



(12) **PATENT**

(19) NO

(11) **328884**

(13) **B1**

NORGE

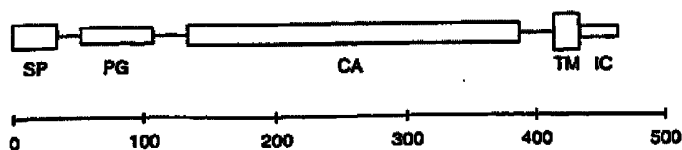
(51) Int Cl.

C07K 7/06 (2006.01)
C07K 14/47 (2006.01)
A61K 38/08 (2006.01)
C12N 9/12 (2006.01)
C12N 15/09 (2006.01)
A61K 39/395 (2006.01)
A61K 48/00 (2006.01)
A61P 35/00 (2006.01)
C07K 14/82 (2006.01)
G01N 33/566 (2006.01)
C12Q 1/02 (2006.01)
C12R 1/91 (2006.01)

Patentstyret

(21)	Søknadsnr	20011926	(86)	Int.inng.dag og søknadsnr	1999.10.22 PCT/US99/024879
(22)	Inng.dag	2001.04.19	(85)	Videreføringsdag	2001.04.19
(24)	Løpedag	1999.10.22	(30)	Prioritet	1998.10.23, US, 177776 1998.10.23, US, 178115
(41)	Alm.tilgj	2001.06.19			
(45)	Meddelt	2010.06.07			
(73)	Innehaver	Institute of Virology - Slovak Academy of Sciences, Dubravská Cesta 9, SK-84246 BRATISLAVA, Slovakia Jan Zavada, Praha, CZ-, Den tsjekkiske republikk Jaromir Pastorek, Bratislava, SK-, Slovakia Silvia Pastorekova, Bratislava, SK-, Slovakia			
(72)	Oppfinner				
(74)	Fullmektig				Bryn Aarflot AS, Postboks 449 Sentrum, 0104 OSLO, Norge
(54)	Benevnelse	Polypeptid samt anvendelse derav og peptidkompleks, og fremgangsmåte for identifisering av et organisk eller uorganisk molekyl.			
(56)	Anførte publikasjoner	WO 9534650 A2, J. SAVADA ET AL.			
(57)	Sammendrag				

Lokaliseringen av MN-protein-bindingssetet og MN-proteiner/polypeptider som konkurrerer om tilkobling til vertebratceller med immobilisert MN-protein, beskrives. Slike MN-proteiner/polypeptider forhindrer celle-celle-adhesjon og dannelsen av intercellulære kontakter. MN-protein-bindingssetet er et terapeutisk mål som kan blokkeres av organiske eller uorganiske molekyler, fortrinnsvis organiske molekyler, mer foretrukket proteiner/polypeptider, som spesifikt binder til dette setet. Terapeutiske metoder for inhibering av veksten av preneoplastiske/neoplastiske vertebratceller som unormalt uttrykker MN-protein, er omtalt. Det tilveiebringes vektorer som koder for de variable domener av MN-spesifikke antistoffer og et fleksibelt linker-polypeptid som adskiller disse domenene. Dessuten beskrives ytterligere vektorer som koder for et cytotoxisk protein/polypeptid operativt koblet til MN-genpromoter og som fortrinnsvis ytterligere koder for et cytokin. MN- genpromoter er karakterisert, og bindingssetet for en repressor av MN- transkripsjon er beskrevet.



Foreliggende oppfinnelse vedrører polypeptid samt anvendelse derav og peptidkompleks, og fremgangsmåte for identifisering av et organisk eller uorganisk molekyl.

OPPFINNELSESOMRÅDE

Foreliggende oppfinnelse angår det generelle området medisinsk genetikk og feltet biokjemisk manipulering, immunkjemi og onkologi. Nærmere bestemt vedrører den MN-genet - et cellulært gen som ansees å være et onkogen som koder for det onkoprotein som nå alternativt er kjent som MN-proteinet, MN/CA IX isoenzymet eller MN/G250-proteinet.

BAKGRUNN FOR OPPFINNELSEN

Zavada *et al.*, Internasjonal publikasjon nummer WO 93/18152 (publisert 16. september, 1993) og US-patent 5.387.676 (av 7. februar, 1996) beskriver utforskningen av den biologiske og molekylære natur av MaTu som resulterte i oppdagelsen av MN-genet og -proteinet. MN-genet ble funnet å forekomme i det kromosomale DNA i alle undersøkte vertebrater, og at dets ekspresjon var sterkt korrelert med tumorigenitet.

WO 9534650 omtaler MN-genet, fragmenter derav og cDNA for produksjon av MN-proteiner og polypeptider.

J. Zavada *et al.*, International Journal of Oncology, vol. 10, nr. 4, 1997, s. 857-863 beskriver antistoffet Mab M75 mot MN-proteinet.

MN-proteinet ble først påvist i HeLa-celler avledet fra et humant livmorhalskarsinom. Det forekommer i mange typer av humane karsinomer (særlig blant annet livmorhals, ovarie, endometriet, nyre, blære, bryst, kolorektalt, lunge, øsofagus og prostata). Meget få normale vev har vist seg å uttrykke MN-protein i vesentlig grad. De MN-uttrykkende normale vev innbefatter den humane maveslimhinne og galleblæreepitel samt enkelte andre normale vev i den alimentære trakt. Paradoksalt nok har MN-genekspresjon vist seg å gå tapt eller være redusert i karsinomer og andre preneoplastiske/neoplastiske sykdommer i enkelte vev som normalt uttrykker MN, f.eks. maveslimhinne.

Generelt kan onkogenese kjennetegnes av den unormale ekspresjon av MN-protein. For eksempel kan onkogenese forventes: (1) når MN-protein forekommer i et vev som normalt ikke uttrykker MN-protein i vesentlig grad; (2) når MN-protein mangler i et vev som normalt uttrykker det; (3) når MN-genekspresjon skjer i

betydelig øket grad eller i betydelig redusert grad i forhold til hva som normalt uttrykkes i et vev; eller (4) når MN-protein uttrykkes på et unormalt sted i en celle.

Zavada *et al.*, WO 93/18152 og Zavada *et al.*, WO 95/34650 (publisert 21. desember, 1995) omtaler hvorledes oppdagelsen av MN-genet og -proteinet og den sterke assosiasjon mellom MN-genekspresjon og tumorigenitet førte til utviklingen av metoder som både er diagnostisk/prognostiske og terapeutiske for cancer og precancerøse tilstander. Det ble i disse angitt metoder og sammensetninger for påvisning av inntreden og nærvær av neoplastisk sykdom ved påvisning, eller påvisning og kvantifisering av unormal MN-genekspresjon i vertebrater. Unormal MN-genekspresjon kan påvises, eller påvises og kvantifiseres gjennom en rekke konvensjonelle tester i vertebratprøver, for eksempel ved immunoassay under bruk av MN-spesifikke antistoffer for å påvise, eller påvise og kvantifisere MN-antigen, ved hybridiseringstester eller ved PCR-tester, så som RT-PCR, ved bruk av MN-nukleinsyrer, så som MN cDNA, for å påvise, eller påvise og kvantifisere, MN-nukleinsyrer, så som MN mRNA.

Zavada *et al.*, WO 93/18152 og WO 95/34650 beskriver produksjonen av MN-spesifikke antistoffer. Et representativt og foretrukket MN-spesifikt antistoff, det monoklonale antistoff M75 (Mab M75), ble deponert ved the American Type Culture Collection (ATCC) i Manassus, VA (USA) under ATCC nummer HB 11128. M75-antistoffet ble benyttet til oppdagelsen og identifiseringen av MN-proteinet, og kan benyttes til lett å identifisere MN-antigen i western blott, i radioimmunoassays og immunhistokjemisk, for eksempel i vevsprøver som er ferske, frosne eller formalin-, alkohol-, aceton- eller på annen måte fiksert og/eller parafin-innleiret og deparafinisert. Et annet representativt og foretrukket MN-spesifikt antistoff, Mab MN12, utskilles av hybridomet MN12.2.2, som ble deponert ved ATCC under betegnelsen HB 11647. Eksempel 1 i Zavada *et al.*, WO 95/34650, gir representative resultater fra immunhistokjemisk farving av vev ved bruk av MAb M75, hvis resultater støtter betegnelsen av MN-genet som et onkogen.

Mange undersøkelser har bekreftet den diagnostiske/prognostiske nytte av MN. Følgende artikler diskuterer anvendelsen av det MN-spesifikke MAb M75 ved diagnostisering/prognostisering av precancerøse og cancerøse cervikale lesjoner: Leff, D.N., «Half a Century of HeLa Cells: Transatlantic Antigen Enhances Reliability of Cervical Cancer Pap Test, Clinical Trials Pending,» **BioWorld® Today: The Daily**

Biotechnology Newspaper, 9(55) (24. mars, 1998); Stanbridge, E.J., «Cervical marker can help resolve ambiguous Pap smears», **Diagnostics Intelligence**, 10(5): 11 (1998); Liao & Stanbridge, «Expression of the MN Antigen in Cervical Papanicolaou Smears Is an Early Diagnostic Biomarker of Cervical Dysplasia,» **Cancer Epidemiology, Biomarkers & Prevention**, 5: 549-557 (1996); Brewer *et al.*, «A Study of Biomarkers in Cervical Carcinoma and Clinical Correlation of the Novel Biomarker MN», **Gynecologic Oncology**, 63:337-344 (1996); og Liao *et al.*, «Identification of the MN Antigen as a Diagnostic Biomarker of Cervical Intraepithelial Squamous and Glandular Neoplasia and Cervical Carcinomas,» **American Journal of Pathology**, 145(3):598-609 (1994).

Premaligne og maligne kolorektale lesjoner. MN er blitt påvist i normal gastrisk, intestinal og biliær mukosa [Pastorekova *et al.*, **Gastroenterology**, 112, 398-408 (1997)]. Immunhistokjemisk analyse av den normale tykktarm har vist moderat farving i den proksimale kolon, mens reaksjonen blir svakere distalt. Farvingen var begrenset til de basolaterale overflatene av de kryptale epitelcellene, området med høyest proliferativ evne. Siden MN er betydelig rikeligere i det prolifererende kryptale epitel enn i den øvre del av mukosa, kan den spille en rolle ved kontroll av proliferasjonen og differensieringen av intestinale epitelceller. Celleproliferasjon tiltar unormalt i premaligne og maligne lesjoner av det kolorektale epitel og ansees derfor som en indikator av kolorektal tumorprogresjon. [Risio, M., **J. Cell Biochem.** 16G:79-87 (1992); og Moss *et al.*, **Gastroenterology** 111:1425-1432 (1996)].

MN-proteinet ansees nå som det første tumor-assosierte karboanhydrase (CA) isoenzym som er blitt beskrevet. Karboanhydraser (CA) utgjør en stor familie av gener som koder for sink-metalloenzymer av stor fysiologisk betydning. Som katalysatorer av reversibel hydratisering av karbondioksyd, deltar disse enzymene i en rekke biologiske prosesser, inkludert respirasjon, forkalkning, syre/base-balanse, benresorpsjon, dannelse av øyekammervæske, cerebrospinalvæske, spytt og mavesyre (oversikt av Dodgson *et al.*, **The Carbonic Anhydrases**, Plenum Press, New York-London, s. 398 (1991)). CA er vidt utbredt i forskjellige levende organismer.

I pattedyr var det blitt påvist minst syv isoenzymer (CA I-VII) og noen få CA-relaterte proteiner (CARP/CA VIII, RPTP- β , RPTP- τ) [Hewlett-Emmett and Tashian, **Mol. Phyl. Evol.**, 5:50-77 (1996)] når analyse av den MN-deduserte aminosyre-

sekvens avdekket en slående homologi mellom den sentrale del av MN-proteinet og karboanhydraser, med det konserverte sink-bindingssettet så vel som enzymets aktive senter. Deretter ble MN-protein funnet å binde sink og å ha CA-aktivitet. På grunnlag av disse data ansees MN-proteinet nå å være det niende karboanhydrase isoenzym - MN/CA IX. [Opavsky *et al.*, **Genomics**, **33**:480-487 (mai, 1996)]. [Se også Hewlett-Emmett, *supra*, hvor CA IX foreslås som en nomenklaturbetegnelse].

CA og CA-relaterte proteiner oppviser omfattende diversitet både med hensyn til deres vevsfordeling og deres antatte eller etablerte biologiske funksjoner (Tashian, R.E., **Adv. in Genetics**, **30**:321-356 (1992)). Enkelte av disse CA uttrykkes i nesten alle vev (CA II), mens ekspresjonen av andre synes å være mer begrenset (CA VI og CA VII i spyttkjertler). I celler kan de befinne seg i cytoplasmaet (CA I, CA II, CA III og CA VII), i mitokondrier (CA V), i sekretoriske granuler (CA VI) eller de kan assosiere med membran (CA IV). Ibland er det registrert nukleær lokalisering av enkelte isoenzymer [Parkkila *et al.*, **Gut**, **35**:646-650 (1994); Parkkila *et al.*, **Histochem. J.**, **27**:133-138 (1995); Mori *et al.*, **Gastroenterol.**, **105**:820-826 (1993)].

CA og CA-relaterte proteiner adskiller seg også med hensyn til kinetiske egenskaper og følsomhet overfor inhibitorer [Sly & Hu, **Annu. Rev. Biochem.**, **64**:375-401 (1995)]. I alimentærtrakten er karboanhydraseaktivitet involvert i mange viktige funksjoner, så som spyttsekresjon, produksjon av mavesyre, pankreassaft og galle, intestinal vann- og ionetransport, fettsyreopptak og biogenese i leveren. Minst syv CA-isoenzymer er blitt påvist i forskjellige områder av den alimentære trakt. Biokjemiske, histokjemiske og immuncytokjemiske studier har imidlertid avdekket en betydelig heterogenitet med hensyn til deres nivåer og fordeling [Swensen, E.R., «Distribution and Functions of carbonic anhydrase in the gastrointestinal tract», i: **The Carbonic Anhydrases. Cellular Physiology and Molecular Genetics**, (Dodgson *et al.*, red.) Plenum Press, New York, side 265-287 (1991); og Parkkila & Parkkila, **Scan J. Gastroenterol.**, **31**:305-317 (1996)]. Mens CA II finnes langs hele den alimentære kanal, er CA IV koblet til den nedre gastrointestinaltrakt, CA I, III og V forekommer bare i noen få vev, og ekspresjonen av CA VI og VII er begrenset til spyttkjertlene [Parkkila *et al.*, **Gut**, **35**:646-650 (1994); Fleming *et al.*, **J. Clin. Invest.**, **96**:2907-2913 (1995); Parkkila *et al.*, **Hepatology**, **24**:104 (1996)].

MN/CA IX har en rekke egenskaper som adskiller det fra andre kjente CA-isoenzymer og viser dets relevans for onkogenese. Disse egenskapene innbefatter dets tetthetsavhengige ekspresjon i cellekultur (f.eks. HeLa-celler), dets korrelasjon med den tumorigene fenotype av somatiske cellehybrider mellom HeLa og normale humane fibroblaster, dets nære assosiasjon med en rekke humane karsinomer og dets fravær i tilsvarende normale vev [f.eks. Zavada *et al.*, **Int. J. Cancer**, **54**:268-274 (1993); Pastorekova *et al.*, **Virology**, **187**:620-626 (1992); Liao *et al.*, **Am. J. Pathol.**, **145**:598-609 (1994); Pastorek *et al.*, **Oncogene**, **9**:2788-2888 (1994); Cote, **Women's Health Weekly: News Section**, s. 7 (30. mars, 1998); Liao *et al.*, **Cancer Res.**, **57**:2827 (1997), Vermeylen *et al.*, «Expression of the MN antigen as a biomarker of lung carcinoma and associated precancerous conditions,» **Proceedings AACR**, **39**:334 (1998); McKiernan *et al.*, **Cancer Res.**, **57**:2362 (1997); og Turner *et al.*, **Hum. Pathol.**, **28**(6):740 (1997)]. I tillegg har *in vitro* transformasjonsevnen til MN/CA IX cDNA vært påvist i NIH 3T3 fibroblaster [Pastorek *et al.*, *id*].

MN-proteinet har også vært påvist med G260-antigenet. Uemura *et al.*, «Expression of Tumor-Associated Antigen MN/G250 in Urologic Carcinoma: Potential Therapeutic Target,» **J. Urol.** **157** (4. Suppl.):377 (Abstract 1475; 1997) angir: «Sequence analysis and database searching revealed that G250 antigen is identical to MN, a human tumor-associated antigen identified in cervical carcinoma (Pastorek *et al.*, 1994).».

SAMMENFATNING AV OPPFINNELSEN

Foreliggende oppfinnelse vedrører følgende polypeptid, kjennetegnet ved at det er valgt fra SEKV ID NR: 137 og 138.

Videre er nevnte polypeptid for bruk innen medisin.

Det som her påvises er lokaliseringen av MN-proteinets bindingssete. Av særlig betydning er den region innen det protoglykan-lignende domene, aa 61-96 (SEKV. ID. NR:97) som inneholder en 6-gangers tandemrepetisjon av seks aminosyrer, innenfor hvilke epitopen for M75 MAb befinner seg i minst to kopier, og innenfor hvilke MN-bindingssetet ansees å være lokalisert. Et alternativt MN-bindingssete kan være lokalisert i CA-domenet.

Dessuten er også MN-proteiner og MN-polypeptider som konkurrerer om celledilkobling med immobilisert MN-protein, påvist. Slike MN-protein/polypeptider forhindrer celle-celle adhesjon og dannelsen av intercellulære kontakter.

Det beskrives her celleadhesjons-testmetoder som benyttes for å identifisere bindingssteder på det MN-protein som vertebratceller, fortrinnsvis pattedyrceller, mer foretrukket humanceller, binder til. Et slikt MN-bindingssete identifiseres deretter som et terapeutisk mål som kan blokkeres med MN-spesifikke antistoffer eller uorganiske eller organiske molekyler, fortrinnsvis organiske molekyler, mer foretrukket proteiner/polypeptider som spesifikt binder til nevnte sete.

Aspekter ved foreliggende oppfinnelse er mer detaljert beskrevet i det følgende. Den terapeutiske anvendelse av organiske eller uorganiske molekyler, fortrinnsvis organiske molekyler, beskrives. Fortrinnsvis binder slike molekyler spesifikt til et sete på MN-proteinet som vertebratceller adhererer til i en celle-adhesjonstest, hvor nevnte molekyl når det testes *in vitro*, inhiberer adhesjonen av celler til MN-protein. Videre foretrekkes slike molekyler som når de er i kontakt med en vertebrat preneoplastisk eller neoplastisk celle som unormalt uttrykker MN-protein, inhiberer veksten av nevnte celle. Nevnte vertebratceller er fortrinnsvis pattedyr- og mer foretrukket humane celler.

Fortrinnsvis er et slikt molekyl organisk, og mer foretrukket er et slikt organisk molekyl et protein eller et polypeptid. Enda mer foretrukket omfatter nevnte protein eller polypeptid en aminosyresekvens valgt fra gruppen bestående av SEKV. ID. NR: 137 og 138.

