Gly Gly Asn Pro Cys Tyr Asn Gln Gly Thr Cys Glu Pro Thr Ser
1395 1400 1405

Glu Ser Pro Phe Tyr Arg Cys Leu Cys Pro Ala Lys Phe Asn Gly Leu
1410 1415 1420

Leu Cys His Ile Leu Asp Tyr Ser Phe Gly Gly Gly Ala Gly Arg Asp
1425 1430 1435 1440

Ile Pro Pro Pro Leu Ile Glu Glu Ala Cys Glu Leu Pro Glu Cys Gln
1445 1450 1455

Glu Asp Ala Gly Asn Lys Val Cys Ser Leu Gln Cys Asn Asn His Ala
1460 1465 1470

Cys Gly Trp Asp Gly Gly Asp Cys Ser Leu Asn Phe Asn Asp Pro Trp
1475 1480 1485

Lys Asn Cys Thr Gln Ser Leu Gln Cys Trp Lys Tyr Phe Ser Asp Gly
1490 1495 1500

His Cys Asp Ser Gln Cys Asn Ser Ala Gly Cys Leu Phe Asp Gly Phe
1505 1510 1515 1520

Asp Cys Gln Arg Ala Glu Gly Gln Cys Asn Pro Leu Tyr Asp Gln Tyr
1525 1530 1535

Cys Lys Asp His Phe Ser Asp Gly His Cys Asp Gln Gly Cys Asn Ser
1540 1545 1550

Ala Glu Cys Glu Trp Asp Gly Leu Asp Cys Ala Glu His Val Pro Glu
1555 1560 1565

Arg Leu Ala Ala Gly Thr Leu Val Val Val Val Leu Met Pro Pro Glu
1570 1575 1580

Gln Leu Arg Asn Ser Ser Phe His Phe Leu Arg Glu Leu Ser Arg Val
1585 1590 1595 1600

Leu His Thr Asn Val Val Phe Lys Arg Asp Ala His Gly Gln Gln Met
1605 1610 1615

Ile Phe Pro Tyr Tyr Gly Arg Glu Glu Glu Leu Arg Lys His Pro Ile
1620 1625 1630

Lys Arg Ala Ala Glu Gly Trp Ala Ala Pro Asp Ala Leu Leu Gly Gln
1635 1640 1645

Val Lys Ala Ser Leu Leu Pro Gly Gly Ser Glu Gly Gly Arg Arg Arg
1650 1655 1660

Arg Glu Leu Asp Pro Met Asp Val Arg Gly Ser Ile Val Tyr Leu Glu
1665 1670 1675 1680

Ile Asp Asn Arg Gln Cys Val Gln Ala Ser Ser Gln Cys Phe Gln Ser
1685 1690 1695

Ala Thr Asp Val Ala Ala Phe Leu Gly Ala Leu Ala Ser Leu Gly Ser
1700 1705 1710

Leu Asn Ile Pro Tyr Lys Ile Glu Ala Val Gln Ser Glu Thr Val Glu
1715 1720 1725

Pro Pro Pro Pro Ala Gln Leu His Phe Met Tyr Val Ala Ala Ala Ala

1730 1735 1740

Phe Val Leu Leu Phe Phe Val Gly Cys Gly Val Leu Leu Ser Arg Lys
1745 1750 1755 1760

Arg Arg Arg Gln His Gly Gln Leu Trp Phe Pro Glu Gly Phe Lys Val
1765 1770 1775

Ser Glu Ala Ser Lys Lys Lys Arg Arg Glu Glu Leu Gly Glu Asp Ser
1780 1785 1790

Val Gly Leu Lys Pro Leu Lys Asn Ala Ser Asp Gly Ala Leu Met Asp
1795 1800 1805

Asp Asn Gln Asn Glu Trp Gly Asp Glu Asp Leu Glu Thr Lys Lys Phe
1810 1815 1820

Arg Phe Glu Glu Pro Val Val Leu Pro Asp Leu Asp Asp Gln Thr Asp
1825 1830 1835 1840

His Arg Gln Trp Thr Gln Gln His Leu Asp Ala Ala Asp Leu Arg Met
1845 1850 1855

Ser Ala Met Ala Pro Thr Pro Pro Gln Gly Glu Val Asp Ala Asp Cys
1860 1865 1870

Met Asp Val Asn Val Arg Gly Pro Asp Gly Phe Thr Pro Leu Met Ile
1875 1880 1885

Ala Ser Cys Ser Gly Gly Gly Leu Glu Thr Gly Asn Ser Glu Glu Glu
1890 1895 1900

Glu Asp Ala Pro Ala Val Ile Ser Asp Phe Ile Tyr Gln Gly Ala Ser
1905 1910 1915 1920

Leu His Asn Gln Thr Asp Arg Thr Gly Glu Thr Ala Leu His Leu Ala
1925 1930 1935

Ala Arg Tyr Ser Arg Ser Asp Ala Ala Lys Arg Leu Leu Glu Ala Ser
1940 1945 1950

Ala Asp Ala Asn Ile Gln Asp Asn Met Gly Arg Thr Pro Leu His Ala
1955 1960 1965

Ala Val Ser Ala Asp Ala Gln Gly Val Phe Gln Ile Leu Ile Arg Asn
1970 1975 1980

Arg Ala Thr Asp Leu Asp Ala Arg Met His Asp Gly Thr Thr Pro Leu
1985 1990 1995 2000

Ile Leu Ala Ala Arg Leu Ala Val Glu Gly Met Leu Glu Asp Leu Ile
2005 2010 2015

Asn Ser His Ala Asp Val Asn Ala Val Asp Asp Leu Gly Lys Ser Ala
2020 2025 2030

Leu His Trp Ala Ala Ala Val Asn Asn Val Asp Ala Ala Val Val Leu
2035 2040 2045

Leu Lys Asn Gly Ala Asn Lys Asp Met Gln Asn Asn Arg Glu Glu Thr
2050 2055 2060

Pro Leu Phe Leu Ala Ala Arg Glu Gly Ser Tyr Glu Thr Ala Lys Val
2065 2070 2075 2080

Leu Leu Asp His Phe Ala Asn Arg Asp Ile Thr Asp His Met Asp Arg
2085 2090 2095

Leu Pro Arg Asp Ile Ala Gln Glu Arg Met His His Asp Ile Val Arg
2100 2105 2110

Leu Leu Asp Glu Tyr Asn Leu Val Arg Ser Pro Gln Leu His Gly Ala
2115 2120 2125

Pro Leu Gly Gly Thr Pro Thr Leu Ser Pro Pro Leu Cys Ser Pro Asn
2130 2135 2140

Gly Tyr Leu Gly Ser Leu Lys Pro Gly Val Gln Gly Lys Lys Val Arg
2145 2150 2155 2160

Lys Pro Ser Ser Lys Gly Leu Ala Cys Gly Ser Lys Glu Ala Lys Asp
2165 2170 2175

Leu Lys Ala Arg Arg Lys Lys Ser Gln Asp Gly Lys Gly Cys Leu Leu
2180 2185 2190

Asp Ser Ser Gly Met Leu Ser Pro Val Asp Ser Leu Glu Ser Pro His
2195 2200 2205

Gly Tyr Leu Ser Asp Val Ala Ser Pro Pro Leu Leu Pro Ser Pro Phe
2210 2215 2220

Gln Gln Ser Pro Ser Val Pro Leu Asn His Leu Pro Gly Met Pro Asp
2225 2230 2235 2240

Thr His Leu Gly Ile Gly His Leu Asn Val Ala Ala Lys Pro Glu Met
2245 2250 2255

Ala Ala Leu Gly Gly Gly Gly Arg Leu Ala Phe Glu Thr Gly Pro Pro
2260 2265 2270

Arg Leu Ser His Leu Pro Val Ala Ser Gly Thr Ser Thr Val Leu Gly
2275 2280 2285

Ser Ser Ser Gly Gly Ala Leu Asn Phe Thr Val Gly Gly Ser Thr Ser
2290 2295 2300

Leu Asn Gly Gln Cys Glu Trp Leu Ser Arg Leu Gln Ser Gly Met Val
2305 2310 2315 2320

Pro Asn Gln Tyr Asn Pro Leu Arg Gly Ser Val Ala Pro Gly Pro Leu
2325 2330 2335

Ser Thr Gln Ala Pro Ser Leu Gln His Gly Met Val Gly Pro Leu His
2340 2345 2350

Ser Ser Leu Ala Ala Ser Ala Leu Ser Gln Met Met Ser Tyr Gln Gly
2355 2360 2365

Leu Pro Ser Thr Arg Leu Ala Thr Gln Pro His Leu Val Gln Thr Gln
2370 2375 2380

Gln Val Gln Pro Gln Asn Leu Gln Met Gln Gln Gln Asn Leu Gln Pro
2385 2390 2395 2400

Ala Asn Ile Gln Gln Gln Gln Ser Leu Gln Pro Pro Pro Pro Pro Pro
2405 2410 2415

Gln Pro His Leu Gly Val Ser Ser Ala Ala Ser Gly His Leu Gly Arg
2420 2425 2430

Ser Phe Leu Ser Gly Glu Pro Ser Gln Ala Asp Val Gln Pro Leu Gly
2435 2440 2445

Pro Ser Ser Leu Ala Val His Thr Ile Leu Pro Gln Glu Ser Pro Ala

```
      2450                2455                2460

Leu Pro Thr Ser Leu Pro Ser Ser Leu Val Pro Pro Val Thr Ala Ala
2465                2470                2475                2480

Gln Phe Leu Thr Pro Pro Ser Gln His Ser Tyr Ser Ser Pro Val Glu
                    2485                2490                2495

Asn Thr Pro Ser His Gln Leu Gln Val Pro Glu His Pro Phe Leu Thr
               2500                2505                2510

Pro Ser Pro Glu Ser Pro Asp Gln Trp Ser Ser Ser Ser Pro His Ser
          2515                2520                2525

Asn Val Ser Asp Trp Ser Glu Gly Val Ser Ser Pro Pro Thr Ser Met
      2530                2535                2540

Gln Ser Gln Ile Ala Arg Ile Pro Glu Ala Phe Lys
2545                2550                2555
```

(2) ANGABEN ZU SEQ ID NO:21:

(i) SEQUENZKENNZEICHEN:
- (A) LÄNGE: 9723 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: unbekannt

(ii) MOLEKULARTYP: cDNA

(ix) MERKMAL:
- (A) NAME/SCHLÜSSEL: CDS
- (B) LAGE: 10..7419

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:21:

```
GGAATTCCG CCC GCC CTG CGC CCC GCT CTG CTG TGG GCG CTG CTG GCG          48
         Pro Ala Leu Arg Pro Ala Leu Leu Trp Ala Leu Leu Ala
           1               5                  10

CTC TGG CTG TGC TGC GCG GCC CCC GCG CAT GCA TTG CAG TGT CGA GAT        96
Leu Trp Leu Cys Cys Ala Ala Pro Ala His Ala Leu Gln Cys Arg Asp
     15                  20                  25

GGC TAT GAA CCC TGT GTA AAT GAA GGA ATG TGT GTT ACC TAC CAC AAT       144
Gly Tyr Glu Pro Cys Val Asn Glu Gly Met Cys Val Thr Tyr His Asn
 30                  35                  40                  45

GGC ACA GGA TAC TGC AAA TGT CCA GAA GGC TTC TTG GGG GAA TAT TGT       192
Gly Thr Gly Tyr Cys Lys Cys Pro Glu Gly Phe Leu Gly Glu Tyr Cys
                 50                  55                  60

CAA CAT CGA GAC CCC TGT GAG AAG AAC CGC TGC CAG AAT GGT GGG ACT       240
Gln His Arg Asp Pro Cys Glu Lys Asn Arg Cys Gln Asn Gly Gly Thr
           65                  70                  75

TGT GTG GCC CAG GCC ATG CTG GGG AAA GCC ACG TGC CGA TGT GCC TCA       288
Cys Val Ala Gln Ala Met Leu Gly Lys Ala Thr Cys Arg Cys Ala Ser
         80                  85                  90

GGG TTT ACA GGA GAG GAC TGC CAG TAC TCA ACA TCT CAT CCA TGC TTT       336
Gly Phe Thr Gly Glu Asp Cys Gln Tyr Ser Thr Ser His Pro Cys Phe
     95                  100                 105

GTG TCT CGA CCC TGC CTG AAT GGC GGC ACA TGC CAT ATG CTC AGC CGG       384
```

```
Val Ser Arg Pro Cys Leu Asn Gly Gly Thr Cys His Met Leu Ser Arg
110                 115                 120                 125

GAT ACC TAT GAG TGC ACC TGT CAA GTC GGG TTT ACA GGT AAG GAG TGC      432
Asp Thr Tyr Glu Cys Thr Cys Gln Val Gly Phe Thr Gly Lys Glu Cys
                130                 135                 140

CAA TGG ACG GAT GCC TGC CTG TCT CAT CCC TGT GCA AAT GGA AGT ACC      480
Gln Trp Thr Asp Ala Cys Leu Ser His Pro Cys Ala Asn Gly Ser Thr
            145                 150                 155

TGT ACC ACT GTG GCC AAC CAG TTC TCC TGC AAA TGC CTC ACA GGC TTC      528
Cys Thr Thr Val Ala Asn Gln Phe Ser Cys Lys Cys Leu Thr Gly Phe
        160                 165                 170

ACA GGG CAG AAA TGT GAG ACT GAT GTC AAT GAG TGT GAC ATT CCA GGA      576
Thr Gly Gln Lys Cys Glu Thr Asp Val Asn Glu Cys Asp Ile Pro Gly
    175                 180                 185

CAC TGC CAG CAT GGT GGC ACC TGC CTC AAC CTG CCT GGT TCC TAC CAG      624
His Cys Gln His Gly Gly Thr Cys Leu Asn Leu Pro Gly Ser Tyr Gln
190                 195                 200                 205

TGC CAG TGC CCT CAG GGC TTC ACA GGC CAG TAC TGT GAC AGC CTG TAT      672
Cys Gln Cys Pro Gln Gly Phe Thr Gly Gln Tyr Cys Asp Ser Leu Tyr
                210                 215                 220

GTG CCC TGT GCA CCC TCA CCT TGT GTC AAT GGA GGC ACC TGT CGG CAG      720
Val Pro Cys Ala Pro Ser Pro Cys Val Asn Gly Gly Thr Cys Arg Gln
            225                 230                 235

ACT GGT GAC TTC ACT TTT GAG TGC AAC TGC CTT CCA GGT TTT GAA GGG      768
Thr Gly Asp Phe Thr Phe Glu Cys Asn Cys Leu Pro Gly Phe Glu Gly
        240                 245                 250

AGC ACC TGT GAG AGG AAT ATT GAT GAC TGC CCT AAC CAC AGG TGT CAG      816
Ser Thr Cys Glu Arg Asn Ile Asp Asp Cys Pro Asn His Arg Cys Gln
    255                 260                 265

AAT GGA GGG GTT TGT GTG GAT GGG GTC AAC ACT TAC AAC TGC CGC TGT      864
Asn Gly Gly Val Cys Val Asp Gly Val Asn Thr Tyr Asn Cys Arg Cys
270                 275                 280                 285

CCC CCA CAA TGG ACA GGA CAG TTC TGC ACA GAG GAT GTG GAT GAA TGC      912
Pro Pro Gln Trp Thr Gly Gln Phe Cys Thr Glu Asp Val Asp Glu Cys
                290                 295                 300

CTG CTG CAG CCC AAT GCC TGT CAA AAT GGG GGC ACC TGT GCC AAC CGC      960
Leu Leu Gln Pro Asn Ala Cys Gln Asn Gly Gly Thr Cys Ala Asn Arg
            305                 310                 315

AAT GGA GGC TAT GGC TGT GTA TGT GTC AAC GGC TGG AGT GGA GAT GAC     1008
Asn Gly Gly Tyr Gly Cys Val Cys Val Asn Gly Trp Ser Gly Asp Asp
        320                 325                 330

TGC AGT GAG AAC ATT GAT GAT TGT GCC TTC GCC TCC TGT ACT CCA GGC     1056
Cys Ser Glu Asn Ile Asp Asp Cys Ala Phe Ala Ser Cys Thr Pro Gly
    335                 340                 345

TCC ACC TGC ATC GAC CGT GTG GCC TCC TTC TCT TGC ATG TGC CCA GAG     1104
Ser Thr Cys Ile Asp Arg Val Ala Ser Phe Ser Cys Met Cys Pro Glu
350                 355                 360                 365

GGG AAG GCA GGT CTC CTG TGT CAT CTG GAT GAT GCA TGC ATC AGC AAT     1152
Gly Lys Ala Gly Leu Leu Cys His Leu Asp Asp Ala Cys Ile Ser Asn
                370                 375                 380
```

```
CCT TGC CAC AAG GGG GCA CTG TGT GAC ACC AAC CCC CTA AAT GGG CAA      1200
Pro Cys His Lys Gly Ala Leu Cys Asp Thr Asn Pro Leu Asn Gly Gln
            385                 390                 395

TAT ATT TGC ACC TGC CCA CAA GGC TAC AAA GGG GCT GAC TGC ACA GAA      1248
Tyr Ile Cys Thr Cys Pro Gln Gly Tyr Lys Gly Ala Asp Cys Thr Glu
        400                 405                 410

GAT GTG GAT GAA TGT GCC ATG GCC AAT AGC AAT CCT TGT GAG CAT GCA      1296
Asp Val Asp Glu Cys Ala Met Ala Asn Ser Asn Pro Cys Glu His Ala
    415                 420                 425

GGA AAA TGT GTG AAC ACG GAT GGC GCC TTC CAC TGT GAG TGT CTG AAG      1344
Gly Lys Cys Val Asn Thr Asp Gly Ala Phe His Cys Glu Cys Leu Lys
430                 435                 440                 445

GGT TAT GCA GGA CCT CGT TGT GAG ATG GAC ATC AAT GAG TGC CAT TCA      1392
Gly Tyr Ala Gly Pro Arg Cys Glu Met Asp Ile Asn Glu Cys His Ser
                450                 455                 460

GAC CCC TGC CAG AAT GAT GCT ACC TGT CTG GAT AAG ATT GGA GGC TTC      1440
Asp Pro Cys Gln Asn Asp Ala Thr Cys Leu Asp Lys Ile Gly Gly Phe
            465                 470                 475

ACA TGT CTG TGC ATG CCA GGT TTC AAA GGT GTG CAT TGT GAA TTA GAA      1488
Thr Cys Leu Cys Met Pro Gly Phe Lys Gly Val His Cys Glu Leu Glu
        480                 485                 490

ATA AAT GAA TGT CAG AGC AAC CCT TGT GTG AAC AAT GGG CAG TGT GTG      1536
Ile Asn Glu Cys Gln Ser Asn Pro Cys Val Asn Asn Gly Gln Cys Val
    495                 500                 505

GAT AAA GTC AAT CGT TTC CAG TGC CTG TGT CCT CCT GGT TTC ACT GGG      1584
Asp Lys Val Asn Arg Phe Gln Cys Leu Cys Pro Pro Gly Phe Thr Gly
510                 515                 520                 525

CCA GTT TGC CAG ATT GAT ATT GAT GAC TGT TCC AGT ACT CCG TGT CTG      1632
Pro Val Cys Gln Ile Asp Ile Asp Asp Cys Ser Ser Thr Pro Cys Leu
                530                 535                 540

AAT GGG GCA AAG TGT ATC GAT CAC CCG AAT GGC TAT GAA TGC CAG TGT      1680
Asn Gly Ala Lys Cys Ile Asp His Pro Asn Gly Tyr Glu Cys Gln Cys
            545                 550                 555

GCC ACA GGT TTC ACT GGT GTG TTG TGT GAG GAG AAC ATT GAC AAC TGT      1728
Ala Thr Gly Phe Thr Gly Val Leu Cys Glu Glu Asn Ile Asp Asn Cys
        560                 565                 570

GAC CCC GAT CCT TGC CAC CAT GGT CAG TGT CAG GAT GGT ATT GAT TCC      1776
Asp Pro Asp Pro Cys His His Gly Gln Cys Gln Asp Gly Ile Asp Ser
    575                 580                 585

TAC ACC TGC ATC TGC AAT CCC GGG TAC ATG GGC GCC ATC TGC AGT GAC      1824
Tyr Thr Cys Ile Cys Asn Pro Gly Tyr Met Gly Ala Ile Cys Ser Asp
590                 595                 600                 605

CAG ATT GAT GAA TGT TAC AGC AGC CCT TGC CTG AAC GAT GGT CGC TGC      1872
Gln Ile Asp Glu Cys Tyr Ser Ser Pro Cys Leu Asn Asp Gly Arg Cys
                610                 615                 620

ATT GAC CTG GTC AAT GGC TAC CAG TGC AAC TGC CAG CCA GGC ACG TCA      1920
Ile Asp Leu Val Asn Gly Tyr Gln Cys Asn Cys Gln Pro Gly Thr Ser
            625                 630                 635

GGG GTT AAT TGT GAA ATT AAT TTT GAT GAC TGT GCA AGT AAC CCT TGT      1968
Gly Val Asn Cys Glu Ile Asn Phe Asp Asp Cys Ala Ser Asn Pro Cys
        640                 645                 650
```

```
ATC CAT GGA ATC TGT ATG GAT GGC ATT AAT CGC TAC AGT TGT GTC TGC      2016
Ile His Gly Ile Cys Met Asp Gly Ile Asn Arg Tyr Ser Cys Val Cys
    655                 660                 665

TCA CCA GGA TTC ACA GGG CAG AGA TGT AAC ATT GAC ATT GAT GAG TGT      2064
Ser Pro Gly Phe Thr Gly Gln Arg Cys Asn Ile Asp Ile Asp Glu Cys
670                 675                 680                 685

GCC TCC AAT CCC TGT CGC AAG GGT GCA ACA TGT ATC AAC GGT GTG AAT      2112
Ala Ser Asn Pro Cys Arg Lys Gly Ala Thr Cys Ile Asn Gly Val Asn
                690                 695                 700

GGT TTC CGC TGT ATA TGC CCC GAG GGA CCC CAT CAC CCC AGC TGC TAC      2160
Gly Phe Arg Cys Ile Cys Pro Glu Gly Pro His His Pro Ser Cys Tyr
            705                 710                 715

TCA CAG GTG AAC GAA TGC CTG AGC AAT CCC TGC ATC CAT GGA AAC TGT      2208
Ser Gln Val Asn Glu Cys Leu Ser Asn Pro Cys Ile His Gly Asn Cys
        720                 725                 730

ACT GGA GGT CTC AGT GGA TAT AAG TGT CTC TGT GAT GCA GGC TGG GTT      2256
Thr Gly Gly Leu Ser Gly Tyr Lys Cys Leu Cys Asp Ala Gly Trp Val
    735                 740                 745

GGC ATC AAC TGT GAA GTG GAC AAA AAT GAA TGC CTT TCG AAT CCA TGC      2304
Gly Ile Asn Cys Glu Val Asp Lys Asn Glu Cys Leu Ser Asn Pro Cys
750                 755                 760                 765

CAG AAT GGA GGA ACT TGT GAC AAT CTG GTG AAT GGA TAC AGG TGT ACT      2352
Gln Asn Gly Gly Thr Cys Asp Asn Leu Val Asn Gly Tyr Arg Cys Thr
                770                 775                 780

TGC AAG AAG GGC TTT AAA GGC TAT AAC TGC CAG GTG AAT ATT GAT GAA      2400
Cys Lys Lys Gly Phe Lys Gly Tyr Asn Cys Gln Val Asn Ile Asp Glu
            785                 790                 795

TGT GCC TCA AAT CCA TGC CTG AAC CAA GGA ACC TGC TTT GAT GAC ATA      2448
Cys Ala Ser Asn Pro Cys Leu Asn Gln Gly Thr Cys Phe Asp Asp Ile
        800                 805                 810

AGT GGC TAC ACT TGC CAC TGT GTG CTG CCA TAC ACA GGC AAG AAT TGT      2496
Ser Gly Tyr Thr Cys His Cys Val Leu Pro Tyr Thr Gly Lys Asn Cys
    815                 820                 825

CAG ACA GTA TTG GCT CCC TGT TCC CCA AAC CCT TGT GAG AAT GCT GCT      2544
Gln Thr Val Leu Ala Pro Cys Ser Pro Asn Pro Cys Glu Asn Ala Ala
830                 835                 840                 845

GTT TGC AAA GAG TCA CCA AAT TTT GAG AGT TAT ACT TGC TTG TGT GCT      2592
Val Cys Lys Glu Ser Pro Asn Phe Glu Ser Tyr Thr Cys Leu Cys Ala
                850                 855                 860

CCT GGC TGG CAA GGT CAG CGG TGT ACC ATT GAC ATT GAC GAG TGT ATC      2640
Pro Gly Trp Gln Gly Gln Arg Cys Thr Ile Asp Ile Asp Glu Cys Ile
            865                 870                 875

TCC AAG CCC TGC ATG AAC CAT GGT CTC TGC CAT AAC ACC CAG GGC AGC      2688
Ser Lys Pro Cys Met Asn His Gly Leu Cys His Asn Thr Gln Gly Ser
        880                 885                 890

TAC ATG TGT GAA TGT CCA CCA GGC TTC AGT GGT ATG GAC TGT GAG GAG      2736
Tyr Met Cys Glu Cys Pro Pro Gly Phe Ser Gly Met Asp Cys Glu Glu
    895                 900                 905

GAC ATT GAT GAC TGC CTT GCC AAT CCT TGC CAG AAT GGA GGT TCC TGT      2784
Asp Ile Asp Asp Cys Leu Ala Asn Pro Cys Gln Asn Gly Gly Ser Cys
910                 915                 920                 925
```

```
ATG GAT GGA GTG AAT ACT TTC TCC TGC CTC TGC CTT CCG GGT TTC ACT      2832
Met Asp Gly Val Asn Thr Phe Ser Cys Leu Cys Leu Pro Gly Phe Thr
                930                 935                 940

GGG GAT AAG TGC CAG ACA GAC ATG AAT GAG TGT CTG AGT GAA CCC TGT      2880
Gly Asp Lys Cys Gln Thr Asp Met Asn Glu Cys Leu Ser Glu Pro Cys
            945                 950                 955

AAG AAT GGA GGG ACC TGC TCT GAC TAC GTC AAC AGT TAC ACT TGC AAG      2928
Lys Asn Gly Gly Thr Cys Ser Asp Tyr Val Asn Ser Tyr Thr Cys Lys
        960                 965                 970

TGC CAG GCA GGA TTT GAT GGA GTC CAT TGT GAG AAC AAC ATC AAT GAG      2976
Cys Gln Ala Gly Phe Asp Gly Val His Cys Glu Asn Asn Ile Asn Glu
    975                 980                 985

TGC ACT GAG AGC TCC TGT TTC AAT GGT GGC ACA TGT GTT GAT GGG ATT      3024
Cys Thr Glu Ser Ser Cys Phe Asn Gly Gly Thr Cys Val Asp Gly Ile
990                 995                 1000                1005

AAC TCC TTC TCT TGC TTG TGC CCT GTG GGT TTC ACT GGA TCC TTC TGC      3072
Asn Ser Phe Ser Cys Leu Cys Pro Val Gly Phe Thr Gly Ser Phe Cys
                1010                1015                1020

CTC CAT GAG ATC AAT GAA TGC AGC TCT CAT CCA TGC CTG AAT GAG GGA      3120
Leu His Glu Ile Asn Glu Cys Ser Ser His Pro Cys Leu Asn Glu Gly
            1025                1030                1035

ACG TGT GTT GAT GGC CTG GGT ACC TAC CGC TGC AGC TGC CCC CTG GGC      3168
Thr Cys Val Asp Gly Leu Gly Thr Tyr Arg Cys Ser Cys Pro Leu Gly
        1040                1045                1050

TAC ACT GGG AAA AAC TGT CAG ACC CTG GTG AAT CTC TGC AGT CGG TCT      3216
Tyr Thr Gly Lys Asn Cys Gln Thr Leu Val Asn Leu Cys Ser Arg Ser
    1055                1060                1065

CCA TGT AAA AAC AAA GGT ACT TGT GTT CAG AAA AAA GCA GAG TCC CAG      3264
Pro Cys Lys Asn Lys Gly Thr Cys Val Gln Lys Lys Ala Glu Ser Gln
1070                1075                1080                1085

TGC CTA TGT CCA TCT GGA TGG GCT GGT GCC TAT TGT GAC GTG CCC AAT      3312
Cys Leu Cys Pro Ser Gly Trp Ala Gly Ala Tyr Cys Asp Val Pro Asn
                1090                1095                1100

GTC TCT TGT GAC ATA GCA GCC TCC AGG AGA GGT GTG CTT GTT GAA CAC      3360
Val Ser Cys Asp Ile Ala Ala Ser Arg Arg Gly Val Leu Val Glu His
            1105                1110                1115

TTG TGC CAG CAC TCA GGT GTC TGC ATC AAT GCT GGC AAC ACG CAT TAC      3408
Leu Cys Gln His Ser Gly Val Cys Ile Asn Ala Gly Asn Thr His Tyr
        1120                1125                1130

TGT CAG TGC CCC CTG GGC TAT ACT GGG AGC TAC TGT GAG GAG CAA CTC      3456
Cys Gln Cys Pro Leu Gly Tyr Thr Gly Ser Tyr Cys Glu Glu Gln Leu
    1135                1140                1145

GAT GAG TGT GCG TCC AAC CCC TGC CAG CAC GGG GCA ACA TGC AGT GAC      3504
Asp Glu Cys Ala Ser Asn Pro Cys Gln His Gly Ala Thr Cys Ser Asp
1150                1155                1160                1165

TTC ATT GGT GGA TAC AGA TGC GAG TGT GTC CCA GGC TAT CAG GGT GTC      3552
Phe Ile Gly Gly Tyr Arg Cys Glu Cys Val Pro Gly Tyr Gln Gly Val
                1170                1175                1180

AAC TGT GAG TAT GAA GTG GAT GAG TGC CAG AAT CAG CCC TGC CAG AAT      3600
Asn Cys Glu Tyr Glu Val Asp Glu Cys Gln Asn Gln Pro Cys Gln Asn
            1185                1190                1195
```

```
GGA GGC ACC TGT ATT GAC CTT GTG AAC CAT TTC AAG TGC TCT TGC CCA      3648
Gly Gly Thr Cys Ile Asp Leu Val Asn His Phe Lys Cys Ser Cys Pro
        1200                1205                1210

CCA GGC ACT CGG GGC CTA CTC TGT GAA GAG AAC ATT GAT GAC TGT GCC      3696
Pro Gly Thr Arg Gly Leu Leu Cys Glu Glu Asn Ile Asp Asp Cys Ala
    1215                1220                1225

CGG GGT CCC CAT TGC CTT AAT GGT GGT CAG TGC ATG GAT AGG ATT GGA      3744
Arg Gly Pro His Cys Leu Asn Gly Gly Gln Cys Met Asp Arg Ile Gly
1230                1235                1240                1245

GGC TAC AGT TGT CGC TGC TTG CCT GGC TTT GCT GGG GAG CGT TGT GAG      3792
Gly Tyr Ser Cys Arg Cys Leu Pro Gly Phe Ala Gly Glu Arg Cys Glu
                1250                1255                1260

GGA GAC ATC AAC GAG TGC CTC TCC AAC CCC TGC AGC TCT GAG GGC AGC      3840
Gly Asp Ile Asn Glu Cys Leu Ser Asn Pro Cys Ser Ser Glu Gly Ser
            1265                1270                1275

CTG GAC TGT ATA CAG CTC ACC AAT GAC TAC CTG TGT GTT TGC CGT AGT      3888
Leu Asp Cys Ile Gln Leu Thr Asn Asp Tyr Leu Cys Val Cys Arg Ser
        1280                1285                1290

GCC TTT ACT GGC CGG CAC TGT GAA ACC TTC GTC GAT GTG TGT CCC CAG      3936
Ala Phe Thr Gly Arg His Cys Glu Thr Phe Val Asp Val Cys Pro Gln
    1295                1300                1305

ATG CCC TGC CTG AAT GGA GGG ACT TGT GCT GTG GCC AGT AAC ATG CCT      3984
Met Pro Cys Leu Asn Gly Gly Thr Cys Ala Val Ala Ser Asn Met Pro
1310                1315                1320                1325

GAT GGT TTC ATT TGC CGT TGT CCC CCG GGA TTT TCC GGG GCA AGG TGC      4032
Asp Gly Phe Ile Cys Arg Cys Pro Pro Gly Phe Ser Gly Ala Arg Cys
                1330                1335                1340

CAG AGC AGC TGT GGA CAA GTG AAA TGT AGG AAG GGG GAG CAG TGT GTG      4080
Gln Ser Ser Cys Gly Gln Val Lys Cys Arg Lys Gly Glu Gln Cys Val
            1345                1350                1355

CAC ACC GCC TCT GGA CCC CGC TGC TTC TGC CCC AGT CCC CGG GAC TGC      4128
His Thr Ala Ser Gly Pro Arg Cys Phe Cys Pro Ser Pro Arg Asp Cys
        1360                1365                1370

GAG TCA GGC TGT GCC AGT AGC CCC TGC CAG CAC GGG GGC AGC TGC CAC      4176
Glu Ser Gly Cys Ala Ser Ser Pro Cys Gln His Gly Gly Ser Cys His
    1375                1380                1385

CCT CAG CGC CAG CCT CCT TAT TAC TCC TGC CAG TGT GCC CCA CCA TTC      4224
Pro Gln Arg Gln Pro Pro Tyr Tyr Ser Cys Gln Cys Ala Pro Pro Phe
1390                1395                1400                1405

TCG GGT AGC CGC TGT GAA CTC TAC ACG GCA CCC CCC AGC ACC CCT CCT      4272
Ser Gly Ser Arg Cys Glu Leu Tyr Thr Ala Pro Pro Ser Thr Pro Pro
                1410                1415                1420

GCC ACC TGT CTG AGC CAG TAT TGT GCC GAC AAA GCT CGG GAT GGC GTC      4320
Ala Thr Cys Leu Ser Gln Tyr Cys Ala Asp Lys Ala Arg Asp Gly Val
            1425                1430                1435

TGT GAT GAG GCC TGC AAC AGC CAT GCC TGC CAG TGG GAT GGG GGT GAC      4368
Cys Asp Glu Ala Cys Asn Ser His Ala Cys Gln Trp Asp Gly Gly Asp
        1440                1445                1450

TGT TCT CTC ACC ATG GAG AAC CCC TGG GCC AAC TGC TCC TCC CCA CTT      4416
Cys Ser Leu Thr Met Glu Asn Pro Trp Ala Asn Cys Ser Ser Pro Leu
    1455                1460                1465
```

```
CCC TGC TGG GAT TAT ATC AAC AAC CAG TGT GAT GAG CTG TGC AAC ACG      4464
Pro Cys Trp Asp Tyr Ile Asn Asn Gln Cys Asp Glu Leu Cys Asn Thr
1470                1475                1480                1485

GTC GAG TGC CTG TTT GAC AAC TTT GAA TGC CAG GGG AAC AGC AAG ACA      4512
Val Glu Cys Leu Phe Asp Asn Phe Glu Cys Gln Gly Asn Ser Lys Thr
                1490                1495                1500

TGC AAG TAT GAC AAA TAC TGT GCA GAC CAC TTC AAA GAC AAC CAC TGT      4560
Cys Lys Tyr Asp Lys Tyr Cys Ala Asp His Phe Lys Asp Asn His Cys
            1505                1510                1515

AAC CAG GGG TGC AAC AGT GAG GAG TGT GGT TGG GAT GGG CTG GAC TGT      4608
Asn Gln Gly Cys Asn Ser Glu Glu Cys Gly Trp Asp Gly Leu Asp Cys
        1520                1525                1530

GCT GCT GAC CAA CCT GAG AAC CTG GCA GAA GGT ACC CTG GTT ATT GTG      4656
Ala Ala Asp Gln Pro Glu Asn Leu Ala Glu Gly Thr Leu Val Ile Val
    1535                1540                1545

GTA TTG ATG CCA CCT GAA CAA CTG CTC CAG GAT GCT CGC AGC TTC TTG      4704
Val Leu Met Pro Pro Glu Gln Leu Leu Gln Asp Ala Arg Ser Phe Leu
1550                1555                1560                1565

CGG GCA CTG GGT ACC CTG CTC CAC ACC AAC CTG CGC ATT AAG CGG GAC      4752
Arg Ala Leu Gly Thr Leu Leu His Thr Asn Leu Arg Ile Lys Arg Asp
                1570                1575                1580

TCC CAG GGG GAA CTC ATG GTG TAC CCC TAT TAT GGT GAG AAG TCA GCT      4800
Ser Gln Gly Glu Leu Met Val Tyr Pro Tyr Tyr Gly Glu Lys Ser Ala
            1585                1590                1595

GCT ATG AAG AAA CAG AGG ATG ACA CGC AGA TCC CTT CCT GGT GAA CAA      4848
Ala Met Lys Lys Gln Arg Met Thr Arg Arg Ser Leu Pro Gly Glu Gln
        1600                1605                1610

GAA CAG GAG GTG GCT GGC TCT AAA GTC TTT CTG GAA ATT GAC AAC CGC      4896
Glu Gln Glu Val Ala Gly Ser Lys Val Phe Leu Glu Ile Asp Asn Arg
    1615                1620                1625

CAG TGT GTT CAA GAC TCA GAC CAC TGC TTC AAG AAC ACG GAT GCA GCA      4944
Gln Cys Val Gln Asp Ser Asp His Cys Phe Lys Asn Thr Asp Ala Ala
1630                1635                1640                1645

GCA GCT CTC CTG GCC TCT CAC GCC ATA CAG GGG ACC CTG TCA TAC CCT      4992
Ala Ala Leu Leu Ala Ser His Ala Ile Gln Gly Thr Leu Ser Tyr Pro
                1650                1655                1660

CTT GTG TCT GTC GTC AGT GAA TCC CTG ACT CCA GAA CGC ACT CAG CTC      5040
Leu Val Ser Val Val Ser Glu Ser Leu Thr Pro Glu Arg Thr Gln Leu
            1665                1670                1675

CTC TAT CTC CTT GCT GTT GCT GTT GTC ATC ATT CTG TTT ATT ATT CTG      5088
Leu Tyr Leu Leu Ala Val Ala Val Val Ile Ile Leu Phe Ile Ile Leu
        1680                1685                1690

CTG GGG GTA ATC ATG GCA AAA CGA AAG CGT AAG CAT GGC TCT CTC TGG      5136
Leu Gly Val Ile Met Ala Lys Arg Lys Arg Lys His Gly Ser Leu Trp
    1695                1700                1705

CTG CCT GAA GGT TTC ACT CTT CGC CGA GAT GCA AGC AAT CAC AAG CGT      5184
Leu Pro Glu Gly Phe Thr Leu Arg Arg Asp Ala Ser Asn His Lys Arg
1710                1715                1720                1725

CGT GAG CCA GTG GGA CAG GAT GCT GTG GGG CTG AAA AAT CTC TCA GTG      5232
Arg Glu Pro Val Gly Gln Asp Ala Val Gly Leu Lys Asn Leu Ser Val
                1730                1735                1740
```

```
CAA GTC TCA GAA GCT AAC CTA ATT GGT ACT GGA ACA AGT GAA CAC TGG      5280
Gln Val Ser Glu Ala Asn Leu Ile Gly Thr Gly Thr Ser Glu His Trp
            1745                1750                1755

GTC GAT GAT GAA GGG CCC CAG CCA AAG AAA GTA AAG GCT GAA GAT GAG      5328
Val Asp Asp Glu Gly Pro Gln Pro Lys Lys Val Lys Ala Glu Asp Glu
        1760                1765                1770

GCC TTA CTC TCA GAA GAA GAT GAC CCC ATT GAT CGA CGG CCA TGG ACA      5376
Ala Leu Leu Ser Glu Glu Asp Asp Pro Ile Asp Arg Arg Pro Trp Thr
    1775                1780                1785

CAG CAG CAC CTT GAA GCT GCA GAC ATC CGT AGG ACA CCA TCG CTG GCT      5424
Gln Gln His Leu Glu Ala Ala Asp Ile Arg Arg Thr Pro Ser Leu Ala
1790                1795                1800                1805

CTC ACC CCT CCT CAG GCA GAG CAG GAG GTG GAT GTG TTA GAT GTG AAT      5472
Leu Thr Pro Pro Gln Ala Glu Gln Glu Val Asp Val Leu Asp Val Asn
                1810                1815                1820

GTC CGT GGC CCA GAT GGC TGC ACC CCA TTG ATG TTG GCT TCT CTC CGA      5520
Val Arg Gly Pro Asp Gly Cys Thr Pro Leu Met Leu Ala Ser Leu Arg
            1825                1830                1835

GGA GGC AGC TCA GAT TTG AGT GAT GAA GAT GAA GAT GCA GAG GAC TCT      5568
Gly Gly Ser Ser Asp Leu Ser Asp Glu Asp Glu Asp Ala Glu Asp Ser
        1840                1845                1850

TCT GCT AAC ATC ATC ACA GAC TTG GTC TAC CAG GGT GCC AGC CTC CAG      5616
Ser Ala Asn Ile Ile Thr Asp Leu Val Tyr Gln Gly Ala Ser Leu Gln
    1855                1860                1865

GCC CAG ACA GAC CGG ACT GGT GAG ATG GCC CTG CAC CTT GCA GCC CGC      5664
Ala Gln Thr Asp Arg Thr Gly Glu Met Ala Leu His Leu Ala Ala Arg
1870                1875                1880                1885

TAC TCA CGG GCT GAT GCT GCC AAG CGT CTC CTG GAT GCA GGT GCA GAT      5712
Tyr Ser Arg Ala Asp Ala Ala Lys Arg Leu Leu Asp Ala Gly Ala Asp
                1890                1895                1900

GCC AAT GCC CAG GAC AAC ATG GGC CGC TGT CCA CTC CAT GCT GCA GTG      5760
Ala Asn Ala Gln Asp Asn Met Gly Arg Cys Pro Leu His Ala Ala Val
            1905                1910                1915

GCA GCT GAT GCC CAA GGT GTC TTC CAG ATT CTG ATT CGC AAC CGA GTA      5808
Ala Ala Asp Ala Gln Gly Val Phe Gln Ile Leu Ile Arg Asn Arg Val
        1920                1925                1930

ACT GAT CTA GAT GCC AGG ATG AAT GAT GGT ACT ACA CCC CTG ATC CTG      5856
Thr Asp Leu Asp Ala Arg Met Asn Asp Gly Thr Thr Pro Leu Ile Leu
    1935                1940                1945

GCT GCC CGC CTG GCT GTG GAG GGA ATG GTG GCA GAA CTG ATC AAC TGC      5904
Ala Ala Arg Leu Ala Val Glu Gly Met Val Ala Glu Leu Ile Asn Cys
1950                1955                1960                1965

CAA GCG GAT GTG AAT GCA GTG GAT GAC CAT GGA AAA TCT GCT CTT CAC      5952
Gln Ala Asp Val Asn Ala Val Asp Asp His Gly Lys Ser Ala Leu His
                1970                1975                1980

TGG GCA GCT GCT GTC AAT AAT GTG GAG GCA ACT CTT TTG TTG TTG AAA      6000
Trp Ala Ala Ala Val Asn Asn Val Glu Ala Thr Leu Leu Leu Leu Lys
            1985                1990                1995

AAT GGG GCC AAC CGA GAC ATG CAG GAC AAC AAG GAA GAG ACA CCT CTG      6048
Asn Gly Ala Asn Arg Asp Met Gln Asp Asn Lys Glu Glu Thr Pro Leu
        2000                2005                2010
```

```
TTT CTT GCT GCC CGG GAG GGG AGC TAT GAA GCA GCC AAG ATC CTG TTA      6096
Phe Leu Ala Ala Arg Glu Gly Ser Tyr Glu Ala Ala Lys Ile Leu Leu
    2015                2020                2025

GAC CAT TTT GCC AAT CGA GAC ATC ACA GAC CAT ATG GAT CGT CTT CCC      6144
Asp His Phe Ala Asn Arg Asp Ile Thr Asp His Met Asp Arg Leu Pro
2030                2035                2040                2045

CGG GAT GTG GCT CGG GAT CGC ATG CAC CAT GAC ATT GTG CGC CTT CTG      6192
Arg Asp Val Ala Arg Asp Arg Met His His Asp Ile Val Arg Leu Leu
                2050                2055                2060

GAT GAA TAC AAT GTG ACC CCA AGC CCT CCA GGC ACC GTG TTG ACT TCT      6240
Asp Glu Tyr Asn Val Thr Pro Ser Pro Pro Gly Thr Val Leu Thr Ser
            2065                2070                2075

GCT CTC TCA CCT GTC ATC TGT GGG CCC AAC AGA TCT TTC CTC AGC CTG      6288
Ala Leu Ser Pro Val Ile Cys Gly Pro Asn Arg Ser Phe Leu Ser Leu
        2080                2085                2090

AAG CAC ACC CCA ATG GGC AAG AAG TCT AGA CGG CCC AGT GCC AAG AGT      6336
Lys His Thr Pro Met Gly Lys Lys Ser Arg Arg Pro Ser Ala Lys Ser
    2095                2100                2105

ACC ATG CCT ACT AGC CTC CCT AAC CTT GCC AAG GAG GCA AAG GAT GCC      6384
Thr Met Pro Thr Ser Leu Pro Asn Leu Ala Lys Glu Ala Lys Asp Ala
2110                2115                2120                2125

AAG GGT AGT AGG AGG AAG AAG TCT CTG AGT GAG AAG GTC CAA CTG TCT      6432
Lys Gly Ser Arg Arg Lys Lys Ser Leu Ser Glu Lys Val Gln Leu Ser
                2130                2135                2140

GAG AGT TCA GTA ACT TTA TCC CCT GTT GAT TCC CTA GAA TCT CCT CAC      6480
Glu Ser Ser Val Thr Leu Ser Pro Val Asp Ser Leu Glu Ser Pro His
            2145                2150                2155

ACG TAT GTT TCC GAC ACC ACA TCC TCT CCA ATG ATT ACA TCC CCT GGG      6528
Thr Tyr Val Ser Asp Thr Thr Ser Ser Pro Met Ile Thr Ser Pro Gly
        2160                2165                2170

ATC TTA CAG GCC TCA CCC AAC CCT ATG TTG GCC ACT GCC GCC CCT CCT      6576
Ile Leu Gln Ala Ser Pro Asn Pro Met Leu Ala Thr Ala Ala Pro Pro
    2175                2180                2185

GCC CCA GTC CAT GCC CAG CAT GCA CTA TCT TTT TCT AAC CTT CAT GAA      6624
Ala Pro Val His Ala Gln His Ala Leu Ser Phe Ser Asn Leu His Glu
2190                2195                2200                2205

ATG CAG CCT TTG GCA CAT GGG GCC AGC ACT GTG CTT CCC TCA GTG AGC      6672
Met Gln Pro Leu Ala His Gly Ala Ser Thr Val Leu Pro Ser Val Ser
                2210                2215                2220

CAG TTG CTA TCC CAC CAC CAC ATT GTG TCT CCA GGC AGT GGC AGT GCT      6720
Gln Leu Leu Ser His His His Ile Val Ser Pro Gly Ser Gly Ser Ala
            2225                2230                2235

GGA AGC TTG AGT AGG CTC CAT CCA GTC CCA GTC CCA GCA GAT TGG ATG      6768
Gly Ser Leu Ser Arg Leu His Pro Val Pro Val Pro Ala Asp Trp Met
        2240                2245                2250

AAC CGC ATG GAG GTG AAT GAG ACC CAG TAC AAT GAG ATG TTT GGT ATG      6816
Asn Arg Met Glu Val Asn Glu Thr Gln Tyr Asn Glu Met Phe Gly Met
    2255                2260                2265

GTC CTG GCT CCA GCT GAG GGC ACC CAT CCT GGC ATA GCT CCC CAG AGC      6864
Val Leu Ala Pro Ala Glu Gly Thr His Pro Gly Ile Ala Pro Gln Ser
2270                2275                2280                2285
```

```
AGG CCA CCT GAA GGG AAG CAC ATA ACC ACC CCT CGG GAG CCC TTG CCC      6912
Arg Pro Pro Glu Gly Lys His Ile Thr Thr Pro Arg Glu Pro Leu Pro
                2290                2295                2300

CCC ATT GTG ACT TTC CAG CTC ATC CCT AAA GGC AGT ATT GCC CAA CCA      6960
Pro Ile Val Thr Phe Gln Leu Ile Pro Lys Gly Ser Ile Ala Gln Pro
            2305                2310                2315

GCG GGG GCT CCC CAG CCT CAG TCC ACC TGC CCT CCA GCT GTT GCG GGC      7008
Ala Gly Ala Pro Gln Pro Gln Ser Thr Cys Pro Pro Ala Val Ala Gly
        2320                2325                2330

CCC CTG CCC ACC ATG TAC CAG ATT CCA GAA ATG GCC CGT TTG CCC AGT      7056
Pro Leu Pro Thr Met Tyr Gln Ile Pro Glu Met Ala Arg Leu Pro Ser
    2335                2340                2345

GTG GCT TTC CCC ACT GCC ATG ATG CCC CAG CAG GAC GGG CAG GTA GCT      7104
Val Ala Phe Pro Thr Ala Met Met Pro Gln Gln Asp Gly Gln Val Ala
2350                2355                2360                2365

CAG ACC ATT CTC CCA GCC TAT CAT CCT TTC CCA GCC TCT GTG GGC AAG      7152
Gln Thr Ile Leu Pro Ala Tyr His Pro Phe Pro Ala Ser Val Gly Lys
                2370                2375                2380

TAC CCC ACA CCC CCT TCA CAG CAC AGT TAT GCT TCC TCA AAT GCT GCT      7200
Tyr Pro Thr Pro Pro Ser Gln His Ser Tyr Ala Ser Ser Asn Ala Ala
            2385                2390                2395

GAG CGA ACA CCC AGT CAC AGT GGT CAC CTC CAG GGT GAG CAT CCC TAC      7248
Glu Arg Thr Pro Ser His Ser Gly His Leu Gln Gly Glu His Pro Tyr
        2400                2405                2410

CTG ACA CCA TCC CCA GAG TCT CCT GAC CAG TGG TCA AGT TCA TCA CCC      7296
Leu Thr Pro Ser Pro Glu Ser Pro Asp Gln Trp Ser Ser Ser Ser Pro
    2415                2420                2425

CAC TCT GCT TCT GAC TGG TCA GAT GTG ACC ACC AGC CCT ACC CCT GGG      7344
His Ser Ala Ser Asp Trp Ser Asp Val Thr Thr Ser Pro Thr Pro Gly
2430                2435                2440                2445

GGT GCT GGA GGA GGT CAG CGG GGA CCT GGG ACA CAC ATG TCT GAG CCA      7392
Gly Ala Gly Gly Gly Gln Arg Gly Pro Gly Thr His Met Ser Glu Pro
                2450                2455                2460

CCA CAC AAC AAC ATG CAG GTT TAT GCG TGAGAGAGTC CACCTCCAGT           7439
Pro His Asn Asn Met Gln Val Tyr Ala
            2465                2470

GTAGAGACAT AACTGACTTT TGTAAATGCT GCTGAGGAAC AAATGAAGGT CATCCGGGAG   7499

AGAAATGAAG AAATCTCTGG AGCCAGCTTC TAGAGGTAGG AAAGAGAAGA TGTTCTTATT   7559

CAGATAATGC AAGAGAAGCA ATTCGTCAGT TTCACTGGGT ATCTGCAAGG CTTATTGATT   7619

ATTCTAATCT AATAAGACAA GTTTGTGGAA ATGCAAGATG AATACAAGCC TTGGGTCCAT   7679

GTTTACTCTC TTCTATTTGG AGAATAAGAT GGATGCTTAT TGAAGCCCAG ACATTCTTGC   7739

AGCTTGGACT GCATTTTAAG CCCTGCAGGC TTCTGCCATA TCCATGAGAA GATTCTACAC   7799

TAGCGTCCTG TTGGGAATTA TGCCCTGGAA TTCTGCCTGA ATTGACCTAC GCATCTCCTC   7859

CTCCTTGGAC ATTCTTTTGT CTTCATTTGG TGCTTTTGGT TTTGCACCTC TCCGTGATTG   7919

TAGCCCTACC AGCATGTTAT AGGGCAAGAC CTTTGTGCTT TTGATCATTC TGGCCCATGA   7979

AAGCAACTTT GGTCTCCTTT CCCCTCCTGT CTTCCCGGTA TCCCTTGGAG TCTCACAAGG   8039
```

```
TTTACTTTGG TATGGTTCTC AGCACAAACC TTTCAAGTAT GTTGTTTCTT TGGAAAATGG   8099
ACATACTGTA TTGTGTTCTC CTGCATATAT CATTCCTGGA GAGAGAAGGG GAGAAGAATA   8159
CTTTTCTTCA ACAAATTTTG GGGGCAGGAG ATCCCTTCAA GAGGCTGCAC CTTAATTTTT   8219
CTTGTCTGTG TGCAGGTCTT CATATAAACT TTACCAGGAA GAAGGGTGTG AGTTTGTTGT   8279
TTTTCTGTGT ATGGGCCTGG TCAGTGTAAA GTTTTATCCT TGATAGTCTA GTTACTATGA   8339
CCCTCCCCAC TTTTTTAAAA CCAGAAAAAG GTTTGGAATG TTGGAATGAC CAAGAGACAA   8399
GTTAACTCGT GCAAGAGCCA GTTACCCACC CACAGGTCCC CCTACTTCCT GCCAAGCATT   8459
CCATTGACTG CCTGTATGGA ACACATTTGT CCCAGATCTG AGCATTCTAG GCCTGTTTCA   8519
CTCACTCACC CAGCATATGA AACTAGTCTT AACTGTTGAG CCTTTCCTTT CATATCCACA   8579
GAAGACACTG TCTCAAATGT TGTACCCTTG CCATTTAGGA CTGAACTTTC CTTAGCCCAA   8639
GGGACCCAGT GACAGTTGTC TTCCGTTTGT CAGATGATCA GTCTCTACTG ATTATCTTGC   8699
TGCTTAAAGG CCTGCTCACC AATCTTTCTT TCACACCGTG TGGTCCGTGT TACTGGTATA   8759
CCCAGTATGT TCTCACTGAA GACATGGACT TTATATGTTC AAGTGCAGGA ATTGGAAAGT   8819
TGGACTTGTT TTCTATGATC CAAAACAGCC CTATAAGAAG GTTGGAAAAG GAGGAACTAT   8879
ATAGCAGCCT TTGCTATTTT CTGCTACCAT TTCTTTTCCT CTGAAGCGGC CATGACATTC   8939
CCTTTGGCAA CTAACGTAGA AACTCAACAG AACATTTTCC TTTCCTAGAG TCACCTTTTA   8999
GATGATAATG GACAACTATA GACTTGCTCA TTGTTCAGAC TGATTGCCCC TCACCTGAAT   9059
CCACTCTCTG TATTCATGCT CTTGGCAATT TCTTTGACTT TCTTTTAAGG GCAGAAGCAT   9119
TTTAGTTAAT TGTAGATAAA GAATAGTTTT CTTCCTCTTC TCCTTGGGCC AGTTAATAAT   9179
TGGTCCATGG CTACACTGCA ACTTCCGTCC AGTGCTGTGA TGCCCATGAC ACCTGCAAAA   9239
TAAGTTCTGC CTGGGCATTT TGTAGATATT AACAGGTGAA TTCCCGACTC TTTTGGTTTG   9299
AATGACAGTT CTCATTCCTT CTATGGCTGC AAGTATGCAT CAGTGCTTCC CACTTACCTG   9359
ATTTGTCTGT CGGTGGCCCC ATATGGAAAC CCTGCGTGTC TGTTGGCATA ATAGTTTACA   9419
AATGGTTTTT TCAGTCCTAT CCAAATTTAT TGAACCAACA AAAATAATTA CTTCTGCCCT   9479
GAGATAAGCA GATTAAGTTT GTTCATTCTC TGCTTTATTC TCTCCATGTG GCAACATTCT   9539
GTCAGCCTCT TTCATAGTGT GCAAACATTT TATCATTCTA AATGGTGACT CTCTGCCCTT   9599
GGACCCATTT ATTATTCACA GATGGGGAGA ACCTATCTGC ATGGACCCTC ACCATCCTCT   9659
GTGCAGCACA CACAGTGCAG GGAGCCAGTG GCGATGGCGA TGACTTTCTT CCCCTGGGAA   9719
TTCC                                                               9723
```