Det setet på MN-proteiner som vertebratceller adhererer til ved nevnte celle-adhesjonstest er fortrinnsvis innenfor det proteoglykan-lignende domene [SEKV. ID. NR:50] eller innenfor karboanhydrasedomenet [SEKV. ID. NR:51] av MN-proteinet. Fortrinnsvis omfatter vedkommende sete en aminosyresekvens som er valgt fra gruppen bestående av SEKV. ID. NR:10 og 97-106. Enda mer foretrukket har setet en aminosyresekvens som er valgt fra gruppen bestående av SEKV. ID. NR:10 og 97-106.

MN-proteiner og MN-polypeptider medierer tilkobling av vertebratceller i en celleadhesjonsbestemmelse, hvor nevnte MN-protein eller MN-polypeptid når det innføres i det ekstracellulære væskemiljø i vertebratceller, forhindrer dannelsen av intercellulære kontakter og adhesjonen av nevnte vertebratceller til hverandre. Slike MN-proteiner og MN-polypeptider kan være anvendelige til inhibering av veksten av

vertebrate preneoplastiske eller neoplastiske celler som unormalt uttrykker MN-protein, når slike MN-proteiner eller MN-polypeptider innføres i det ekstracellulære væskemiljø i slike vertebratceller. Slike vertebratceller er fortrinnsvis pattedyr- og mer foretrukket humane celler.

5 Nevnte MN-proteiner eller MN-polypeptider som medierer tilkobling av vertebratceller i en celleadhesjonstest har fortrinnsvis aminosyresekvenser valgt fra gruppen bestående av SEKV ID NR:10 og 97-106.

Representative MN-proteiner og MN-polypeptider som medierer tilkobling av vertebratceller i en celleadhesjonstest, bindes spesifikt enten av det M75
10 monoklonale antistoff som utskilles fra hybridomet VU-M75 som ble deponert ved American Type Culture Collection under ATCC nummer HB 11128, eller av det MN12 monoklonale antistoff som utskilles fra hybridomet MN12.2.2, som ble deponert ved American Type Culture Collection under ATCC nummer HB 11647, eller av begge nevnte monoklonale antistoffer.

15 Et annet aspekt ved foreliggende oppfinnelse er en fremgangsmåte for identifisering av et organisk eller et uorganisk molekyl som spesifikt bindes til et sete på et MN protein, hvortil vertebrat celler adhererer i en celleadhesjonsanalyse, idet nevnte sete består av en aminosyresekvens valgt fra SEKV ID NR: 10 og 97-106, kjennetegnet ved at den omfatter testing av organiske og uorganiske molekyler i
20 nevnte celleadhesjonsanalyse og identifisering av molekyler som inhiberer adhesjonen av vertebratcellene til nevnte MN protein som spesifikt bindende til nevnte sete, og nevnte molekyl er nyttig for behandling av preneoplastisk og/eller neoplastisk sykdom, som er kjennetegnet ved unormal ekspresjon av MN protein, idet nevnte celleadhesjonsanalyse omfatter:

25 (a) muliggjøre at MN proteinet bindes til et substrat hvortil nevnte vertebratceller ikke binder;

(b) skylle ubundet MN protein fra nevnte substrat;

30 (c) inkubering av nevnte bundete MN protein med nevnte molekyl og med nevnte vertebratceller;

(d) skylling av ubundede vertebratceller fra nevnte MN protein; og

(e) bestemme om nevnte molekyl inhiberer adhesjonen av nevnte vertebratceller til nevnte MN protein.

Enda et annet aspekt ved foreliggende oppfinnelse er anvendelse av polypeptid ifølge krav 1, for fremstilling av et medikament for behandling av preneoplastisk og/eller neoplastisk sykdom som er kjennetegnet ved unormal ekspresjon av MN protein, ved inhibering av veksten av vertebrate preneoplastiske og/eller neoplastiske celler som uttrykker MN protein unormalt.

I henhold til et ytterligere aspekt angår foreliggende oppfinnelse et peptidkompleks, kjennetegnet ved at det omfatter et polypeptid ifølge krav 1, kovalentlig bundet til polylysin hvortil det er bundet en nukleinsyre som koder for et cytotoxisk protein eller cytotoxisk polypeptid operativt koblet til MN-genpromoter, hvor nevnte peptidkompleks når administrert inn i en vertebrat preneoplastisk eller neoplastisk celle som unormalt uttrykker MN-protein, inhiberer veksten av nevnte celle.

I henhold til en annen foretrukket utførelsesform er nevnte cytotoxiske protein HSV tymidinkinase. Fortrinnsvis omfatter nevnte nukleinsyre ytterligere en nukleinsyre som koder for et cytokin operativt koblet til nevnte MN-genpromoter.

MN-genpromoter karakteriseres her. Identifiseringen av bindingssetet for en repressor av MN-transkripsjon er beskrevet. Mutasjonsanalyser tydet på at den direkte repetisjon AGGGCacAGGGC [SEKV. ID. NR:143] fordres for effektiv repressorbinding.

Identifisering av det protein som binder til repressoren og modifikasjon av dets bindingsegenskaper er en annen måte for å modulere MN-ekspresjon som fører til cancerterapi. Undertrykkelse av MN-ekspresjon i tumorceller ved overekspresjon av en negativ regulator forventes å føre til nedsatt tumorvekst. Et repressorkompleks som omfatter minst to subenheter, ble funnet å binde til SEKV. ID. NR:115 av MN-genpromoter. Et repressorkompleks, funnet å være i direkte kontakt med SEKV. ID. NR:115, ved UV-krysskobling, omfattet to proteiner som har molekylvekter på henholdsvis 35 og 42 kilodalton.

Forkortelser

Følgende forkortelser benyttes her:

aa - aminosyre

	ATCC	-	American Type Culture Collection
	bp	-	basepar
	BLV	-	bovint leukemivirus
	BSA	-	bovint serumalbumin
5	BRL	-	Bethesda Research Laboratories
	CA	-	karboanhydrase
	CAM	-	celleadhesjonsmolekyl
	CARP	-	karboanhydrase-relatert protein
	CAT	-	kloramfenikol-acetyltransferase
10	Ci	-	Curie
	cm	-	centimeter
	CMV	-	cytomegalovirus
	cpm	-	telling per minutt
	C-terminal	-	karboksyl-terminal
15	CTL	-	cytotoksiske T-lymfocytter
	°C	-	grader Celsius
	DEAE	-	dietylaminoetyl
	DMEM	-	Dulbeccos modifiserte Eagle medium
	ds	-	dobbeltkjedet
20	EDTA	-	etylendiamintetraacetat
	EGF	-	epidermal vekstfaktor
	EIA	-	enzym-immunoassay
	ELISA	-	enzym-linked immunosorbent assay
	EMSA	-	electrophoretic mobility shift assay
25	F	-	fibroblaster
	FACS	-	cytofluorometrisk undersøkelse
	FCS	-	føtalt kalveserum
	FITC	-	fluorescein-isotiocyanat
	FTP	-	DNase 1 footprinting analysis
30	GST-MN	-	fusjonsprotein MN glutation-S-transferase
	GVC	-	ganciklovir
	H	-	HeLa-celler
	H-E	-	hematoksylin-eosin
	HEF	-	humane embryo fibroblast

	HeLa K	-	standardtype HeLa-celler
	HeLa S	-	Stanbridge's mutante HeLa D98/AH.2
	H/F-T	-	hybrid HeLa fibroblastceller som er tumorigene; avledet fra HeLa D98/AH.2
5	H/F-N	-	hybrid HeLa fibroblastceller som ikke er tumorigene; avledet fra HeLa D98/AH.2
	HPV	-	humant papillomavirus
	HRP	-	pepperrot-peroksydase
	HSV	-	Herpes simplex-virus
10	IC	-	intracellulær
	IFN	-	interferon
	IL-2	-	interleukin-2
	Inr	-	initiator
	IPTG	-	isopropyl-beta-D-tiogalakto-pyranosid
15	kb	-	kilobase
	kbp	-	kilobasepar
	kd eller kDa	-	kilodalton
	KS	-	keratansulfat
	LCMV	-	lymfocytisk choriomeningitt-virus
20	LTR	-	long terminal repeat
	M	-	molar
	mA	-	milliamper
	MAb	-	monoklonalt antistoff
	MCSF	-	makrofag kolonistimulerende faktor
25	ME	-	merkaptoetanol
	MEM	-	minimalt essensielt medium
	min.	-	minutt(er)
	mg	-	milligram
	mL	-	milliliter
30	mM	-	millimolar
	MMC	-	mitomycin C
	mmol	-	millimol
	MLV	-	murint leukemivirus
	N	-	normal konsentrasjon

	NEG	-	negativ
	ng	-	nanogram
	nm	-	nanometer
	nt	-	nukleotid
5	N-terminus	-	aminoterminal
	ODN	-	oligodeoksynukleotid
	ORF	-	åpen leseramme
	PA	-	Protein A
	PBS	-	fosfat-buffret saltvann
10	PCR	-	polymerasekjedereaksjon
	PEST	-	kombinasjon av énbokstavs forkortelser for prolin, glutaminsyre, serin, treonin,
	PG	-	proteoglykan
	pI	-	isoelektrisk punkt
15	PMA	-	forbol 12-myristat 13-acetat
	POS	-	positiv
	Py	-	pyrimidin
	RACE	-	hurtig amplifikasjon av cDNA-ender
	RCC	-	nyrecellekarsinom
20	RIA	-	radioimmunoassay
	RIP	-	radioimmunopresipitasjon
	RIPA	-	radioimmunopresipitasjonstest
	RNP	-	RNase-beskyttelsestest
	RT-PCT	-	revers transkripsjons polymerasekjedereaksjon
25	SAC	-	<i>Staphylococcus aureus</i> -celler
	<i>S. aureus</i>	-	<i>Staphylococcus aureus</i>
	sc	-	subkutan
	SDRE	-	serum doseresponselement
	SDS	-	natriumdodekylsulfat
30	SDS-PAGE	-	natriumdodekylsulfat-polyakrylamid gelelektroforese
	SINE	-	short interspersed repeated sequence
	SP	-	signalpeptid
	SP-RIA	-	fastfase radioimmunoassay
	SSDS	-	syntetisk spleise-donorsete

	SSH	-	subtractive suppressive PCR
	SSPE	-	NaCl (0,18M), natriumfosfat (0,01M), EDTA (0,001M)
	TBE	-	Tris-borat/EDTA elektroforesebuffer
	TC	-	vevskultur
5	TCA	-	triklorediksyre
	TC-medium	-	vevskulturmedium
	tk	-	tymidinkinase
	TM	-	transmembran
	TMB	-	tetrametylbenzidin
10	Tris	-	tris(hydroksymetyl)aminometan
	μCi	-	mikrocurie
	μg	-	mikrogram
	μL	-	mikroliter
	μM	-	mikromolar
15	VSV	-	vesikulært stomatitt-virus
	VV	-	vaccinia-virus
	X-MLV	-	xenotropisk murint leukemi-virus

Cellelinjer

20	AGS	-	cellelinje avledet fra et primært adenogastrisk karsinom [Barranco & Townsend, Cancer Res. , 43 :1703 (1983) og Invest. New Drugs , 1 :117 (1983)]; tilgjengelig fra ATCC under CRL-1739;
	BL-3	-	bovine B-lymfocytter [ATCC CRL-8037; leukemicellesuspensjon; J. Natl. Cancer Inst. (Bethesda) 40 :737 (1968)];
25	C33	-	en cellelinje avledet fra en human cervikal karsinombiopsi [Auersperg, N., J. Natl. Cancer Inst. (Bethesda), 32 :135-148 (1964)]; tilgjengelig fra ATCC under HTB-31;
	C33A	-	humane cervikale karsinomceller [ATCC HTB-31; J. Natl. Cancer Inst. (Bethesda) 32 :135 (1964)];
30	COS	-	simian-viruscellelinje [Gluzman, Y., <i>Cell</i> , 23 :175 (1981)];
	HeLa K	-	standardtype HeLa-celler; aneuploid, epitel-lignende cellelinje isolert fra et humant cervikalt adenokarsinom [Gey <i>et al.</i> ,

Cancer Res., 12:264 (1952)]; Jones *et al.*, **Obstet. Gynecol.** 38: 945-949 (1971)] mottatt fra Professor B. Korych, [Institute of Medical Microbiology and Immunology, Charles University; Prag, Tsjekkia];

- 5 HeLa - mutant HeLa klon som mangler hypoksantin-guanin-fosforibosyl-
D98/AH.2 transferase (HGPRT⁻) velvillig stillet til rådighet av
(også HeLa s) Eric J. Stanbridge [Department of Microbiology,
College of Medicine, University of California, Irvine, CA (USA)]
og rapportert i Stanbridge *et al.*, **Science**, 215:252-259
10 (15. januar, 1982); opphav til hybridceller H/F-N og H/F-T,
også mottatt fra E.J. Stanbridge;
- KATO III - cellelinje fremstillet fra en metastatisk form av et gastrisk
karsinom [Sekiguichi *et al.*, **Japan J. Exp. Med.**, 48:61 (1978)];
tilgjengelig fra ATCC under HTB-103;
- 15 NIH-3T3 - murin fibroblast-cellelinje rapportert i Aaronson, **Science**, 237:
178 (1987);
- QT35 - fibrosarkomceller fra vaktel [ECACC:93120832; **Cell**, 11:95
(1977)];
- Raj - human Burkitts lymfocellelinje [ATCC CCL-86;
20 **Lancet**, 1:238 (1964)];
- rat2TK⁻ - cellelinje (rotteembryo, tymidinkinasemutant) ble avledet fra en
subklon av en 5'-brom-deoksyuridin resistent stamme av den
Fischer rottefibroblast 3T3-lignende cellelinje Rat1; cellene
mangler betydelig grad av nukleær tymidinkinase [Ahrens, B.,
25 **Virology**, 113: 408 (1981)];
- SiHa - human cervikal skvamøs karsinomcellelinje [ATCC HTB-35;
Friedl *et al.*, **Proc. Soc. Exp. Biol. Med.**, 135:543 (1990)];
- XC - celler avledet fra et rotte-rhabdomyosarkom indusert med Rous
sarkom-virus-indusert rotte-sarkom [Svoboda, J., **Natl. Cancer**
30 **Center Institute Monograph No. 17**, i: «International
Conference on Avian Tumor Viruses» (J.W. Beard red.), s. 277-
298 (1964)], velvillig stillet til rådighet av Jan Svoboda
[Institute of Molecular Genetics, Czechoslovak Academy of

Sciences; Prag, Tsjekkia]; og

CGL1 - H/F-N hybridceller (HeLa D98/AH.2 derivat);

CGL2 - H/F-N hybridceller (HeLa D98/AH.2 derivat);

CGL3 - H/F-T hybridceller (HeLa D98/AH.2 derivat);

5 CGL4 - H/F-T hybridceller (HeLa D98/AH.2 derivat).

Nukleotid- og aminosyre-søkvenssymboler

Følgende symboler benyttes her for nukleotider:

10	Basesymbol	Betydning
	A	adenin
	C	cytosin
	G	guanin
	T	tymin
15	U	uracil
	I	inosin
	M	A eller C
	R	A eller G
	W	A eller T/U
20	S	C eller G
	Y	C eller T/U
	K	G eller T/U
	V	A eller C eller G
	H	A eller C eller T/U
25	D	A eller G eller T/U
	B	C eller G eller T/U
	N/X	A eller C eller G eller T/U

Det finnes 20 hoved-aminosyrer som hver spesifiseres av et ulikt arrangement
 30 av 3 nabo-nukleotider (triplett-kode eller kodon) og som kobles sammen i en bestemt
 rekkefølge for å danne et karakteristisk protein. Det benyttes her en trebokstavs eller
 énbokstavs konvensjon for å identifisere nevnte aminosyrer, for eksempel i Figur 1,
 som følger:

	Aminosyrenavn	3 bokst. forkort.	1 bokst. forkort.
	Alanin	Ala	A
	Arginin	Arg	R
	Asparagin	Asn	N
5	Asparaginsyre	Asp	D
	Cystein	Cys	C
	Glutaminsyre	Glu	E
	Glutamin	Gln	Q
	Glycin	Gly	G
10	Histidin	His	H
	Isoleucin	Ile	I
	Leucin	Leu	L
	Lysin	Lys	K
	Metionin	Met	M
15	Fenylalanin	Phe	F
	Prolin	Pro	P
	Serin	Ser	S
	Treonin	Thr	T
	Tryptofan	Trp	W
20	Tyrosin	Tyr	Y
	Valin	Val	V
	Ukjent eller annet		

KORT BESKRIVELSE AV FIGURENE

25 Figur 1A-C angir nukleotidsekvensen for en MN cDNA-klon [SEKV. ID. NR:1] isolert som her beskrevet. Figur 1A-C angir også den predikerte aminosyresekvens [SEKV. ID. NR:2] som kodes av cDNA.

 Figur 2A-F angir en 10.898 bp komplett genomisk sekvens av MN [SEKV. ID. NR:5]. Baseantallet er som følger: 2654 A, 2739 C; 2645 G; og 2859 T. De 11
30 exonene er generelt angitt i kapiteler, men exon 1 ansees ifølge RNase-
beskyttelsestest å begynne i posisjon 3507.

 Figur 3 er et restriksjonskart over den fulle lengde av MN cDNA. Den åpne leseramme er vist som en åpen boks. De brede linjene under restriksjonskartet illustrerer størrelse og posisjon av to overlappende cDNA-kloner. De horisontale

pilene angir posisjonene av primerne R1 [SEKV. ID. NR:7] og R2 [SEKV. ID. NR:8] benyttet for 5'-ende RACE. Relevante restriksjonsseter er *Bam*HI (B), *Eco*RV (V), *Eco*RI (E), *Pst*II (Ps), *Pvu*II (Pv).

Figur 4 fremstiller skjematisk den 5' MN genomiske region av en MN genomisk klon, hvor nummereringen tilsvarer transkripsjonsinitieringssetene bestemt ved RACE.

Figur 5 er et exon-intron kart over det humane MN/CA IX-gen. Posisjoner og størrelser av exonene (nummererte, kryss-skraverte bokser) Alu-repetisjons-elementer (åpne bokser) og en LTR-relatert sekvens (første unummererte prikkede boks) er justert til den angitte skala. De exonene som tilsvarer individuelle MN/CA IX proteindomenene er innesluttet mellom stiplede rammer betegnet PG (proteoglykan-lignende domene), CA (karboanhydrase-domene), TM (transmembranforankring) og IC (intracytoplasmatisk hale). Under kartet illustrerer sammenligningsoppstillingen av aminosyresekvenser graden av homologi mellom MN/CA IX-protein PG-regionen (aa 53-111) [SEKV. ID. NR:50] og det humane aggrecan (aa 781-839) [SEKV. ID. NR:54].

Figur 6 er en nukleotidsekvens for den foreslåtte promoter av det humane MN-gen [SEKV. ID. NR:27]. Nukleotidene er nummerert fra transkripsjons-initieringssetet i henhold til RNase-beskyttelsestest. Potensielle regulelementer er overstreket. Transkripsjonsstartseter er angitt med stjerner (RNase-beskyttelse) og prikker (RACE) over de tilsvarende nukleotider. Sekvensen for det første exon begynner under stjernene. FTP-analyse av MN4-promoterfragmentet avdekket 5 regioner (I-V) beskyttet både i den kodende og den ikke-kodende kjede, og to regioner (VI og VII) beskyttet i den kodende kjede, men ikke i den ikke-kodende kjede.