Patentansprüche

1. Pharmazeutische Zusammensetzung, die einen neutralisierenden Anti-Notch Antikörper und einen pharmazeutisch annehmbaren Träger enthält.

2. Pharmazeutische Zusammensetzung, die ein Fragment oder Derivat eines Antikörpers zu einem Notch-Protein, das den Idiotyp des Antikörpers enthält, und einen pharmazeutisch annehmbaren Träger enthält.

3. Verwendung einer Zusammensetzung, enthaltend ein Molekül, das im Hinblick auf die Funktion eines Notch-Proteins antagonistisch wirkt, zur Herstellung eines Medikaments zur Behandlung einer krebsartigen Er-

krankung oder zur Verhinderung einer Fortentwicklung eines präneoplastischen oder nicht-malignen Zustands zu einem neoplastischen oder malignen Zustand, wobei das Molekül ausgewählt ist aus der Gruppe, die besteht aus einer Notch Antisense-Nukleinsäure, einem neutralisierenden Anti-Notch Antikörper, einem Fragement, das die bindende Domäne eines gegen Notch gerichteten Antikörpers enthält, und einem Protein, das Notch ELR-11 und ELR-12 aufweist.

4. Verwendung nach Anspruch 3, bei der die krebsartige Erkrankung oder der maligne Zustand Gebärmutterhalskrebs, Brustkrebs oder Dickdarmkrebs ist.

5. Verwendung eines Moleküls, wobei das Molekül im Hinblick auf die Funktion eines Notch-Proteins antagonistisch wirkt, zur Herstellung eines Medikaments zur Behandlung einer bösartigen Erkrankung in einem Patienten, wobei die bösartige Erkrankung gekennzeichnet ist durch (i) eine erhöhte Notch-Aktivität oder (ii) eine erhöhte Expression eines Notch-Proteins oder (iii) eine erhöhte Expression eines Notch-Derivats, das in der Lage ist, von einem Anti-Notch-Antikörper gebunden zu werden, und zwar bezogen auf die genannte Notch-Aktivität oder -Expression in einer analogen nicht-malignen Probe, wobei das Molekül ausgewählt ist aus der Gruppe, die besteht aus einer Notch Antisense-Nukleinsäure, einem neutralisierenden Anti-Notch Antikörper, einem Fragement, das die bindende Domäne eines gegen Notch gerichteten Antikörpers enthält, und einem Protein, das Notch ELR-11 und ELR-12 aufweist.

6. Verwendung nach Anspruch 5, wobei der Patient ein Mensch ist.

7. Verwendung nach Anspruch 4, wobei das Molekül ein neutralisierender Anti-Notch Antikörper oder ein Teil des genannten Antikörpers ist, der dessen Bindungsdomäne enthält.

8. Verwendung nach Anspruch 4, bei der das Protein, das Notch ELR-11 und ELR-12 aufweist, ein Notch-Fragment ist, das die extrazelluläre Domäne eines Notch-Proteins oder einen Teil davon aufweist, der in der Lage ist, an einen Notch-Liganden zu binden.

9. Verwendung nach Anspruch 4, bei der das Protein, das Notch ELR-11 und ELR-12 aufweist, ein Notch-Fragment ist, das die EGF-homologen Wiederholungseinheiten eines Notch-Proteins aufweist.

10. Verwendung nach Anspruch 4, bei der das Protein, das Notch ELR-11 und ELR-12 aufweist, ein Notch-Fragment ist, das ein adhäsives Fragment eines Notch-Proteins aufweist.

11. Verwendung nach Anspruch 4, bei der die Notch Antisense Nukleinsäure ein Oligonukleotid ist, das (a) wenigstens aus sechs Nukleotiden besteht, (b) eine Sequenz umfasst, die komplementär zu wenigstens einem Teil eines RNA-Transkripts eines Notch-Gens ist; und (c) mit dem RNA-Transkript hybridisierbar ist.

12. Verwendung eines neutralisierenden Anti-Notch Antikörpers zur Herstellung eines Medikaments zur Verhinderung oder zur Behandlung einer malignen Erkrankung bei einem Patienten.

13. Verwendung nach Anspruch 12, bei der der Antikörper monoklonal ist.

14. Verfahren zur Diagnose einer malignen Erkrankung, die gekennzeichnet ist durch eine erhöhte Menge eines Notch-Proteins oder eines Notch-Derivats, das in der Lage ist, von einem Anti-Notch Antikörper gebunden zu werden, das das Messen der Menge eines Notch-Proteins oder eines Notch-Derivats, das von einem Anti-Notch Antikörper gebunden werden kann, in einer Probe umfasst, die maligne Zellen eines Patienten enthält oder im Verdacht steht, derartige Zellen zu enthalten, wobei eine Zunahme des Notch-Proteins oder des Notch-Derivats, das von einem Anti-Notch Antikörper gebunden werden kann, in der Probe, bezogen auf die entsprechende Menge in einer analogen Probe von nicht-malignen Zellen, das Vorliegen der Erkrankung oder Gesundheitsstörung bei dem Patienten anzeigt.

15. Verfahren nach Anspruch 14, bei dem die bösartige Erkrankung Gebärmutterhalskrebs ist.

16. Verfahren nach Anspruch 14, bei dem die bösartige Erkrankung Brustkrebs ist.

17. Verfahren nach Anspruch 14, bei dem die bösartige Erkrankung Dickdarmkrebs ist.

18. Verfahren nach Anspruch 14, bei dem die Menge des Notch-Proteins oder des Derivats mittels eines Verfahrens gemessen wird, das das Inkontaktbringen der Probe mit einem Anti-Notch-Antikörper auf eine Wei-

se, dass es zu einer immunspezifischen Bindung kommen kann, und das Messen der Menge jeglicher immunspezifischer Bindung des Antikörpers, zu der es kommt, umfasst.

19. Zusammensetzung, die eine therapeutisch wirksame Menge eines Moleküls aufweist, das im Hinblick auf die Funktion eines Notch-Proteins antagonistisch wirkt, zur Verwendung als ein Medikament, wobei das Molekül aus der Gruppe ausgewählt ist, die besteht aus einer Notch Antisense-Nukleinsäure, einem neutralisierenden Anti-Notch Antikörper, einem Fragement, das die bindende Domäne eines gegen Notch gerichteten Antikörpers enthält, und einem Protein, das Notch ELR-11 und Notch ELR-12 aufweist.

Es folgen 68 Blatt Zeichnungen

Anhängende Zeichnungen

```
GAATTCGGAG GAATTATTCA AAACATAAAC ACAATAAACA ATTTGAGTAG TTGCCGCACA   60

CACACACACA CACAGCCCGT GGATTATTAC ACTAAAAGCG ACACTCAATC CAAAAAATCA  120

GCAACAAAAA CATCAATAAA C ATG CAT TGG ATT AAA TGT TTA TTA ACA GCA     171
                        Met His Trp Ile Lys Cys Leu Leu Thr Ala
                         1               5                  10

TTC ATT TGC TTC ACA GTC ATC GTG CAG GTT CAC AGT TCC GGC AGC TTT     219
Phe Ile Cys Phe Thr Val Ile Val Gln Val His Ser Ser Gly Ser Phe
            15                  20                  25

GAG TTG CGC CTG AAG TAC TTC AGC AAC GAT CAC GGG CGG GAC AAC GAG     267
Glu Leu Arg Leu Lys Tyr Phe Ser Asn Asp His Gly Arg Asp Asn Glu
        30                  35                  40

GGT CGC TGC TGC AGC GGG GAG TCG GAC GGA GCG ACG GGC AAG TGC CTG     315
Gly Arg Cys Cys Ser Gly Glu Ser Asp Gly Ala Thr Gly Lys Cys Leu
    45                  50                  55

GGC AGC TGC AAG ACG CGG TTT CGC GTC TGC CTA AAG CAC TAC CAG GCC     363
Gly Ser Cys Lys Thr Arg Phe Arg Val Cys Leu Lys His Tyr Gln Ala
60                  65                  70

ACC ATC GAC ACC ACC TCC CAG TGC ACC TAC GGG GAC GTG ATC ACG CCC     411
Thr Ile Asp Thr Thr Ser Gln Cys Thr Tyr Gly Asp Val Ile Thr Pro
75                  80                  85                  90

ATT CTC GGC GAG AAC TCG GTC AAT CTG ACC GAC GCC CAG CGC TTC CAG     459
Ile Leu Gly Glu Asn Ser Val Asn Leu Thr Asp Ala Gln Arg Phe Gln
                95                  100                 105

AAC AAG GGC TTC ACG AAT CCC ATC CAG TTC CCC TTC TCG TTC TCA TGG     507
Asn Lys Gly Phe Thr Asn Pro Ile Gln Phe Pro Phe Ser Phe Ser Trp
            110                 115                 120
```

FIG.1A

```
CCG GGT ACC TTC TCG CTG ATC GTC GAG GCC TGG CAT GAT ACG AAC AAT   555
Pro Gly Thr Phe Ser Leu Ile Val Glu Ala Trp His Asp Thr Asn Asn
        125                 130                 135

AGC GGC AAT GCG CGA ACC AAC AAG CTC CTC ATC CAG CGA CTC TTG GTG   603
Ser Gly Asn Ala Arg Thr Asn Lys Leu Leu Ile Gln Arg Leu Leu Val
    140                 145                 150

CAG CAG GTA CTG GAG GTG TCC TCC GAA TGG AAG ACG AAC AAG TCG GAA   651
Gln Gln Val Leu Glu Val Ser Ser Glu Trp Lys Thr Asn Lys Ser Glu
155                 160                 165                 170

TCG CAG TAC ACG TCG CTG GAG TAC GAT TTC CGT GTC ACC TGC GAT CTC   699
Ser Gln Tyr Thr Ser Leu Glu Tyr Asp Phe Arg Val Thr Cys Asp Leu
                175                 180                 185

AAC TAC TAC GGA TCC GGC TGT GCC AAG TTC TGC CGG CCC CGC GAC GAT   747
Asn Tyr Tyr Gly Ser Gly Cys Ala Lys Phe Cys Arg Pro Arg Asp Asp
            190                 195                 200

TCA TTT GGA CAC TCG ACT TGC TCG GAG ACG GGC GAA ATT ATC TGT TTG   795
Ser Phe Gly His Ser Thr Cys Ser Glu Thr Gly Glu Ile Ile Cys Leu
        205                 210                 215

ACC GGA TGG CAG GGC GAT TAC TGT CAC ATA CCC AAA TGC GCC AAA GGC   843
Thr Gly Trp Gln Gly Asp Tyr Cys His Ile Pro Lys Cys Ala Lys Gly
    220                 225                 230

TGT GAA CAT GGA CAT TGC GAC AAA CCC AAT CAA TGC GTT TGC CAA CTG   891
Cys Glu His Gly His Cys Asp Lys Pro Asn Gln Cys Val Cys Gln Leu
235                 240                 245                 250

GGC TGG AAG GGA GCC TTG TGC AAC GAG TGC GTT CTG GAA CCG AAC TGC   939
Gly Trp Lys Gly Ala Leu Cys Asn Glu Cys Val Leu Glu Pro Asn Cys
                255                 260                 265
```

FIG.1B

```
ATC CAT GGC ACC TGC AAC AAA CCC TGG ACT TGC ATC TGC AAC GAG GGT      987
Ile His Gly Thr Cys Asn Lys Pro Trp Thr Cys Ile Cys Asn Glu Gly
            270                 275                 280

TGG GGA GGC TTG TAC TGC AAC CAG GAT CTG AAC TAC TGC ACC AAC CAC     1035
Trp Gly Gly Leu Tyr Cys Asn Gln Asp Leu Asn Tyr Cys Thr Asn His
        285                 290                 295

AGA CCC TGC AAG AAT GGC GGA ACC TGC TTC AAC ACC GGC GAG GGA TTG     1083
Arg Pro Cys Lys Asn Gly Gly Thr Cys Phe Asn Thr Gly Glu Gly Leu
    300                 305                 310

TAC ACA TGC AAA TGC GCT CCA GGA TAC AGT GGT GAT GAT TGC GAA AAT     1131
Tyr Thr Cys Lys Cys Ala Pro Gly Tyr Ser Gly Asp Asp Cys Glu Asn
315                 320                 325                 330

GAG ATC TAC TCC TGC GAT GCC GAT GTC AAT CCC TGC CAG AAT GGT GGT     1179
Glu Ile Tyr Ser Cys Asp Ala Asp Val Asn Pro Cys Gln Asn Gly Gly
                335                 340                 345

ACC TGC ATC GAT GAG CCG CAC ACA AAA ACC GGC TAC AAG TGT CAT TGC     1227
Thr Cys Ile Asp Glu Pro His Thr Lys Thr Gly Tyr Lys Cys His Cys
            350                 355                 360

GCC AAC GGC TGG AGC GGA AAG ATG TGC GAG GAG AAA GTG CTC ACG TGT     1275
Ala Asn Gly Trp Ser Gly Lys Met Cys Glu Glu Lys Val Leu Thr Cys
        365                 370                 375

TCG GAC AAA CCC TGT CAT CAG GGA ATC TGC CGC AAC GTT CGT CCT GGC     1323
Ser Asp Lys Pro Cys His Gln Gly Ile Cys Arg Asn Val Arg Pro Gly
    380                 385                 390

TTG GGA AGC AAG GGT CAG GGC TAC CAG TGC GAA TGT CCC ATT GGC TAC     1371
Leu Gly Ser Lys Gly Gln Gly Tyr Gln Cys Glu Cys Pro Ile Gly Tyr
395                 400                 405                 410
```

FIG.1C

```
AGC GGA CCC AAC TGC GAT CTC CAG CTG GAC AAC TGC AGT CCG AAT CCA   1419
Ser Gly Pro Asn Cys Asp Leu Gln Leu Asp Asn Cys Ser Pro Asn Pro
                415                 420                 425

TGC ATA AAC GGT GGA AGC TGT CAG CCG AGC GGA AAG TGT ATT TGC CCA   1467
Cys Ile Asn Gly Gly Ser Cys Gln Pro Ser Gly Lys Cys Ile Cys Pro
            430                 435                 440

GCG GGA TTT TCG GGA ACG AGA TGC GAG ACC AAC ATT GAC GAT TGT CTT   1515
Ala Gly Phe Ser Gly Thr Arg Cys Glu Thr Asn Ile Asp Asp Cys Leu
        445                 450                 455

GGC CAC CAG TGC GAG AAC GGA GGC ACC TGC ATA GAT ATG GTC AAC CAA   1563
Gly His Gln Cys Glu Asn Gly Gly Thr Cys Ile Asp Met Val Asn Gln
    460                 465                 470

TAT CGC TGC CAA TGC GTT CCC GGT TTC CAT GGC ACC CAC TGT AGT AGC   1611
Tyr Arg Cys Gln Cys Val Pro Gly Phe His Gly Thr His Cys Ser Ser
475                 480                 485                 490

AAA GTT GAC TTG TGC CTC ATC AGA CCG TGT GCC AAT GGA GGA ACC TGC   1659
Lys Val Asp Leu Cys Leu Ile Arg Pro Cys Ala Asn Gly Gly Thr Cys
                495                 500                 505

TTG AAT CTC AAC AAC GAT TAC CAG TGC ACC TGT CGT GCG GGA TTT ACT   1707
Leu Asn Leu Asn Asn Asp Tyr Gln Cys Thr Cys Arg Ala Gly Phe Thr
            510                 515                 520

GGC AAG GAT TGC TCT GTG GAC ATC GAT GAG TGC AGC AGT GGA CCC TGT   1755
Gly Lys Asp Cys Ser Val Asp Ile Asp Glu Cys Ser Ser Gly Pro Cys
        525                 530                 535

CAT AAC GGC GGC ACT TGC ATG AAC CGC GTC AAT TCG TTC GAA TGC GTG   1803
His Asn Gly Gly Thr Cys Met Asn Arg Val Asn Ser Phe Glu Cys Val
    540                 545                 550
```

FIG.1D

```
TGT GCC AAT GGT TTC AGG GGC AAG CAG TGC GAT GAG GAG TCC TAC GAT   1851
Cys Ala Asn Gly Phe Arg Gly Lys Gln Cys Asp Glu Glu Ser Tyr Asp
555                 560                 565                 570

TCG GTG ACC TTC GAT GCC CAC CAA TAT GGA GCG ACC ACA CAA GCG AGA   1899
Ser Val Thr Phe Asp Ala His Gln Tyr Gly Ala Thr Thr Gln Ala Arg
                575                 580                 585

GCC GAT GGT TTG ACC AAT GCC CAG GTA GTC CTA ATT GCT GTT TTC TCC   1947
Ala Asp Gly Leu Thr Asn Ala Gln Val Val Leu Ile Ala Val Phe Ser
            590                 595                 600

GTT GCG ATG CCT TTG GTG GCG GTT ATT GCG GCG TGC GTG GTC TTC TGC   1995
Val Ala Met Pro Leu Val Ala Val Ile Ala Ala Cys Val Val Phe Cys
        605                 610                 615

ATG AAG CGC AAG CGT AAG CGT GCT CAG GAA AAG GAC GAC GCG GAG GCC   2043
Met Lys Arg Lys Arg Lys Arg Ala Gln Glu Lys Asp Asp Ala Glu Ala
    620                 625                 630

AGG AAG CAG AAC GAA CAG AAT GCG GTG GCC ACA ATG CAT CAC AAT GGC   2091
Arg Lys Gln Asn Glu Gln Asn Ala Val Ala Thr Met His His Asn Gly
635                 640                 645                 650

AGT GGG GTG GGT GTA GCT TTG GCT TCA GCC TCT CTG GGC GGC AAA ACT   2139
Ser Gly Val Gly Val Ala Leu Ala Ser Ala Ser Leu Gly Gly Lys Thr
                655                 660                 665

GGC AGC AAC AGC GGT CTC ACC TTC GAT GGC GGC AAC CCG AAT ATC ATC   2187
Gly Ser Asn Ser Gly Leu Thr Phe Asp Gly Gly Asn Pro Asn Ile Ile
            670                 675                 680

AAA AAC ACC TGG GAC AAG TCG GTC AAC AAC ATT TGT GCC TCA GCA GCA   2235
Lys Asn Thr Trp Asp Lys Ser Val Asn Asn Ile Cys Ala Ser Ala Ala
        685                 690                 695
```