Figur 7 angir et skjema av sammenligningsoppstillingen av MN genomiske kloner i henhold til deres posisjon i forhold til transkripsjonsinitieringssetet. Alle de genomiske fragmentene bortsett fra Bd3 ble isolert fra et lambda FIX II genomisk bibliotek avledet fra HeLa-celler. Klon Bd3 ble avledet fra et humant føtalhjernebibliotek.

Figur 8 fremstiller skjematisk MN-proteinstrukturen. Forkortelsene er de samme som i Figur 5. Skalaen indikerer antallet aminosyrer.

DETALJERT BESKRIVELSE

Betegnelsene «MN/CA» og «MN/CA9» ansees her som synonymer for MN. G250 antigenet ansees dessuten å referere til MN-protein/polypeptid [Uemura *et al.*, *J. Urol.*, **154** (4 Suppl.):377 (Abstract 1475; 1997)].

5 MN/CA IX ble først påvist i HeLa-celler avledet fra humant livmorhalskarsinom både som et plasmamembran- og nukleært protein med en tilsynelatende molekylvekt på 58 og 54 kilodalton (kDa) ifølge western blotting. Det er N-glykosylert med en enkelt 3 kDa karbohydratkjede og danner under ikke-reduserende betingelser, S-S-koblede oligomerer (Pastorekova *et al.*, *Virology*, **187**:620-626 (1992); Pastorek *et al.*, *Oncogene*, **9**:2788-2888 (1994)]. MN/CA IX er et transmembranprotein lokalisert 10 ved celleoverflaten, men det har i enkelte tilfeller vært påvist i kjernen (Zavada *et al.*, *Int. J. Cancer*, **54**:268-274 (1993); Pastorekova *et al.*, *supra*].

MN manifesteres i HeLa-celler med et tvillingprotein, p54/58N. Immunblotting ved bruk av et monoklonalt antistoff som er reaktivt med p54/58N (MAb M75) viste to 15 bånd ved 54 kd og 58 kd. Disse to båndene kan tilsvare én type protein som mest sannsynlig adskiller seg ved post-translasjonell prosessering. Her angir vendingen «tvillingprotein» p54/58N.

Zavada *et al.*, WO 93/18152 og/eller WO 95/34650, omtaler MN cDNA-sekvensen (SEKV. ID. NR:1) som her er vist i Figur 1A-1C, MN-aminosyresekvensen 20 (SEKV. ID. NR:2) også vist i Figur 1A-1C, og MN-genomsekvensen (SEKV. ID. NR:5) her vist i Figur 2A-2F. MN-genet er organisert i 11 exoner og 10 introner.

De første 37 aminosyrene av MN-proteinet vist i Figur 1A-1C, er det antatte MN-signalpeptid [SEKV. ID. NR:6]. MN-proteinet har et ekstracellulært domene [aminosyrene (aa) 38-414 i Figur 1A-1C (SEKV. ID. NR:87)], et transmembran- 25 domene [aa 415-434 (SEKV. ID. NR:52)] og et intracellulært domene [aa 435-459 (SEKV. ID. NR:53)]. Det ekstracellulære domenet inneholder det proteoglykanlignende domene [aa 53-111 (SEKV. ID. NR:50)] og karboanhydrase- (CA) domenet [aa 135-391 (SEKV. ID. NR:51)].

30 Anticancermedikamenter og antistoffer som blokkerer interaksjon mellom MN-protein og reseptormolekyler

MN-protein ansees å være et unikt egnet mål for cancerterapi av en rekke grunner, herunder følgende:

(1) det er lokalisert på celleoverflaten, noe som gjør det tilgjengelig.

(2) Det uttrykkes i en høy prosentandel av humane karsinomer (f.eks. livmorhals-, nyre-, kolon-, bryst-, spiserør-, lunge-, hode- og hals-karsinomer), men uttrykkes vanligvis ikke i vesentlig grad i det normale vev som slike karsinomer skriver seg fra.

(3) Det uttrykkes vanligvis bare i mavesekkens mukosa og i enkelte epitel i fordøyelsestrakten (galleblære- og tynntarm epitel). Det eksisterer derved en anatomisk barriere mellom de MN-uttrykkende preneoplastiske/neoplastiske og MN-uttrykkende normale vev. Det kan således administreres medikamenter, inklusiv antistoffer, som kan nå frem til tumorer uten å interferere med MN-uttrykkende normale vev.

(4) MAb M75 har høy affinitet og spesifisitet for MN-protein.

(5) Det er blitt isolert MN cDNA og MN genomiske kloner som omfatter de proteinkodende og genregulatoriske sekvenser.

(6) MN-spesifikke antistoffer har vært vist å ha blant de høyeste tumoropptak rapportert i kliniske studier med anti-tumor antistoffer i solide tumorer, som vist for det MN-spesifikke kimære antistoff G250 i forsøksdyr og i fase I kliniske forsøk med nyrekarsinompasienter [Steffens *et al.*, *J. Clin. Oncol.*, **15**:1529 (1997)]. Dessuten har MN-spesifikke antistoffer lavt opptak i normale vev.

Data, f.eks. som her angitt, stemmer overens med følgende teori angående hvorledes MN-protein virker i normale vev i preneoplastiske/neoplastiske vev. I normale vev (f.eks. i mavesekk-mukosa), ansees MN-protein å være en differensieringsfaktor. Det binder med dets normale reseptor S (for mavesekk). Mavesekk-karsinomer har vært vist å ikke inneholde MN-protein.

Ektopisk ekspresjon av MN-protein i andre vev forårsaker malign omdannelse av celler. Slik ektopisk ekspresjon ansees å forårsakes av bindingen av MN-protein til en alternativ reseptor H (for HeLa-celler), koblet til en signaloverføringsvei som fører til malignitet. Medikamenter eller antistoffer som blokkerer bindingsstedet av MN-protein for reseptor H vil kunne forventes å forårsake reversjon av preneoplastiske/-neoplastiske celler til det normale eller indusere deres død.

Design og utvikling av MN-blokkerende medikamenter eller antistoffer

En fremgangsmåte for design og utvikling av MN-blokkerende medikamenter, f.eks. peptider med høy affinitet for MN-protein, eller antistoffer, har flere trinn. Det

første er å teste med henblikk på bindingen av MN-protein til reseptorer, basert på celleadhesjonsbestemmelser beskrevet nedenfor. Den samme fremgangsmåte vil også kunne benyttes til å bestemme medikamenter som blokkerer MN-protein-bindingsssetet. I betraktning av de alternative reseptorene S og H, vil mavesekk-epitelceller eller revertanter (inneholdende fortrinnsvis S-reseptorer), HeLa-celler (som inneholder H-reseptoren og mangler S-reseptoren) bli benyttet i celleadhesjonsbestemmelsen.

For å identifisere reseptorbindingssetet til MN-protein, kan delesjonsvarianter av MN-protein som mangler forskjellige domener for å identifisere regioner som er ansvarlige for interaksjonen av MN-protein med en reseptor, benyttes. Eksempel 2 identifiserer og illustrerer hvorledes andre bindingsseter på MN-protein kan påvises. Et foretrukket MN-bindingssete ansees å være nært beslektet eller identisk med epitopen for MAb M75, som er lokalisert i minst to kopier innenfor 6-gangers tandemrepetisjonen av 6 aminosyrer [aa 61-96 (SEKV. ID. NR:97)] i det proteoglykan-lignende domenet av MN-proteinet. Mindre delesjonsvarianter kan fremstilles innenfor dette relevante domenet, f.eks. kan det fremstilles fusjonsproteiner med kun små segmenter av MN-protein. Det kan dessuten foretas kontrollert oppslutning av MN-protein med spesifikke proteaser med påfølgende separasjon av produktene.

Videre kan det syntetiseres peptider som omfatter det forventede bindingssetet. Alle disse produktene kan testes ved celleadhesjonsbestemmelser, som eksemplifisert nedenfor [se f.eks. Pierschbacher and Ruoslahti, **PNAS**, **81**:5985 (1984); Ruoslahti and Pierschbacher, **Science**, **238**:491].

Molekyler kan konstrueres for å blokkere MN-reseptorbindingssetet. For eksempel kan bruk av et fagdisplay-peptidbibliotek-sett [som Ph.D®-7 Peptide 7-Mer Library Kit fra New England Biolabs; Beverly, MA (USA)] som eksemplifisert i Eksempel 2 og 3, benyttes til å finne peptider med høy affinitet overfor målmolekyler. Biologisk aktivitet av de identifiserte peptidene vil kunne testes *in vitro* ved inhibering av celleadhesjon til MN-protein, ved effekter på cellemorfologi og vekstkarakteristika av MN-relaterte tumorceller (HeLa) og av kontrollceller [Symington, **J. Biol. Chem.**, **267**:25744 (1992)]. *In vivo* screening vil kunne foretas på nakne mus som har fått injisert HeLa-celler.

Peptider som inneholder bindingssetet for MN-proteinet vil bli fremstillet [f.eks. MAP (multiple antigenpeptider); Tam., J.P. **PNAS** (USA) **85**:5409 (1988); Butz *et al.*, **Peptide Res.**, **7**:20 (1994)]. Disse MAP vil bli benyttet til å immunisere dyr for å oppnå antistoffer (polyklonale og/eller monoklonale) som gjenkjenner og blokkerer bindingssetet. [Se f.eks. Brooks *et al.*, **Cell**, **79**:1157 (1994)]. «Vaksinasjon» vil deretter kunne benyttes for å undersøke om dyr er beskyttet. Antistoffer mot MN-bindingsssetet vil eventuelt kunne benyttes til å blokkere MN-proteiners interaksjon med andre molekyler.

Datamaskinmodellering kan også benyttes for å utvikle molekyler med spesifikk affinitet for MN-protein som vil mediere sterisk inhibering mellom MN-protein og dets reseptor. En datamaskinmodell av MN-bindingsssetet for reseptoren vil inneholde romlige, elektrostatiske, hydrofobe og andre karakteristika for denne struktur. Organiske molekyler som er komplementære til strukturen, som best passer inn i bindingssetet, vil kunne utvikles. Også uorganiske molekyler som kan blokkere MN-bindingsssetet kan testes tilsvarende.

Anvendelsen av onkoproteiner som mål for utvikling av nye cancerterapeutika ansees av fagmannen som vanlig. [Se f.eks. Mendelsohn & Lippman, «Growth Factors», s. 114-133, i: DeVita *et al.* (red). **Cancer: Principles and Practice of Oncology** (4. utg.; Lippincott; Philadelphia, (1993)]. I videste forstand kan utviklingen av blokkerende medikamenter baseres på kompetitive inhiberingsforsøk. Slike forsøk har vært benyttet for å finne frem til nye medikamenter siden oppdagelsen av sulfonamider (kompetitive inhibitorer av para-aminobenzosyre, en forløper for folsyre). Enkelte cytostatika er også kompetitive inhibitorer (blant annet f.eks. halogenerte pyrimidiner).

Anvendelsen av slike tilnærminger til MN er imidlertid ny. Sammenlignet med andre tumor-relaterte molekyler (f.eks. vekstfaktorer og deres reseptorer) har MN den unike egenskap at det uttrykkes ulikt i preneoplastiske/neoplastiske vev og normale vev som er adskilt av en anatomisk barriere.

MN-gen - Kloning og sekvensering

Figur 1A-C angir nukleotidsekvensen for en full-lengdes MN cDNA-klon isolert som beskrevet nedenfor [SEKV. ID. NR:1]. Figur 2A-F angir en fullstendig MN

genomisk sekvens [SEKV. ID. NR:5]. Figur 6 viser nukleotidsekvensen for en fore-
slått MN-promoter [SEKV. ID. NR:27].

Det er på grunn av den genetiske kodes degenererthet, det vil si at mer enn
ett kodon koder for én aminosyre [for eksempel koder hver av kodonene TTA, TTG,
5 CTT, CTC, CTA og CTG for aminosyren leucin (leu)], at variasjoner av nukleotid-
sekvensen i for eksempel SEKV. ID. NR:1 og 5, hvor ett kodon er benyttet i stedet
for et annet, vil produsere tilnærmet ekvivalent protein eller polypeptid i henhold til
denne oppfinnelse. Alle slike variasjoner i nukleotidsekvensene av MN cDNA og
komplementære nukleinsyresekvenser faller inn under rammen for denne opp-
10 finnelse.

Det er videre klart at de nukleotidsekvenser som her er beskrevet og vist i
Figurene 1, 2 og 6, bare representerer de nøyaktige strukturene av cDNA,
genomiske og promoternukleotidsekvenser som her er isolert og beskrevet. Det kan
forventes at svakt modifiserte nukleotidsekvenser vil bli funnet eller ved hjelp av
15 kjente teknikker, kunne modifiseres til å kode for tilnærmet lignende eller homologe
MN-proteiner og polypeptider, for eksempel de som har lignende epitoper, og slike
nukleotidsekvenser og protein/polypeptider ansees å være likeverdige for fore-
liggende oppfinnelses formål. DNA eller RNA som har ekvivalente kodoner ansees å
ligge innenfor rammen av oppfinnelsen i likhet med syntetiske nukleinsyresekvenser
20 som koder for protein/polypeptid-homologer eller tilnærmede homologer til MN-
proteiner/polypeptider, så vel som de nukleinsyresekvenser som vil hybridisere til
nevnte sekvenseksempler [SEKV. ID. NR:1, 5 og 27] under stringente betingelser
eller som, bortsett fra den genetiske kodes degenererthet, vil hybridisere til nevnte
cDNA nukleotidsekvenser under stringente hybridiseringsbetingelser. Modifikasjoner
25 og variasjoner av nukleinsyresekvenser som her antydnet, ansees å resultere i
sekvenser som i det vesentlige er de samme som MN-eksempelsekvenser og
fragmenter derav.

Stringente hybridiseringsbetingelser ansees her å bestå i standard
hybridiseringsbetingelser som på området forstås som stringente. For eksempel er
30 det i alminnelighet underforstått at stringente betingelser omfatter relativt lave salt-
og/eller høye temperaturbetingelser, så som oppnådd med 0,02M til 0,15M NaCl ved
temperaturer på 50°C til 70°C. Mindre stringente betingelser, så som 0,15M til 0,9M
salt ved temperaturer varierende fra 20°C til 55°C kan gjøres mer stringente ved å

tilsette økende mengder formamid som i likhet med øket temperatur tjener til å destabilisere hybrid-duplekser.

Eksempler på stringente hybridiseringsbetingelser er beskrevet i Sambrook *et al.*, **Molecular Cloning: A Laboratory Manual**, s. 1.91 og 9.47-9.51 (2. utg., Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY; 1989); Maniatis *et al.*, **Molecular Cloning: A Laboratory Manual**, s. 387-389 (Cold Spring Harbor Laboratory; Cold Spring Harbor, NY; 1982; Tsuchiya *et al.*, **Oral Surgery, Oral Medicine, Oral Pathology**, 71(6):721-725 (juni, 1991).

Zavada *et al.*, WO 95/34650, beskrev hvorledes en partiell MN cDNA-klon, en full-lengde MN cDNA-klon og MN genomiske kloner ble isolert og sekvensert. Dessuten beskriver Zavada *et al.*, **Int. J. Cancer**, 54:268 (1993) isoleringen og sekvenseringen av en partiell MN cDNA med en lengde på 1397 bp. Kort sagt har forsøk på å isolere en klon av full lengde fra det opprinnelige cDNA-bibliotek vært mislykket. Oppfinnerne foretok derfor en hurtig amplifikasjon av cDNA-ender (RACE) ved å benytte MN-spesifikke primere, R1 og R2 [SEKV. ID. NR:7 og 8], avledet fra 5'-regionen av den opprinnelige cDNA-klon. RACE-produktet ble innført i pBluescript, og hele populasjonen av rekombinante plasmider ble sekvensert med en MN-spesifikk primer ODN1 [SEKV. ID. NR:3]. På denne måte ble det oppnådd en pålitelig sekvens av hele 5'-enden av MN cDNA som vist i Figur 1 [SEKV. ID. NR:1].

Nærmere bestemt ble RACE foretatt ved å benytte 5' RACE System [GIBCO BRL; Gaithersburg, MD (USA)] som følger. 1 µg mRNA (samme som ovenfor) ble benyttet som templat for førstekjede cDNA-syntesen som ble utløst av det MN-spesifikke antisense oligonukleotid, R1

(5'-TGGGGTTCTTGAGGATCTCCAGGAG-3') [SEKV. ID. NR:7]. Førstekjede

produktet ble utfelt to ganger i nærvær av ammoniumacetat og en homopolymer C-hale forbundet med dens 3'-ende med TdT. Denne cDNA med hale ble deretter amplifisert ved PCR ved å benytte en nestet primer, R2

(5'-CTCTAACTTCAGGGAGCCCTCTTCTT-3') [SEKV. ID. NR:8] og en forankringsprimer som renaturerer til den homopolymere hale

(5'-CUACUACUACUAGGCCACGCGTCGACTAGTACGGGI IGGGIIGGGIIG-3')

[SEKV. ID. NR:9]. Det amplifiserte produkt ble kuttet med *Bam*HI og *Sal*I

restriksjonsenzymmer og klonet inn i pBluescript II KS plasmid. Etter transformasjon ble plasmid DNA rensset fra hele populasjonen av transformerte celler og benyttet

som templat for sekvensering med den MN-spesifikke primer ODN1 [SEKV. ID. NR:3; en 29-mer 5' CGCCCAGTGGGTCATCTTCCCCAGAAGAG 3'].

For å studere MN-regulering ble MN genomiske kloner isolert. En MN genomisk klon (Bd3) ble isolert fra et humant cosmidbibliotek fremstillet fra foster-
 5 hjerne ved å benytte både MN cDNA som probe og de MN-spesifikke primerne avledet fra 5'-enden av cDNA ODN1 [SEKV. ID. NR.3, *supra*] og ODN2 [SEKV. ID. NR:4; 19-mer (5' GGAATCCTCCTGCATCCGG 3')]. Sekvensanalyse viste at den genomiske klon dekket en region oppstrøms fra et MN-transkripsjonsstartsete og endte med *Bam*HI restriksjonssetet lokalisert innenfor MN cDNA. Andre MN
 10 genomiske kloner kan isoleres tilsvarende.

Figur 7 gir et skjema over sammenligningsoppstilling av MN genomiske kloner i henhold til transkripsjons-initieringssetet. Plasmider som inneholder A4a-klonen og XE1- og XE3-subklonene er blitt deponert ved American Type Culture Collection (ATCC) 6. juni, 1995, under ATCC deponeringsnummerne henholdsvis 97199,
 15 97200 og 97198.

Exon-intronstruktur av komplett MN genomisk region

Den fullstendig sekvens av de overlappende klonene inneholder 10.898 bp [SEKV. ID. NR:5]. Figur 5 viser organiseringen av det humane MN-gen som viser
 20 lokaliseringen av alle 11 exonene så vel som de 2 oppstrøms og 6 introniske Alu-repetisjonselementene. Alle exonene er små, varierende fra 27 til 191 bp, med unntak av det første exon som er 445 bp. Intronstørrelsen varierer fra 89 til 1400 bp. CA-domenet kodes av exonene 2-8, mens exonene 1, 10 og 11 tilsvarer henholdsvis det proteoglykan-lignende domenet, transmembranforankringen og den cyto-
 25 plasmatiske hale av MN/CA IX-proteinet. Tabell 1 nedenfor angir de spleise-donor- og akseptorsekvenser som stemmer overens med konsensus spleisesekvensene, inklusivt AG-GT-motivet [Mount, **Nucleic Acids Res.**, 10:459-472 (1982)].

Tabell 1

Exon-intron struktur av det humane MN-gen

	Exon	Størrelse	Genomisk Posisjon**	SEKV. ID. NR.	5'-spleise-donor	SEKV. ID. NR.
5	1	445	*3507-3951	28	AGAAG gtaagt	67
	2	30	5126-5155	29	TGGAG gtgaga	68
	3	171	5349-5519	30	CAGTC gtgagg	69
10	4	143	5651-5793	31	CCGAG gtgagc	70
	5	93	5883-5975	32	TGGAG gtacca	71
	6	67	7376-7442	33	GGAAG gtcagt	72
	7	158	8777-8934	34	AGCAG gtgggc	73
	8	145	9447-9591	35	GCCAG gtacag	74
15	9	27	9706-9732	36	TGCTG gtgagt	75
	10	82	10350-10431	37	CACAG gtatta	76
	11	191	10562-10752	38	ATAAT-ende	
	Intron	Størrelse	Genomisk Posisjon**	SEKV. ID. NR.	3'-spleise-akseptor	SEKV. ID. NR.
20	1	1174	3952-5125	39	atacag GGGAT	77
	2	193	5156-5348	40	ccccag GCGAC	78
25	3	131	5520-5650	41	acgcag TGCAA	79
	4	89	5794-5882	42	tttcag ATCCA	80
	5	1400	5976-7375	43	ccccag GAGGG	81
	6	1334	7443-8776	44	tcacag GCTCA	82
	7	512	8935-9446	45	ccctag CTCCA	83
30	8	114	9592-9705	46	ctccag TCCAG	84
	9	617	9733-10349	47	tcgcag GTGACA	85
	10	130	10432-10561	48	acacag AAGGG	86

** Posisjoner er relatert til nt-nummerering i hele genomsekvensen, inklusivt den 5'-flankerende region [Figur 2A-F]

* Tallet tilsvarer transkripsjons-initieringssetet bestemt nedenfor ved RNase-beskyttelsestest

Kartlegging av MN-gentranskripsjonsinitierings- og termineringsseter

Zavada *et al.*, WO 95/34650, beskriver kartleggingsprosessen av MN-gentranskripsjonsinitierings- og termineringssetene. En RNase-beskyttelsestest ble benyttet for finkartlegging av 5'-enden av MN-genet. Proben var en uniformt merket 470 nukleotid kopi RNA (nt -205 til +265) [SEKV. ID. NR:55] som var hybridisert til total RNA fra MN-uttrykkende HeLa- og CGI.3-celler og analysert på en sekvenserende gel. Denne analysen har vist at MN-gentranskripsjonen i multiple seter, initierer 5'-enden av det lengste MN-transkript som er 30 nt lenger enn det tidligere karakterisert ved RACE.

Karakterisering av den 5'-flankerende region

Den Bd3 genomiske klon isolert fra det humane fosterhjerne-cosmidbibliotek, ble funnet å dekke en region på 3,5 kb oppstrøms fra transkripsjonsstartsetet for MN-genet. Det inneholder ingen signifikant kodende region. To Alu-repetisjoner befinner seg i posisjonene -2587 til -2296 [SEKV. ID. NR:56] og -1138 til -877 [SEKV. ID. NR:57] (med hensyn til transkripsjonsstart bestemt ved RNP).

Nukleotidsekvensanalyse av DNA 5' til transkripsjonsstart (fra nt -507) viste ingen gjenkjennelig TATA-boks innenfor den forventede avstand fra begynnelsen av det første exon. Forekomsten av potensielle bindingssteder for transkripsjonsfaktorer tyder imidlertid på at denne region kan inneholde en promoter for MN-genet. Det forekommer flere konsensussekvenser for transkripsjonsfaktorene AP1 og AP2, så vel som for andre regulelementer, inklusivt et p53-bindingsete [Locker & Buzard, J., **DNA Sequencing and Mapping**, 1:3-11 (1990); Imagawa *et al.*, **Cell**, 51:251-260 (1987); El Deiry *et al.*, **Nat. Genet.**, 1:44-49 (1992)]. Selv om den antatte promoterregion inneholder 59,3% C + G, har den ikke ytterligere kjennetegn på CpG-rike øyer som er typiske for TATA-manglende promotere av husholdningsgener [Bird, **Nature**, 321:209-213 (1986)]. En annen klasse av gener som mangler TATA-boks gjør bruk av initiator- (Inr) elementet som en promoter. Mange av disse genene er ikke konstitutivt aktive, men reguleres derimot under differensiering eller utvikling. Inr har en konsensussekvens av PyPyPyCAPyPyPyPyPy [SEKV. ID. NR:23] og omfatter transkripsjonsstartsetet [Smale & Baltimore, **Cell**, 57:103-113 (1989)]. Det finnes to slike konsensussekvenser i den antatte MN-promoter, men de overlapper imidlertid ikke transkripsjonsstarten (Figur 6).

En interessant region ble funnet i midten av MN-genet. Regionen er ca. 1,4 kb lang [nt 4.600-6.000 av den genomiske sekvens; SEKV. ID. NR:49] og strekker seg fra 3'-delen av det 1. intron til enden av det 5. exon. Regionene har karakteren av en typisk CpG-rik øy, med 62,8% C + G-innhold og 82 CpG:131 GpC-dinukleotider. Det forekommer dessuten flere antatte bindingssteder for transkripsjonsfaktorene AP2 og Sp1 [Locker & Buzard, *supra*; Briggs *et al.*, **Science**, **234**:47-52 (1986)] konsentrert i senteret for dette området. Særlig inneholder 3. intronet med lengde 131 bp tre Sp1- og tre AP2-konsensussekvenser. Disse data tyder på mulig medvirkning av vedkommende region i reguleringen av MN-genekspresjon. Funksjonaliteten av denne region, så vel som andre reguletoelementer funnet i den foreslåtte 5' MN-promoter, gjenstår imidlertid å bestemme.

MN-promoter

Studiet av MN-promoterer har vist at den er uten TATA og inneholder regulatorsekvenser for AP-1, AP-2 så vel som to p53-bindingssteder. Sekvensen av 5'-enden av den 3,5 kb flankerende region oppstrøms for MN-genet, har vist omfattende homologi med LTR av HERV-K endogene retrovirus. Basal transkripsjonsaktivitet av promoterer er meget svak, hvilket vises av analyse ved bruk av CAT og neo-rapportørgener. Ekspresjon av rapportørgenet økes imidlertid flere ganger når det styres av den 3,5 kb flankerende region, hvilket indikerer at antatte enhancere er involvert.

Funksjonell karakterisering av den 3,5 kb MN 5' oppstrøms region ved delesjonsanalyse førte til identifisering av [-173, +31] fragmentet [SEKV. ID. NR:21] (også alternativt, men mindre foretrukket, det nesten identiske -172, +31 fragment [SEKV. ID: NR:91] som MN-promoterer. *In vitro* DNase I footprinting viste nærværet av 5 beskyttede regioner (PR) innen MN-promoterer. Detaljert delesjonsanalyse av promoterer identifiserte PR1 og 2 (nummeret fra transkripsjonsstarten) som den mest avgjørende for transkripsjonsaktivitet. PR4 [SEKV. ID. NR:115] påvirket transkripsjon negativt idet dens delesjon førte til øket promoteraktivitet og ble bekreftet å virke som et promoter-, posisjons- og orienteringsuavhengig nøytraliserende element (silencer element). Mutasjonsanalyse tydet på at den direkte repetisjon AGGGCacAGGGC [SEKV. ID NR:143] fordres for effektiv repressorbinding. To komponenter av repressorkomplekset (35 og 42 kDa) ble ved UV kryssbinding funnet å være i direkte kontakt med PR4. Øket celletetthet, kjent for

å indusere MN-ekspresjon, påvirket ikke nivåene av PR4-binding i HeLa-celler. Signifikant redusert repressornivå synes å være ansvarlig for MN-oppregulering når det gjelder tumorigen CGL3 sammenlignet med ikke-tumorigen CGL1 HeLa x normale fibroblast hybridceller.

5

Anvendelse av MN-promoter som en tumorspesifikk promoter for genterapi

Det som undersøkes er hvorvidt MN-genpromoterer kan benyttes som en tumorspesifikk promoter for å styre ekspresjonen av et selvmords-gen [tymidinkinase (tk) av HSV] og mediere den direkte og «bystander» dreping av tumorceller. HSVtk-gen overført til tumorceller, omdanner nukleosidanalogen ganciklovir (GCV) til toksiske trifosfater og medierer døden for transduserte og også nabo-tumorceller. Styringen av HSVtk gjennom MN-genpromoterer ville tillate dens ekspresjon bare i tumorceller, hvilket gir mulighet for biosyntesen av MN-protein, og selektivt drepe tumorceller, men ikke normale celler som MN-ekspresjon er undertrykket i.

En plasmidkonstruksjon hvor HSVtk ble klonet nedstrøms for MN-promoter-regionen Bd3, inneholdende både proksimale og fjerntliggende regulelementer av MN, ble fremstillet. Dette plasmid pMN-HSVtk, ble transfektert til rat2TK-celler og C33 humane livmorhals-karsinomceller ved å benytte henholdsvis kalsiumfosfat-presipitasjon og lipofeksjon. Transfektanter ble testet med henblikk på ekspresjon av HSVtk- og GVC-følsomhet. Analyse av transfektantene har vist den bemerkelsesverdige cytotoxiske *in vitro* effekt av GVC endog i lave konsentrasjoner (opptil 95% av celler drept).

Polyklonalt kanin antiserum mot HSVtk under bruk av fusjonsprotein med GST i pGEX-3X, er fremstillet for å immunpåvise HSVtk syntetisert i transfekterte celler. Dette modellsystem studeres for å anslå «bystander»-effekten, inhiberingen av kloningseffektiviteten og invasiviteten av transduserte og GVC-behandlede celler til kollagenmatrikser. En rekombinant retroviral vektor med det MN-promoterstyrte HSVtk, bør fremstilles for å teste dets *in vivo* effekt ved å benytte en dyremodell (f.eks. SCID-mus).

30

MN-promoteranalyse

Siden MN-promoteren er svak ville en klassisk måte for å studere den være begrenset som følge av den relativt lave effektivitet av forbigående transfeksjoner (opptil 10%). Stabile klonale cellelinjer som uttrykker konstruksjoner inneholdende MN-promoteren fusjonert til CAT-genet, ble derfor fremstillet. I slike klonale linjer uttrykker 100% av cellene CAT-genet som styres fra MN-promoteren, og aktiviteten av promoteren er således lettere påvisbar enn i transiente forsøk. Promoter-aktiviteten kan også analyseres gjentatte ganger i de samme celler under ulike betingelser eller behandles med forskjellige faktorer og medikamenter. Denne tilnærming muliggjør studiet av de mekanismer som ligger til grunn for MN-regulering på transkripsjonsinitieringsnivået.

Flere typer transfeksjoner med promoterkonstruksjoner koblet til et rapportør CAT-gen (kalsiumpresipitasjon, DEAE-dekstran kombinert med DMSO-sjokk og/eller klorokin, så vel som elektroporering), forskjellige metoder for CAT-aktivitetsbestemmelse (scintillasjonsmetode, tynnskiktkromatografi) og flere resipient-cellelinjer adskiller seg i graden av MN-ekspresjon og i transfeksjonseffektivitet (HeLa-, SiHa-, CGL3-, KATO III-, Rat2TK- og C33-celler). Aktiviteten av MN-promoteren ble påvist, fortrinnsvis ved elektroporering av CGL3-celler og tynnskiktkromatografi. Videre ble fortrinnsvis C33-celler, kotransfektet med MN-promoter CAT-konstruksjoner og pSV2neo benyttet.

1. For å påvise basalaktivitet av MN-promoteren og anslå posisjonen av kjernepromoteren, ble ekspresjonen av CAT-genet fra konstruksjonene pMN1 til pMN7 etter transfeksjon til CGL3-celler, analysert. Plasmider med progressive 5'-delesjoner ble transfektet inn i CGL3-celler, og aktiviteten ble analysert ved CAT-bestemmelse. [8 µg DNA ble i alle tilfellene unntatt pBLV-LTR (2 µg) benyttet for transfeksjon].

Kun meget svak CAT-aktivitet ble påvist i celler transfektet med pMN1 og pMN2 (inneholdende henholdsvis 933 bp og 600 bp av promotersekvensen). En litt høyere aktivitet ble oppvist med konstruksjonene pMN3, pMN4 og pMN6 (inneholdende henholdsvis 446 bp, 243 bp og 58 bp av promoteren). En svak aktivitets-topp ble oppnådd med pMN5 (startende i posisjon -172 i forhold til transkripsjons-start). Virkningen av MN-kjernepromoteren kan således tilordnes til en region på ca. 500 bp umiddelbart oppstrøms fra MN-transkripsjonsinitieringssetet.

Det er interessant at aktiviteten av den større Bd3-region (som dekker 3,5 kbp oppstrøms for transkripsjonsstart) var flere ganger høyere enn aktiviteten av kjerne-promoteren. Nivået var imidlertid fremdeles meget lavere enn det som oppvises av en positiv kontroll, dvs. BLV-LTR transaktivert med Tax, og enda lavere enn aktiviteten av BLV-LTR uten transaktivering. At aktiviteten av Bd3 var høyere sammenlignet med kjernepromoteren tyder på tilstedeværelse av et eller annet regulerende element. Slike elementer befinner seg mest sannsynlig i sekvensen mellom pMN1 og Bd3 (dvs. fra -1 kbp til -3,5 kbp) [SEKV. ID. NR.58]. Kloningen og transfeksjonen av flere delesjonsversjoner av Bd3 som dekker den angitte region, kan benyttes for å bestemme lokaliseringen av de antatte reguletoelementene.

Lignende resultater ble oppnådd ved å transfekte KATO III-celler med Bd3 og pMN4. De transfekte cellene uttrykte et lavere nivå av MN enn CGL3-cellene. Aktiviteten av MN-promoteren ble derfor funnet å være lavere enn i CGL3-celler.

2. I et parallelt forsøk på å studere MN-promoteren, ble det foretatt en analyse på grunnlag av G418-seleksjon av celler transfektert med plasmider som inneholdt den aktuelle promoter klonet oppstrøms fra neo-genet. Denne tilnærming egner seg for å studere svake promotere, siden følsomheten for denne er meget høyere enn den for en standard CAT-bestemmelse. Det prinsipp som ligger til grunn for metoden er følgende: en aktiv promoter styrer ekspresjon av neo-genet som beskytter transfekte celler fra den toksiske effekt av G418, mens en inaktiv promoter resulterer i at det ikke dannes noe neo-produkt og at de cellene som derved transfekteres, dør etter virkningen av G418. Aktiviteten av promoteren kan derfor anslås ut fra det antall cellekolonier som oppnås etter to ukers seleksjon med G418. Tre konstruksjoner ble benyttet i de innledende forsøk - pMN1neo, pMN4neo og pMN7neo. Siden pMN7neo bare inneholder 30 bp oppstrøms for transkripsjonsstartsetet, ble den betraktet som en negativ kontroll. Som positiv kontroll ble pSV2neo med en promoter avledet fra SV40 benyttet. Rat2TK⁻-celler ble valgt som mottagerceller siden de kan transfekteres med høy effektivitet ved hjelp av kalsium-presipitasjonsmetoden.

Etter transfeksjon ble cellene underkastet seleksjon i to uker. Deretter ble mediet fjernet, cellene rensset med PBS og koloniene synliggjort med farving med metylenblått. Resultatene oppnådd fra tre uavhengige forsøk understøttet datene fra CAT-bestemmelsene. Promoterkonstruksjonen pMN1neo4 oppviste høyere transkripsjonsaktivitet enn pMN1neo. Forskjellen mellom den positive kontroll og

pMN4neo var imidlertid ikke så slående som i CAT-bestemmelsen. Dette kan ha vært en følge av både lavere promoteraktivitet av pSV2neo sammenlignet med Tax-transaktivert pBLV-LTR og forskjellige betingelser for cellevekst etter transfeksjon. Ut fra dette er stabil transfeksjon sannsynligvis mer fordelaktig for MN-ekspressjon, siden

5 cellene vokser i kolonier med tett celle-til-celle kontakt og fordi forsøket varer meget lenger og gir en bedre mulighet for å påvise promoteraktivitet.

3. Stabile transfektanter som uttrykker MN-promoter CAT kimære gener ble fremstillet ved kotransfeksjon av relevante plasmider med pSV2neo. Som mottagerceller ble det først benyttet HeLa-celler. Det ble imidlertid ikke oppnådd

10 noen kloner som uttrykte promoter CAT-konstruksjonene. Dette negative resultat var sannsynligvis forårsaket av homolog rekombinasjon av den transfekteerte genomiske region av MN (f.eks. promoteren) med den tilsvarende endogene sekvens. Basert på denne erfaring ble C33-celler avledet fra et HPV-negativt livmorhalskarsinom benyttet. C33-celler uttrykker ikke MN siden de under tumorigeneseprosessen tapte

15 genetisk materiale, herunder den kromosomale region 9p som inneholder MN-genet. I disse forsøkene kan fraværet av MN-genet representere en fordel siden muligheten for homologe rekombinasjoner unngås.

C33-celler transfektert med MN-promoter CAT-konstruksjoner

20 C33-celler som uttrykker CAT-genet under MN-promoterregionene Bd3 (-3500/+331) [SEKV. ID. NR:90] og MN5 (-172/+31) [SEKV. ID. NR:91] ble benyttet i de innledende forsøk for å analysere påvirkningen av celletetthet på den transkripsjonelle aktivitet av MN-promoteren. Resultatene tydet på at signaler utviklet etter at cellene kommer i nær kontakt, aktiverer transkripsjon av CAT-proteinet fra

25 MN-promoteren i forhold til tettheten av cellekulturen. Det er interessant at dataene antyder at MN-proteinet ikke fordres for denne signaltransduksjonsfase, siden påvirkningen av tetthet klart demonstreres i MN-negative C33-celler. Det synes snarere som at MN-protein virker som et effektormolekyl produsert i tette celler for å bevirke en viss biologisk funksjon (dvs. for å forstyrre kontaktinhibering). Det er også

30 interessant at MN-promoteraktiviteten er påvisbar selv i meget glisne cellekulturer, hvilket tyder på at MN uttrykkes i meget lav grad også i sparsomt subkonfluent kultur.