FIG. 1E

```
GCA GCG GCG GCG GCG GCA GCA GCG GCG GAC GAG TGT CTC ATG TAC GGC   2283
Ala Ala Ala Ala Ala Ala Ala Ala Ala Asp Glu Cys Leu Met Tyr Gly
    700                 705                 710

GGA TAT GTG GCC TCG GTG GCG GAT AAC AAC AAT GCC AAC TCA GAC TTT   2331
Gly Tyr Val Ala Ser Val Ala Asp Asn Asn Asn Ala Asn Ser Asp Phe
715                 720                 725                 730

TGT GTG GCT CCG CTA CAA AGA GCC AAG TCG CAA AAG CAA CTC AAC ACC   2379
Cys Val Ala Pro Leu Gln Arg Ala Lys Ser Gln Lys Gln Leu Asn Thr
                735                 740                 745

GAT CCC ACG CTC ATG CAC CGC GGT TCG CCG GCA GGC AGC TCA GCC AAG   2427
Asp Pro Thr Leu Met His Arg Gly Ser Pro Ala Gly Ser Ser Ala Lys
            750                 755                 760

GGA GCG TCT GGC GGA GGA CCG GGA GCG GCG GAG GGC AAG AGG ATC TCT   2475
Gly Ala Ser Gly Gly Gly Pro Gly Ala Ala Glu Gly Lys Arg Ile Ser
        765                 770                 775

GTT TTA GGC GAG GGT TCC TAC TGT AGC CAG CGT TGG CCC TCG TTG GCG   2523
Val Leu Gly Glu Gly Ser Tyr Cys Ser Gln Arg Trp Pro Ser Leu Ala
    780                 785                 790

GCG GCG GGA GTG GCC GGA GCC TGT TCA TCC CAG CTA ATG GCT GCA GCT   2571
Ala Ala Gly Val Ala Gly Ala Cys Ser Ser Gln Leu Met Ala Ala Ala
795                 800                 805                 810

TCG GCA GCG GGC AGC GGA GCG GGG ACG GCG CAA CAG CAG CGA TCC GTG   2619
Ser Ala Ala Gly Ser Gly Ala Gly Thr Ala Gln Gln Gln Arg Ser Val
                815                 820                 825

GTC TGC GGC ACT CCG CAT ATG TAACTCCAAA AATCCGGAAG GGCTCCTGGT      2670
Val Cys Gly Thr Pro His Met
            830

AAATCCGGAG AAATCCGCAT GGAGGAGCTG ACAGCACATA CACAAAGAAA AGACTGGGTT  2730
GGGTTCAAAA TGTGAGAGAG ACGCCAAAAT GTTGTTGTTG ATTGAAGCAG TTTAGTCGTC  2790
ACGAAAAATG AAAAATCTGT AACAGGCATA ACTCGTAAAC TCCCTAAAAA ATTTGTATAG  2850
TAATTAGCAA AGCTGTGACC CAGCCGTTTC GATCCCGAAT TC                     2892
```

FIG.1F

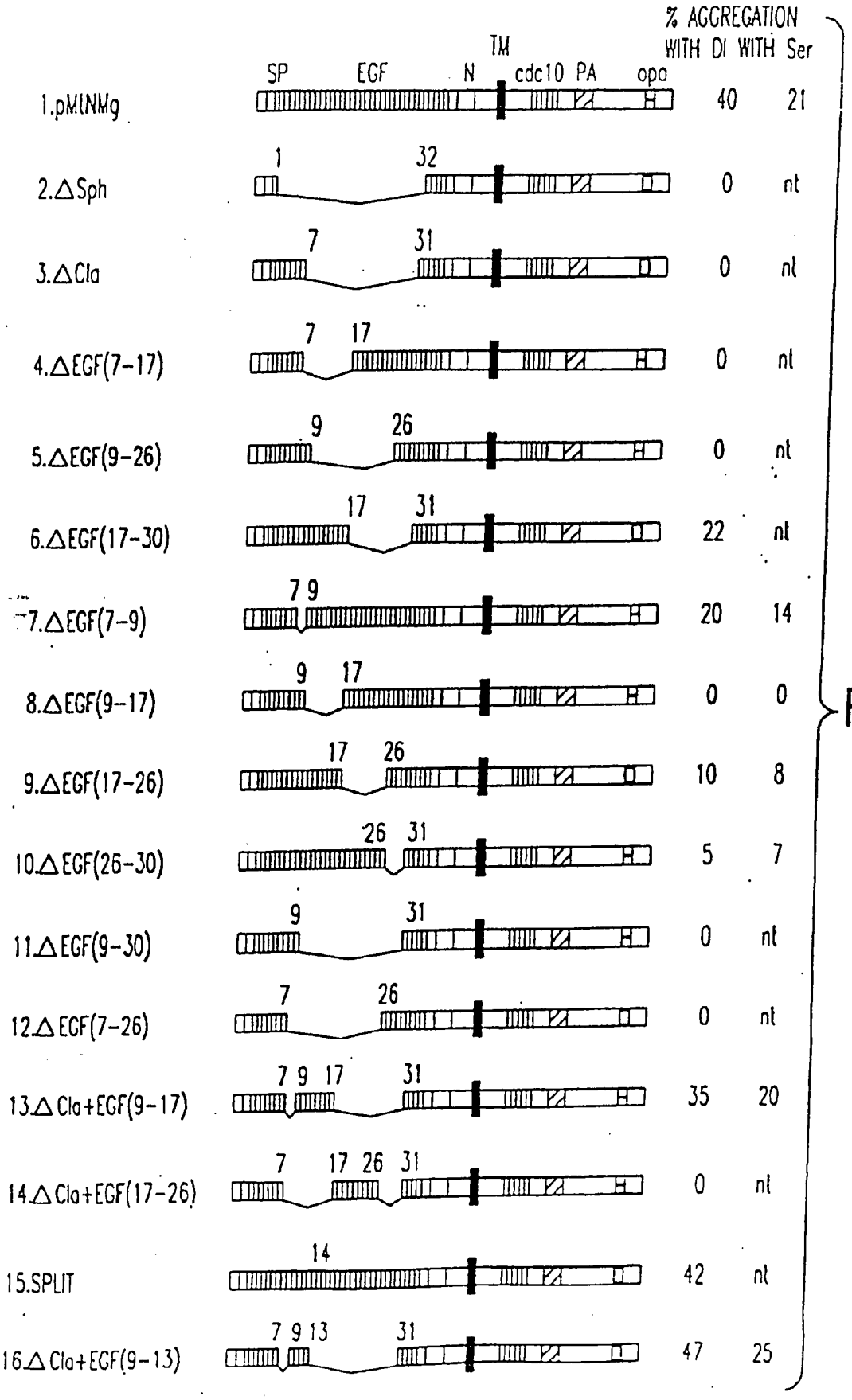

% AGGREGATION
WITH DI WITH Ser
TM
SP
EGF
N
cdc10
PA
opa
1.pMINMg
40 21
2.ΔSph
1 32
0 nt
3.ΔCla
7 31
0 nt
4.ΔEGF(7-17)
7 17
0 nt
5.ΔEGF(9-26)
9 26
0 nt
6.ΔEGF(17-30)
17 31
22 nt
7.ΔEGF(7-9)
7 9
20 14
8.ΔEGF(9-17)
9 17
0 0
9.ΔEGF(17-26)
17 26
10 8
10.ΔEGF(26-30)
26 31
5 7
11.ΔEGF(9-30)
9 31
0 nt
12.ΔEGF(7-26)
7 26
0 nt
13.ΔCla+EGF(9-17)
7 9 17 31
35 20
14.ΔCla+EGF(17-26)
7 17 26 31
0 nt
15.SPLIT
14
42 nt
16.ΔCla+EGF(9-13)
7 9 13 31
47 25

FIG.2A

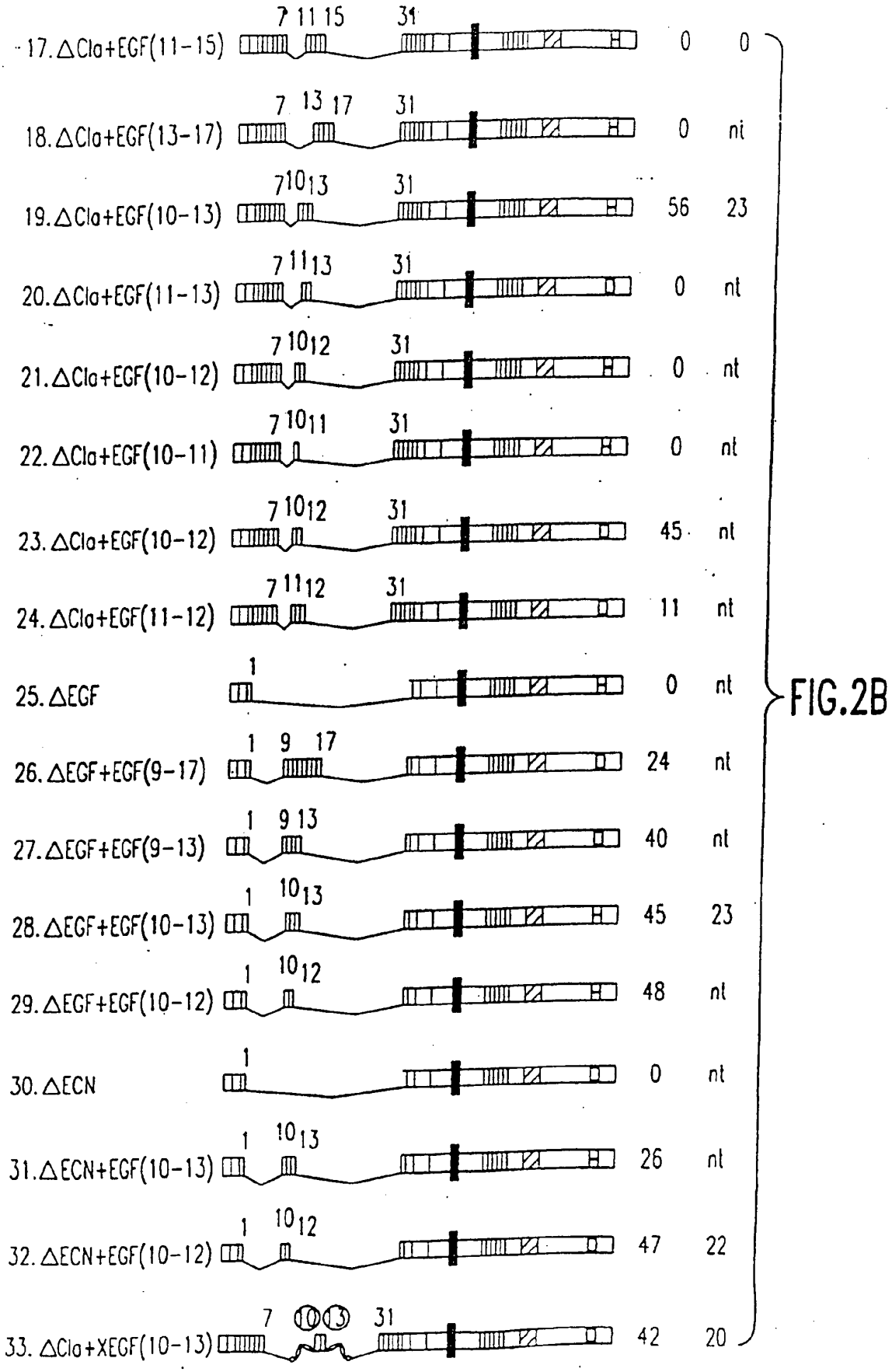

17. ΔCla+EGF(11-15) 7 11 15 31 0 0
18. ΔCla+EGF(13-17) 7 13 17 31 0 nt
19. ΔCla+EGF(10-13) 7 10 13 31 56 23
20. ΔCla+EGF(11-13) 7 11 13 31 0 nt
21. ΔCla+EGF(10-12) 7 10 12 31 0 nt
22. ΔCla+EGF(10-11) 7 10 11 31 0 nt
23. ΔCla+EGF(10-12) 7 10 12 31 45 nt
24. ΔCla+EGF(11-12) 7 11 12 31 11 nt
25. ΔEGF 1 0 nt
26. ΔEGF+EGF(9-17) 1 9 17 24 nt
27. ΔEGF+EGF(9-13) 1 9 13 40 nt
28. ΔEGF+EGF(10-13) 1 10 13 45 23
29. ΔEGF+EGF(10-12) 1 10 12 48 nt
30. ΔECN 1 0 nt
31. ΔECN+EGF(10-13) 1 10 13 26 nt
32. ΔECN+EGF(10-12) 1 10 12 47 22
33. ΔCla+XEGF(10-13) 7 10 13 31 42 20

FIG.2B

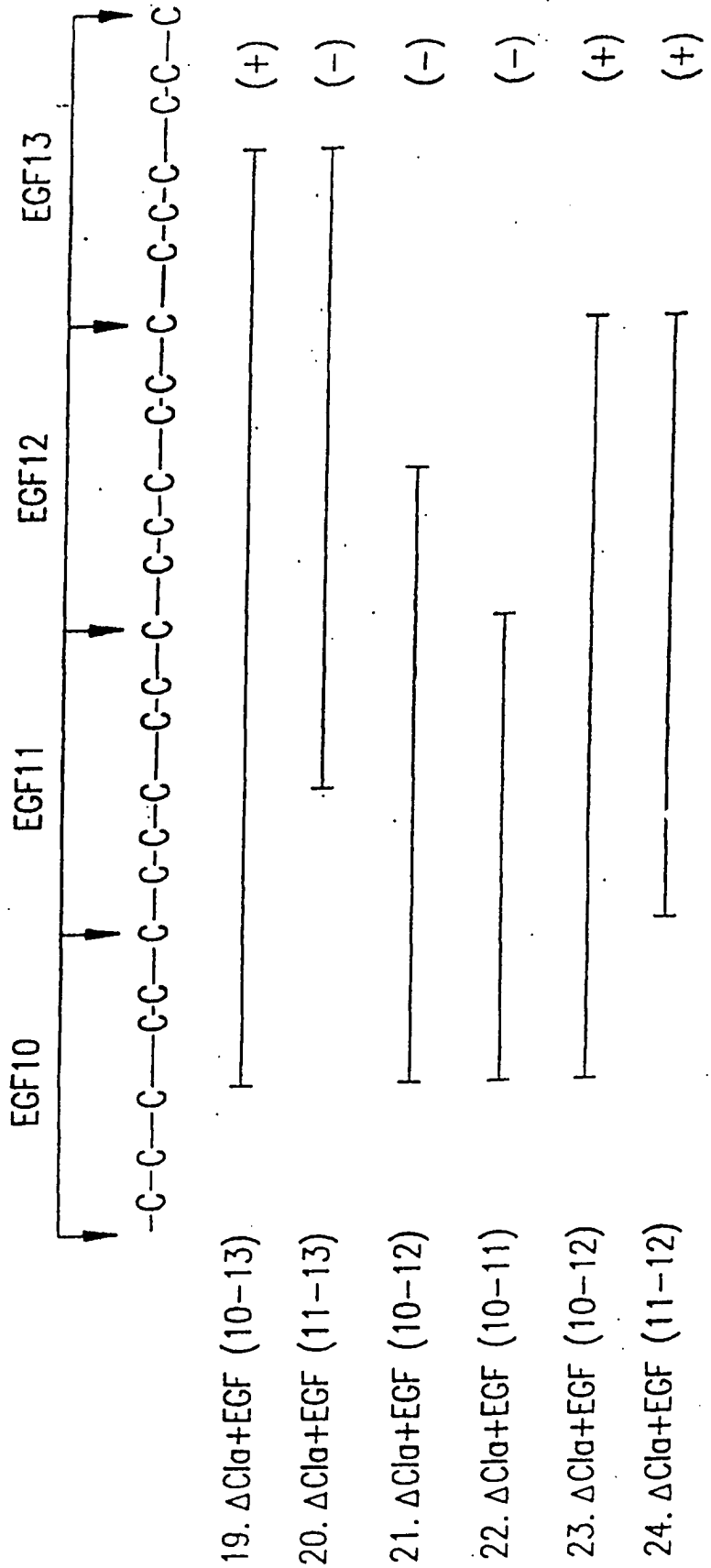

EGF10
EGF11
EGF12
EGF13
19. ΔCla+EGF (10–13) (+)
20. ΔCla+EGF (11–13) (–)
21. ΔCla+EGF (10–12) (–)
22. ΔCla+EGF (10–11) (–)
23. ΔCla+EGF (10–12) (+)
24. ΔCla+EGF (11–12) (+)

FIG.3

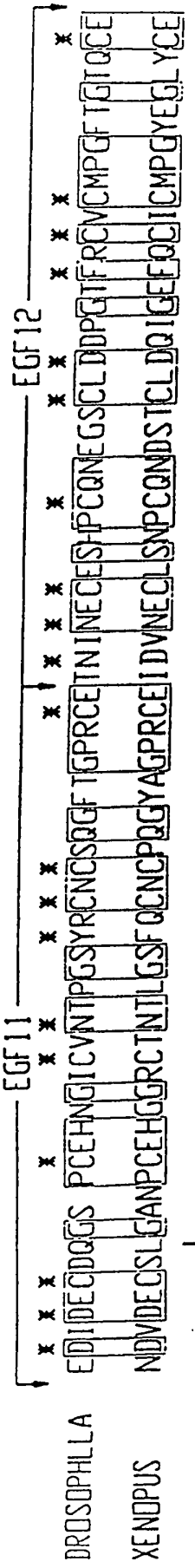


FIG. 4

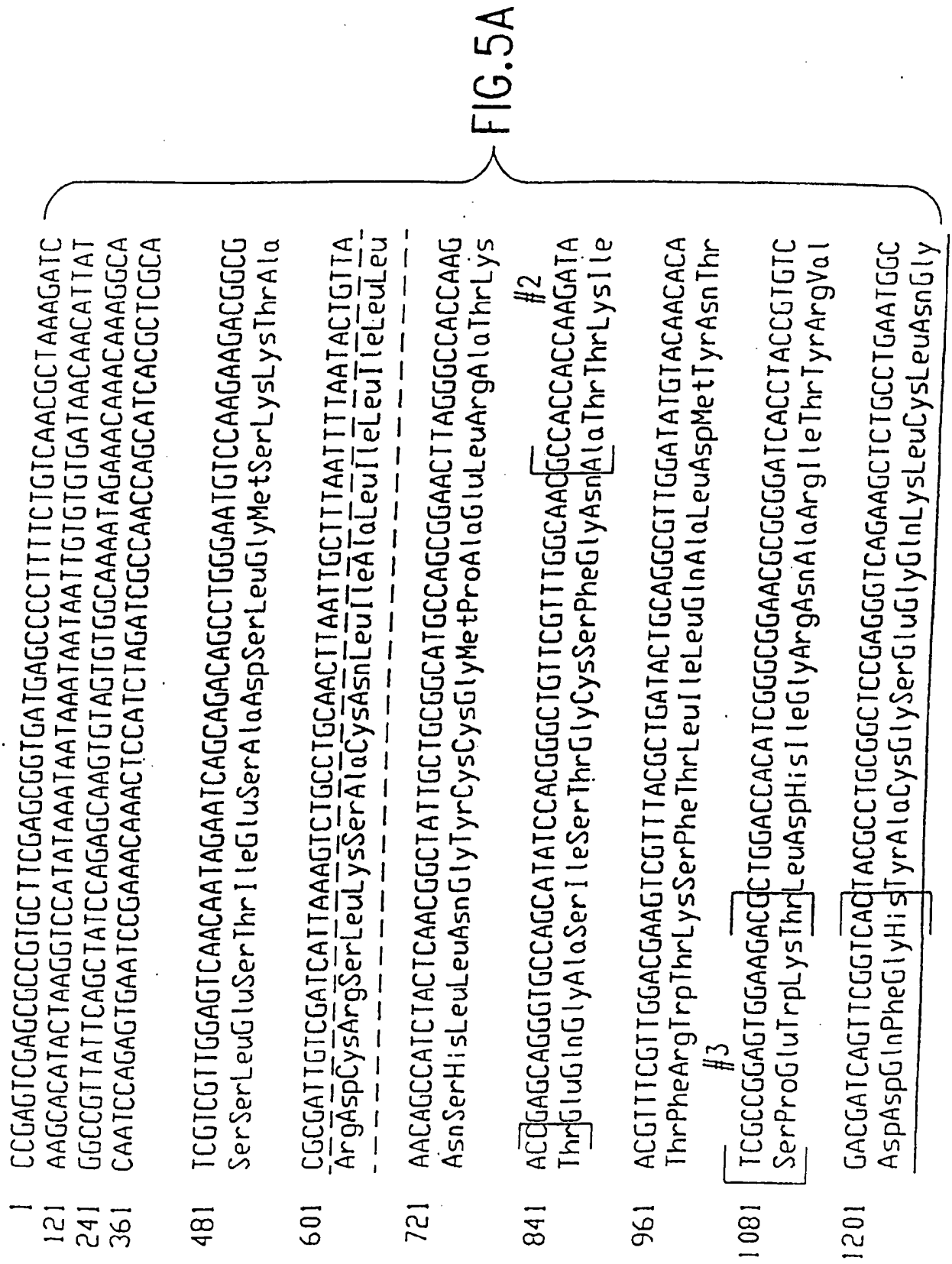

1 CCGAGTCGAGCGCCGTGCTTCGAGCGGTGATGAGCCCCTTTTCTGTCAACGCTAAAGATC
121 AAGCACATACTAAGGTCCATATAAATAATAAATAATTGTGTGTGATAACAACATTAT
241 GGCCGTTATTCAGCTATCCAGAGCAAGTGTAGTGTGGCAAAATAGAAACAAACAAAGGCA
361 CAATCCAGAGTGAATCCGAAACAAACTCCATCTAGATCGCCAACCAGCATCACGCTCGCA
481 TCGTCGTTGGAGTCAACAATAGAATCAGCAGACAGCCTGGGAATGTCCAAGAAGACGGCG
SerSerLeuGluSerThrIleGluSerAlaAspSerLeuGlyMetSerLysLysThrAla
601 CGCGATTGTCGATCATTAAAGTCTGCCTGCAACTTAATTGCTTTAATTTTAATACTGTTA
ArgAspCysArgSerLeuLysSerAlaCysAsnLeuIleAlaLeuIleLeuIleLeuLeu
721 AACAGCCATCTACTCAACGGCTATTGCTGCGGCATGCCAGCGGAACTTAGGGCCACCAAG
AsnSerHisLeuLeuAsnGlyTyrCysCysGlyMetProAlaGluLeuArgAlaThrLys
#2
841 ACCGAGCAGGGTGCCAGCATATCCACGGGCTGTTCGTTGGCAACGCCACCACCAAGATA
ThrGluGlnGlyAlaSerIleSerThrGlyCysSerPheGlyAsnAlaThrThrLysIle
961 ACGTTCGTTGGACGAAGTCGTTACGCTGATACTGCAGGCGTTGGATATGTACAACACA
ThrPheArgTrpThrLysSerPheThrLeuIleLeuGlnAlaLeuAspMetTyrAsnThr
#3
1081 TCGCCGGAGTGGAAGACGCTGGACCACATCGGGCGGAACGCGCGGATCACCTACCGTGTC
SerProGluTrpLysThrLeuAspHisIleGlyArgAsnAlaArgIleThrTyrArgVal
1201 GACGATCAGTTCGGTCAGTACGCCTGCGGCTCCGAGGGTCAGAAGCTCTGCCTGAATGGC
AspAspGlnPheGlyHisTyrAlaCysGlySerGluGlyGlnLysLeuCysLeuAsnGly

FIG.5A

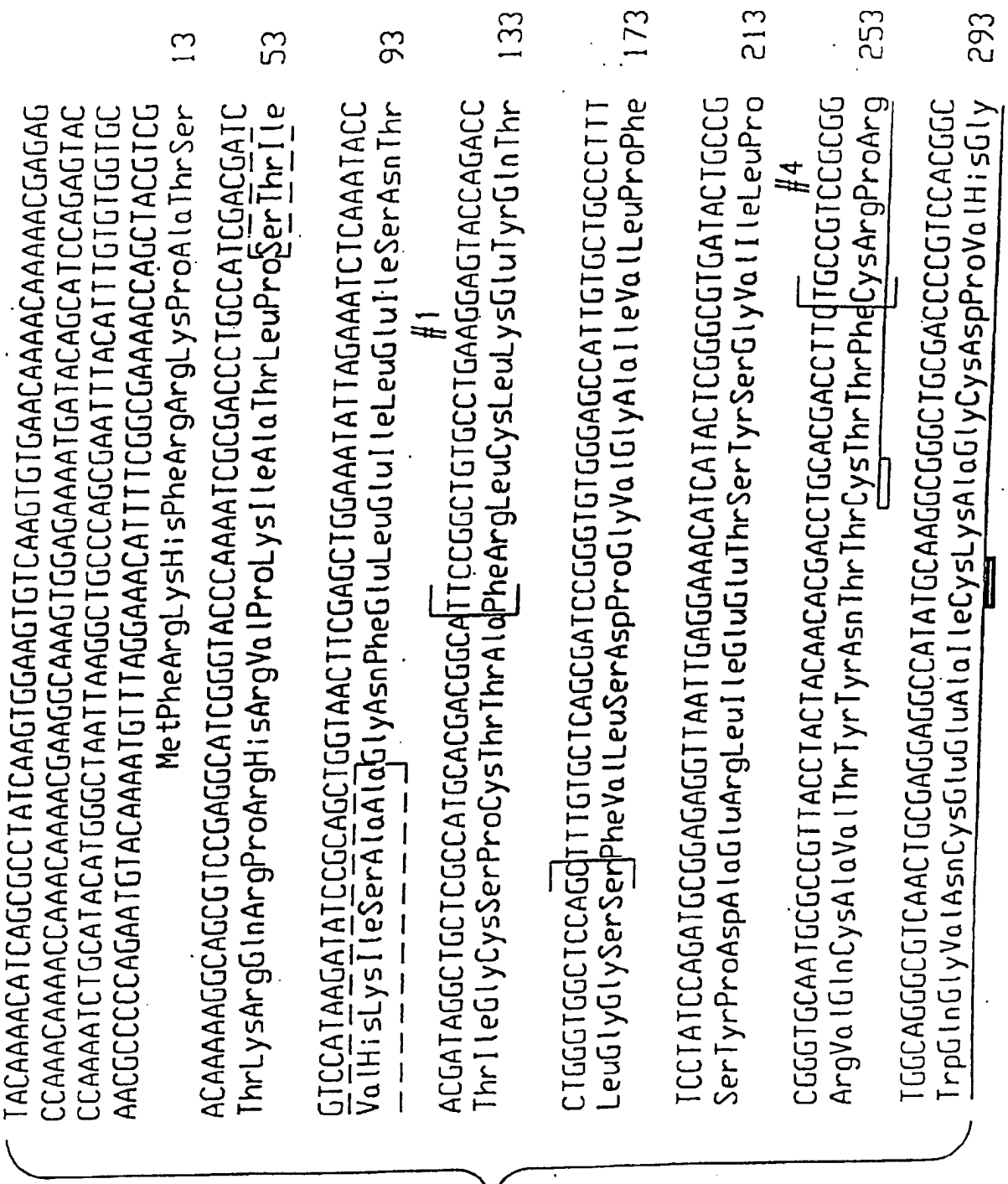

TACAAAACATCAGCGCCTATCAAGTGGAAGTGTCAAGTGTGAACAAAACAAAAACGAGAG
CCAAACAAAACCAAACAAAACGAAGGCAAAGTGGAGAAAATGATACAGCATCCAGAGTAC
CCAAAATCTGCATACATGGGCTAATTAAGGCTGCCCAGCGAATTTACATTTGTGTGGTGC
AACGCCCCCAGAATGTTACAAAATGTTTAGGAAACATTTTCGGCGAAAACCAGCTACGTCG
MetPheArgLysHisPheArgArgLysProAlaThrSer
13
ACAAAAAGGCAGCGTCCGAGGCATCGGGTACCCAAAATCGCGACCCTGCCATCGACGATC
ThrLysArgGlnArgProArgHisArgValProLysIleAlaThrLeuProSerThrIle
53
GTCCATAAGATATCCGCAGCTGGTAACTTCGAGCTGGAAATATTAGAAATCTCAAATACC
ValHisLysIleSerAlaAlaGlyAsnPheGluLeuGluIleLeuGluIleSerAsnThr
93
#1
ACGATAGGCTGCTCGCCATGCACGACGGCATTCCGGCTGTGCCTGAAGGAGTACCAGACC
ThrIleGlyCysSerProCysThrThrAlaPheArgLeuCysLeuLysGluTyrGlnThr
133
CTGGGTGGCTCCAGCTTTGTGCTCAGCGATCCGGGTGTGGGAGCCATTGTGCTGCCCTTT
LeuGlyGlySerSerPheValLeuSerAspProGlyValGlyAlaIleValLeuProPhe
173
TCCTATCCAGATGCGGAGAGGTTAATTGAGGAAACATCATACTCGGGCGTGATACTGCCG
SerTyrProAspAlaGluArgLeuIleGluGluThrSerTyrSerGlyValIleLeuPro
213
#4
CGGGTGCAATGCGCCGTTACCTACTACAACACGACCTGCACGACCTTCTGCCGTCCGCGG
ArgValGlnCysAlaValThrTyrTyrAsnThrThrCysThrThrPheCysArgProArg
253
TGGCAGGGCGTCAACTGCGAGGAGGCCATATGCAAGGCGGGCTGCGACCCCGTCCACGGC
TrpGlnGlyValAsnCysGluGluAlaIleCysLysAlaGlyCysAspProValHisGly
293

FIG.5B

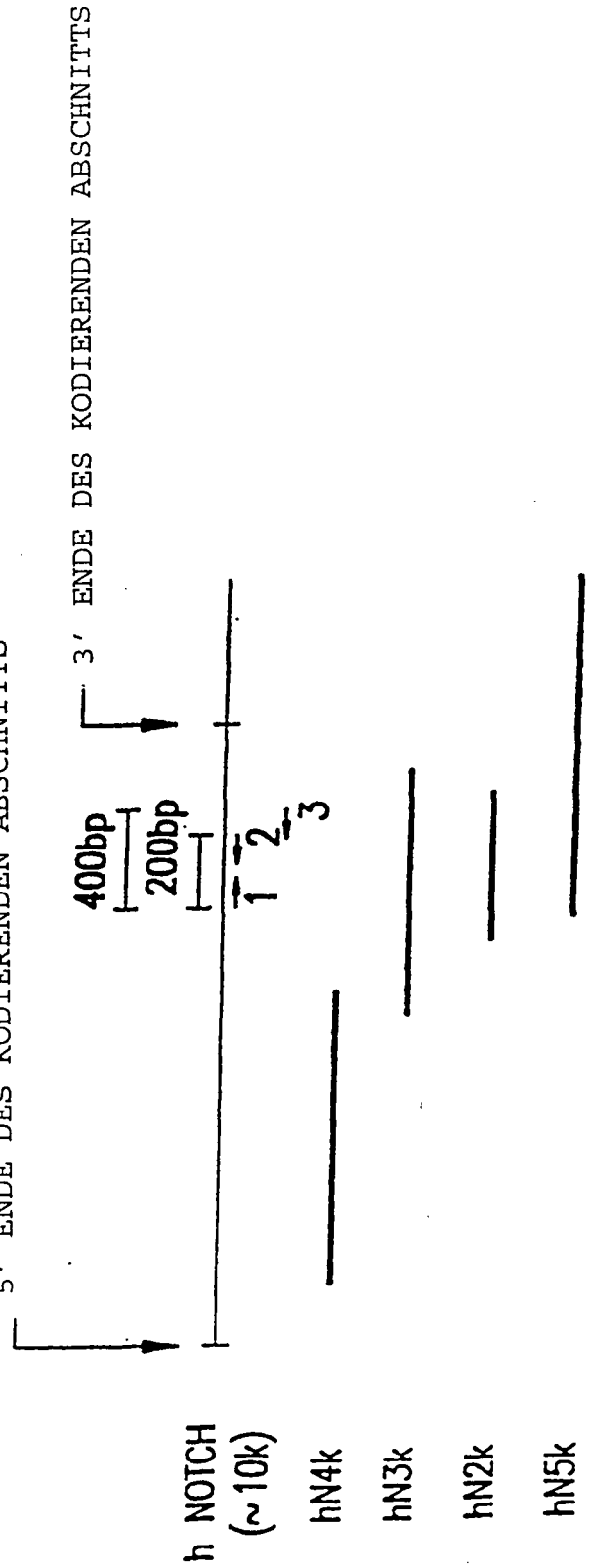
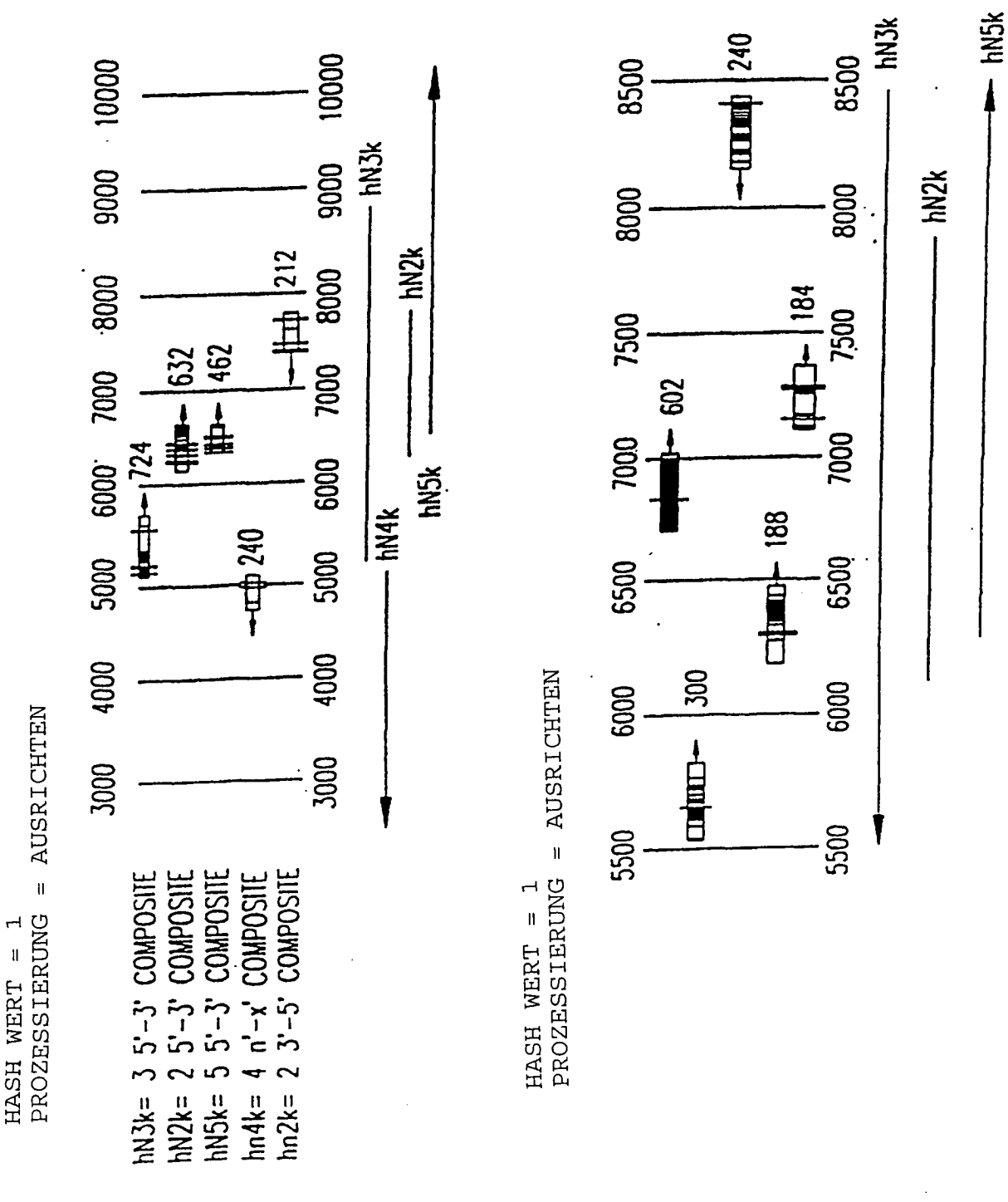

5' ENDE DES KODIERENDEN ABSCHNITTS
3' ENDE DES KODIERENDEN ABSCHNITTS
400bp
200bp
1
2
3
h NOTCH (~10k)
hN4k
hN3k
hN2k
hN5k

FIG.6

FIG.7


HASH WERT = 1
PROZESSIERUNG = AUSRICHTEN
hN3k= 3 5'-3' COMPOSITE
hN2k= 2 5'-3' COMPOSITE
hN5k= 5 5'-3' COMPOSITE
hn4k= 4 n'-x' COMPOSITE
hn2k= 2 3'-5' COMPOSITE
3000
4000
5000
6000
7000
8000
9000
10000
724
632
462
240
212
hN4k
hN3k
hN2k
hN5k
HASH WERT = 1
PROZESSIERUNG = AUSRICHTEN
5500
6000
6500
7000
7500
8000
8500
300
188
602
184
240
hN3k
hN2k
hN5k

```
  1  GAATTCCGCT GGGAGAATGG TCTGAGCTAC CTGCCCGTCC TGCTGGGGCA TCAATGGCAA
 61  GTGGGGAAAG CCACACTGGG CAAACGGGCC AGGCCATTTC TGGAATGTGG TACATGGTGG
121  GCAGGGGGCC CGCAACAGCT GGAGGGCAGG TGGACTGAGG CTGGGGATCC CCCGCTGGTT
181  GGGCAATACT GCCTTTACCC ATGAGCTGGA AAGTCACAAT GGGGGGCAAG GGCTCCCGAG
241  GGTGGTTATG TGCTTCCTTC AGGTGGC
```