Delesjonsvarianter. Delesjonsvarianter av Bd3-CAT-promoter-konstruksjonen ble deretter fremstillet. Konstruksjonene ble kotransfektet med pSV2neo inn i C33 livmorhalsceller. Etter seleksjon med G418 ble hele populasjonen

av stabilt transfekterte celler underkastet CAT ELISA-analyse. Ekspresjon av delesjonskonstruksjonene resulterte i syntese av lignende nivåer av CAT-protein som de oppnådd med Bd3-CAT-konstruksjonen. På grunnlag av disse preliminare data har oppfinnerne ment at sekvenser som stimulerer transkripsjon av MN befinner seg mellom -3506 og -3375 bp [SEKV. ID. NR:92] oppstrøms fra transkripsjonsstart. Det er den sekvens om oppviser homologi med HERV-K LTR.

Transiente transfeksjonsstudier på CGL3-celler viste imidlertid gjentatte ganger at LTR-regionen ikke fordres for økningen av basal MN-promoteraktivitet. Resultater oppnådd i CGL3-celler tydet videre på at det aktiverende element er lokalisert i regionen fra -933 til -2179 [SEKV. ID. NR:110] når det gjelder transkripsjonsinitieringssetet (den posisjon av regionen som er utledet av overlappende sekvenser i Bd3-delesjonsmutantene).

Interaksjon mellom nukleære proteiner og MN-promotersekvenser

For å identifisere transkripsjonsfaktorer som binder til MN-promoteren og eventuelt regulerer dens aktivitet, ble det foretatt en rekke analyser ved bruk av EMSA (electrophoretic mobility shift assay) og DNase I FTP (footprinting) analyse.

EMSA

Ved EMSA fikk rensede promoterfragmenter MN4 (-243/+31) [SEKV. ID. NR:93], MN5 (-172/+31) [SEKV. ID. NR:91], MN6 (-58/+31) [SEKV. ID. NR:94] og pMN7 (-30/+31) [SEKV. ID. NR:95], merket med Klenow-enzymet i 3'-endene, anledning til interaksjon med proteiner i kjerneekstrakter fremstillet fra CGL1- og CGL3-celler. [40 µg kjerneproteiner ble inkubert med 30.000 cpm endemerkede DNA-fragmenter i nærvær av 2 µg poly(dIdC)]. DNA-proteinkomplekser ble analysert ved PAGE (nativ 6%), hvor kompleksene dannet ekstra bånd som vandret langsommere enn de frie DNA-fragmentene som følge av den mobilitetsforskyvning som avhenger av den bundne proteindel.

EMSA av MN4- og MN5-promoterfragmenter avdekket flere DNA-protein-komplekser. De respektivt oppnådde bindingsmønstre med CGL1 og CGL3 kjerneekstrakter var imidlertid ikke identiske. Det foreligger et enkelt CGL1-spesifikt kompleks.

EMSA av MN6-promoterfragmentet resulterte i dannelse av tre identiske komplekser både med CGL1- og CGL3-kjerneekstrakter, mens MN7-promoterfragmentet ikke bandt noen nukleære proteiner.

EMSA-resultatene tydet på at CGL1-kjerneekstraktet inneholder en spesifikk faktor som kunne delta i den negative regulering av MN-ekspressjon i CGL1-celler. Siden det spesifikke DNA-proteinkompleks dannes med MN4 (-243/+31) [SEKV. ID. NR:93] og MN5 (-172/+31) [SEKV. ID. NR:91] promoterfragmentene, men ikke med MN6 (-58/+31) [SEKV. ID. NR:94], synes det som om bindingssetet av protein-komponenten av vedkommende spesifikke kompleks er lokalisert mellom -173 og -58 bp [SEKV. ID. NR:96] i forhold til transkripsjonsinitiering.

Det neste trinn var en rekke EMSA-analyser ved bruk av dobbeltkjadede (ds) oligonukleotider utviklet i henhold til de beskyttede regionene i FTP-analysen. Et ds-oligonukleotid avledet fra den beskyttede region PR2 [som dekker sekvensen fra -72 til -56 bp (SEKV. ID. NR:111)] av MN-promoteren, bekreftet bindingen av AP-1-transkripsjonsfaktoren i kompetitiv EMSA ved anvendelse av kommersielle ds-oligonukleotider som representerer bindingssetet for AP-1.

EMSA av ds-oligonukleotider avledet fra de beskyttede regioner av PR1 [-46 til -24 bp (SEKV. ID. NR:112)], PR2 [-72 til -56 bp (SEKV. ID. NR:111)], PR3 [-102 til -85 (SEKV. ID. NR:113)] og PR5 [-163 til -144 (SEKV. ID. NR:114)] avdekket ingen forskjeller i bindingsmønsteret av nukleære proteiner ekstrahert fra CGL1- og CGL3-celler, hvilket indikerer at disse regionene ikke binder avgjørende transkripsjonsfaktorer som kontrollerer aktiveringen av MN-genet i CGL3, eller dets negative regulering i CGL1. EMSA av ds-oligonukleotidene fra den beskyttede region PR4 [-133 til -108; SEKV. ID. NR:115] oppviste imidlertid gjentatt bemerkelsesverdige kvantitative forskjeller mellom binding av CGL1- og CGL3-nukleære proteiner. CGL1-nukleære proteiner dannet en betydelig høyere mengde DNA-proteinkomplekser, hvilket indikerer at PR4-regionen inneholder et bindingssete for spesifikke transkripsjonsfaktorer som kan representere en negativ regulator av MN-gen-transkripsjon i CGL1-celler. Dette stemmer overens med de tidligere EMSA-data som oppviste CGL1-spesifikt DNA-proteinkompleks med promoterfragmentene pMN4 (-243/+31; SEKV. ID. NR:93) og pMN5 (-172/+31; SEKV. ID. NR:91), men ikke med pMN6 (-58/+31; SEKV. ID. NR:94).

For å identifisere det involverte protein eller dannelsen av et spesifikt kompleks med MN-promoteren i PR4-regionen, vil det bli benyttet relevante ds-

oligonukleotider som kovalent er bundet til magnetkuler, for å rense den tilsvarende transkripsjonsfaktor. Alternativt vil ONE Hybrid System® [Clontech (Palo Alto, CA (USA))] bli benyttet til å lete etter og klonere transkripsjonsfaktorer som er involvert i reguleringen av den analyserte promoterregion. Et cDNA-bibliotek fra HeLa-celler vil bli benyttet for undersøkelsen.

FTP

For å bestemme den nøyaktige lokalisering av Cis-regulatorelementer som deltar i den transkripsjonelle regulering av MN-genet, ble FTP benyttet. Proteiner i kjerneekstrakter fremstillet henholdsvis fra CGL1- og CGL3-celler, fikk interagere med et renset ds-DNA-fragment av MN-promoter (MN4, -243/+31) [SEKV. ID. NR:93] som ble merket i 5'-enden av den ene kjede. [MN4-fragmenter ble merket enten i *Xho* 1-setet (-243/+31*) eller i *Xba* 1-setet (*-243/+31)]. DNA-protein-komplekset ble deretter underkastet DNase I-angrep, som forårsaket at DNA-kjeden ble brutt i visse baser dersom de ikke var i kontakt med proteiner. [En kontroll benyttet BSA i stedet for DNase]. Undersøkelse av båndmønsteret av den denaturerte DNA etter gelelektroforese [8% denaturerende gel] angir hvilke av basene på den merkede kjede som ble beskyttet av protein.

FTP-analyse av MN4-promoterfragmentet avdekket 5 områder (I-V) beskyttet i både den kodende og den ikke-kodende kjede, så vel som 2 regioner (VI og VII) beskyttet i den kodende kjede, men ikke i den ikke-kodende kjede. Figur 6 indikerer de generelle regioner på MN-promoter som var beskyttet.

Sekvensene av de identifiserte beskyttede regionene (PR) ble underkastet datamaskinanalyse ved å benytte SIGNALSCAN-programmet for å se om de tilsvarte kjente konsensussekvenser for transkripsjonsfaktorer. Dataene oppnådd gjennom denne datamaskinanalyse er som følger:

- PR I - kodende kjede - AP-2, p53, GAL4 ikke-kodende kjede - JCV-repetert
- PR II - kodende kjede - AP-1, CGN4 ikke-kodende kjede - TCF-1, dFRA, CGN4
- PR III - kodende kjede - ingen kjent konsensussekvens, bare delvis overlapping av AP1 ikke-kodende kjede - 2 TCF-1 seter
- PR IV - kodende kjede - TCF-1, ADR-1 ikke-kodende kjede - CTCF, LF-A1, LBP-1
- PR V - kodende kjede - intet kjent konsensusmotiv ikke-kodende kjede - JCV-repetert

PRVI - kodende kjede - intet kjent konsensusmotiv ikke-kodende kjede - T-antigen av SV40, GAL4

PRVII - kodende kjede - NF- κ B, U2snRNA.2 ikke-kodende kjede - AP-2, IgHC.12, MyoD.

5

I motsetning til EMSA, fant FTP-analyse ingen forskjeller mellom CGL1- og CGL3-kjerneekstrakter. Nærværet av spesifikke DNA-proteininteraksjoner påvist i CGL1-kjerneekstrakter ved EMSA kunne imidlertid ha vært resultat av bindingen av ytterligere protein for å danne DNA-protein-proteinkompleks. Dersom det spesifikke protein ikke kontaktet DNA-sekvensen direkte, ville dets nærvær ikke være påvisbart ved FTP.

10

EMSA superskift-analyse

Resultatene av FTP tyder på at transkripsjonsfaktorene AP-1, AP-2 så vel som tumorsuppressorprotein p53 eventuelt er involvert i reguleringen av MN-ekspressjon. For å bekrefte binding av disse særlige proteinene til MN-promoter ble det foretatt en superskift-analyse ved å benytte antistoffer som var spesifikke for disse proteinene. For denne analyse fikk DNA-proteinkomplekser, fremstillet som beskrevet for EMSA, interagere med MAb eller polyklonale antistoffer, spesifikke for proteiner som eventuelt var inkludert i komplekset. Bindingen av antistoff til det tilsvarende protein resulterer i en ytterligere forskyvning (superskift) i mobiliteten av DNA-protein-antistoffkomplekset som ved PAGE visualiseres som et ytterligere, mer langsomt vandrende bånd.

20

Ved hjelp av denne metode ble bindingen av AP-2 til MN-promoter bekreftet. Denne metode beviste imidlertid ikke binding av AP-1-transkripsjonsfaktoren. Det er mulig at MN-proteiner binder AP-1-relatert protein som er antagelig forskjellig fra det AP-1 som gjenkjennes av de antistoffer som benyttes i denne test.

25

Av stor interesse er også den mulige binding av p53 tumorsuppressorproteinet til MN-promoter. Det er vel kjent at wt p53 virker som en transkripsjonsfaktor, som aktiverer ekspressjon av vekstrestriksjonsgener og direkte eller indirekte, nedmodulerer ekspressjonen av gener som fordres for pågående celleproliferasjon. Forsøk med transient kotransfeksjon ved bruk av pMN4-CAT-promoterkonstruksjonen i kombinasjon med henholdsvis wt p53 cDNA og mut p53 cDNA, tydet på at wt p53, men ikke mut p53, negativt regulerer ekspressjon av MN. I tillegg beskyttes ett av to

30

p53-bindingssteder i MN-promoteren ved FTP-analyse (Figur 6), hvilket indikerer at det bindes til det tilsvarende protein. Supershift-analyse for å bevise at p53 binder til MN-promoteren med to p53-spesifikke antistoffer, f.eks. Mab 421 og DO-1 [sistnevnte stillet til disposisjon av Dr. Vojtesek fra Masaryk Memorial Cancer Institute i Brno, Tsjekkia] må derfor foretas med passende kjerneekstrakter, f.eks. fra MCF-7 brystkarsinomceller som i tilstrekkelig grad uttrykker wt p53.

Regulering av MN-ekspresjon og MN-promoter

MN viser seg å være et nytt regulatorprotein som direkte er involvert i kontrollen av celleproliferasjon og i cellulær transformasjon. I HeLa-celler reguleres ekspresjonen av MN positivt av celletetthet. Nivået økes ved vedvarende infeksjon med LCMV. I hybridceller mellom HeLa og normale fibroblaster korrelerer MN-ekspresjon med tumorigenitet. Det faktum at MN ikke forekommer i ikke-tumorigene hybridceller (CGL1), men uttrykkes i en tumorigen segregant som mangler kromosom 11, indikerer at MN negativt reguleres av en antatt suppressor i kromosom 11.

Indikasjoner som underbygger den regulerende rolle av MN-protein, ble funnet i utviklingen av stabile transfectanter av NIH 3T3-celler som konstitutivt uttrykker MN-protein. Som en følge av MN-ekspresjon fikk NIH 3T3-cellene trekk som var assosiert med en transformert fenotype: endret morfologi, øket tetthet, proliferativ fordel i serumreduserte medier, øket DNA-syntese og evne til forankrings-uavhengig vekst. Videre indikerte strømningscytometriske analyser av asynkrone cellepopulasjoner at ekspresjonen av MN-protein fører til aksellerert progresjon av celler gjennom G1-fase, reduksjon av cellediameter og tap av evnen til vekststans under upassende betingelser. Dessuten oppviste MN-uttrykkende celler en nedsatt følsomhet for det DNA-ødeleggende medikament, mitomycin C.

Ikke-tumorigene humanceller, CGL1-celler, ble også transfectert med det full-lengdes MN cDNA. Den samme pSG5C-MN-konstruksjon i kombinasjon med pSV2neo plasmid, som benyttet til å transfectere NIH 3T3-celler, ble benyttet. Av i alt 15 MN-positive kloner (testet ved SP-RIA og western blotting), ble 3 valgt for ytterligere analyse. To MN-negative kloner, isolert fra CGL1-celler transfectert med tomt plasmid, ble tilsatt som kontroller. Innledende analyse tydet på at morfologien og vekstvaner til MN-transfecterte CGL1-celler ikke endres dramatisk, men at deres proliferasjonshastighet og utplatingseffektivitet økes.

MN-promoter - sense/antisense konstruksjoner

Når promoterregionen fra den MN genomiske klon, isolert som beskrevet ovenfor, ble koblet til MN cDNA og transfektert inn i CGL1-hybridceller var ekspresjonen av MN-protein påvisbar umiddelbart etter seleksjon. Deretter opphørte den imidlertid gradvis, hvilket indikerer en virkning av en tilbakekoblingsregulator. Det antatte regulelement syntes å virke via MN-promoter, fordi det ikke ble observert noen lignende effekt når full-lengdes cDNA (som ikke inneholdt promoteren) ble benyttet for transfeksjon.

En «antisense» MN cDNA/MN-promoter-konstruksjon ble benyttet til å transfektere CGL3-celler. Effekten var den motsatte av den for CGL1-celler transfektert med «sense» konstruksjonen. Mens de transfekterte CGL1-cellene dannet kolonier som var flere ganger større enn kontroll CGL1, dannet transfekterte CGL3-celler meget mindre kolonier enn kontroll CGL3-cellene. Det samme resultat ble oppnådd med antisense MN cDNA transfeksjon i SiHa- og HeLa-celler.

For disse forsøkene ble den del av promoterregionen som var koblet til MN cDNA gjennom et *Bam*HI-sete, avledet fra et *Nco*I-*Bam*HI-fragment av den MN genomiske klon [Bd3], og den representerer en region noen få hundre bp oppstrøms fra transkripsjonsinitieringssetet. Etter ligeringen ble det forente DNA innskutt i en pBK-CMV ekspresjonsvektor [Stratagene]. Orienteringen som den innskutte sekvens fordret ble sikret ved rettet kloning og deretter verifisert ved restriksjonsanalyse. Transfeksjonsprosedyren var den samme som benyttet ved transfektering av NIH 3T3-celler, men kotransfeksjon med pSV2neo plasmidet var ikke nødvendig siden neo-seleksjonsmarkøren allerede var inkludert i pBK-CMV-vektoren.

Etter to ukers seleksjon i et medium inneholdende G418, var bemerkelsesverdige forskjeller mellom antall og størrelse av de dyrkede kolonier tydelige, slik som nevnt ovenfor. Umiddelbart etter seleksjonen og kloningen, ble de MN-transfektete CGL1- og CGL3-cellene testet ved SP-RIA med henblikk på henholdsvis ekspresjon og represjon av MN. De isolerte transfektete CGL1-klonene var MN-positive (selv om nivået var lavere enn hva som ble oppnådd med full-lengde cDNA), mens MN-protein var nesten fraværende fra de transfektete CGL3-klonene. I senere passasjer begynte ekspresjonen av MN i transfektete CGL1-celler å opphøre og ble deretter blokkert, hvilket muligens er tegn på en tilbakekoblingskontrollmekanisme.

Som resultat av den svært meget nedsatte proliferasjon av de transfekterte CGL3-cellene, var det vanskelig å ekspandere størstedelen av klonede celler (i henhold til SP-RIA, de med det laveste MN-nivå), og de gikk tapt under passering. Enkelte kloner tålte problemet og uttrykte igjen MN. Det er når disse cellene først
 5 kom opp i en høyere mengde, mulig at nivået av endogent produsert MN mRNA steg over mengden av ektopisk uttrykt antisense mRNA.

Identifisering av spesifikke transkripsjonsfaktorer involvert i kontroll av MN-ekspresjon

10 Kontroll av MN-ekspresjon på transkripsjonsnivået involverer regulator-
elementer av MN-promoteren. Disse elementene binder transkripsjonsfaktorer som
er ansvarlige for MN-aktivering i tumorceller og/eller for represjon i normale celler.
Identifiseringen og isolering av disse spesifikke transkripsjonsfaktorene og en
forståelse av hvorledes de regulerer MN-ekspresjon, kunne gjøre dem terapeutisk
15 anvendelige til modulering av MN-ekspresjon.

EMSA-forsøk tyder på eksistensen av en MN-genrepressor. Ved å benytte
One Hybrid System® [Clontech (Palo Alto, CA); en genetisk *in vivo* test på gjær for å
isolere gener som koder for proteiner som binder til et mål, et cis-virkende regulator-
element eller en hvilken som helst annen kort DNA-bindingssekvens; Fields & Song,
20 **Nature**, 340:245 (1989); Wu *et al.*, **EMBO J.**, 13:4823 (1994)] og subtraktiv
suppressiv PCR (SSH). SSH muliggjør kloning av gener som er uttrykt ulikt under
betingelser som er kjent for å opp- eller nedregulere MN-ekspresjon, så som tetthet
versus glissenhets av HeLa-celler, og suspensjon versus adherente HeLa-celler.

I forsøk med HPV immobiliserte livmorhalsceller (HCE 16/3) ble det funnet at
25 reguleringen av MN-ekspresjon avviker fra den i fullstendig transformerte karsinom-
celler. For eksempel regulerer glukokortikoidhormoner, som aktiverer HPV-
transkripsjon, negativt MN-ekspresjon i HCE, men stimulerer MN i HeLa og SiHa.
Ytterligere keratinocyttvekstfaktorer som nedregulerer transkripsjon av HPV-
onkogener, stimulerer MN-ekspresjon i suspensjon HCE, men ikke i adherente celler.

30 EGF og insulin er involvert i aktiveringen av MN-ekspresjon i både
immortaliserte celler og karsinomceller. Alle de angitte fakta kan benyttes ved leting
etter MN-spesifikke transkripsjonsfaktorer og moduleringen av MN-ekspresjon for
terapeutiske formål.

Dedusert aminosyresekvens

ORF av MN cDNA vist i Figur 1 har kodende evne for et 459 aminosyreprotein med en beregnet molekylvekt på 49,7 kd. Den totale aminosyresammensetning av MN/CA IX proteinet er heller surt, og forventes å ha en pI på 4,3. Analyse av nativt MN/CA IX protein fra CGL3-celler ved todimensjonal elektroforese, etterfulgt av immunblotting, har vist at MN/CA IX, i overensstemmelse med datamaskin-prediksjon, er et surt protein som eksisterer i flere isoelektriske former med pI varierende fra 4,7 til 6,3.

Som bestemt ved aminosyresekvensanalyse kan den deduserte primærstruktur av MN-proteinet inndeles i fire distinkte regioner. Den innledende hydrofobe region på 37 aminosyrer (aa) tilsvarer et signalpeptid. Det modne protein har en N-terminal eller ekstracellulær del på 377 aminosyrer [aa 38-414 (SEKV. ID. NR:87)], et hydrofobt transmembransegment på 20 aminosyrer [aa 415-434 (SEKV. ID. NR:52)] og en C-terminal region på 25 aminosyrer [aa 435-459 (SEKV. ID. NR:53)].

Den ekstracellulære del er sammensatt av to distinkte domener: (1) et proteoglykan-lignende domene [aa 53-111 (SEKV. ID. NR:50)]; og (2) et CA-domene, lokalisert nær plasmamembranen [aa 135-391 (SEKV. ID. NR:51)]. [Aminosyrenummerne er låst til de i Figur 1].

Et mer detaljert innsyn i MN-proteinets primærstruktur viser nærvær av diverse konsensussekvenser. Et potensielt N-glykosyleringsete ble funnet i posisjon 346 i Figur 1. Dette trekk, sammen med en predikert membran-overspennende region, stemmer overens med de resultater hvor MN ble vist å være et N-glykosylert protein lokalisert i plasmamembranen. MN-proteinsekvens dedusert fra cDNA ble også funnet å inneholde syv S/TPXX-sekvenselementer [SEKV. ID. NR:25 og 26] (den ene av disse er signalpeptidet) definert av Suzuki, *J. Mol. Biol.*, 207:61-84 (1989) som motiver som ofte finnes i genregulerende proteiner. Bare to av disse er imidlertid sammensatt av de foreslåtte konsensus-aminosyrene.

Forsøk med affinitetskromatografi under bruk av Zn-ladet chelaterende Sepharose har vist at MN-proteinet er i stand til å binde sink-kationer. MN-protein immunpresipitert fra HeLa-celler med Mab M75 ble funnet å ha svak katalytisk CA-aktivitet. Det CA-lignende domenet av MN er strukturelt predisponert til å tjene som et bindingssete for små løselige domener. MN-proteinet kunne således mediere en eller annen signaloverføring.

MN-protein fra LCMV-infiserte HeLa-celler ble ved bruk av DNA cellulose-affinitetskromatografi vist å binde til immobilisert dobbeltkjedet laksesperma DNA. Bindingsaktiviteten fordret både nærvær av sink-kationer og fravær av et reduksjonsmiddel i bindingsbufferen.

CA-domene nødvendig for forankrings-uavhengighet unntatt for øket proliferasjon av transfekterte

NIH 3T3 fibroblaster

I transfekterte NIH 3T3 fibroblaster inducerer MN-protein morfologisk transformasjon, øket proliferasjon og forankrings-uavhengighet. Følgene av konstitutiv ekspresjon av to MN-trunkerte varianter i NIH 3T3 celler ble studert. Det viste seg at den proteoglykan-lignende region er tilstrekkelig for den morfologiske endring av transfekterte celler og oppviser den vekstfremmende aktivitet som formodentlig er relatert til forstyrrelse av kontaktinhibering.

CA-domenet er essensielt for induksjon av forankrings-uavhengighet, mens TM-forankringen og IC-halen ikke trengs for denne biologiske effekt. MN-proteinet kan også forårsake plasmamembran-rynking i de transfekterte cellene og synes å bidra til deres festing til den faste bærer. Dataene tilkjenner at MN er involvert i reguleringen av celleproliferasjon, adhesjon og intercellulær kommunikasjon.

Sekvenslikheter

Datamaskinanalyse av MN cDNA-sekvensen ble foretatt ved å benytte DNASIS og PROSIS (Pharmacia Software pakker). GenBank, EMBL, Protein Identification Resource og SWISS-PROT databaser ble gjennomført for alle mulige sekvenslikheter. I tillegg ble det foretatt et søk med hensyn til proteiner som hadde sekvenslikheter tilfelles med MN, i MIPS databank med FastA-programmet [Pearson & Lipman, **PNAS**, (USA), **85**:2444 (1988)].

Det proteoglykan-lignende domenet [aa 53-111 (SEKV. ID. NR:50)], som er mellom signalpeptidet og CA-domenet, oppviser signifikant homologi (38% identitet og 44% positivitet) med et keratansulfat-tilkoblingsdomene i et humant stort aggregierende proteoglykan-aggrekan [Doege *et al.*, **J. Biol. Chem.**, **266**:894-902 (1991)].

CA-domenet [aa 135-391 (SEKV. ID. NR:51)] er utbredt over 265 aa og viser 38,9% aminosyreidentitet med det humane CA VI isoenzym [Aldred *et al.*,

Biochemistry, 30:569-575 (1991)]. Homologien mellom MN/CA IX og andre isoenzymer er som følger: 35,2% med CA II i en 261 aa overlapping [Montgomery *et al.*, **Nucl. Acids Res.**, 15:4687 (1987)], 31,8% med CA I i en 261 aa overlapping [Barlow *et al.*, **Nucl. Acids Res.**, 15:2386 (1987)], 31,6% med CA IV i en 266 aa overlapping [Okuyama *et al.*, PNAS (USA) 89:1315-1319 (1992)] og 30,5% med CA III i en 259 aa overlapping (Lloyd *et al.*, **Genes. Dev.**, 1:594-602 (1987)).

I tillegg til CA-domenet har MN/CA IX fått både N-terminale og C-terminale forlengelser som ikke er relatert til de øvrige CA-isoenzymene. Aminosyresekvensen av den C-terminale del, bestående av den transmembrane forankring og den intracytoplasmatiske hale, viser ingen signifikant homologi med noen kjent protein-sekvens.

MN-genet ble klart funnet å være en ny sekvens avledet fra det humane genom. Den totale sekvenshomologi mellom cDNA MN-sekvensen og cDNA-sekvenser som koder for forskjellige CA-isoenzymene ligger innen homologiområdet 48-50% som innen fagområdet ansees som lavt. MN cDNA-sekvensen er derfor ikke nært beslektet med noen CA cDNA-sekvens.

Bare meget nær beslektede nt-sekvenser som oppviser en homologi på minst 80-90% ville hybridisere med hverandre under stringente betingelser. En sekvens-sammenligning av MN cDNA-sekvensen vist i Figur 1, og en tilsvarende cDNA av den humane karboanhydrase II (CA II) viste at det ikke er noen identitetsstreknings mellom de to sekvensene som ville være tilstrekkelig lang til å la et segment av CA II cDNA-sekvensen som har 25 eller flere nukleotider, til å hybridisere under stringente hybridiseringsbetingelser til MN cDNA eller *vice versa*.

Et søk med henblikk på nt-sekvenser relatert til MN-genet i EMBL data-biblioteket avdekket ingen spesifikk homologi bortsett fra 6 komplette og 2 partielle Alu-type repetisjoner med homologi til Alu-sekvenser varierende fra 69,8% til 91% [Jurka & Milosavljevic, **J. Mol. Evol.**, 32:105-121 (1991)]. Også en 222 bp sekvens proksimalt til 5'-enden av den genomiske region er vist å være nært homolog med en region til HERV-K LTR.

Generelt kan nukleotidsekvenser som ikke er i Alu eller LTR-lignende regioner, på fortrinnsvis 25 baser eller mer, eller enda mer foretrukket på 50 baser eller mer, testes rutinemessig og underkastes screening, og er funnet å hybridisere under stringente betingelser til kun MN-nukleotidsekvenser. Dessuten er ikke alle

homologier innen de Alu-lignende MN genomiske sekvenser så nær Alu-repetisjoner at de gir et hybridiseringssignal under stringente hybridiseringsbetingelser. Prosent-andelen homologi mellom MN Alu-lignende regioner og en standard Alu-J-sekvens er som følger:

	Homologiregion innen MN genomisk sekvens [SEKV. ID. NR:5; Figur 2A-F]	SEKV. ID. NR:	% homologi med hel Alu-I sekvens
5	921-1212	59	89,1%
10	2370 - 2631	60	78,6%
	4587 - 4880	61	90,1%
	6463 - 6738	62	85,4%
	7651 - 7939	63	91,0%
	9020 - 9317	64	69,8%
15			% homologi med halvparten av Alu-J-sekvensen
	8301 - 8405	65	88,8%
	10040 - 10122	66	73,2%

20

MN-proteiner og/eller polypeptider

Uttrykket «MN-proteiner og/eller polypeptider» (MN-proteiner/polypeptider) er her ment å bety proteiner og/eller polypeptider som kodes av et MN-gen eller fragment derav. Et eksempel og et foretrukket MN-protein i henhold til denne oppfinnelse har den deduserte aminosyresekvens vist i Figur 1. Foretrukne MN-proteiner/polypeptider er de proteiner og/eller polypeptider som har vesentlig homologi med MN-proteinet vist i Figur 1. For eksempel er slike hovedsakelig homologe MN-proteiner/polypeptider de som er reaktive med de MN-spesifikke antistoffene ifølge denne oppfinnelse, fortrinnsvis Mab M75, MN12, MN9 og MN7 eller ekvivalenter derav.

30

«Polypeptid» eller «peptid» er en kjede av aminosyrer kovalent bundet av peptidbindinger og ansees her å være sammensatt av 50 eller færre aminosyrer. Et

«protein» er her ment å være et polypeptid sammensatt av mer enn 50 aminosyrer. Betegnelsen polypeptid omfatter betegnelse peptid og oligopeptid.

MN-proteiner oppviser flere interessante trekk: cellemembranlokalisering, celletetthet-avhengig ekspresjon i HeLa celler, korrelasjon med den tumorigene fenotype av HeLa x fibroblast somatiske cellehybrider og ekspresjon i flere humane karsinomer blant andre vev. MN-protein kan finnes direkte i tumorvevsnitt, men i alminnelighet ikke i tilsvarende normale vev (unntak angitt nedenfor som i normal gastrisk mukosa og galleblærevev). MN uttrykkes også iblant i morfologisk normalt utseende områder av vevsprøver som oppviser dysplasi og/eller malignitet. Alt i alt tyder disse trekk på en mulig medvirkning av MN i reguleringen av celleproliferasjon, differensiering og/eller transformasjon.

Det vil være klart at et protein eller polypeptid fremstillet av en neoplastisk celle *in vivo*, vil kunne være endret i sekvens fra det som produseres av en tumorcelle i cellekultur eller av en transformert celle. MN-proteiner og/eller proteiner som har avvikende aminosyresekvenser, inklusivt uten begrensning, aminosyre-substitusjoner, ekstensjoner, delesjoner, trunkeringer og kombinasjoner derav, faller således innenfor rammen av denne oppfinnelse. Det vil også være klart at et protein som forefinnes i kroppsvæsker er utsatt for nedbrytende prosesser, som f.eks. proteolytiske prosesser. MN-proteiner som er betydelig trunkert og MN-polypeptider kan således finnes i kroppsvæsker, så som serum. Betegnelsen «MN-antigen» benyttes her om MN-proteiner og/eller polypeptider.

Det vil videre være klart at aminosyresekvensen av MN-proteiner og polypeptider kan modifiseres ved genetiske teknikker. En eller flere aminosyrer kan være deletert eller substituert. Slike aminosyreendringer behøver ikke forårsake noen målbar endring i den biologiske aktivitet av proteinet eller polypeptidet og kan resultere i proteiner eller polypeptider som ligger innenfor rammen av oppfinnelsen, så vel som MN-muteiner.

MN-proteinene og polypeptidene ifølge oppfinnelsen kan fremstilles på en rekke måter i henhold til foreliggende oppfinnelse, for eksempel rekombinant, syntetisk eller på biologisk måte, dvs. ved å spalte lengre proteiner og polypeptider enzymatisk og/eller kjemisk. En foretrukket fremgangsmåte for fremstilling av MN-proteiner er ad rekombinant vei. Særlig foretrukne fremgangsmåter for rekombinant fremstilling av MN-proteiner er beskrevet nedenfor for GST-MN-, MN 20-19-, MN-Fc- og MN-PA-proteiner.

Rekombinant produksjon av MN-proteiner og polypeptider

En representativ fremgangsmåte for å fremstille MN-proteinene vist i Figur 1, eller fragmenter derav, vil være å innføre den fulle lengde eller et passende fragment av MN cDNA i en passende ekspresjonsvektor slik som eksemplifisert nedenfor. I

5 Zavada *et al.*, WO 93/18152, ovenfor, beskrives produksjon av et fusjonsprotein GEX-3X-MN (nå betegnet GST-MN) ved å benytte den partielle cDNA-klon (beskrevet ovenfor) i vektoren pGEX-3X (Pharmacia). Ikke-glykosylert GST-MN (MN-fusjonsproteinet MN glutation-S-transferase) fra XL1-Blue-celler.

10 Zavada *et al.*, WO 95/34650 beskriver den rekombinante fremstilling av både et glykosylert MN-protein uttrykt fra insektceller, og et ikke-glykosylert MN-protein uttrykt fra *E. coli*, ved å benytte ekspresjonsplasmidet pEt-22b [Novagen Inc.; Madison, WI (USA)]. Rekombinante baculovirus-ekspresjonsvektorer ble benyttet til å infisere insektceller. Det glykosylerte MN 20-19-protein ble fremstillet rekombinant i baculovirus-infiserte sf9-celler [Clontech; Palo Alto, CA (USA)]. MN 20-19-proteinet 15 mangler det antatte signalpeptid (aa 1-37) av SEKV. ID. NR:6 (Figur 1), har et metionin (met) i N-enden for ekspresjon, og et Leu-Glu-His-His-His-His-His-[SEKV. ID. NR:22] addert til C-enden, for rensing.

For å innføre delen av den MN-kodende sekvens for GST-MN-fusjonsproteinet inn i alternative ekspresjonssystemer, ble det konstruert et sett av 20 primere for PCR. Primerne ble konstruert for å gi restriksjonssteder i hver ende av den kodende sekvens, så vel som start- og stopp-kodoner i ramme. De sekvensene av primerne som indikerer restriksjonsenzym-spaltningsseter og ekspresjonskjennetegn, er vist nedenfor.

Primer nummer 20:N-terminal

25 -Translasjonsstart
5'-GTCGCTAGCTCCATGGGTCATATGCAGAGGTTGCCCCGGATGCAG 3'
NheI NcoI NdeI -MN cDNA# 1 [SEKV. ID. NR:17]

Primer nummer 19: C-terminal

30 -Translasjonsstopp
5' GAAGATCTCTTACTCGAGCATTCTCCAAGATCCAGCCTCTAGG 3'
BglII XhoI -MN cDNA [SEKV. ID. NR:18]

Primerne SEKV. ID. NR:17 og 18 ble benyttet til å amplifisere den MN-kodende sekvens som forekom i GEX-3X-MN-vektoren ved å benytte standard PCR-teknikker.

Det resulterende PCR-produkt (betegnet MN 20-19) ble underkastet elektroforese på en 0,5% agarose/1X TBE gel; 1,3 kb båndet ble snittet ut og DNA isolert ved å benytte Gene Clean II settet i henhold til produsentens instruksjoner [Bio101; LaJolla, CA (USA)].

5

Påvisning av MN-proteinpartnere

Et søk med henblikk på proteiner som interagerer med MN ble igangsatt ved å benytte ekspresjonskloning av de korresponderende cDNA og et MN-Fc fusjonsprotein som probe. Det kimære MN-Fc cDNA ble konstruert i pSG5C-vektor ved
 10 substitusjon av MN cDNA-sekvenser som kodet både den transmembrane forankring og den intracellulære hale av MN-proteinet med den cDNA som koder for Fc-fragmentet av mus IgG. Denne Fc-fragment cDNA ble fremstillet ved RT-PCR fra det muse hybridom som produserer IgG2a antistoff.

Den kimære MN-Fc cDNA ble uttrykt ved transient transfeksjon i COS-celler.
 15 COS-cellene ble transfektert ved å benytte leptofeksjon. Rekombinant MN-Fc-protein ble frigitt til TC-medium av de transfekterte cellene (som følge av den manglende transmembrane region), renses ved affinitetskromatografi på en Protein A Sepharose og benyttet for videre forsøk.

Proteinekstrakter fra placebo-transfekteerte celler og cellene transfektert med
 20 pSG5C-MN-Fc ble analysert ved immunblotting under bruk av M75 MAb, Sw α M-Px og ECL deteksjon [ECL® - forsterket kjemiluminescenssystem for å påvise fosforylerte tyrosinrester; Amersham; Arlington, Hts., IL (USA)]. Størrelsen av MN-Fc-proteinet uttrykt fra pSG5C-vektoren tilsvarer dens datamaskinberegnete molekulvekt.

25 ³⁵S-merket MN-Fc-protein ble benyttet i bestemmelser av celleoverflate-binding. Det viste seg å binde til flere pattedyrceller, f.eks. HeLa, Raji, COS, QT35, BL3. Lignende resultater ble oppnådd i celleadhesjonstester ved å benytte MN-Fc-protein avsatt på petriskåler. Disse testene viste at KATO III human mavesekk adenokarsinomcellelinjen mangler evnen til interaksjon med MN-Fc-protein. Dette
 30 funn satte oss i stand til å benytte KATO III celler for ekspresjonskloning og screening av den cDNA som koder for MN-bindingsprotein.

cDNA ekspresjonsbiblioteket i pBK-CMV-vektoren ble fremstillet fra tette HeLa-celler og benyttet for transfeksjon av KATO III celler. Ved den første screening-

runde ble KATO III celler transfektert ved elektroporering. Etter to dagers inkubering fikk de ligand-uttrykkende cellene anledning til å binde til MN-Fc-protein, deretter til protein A konjugert med biotin og ble tilslutt selektert ved å trekkes ut med Streptavidin-dekkede magnetiske kuler. Plasmid DNA ble ekstrahert fra de selekterte cellene og transformert til *E. coli*. Individuelle *E. coli* kolonier ble plukket og ansamlinger av 8-10 kloner fremstillet. Plasmid DNA fra samlingene ble isolert og benyttet i den andre screeningrunde.

I den andre screeningrunde ble KATO III celler transfektert ved hjelp av DEAE-dekstranmetoden. For å identifisere den samling som inneholdt cDNA for MN-bindingsprotein ble det benyttet en ELISA-metode basert på bindingen av MN-Fc til de transfekterte cellene, og påvisning ved å benytte peroksydasemerket protein A. Samlinger ble selektert ved hjelp av evnen til å binde MN-Fc.