FIG.8A

```
  1  GAATTCCTTC CATTATACGT GACTTTTCTG AAACTGTAGC CACCCTAGTG TCTCTAACTC
 61  CCTCTGGAGT TTGTCAGCTT TGGTCTTTTC AAAGAGCAGG CTCTCTTCAA GCTCCTTAAT
121  GCGGGCATGC TCCAGTTTGG TCTGCGTCTC AAGATCACCT TTGGTAATTG ATTCTTCTTC
181  AACCCGGAAC TGAAGGCTGG CTCTCACCCT CTAGGCAGAG CAGGAATTCC GAGGTGGATG
241  TGTTAGATGT GAATGTCCGT GGCCCAGATG GCTGCACCCC ATTGATGTTG GCTTCTCTCC
301  GAGGAGGCAG CTCAGATTTG AGTGATGAAG ATGAAGATGC AGAGGACTGT TCTGCTAACA
361  TCATCACAGA CTTGGTCTAC CAGGGTGCCA GCCTCCAGAC CAGACAGACC GGACTGGTGA
421  GATGGCCCTG CACCTTGCAG CCCGCTACTC ACGGGCTGAT GCTGCCAAGC GTCTCCTGGA
481  TGCAGGTGCA GATGCCAATG CCCAGGACAA CATGGGCCGC TGTCCACTCC ATGCTGCAGT
541  GGCACGTGAT GCCAAGGTGT ATTCAGATCT GTTA
```

FIG.8B

```
  1  TCCAGATTCT GATTCGCAAC CGAGTAACTG ATCTAGATGC CAGGATGAAT GATGGTACTA
 61  CACCCCTGAT CCTGGCTGCC CGCCTGGCTG TGGAGGGAAT GGTGGCAGAA CTGATCAACT
121  GCCAAGCGGA TGTGAATGCA GTGGATGACC ATGGAAAATC TGCTCTTCAC TGGGCAGCTG
181  CTGTCAATAA TGTGGAGGCA ACTCTTTTGT TGTTGAAAAA TGGGGCCAAC CGAGACATGC
241  AGGACAACAA GGAAGAGACA CCTCTGTTTC TTGCTGCCCG GGAGGAGCTA TAAGC
```

FIG.8C

```
  1  GAATTCCATT CAGGAGGAAA GGGTGGGGAG AGAAGCAGGC ACCCACTTTC CCGTGGCTGG
 61  ACTCGTTCCC AGGTGGCTCC ACCGGCAGCT GTGACCGCCG CAGGTGGGGG CGGAGTGCCA
121  TTCAGAAAAT TCCAGAAAAG CCCTACCCCA ACTCGGACGG CAACGTCACA CCCGTGGGTA
181  GCAACTGGCA CACAAACAGC CAGCGTGTCT GGGGCACGGG GGGATGGCAC CCCCTGCAGG
241  CAGAGCTG
```

FIG.9A

```
  1  CTAAAGGGAA CAAAAGCNGG AGCTCCACCG CGGGCGGCNC NGCTCTAGAA CTAGTGGANN
 61  NCCCGGGCTG CAGGAATTCC GGCGGACTGG GCTCGGGCTC AGAGCGGCGC TGTGGAAGAG
121  ATTCTAGACC GGGAGAACAA GCGAATGGCT GACAGCTGGC CTCCAAAGTC ACCAGGCTCA
181  AATCGCTCGC CCTGGACATC GAGGGATGCA GAGGATCAGA ACCGGTACCT GGATGGCATG
241  ACTCGGATTT ACAAGCATGA CCAGCCTGCT TACAGGGAGC GTGANNTTTT CACATGCAGT
301  CGACAGACAC GAGCTCTATG CAT
```

FIG.9B

```
          10                  20                  30                  40
     *    *    *    *    *    *    *    *    *    *    *    *    *
TGC CAG GAG GAC GCG GGC AAC AAG GTC TGC AGC CTG CAG TGC AAC AAC
 C   Q   E   D   A   G   N   K   V   C   S   L   Q   C   N   N>

50                  60                  70                  80                  90
*    *    *    *    *    *    *    *    *    *    *    *    *    *    *
CAC GCG TGC GGC TGG GAC GGC GGT GAC TGC TCC CTC AAC TTC AAT GAC
 H   A   C   G   W   D   G   G   D   C   S   L   N   F   N   D>

  100                 110                 120                 130                 140
    *    *    *    *    *    *    *    *    *    *    *    *    *
CCC TGG AAG AAC TGC ACG CAG TCT CTG CAG TGC TGG AAG TAC TTC AGT
 P   W   K   N   C   T   Q   S   L   Q   C   W   K   Y   F   S>

     150                 160                 170                 180                 190
*    *    *    *    *    *    *    *    *    *    *    *    *    *    *
GAC GGC CAC TGT GAC AGC CAG TGC AAC TCA GCC GGC TGC CTC TTC GAC
 D   G   H   C   D   S   Q   C   N   S   A   G   C   L   F   D>
```

FIG. 10A

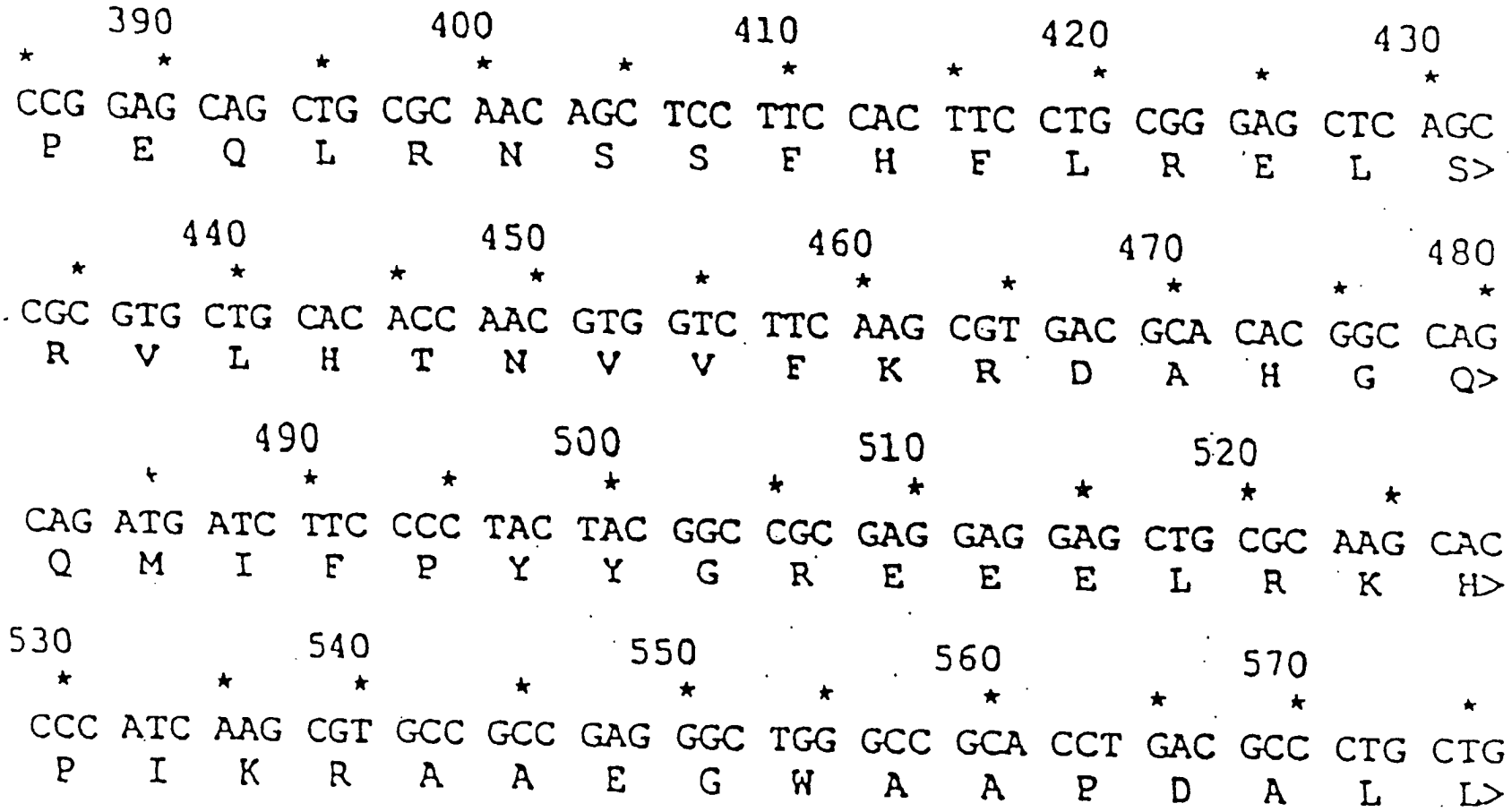
```
        200         210         220         230         240
  *     *     *     *     *     *     *     *     *     *
GGC TTT GAC TGC CAG CGT GCG GAA GGC CAG TGC AAC CCC CTG TAC GAC
 G   F   D   C   Q   R   A   E   G   Q   C   N   P   L   Y   D>

          250         260         270         280
     *     *     *     *     *     *     *     *     *
CAG TAC TGC AAG GAC CAC TTC AGC GAC GGG CAC TGC GAC CAG GGC TGC
 Q   Y   C   K   D   H   F   S   D   G   H   C   D   Q   G   C>

290         300         310         320         330
 *     *     *     *     *     *     *     *     *     *
AAC AGC GCG GAG TGC GAG TGG GAC GGG CTG GAC TGT GCG GAG CAT GTA
 N   S   A   E   C   E   W   D   G   L   D   C   A   E   H   V>

  340         350         360         370         380
    *     *     *     *     *     *     *     *     *
CCC GAG AGG CTG GCG GCC GGC ACG CTG GTG GTG GTG GTG CTG ATG CCG
 P   E   R   L   A   A   G   T   L   V   V   V   V   L   M   P>
```

FIG.10B



DE 693 31 832 T3 2009.07.16

```
  580         590         600         610         620
   *     *     *     *     *     *     *     *     *
GGC CAG GTG AAG GCC TCG CTG CTC CCT GGT GGC AGC GAG GGT GGG CGG
 G   Q   V   K   A   S   L   L   P   G   G   S   E   G   G   R>

    630         640         650         660         670
*    *     *     *     *     *     *     *     *     *
CGG CGG AGG GAG CTG GAC CCC ATG GAC GTC CGC GGC TCC ATC GTC TAC
 R   R   R   E   L   D   P   M   D   V   R   G   S   I   V   Y>

      680         690         700         710         720
  *     *     *     *     *     *     *     *     *     *
CTG GAG ATT GAC AAC CGG CAG TGT GTG CAG GCC TCC TCG CAG TGC TTC
 L   E   I   D   N   R   Q   C   V   Q   A   S   S   Q   C   F>

        730         740         750         760
    *     *     *     *     *     *     *     *     *
CAG AGT GCC ACC GAC GTG GCC GCA TTC CTG GGA GCG CTC GCC TCG CTG
 Q   S   A   T   D   V   A   A   F   L   G   A   L   A   S   L>
```

FIG. 10D

```
770         780         790         800         810
 *    *    *    *    *    *    *    *    *    *
GGC AGC CTC AAC ATC CCC TAC AAG ATC GAG GCC GTG CAG AGT GAG ACC
 G   S   L   N   I   P   Y   K   I   E   A   V   Q   S   E   T>

  820         830         840         850         860
   *    *    *    *    *    *    *    *    *    *
GTG GAG CCG CCC CCG CCG GCG CAG CTG CAC TTC ATG TAC GTG GCG GCG
 V   E   P   P   P   P   A   Q   L   H   F   M   Y   V   A   A>

    870         880         890         900         910
 *    *    *    *    *    *    *    *    *    *
GCC GCC TTT GTG CTT CTG TTC TTC GTG GGC TGC GGG GTG CTG CTG TCC
 A   A   F   V   L   L   F   F   V   G   C   G   V   L   L   S>

       920         930         940         950         960
  *    *    *    *    *    *    *    *    *    *
CGC AAG CGC CGG CGG CAG CAT GGC CAG CTC TGG TTC CCT GAG GGC TTC
 R   K   R   R   R   Q   H   G   Q   L   W   F   P   E   G   F>
```

FIG.10E

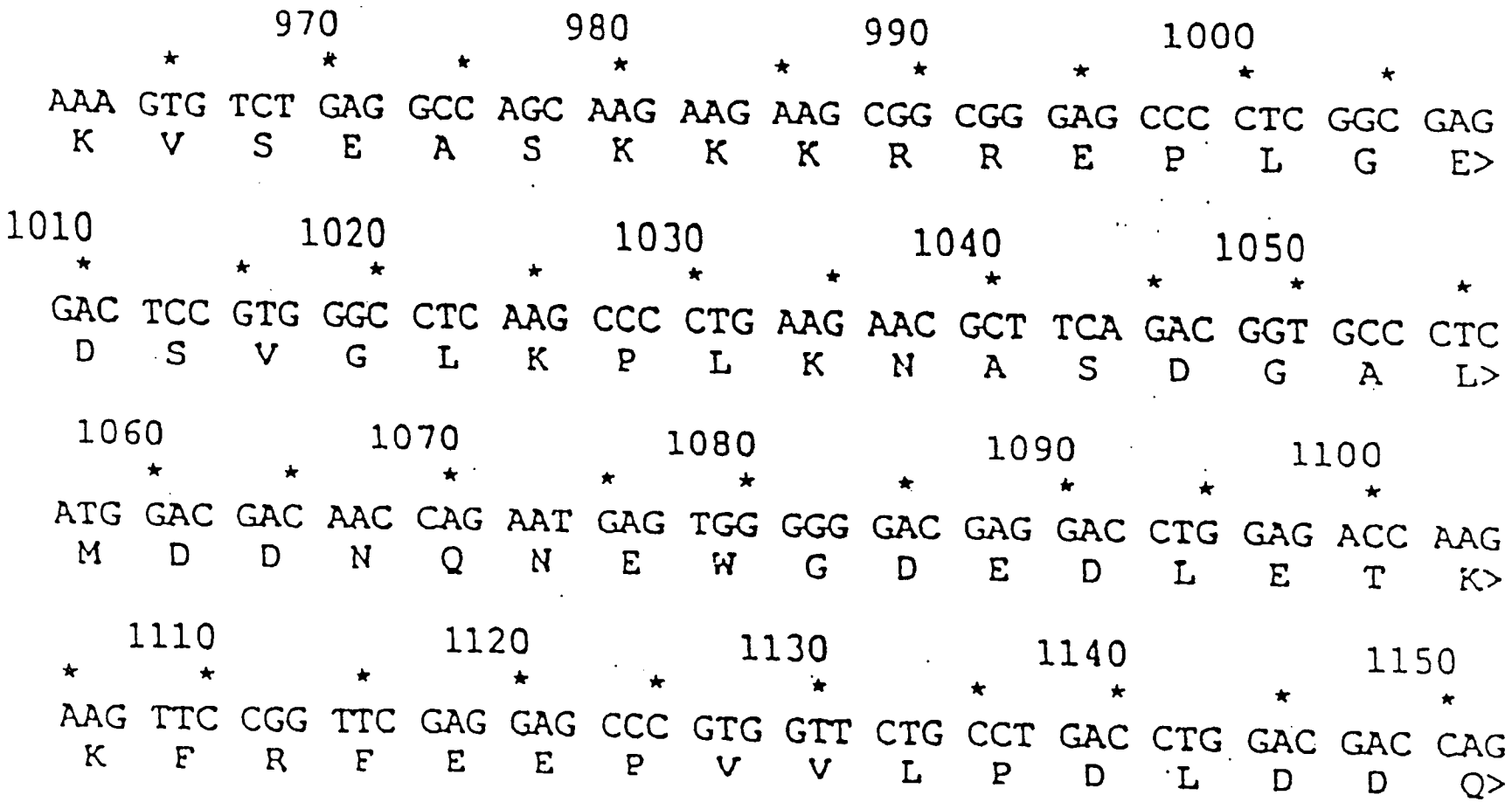

970 980 990 1000
AAA GTG TCT GAG GCC AGC AAG AAG AAG CGG CGG GAG CCC CTC GGC GAG
K V S E A S K K K R R E P L G E>
1010 1020 1030 1040 1050
GAC TCC GTG GGC CTC AAG CCC CTG AAG AAC GCT TCA GAC GGT GCC CTC
D S V G L K P L K N A S D G A L>
1060 1070 1080 1090 1100
ATG GAC GAC AAC CAG AAT GAG TGG GGG GAC GAG GAC CTG GAG ACC AAG
M D D N Q N E W G D E D L E T K>
1110 1120 1130 1140 1150
AAG TTC CGG TTC GAG GAG CCC GTG GTT CTG CCT GAC CTG GAC GAC CAG
K F R F E E P V V L P D L D D Q>

FIG.10F

```
      1160         1170         1180         1190         1200
 *     *     *     *     *     *     *     *     *     *
ACA GAC CAC CGG CAG TGG ACT CAG CAG CAC CTG GAT GCC GCT GAC CTG
 T   D   H   R   Q   W   T   Q   Q   H   L   D   A   A   D   L>

          1210         1220         1230         1240
     *     *     *     *     *     *     *     *     *
CGC ATG TCT GCC ATG GCC CCC ACA CCG CCC CAG GGT GAG GTT GAC GCC
 R   M   S   A   M   A   P   T   P   P   Q   G   E   V   D   A>

1250         1260         1270         1280         1290
 *     *     *     *     *     *     *     *     *     *
GAC TGC ATG GAC GTC AAT GTC CGC GGG CCT GAT GGC TTC ACC CCG CTC
 D   C   M   D   V   N   V   R   G   P   D   G   F   T   P   L>

 1300         1310         1320         1330         1340
    *     *     *     *     *     *     *     *     *
ATG ATC GCC TCC TGC AGC GGG GGC GGC CTG GAG ACG GGC AAC AGC GAG
 M   I   A   S   C   S   G   G   G   L   E   T   G   N   S   E>
```

FIG.10G

```
   1350          1360          1370          1380          1390
*     *     *     *     *     *     *     *     *     *     *
GAA GAG GAG GAC GCG CCG GCC GTC ATC TCC GAC TTC ATC TAC CAG GGC
 E   E   E   D   A   P   A   V   I   S   D   F   I   Y   Q   G>

      1400          1410          1420          1430          1440
 *     *     *     *     *     *     *     *     *     *     *
GCC AGC CTG CAC AAC CAG ACA GAC CGC ACG GGC GAG ACC GCC TTG CAC
 A   S   L   H   N   Q   T   D   R   T   G   E   T   A   L   H>

          1450          1460          1470          1480
      *     *     *     *     *     *     *     *     *     *
CTG GCC GCC CGC TAC TCA CGC TCT GAT GCC GCC AAG CGC CTG CTG GAG
 L   A   A   R   Y   S   R   S   D   A   A   K   R   L   L   E>

1490          1500          1510          1520          1530
 *     *     *     *     *     *     *     *     *     *     *
GCC AGC GCA GAT GCC AAC ATC CAG GAC AAC ATG GGC CGC ACC CCG CTG
 A   S   A   D   A   N   I   Q   D   N   M   G   R   T   P   L>
```

FIG. 10H

```
1540          1550          1560          1570          1580
   *     *       *     *       *     *       *     *       *
CAT GCG GCT GTG TCT GCC GAC GCA CAA GGT GTC TTC CAG ATC CTG ATC
 H   A   A   V   S   A   D   A   Q   G   V   F   Q   I   L   I>

   1590          1600          1610          1620          1630
*     *       *     *       *     *       *     *       *     *
CGG AAC CGA GCC ACA GAC CTG GAT GCC CGC ATG CAT GAT GGC ACG ACG
 R   N   R   A   T   D   L   D   A   R   M   H   D   G   T   T>

      1640          1650          1660          1670          1680
  *     *       *     *       *     *       *     *       *     *
CCA CTG ATC CTG GCT GCC CGC CTG GCC GTG GAG GGC ATG CTG GAG GAC
 P   L   I   L   A   A   R   L   A   V   E   G   M   L   E   D>

         1690          1700          1710          1720
     *     *       *     *       *     *       *     *       *
CTC ATC AAC TCA CAC GCC GAC GTC AAC GCC GTA GAT GAC CTG GGC AAG
 L   I   N   S   H   A   D   V   N   A   V   D   D   L   G   K>

1730          1740          1750          1760          1770
   *     *       *     *       *     *       *     *       *
TCC GCC CTG CAC TGG GCC GCC GCC GTG AAC AAT GTG GAT GCC GCA GTT
 S   A   L   H   W   A   A   A   V   N   N   V   D   A   A   V>
```

FIG. 10I

```
1780          1790          1800          1810          1820
   *     *       *     *       *     *       *     *       *
GTG CTC CTG AAG AAC GGG GCT AAC AAA GAT ATG CAG AAC AAC AGG GAG
 V   L   L   K   N   G   A   N   K   D   M   Q   N   N   R   E>

  1830          1840          1850          1860          1870
*     *       *     *       *     *       *     *       *     *
GAG ACA CCC CTG TTT CTG GCC GCC CGG GAG GGC AGC TAC GAG ACC GCC
 E   T   P   L   F   L   A   A   R   E   G   S   Y   E   T   A>

      1880          1890          1900          1910          1920
  *       *     *       *     *       *     *       *     *       *
AAG GTG CTG CTG GAC CAC TTT GCC AAC CGG GAC ATC ACG GAT CAT ATG
 K   V   L   L   D   H   F   A   N   R   D   I   T   D   H   M>

          1930          1940          1950          1960
    *       *     *       *     *       *     *       *     *
GAC CGC CTG CCG CGC GAC ATC GCA CAG GAG CGC ATG CAT CAC GAC ATC
 D   R   L   P   R   D   I   A   Q   E   R   M   H   H   D   I>
```

FIG.10J

```
1970         1980         1990         2000         2010
   *    *    *    *    *    *    *    *    *    *    *
GTG AGG CTG CTG GAC GAG TAC AAC CTG GTG CGC AGC CCG CAG CTG CAC
 V   R   L   L   D   E   Y   N   L   V   R   S   P   Q   L   H>

2020         2030         2040         2050         2060
   *    *    *    *    *    *    *    *    *    *    *
GGA GCC CCG CTG GGG GGC ACG CCC ACC CTG TCG CCC CCG CTC TGC TCG
 G   A   P   L   G   G   T   P   T   L   S   P   P   L   C   S>

2070         2080         2090         2100         2110
   *    *    *    *    *    *    *    *    *    *    *
CCC AAC GGC TAC CTG GGC AGC CTC AAG CCC GGC GTG CAG GGC AAG AAG
 P   N   G   Y   L   G   S   L   K   P   G   V   Q   G   K   K>

2120         2130         2140         2150         2160
   *    *    *    *    *    *    *    *    *    *    *
GTC CGC AAG CCC AGC AGC AAA GGC CTG GCC TGT GGA AGC AAG GAG GCC
 V   R   K   P   S   S   K   G   L   A   C   G   S   K   E   A>
```

FIG. 10K

```
          2170          2180          2190          2200
      *     *     *     *     *     *     *     *     *
AAG GAC CTC AAG GCA CGG AGG AAG AAG TCC CAG GAT GGC AAG GGC TGC
 K   D   L   K   A   R   R   K   K   S   Q   D   G   K   G   C>

2210          2220          2230          2240          2250
   *     *     *     *     *     *     *     *     *     *
CTG CTG GAC AGC TCC GGC ATG CTC TCG CCC GTG GAC TCC CTG GAG TCA
 L   L   D   S   S   G   M   L   S   P   V   D   S   L   E   S>

 2260          2270          2280          2290          2300
    *     *     *     *     *     *     *     *     *     *
CCC CAT GGC TAC CTG TCA GAC GTG GCC TCG CCG CCA CTG CTG CCC TCC
 P   H   G   Y   L   S   D   V   A   S   P   P   L   L   P   S>

    2310          2320          2330          2340          2350
*     *     *     *     *     *     *     *     *     *     *
CCG TTC CAG CAG TCT CCG TCC GTG CCC CTC AAC CAC CTG CCT GGG ATG
 P   F   Q   Q   S   P   S   V   P   L   N   H   L   P   G   M>
```

FIG.10L

```
        2360          2370          2380          2390          2400
   *     *     *     *     *     *     *     *     *     *     *
CCC GAC ACC CAC CTG GGC ATC GGG CAC CTG AAC GTG GCG GCC AAG CCC
 P   D   T   H   L   G   I   G   H   L   N   V   A   A   K   P>

          2410          2420          2430          2440
     *     *     *     *     *     *     *     *     *     *
GAG ATG GCG GCG CTG GGT GGG GGC GGC CGG CTG GCC TTT GAG ACT GGC
 E   M   A   A   L   G   G   G   G   R   L   A   F   E   T   G>

2450          2460          2470          2480          2490
   *     *     *     *     *     *     *     *     *     *     *
CCA CCT CGT CTC TCC CAC CTG CCT GTG GCC TCT GGC ACC AGC ACC GTC
 P   P   R   L   S   H   L   P   V   A   S   G   T   S   T   V>

 2500          2510          2520          2530          2540
    *     *     *     *     *     *     *     *     *     *
CTG GGC TCC AGC AGC GGA GGG GCC CTG AAT TTC ACT GTG GGC GGG TCC
 L   G   S   S   S   G   G   A   L   N   F   T   V   G   G   S>
```

FIG.10M

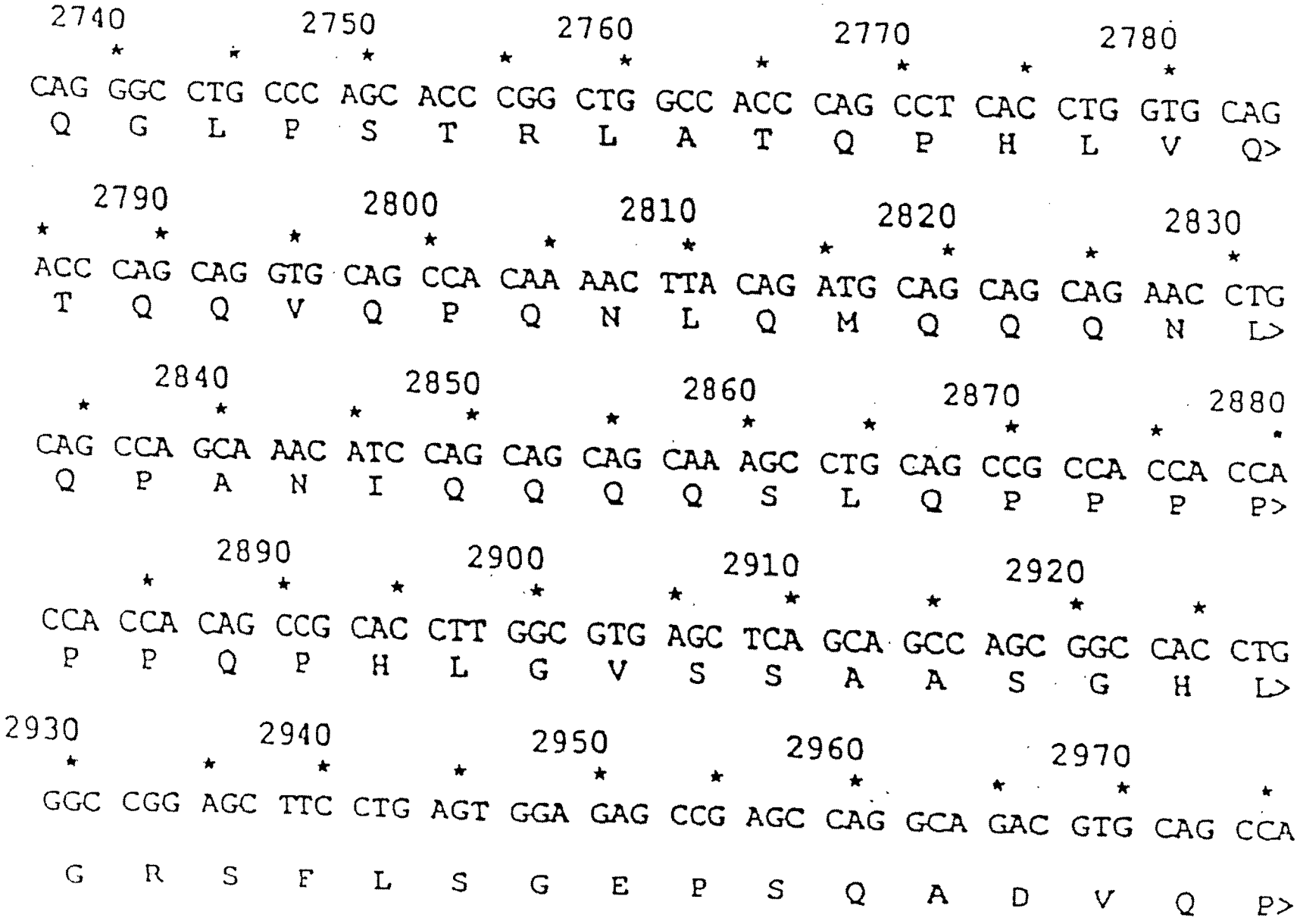
```
   2550          2560          2570          2580          2590
*     *       *     *       *     *       *     *       *     *
ACC AGT TTG AAT GGT CAA TGC GAG TGG CTG TCC CGG CTG CAG AGC GGC
 T   S   L   N   G   Q   C   E   W   L   S   R   L   Q   S   G>

      2600          2610          2620          2630          2640
  *     *       *     *       *     *       *     *       *     *
ATG GTG CCG AAC CAA TAC AAC CCT CTG CGG GGG AGT GTG GCA CCA GGC
 M   V   P   N   Q   Y   N   P   L   R   G   S   V   A   P   G>

          2650          2660          2670          2680
      *     *       *     *       *     *       *     *       *
CCC CTG AGC ACA CAG GCC CCC TCC CTG CAG CAT GGC ATG GTA GGC CCG
 P   L   S   T   Q   A   P   S   L   Q   H   G   M   V   G   P>

2690          2700          2710          2720          2730
  *     *       *     *       *     *       *     *       *     *
CTG CAC AGT AGC CTT GCT GCC AGC GCC CTG TCC CAG ATG ATG AGC TAC
 L   H   S   S   L   A   A   S   A   L   S   Q   M   M   S   Y>
```

FIG. 10N


2740 2750 2760 2770 2780
CAG GGC CTG CCC AGC ACC CGG CTG GCC ACC CAG CCT CAC CTG GTG CAG
Q G L P S T R L A T Q P H L V Q>
2790 2800 2810 2820 2830
ACC CAG CAG GTG CAG CCA CAA AAC TTA CAG ATG CAG CAG CAG AAC CTG
T Q Q V Q P Q N L Q M Q Q Q N L>
2840 2850 2860 2870 2880
CAG CCA GCA AAC ATC CAG CAG CAG CAA AGC CTG CAG CCG CCA CCA CCA
Q P A N I Q Q Q Q S L Q P P P P>
2890 2900 2910 2920
CCA CCA CAG CCG CAC CTT GGC GTG AGC TCA GCA GCC AGC GGC CAC CTG
P P Q P H L G V S S A A S G H L>
2930 2940 2950 2960 2970
GGC CGG AGC TTC CTG AGT GGA GAG CCG AGC CAG GCA GAC GTG CAG CCA
G R S F L S G E P S Q A D V Q P>
FIG.10 0

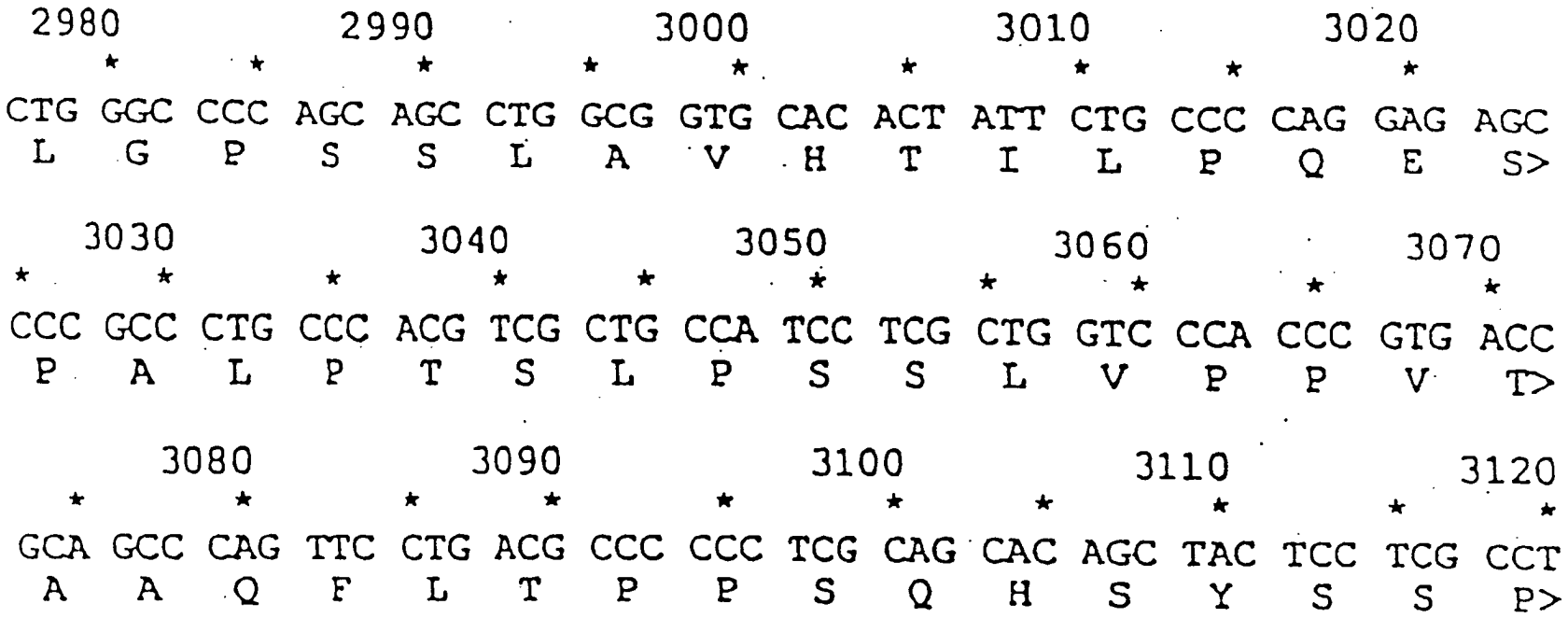

2980 2990 3000 3010 3020
CTG GGC CCC AGC AGC CTG GCG GTG CAC ACT ATT CTG CCC CAG GAG AGC
L G P S S L A V H T I L P Q E S>
3030 3040 3050 3060 3070
CCC GCC CTG CCC ACG TCG CTG CCA TCC TCG CTG GTC CCA CCC GTG ACC
P A L P T S L P S S L V P P V T>
3080 3090 3100 3110 3120
GCA GCC CAG TTC CTG ACG CCC CCC TCG CAG CAC AGC TAC TCC TCG CCT
A A Q F L T P P S Q H S Y S S P>

FIG.10P

```
          3130          3140          3150          3160
     *     *     *     *     *     *     *     *     *
GTG GAC AAC ACC CCC AGC CAC CAG CTA CAG GTG CCT GTT CCT GTA ATG
 V   D   N   T   P   S   H   Q   L   Q   V   P   V   P   V   M>

3170          3180          3190          3200          3210
 *     *     *     *     *     *     *     *     *     *     *
GTA ATG ATC CGA TCT TCG GAT CCT TCT AAA GGC TCA TCA ATT TTG ATC
 V   M   I   R   S   S   D   P   S   K   G   S   S   I   L   I>

 3220          3230
    *     *     *
GAA GCT CCC GAC TCA TGG
 E   A   P   D   S   W>
```

FIG. 10Q

```
G GAG GTG GAT GTG TTA GAT GTG AAT GTC CGT GGC CCA GAT GGC TGC        46
  Glu Val Asp Val Leu Asp Val Asn Val Arg Gly Pro Asp Gly Cys
   1               5                  10                  15

ACC CCA TTG ATG TTG GCT TCT CTC CGA GGA GGC AGC TCA GAT TTG AGT      94
Thr Pro Leu Met Leu Ala Ser Leu Arg Gly Gly Ser Ser Asp Leu Ser
            20                  25                  50

GAT GAA GAT GAA GAT GCA GAG GAC TCT TCT GCT AAC ATC ATC ACA GAC      142
Asp Glu Asp Glu Asp Ala Glu Asp Ser Ser Ala Asn Ile Ile Thr Asp
         35                  40                  45

TTG GTC TAC CAG GGT GCC AGC CTC CAG GCC CAG ACA GAC CGG ACT GGT      190
Leu Val Tyr Gln Gly Ala Ser Leu Gln Ala Gln Thr Asp Arg Thr Gly
      50                  55                  60

GAG ATG GCC CTG CAC CTT GCA GCC CGC TAC TCA CGG GCT GAT GCT GCC      238
Glu Met Ala Leu His Leu Ala Ala Arg Tyr Ser Arg Ala Asp Ala Ala
   65                  70                  75

AAG CGT CTC CTG GAT GCA GGT GCA GAT GCC AAT GCC CAG GAC AAC ATG      286
Lys Arg Leu Leu Asp Ala Gly Ala Asp Ala Asn Ala Gln Asp Asn Met
80                  85                  90                  95

GGC CGC TGT CCA CTC CAT GCT GCA GTG GCA GCT GAT GCC CAA GGT GTC      334
Gly Arg Cys Pro Leu His Ala Ala Val Ala Ala Asp Ala Gln Gly Val
            100                 105                 110

TTC CAG ATT CTG ATT CGC AAC CGA GTA ACT GAT CTA GAT GCC AGG ATG      382
Phe Gln Ile Leu Ile Arg Asn Arg Val Thr Asp Leu Asp Ala Arg Met
         115                 120                 125

AAT GAT GGT ACT ACA CCC CTG ATC CTG GCT GCC CGC CTG GCT GTG GAG      430
Asn Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu
      130                 135                 140
```

FIG.11A

```
GGA ATG GTG GCA GAA CTG ATC AAC TGC CAA GCG GAT GTG AAT GCA GTG   478
Gly Met Val Ala Glu Leu Ile Asn Cys Gln Ala Asp Val Asn Ala Val
    145                 150                 155

GAT GAC CAT GGA AAA TCT GCT CTT CAC TGG GCA GCT GCT GTC AAT AAT   526
Asp Asp His Gly Lys Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn
160                 165                 170                 175

GTG GAG GCA ACT CTT TTG TTG TTG AAA AAT GGG GCC AAC CGA GAC ATG   574
Val Glu Ala Thr Leu Leu Leu Leu Lys Asn Gly Ala Asn Arg Asp Met
                180                 185                 190

CAG GAC AAC AAG GAA GAG ACA CCT CTG TTT CTT GCT GCC CGG GAG GGG   622
Gln Asp Asn Lys Glu Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly
            195                 200                 205

AGC TAT GAA GCA GCC AAG ATC CTG TTA GAC CAT TTT GCC AAT CGA GAC   670
Ser Tyr Glu Ala Ala Lys Ile Leu Leu Asp His Phe Ala Asn Arg Asp
        210                 215                 220

ATC ACA GAC CAT ATG GAT CGT CTT CCC CGG GAT GTG GCT CGG GAT CGC   718
Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp Val Ala Arg Asp Arg
    225                 230                 235

ATG CAC CAT GAC ATT GTG CGC CTT CTG GAT GAA TAC AAT GTG ACC CCA   766
Met His His Asp Ile Val Arg Leu Leu Asp Glu Tyr Asn Val Thr Pro
240                 245                 250                 255

AGC CCT CCA GGC ACC GTG TTG ACT TCT GCT CTC TCA CCT GTC ATC TGT   814
Ser Pro Pro Gly Thr Val Leu Thr Ser Ala Leu Ser Pro Val Ile Cys
                260                 265                 270

GGG CCC AAC AGA TCT TTC CTC AGC CTG AAG CAC ACC CCA ATG GGC AAG   862
Gly Pro Asn Arg Ser Phe Leu Ser Leu Lyn His Thr Pro Met Gly Lys
            275                 280                 285
```

FIG.11B

```
AAG TCT AGA CGG CCC AGT GCC AAG AGT ACC ATG CCT ACT AGC CTC CCT   910
Lys Ser Arg Arg Pro Ser Ala Lys Ser Thr Met Pro Thr Ser Leu Pro
        290                 295                 300

AAC CTT GCC AAG GAG GCA AAG GAT GCC AAG GGT AGT AGG AGG AAG AAG   958
Asn Leu Ala Lys Glu Ala Lys Asp Ala Lys Gly Ser Arg Arg Lys Lys
    305                 310                 315

TCT CTG AGT GAG AAG GTC CAA CTG TCT GAG AGT TCA GTA ACT TTA TCC  1006
Ser Leu Ser Glu Lys Val Gln Leu Ser Glu Ser Ser Val Thr Leu Ser
320                 325                 330                 335

CCT GTT GAT TCC CTA GAA TCT CCT CAC ACG TAT GTT TCC GAC ACC ACA  1054
Pro Val Asp Ser Leu Glu Ser Pro His Thr Tyr Val Ser Asp Thr Thr
                340                 345                 350

TCC TCT CCA ATG ATT ACA TCC CCT GGG ATC TTA CAG GCC TCA CCC AAC  1102
Ser Ser Pro Met Ile Thr Ser Pro Gly Ile Leu Gln Ala Ser Pro Asn
            355                 360                 365

CCT ATG TTG GCC ACT GCC GCC CCT CCT GCC CCA GTC CAT GCC CAG CAT  1150
Pro Met Leu Ala Thr Ala Ala Pro Pro Ala Pro Val His Ala Gln His
        370                 375                 380

GCA CTA TCT TTT TCT AAC CTT CAT GAA ATG CAG CCT TTG GCA CAT GGG  1198
Ala Leu Ser Phe Ser Asn Leu His Glu Met Gln Pro Leu Ala His Gly
    385                 390                 395

GCC AGC ACT GTG CTT CCC TCA GTG AGC CAG TTG CTA TCC CAC CAC CAC  1246
Ala Ser Thr Val Leu Pro Ser Val Ser Gln Leu Leu Ser His His His
400                 405                 410                 415

ATT GTG TCT CCA GGC AGT GGC AGT GCT GGA AGC TTG AGT AGG CTC CAT  1294
Ile Val Ser Pro Gly Ser Gly Ser Ala Gly Ser Leu Ser Arg Leu His
                420                 425                 430

CCA GTC CCA GTC CCA GCA GAT TGG ATG AAC CGC ATG GAG GTG AAT GAG  1342
Pro Val Pro Val Pro Ala Asp Trp Met Asn Arg Met Glu Val Asn Glu
            435                 440                 445
```

FIG.11C

```
ACC CAG TAC AAT GAG ATG TTT GGT ATG GTC CTG GCT CCA GCT GAG GGC  1390
Thr Gln Tyr Asn Glu Met Phe Gly Met Val Leu Ala Pro Ala Glu Gly
        450                 455                 460

ACC CAT CCT GGC ATA GCT CCC CAG AGC AGG CCA CCT GAA GGG AAG CAC  1438
Thr His Pro Gly Ile Ala Pro Gln Ser Arg Pro Pro Glu Gly Lys His
    465                 470                 475

ATA ACC ACC CCT CGG GAG CCC TTG CCC CCC ATT GTG ACT TTC CAG CTC  1486
Ile Thr Thr Pro Arg Glu Pro Leu Pro Pro Ile Val Thr Phe Gln Leu
480                 485                 490                 495

ATC CCT AAA GGC AGT ATT GCC CAA CCA GCG GGG GCT CCC CAG CCT CAG  1534
Ile Pro Lys Gly Ser Ile Ala Gln Pro Ala Gly Ala Pro Gln Pro Gln
                500                 505                 510

TCC ACC TGC CCT CCA GCT GTT GCG GGC CCC CTG CCC ACC ATG TAC CAG  1582
Ser Thr Cys Pro Pro Ala Val Ala Gly Pro Leu Pro Thr Met Tyr Gln
            515                 520                 525

ATT CCA GAA ATG GCC CGT TTG CCC AGT GTG GCT TTC CCC ACT GCC ATG  1630
Ile Pro Glu Met Ala Arg Leu Pro Ser Val Ala Phe Pro Thr Ala Met
        530                 535                 540

ATG CCC CAG CAG GAC GGG CAG GTA GCT CAG ACC ATT CTC CCA GCC TAT  1678
Met Pro Gln Gln Asp Gly Gln Val Ala Gln Thr Ile Leu Pro Ala Tyr
    545                 550                 555

CAT CCT TTC CCA GCC TCT GTG GGC AAG TAC CCC ACA CCC CCT TCA CAG  1726
His Pro Phe Pro Ala Ser Val Gly Lys Tyr Pro Thr Pro Pro Ser Gln
560                 565                 570                 575

CAC AGT TAT GCT TCC TCA AAT GCT GCT GAG CGA ACA CCC AGT CAC AGT  1774
His Ser Tyr Ala Ser Ser Asn Ala Ala Glu Arg Thr Pro Ser His Ser
                580                 585                 590

GGT CAC CTC CAG GGT GAG CAT CCC TAC CTG ACA CCA TCC CCA GAG TCT  1822
Gly His Leu Gln Gly Glu His Pro Tyr Leu Thr Pro Ser Pro Glu Ser
            595                 600                 605
```

FIG.11D

```
CCT GAC CAG TGG TCA AGT TCA TCA CCC CAC TCT GCT TCT GAC TGG TCA      1870
Pro Asp Gln Trp Ser Ser Ser Ser Pro His Ser Ala Ser Asp Trp Ser
        610                 615                 620

GAT GTG ACC ACC AGC CCT ACC CCT GGG GGT GCT GGA GGA GGT CAG CGG      1918
Asp Val Thr Thr Ser Pro Thr Pro Gly Gly Ala Gly Gly Gly Gln Arg
    625                 630                 635

GGA CCT GGG ACA CAC ATG TCT GAG CCA CCA CAC AAC AAC ATG CAG GTT      1966
Gly Pro Gly Thr His Met Ser Glu Pro Pro His Asn Asn Met Gln Val
640                 645                 650                 655

TAT GCG TGAGAGAGTC CACCTCCAGT GTAGAGACAT AACTGACTTT TGTAAATGCT       2022
Tyr Ala

GCTGAGGAAC AAATGAAGGT CATCCGGGAG AGAAATGAAG AAATCTCTGG AGCCAGCTTC   2082

TAGAGGTAGG AAAGAGAAGA TGTTCTTATT CAGATAATGC AAGAGAAGCA ATTCGTCAGT   2142

TTCACTGGGT ATCTGCAAGG CTTATTGATT ATTCTAATCT AATAAGACAA GTTTGTGGAA   2202

ATGCAAGATG AATACAAGCC TTGGGTCCAT GTTTACTCTC TTCTATTTGG AGAATAAGAT   2262

GGATGCTTAT TGAAGCCCAG ACATTCTTGC AGCTTGGACT GCATTTTAAG CCCTGCAGGC   2322

TTCTGCCATA TCCATGAGAA GATTCTACAC TAGCGTCCTG TTGGGAATTA TGCCCTGGAA   2382

TTCTGCCTGA ATTGACCTAC GCATCTCCTC CTCCTTGGAC ATTCTTTTGT CTTCATTTGG   2442

TGCTTTTGGT TTTGCACCTC TCCGTGATTG TAGCCCTACC AGCATGTTAT AGGGCAAGAC   2502

CTTTGTGCTT TTGATCATTC TGGCCCATGA AAGCAACTTT GGTCTCCTTT CCCCTCCTGT   2562

CTTCCCGGTA TCCCTTGGAG TCTCACAAGG TTTACTTTGG TATGGTTCTC AGCACAAACC   2622

TTTCAAGTAT GTTGTTTCTT TGGAAAATGG ACATACTGTA TTGTGTTCTC CTGCATATAT   2682

CATTCCTGGA GAGAGAAGGG GAGAAGAATA CTTTTCTTCA ACAAATTTTG GGGGCAGGAG   2742

ATCCCTTCAA GAGGCTGCAC CTTAATTTTT CTTGTCTGTG TGCAGGTCTT CATATAAACT   2802
```

FIG.11E

TTACCAGGAA GAAGGGTGTG AGTTTGTTGT TTTTCTGTGT ATGGGCCTGG TCAGTGTAAA 2862

GTTTTATCCT TGATAGTCTA GTTACTATGA CCCTCCCCAC TTTTTTAAAA CCAGAAAAAG 2922

GTTTGGAATG TTGGAATGAC CAAGAGACAA GTTAACTCGT GCAAGAGCCA GTTACCCACC 2982

CACAGGTCCC CCTACTTCCT GCCAAGCATT CCATTGACTG CCTGTATGGA ACACATTTGT 3042

CCCAGATCTG AGCATTCTAG GCCTGTTTCA CTCACTCACC CAGCATATGA AACTAGTCTT 3102

AACTGTTGAG CCTTTCCTTT CATATCCACA GAAGACACTG TCTCAAATGT TGTACCCTTG 3162

CCATTTAGGA CTGAACTTTC CTTAGCCCAA GGGACCCAGT GACAGTTGTC TTCCGTTTGT 3222

CAGATGATCA GTCTCTACTG ATTATCTTGC TGCTTAAAGG CCTGCTCACC AATCTTTCTT 3282

TCACACCGTG TGGTCCGTGT TACTGGTATA CCCAGTATGT TCTCACTGAA GACATGGACT 3342

TTATATGTTC AAGTGCAGGA ATTGGAAAGT TGGACTTGTT TTCTATGATC CAAAACAGCC 3402

CTATAAGAAG GTTGGAAAAG GAGGAACTAT ATAGCAGCCT TTGCTATTTT CTGCTACCAT 3462

TTCTTTTCCT CTGAAGCGGC CATGACATTC CCTTTGGCAA CTAACGTAGA AACTCAACAG 3522

FIG.11F

AACATTTTCC TTTCCTAGAG TCACCTTTTA GATGATAATG GACAACTATA GACTTGCTCA 3582

TTGTTCAGAC TGATTGCCCC TCACCTGAAT CCACTCTCTG TATTCATGCT CTTGGCAATT 3642

TCTTTGACTT TCTTTTAAGG GCAGAAGCAT TTTAGTTAAT TGTAGATAAA GAATAGTTTT 3702

CTTCCTCTTC TCCTTGGGCC AGTTAATAAT TGGTCCATGG CTACACTGCA ACTTCCGTCC 3762

AGTGCTGTGA TGCCCATGAC ACCTGCAAAA TAAGTTCTGC CTGGGCATTT TGTAGATATT 3822

AACAGGTGAA TTCCCGACTC TTTTGGTTTG AATGACAGTT CTCATTCCTT CTATGGCTGC 3882

AAGTATGCAT CAGTGCTTCC CACTTACCTG ATTTGTCTGT CGGTGGCCCC ATATGGAAAC 3942

CCTGCGTGTC TGTTGGCATA ATAGTTTACA AATGGTTTTT TCAGTCCTAT CCAAATTTAT 4002

TGAACCAACA AAAATAATTA CTTCTGCCCT GAGATAAGCA GATTAAGTTT GTTCATTCTC 4062

TGCTTTATTC TCTCCATGTG GCAACATTCT GTCAGCCTCT TTCATAGTGT GCAAACATTT 4122

TATCATTCTA AATGGTGACT CTCTGCCCTT GGACCCATTT ATTATTCACA GATGGGGAGA 4182

ACCTATCTGC ATGGACCCTC ACCATCCTCT GTGCAGCACA CACAGTGCAG GGAGCCAGTG 4242

GCGATGGCGA TGACTTTCTT CCCCTG 4268

FIG. 11G

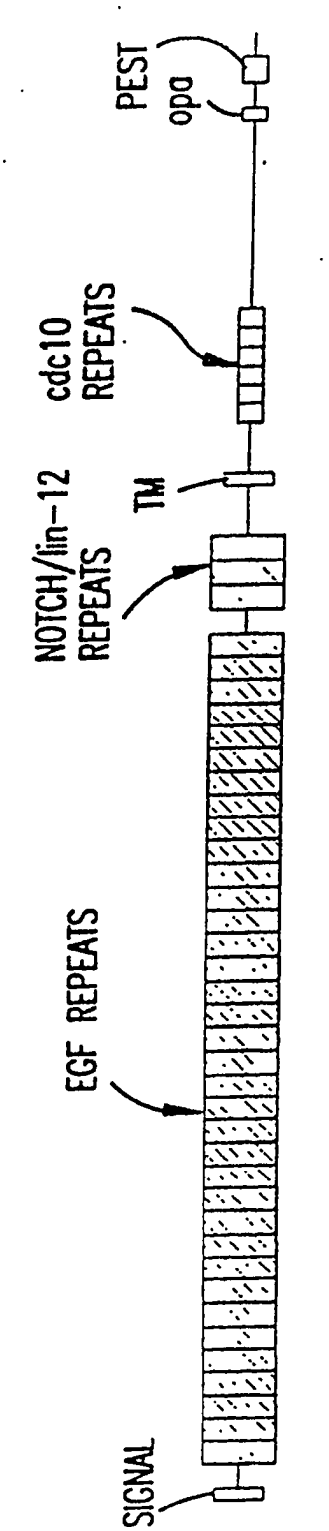

PEST
opa
cdc10
REPEATS
TM
NOTCH/lin-12
REPEATS
EGF REPEATS
SIGNAL

FIG. 12A

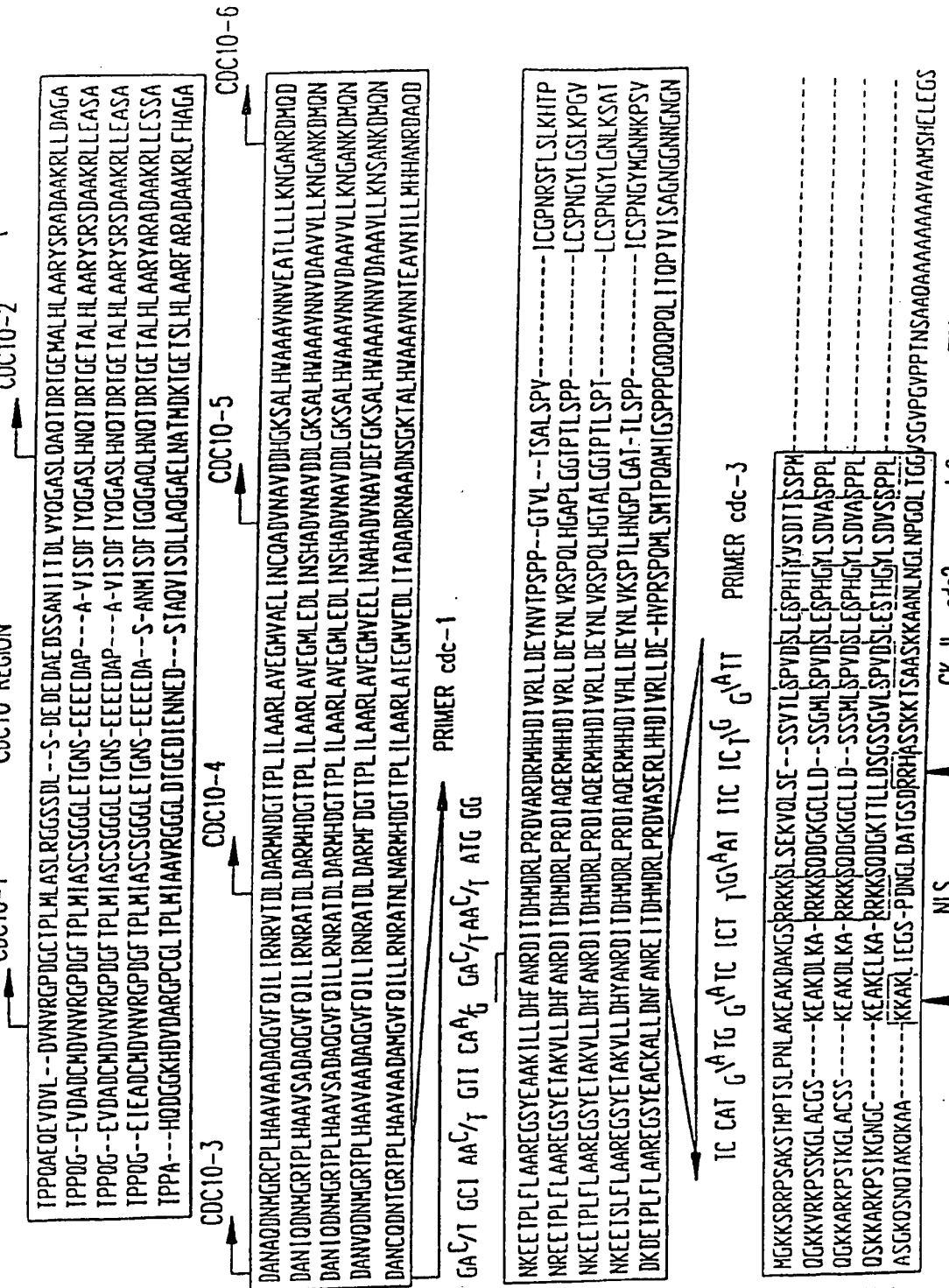

CDC10-1
CDC10 REGION
CDC10-2
CDC10-3
CDC10-4
CDC10-5
CDC10-6
PRIMER cdc-1
PRIMER cdc-3
NLS
CK-II
cdc2
cdc2
BNTS
hN5k
TAN-1
rat NOTCH
XENOPUS NOTCH
DROSOPH NOTCH

FIG.12B

```
hN5k                 ------------------------------------------------------------------------ITSPGIDQASPNPML--ATAAPPAPVHAQHALSF
TAN-1          2218  ------------------------------------------------------------------------LRSPR--QQSPSVPLNHLPGMPDTHLGIGHLNVA
rat NOTCH      2209  ------------------------------------------------------------------------LRSPR--QQSPSMPLSHLPGMPDTHLGISHLNVA
XENOPUS NOTCH  2214  ------------------------------------------------------------------------MTSPR--QQSPSMPLNHLTSMPESQLGMNHINMA
DROSOPH NOTCH  2285  PVGVGMGGNLPSPYDTSSMYSNAMAAPLANGNPNTGAKQPRSYEDCIKNAQSMQSLQGNGLDMIKLDNYAYSMGSPR--QQELLNGQGLGMNGNGQRNGVGPGVLP
                                                        CK-II                                  CK-II

hN5k                 SNLHEMQ--------------------PLAHGASTVLPSVSQLLSHHHIVSPGS--GSAGSLSRLHPVPVPADW--MNRMEVNETQYNEMFGMVLAPAEG-THPGI
TAN-1          2250  A-KPEMAALGGGGRLAFETGPPRLSHLPVASGTSTVLGSSSGGALNFTVGGSTSLNGQCEWLSRLQSGMVPNQYNPLRGSVAPGPLSTQAPSLQHG-MVGPLHSSL
rat NOTCH      2242  A-KPEMAALAGGSRLAFEPPPPRLSHLPVASSASTVLSTNGTGAMNFTVGAPASLNGQCEWLPRLQNGMVPSQYNPLRPGVTPGTLSTQAAGLQHGMM-SPIHSSL
XENOPUS NOTCH  2247  T-KQEMAA--GSNRMAFDAMVPRLTHL-NASSPNTIMS---NGSMHFTVGGAPTMNSQCDWLARLQNGMVQNQYDPIRNGIQQGN-AQQAQALQHGLMTS-LHNGL
DROSOPH NOTCH  2390  GGLCGMGGLSGAGNGNSHEQGLSPPYS-NQSPPHSVQSSLALSPHAYLGSPSPAKSRPSLPTSPTHIQAMRHATQQKQFGGSNLNSLLGGANGGGVVGGGGGGGGV

hN5k                 APQSRPPEGK--------------HITTPREPLPP-IV-TFQLIPKGSIAQPAG-----------------APQPQSTCPPAVAGPLPTMYQIP-----EMARL-P
TAN-1          2354  AASALSQMMS------------YQGLPSTRLATQPHLVQTQQVQPQNLQMQQQNLQPANIQQQQSLQPPPPPPQPHLGVSSAASGHLGRSFLSGEPSQADVQPLGP
rat NOTCH      2344  STNTLSPII-------------YQGLPNTRLATQPHLVQTQQVQPQNLQIQPQN-----------LQPPS---QPHLSVSSAANGHLGRSFLSGEPSQADVQPLGP
XENOPUS NOTCH  2343  PATTLSQMMT------------YQAMPNTRLANQPHLMQAQQMQQQQN-----------------LQLHQSMQQQHHN-SSTTSTHINSPFCSSDISQTDLQQM--
DROSOPH NOTCH  2495  VGQGPQNSPVSLGIISPTGSDMGIMLAPPQSSKNSAIMQTISPQQQQQQQQQQQQQHQQQQQQQQQQQQQQQQQLGGLEFGSAGLDLNG-FCGSPDSFHSGQMNPP

hN5k                 SVAFPTAMMPQQDGQVAQTILPAYHPFPASVGKYHTPPSQHSYASSNAAERTPSHSGHLQGEHPYLTPSPESPDQWSSSSPHSA-SDWSDVTTSPTP
TAN-1          2448  SSLAVHTILPQ-ESPALPTSLPSSLVPPVTAAQFLTPPSQHSY-SS-PVENTPSHQLQVP-EHPFLTPSPESPDQWSSSSPHSNVSDWSEGVSSPPT
rat NOTCH      2423  SSLPVHTILPQ-ESQALPTSLPSSMVPPMTTTQFLTPPSQHSY-SSSPVDNTPSHQLQVP-EHPFLTPSPESPDQWSSSSRHSNISDWSEGISSPPT
XENOPUS NOTCH  2416  SSNNIHSVMPQ-DTQIFAASLPSNLTQSMTTAQFLTPPSQHSY-SS-PMDNTPSHQLQVP-DHPFLTPSPESPDQWSSSSPHSNMSDWSEGISSPPT
DROSOPH NOTCH  2599  S---IQSSMSG-SSPSTNMLSPSSQHNQQAFYQYLTPSSQHS------GGHTPQHLVQTL-D-SYPTPSPESPGHWSSSSPRSN-SDWSEGVQSPAA
```

PEST ENTHALTENDE REGION

FIG.12C

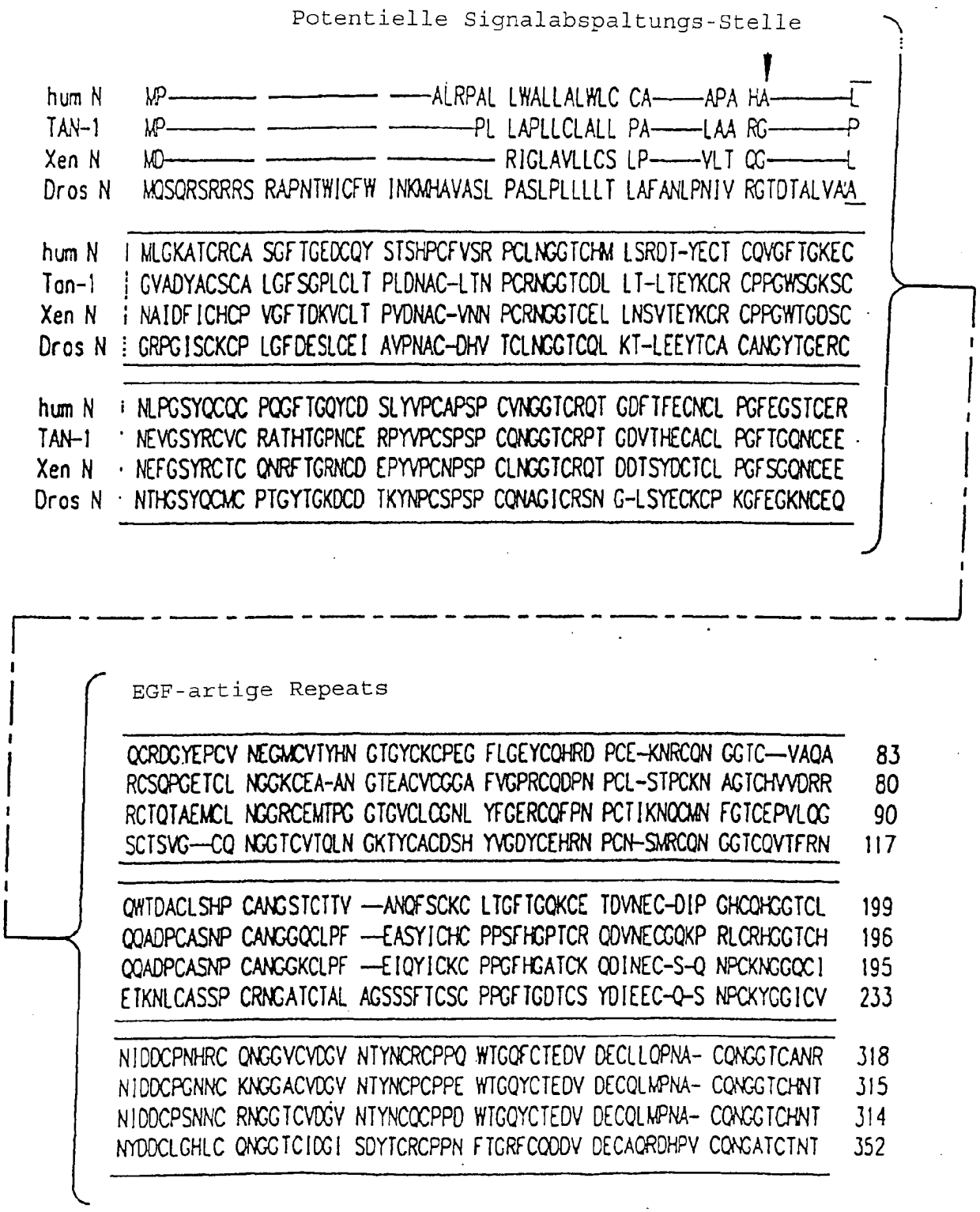

Potentielle Signalabspaltungs-Stelle
hum N MP——— ——— ——ALRPAL LWALLALWLC CA——APA HA———L
TAN-1 MP——— ——— ———PL LAPLLCLALL PA——LAA RG———P
Xen N MD——— ——— ——— RIGLAVLLCS LP——VLT QG———L
Dros N MQSQRSRRRS RAPNTWICFW INKMHAVASL PASLPLLLLT LAFANLPNIV RGTDTALVAA
hum N MLGKATCRCA SGFTGEDCQY STSHPCFVSR PCLNGGTCHM LSRDT-YECT CQVGFTGKEC
Tan-1 GVADYACSCA LGFSGPLCLT PLDNAC-LTN PCRNGGTCDL LT-LTEYKCR CPPGWSGKSC
Xen N NAIDFICHCP VGFTDKVCLT PVDNAC-VNN PCRNGGTCEL LNSVTEYKCR CPPGWTGDSC
Dros N GRPGISCKCP LGFDESLCEI AVPNAC-DHV TCLNGGTCQL KT-LEEYTCA CANGYTGERC
hum N NLPGSYQCQC PQGFTGQYCD SLYVPCAPSP CVNGGTCRQT GDFTFECNCL PGFEGSTCER
TAN-1 NEVGSYRCVC RATHTGPNCE RPYVPCSPSP CQNGGTCRPT GDVTHECACL PGFTGQNCEE
Xen N NEFGSYRCTC QNRFTGRNCD EPYVPCNPSP CLNGGTCRQT DDTSYDCTCL PGFSGQNCEE
Dros N NTHGSYQCMC PTGYTGKDCD TKYNPCSPSP CQNAGICRSN G-LSYECKCP KGFEGKNCEQ
EGF-artige Repeats
QCRDGYEPCV NEGMCVTYHN GTGYCKCPEG FLGEYCQHRD PCE-KNRCQN GGTC—VAQA 83
RCSQPGETCL NGGKCEA-AN GTEACVCGGA FVGPRCQDPN PCL-STPCKN AGTCHVVDRR 80
RCTQTAEMCL NGGRCEMTPG GTGVCLCGNL YFGERCQFPN PCTIKNQCMN FGTCEPVLQG 90
SCTSVG—CQ NGGTCVTQLN GKTYCACDSH YVGDYCEHRN PCN-SMRCQN GGTCQVTFRN 117
QWTDACLSHP CANGSTCTTV —ANQFSCKC LTGFTGQKCE TDVNEC-DIP GHCQHGGTCL 199
QQADPCASNP CANGGQCLPF —EASYICHC PPSFHGPTCR QDVNECGQKP RLCRHGGTCH 196
QQADPCASNP CANGGKCLPF —EIQYICKC PPGFHGATCK QDINEC-S-Q NPCKNGGQCI 195
ETKNLCASSP CRNGATCTAL AGSSSFTCSC PPGFTGDTCS YDIEEC-Q-S NPCKYGGICV 233
NIDDCPNHRC QNGGVCVDGV NTYNCRCPPQ WTGQFCTEDV DECLLQPNA- CQNGGTCANR 318
NIDDCPGNNC KNGGACVDGV NTYNCPCPPE WTGQYCTEDV DECQLMPNA- CQNGGTCHNT 315
NIDDCPSNNC RNGGTCVDGV NTYNCQCPPD WTGQYCTEDV DECQLMPNA- CQNGGTCHNT 314
NYDDCLGHLC QNGGTCIDGI SDYTCRCPPN FTGRFCQDDV DECAQRDHPV CQNGATCTNT 352

FIG.13A

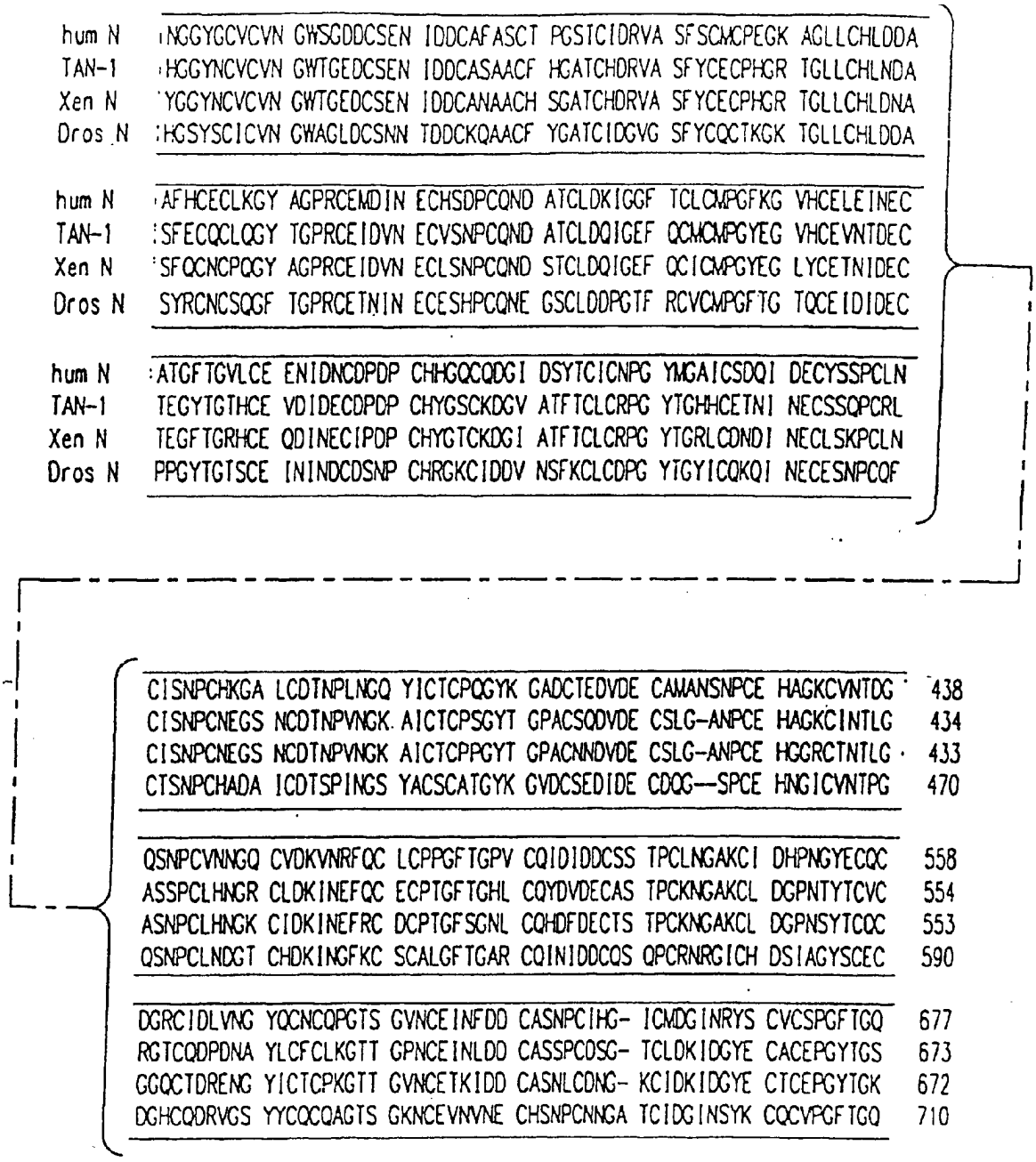

hum N NGGYGCVCVN GWSGDDCSEN IDDCAFASCT PGSTCIDRVA SFSCMCPEGK AGLLCHLDDA
TAN-1 HGGYNCVCVN GWTGEDCSEN IDDCASAACF HGATCHDRVA SFYCECPHGR TGLLCHLNDA
Xen N YGGYNCVCVN GWTGEDCSEN IDDCANAACH SGATCHDRVA SFYCECPHGR TGLLCHLDNA
Dros N HGSYSCICVN GWAGLDCSNN TDDCKQAACF YGATCIDGVG SFYCQCTKGK TGLLCHLDDA
hum N AFHCECLKGY AGPRCEMDIN ECHSDPCQND ATCLDKIGGF TCLCMPGFKG VHCELEINEC
TAN-1 SFECQCLQGY TGPRCEIDVN ECVSNPCQND ATCLDQIGEF QCMCMPGYEG VHCEVNTDEC
Xen N SFQCNCPQGY AGPRCEIDVN ECLSNPCQND STCLDQIGEF QCICMPGYEG LYCETNIDEC
Dros N SYRCNCSQGF TGPRCETNIN ECESHPCQNE GSCLDDPGTF RCVCMPGFTG TQCEIDIDEC
hum N ATGFTGVLCE ENIDNCDPDP CHHGQCQDGI DSYTCICNPG YMGAICSDQI DECYSSPCLN
TAN-1 TEGYTGTHCE VDIDECDPDP CHYGSCKDGV ATFTCLCRPG YTGHHCETNI NECSSQPCRL
Xen N TEGFTGRHCE QDINECIPDP CHYGTCKDGI ATFTCLCRPG YTGRLCDNDI NECLSKPCLN
Dros N PPGYTGTSCE ININDCDSNP CHRGKCIDDV NSFKCLCDPG YTGYICQKQI NECESNPCQF
CISNPCHKGA LCDTNPLNGQ YICTCPQGYK GADCTEDVDE CAMANSNPCE HAGKCVNTDG 438
CISNPCNEGS NCDTNPVNGK AICTCPSGYT GPACSQDVDE CSLG-ANPCE HAGKCINTLG 434
CISNPCNEGS NCDTNPVNGK AICTCPPGYT GPACNNDVDE CSLG-ANPCE HGGRCTNTLG 433
CTSNPCHADA ICDTSPINGS YACSCATGYK GVDCSEDIDE CDQG--SPCE HNGICVNTPG 470
QSNPCVNNGQ CVDKVNRFQC LCPPGFTGPV CQIDIDDCSS TPCLNGAKCI DHPNGYECQC 558
ASSPCLHNGR CLDKINEFQC ECPTGFTGHL CQYDVDECAS TPCKNGAKCL DGPNTYTCVC 554
ASNPCLHNGK CIDKINEFRC DCPTGFSGNL CQHDFDECTS TPCKNGAKCL DGPNSYTCQC 553
QSNPCLNDGT CHDKINGFKC SCALGFTGAR CQINIDDCQS QPCRNRGICH DSIAGYSCEC 590
DGRCIDLVNG YQCNCQPGTS GVNCEINFDD CASNPCIHG- ICMDGINRYS CVCSPGFTGQ 677
RGTCQDPDNA YLCFCLKGTT GPNCEINLDD CASSPCDSG- TCLDKIDGYE CACEPGYTGS 673
GGQCTDRENG YICTCPKGTT GVNCETKIDD CASNLCDNG- KCIDKIDGYE CTCEPGYTGK 672
DGHCQDRVGS YYCQCQAGTS GKNCEVNVNE CHSNPCNNGA TCIDGINSYK CQCVPGFTGQ 710