I den tredje screeningrunde ble plasmid DNA isolert fra individuelle bakteriekolonier av selekterte ansamlinger transfektert til KATO III celler. De transfekterte cellene ble underkastet binding med MN-Fc og påvisning med Protein A som tidligere. Et slikt eksempel på screening forventes å identifisere en klon som inneholder den cDNA som koder for den antatte MN-proteinpartner. Denne klon vil deretter kunne sekvenseres og ekspresjonsproduktet bekreftes som bindende til MN-protein ved celleadhesjonsbestemmelse. (Far-Western blotting, kopresipitasjon, etc.). Det vil deretter kunne fremstilles hybridomer som produserer Mab til ekspresjonsproduktet, hvilket ville muliggjøre analyse av de biologiske karakteristika til proteinpartneren av MN.

Fremstilling av MN-spesifikke antistoffer

Betegnelsen «antistoffer» er her definert slik at ikke bare hele antistoffer, men også biologisk aktive fragmenter av antistoffer, fortrinnsvis fragmenter inneholdende de antigen-bindende regioner, inkluderes. Videre er bispesifikke antistoffer som er spesifikke for MN-protein og annet vevsspesifikt antigen, inkludert i definisjonen av antistoffer.

Zavada *et al.*, WO 93/18152 og WO 95/34650 beskriver i detalj fremgangsmåter for fremstilling av MN-spesifikke antistoffer og detaljerte trinn for fremstilling av representative MN-spesifikke antistoffer som M75, MN7, MN9 og MN12 monoklonale antistoffer. Foretrukne MN-antigenepitoper omfatter: aa 62-67 (SEKV. ID. NR:10); aa 61-66, aa 79-84, aa 85-90 og aa 91-96 (SEKV. ID. NR:98); aa 62-65, aa 80-83,

aa 86-89 og aa 92-95 (SEKV. ID. NR:99); aa 62-66, aa 80-84, aa 86-90 og aa 92-96 (SEKV. ID. NR:100); aa 63-68 (SEKV. ID. NR:101); aa 62-68 (SEKV. ID. NR:102); aa 82-87 og aa 88-93 (SEKV. ID. NR:103); aa 55-60 (SEKV. ID. NR:11); aa 127-147 (SEKV. ID. NR:12); aa 36-51 (SEKV. ID. NR:13); aa 68-91 (SEKV. ID. NR:14);
 5 aa 279-291 (SEKV. ID. NR:15); og aa 435-450 (SEKV. ID. NR:16). Eksempel 2 gir ytterligere beskrivelse angående foretrukne MN-antigenepitoper.

Bispesifikke antistoffer. Bispesifikke antistoffer kan fremstilles ved kjemisk å koble to antistoffer av den ønskede spesifisitet. Bispesifikke MAb kan fortrinnsvis utvikles ved somatisk hybridisering av 2 hybridomer. Bispesifikke MAb for målsøking
 10 av MN-protein og et annet antigen kan fremstilles ved å fusjonere et hybridom som danner MN-spesifikke MAb med et hybridom som danner MAb som er spesifikke for et annet antigen. For eksempel vil en celle (et kvadrom), dannet ved fusjon av et hybridom som danner et MN-spesifikt MAb og et hybridom som danner et anti-cytotoksisk celle-antistoff, føre til hybrid-antistoff som har spesifisiteten av de opp-
 15 havelige antistoffene. [Se f.eks. *Immol.- Rev.*, (1979); **Cold Spring Harbor Symposium Quant. Biol.**, 41:793 (1977); van Dijk *et al.*, *Int. J. Cancer*, 43:344-349 (1989)]. Et hybridom som danner et MN-spesifikt MAb kan deretter fusjoneres med et hybridom som for eksempel danner et anti-T3 antistoff for å gi en cellelinje som fører til et MN/T3-bispesifikt antistoff som kan styre cytotoksiske T-celler til MN-
 20 uttrykkende tumorceller.

For terapeutiske anvendelser og/eller avbildning kan det være å foretrekke at antistoffene er biologisk aktive antistoff-fragmenter, fortrinnsvis genetisk manipulerede fragmenter, mer foretrukket genetisk manipulerede fragmenter fra fra V_H - og/eller V_L -regionene, og enda mer foretrukket at de omfatter de hypervariable regioner derav.
 25 For enkelte terapeutiske anvendelser vil imidlertid bispesifikke antistoffer som mål-søker MN-protein og cytotoksiske celler, være å foretrekke.

Epitoper

Affiniteten av et MAb til peptider som inneholder en epitop avhenger av
 30 sammenhengen, f.eks. hvorvidt peptidet er en kort sekvens (4-6 aa), eller hvorvidt et slikt kort peptid er flankert av lengre aa-sekvenser på den ene eller begge sider, eller hvorvidt peptidene ved testing med henblikk på epitop, er i løsning eller immobilisert på en overflate. Fagmannen ville derfor ha forventet at de her beskrevne

representative epitoper for de MN-spesifikke MAb ville variere med hensyn til anvendelsen av disse MAb.

Betegnelsen «som tilsvarer en epitop av et MN-protein/polypeptid» innbefatter den praktiske mulighet at aminosyresekvensvariasjoner i et naturlig forekommende protein eller polypeptid i enkelte tilfeller kan være antigene og gi beskyttende immunitet mot neoplastisk sykdom og/eller antitumorigene effekter. Mulige sekvensvariasjoner innbefatter, uten begrensning, aminosyresubstitusjoner, ekstensjoner, delesjoner, trunkeringer, interpolasjoner og kombinasjoner derav. Slike variasjoner faller innenfor oppfinnelsens ramme, forutsatt at proteinet eller polypeptidet som inneholder disse, er immunogent og at antistoffer som utløses av et slikt polypeptid eller protein, i tilstrekkelig grad kryssreagerer med naturlig forekommende MN-proteiner og polypeptider til å gi beskyttende immunitet og/eller antitumorigen aktivitet ved administrering som en vaksine.

Epitop for M75 MAb

M75 epitopen ansees å forekomme i minst to kopier innenfor 6X tandem repetisjonen av 6 aminosyrer [aa 61-96 (SEKV. ID. NR:97)] i MN-proteinets proteoglykandomene. Eksempler på peptider som representerer epitopen, kan avhengig av sammenhengen, innbefatte følgende peptider fra vedkommende tandemrepetisjon: EEDLPS (SEKV. ID. NR:10; aa 62-67); GEEDLP (SEKV. ID. NR:98; aa 61-66; aa 79-84; aa 85-90; aa 91-96); EEDL (SEKV. ID. NR:99; aa 62-65; aa 80-83; aa 86-89; aa 92-95); EEDLP (SEKV. ID. NR:100; aa 62-66; aa 80-84; aa 86-90; aa 92-96); EDLPSE (SEKV. ID. NR:101; aa 63-68); EEDLPSE (SEKV. ID. NR:102; aa 62-68); og DLPGEE (SEKV. ID. NR:103; aa 82-87, aa 88-93).

Tre syntetiske peptider fra den deduserte aa-sekvens for EC-domenet av MN-protein vist i Figur 1, ble fremstillet. Disse syntetiske peptidene er representert ved aa 51-72 (SEKV. ID. NR:104); aa 61-85 (SEKV. ID. NR:105) og aa 75-98 (SEKV. ID. NR:106). Disse syntetiske peptidene inneholder alle motivet EEDLP (SEKV. ID. NR:100) og viste seg å være reaktive med M75 MAb.

Andre epitoper

Mab MN9. Monoklonalt antistoff MN9 (Mab MN9) reagerer med den samme epitop som Mab M75, som beskrevet ovenfor. Som Mab M75 gjenkjenner Mab MN9 både GST-MN-fusjonsproteinet og nativt MN-protein like godt.

Mab som tilsvarer Mab MN9, kan fremstilles reproducerbart ved screening av en rekke Mab fremstillet mot et MN-protein/polypeptid, så som GST-MN-fusjonsproteinet, mot de peptider som representerer epitopen for Mab M75 og MN9. Som et alternativ kan Novatope-systemet [Novagen] eller konkurranse med det deponerte Mab M75 benyttes til selektering av Mab som er sammenlignbare med Mab M75 og MN9.

Mab MN12. Monoklonalt antistoff MN12 (Mab MN12) produseres av det lymfocytiske musehybridom MN 12.2.2 som ble deponert under ATCC HB 11647. Antistoffer som tilsvarer Mab MN12 kan også fremstilles, analogt med den metode som ovenfor er skissert for Mab MN9, ved screening av en rekke antistoffer fremstillet mot et MN-protein/polypeptid, mot det peptid som representerer epitopen for Mab MN12. Dette peptid er aa 55 - aa 60 i Figur 1 [SEKV. ID. NR:11]. Novatope-systemet ville også kunne benyttes for å finne antistoffer som er spesifikke for nevnte epitop.

Mab MN7. Monoklonalt antistoff MN7 (Mab MN7) ble selektert fra Mab fremstillet mot ikke-glykosylert GST-MN slik som beskrevet ovenfor. Det gjenkjenner epitopen representert ved aminosyresekvensen fra aa 127 til aa 147 [SEKV. ID. NR:12] i Figur 1 av MN-proteinet. Analogt med fremgangsmåtene beskrevet ovenfor for Mab MN9 og MN12, kan Mab tilsvarende Mab MN7 fremstilles ved selektering av Mab fremstillet mot et MN-protein/polypeptid som er reaktivt med det peptid som har SEK. ID. NR:12, eller ved hjelp av de angitte alternative måtene.

MN-spesifikke antistoffer - Målrettet tumordreping via intracellulær ekspresjon av MN-spesifikke antistoffer for å blokkere transport av MN-protein til celleoverflate

Genet som koder for antistoffer kan manipuleres slik at det antigen-bindende domenet kan uttrykkes intracellulært. Slike «intrastoffer» som målrettes til lumenet av det endoplasmatiske retikulum, utgjør en enkel og effektiv mekanisme for inhibering av transporten av plasmamembranproteiner til celleoverflaten. [Marasco, W.A., «Review - Intrabodies: turning the humoral immune system outside in or intracellular

immunization», **Gene Therapy**, 4:11-15 (1997); Chen *et al.*, «Intracellular antibodies as a new class of therapeutic molecules for gene therapy», **Hum. Gene Ther.**, 5(5):595-601 (1994); Mhashilkar *et al.*, **EMBO J.**, 14:1542-1551 (1995); Mhashilkar *et al.*, **J. Virol.**, 71:6486-6494 (1997); Marasco (red.), **Intrabodies: Basic Research and Clinical Gene Therapy Applications**, (Springer Life Sciences 1998; ISBN 3-540-64151-3) (sammenfatter prekliniske undersøkelser fra laboratorier over hele verden som har benyttet «intrastoffer»); Zanetti & Capra (red.), «Intrabodies: From Antibody Genes to Intracellular Communication,» **The Antibodies: Volume 4**, [Harwood Academic Publishers; ISBN 90-5702-559-0 (des. 1997)]; Jones & Marasco, **Advanced Drug Delivery Reviews**, 31(1-2):153-170 (1998); Pumphrey & Marasco, **Biodrugs**, 9(3):179-185 (1998); Dachs *et al.*, **Oncology Res.**, 9(6-7):313-325 (1997); Rondon & Marasco, **Ann. Rev. Microbiol.**, 51:257-283 (1997)]; Marasco, W.A., **Immunotechnology**, 1(1):1-19 (1995); og Richardson & Marasco, **Trend in Biotechnology**, 13(8):306-310 (1995)].

MN-spesifikke intrastoffer kan forhindre modningen og transporten av MN-protein til celleoverflaten og derved hindre MN-proteinet i å virke i en onkogen prosess. Antistoffer rettet mot MN's EC-, TM- eller IC-domener kan være egnet i denne sammenheng. MN-protein ansees å mediere signaltransduksjon ved overføringssignaler fra EC-domenet til IC-halen og deretter ved å assosiere med andre intracellulære proteiner inne i cellens indre. MN-spesifikke intrastoffer kunne avbryte denne assosiasjon og forstyrre denne MN-funksjon.

Inaktivering av MN-proteinets funksjon kunne resultere i reversering av tumorceller til en ikke-transformert fenotype [Marasco *et al.*, (1997), *supra*]. Antisense ekspresjon av MN cDNA i livmorhals-karsinomceller har som demonstrert her, vist at tap av MN-protein har ført til vekstundertrykkelse av de transfektete cellene. Det forventes likeledes at inhibering av MN-proteintransport til celleoverflaten ville ha lignende virkninger. Kloning og intracellulær ekspresjon av M75 MAb's variable region bør undersøkes for å bekrefte denne forventning.

Fortrinnsvis er de intracellulært produserte MN-spesifikke antistoffene enkeltkjede antistoffer, nærmere bestemt enkeltkjede variable region-fragmenter eller sFv, hvor de tunge og lette kjeders variable domene syntetiseres som et enkelt polypeptid og separeres av et fleksibelt linkerpeptid, fortrinnsvis (Gly₄-Ser)₃ [SEKV. ID. NR:116].

MN-spesifikke intracellulært produserte antistoffer kan benyttes terapeutisk til å behandle preneoplastisk/neoplastisk sykdom ved å transfektore preneoplastiske/-neoplastiske celler som unormalt uttrykker MN-protein, med en vektor som omfatter en nukleinsyre som koder for MN-spesifikke antistoff-fragmenter av variabel region, operativt koblet til en ekspresjonskontrollsekvens. Fortrinnsvis omfatter nevnte ekspresjonskontrollsekvens MN-genpromoteren.

Antistoff-mediert genoverføring ved bruk av MN-spesifikke antistoffer eller peptider for målsøking av MN-uttrykkende tumorceller

Et MN-spesifikt antistoff eller peptid som kovalent er koblet til polylysin, et polykation som er i stand til å kompaktere DNA og nøytralisere dets negative ladninger, ville forventes å effektivt avgis biologisk aktivt DNA inn i en MN-uttrykkende tumorcelle. Dersom den pakke DNA inneholder HSVtk-genet under kontroll av MN-promoteren, ville systemet ha dobbelt spesifisitet for gjenkjenning og ekspresjon bare i MN-uttrykkende tumorceller. Den pakke DNA kunne også kode for cytokiner for å indusere CTL-aktivitet eller for andre biologisk aktive molekyler. M75 MAb (eller for eksempel, som et enkeltkjedet antistoff, eller som dets variable region) utgjør eksempel på et slikt MN-spesifikt antistoff.

De følgende eksempler har kun det formål å illustrere og er ikke på noen måte ment å begrense oppfinnelsen.

Eksempel 1

Transient transformasjon av pattedyrceller med MN-protein

Dette eksempel (1) undersøker de biologiske konsekvenser ved transfektering av humane celler eller museceller med MN-cDNA innskutt i ekspresjonsvektorer, hovedsakelig ut fra synspunktet involvering av MN-protein i onkogenese; (2) bestemme om MN-protein utøver karboanhydraseaktivitet, og hvorvidt slik aktivitet er relevant for morfologisk transformasjon av celler; og (3) teste hvorvidt MN-proteinet er et celledhesjonsmolekyl (CAM).

Oversikt

Metoder: MN-cDNA ble innskutt i 3 ekspresjonsvektorer og benyttet for transfektering av humanceller eller museceller. MN-protein ble påvist ved western blotting, radioimmunbestemmelser eller immunperoksydasefarging. For alle testene ble det MN-spesifikke monoklonale antistoff M75 (MAb M75) benyttet. Karboanhydraseaktivitet ble bestemt ved surgjøringshastigheten av karbonatbuffer i CO₂-atmosfære.

Resultater: (1) celler (humane CGL-1-celler og museceller NIH3T3) transfektert med MN-cDNA, oppviste morfologisk transformasjon, men vendte tilbake til normal fenotype etter 4-5 uker. (2) Denne tilbakevending skyldtes ikke tap, nøytralisering (silencing) eller mutasjon av MN-innskudd. (3) MN-protein har enzymaktivitet av en karboanhydrase som kan inhiberes med acetazolamid, men inhiberingen av karboanhydrase-enzymaktiviteten påvirket ikke transformasjon. (4) MN-proteinet er et adhesjonsprotein som er involvert i celle-til-celle kontakter.

Bakgrunn

Dette eksempel vedrører transformasjon av pattedyrceller med MN-cDNA innskutt i ekspresjonsvektorer avledet fra retrovirus. Slike vektorer er egnet for effektiv og stabil integrering i cellulært DNA og for kontinuerlig ekspresjon av MN-protein. Celler transfektert med disse konstruksjonene oppviste morfologisk transformasjon, men etter en tid vendte de tilbake til normal fenotype.

Sulfonamider, inklusivt acetazolamid, er meget potente inhibitorer av kjente karboanhydraser [Maren & Ellison, **Mol. Pharmacol.** 3:503-508 (1967)]. Acetazolamid ble undersøkt for å bestemme om det også inhiberte MN-karboanhydrasen, og i så fall, hvorvidt inhiberingen av enzymet påvirket celletransformasjon.

Det er grunn til å tro at MN-proteinet kunne være involvert i direkte celle-til-celle interaksjoner: A) forutgående observasjoner tyder på en funksjonell likhet mellom MN-protein og overflate-glykoproteiner av kappe-holdige virus som medierer virus-adsorpsjon til celleoverflatereseptorer, og MN deltok i dannelsen av fenotypisk blandede virioner av vesikulær stomatitt-virus. B) Induserbarhet av MN-protein-ekspresjon ved dyrkning av HeLa-celler i tettpakkede monolag tydet på at de kan være involvert i direkte interaksjoner mellom celler. C) Endelig er det en strukturell

likhet mellom MN-proteinet og reseptor-tyrosinfosfatase β , som også inneholder proteoglykan- og karboanhydrasedomener. Disse domenene medierer direkte kontakter mellom celler i det utviklende nervesystem [Peles *et al.*, **Cell**, **82**:251-260 (1995)]. MN-proteinet ble derfor testet for å se om det bandt til celleoverflate-
 5 reseptorer, og resultatet var klart positivt for at det gjorde det.

Materialer og metoder

Cellelinjer

Celler benyttet i dette eksempel var: CGL1 og CGL3 - henholdsvis ikke-
 10 tumorigene og tumorigene HeLa x fibroblasthybrider [Stanbridge *et al.*, **Somat. Cell Genet.**, **7**:699-712 (1981)], musecellelinje NIH3T3, HeLa-celler og ape Vero-celler. NIH3T3-cellene ble utsådd i meget lav tetthet for å oppnå kolonier som startet fra enkeltceller. De mest normalt utseende kolonier, betegnet subklon 2, ble plukket for bruk i de forsøk som er omtalt i dette eksempel.