FIG.13B

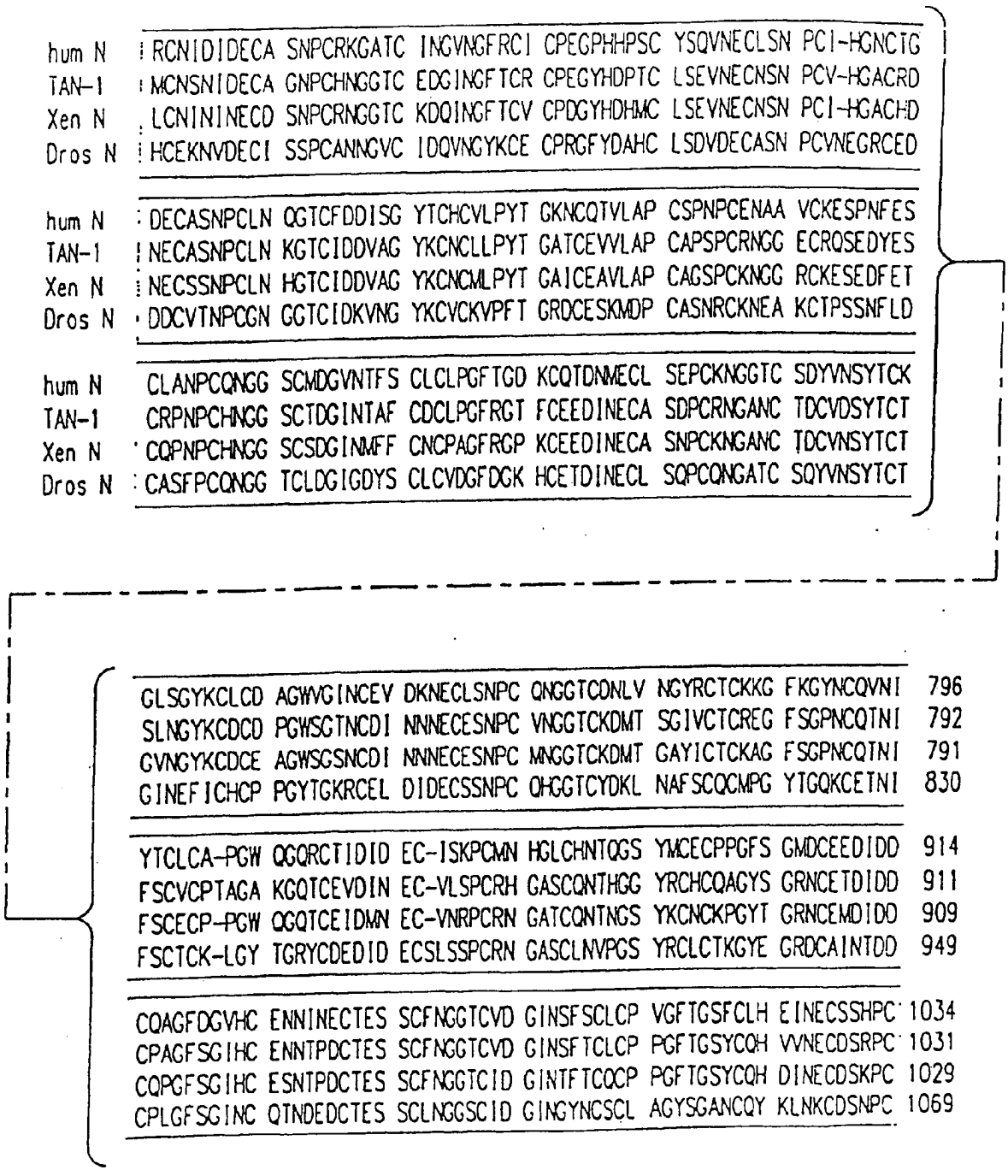

hum N RCNIDIDECA SNPCRKGATC INGVNGFRCI CPEGPHHPSC YSQVNECLSN PCI-HGNCTG
TAN-1 MCNSNIDECA GNPCHNGGTC EDGINGFTCR CPEGYHDPTC LSEVNECNSN PCV-HGACRD
Xen N LCNININECD SNPCRNGGTC KDQINGFTCV CPDGYHDHMC LSEVNECNSN PCI-HGACHD
Dros N HCEKNVDECI SSPCANNGVC IDQVNGYKCE CPRGFYDAHC LSDVDECASN PCVNEGRCED
hum N DECASNPCLN QGTCFDDISG YTCHCVLPYT GKNCQTVLAP CSPNPCENAA VCKESPNFES
TAN-1 NECASNPCLN KGTCIDDVAG YKCNCLLPYT GATCEVVLAP CAPSPCRNGG ECRQSEDYES
Xen N NECSSNPCLN HGTCIDDVAG YKCNCMLPYT GAICEAVLAP CAGSPCKNGG RCKESEDFET
Dros N DDCVTNPCGN GGTCIDKVNG YKCVCKVPFT GRDCESKMDP CASNRCKNEA KCTPSSNFLD
hum N CLANPCQNGG SCMDGVNTFS CLCLPGFTGD KCQTDNMECL SEPCKNGGTC SDYVNSYTCK
TAN-1 CRPNPCHNGG SCTDGINTAF CDCLPGFRGT FCEEDINECA SDPCRNGANC TDCVDSYTCT
Xen N CQPNPCHNGG SCSDGINMFF CNCPAGFRGP KCEEDINECA SNPCKNGANC TDCVNSYTCT
Dros N CASFPCQNGG TCLDGIGDYS CLCVDGFDGK HCETDINECL SQPCQNGATC SQYVNSYTCT
GLSGYKCLCD AGWVGINCEV DKNECLSNPC QNGGTCDNLV NGYRCTCKKG FKGYNCQVNI 796
SLNGYKCDCD PGWSGTNCDI NNNECESNPC VNGGTCKDMT SGIVCTCREG FSGPNCQTNI 792
GVNGYKCDCE AGWSGSNCDI NNNECESNPC MNGGTCKDMT GAYICTCKAG FSGPNCQTNI 791
GINEFICHCP PGYTGKRCEL DIDECSSNPC QHGGTCYDKL NAFSCQCMPG YTGQKCETNI 830
YTCLCA-PGW QGQRCTIDID EC-ISKPCMN HGLCHNTQGS YMCECPPGFS GMDCEEDIDD 914
FSCVCPTAGA KGQTCEVDIN EC-VLSPCRH GASCQNTHGG YRCHCQAGYS GRNCETDIDD 911
FSCECP-PGW QGQTCEIDMN EC-VNRPCRN GATCQNTNGS YKCNCKPGYT GRNCEMDIDD 909
FSCTCK-LGY TGRYCDEDID ECSLSSPCRN GASCLNVPGS YRCLCTKGYE GRDCAINTDD 949
CQAGFDGVHC ENNINECTES SCFNGGTCVD GINSFSCLCP VGFTGSFCLH EINECSSHPC 1034
CPAGFSGIHC ENNTPDCTES SCFNGGTCVD GINSFTCLCP PGFTGSYCQH VVNECDSRPC 1031
CQPGFSGIHC ESNTPDCTES SCFNGGTCID GINTFTCQCP PGFTGSYCQH DINECDSKPC 1029
CPLGFSGINC QTNDEDCTES SCLNGGSCID GINGYNCSCL AGYSGANCQY KLNKCDSNPC 1069

FIG.13C

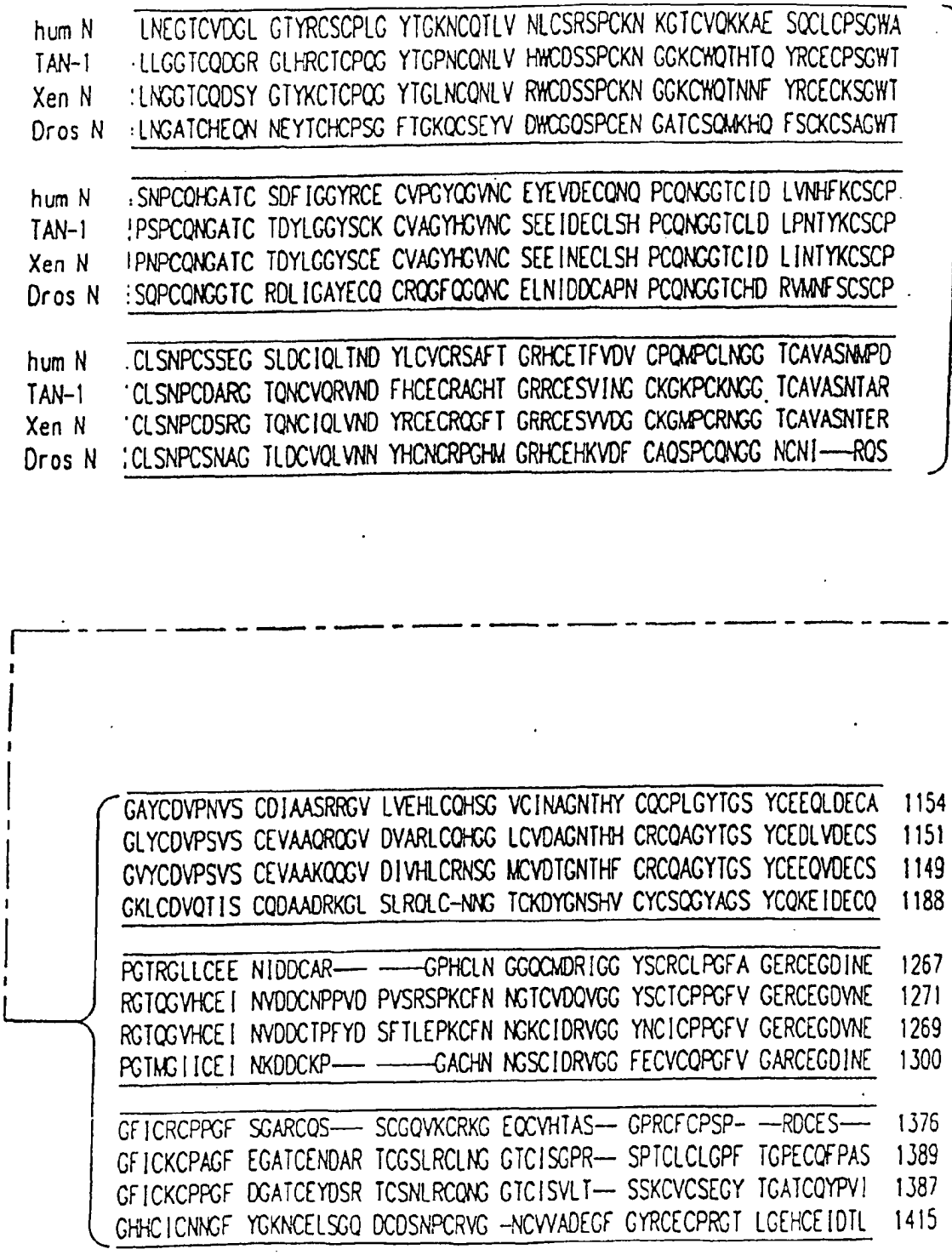

hum N LNEGTCVDGL GTYRCSCPLG YTGKNCQTLV NLCSRSPCKN KGTCVQKKAE SQCLCPSGWA
TAN-1 LLGGTCQDGR GLHRCTCPQG YTGPNCQNLV HWCDSSPCKN GGKCWQTHTQ YRCECPSGWT
Xen N LNGGTCQDSY GTYKCTCPQG YTGLNCQNLV RWCDSSPCKN GGKCWQTNNF YRCECKSGWT
Dros N LNGATCHEQN NEYTCHCPSG FTGKQCSEYV DWCGQSPCEN GATCSQMKHQ FSCKCSAGWT
hum N SNPCQHGATC SDFIGGYRCE CVPGYQGVNC EYEVDECQNQ PCQNGGTCID LVNHFKCSCP
TAN-1 PSPCQNGATC TDYLGGYSCK CVAGYHGVNC SEEIDECLSH PCQNGGTCLD LPNTYKCSCP
Xen N PNPCQNGATC TDYLGGYSCE CVAGYHGVNC SEEINECLSH PCQNGGTCID LINTYKCSCP
Dros N SQPCQNGGTC RDLIGAYECQ CRQGFQGQNC ELNIDDCAPN PCQNGGTCHD RVNNFSCSCP
hum N CLSNPCSSEG SLDCIQLTND YLCVCRSAFT GRHCETFVDV CPQMPCLNGG TCAVASNMPD
TAN-1 CLSNPCDARG TQNCVQRVND FHCECRAGHT GRRCESVING CKGKPCKNGG TCAVASNTAR
Xen N CLSNPCDSRG TQNCIQLVND YRCECRQGFT GRRCESVVDG CKGMPCRNGG TCAVASNTER
Dros N CLSNPCSNAG TLDCVQLVNN YHCNCRPGHM GRHCEHKVDF CAQSPCQNGG NCNI——RQS
GAYCDVPNVS CDIAASRRGV LVEHLCQHSG VCINAGNTHY CQCPLGYTGS YCEEQLDECA 1154
GLYCDVPSVS CEVAAQRQGV DVARLCQHGG LCVDAGNTHH CRCQAGYTGS YCEDLVDECS 1151
GVYCDVPSVS CEVAAKQQGV DIVHLCRNSG MCVDTGNTHF CRCQAGYTGS YCEEQVDECS 1149
GKLCDVQTIS CQDAADRKGL SLRQLC-NNG TCKDYGNSHV CYCSQGYAGS YCQKEIDECQ 1188
PGTRGLLCEE NIDDCAR—— ——GPHCLN GGQCMDRIGG YSCRCLPGFA GERCEGDINE 1267
RGTQGVHCEI NVDDCNPPVD PVSRSPKCFN NGTCVDQVGG YSCTCPPGFV GERCEGDVNE 1271
RGTQGVHCEI NVDDCTPFYD SFTLEPKCFN NGKCIDRVGG YNCICPPGFV GERCEGDVNE 1269
PGTMGIICEI NKDDCKP—— ——GACHN NGSCIDRVGG FECVCQPGFV GARCEGDINE 1300
GFICRCPPGF SGARCQS—— SCGQVKCRKG EQCVHTAS— GPRCFCPSP- —RDCES—— 1376
GFICKCPAGF EGATCENDAR TCGSLRCLNG GTCISGPR— SPTCLCLGPF TGPECQFPAS 1389
GFICKCPPGF DGATCEYDSR TCSNLRCQNG GTCISVLT— SSKCVCSEGY TGATCQYPVI 1387
GHHCICNNGF YGKNCELSGQ DCDSNPCRVG -NCVVADEGF GYRCECPRGT LGEHCEIDTL 1415

FIG.13D

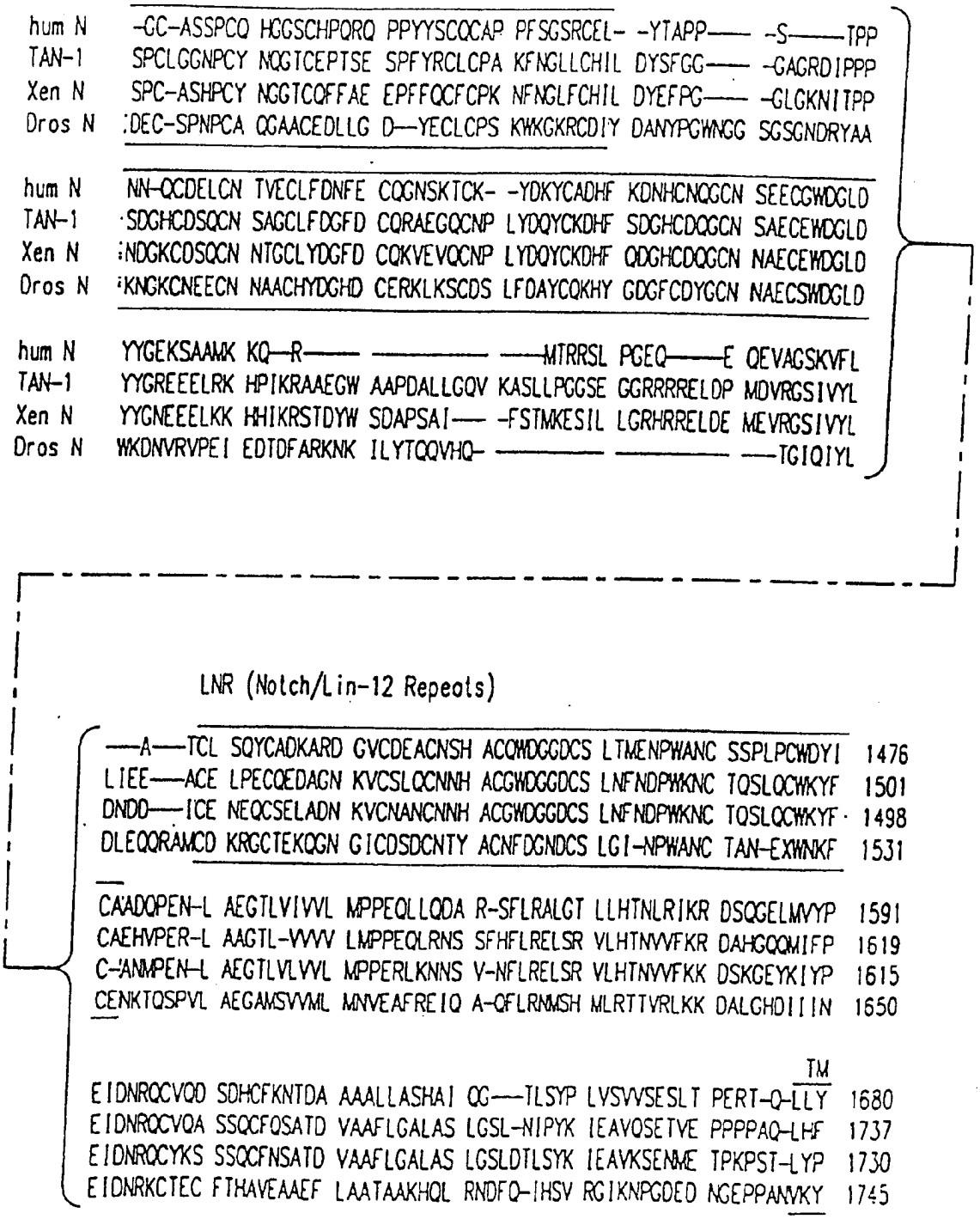

hum N -GC-ASSPCQ HGGSCHPQRQ PPYYSCQCAP PFSGSRCEL- -YTAPP—— -S——TPP
TAN-1 SPCLGGNPCY NQGTCEPTSE SPFYRCLCPA KFNGLLCHIL DYSFGG—— -GAGRDIPPP
Xen N SPC-ASHPCY NGGTCQFFAE EPFFQCFCPK NFNGLFCHIL DYEFPG—— -GLGKNITPP
Dros N DEC-SPNPCA QGAACEDLLG D—YECLCPS KWKGKRCDIY DANYPGWNGG SGSGNDRYAA
hum N NN-QCDELCN TVECLFDNFE CQGNSKTCK- -YDKYCADHF KDNHCNQGCN SEECGWDGLD
TAN-1 SDGHCDSQCN SAGCLFDGFD CQRAEGQCNP LYDQYCKDHF SDGHCDQGCN SAECEWDGLD
Xen N NDGKCDSQCN NTGCLYDGFD CQKVEVQCNP LYDQYCKDHF QDGHCDQGCN NAECEWDGLD
Dros N KNGKCNEECN NAACHYDGHD CERKLKSCDS LFDAYCQKHY GDGFCDYGCN NAECSWDGLD
hum N YYGEKSAAMK KQ—R—— ———— ——MTRRSL PGEQ——E QEVAGSKVFL
TAN-1 YYGREEELRK HPIKRAAEGW AAPDALLGQV KASLLPGGSE GGRRRRELDP MDVRGSIVYL
Xen N YYGNEEELKK HHIKRSTDYW SDAPSAI—— -FSTMKESIL LGRHRRELDE MEVRGSIVYL
Dros N WKDNVRVPEI EDTDFARKNK ILYTQQVHQ- ———— ———— ——TGIQIYL
LNR (Notch/Lin-12 Repeats)
——A——TCL SQYCADKARD GVCDEACNSH ACQWDGGDCS LTMENPWANC SSPLPCWDYI 1476
LIEE——ACE LPECQEDAGN KVCSLQCNNH ACGWDGGDCS LNFNDPWKNC TQSLQCWKYF 1501
DNDD——ICE NEQCSELADN KVCNANCNNH ACGWDGGDCS LNFNDPWKNC TQSLQCWKYF 1498
DLEQQRAMCD KRGCTEKQGN GICDSDCNTY ACNFDGNDCS LGI-NPWANC TAN-EXWNKF 1531
CAADQPEN-L AEGTLVIVVL MPPEQLLQDA R-SFLRALGT LLHTNLRIKR DSQGELMVYP 1591
CAEHVPER-L AAGTL-VVVV LMPPEQLRNS SFHFLRELSR VLHTNVVFKR DAHGQQMIFP 1619
C-ANMPEN-L AEGTLVLVVL MPPERLKNNS V-NFLRELSR VLHTNVVFKK DSKGEYKIYP 1615
CENKTQSPVL AEGAMSVVML MNVEAFREIQ A-QFLRNMSH MLRTTVRLKK DALGHDIIIN 1650
TM
EIDNRQCVQD SDHCFKNTDA AAALLASHAI QG—TLSYP LVSVVSESLT PERT-Q-LLY 1680
EIDNRQCVQA SSQCFQSATD VAAFLGALAS LGSL-NIPYK IEAVQSETVE PPPPAQ-LHF 1737
EIDNRQCYKS SSQCFNSATD VAAFLGALAS LGSLDTLSYK IEAVKSENME TPKPST-LYP 1730
EIDNRKCTEC FTHAVEAAEF LAATAAKHQL RNDFQ-IHSV RGIKNPGDED NGEPPANVKY 1745

FIG.13E

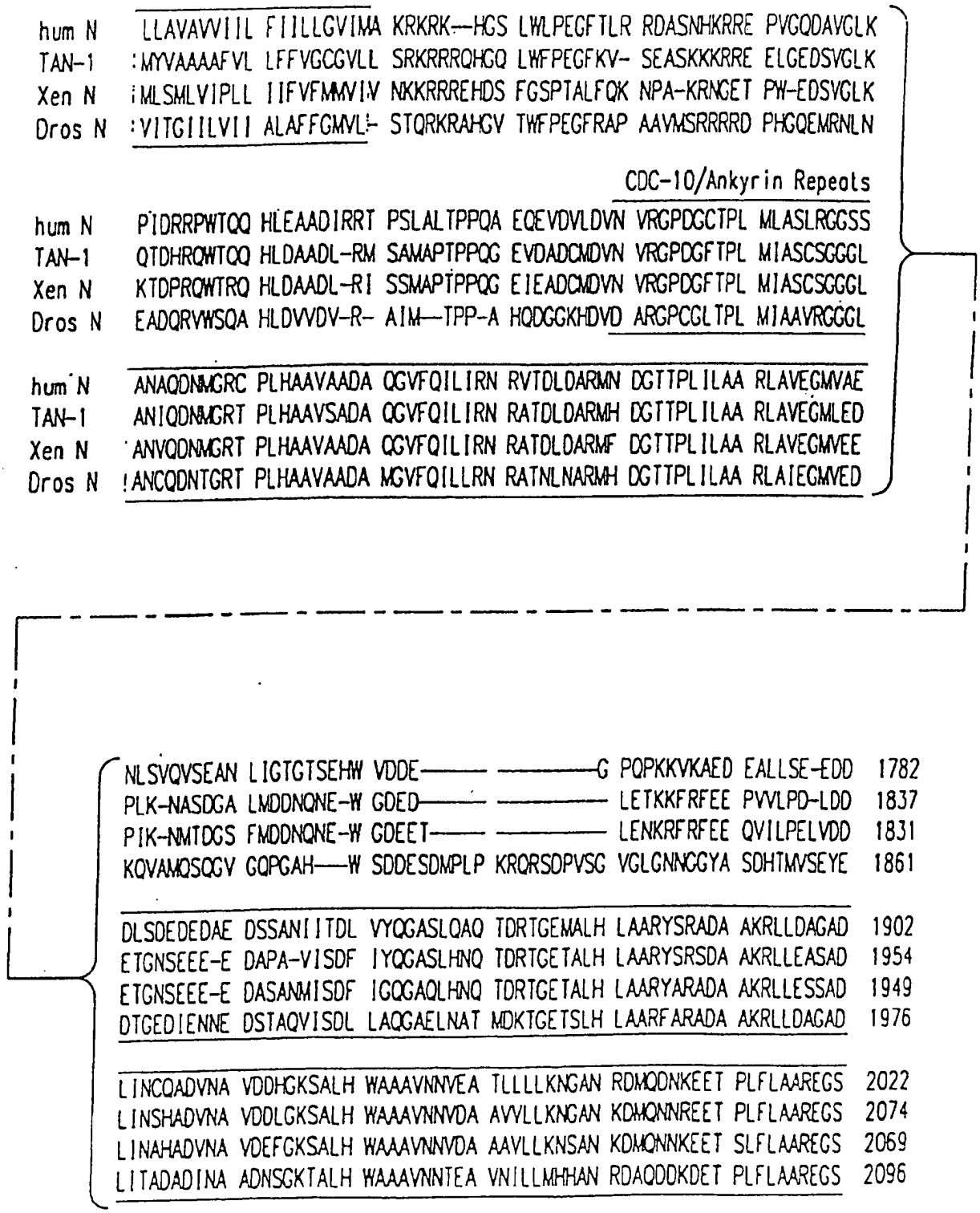

hum N LLAVAVVIIL FIILLGVIMA KRKRK--HGS LWLPEGFTLR RDASNHKRRE PVGQDAVGLK
TAN-1 MYVAAAAFVL LFFVGCGVLL SRKRRRQHGQ LWFPEGFKV- SEASKKKRRE ELGEDSVGLK
Xen N MLSMLVIPLL IIFVFMMVIV NKKRRREHDS FGSPTALFQK NPA-KRNGET PW-EDSVGLK
Dros N VITGIILVII ALAFFGMVL- STQRKRAHGV TWFPEGFRAP AAVMSRRRRD PHGQEMRNLN
CDC-10/Ankyrin Repeats
hum N PIDRRPWTQQ HLEAADIRRT PSLALTPPQA EQEVDVLDVN VRGPDGCTPL MLASLRGGSS
TAN-1 QTDHRQWTQQ HLDAADL-RM SAMAPTPPQG EVDADCMDVN VRGPDGFTPL MIASCSGGGL
Xen N KTDPRQWTRQ HLDAADL-RI SSMAPTPPQG EIEADCMDVN VRGPDGFTPL MIASCSGGGL
Dros N EADQRVWSQA HLDVVDV-R- AIM--TPP-A HQDGGKHDVD ARGPCGLTPL MIAAVRGGGL
hum N ANAQDNMGRC PLHAAVAADA QGVFQILIRN RVTDLDARMN DGTTPLILAA RLAVEGMVAE
TAN-1 ANIQDNMGRT PLHAAVSADA QGVFQILIRN RATDLDARMH DGTTPLILAA RLAVEGMLED
Xen N ANVQDNMGRT PLHAAVAADA QGVFQILIRN RATDLDARMF DGTTPLILAA RLAVEGMVEE
Dros N ANCQDNTGRT PLHAAVAADA MGVFQILLRN RATNLNARMH DGTTPLILAA RLAIEGMVED
NLSVQVSEAN LIGTGTSEHW VDDE---------- ----------G PQPKKVKAED EALLSE-EDD 1782
PLK-NASDGA LMDDNQNE-W GDED---------- ---------- LETKKFRFEE PVVLPD-LDD 1837
PIK-NMTDGS FMDDNQNE-W GDEET---------- ---------- LENKRFRFEE QVILPELVDD 1831
KQVAMQSQGV GQPGAH----W SDDESDMPLP KRQRSDPVSG VGLGNNGGYA SDHTMVSEYE 1861
DLSDEDEDAE DSSANIITDL VYQGASLQAQ TDRTGEMALH LAARYSRADA AKRLLDAGAD 1902
ETGNSEEE-E DAPA-VISDF IYQGASLHNQ TDRTGETALH LAARYSRSDA AKRLLEASAD 1954
ETGNSEEE-E DASANMISDF IGQGAQLHNQ TDRTGETALH LAARYARADA AKRLLESSAD 1949
DTGEDIENNE DSTAQVISDL LAQGAELNAT MDKTGETSLH LAARFARADA AKRLLDAGAD 1976
LINCQADVNA VDDHGKSALH WAAAVNNVEA TLLLLKNGAN RDMQDNKEET PLFLAAREGS 2022
LINSHADVNA VDDLGKSALH WAAAVNNVDA AVVLLKNGAN KDMQNNREET PLFLAAREGS 2074
LINAHADVNA VDEFGKSALH WAAAVNNVDA AAVLLKNSAN KDMQNNKEET SLFLAAREGS 2069
LITADADINA ADNSGKTALH WAAAVNNTEA VNILLMHHAN RDAQDDKDET PLFLAAREGS 2096

FIG.13F

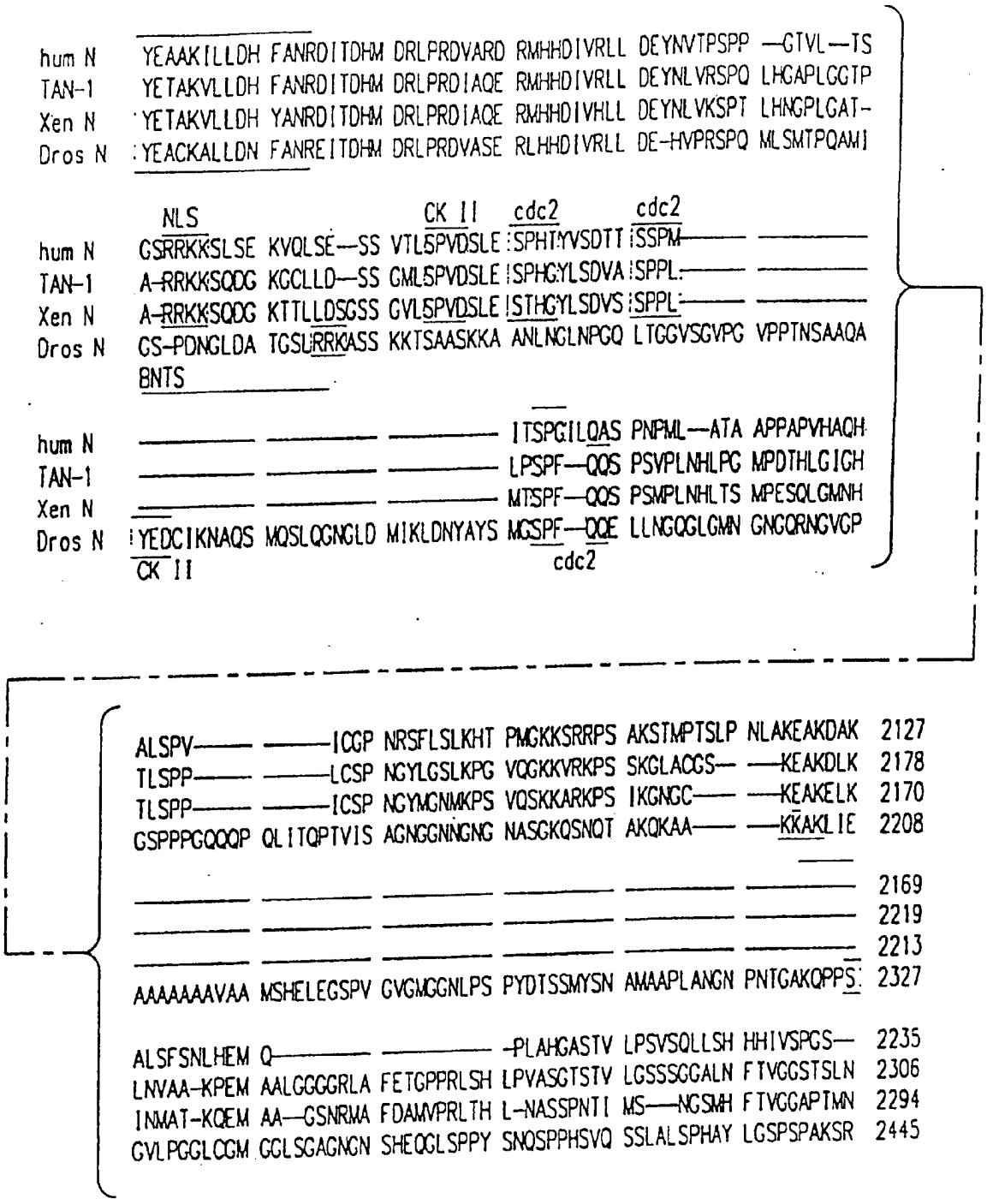

hum N YEAAKILLDH FANRDITDHM DRLPRDVARD RMHHDIVRLL DEYNVTPSPP —GTVL—TS
TAN-1 YETAKVLLDH FANRDITDHM DRLPRDIAQE RMHHDIVRLL DEYNLVRSPQ LHGAPLGGTP
Xen N YETAKVLLDH YANRDITDHM DRLPRDIAQE RMHHDIVHLL DEYNLVKSPT LHNGPLGAT-
Dros N YEACKALLDN FANREITDHM DRLPRDVASE RLHHDIVRLL DE-HVPRSPQ MLSMTPQAMI
NLS
CK II
cdc2
cdc2
hum N GSRRKKSLSE KVQLSE—SS VTLSPVDSLE SPHTMVSDTT SSPM———— ————————
TAN-1 A-RRKKSQDG KGCLLD—SS GMLSPVDSLE SPHGYLSDVA SPPL———— ————————
Xen N A-RRKKSQDG KTTLLDSGSS GVLSPVDSLE STHGYLSDVS SPPL———— ————————
Dros N GS-PDNGLDA TGSLRRKASS KKTSAASKKA ANLNGLNPGQ LTGGVSGVPG VPPTNSAAQA
BNTS
hum N ———— ———— ———— ITSPGILQAS PNPML—ATA APPAPVHAQH
TAN-1 ———— ———— ———— LPSPF—QQS PSVPLNHLPG MPDTHLGIGH
Xen N ———— ———— ———— MTSPF—QQS PSMPLNHLTS MPESQLGMNH
Dros N YEDCIKNAQS MQSLQGNGLD MIKLDNYAYS MGSPF—QQE LLNGQGLGMN GNGQRNGVGP
CK II
cdc2
ALSPV———— ————ICGP NRSFLSLKHT PMGKKSRRPS AKSTMPTSLP NLAKEAKDAK 2127
TLSPP———— ————LCSP NGYLGSLKPG VQGKKVRKPS SKGLACGS— ——KEAKDLK 2178
TLSPP———— ————ICSP NGYMGNMKPS VQSKKARKPS IKGNGC—— ——KEAKELK 2170
GSPPPGQQQP QLITQPTVIS AGNGGNNGNG NASGKQSNQT AKQKAA—— ——KKAKLIE 2208
———— ———— ———— ———— ———— ———— 2169
———— ———— ———— ———— ———— ———— 2219
———— ———— ———— ———— ———— ———— 2213
AAAAAAAVAA MSHELEGSPV GVGMGGNLPS PYDTSSMYSN AMAAPLANGN PNTGAKQPPS 2327
ALSFSNLHEM Q———— ———— -PLAHGASTV LPSVSQLLSH HHIVSPGS— 2235
LNVAA-KPEM AALGGGGRLA FETGPPRLSH LPVASGTSTV LGSSSGGALN FTVGGSTSLN 2306
INMAT-KQEM AA—GSNRMA FDAMVPRLTH L-NASSPNTI MS——NGSMH FTVGGAPTMN 2294
GVLPGGLCGM GGLSGAGNGN SHEQGLSPPY SNQSPPHSVQ SSLALSPHAY LGSPSPAKSR 2445

FIG.13G

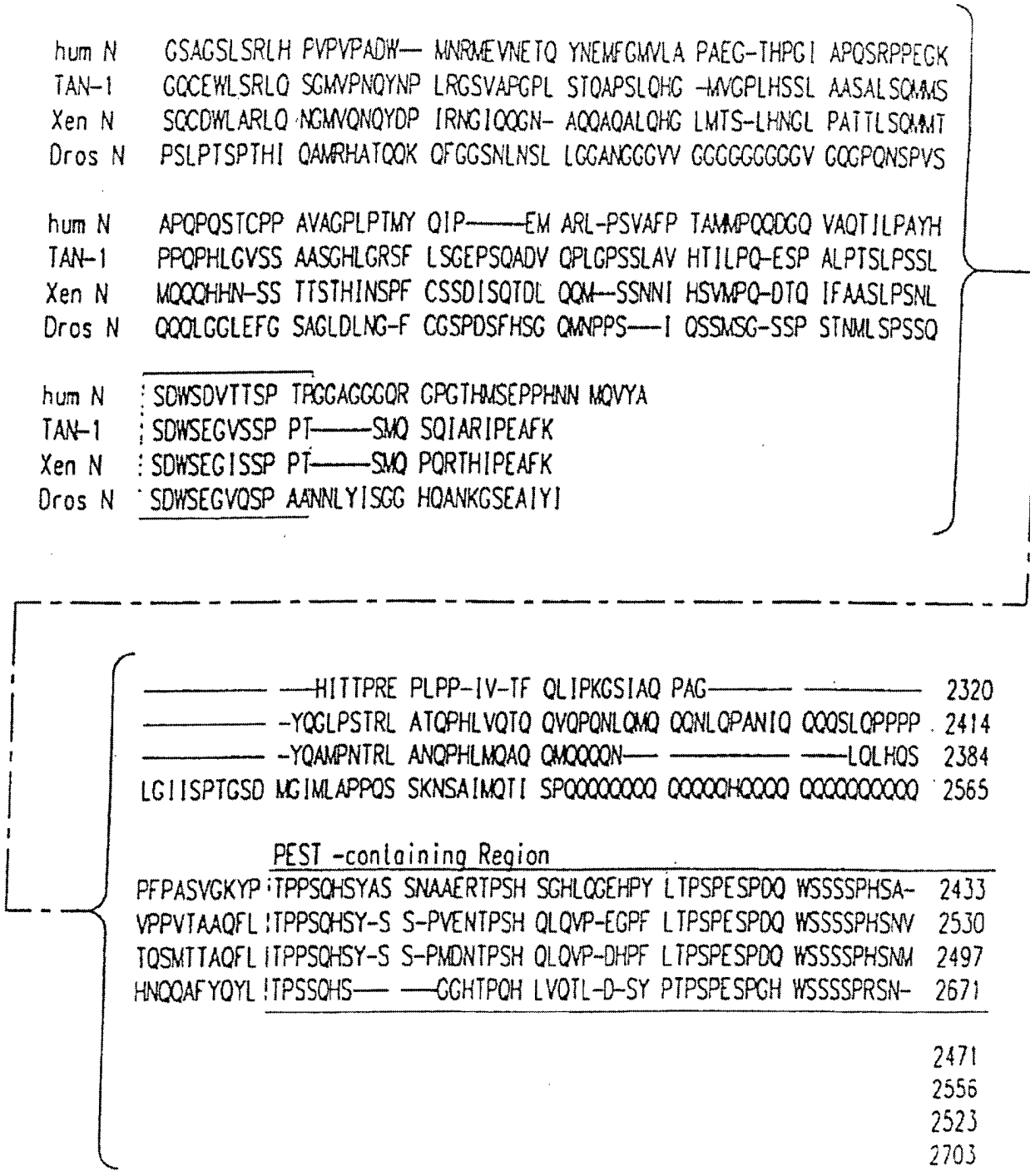

hum N GSAGSLSRLH PVPVPADW— MNRMEVNETQ YNEMFGMVLA PAEG-THPGI APQSRPPEGK
TAN-1 GQCEWLSRLQ SGMVPNQYNP LRGSVAPGPL STQAPSLQHG -MVGPLHSSL AASALSQMMS
Xen N SQCDWLARLQ NGMVQNQYDP IRNGIQQGN- AQQAQALQHG LMTS-LHNGL PATTLSQMMT
Dros N PSLPTSPTHI QAMRHATQQK QFGGSNLNSL LGGANGGVV GGGGGGGGGV GQGPQNSPVS
hum N APQPQSTCPP AVAGPLPTMY QIP———EM ARL-PSVAFP TAMMPQQDGQ VAQTILPAYH
TAN-1 PPQPHLGVSS AASGHLGRSF LSGEPSQADV QPLGPSSLAV HTILPQ-ESP ALPTSLPSSL
Xen N MQQQHHN-SS TTSTHINSPF CSSDISQTDL QQM—SSNNI HSVMPQ-DTQ IFAASLPSNL
Dros N QQQLGGLEFG SAGLDLNG-F CGSPDSFHSG QMNPPS——I QSSMSG-SSP STNMLSPSSQ
hum N SDWSDVTTSP TPGGAGGGQR GPGTHMSEPPHNN MQVYA
TAN-1 SDWSEGVSSP PT———SMQ SQIARIPEAFK
Xen N SDWSEGISSP PT———SMQ PQRTHIPEAFK
Dros N SDWSEGVQSP AANNLYISGG HQANKGSEAIYI
———— ——HITTPRE PLPP-IV-TF QLIPKGSIAQ PAG———— ———— 2320
———— -YQGLPSTRL ATQPHLVQTQ QVQPQNLQMQ QQNLQPANIQ QQQSLQPPPP 2414
———— -YQAMPNTRL ANQPHLMQAQ QMQQQQN—— ———— ——LQLHQS 2384
LGIISPTGSD MGIMLAPPQS SKNSAIMQTI SPQQQQQQQQ QQQQQHQQQQ QQQQQQQQQQ 2565
PEST -containing Region
PFPASVGKYP ITPPSQHSYAS SNAAERTPSH SGHLQGEHPY LTPSPESPDQ WSSSSPHSA- 2433
VPPVTAAQFL ITPPSQHSY-S S-PVENTPSH QLQVP-EGPF LTPSPESPDQ WSSSSPHSNV 2530
TQSMTTAQFL ITPPSQHSY-S S-PMDNTPSH QLQVP-DHPF LTPSPESPDQ WSSSSPHSNM 2497
HNQQAFYQYL ITPSSQHS—— ——GGHTPQH LVQTL-D-SY PTPSPESPGH WSSSSPRSN- 2671
2471
2556
2523
2703

FIG.13H



FIG.14

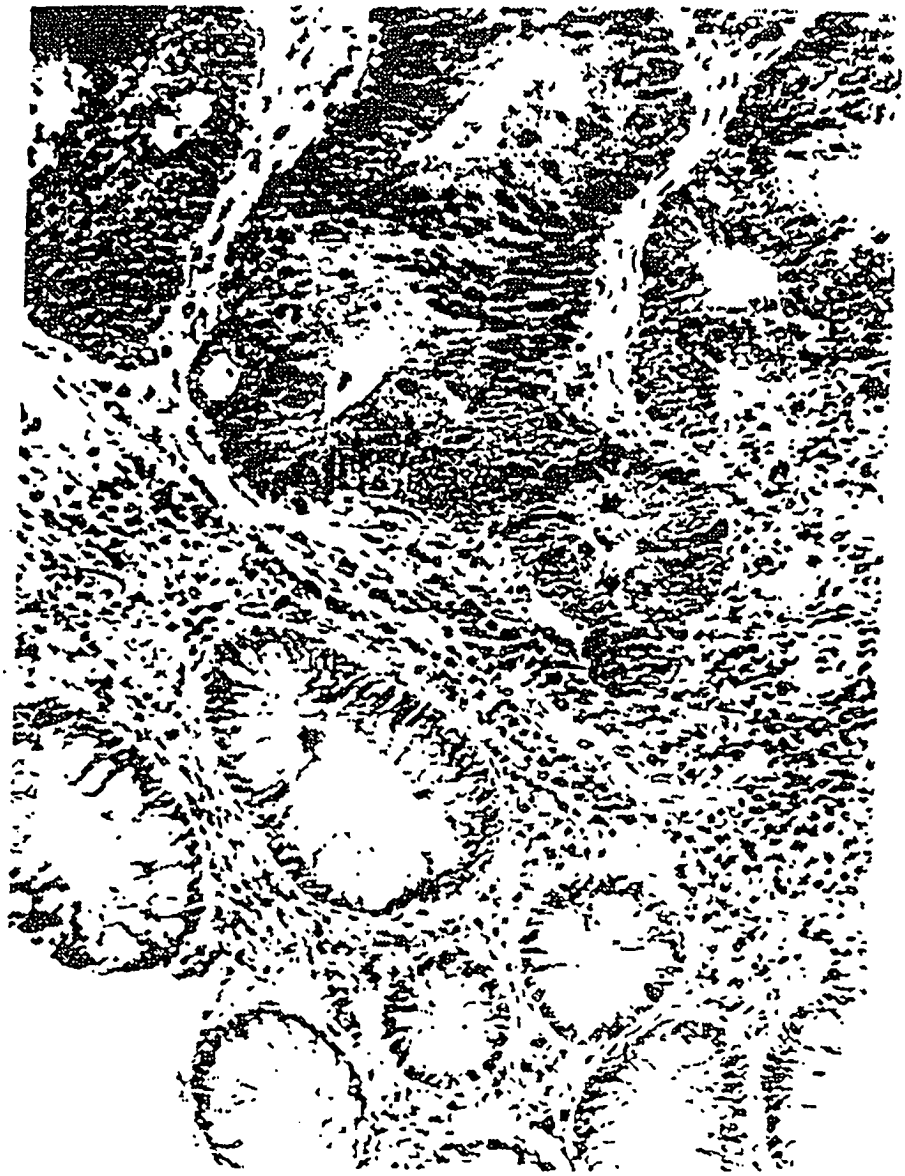


FIG. 15A



FIG. 15B



FIG. 16A



FIG. 16B

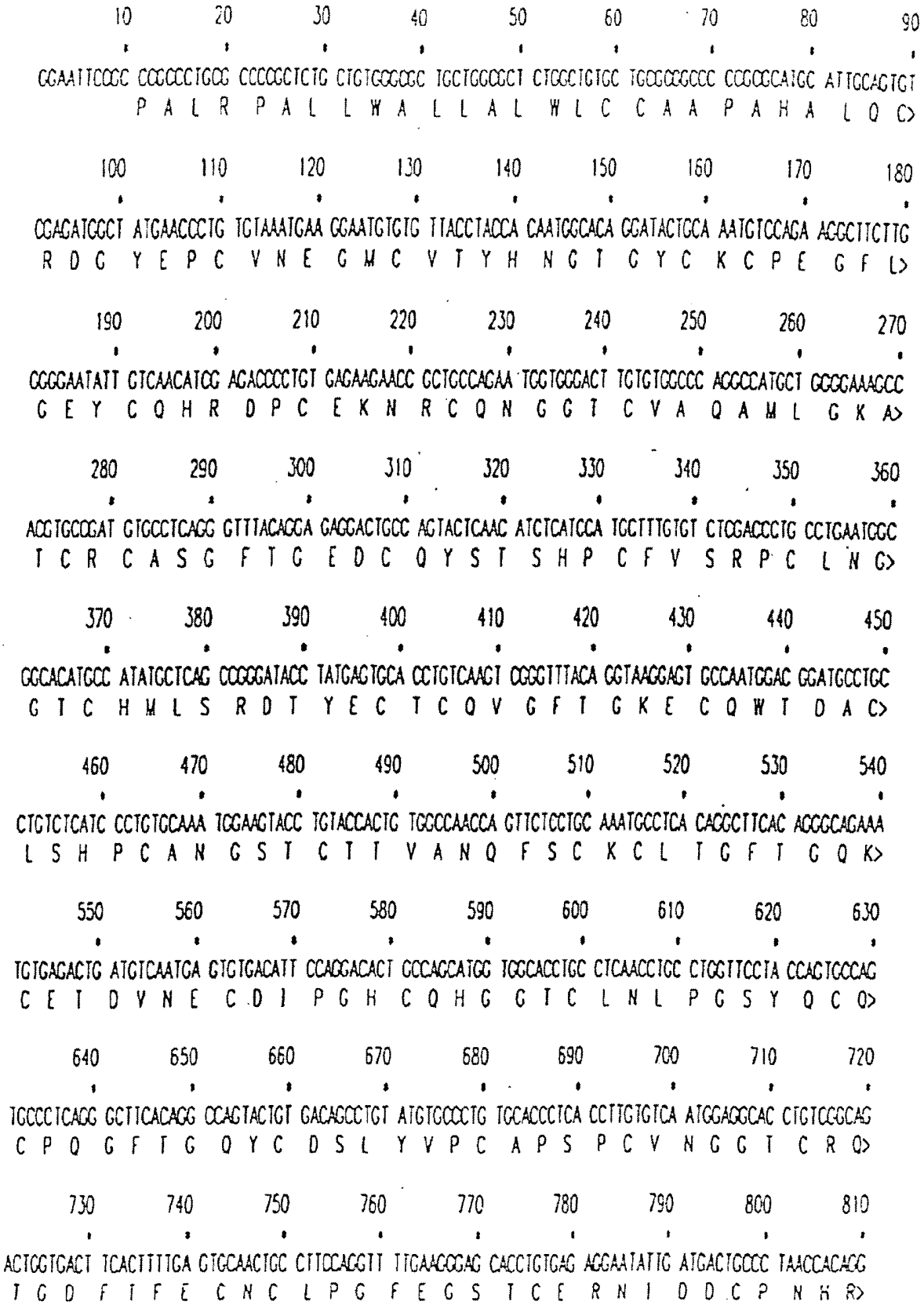
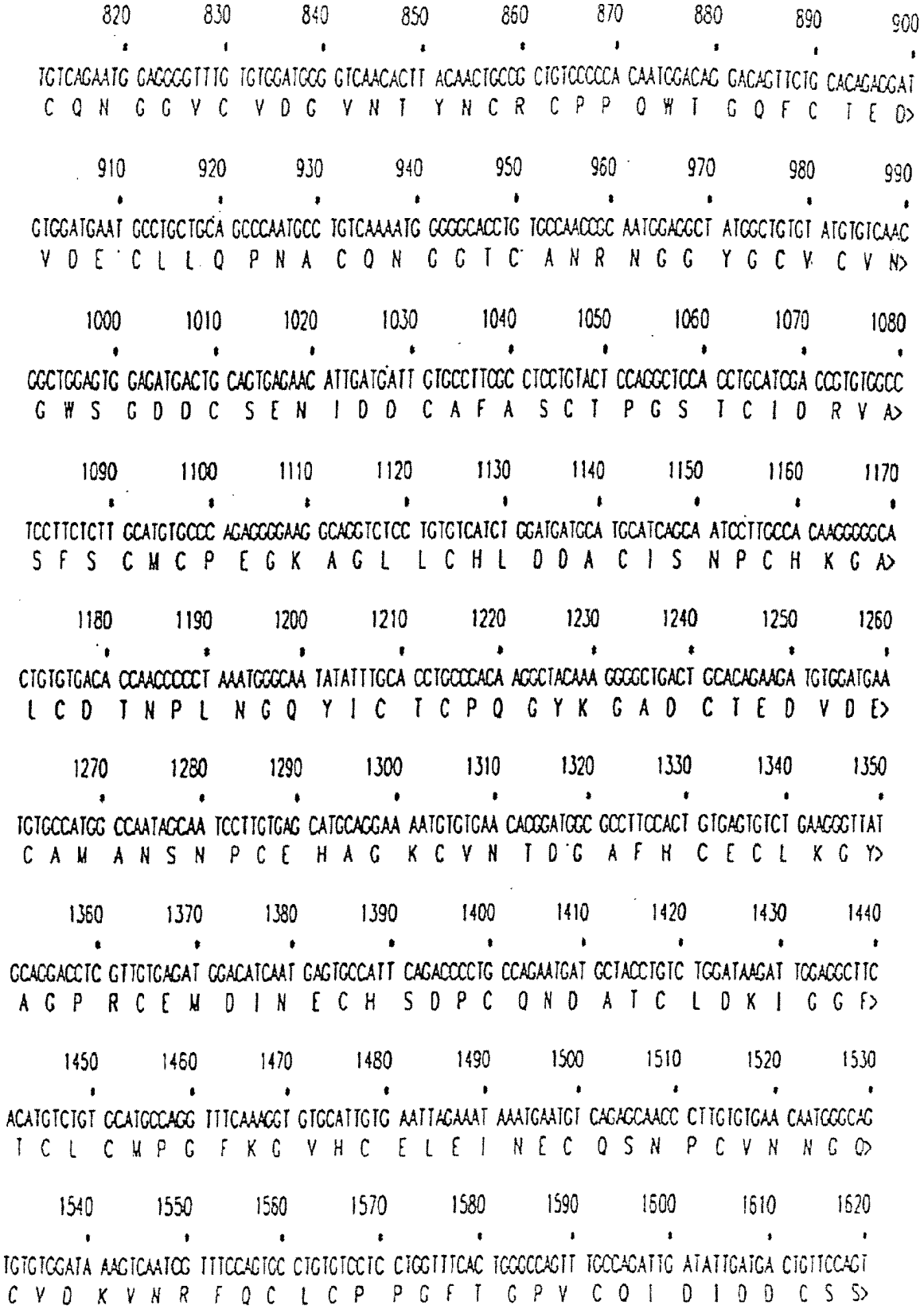

10 20 30 40 50 60 70 80 90
GGAATTCCGC CCGCCCTGCG CCCCGCTCTG CTGTGCGCGC TGCTGGCGCT CTGCCTGTGC TGCCGCGCCC CCGCCCATGC ATTGCAGTGT
P A L R P A L L W A L L A L W L C C A A P A H A L Q C>
100 110 120 130 140 150 160 170 180
CGACATGCCT ATGAACCCTG TGTAAATGAA GCAATGTGTG TTACCTACCA CAATGGCACA GGATACTGCA AATGTCCAGA ACGCTTCTTG
R D G Y E P C V N E G M C V T Y H N G T G Y C K C P E G F L>
190 200 210 220 230 240 250 260 270
CGGGAATATT GTCAACATCG ACACCCCTGT GAGAAGAACC GCTGCCAGAA TGGTGGGACT TGTGTGCCCC ACGCCATGCT GCGCAAAGCC
G E Y C Q H R D P C E K N R C Q N G G T C V A Q A M L G K A>
280 290 300 310 320 330 340 350 360
ACGTGCCGAT GTGCCTCAGG GTTTACAGGA GAGGACTGCC AGTACTCAAC ATCTCATCCA TGCTTTGTGT CTCGACCCTG CCTGAATGGC
T C R C A S G F T G E D C Q Y S T S H P C F V S R P C L N G>
370 380 390 400 410 420 430 440 450
GGCACATGCC ATATGCTCAG CCGGGATACC TATGAGTGCA CCTGTCAAGT CGGGTTTACA GGTAAGGAGT GCCAATGGAC GGATGCCTGC
G T C H M L S R D T Y E C T C Q V G F T G K E C Q W T D A C>
460 470 480 490 500 510 520 530 540
CTGTCTCATC CCTGTGCAAA TGGAAGTACC TGTACCACTG TGGCCAACCA GTTCTCCTGC AAATGCCTCA CAGGCTTCAC AGGGCAGAAA
L S H P C A N G S T C T T V A N Q F S C K C L T G F T G Q K>
550 560 570 580 590 600 610 620 630
TGTGAGACTG ATGTCAATGA GTGTGACATT CCAGGACACT GCCAGCATGG TGGCACCTGC CTCAACCTGC CTGGTTCCTA CCAGTGCCAG
C E T D V N E C D I P G H C Q H G G T C L N L P G S Y Q C Q>
640 650 660 670 680 690 700 710 720
TGCCCTCAGG GCTTCACAGG CCAGTACTGT GACAGCCTGT ATGTGCCCTG TGCACCCTCA CCTTGTGTCA ATGGAGGCAC CTGTCCGCAG
C P Q G F T G Q Y C D S L Y V P C A P S P C V N G G T C R Q>
730 740 750 760 770 780 790 800 810
ACTGGTGACT TCACTTTTGA GTGCAACTGC CTTCCACGTT TTGAAGGGAG CACCTGTGAG AGGAATATTG ATGACTGCCC TAACCACAGG
T G D F T F E C N C L P G F E G S T C E R N I D D C P N H R>

FIG.17A


820 830 840 850 860 870 880 890 900
TGTCAGAATG GACGCGTTTG TGTGGATGCG GTCAACACTT ACAACTGCCG CTGTCCCCCA CAATCGACAC GACAGTTCTG CACAGACGAT
C Q N G G V C V D G V N T Y N C R C P P Q W T G Q F C T E D>
910 920 930 940 950 960 970 980 990
GTGGATGAAT GCCTGCTGCA GCCCAATGCC TGTCAAAATG GGGGCACCTG TGCCAACCGC AATGGACGCT ATGGCTGTGT ATGTGTCAAC
V D E C L L Q P N A C Q N G G T C A N R N G G Y G C V C V N>
1000 1010 1020 1030 1040 1050 1060 1070 1080
GGCTGGAGTG GAGATGACTG CAGTGAGAAC ATTGATGATT GTGCCTTCGC CTCCTGTACT CCAGGCTCCA CCTGCATCGA CCGTGTGGCC
G W S G D D C S E N I D D C A F A S C T P G S T C I D R V A>
1090 1100 1110 1120 1130 1140 1150 1160 1170
TCCTTCTCTT GCATGTGCCC AGAGGGGAAG GCAGGTCTCC TGTGTCATCT GGATGATGCA TGCATCAGCA ATCCTTGCCA CAAGGGGCCA
S F S C M C P E G K A G L L C H L D D A C I S N P C H K G A>
1180 1190 1200 1210 1220 1230 1240 1250 1260
CTGTGTGACA CCAACCCCCT AAATGGGCAA TATATTTGCA CCTGCCCACA AGGCTACAAA GGGGCTGACT GCACAGAAGA TGTGGATGAA
L C D T N P L N G Q Y I C T C P Q G Y K G A D C T E D V D E>
1270 1280 1290 1300 1310 1320 1330 1340 1350
TGTGCCATGG CCAATAGCAA TCCTTGTGAG CATGCAGGAA AATGTGTGAA CAGGGATGGC GCCTTCCACT GTGAGTGTCT GAAGGGTTAT
C A M A N S N P C E H A G K C V N T D G A F H C E C L K G Y>
1360 1370 1380 1390 1400 1410 1420 1430 1440
GCAGGACCTC GTTGTGAGAT GGACATCAAT GAGTGCCATT CAGACCCCTG CCAGAATGAT GCTACCTGTC TGGATAAGAT TGGAGGCTTC
A G P R C E M D I N E C H S D P C Q N D A T C L D K I G G F>
1450 1460 1470 1480 1490 1500 1510 1520 1530
ACATGTCTGT GCATGCCAGG TTTCAAAGGT GTGCATTGTG AATTAGAAAT AAATGAATGT CAGAGCAACC CTTGTGTGAA CAATGGGCAG
T C L C M P G F K G V H C E L E I N E C Q S N P C V N N G Q>
1540 1550 1560 1570 1580 1590 1600 1610 1620
TGTGTGGATA AAGTCAATCG TTTCCAGTGC CTGTGTCCTC CTGGTTTCAC TGGGCCAGTT TGCCAGATTG ATATTGATGA CTGTTCCAGT
C V D K V N R F Q C L C P P G F T G P V C Q I D I D D C S S>
FIG.17B

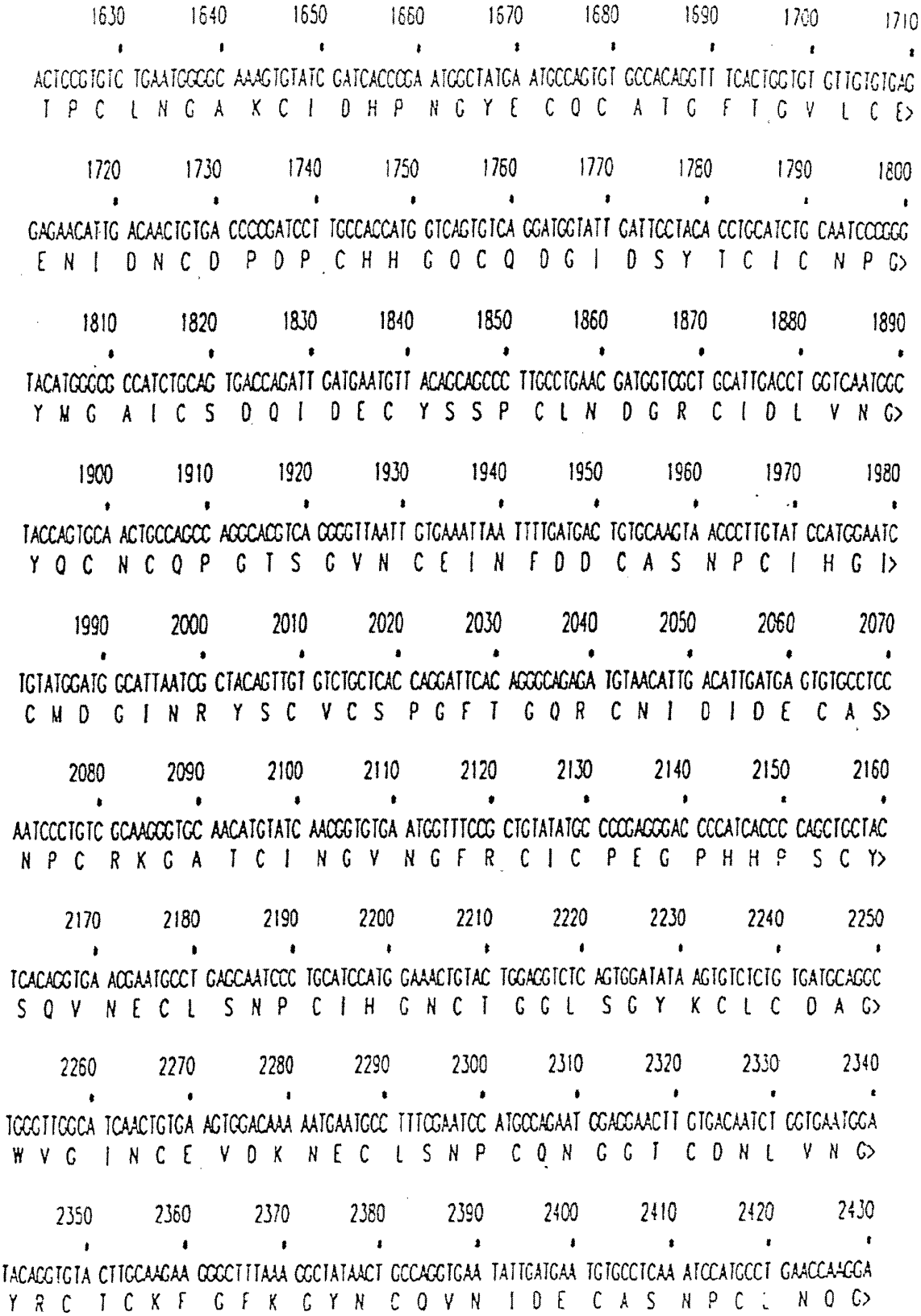

1630 1640 1650 1660 1670 1680 1690 1700 1710
ACTCCGTGTC TGAATGGGGC AAAGTGTATC GATCACCCGA ATGGCTATGA ATGCCAGTGT GCCACAGGTT TCACTGGTGT GTTGTGTGAG
T P C L N G A K C I D H P N G Y E C Q C A T G F T G V L C E>
1720 1730 1740 1750 1760 1770 1780 1790 1800
GAGAACATTG ACAACTGTGA CCCCGATCCT TGCCACCATG GTCAGTGTCA GGATGGTATT GATTCCTACA CCTGCATCTG CAATCCCGGG
E N I D N C D P D P C H H G Q C Q D G I D S Y T C I C N P G>
1810 1820 1830 1840 1850 1860 1870 1880 1890
TACATGGGGG CCATCTGCAG TGACCAGATT GATGAATGTT ACAGCAGCCC TTGCCTGAAC GATGGTCGCT GCATTGACCT GGTCAATGGC
Y M G A I C S D Q I D E C Y S S P C L N D G R C I D L V N G>
1900 1910 1920 1930 1940 1950 1960 1970 1980
TACCAGTGCA ACTGCCAGCC AGGCACGTCA GGGGTTAATT GTGAAATTAA TTTTGATGAC TGTGCAAGTA ACCCTTGTAT CCATGGAATC
Y Q C N C Q P G T S G V N C E I N F D D C A S N P C I H G I>
1990 2000 2010 2020 2030 2040 2050 2060 2070
TGTATGGATG GCATTAATGG CTACAGTTGT GTCTGCTCAC CAGGATTCAC AGGGCAGAGA TGTAACATTG ACATTGATGA GTGTGCCTCC
C M D G I N R Y S C V C S P G F T G Q R C N I D I D E C A S>
2080 2090 2100 2110 2120 2130 2140 2150 2160
AATCCCTGTC GCAAGGGTGC AACATGTATC AACGGTGTGA ATGGTTTCCG CTGTATATGC CCCGAGGGAC CCCATCACCC CAGCTGCTAC
N P C R K G A T C I N G V N G F R C I C P E G P H H P S C Y>
2170 2180 2190 2200 2210 2220 2230 2240 2250
TCACACGTGA ACGAATGCCT GAGCAATCCC TGCATCCATG GAAACTGTAC TGGACGTCTC AGTGGATATA AGTGTCTCTG TGATGCAGGC
S Q V N E C L S N P C I H G N C T G G L S G Y K C L C D A G>
2260 2270 2280 2290 2300 2310 2320 2330 2340
TGGGTTGGCA TCAACTGTGA AGTGGACAAA AATGAATGCC TTTCGAATCC ATGCCAGAAT CGACGAACTT GTGACAATCT CGTGAATGGA
W V G I N C E V D K N E C L S N P C Q N G G T C D N L V N G>
2350 2360 2370 2380 2390 2400 2410 2420 2430
TACAGGTGTA CTTGCAAGAA GGGCTTTAAA GGCTATAACT GCCAGGTGAA TATTGATGAA TGTGCCTCAA ATCCATGCCT GAACCAAGGA
Y R C T C K F G F K G Y N C Q V N I D E C A S N P C L N Q G>
N

FIG.17C

```
2440       2450       2460       2470       2480       2490       2500       2510       2520
ACCTGCTTTG ATGACATAAG TGGCTACACT TGCCACTGTG TGCTGCCATA CACAGGCAAG AATTGTCAGA CAGTATTGGC TCCCTGTTCC
T C F  D D I S  G Y T  C H C  V L P Y  T G K  N C Q  T V L  A P C  S>

2530       2540       2550       2560       2570       2580       2590       2600       2610
CCAAACCCTT GTGAGAATGC TGCTGTTTGC AAAGAGTCAC CAAATTTTGA GAGTTATACT TGCTTGTGTG CTCCTGGCTG GCAAGGTCAG
P N P  C E N A  A V C  K E S  P N F E  S Y T  C L C  A P G W  Q G Q>

2620       2630       2640       2650       2660       2670       2680       2690       2700
CGGTGTACCA TTGACATTGA CGAGTGTATC TCCAAGCCCT GCATGAACCA TGGTCTCTGC CATAACACCC AGGGCAGCTA CATGTGTGAA
R C T  I D I D  E C I  S K P  C M N H  G L C  H N T  Q G S Y  M C E>

2710       2720       2730       2740       2750       2760       2770       2780       2790
TGTCCACCAG GCTTCAGTGG TATGGACTGT GAGGAGGACA TTGATGACTG CCTTGCCAAT CCTTGCCAGA ATGGAGGTTC CTGTATGGAT
C P P  G F S G  M D C  E E D  I D D C  L A N  P C Q  N G G S  C M D>

2800       2810       2820       2830       2840       2850       2860       2870       2880
GGAGTGAATA CTTTCTCCTG CCTCTGCCTT CCGGGTTTCA CTGGGGATAA GTGCCAGACA GACATGAATG AGTGTCTGAG TGAACCCTGT
G V N  T F S C  L C L  P G F  T G D K  C Q T  D M N  E C L S  E P C>

2890       2900       2910       2920       2930       2940       2950       2960       2970
AAGAATGGAG GGACCTGCTC TGACTACGTC AACAGTTACA CTTGCAAGTG CCAGGCAGGA TTTGATGGAG TCCATTGTGA GAACAACATC
K N G  G T C S  D Y V  N S Y  T C K C  Q A G  F D G  V H C E  N N I>

2980       2990       3000       3010       3020       3030       3040       3050       3060
AATGAGTGCA CTGAGAGCTC CTGTTTCAAT GGTGGCACAT GTGTTGATGG GATTAACTCC TTCTCTTGCT TGTGCCCTGT GGGTTTCACT
N E C  T E S S  C F N  G G T  C V D G  I N S  F S C  L C P V  G F T>

3070       3080       3090       3100       3110       3120       3130       3140       3150
GGATCCTTCT GCCTCCATGA GATCAATGAA TGCAGCTCTC ATCCATGCCT GAATGAGGGA ACGTGTGTTG ATGGCCTGGG TACCTACCGC
G S F  C L H E  I N E  C S S  H P C L  N E G  T C V  D G L G  T Y R>

3160       3170       3180       3190       3200       3210       3220       3230       3240
TGCAGCTGCC CCCTGGGCTA CACTGGGAAA AACTGTCAGA CCCTGGTGAA TCTCTGCAGT CGGTCTCCAT GTAAAAACAA AGGTACTTGT
C S C  P L G Y  T G K  N C Q  T L V N  L C S  R S P  C K N K  G T C>
```

FIG.17D

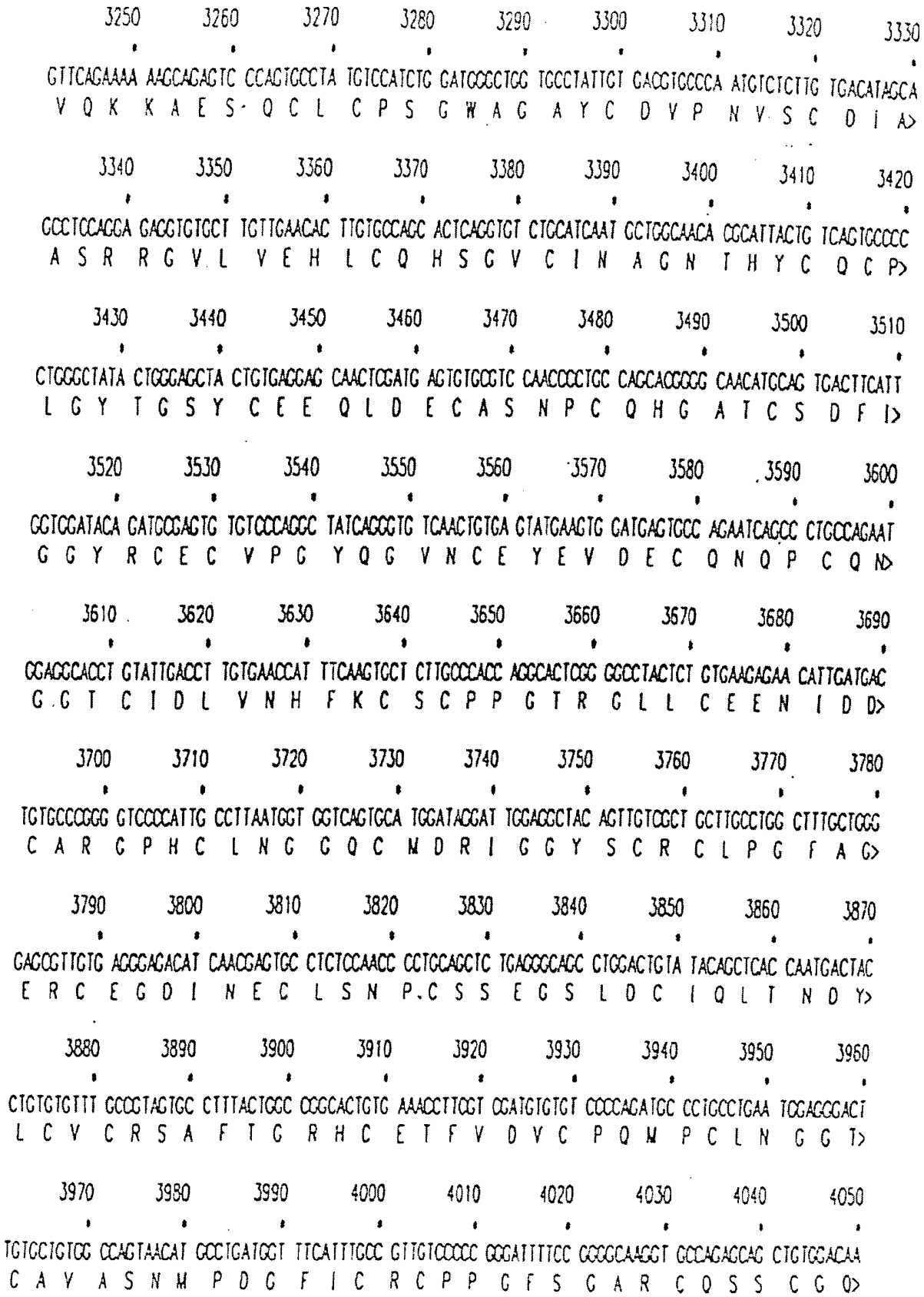

3250 3260 3270 3280 3290 3300 3310 3320 3330
GTTCAGAAAA AAGCAGAGTC CCAGTGCCTA TGTCCATCTG GATGCCCTGG TGCCTATTGT GACCTGCCCA ATGTCTCTTG TGACATAGCA
V Q K K A E S Q C L C P S G W A G A Y C D V P N V S C D I A>
3340 3350 3360 3370 3380 3390 3400 3410 3420
GCCTCCAGGA GACGTGTGCT TGTTGAACAC TTGTGCCAGC ACTCAGGTGT CTGCATCAAT GCTGGCAACA CGCATTACTG TCAGTGCCCC
A S R R G V L V E H L C Q H S G V C I N A G N T H Y C Q C P>
3430 3440 3450 3460 3470 3480 3490 3500 3510
CTGGGCTATA CTGGCAGCTA CTGTGAGGAG CAACTCGATG AGTGTGCGTC CAACCCCTGC CAGCACGGGG CAACATCCAG TGACTTCATT
L G Y T G S Y C E E Q L D E C A S N P C Q H G A T C S D F I>
3520 3530 3540 3550 3560 3570 3580 3590 3600
GGTCGATACA GATGCGAGTG TGTCCCAGGC TATCAGGGTG TCAACTGTGA GTATGAAGTG GATGAGTGCC AGAATCAGCC CTGCCAGAAT
G G Y R C E C V P G Y Q G V N C E Y E V D E C Q N Q P C Q N>
3610 3620 3630 3640 3650 3660 3670 3680 3690
GGAGGCACCT GTATTGACCT TGTGAACCAT TTCAAGTGCT CTTGCCCACC AGGCACTCGG GGCCTACTCT GTGAAGAGAA CATTGATGAC
G G T C I D L V N H F K C S C P P G T R G L L C E E N I D D>
3700 3710 3720 3730 3740 3750 3760 3770 3780
TGTGCCCGGG GTCCCCATTG CCTTAATGGT GGTCAGTGCA TGGATACGAT TGGAGGCTAC AGTTGTCGCT GCTTGCCTGG CTTTGCTGGG
C A R C P H C L N G G Q C M D R I G G Y S C R C L P G F A G>
3790 3800 3810 3820 3830 3840 3850 3860 3870
GAGCGTTGTG AGGGAGACAT CAACGAGTGC CTCTCCAACC CCTGCAGCTC TGAGGGCAGC CTGGACTGTA TACAGCTCAC CAATGACTAC
E R C E G D I N E C L S N P C S S E G S L D C I Q L T N D Y>
3880 3890 3900 3910 3920 3930 3940 3950 3960
CTGTGTGTTT GCCGTAGTGC CTTTACTGGC CGGCACTGTG AAACCTTCGT CGATGTGTGT CCCCAGATGC CCTGCCTGAA TGGAGGGACT
L C V C R S A F T G R H C E T F V D V C P Q M P C L N G G T>
3970 3980 3990 4000 4010 4020 4030 4040 4050
TGTGCTGTGG CCAGTAACAT GCCTGATGGT TTCATTTGCC GTTGTCCCCC GGGATTTTCC GCGGCAAGGT GCCAGAGCAC CTGTGGACAA
C A V A S N M P D G F I C R C P P G F S G A R C Q S S C G Q>

FIG.17E

```
      4060      4070      4080      4090      4100      4110      4120      4130      4140
         *         *         *         *         *         *         *         *         *
GTGAAATGTA GGAAGGGGGA GCAGTGTGTG CACACCGCCT CTGGACCCCG CTGCTTCTGC CCCAGTCCCC GGGACTGCGA GTCAGGCTGT
 V  K  C   R  K  G  E   Q  C  V   H  T  A   S  G  P  R   C  F  C   P  S  P   R  D  C  E   S  G  C>

      4150      4160      4170      4180      4190      4200      4210      4220      4230
         *         *         *         *         *         *         *         *         *
GCCAGTAGCC CCTGCCAGCA CGGGGGCAGC TGCCACCCTC AGCGCCAGCC TCCTTATTAC TCCTGCCAGT GTGCCCCACC ATTCTCGGGT
 A  S  S   P  C  Q  H   G  G  S   C  H  P   Q  R  Q  P   P  Y  Y   S  C  Q   C  A  P  P   F  S  G>

      4240      4250      4260      4270      4280      4290      4300      4310      4320
         *         *         *         *         *         *         *         *         *
AGCCGCTGTG AACTCTACAC GGCACCCCCC AGCACCCCTC CTGCCACCTG TCTGAGCCAG TATTGTGCCG ACAAAGCTCG GGATGGCGTC
 S  R  C   E  L  Y  T   A  P  P   S  T  P   P  A  T  C   L  S  Q   Y  C  A   D  K  A  R   D  G  V>

      4330      4340      4350      4360      4370      4380      4390      4400      4410
         *         *         *         *         *         *         *         *         *
TGTGATGAGG CCTGCAACAG CCATGCCTGC CAGTGGGATG GGGGTGACTG TTCTCTCACC ATGGAGAACC CCTGGGCCAA CTGCTCCTCC
 C  D  E   A  C  N  S   H  A  C   Q  W  D   G  G  D  C   S  L  T   M  E  N   P  W  A  N   C  S  S>

      4420      4430      4440      4450      4460      4470      4480      4490      4500
         *         *         *         *         *         *         *         *         *
CCACTTCCCT GCTGGGATTA TATCAACAAC CAGTGTGATG AGCTGTGCAA CACGGTCGAG TGCCTGTTTG ACAACTTTGA ATGCCAGGGG
 P  L  P   C  W  D  Y   I  N  N   Q  C  D   E  L  C  N   T  V  E   C  L  F   D  N  F  E   C  Q  G>

      4510      4520      4530      4540      4550      4560      4570      4580      4590
         *         *         *         *         *         *         *         *         *
AACAGCAAGA CATGCAAGTA TGACAAATAC TGTGCAGACC ACTTCAAAGA CAACCACTGT AACCAGGGGT GCAACAGTGA GGAGTGTGGT
 N  S  K   T  C  K  Y   D  K  Y   C  A  D   H  F  K  D   N  H  C   N  Q  G   C  N  S  E   E  C  G>

      4600      4610      4620      4630      4640      4650      4660      4670      4680
         *         *         *         *         *         *         *         *         *
TGGGATGGGC TGGACTGTGC TGCTGACCAA CCTGAGAACC TGGCAGAAGG TACCCTGGTT ATTGTGGTAT TGATGCCACC TGAACAACTG
 W  D  G   L  D  C  A   A  D  Q   P  E  N   L  A  E  G   T  L  V   I  V  V   L  M  P  P   E  Q  L>

      4690      4700      4710      4720      4730      4740      4750      4760      4770
         *         *         *         *         *         *         *         *         *
CTCCAGGATG CTCGCAGCTT CTTGCGGGCA CTGGGTACCC TGCTCCACAC CAACCTGCGC ATTAAGCGGG ACTCCCAGGG GGAACTCATG
 L  Q  D   A  R  S  F   L  R  A   L  G  T   L  L  H  T   N  L  R   I  K  R   D  S  Q  G   E  L  M>

      4780      4790      4800      4810      4820      4830      4840      4850      4860
         *         *         *         *         *         *         *         *         *
GTGTACCCCT ATTATGGTGA GAAGTCAGCT GCTATGAAGA AACAGAGGAT GACACGCAGA TCCCTTCCTG GTGAACAAGA ACAGGAGGTG
 V  Y  P   Y  Y  G  E   K  S  A   A  M  K   K  Q  R  M   T  R  R   S  L  P   G  E  Q  E   Q  E  V>
```

FIG.17F

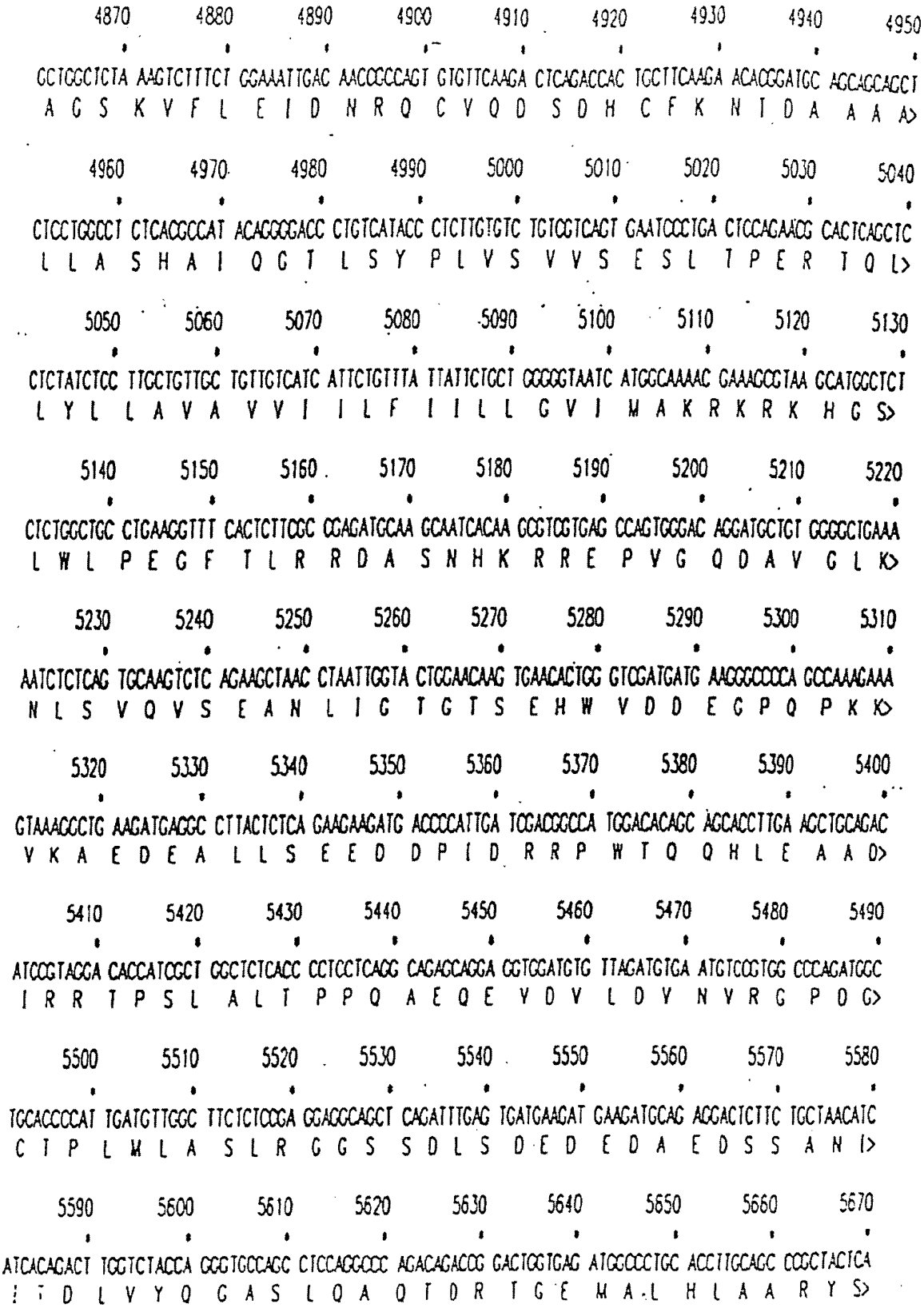

4870 4880 4890 4900 4910 4920 4930 4940 4950
GCTGGCTCTA AAGTCTTTCT GGAAATTGAC AACCGCCAGT GTGTTCAAGA CTCAGACCAC TGCTTCAAGA ACACCGATGC AGCAGCAGCT
A G S K V F L E I D N R Q C V Q D S D H C F K N T D A A A A>
4960 4970 4980 4990 5000 5010 5020 5030 5040
CTCCTGGCCT CTCACGCCAT ACAGGGGACC CTGTCATACC CTCTTGTGTC TGTCGTCAGT GAATCCCTGA CTCCAGAACG CACTCAGCTC
L L A S H A I Q G T L S Y P L V S V V S E S L T P E R T Q L>
5050 5060 5070 5080 5090 5100 5110 5120 5130
CTCTATCTCC TTGCTGTTGC TGTTGTCATC ATTCTGTTTA TTATTCTGCT GGGGGTAATC ATGGCAAAAC GAAAGCGTAA GCATGGCTCT
L Y L L A V A V V I I L F I I L L G V I M A K R K R K H G S>
5140 5150 5160 5170 5180 5190 5200 5210 5220
CTCTGGCTGC CTGAAGGTTT CACTCTTCGC CGAGATGCAA GCAATCACAA GCGTCGTGAG CCAGTGGGAC AGGATGCTGT GGGGCTGAAA
L W L P E G F T L R R D A S N H K R R E P V G Q D A V G L K>
5230 5240 5250 5260 5270 5280 5290 5300 5310
AATCTCTCAG TGCAAGTCTC AGAAGCTAAC CTAATTGGTA CTCGAACAAG TGAACACTGG GTCGATGATG AAGGGCCCCA GCCAAACAAA
N L S V Q V S E A N L I G T G T S E H W V D D E G P Q P K K>
5320 5330 5340 5350 5360 5370 5380 5390 5400
GTAAAGCCTG AAGATGAGGC CTTACTCTCA GAAGAAGATG ACCCCATTGA TCGACGGCCA TGGACACAGC AGCACCTTGA ACCTGCAGAC
V K A E D E A L L S E E D D P I D R R P W T Q Q H L E A A D>
5410 5420 5430 5440 5450 5460 5470 5480 5490
ATCCGTAGCA CACCATCGCT GGCTCTCACC CCTCCTCAGG CAGAGCAGGA CGTGGATGTG TTACATGTGA ATGTCCGTGG CCCAGATGGC
I R R T P S L A L T P P Q A E Q E V D V L D V N V R G P D G>
5500 5510 5520 5530 5540 5550 5560 5570 5580
TGCACCCCAT TGATGTTGGC TTCTCTCCGA GGAGGCAGCT CAGATTTGAG TGATGAAGAT GAAGATGCAG ACGACTCTTC TGCTAACATC
C T P L M L A S L R G G S S D L S D E D E D A E D S S A N I>
5590 5600 5610 5620 5630 5640 5650 5660 5670
ATCACAGACT TGGTCTACCA GGGTGCCAGC CTCCAGGCCC AGACAGACCG GACTGGTGAG ATGGCCCTGC ACCTTGCAGC CGGCTACTCA
I T D L V Y Q G A S L Q A Q T D R T G E M A L H L A A R Y S>

FIG.17G

```
      5680       5690       5700       5710       5720       5730       5740       5750       5760
       .          .          .          .          .          .          .          .          .
CCGGCTGATG CTGCCAAGCG TCTCCTGGAT GCACGTGCAG ATGCCAATGC CCAGGACAAC ATCGGCCGCT GTCCACTCCA TGCTGCAGTG
 R  A  D  A  A  K  R  L  L  D  A  G  A  D  A  N  A  Q  D  N  M  G  R  C  P  L  H  A  A  V>

      5770       5780       5790       5800       5810       5820       5830       5840       5850
       .          .          .          .          .          .          .          .          .
GCAGCTGATG CCCAAGGTGT CTTCCAGATT CTGATTCGCA ACCGAGTAAC TGATCTAGAT GCCAGGATGA ATGATGGTAC TACACCCCTG
 A  A  D  A  Q  G  V  F  Q  I  L  I  R  N  R  V  T  D  L  D  A  R  M  N  D  G  T  T  P  L>

      5860       5870       5880       5890       5900       5910       5920       5930       5940
       .          .          .          .          .          .          .          .          .
ATCCTGCCTG CCCGCCTGGC TGTGGAGGGA ATGGTGGCAG AACTGATCAA CTGCCAAGCG GATGTGAATG CAGTGGATGA CCATGGAAAA
 I  L  A  A  R  L  A  V  E  G  M  V  A  E  L  I  N  C  Q  A  D  V  N  A  V  D  D  H  G  K>

      5950       5960       5970       5980       5990       6000       6010       6020       6030
       .          .          .          .          .          .          .          .          .
TCTGCTCTTC ACTGGGCAGC TGCTGTCAAT AATGTGGAGG CAACTCTTTT GTTGTTGAAA AATGGGGCCA ACCGAGACAT GCACGACAAC
 S  A  L  H  W  A  A  A  V  N  N  V  E  A  T  L  L  L  L  K  N  G  A  N  R  D  M  Q  D  N>

      6040       6050       6060       6070       6080       6090       6100       6110       6120
       .          .          .          .          .          .          .          .          .
AAGGAAGAGA CACCTCTGTT TCTTGCTGCC CGGGAGGGGA GCTATGAAGC AGCCAAGATC CTGTTAGACC ATTTTGCCAA TCGAGACATC
 K  E  E  T  P  L  F  L  A  A  R  E  G  S  Y  E  A  A  K  I  L  L  D  H  F  A  N  R  D  I>

      6130       6140       6150       6160       6170       6180       6190       6200       6210
       .          .          .          .          .          .          .          .          .
ACAGACCATA TGGATCGTCT TCCCCGGGAT GTGGCTCCGG ATCGCATGCA CCATGACATT GTGCGCCTTC TGGATGAATA CAATGTGACC
 T  D  H  M  D  R  L  P  R  D  V  A  R  D  R  M  H  H  D  I  V  R  L  L  D  E  Y  N  V  T>

      6220       6230       6240       6250       6260       6270       6280       6290       6300
       .          .          .          .          .          .          .          .          .
CCAAGCCCTC CAGGCACCGT GTTGACTTCT GCTCTCTCAC CTGTCATCTG TGGGCCCAAC AGATCTTTCC TCAGCCTGAA GCACACCCCA
 P  S  P  P  G  T  V  L  T  S  A  L  S  P  V  I  C  G  P  N  R  S  F  L  S  L  K  H  T  P>

      6310       6320       6340       6350       6360       6370       6380       6390       6400
       .          .          .          .          .          .          .          .          .
ATGGGCAAGA AGTCTAGACG GCCCAGTGCC AAGAGTACCA TGCCTACTAG CCTCCCTAAC CTTGCCAAGG AGGCAAAGGA TGCCAAGGGT
 M  G  K  K  S  R  R  P  S  A  K  S  T  M  P  T  S  L  P  N  L  A  K  E  A  K  D  A  K  G>

      6400       6410       6420       6430       6440       6450       6460       6470       6480
       .          .          .          .          .          .          .          .          .
AGTACGACGA AGAAGTCTCT GAGTGAGAAG GTCCAACTGT CTGAGAGTTC AGTAACTTTA TCCCCTGTTG ATTCCCTAGA ATCTCCTCAC
 S  R  R  K  K  S  L  S  E  K  V  Q  L  S  E  S  S  V  T  L  S  P  V  D  S  L  E  S  P  H>
```

FIG.17H

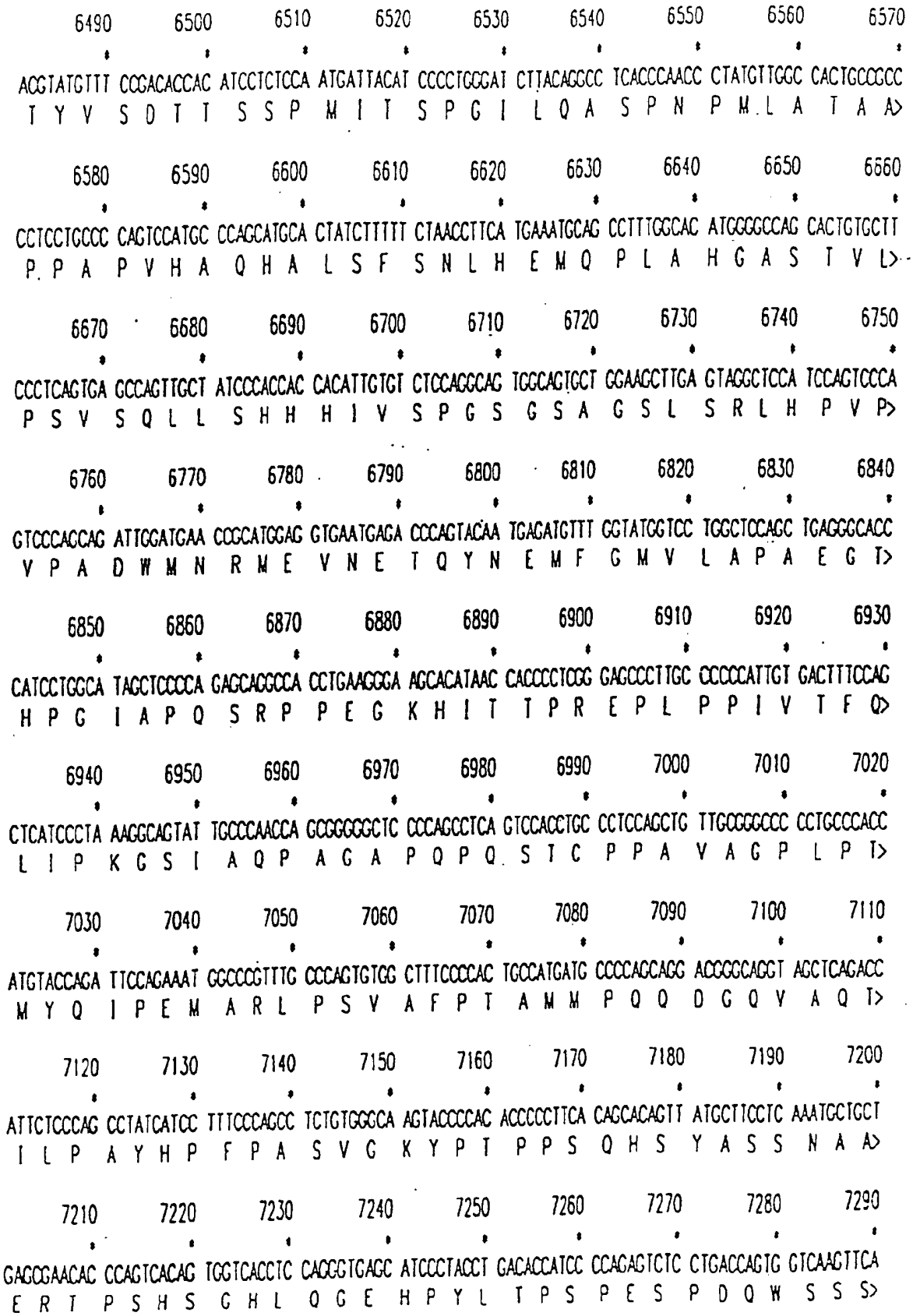

6490 6500 6510 6520 6530 6540 6550 6560 6570
ACGTATGTTT CCGACACCAC ATCCTCTCCA ATGATTACAT CCCCTGCGAT CTTACAGGCC TCACCCAACC CTATGTTGGC CACTGCCCCC
T Y V S D T T S S P M I T S P G I L Q A S P N P M L A T A A>
6580 6590 6600 6610 6620 6630 6640 6650 6660
CCTCCTGCCC CAGTCCATGC CCAGCATGCA CTATCTTTTT CTAACCTTCA TGAAATGCAG CCTTTGGCAC ATGGGGCCAG CACTGTGCTT
P P A P V H A Q H A L S F S N L H E M Q P L A H G A S T V L>
6670 6680 6690 6700 6710 6720 6730 6740 6750
CCCTCAGTGA GCCAGTTGCT ATCCCACCAC CACATTGTGT CTCCAGGCAG TGGCAGTGCT GGAAGCTTGA GTAGGCTCCA TCCAGTCCCA
P S V S Q L L S H H H I V S P G S G S A G S L S R L H P V P>
6760 6770 6780 6790 6800 6810 6820 6830 6840
GTCCCACCAG ATTGGATGAA CCGCATGGAG GTGAATGAGA CCCAGTACAA TGAGATGTTT GGTATGGTCC TGGCTCCAGC TGAGGGCACC
V P A D W M N R M E V N E T Q Y N E M F G M V L A P A E G T>
6850 6860 6870 6880 6890 6900 6910 6920 6930
CATCCTGGCA TAGCTCCCCA GAGCAGGCCA CCTGAAGGGA AGCACATAAC CACCCCTCGG GAGCCCTTGC CCCCCATTGT GACTTTCCAG
H P G I A P Q S R P P E G K H I T T P R E P L P P I V T F Q>
6940 6950 6960 6970 6980 6990 7000 7010 7020
CTCATCCCTA AAGGCAGTAT TGCCCAACCA GCGGGGGCTC CCCAGCCTCA GTCCACCTGC CCTCCAGCTG TTGCGGGCCC CCTGCCCACC
L I P K G S I A Q P A G A P Q P Q S T C P P A V A G P L P T>
7030 7040 7050 7060 7070 7080 7090 7100 7110
ATGTACCAGA TTCCAGAAAT GGCCCGTTTC CCCAGTGTGG CTTTCCCCAC TGCCATGATG CCCCAGCAGG ACGGGCAGGT AGCTCAGACC
M Y Q I P E M A R L P S V A F P T A M M P Q Q D G Q V A Q T>
7120 7130 7140 7150 7160 7170 7180 7190 7200
ATTCTCCCAG CCTATCATCC TTTCCCAGCC TCTGTGGGCA AGTACCCCAC ACCCCCTTCA CAGCACAGTT ATGCTTCCTC AAATGCTGCT
I L P A Y H P F P A S V G K Y P T P P S Q H S Y A S S N A A>
7210 7220 7230 7240 7250 7260 7270 7280 7290
GAGCGAACAC CCAGTCACAG TGGTCACCTC CACGGTGAGC ATCCCTACCT GACACCATCC CCAGAGTCTC CTGACCAGTG GTCAAGTTCA
E R T P S H S G H L Q G E H P Y L T P S P E S P D Q W S S S>

FIG.17I

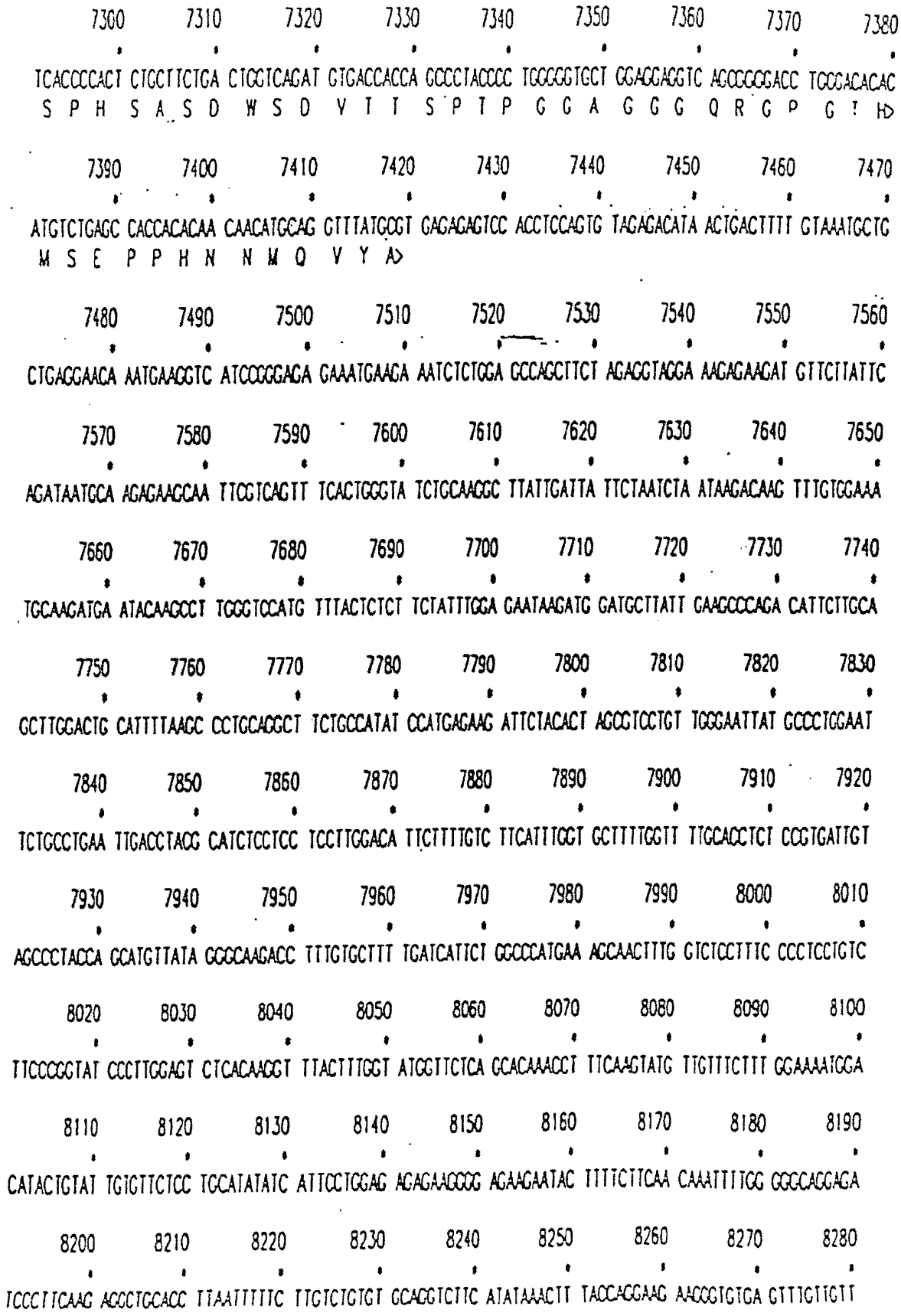

7300 7310 7320 7330 7340 7350 7360 7370 7380
TCACCCCACT CTGCTTCTGA CTGGTCAGAT GTGACCACCA GCCCTACCCC TGGGGGTGCT GGAGGAGGTC AGCGCGGGACC TGCGACACAC
S P H S A S D W S D V T T S P T P G G A G G G Q R G P G T H>
7390 7400 7410 7420 7430 7440 7450 7460 7470
ATGTCTGAGC CACCACACAA CAACATGCAG GTTTATGCGT GAGAGAGTCC ACCTCCAGTG TAGAGACATA ACTGACTTTT GTAAATGCTG
M S E P P H N N M Q V Y A>
7480 7490 7500 7510 7520 7530 7540 7550 7560
CTGAGGAACA AATGAAGGTC ATCCGGGAGA GAAATGAAGA AATCTCTGGA GCCAGCTTCT AGAGGTAGGA AAGAGAAGAT GTTCTTATTC
7570 7580 7590 7600 7610 7620 7630 7640 7650
AGATAATGCA AGAGAAGCAA TTCGTCAGTT TCACTGGGTA TCTGCAAGGC TTATTGATTA TTCTAATCTA ATAAGACAAG TTTGTGGAAA
7660 7670 7680 7690 7700 7710 7720 7730 7740
TGCAAGATGA ATACAAGCCT TGGGTCCATG TTTACTCTCT TCTATTTGGA GAATAAGATG GATGCTTATT GAAGCCCAGA CATTCTTGCA
7750 7760 7770 7780 7790 7800 7810 7820 7830
GCTTGGACTG CATTTTAAGC CCTGCAGGCT TCTGCCATAT CCATGAGAAG ATTCTACACT AGCGTCCTGT TGGGAATTAT GCCCTGGAAT
7840 7850 7860 7870 7880 7890 7900 7910 7920
TCTGCCTGAA TTGACCTACC CATCTCCTCC TCCTTGGACA TTCTTTTGTC TTCATTTGGT GCTTTTGGTT TTGCACCTCT CCGTGATTGT
7930 7940 7950 7960 7970 7980 7990 8000 8010
AGCCCTACCA GCATGTTATA GCGCAAGACC TTTGTGCTTT TGATCATTCT GGCCCATGAA AGCAACTTTG GTCTCCTTTC CCCTCCTGTC
8020 8030 8040 8050 8060 8070 8080 8090 8100
TTCCCGGTAT CCCTTGGAGT CTCACAAGGT TTACTTTGGT ATGGTTCTCA GCACAAACCT TTCAAGTATG TTGTTTCTTT GGAAAATGGA
8110 8120 8130 8140 8150 8160 8170 8180 8190
CATACTGTAT TGTGTTCTCC TGCATATATC ATTCCTGGAG AGAGAAGGGG AGAAGAATAC TTTTCTTCAA CAAATTTTGG GGGCAGGAGA
8200 8210 8220 8230 8240 8250 8260 8270 8280
TCCCTTCAAG AGCCTGCACC TTAATTTTTC TTGTCTGTGT GCAGGTCTTC ATATAAACTT TACCAGGAAG AACCGTGTGA GTTTGTTGTT

FIG.17J

8290 8300 8310 8320 8330 8340 8350 8360 8370

TTTCTGTGTA TGGGCCTGGT CAGTGTAAAG TTTTATCCTT GATAGTCTAG TTACTATGAC CCTCCCCACT TTTTTAAAAC CACAAAAACG

8380 8390 8400 8410 8420 8430 8440 8450 8460

TTTGGAATGT TGGAATGACC AAGAGACAAG TTAACTCGTG CAAGAGCCAG TTACCCACCC ACAGGTCCCC CTACTTCCTG CCAAGCATTC

8470 8480 8490 8500 8510 8520 8530 8540 8550

CATTGACTGC CTGTATGGAA CACATTTGTC CCAGATCTGA GCATTCTAGG CCTGTTTCAC TCACTCACCC AGCATATGAA ACTAGTCTTA

8560 8570 8580 8590 8600 8610 8620 8630 8640

ACTGTTGAGC CTTTCCTTTC ATATCCACAG AAGACACTGT CTCAAATGTT GTACCCTTGC CATTTAGGAC TGAACTTTCC TTACCCCAAG

8650 8660 8670 8680 8690 8700 8710 8720 8730

GGACCCAGTG ACAGTTGTCT TCCGTTTGTC AGATGATCAG TCTCTACTGA TTATCTTGCT GCTTAAAGGC CTGCTCACCA ATCTTTCTTT

8740 8750 8760 8770 8780 8790 8800 8810 8820

CACACCGTGT GGTCCGTGTT ACTGGTATAC CCAGTATGTT CTCACTGAAG ACATGGACTT TATATGTTCA AGTGCACGAA TTGGAAAGTT

8830 8840 8850 8860 8870 8880 8890 8900 8910

GGACTTGTTT TCTATGATCC AAAACAGCCC TATAAGAAGG TTGGAAAAGG AGGAACTATA TACCAGCCTT TGCTATTTTC TGCTACCATT

8920 8930 8940 8950 8960 8970 8980 8990 9000

TCTTTTCCTC TGAAGCGGCC ATGACATTCC CTTTGGCAAC TAACGTAGAA ACTCAACACA ACATTTTCCT TTCCTAGAGT CACCTTTTAC

9010 9020 9030 9040 9050 9060 9070 9080 9090

ATGATAATGG ACAACTATAG ACTTGCTCAT TGTTCAGACT GATTGCCCCT CACCTGAATC CACTCTCTGT ATTCATGCTC TTGGCAATTT

9100 9110 9120 9130 9140 9150 9160 9170 9180

CTTTGACTTT CTTTTAAGGG CAGAAGCATT TTAGTTAATT GTAGATAAAG AATAGTTTTC TTCCTCTTCT CCTTGGGCCA GTTAATAATT

9190 9200 9210 9220 9230 9240 9250 9260 9270

GGTCCATGGC TACACTGCAA CTTCCGTCCA GTGCTGTGAT GCCCATGACA CCTGCAAAAT AAGTTCTGCC TGGGCATTTT GTAGATATTA

FIG.17K

```
      9280      9290      9300      9310      9320      9330      9340      9350      9360
         *         *         *         *         *         *         *         *         *
ACAGGTGAAT TCCCGACTCT TTTGGTTTGA ATGACAGTTC TCATTCCTTC TATGGCTGCA AGTATGCATC AGTGCTTCCC ACTTACCTGA

      9370      9380      9390      9400      9410      9420      9430      9440      9450
         *         *         *         *         *         *         *         *         *
TTTGTCTGTC GGTGGCCCCA TATGGAAACC CTGCGTGTCT GTTGGCATAA TAGTTTACAA ATGGTTTTTT CAGTCCTATC CAAATTTATT

      9460      9470      9480      9490      9500      9510      9520      9530      9540
         *         *         *         *         *         *         *         *         *
GAACCAACAA AAATAATTAC TTCTGCCCTG AGATAAGCAG ATTAAGTTTG TTCATTCTCT GCTTTATTCT CTCCATGTGG CAACATTCTG

      9550      9560      9570      9580      9590      9600      9610      9620      9630
         *         *         *         *         *         *         *         *         *
TCAGCCTCTT TCATAGTGTG CAAACATTTT ATCATTCTAA ATGGTGACTC TCTGCCCTTG GACCCATTTA TTATTCACAG ATGGGGAGAA

      9640      9650      9660      9670      9680      9690      9700      9710      9720
         *         *         *         *         *         *         *         *         *
CCTATCTGCA TGGACCCTCA CCATCCTCTG TGCAGCACAC ACAGTGCAGG GAGCCAGTGG CGATGGCGAT GACTTTCTTC CCCTGGGAAT

TCC
```

FIG.17L