Ekspresjonsvektorer

Full-lengde MN cDNA ble oppnådd fra pBluescript subklon [Pastorek *et al.*, **Oncogene**, **9**:2877-2888 (1994)]. For å fjerne 5'- og 3'-ikke-kodende sekvenser, som kunne redusere påfølgende genekspresjon, ble det foretatt en polymerasekjede-
 20 reaksjon (PCR). 5'-primeren TAGACAGATCTACGATGGCTCCCCTGTGCCCCAG [SEKV. ID. NR:88] omfatter et translasjonsstartsete og *Bgl*II kloningsete, og 3'-primeren ATTCCTCTAGACAGTTACCGGCTCCCCTCAGAT [SEKV. ID. NR:89] omfatter et stoppkodon og *Xba*I kloningssete. Full-lengde MN-cDNA ble benyttet som templat og
 25 Pfu DNA polymerase [Stratagene; LaJolla, CA (USA)] i reaksjonen.

PCR-produktet ble sekvensert og funnet å være identisk med templat og bar ingen mutasjoner. PCR-produktet som utelukkende inneholdt den MN-kodende sekvens, ble innskutt i tre vektorer. 1. pMAMneo [Clontech; Palo Alto, CA (USA)] plasmidet som muliggjorde deksametason-induserbar ekspresjon styrt av MMTV-
 30 LTR-promotoren (Long Terminal Repeat) og som inneholdt et neo-gen for seleksjon av transformanter i medier supplert med Geneticin (G418) antibiotika. 2. Retroviral ekspresjonsvektor pGD [Daley *et al.*, **Science**, **247**:824-829 (1990); stillet til rådighet av Prof. David Baltimore, New York-Cambridge]] som inneholder MLV-LRT-promoter

og neo-genet for G418 antibiotika-seleksjon. 3. Vaccinia-virus ekspresjonsvektor pSC11 [Chakrabarti *et al.*, **Mol. Cell. Biol.**, 5:3403-3409 (1985)]. Transfeksjon ble foretatt via et kalsiumfosfat-presipitat i henhold til Sambrook *et al.*, (red.). **Molecular cloning. A laboratory manual**, 2. Utg., Cold Spring Harbor Laboratory Press (1989).

5 Vaccinia-virus stamme Praha klon 13 ble benyttet som parentalt virus [Kutinova *et al.*, **Vaccine**, 13:487-493 (1995)]. Vaccinia-virus-rekombinant ble fremstillet etter en standardprosedyre [Perkus *et al.*, **Virology**, 152:285-297 (1986)]. Rekombinante virus ble selektert og plakk-renset to ganger i rotte tymidinkinase-manglende RAT2-celler [Topp, W.C., **Virology**, 113:408-411 (1981)] i nærvær av 5'-
10 bromdeoksyuridin (100 µg/mL). Blå plakk ble påvist ved overskikting med agar inneholdende 5-brom-4-klor-3-indolyl-β-D-galaktopyranosid (X-Gal) (200 µg/mL).

CA-bestemmelse

Karboanhydraseaktivitet målt ved hjelp av en mikrometode [Brion *et al.*, **Anal. Biochem.**, 175:289-297 (1988)]. Prinsippet er at hastigheten av reaksjonen
15 $\text{CO}_2 + \text{H}_2\text{O} \rightarrow \text{H}_2\text{CO}_3$ måles ved den tid som fordres for surgjøring av karbonatbuffer, påvist med fenolrødt som pH-indikator. Denne reaksjonen skjer endog uten enzymet, med t_0 = kontrolltid (denne ble satt til 60 sekunder). Karboanhydrase reduserer surgjøringstiden til t ; en enhet av enzymaktiviteten reduserer tiden til halvparten av
20 kontrolltiden: $t/t_0 = 1/2$.

For dette forsøk ble MN-protein immunpresipitert med Mab M75 fra RIPA-buffer- (1% Triton X-100, 0,1% deoksyholat, 1 mM fenylmetylsulfonylfluorid og 200 trypsin-inhiberende enheter/mL Trasylol i PBS, pH 7,2) ekstrakt av Vero-celler infisert med vaccinia-MN-konstruksjon etter at cellene utviklet cytopatisk effekt, eller
25 med «tomme» vaccinia som kontroll. MN + antistoffkomplekset ble deretter adsorbert til protein A - *Staphylococcus aureus*-celler [Kessler, S.W., **I. Immunol.**, 115:1617-1624 (1975)] og rensset 2x med PBS og 2x med 1 mM karbonatbuffer, pH 8,0. Presipitatet ble resuspendert i den samme buffer og tilsatt til reaksjonsblanding. Acetazolamid (Sigma) ble testet med henblikk på inhibering av
30 karboanhydrase [Maren & Ellison, *supra*]. I ekstrakter av infiserte celler benyttet for immunpresipitasjon, ble konsentrasjonen av totale proteiner bestemt etter Lowry-metoden [Lowry *et al.*, **J. Biol. Chem.**, 193:265-275 (1951)] og den for MN-protein

ved en kompetitiv radioimmunassay som beskrevet i Zavada *et al.*, *Int. J. Cancer*, **54**:268-274 (1993).

Western blott

Western blotting og utvikling av blottene ved bruk av ^{125}I -merket M75 og autoradiografi, ble foretatt som tidligere [Pastorekova *et al.*, *Virology*, **187**:620-626 (1992); og Zavada (1993) *supra*].

Adhesjonsbestemmelse

For adhesjonsbestemmelsen [Hoffman, S., «Assays of cell adhesion,» i: **Cell-cell Interactions** (Stevenson *et al.*, red.) s. 1-30 (IRL Press at Oxford University Press; Oxford, N.Y., Tokyo; 1992)], 25 μL alikvoter av MN-protein (affinitetsrenset pGEX-3X MN) [Zavada *et al.*, (1993) *supra*] eller av kontrollproteiner avsatt på bakteriologiske petriskåler med 5 cm diameter, hvor de fikk bindes i to timer ved romtemperatur. Dette førte til sirkulære proteindekkede områder med 4-5 mm diameter. MN-protein ble fortynnet til 10 $\mu\text{g/mL}$ i 50 mM karbonatbuffer, pH 9,2. Lapper av adsorberte kontrollproteiner ble preparert på tilsvarende måte. Disse inkluderte kollagen type I og IV, fibronektin, laminin og gelatin (Sigma-produkter), fortynnet og adsorbent i henhold til produsentens anbefalinger; FCS og BSA ble også inkludert. Etter avsugning av dråpene, ble skålene renset 2x med PBS og mettet i 1 time med DMEM supplert med 5% FCS. Platene ble tilsådd med 5×10^5 celler i 5 mL DMEM + 5% FCS og inkubert over natten ved 37°C. Platene ble renset med PBS og de tilfestede celler fiksert med formaldehyd, etterfiksert med metanol og Giemsa-farvet.

Resultater

1. Transformasjon og reversjon av CGL1-celler transfektert med MN-cDNA

Siden ekspresjonen av MN-protein korrelerte med tumorigeniteten av HeLa x fibroblasthybrider [Zavada *et al.*, (1993) *supra*], ble de ikke-tumorigene hybrid CGL1-cellene undersøkt først. De cellene som var transfektert med pMAM.MN-konstruksjonen, etter seleksjon med Geneticin, dannet kolonier med varierende grad av transformasjon, hvorav enkelte syntes normale. Mens normale CGL1-celler kontakt-inhiberes under vekst i en parallell orientering, dannet de

transformerte cellene meget tette kolonier, som viste tap av kontakt-inhibering. Slike kolonier vokste langsommere enn de originale CGL1.

Etter subkloning utskilte de cellene som var isolert fra transformerte kolonier, revertanter. Reversjonen var en gradvis, trinnvis prosess, hvor kolonier med ulik grad av reversjon forekom. Etter to passeringer fikk hele cellepopulasjonen en morfologi som ikke kunne skjernes fra normal CGL1. Dette skyldtes reversjonen av enkelte celler og den selektive fordel for revertantene, som vokste hurtigere enn de transformerte cellene. På tross av gjentatte forsøk ble det ikke oppnådd selv en eneste stabilt transformert celleklon. Det ble ikke funnet transformerte kolonier i CGL1-celler transfektert med et «tomt» pMAM kontrollplasmid. Vekst av CGL1 + pMAM.MN-revertanter i medier supplert med 5 µg/mL deksametason i 7 dager, øket produksjonen av MN-protein, men morfologien av cellene vendte ikke tilbake til transformert.

2. Befrielse av transformerende MN fra revertantene

Reverseringen av MN-transformerte celler til normal fenotype kunne ha minst fire årsaker: A) tap av MN-innskuddet; B) nøytralisering av MN-innskuddet, f.eks. ved metylering; C) mutering av MN-innskuddet; D) aktivering av et suppressor-gen som koder for et produkt som nøytraliserer transformeringsaktiviteten av MN-protein; E) tap av et MN-bindende protein. For å avgjøre blant disse alternativene ble følgende forsøk utviklet.

MN-cDNA ble innskutt i pGD, en vektor avledet fra muse-leukemivirus - MLV. Det ble derved konstruert mangelfullt virus som inneholdt MN-genet og den selektive markør neo i stedet for gener som koder for virale strukturproteiner. Med denne konstruksjon ble NIH3T3-museceller transfektert. I medier supplert med Geneticin, dannet cellene kolonier med fenotyper varierende fra sterkt transformerte til tilsynelatende normale. Alle de transformerte koloniene og ca. 50% av de normale koloniene uttrykte MN-protein. I motsetning til normale NIH3T3-celler, var transformantene også i stand til å danne kolonier i myk-agar, hvilket reflekterer det tap av forankrings-avhengighet som er karakteristisk for celletransformasjon. Etter passasje vendte cellene isolert fra transformerte kolonier, tilbake til normal morfologi og samtidig mistet de evnen til å danne kolonier i myk-agar, mens de fremdeles uttrykte MN-proteinet. Dette permanente nærvær av MN-protein i revertanter

utelukket alternativer A) og B) ovenfor, dvs. tap eller nøytralisering av MN-genet som årsak til reversjon.

For å velge blant de øvrige 3 alternativene ble revertantene superinfisert med levende, replikasjonskompetent MLV. Dette virus vokser i NIH3T3-celler uten noen morfologiske manifestasjoner og det virker som en «hjelper» for pGD.MN-konstruksjonen. Virusavkom fra MLV-infiserte revertanter representerer et kunstig viruskompleks [pGD.MN + MLV]. Dette består av to typer virioner: en standardtype MLV-partikler og virioner som inneholder pGD.MN-genomet, innkapslet i strukturproteiner frembragt av «hjelper» virus. Dette viruskompleks var infeksiosøst for friske NIH3T3-celler og induserte igjen i disse morfologisk transformasjon og evnen til å danne agarkolonier.

I motsetning til NIH3T3 transfektet med pGD.MN, ble alle koloniene av celler infisert med [pGD.MN + MLV] komplekset, som vokste i nærvær av Geneticin, uniformt transformert og inneholdt MN-proteiner. Transformantene reverterte nok en gang til normal fenotype selv om de fortsatte å produsere infeksiosøst [pGD.MN + MLV] kompleks som induserte transformasjon i friske NIH3T3-celler. Denne syklus av infeksjon-transformasjon-reversjon ble gjentatt tre ganger med det samme resultat. Dette utelukket alternativ C) - mutasjon av MN-cDNA som en årsak til reversjon.

Normale NIH3T3-celler dannet et kontakt-inhibert monolag av flate celler som ikke ble farvet med Mab M75 og immnunperoksydase. Celler infisert med [pGD.MN + MLV] kompleks var klart transformert: de vokste i et kaotisk mønster og oppviste tap av kontakt-inhibering. Noen av cellene viste tegn på apoptose. To passasjer senere reverterte cellepopulasjonen fullstendig til opprinnelig fenotype som resultat av hyppig fremkomst av revertanter og av deres selektive fordeler (hurtigere vekst og en høyere utplatingseffektivitet). Revertantene syntes faktisk å vokse til en noe lavere tetthet enn de opprinnelige NIH3T3-cellene, noe som viser en høyere grad av kontakt-inhibering.

Kontroll NIH3T3-cellene inneholdt ikke noe MN-protein (western blott), mens både transformerte celler og revertanter inneholdt den samme mengde og det samme forhold av 54 og 58 kDa MN-proteinbånd. I en ikke-reduserende gel, forelå MN-proteinet i form av oligomerer på 153 kDa. Ifølge kompetitiv RIA ble ca. 40 ng MN/mg totalt protein konsekvent funnet i både transformerte celler og revertanter.

3. Karboanhydraseaktivitet og dens inhibering

Siden karboanhydrasedomenet representerer en betydelig del av MN-proteinet (se Figur 8), ble det foretatt tester for å bestemme hvorvidt det faktisk er enzymatisk aktivt. Vero-celler infisert med vaccinia MN-konstruksjonen, som
5 inneholdt mer av MN-proteinet enn andre celler benyttet i de foreliggende forsøk, tjente som MN-proteinkilde. Cellene ble ekstrahert med RIPA-buffer, og MN-protein ble konsentrert og delvis rensset ved presipitasjon med MAb M75 og SAC. Immunpresipitatet ble testet med henblikk på CA-aktivitet. 78 µL presipitat inneholdt 1 enhet av enzymet. I ekstraktet ble konsentrasjonen av totale proteiner og av MN-protein
10 bestemt; 1 enhet enzym tilsvarte 145 ng MN-protein eller 0,83 mg totalprotein. Immunpresipitatet fra Vero-celler infisert med kontrollvirus hadde ingen enzymaktivitet. Aktivitet av MN-karboanhydrase ble inhibert av acetazolamid; $1,53 \times 10^{-8}$ M konsentrasjon av medikamentet reduserte enzymaktiviteten til 50%.

Innledende tester viste at konfluente kulturer av HeLa- eller av NIH3T3-celler
15 tolererte 10^{-5} - 10^{-3} M konsentrasjon av acetazolamid i 3 dager uten noen tegn til toksisitet og uten noen effekt på cellemorfologi. I sparsomme kulturer inhiberte ikke 10^{-5} M acetazolamid cellevekst, men allerede 10^{-4} M forårsaket en delvis inhibering. 10^{-5} M acetazolamid ble derfor tilsatt til NIH3T3-celler som nylig var transformert med [pGD.MN + MLV] komplekset. Etter inkubering i 4 dager ble koloniene fiksert og
20 farvet. Det ble ikke observert noen forskjell mellom celler dyrket i eller uten nærvær av acetazolamid, idet ingen av dem kunne skjelnes fra korrekt transformerte NIH3T3-celler. Den enzymatiske aktivitet av karboanhydrase er således ikke relevant for den transformerende virkning av MN-protein.

4. Celleadhesjons-bestemmelse

For å bestemme hvorvidt MN-protein er et celleadhesjonsmolekyl (CAM) eller ikke, ble det foretatt adhesjonsbestemmelser i bakteriologiske petriskåler av plast (ikke behandlet for bruk med vevskultur). Celler adhererer ikke til overflater av slike
skåler med mindre skålene er dekket med et bindingsprotein. NIH3T3-celler
30 adhererte, spredte seg og vokste på lapper av adsorbert MN-protein. Bare meget få celler festet seg utenfor de områder som var dekket med MN-protein.

Andre varianter av forsøket viste at NIH3T3-celler adhererte og spredte seg på lapper av adsorbert kollagen I og IV, fibronectin og laminin. NIH3T3-celler festet seg ikke til flekker av adsorbert gelatin, FCS eller BSA.

CGL1-, HeLa- og Vero-celler adhererte også til MN-protein, men 3 leukemi-cellelinjer viste ingen adheranse. CGL3-celler som sterkt uttrykker MN-protein, adhererte mindre effektivt til MN-proteinflekker enn hva CGL1 gjorde. Nærværet av 10^{-4} M acetazolamid i mediet påvirket ikke celleadhesjonen.

For å beskrive adhesjonsspesifisiteten ble MN-protein adsorbert med SAC ladet med MAb M75 (rettet mot MN) eller MAb M67, rettet mot et ubeslektet antigen (Pastorekova *et al.*, *supra*), før det ble påført overflaten av petriskålene. Absorpsjon med SAC-M75 komplekset opphevet fullstendig cellebindingsaktiviteten, mens absorpsjon med SAC-M67 var uten noen effekt.

Ytterligere celleadhesjonsresultater

Et forkortet MN som mangler TM- og IC-segmenter, utskilles i mediet av 5ET1-celler (et HeLa X fibroblasthybrid, analogt med CGL3-celler som rikelig uttrykker MN-protein) eller av Vero-celler infisert med VV som bærer MN-cDNA med deleterte TM- og IC-sekvenser. Det utskilte MN-protein ble rensset fra mediet og testet ved celleadhesjonsbestemmelser. Cellene adhererte, spredte seg og vokste bare på lapper dekket med adsorbert komplett MN-protein, men ikke på flekker av MN som manglet TM- og IC-regioner. Analoge resultater har også vært beskrevet for enkelte andre adhesjonsmolekyler. En rekke celler (NIH3T3, CGL1, CGL3, HeLa, XC) festet seg til MN-proteinflekker, hvilket tyder på at MN-reseptorer er vanlige på overflaten av vertebratceller.

Det ble også foretatt tester med ekstracellulære matriksproteiner eller kontrollproteiner avsatt som flekker på nitrocellulose. Disse dot-blottene ble behandlet med MN-proteinløsning. Bundet MN-protein ble påvist med MAb M75. MN-protein absorbertes til flekkene av kollagen I og IV, men ikke til fibronectin, laminin, gelatin eller BSA.

Muligheter for terapi. Det finnes mange nye prinsipper for cancerterapi som gjør bruk av onkoproteiner eller molekyler som interagerer med disse som mål [Mendelsohn & Lippman, «Principles of molecular cell biology of cancer; growth factors,» i: DeVita *et al.*, red. **Cancer: principles and practice of oncology**, s. 114-

133, 4. utg., Philadelphia: Lippincott (1993); DeVita *et al.* red. **Biologic therapy of cancer**, 2. utg., Philadelphia: Lippincott (1995)]. MN-proteinet og i det minste noen av dets ligander (eller reseptorer) synes å være særlig egnet for slike formål.

Eksempel 2

Identifisering av MN's bindingssete

MN-protein er et tumor-assosiert celledhesjonsmolekyl (CAM). For å identifisere dets bindingssete ble det syntetisert en rekke overlappende oligopeptider som strakte seg over det N-terminale domenet av MN-proteinet. Det N-terminale domenet er homologt med det for proteoglykaner og inneholder en tandemrepetisjon på 6 aminosyrer.

Denne rekke av oligopeptider ble testet ved celledhesjonsbestemmelsen i det vesentlige som beskrevet ovenfor i Eksempel 1. De syntetiske oligopeptidene ble immobilisert på hydrofobe plastoverflater for å se om de ville mediere tilfestingen, spredningen og veksten av celler. Det ble også undersøkt hvorvidt oligopeptidene eller antistoffer inhiberte tilfestingen av celler (NIH3T3, HeLa og CGL1) til rensset MN-protein belagt på slike plastflater. MN-proteinet ble affinitetsrenset på agarose, kovalent koblet til sulfonamid, idet MN-proteinet omfatter et CA-domene.

Flere av oligopeptidene ble funnet å være biologisk aktive: (i) når de var immobilisert på plasten, medierte de tilfestingen av celler (NIH3T3, HeLa og til CGL1); (ii) ved tilsetning til mediet konkurrerte de om tilfesting til celler med det immobiliserte MN-protein; (iii) disse oligopeptidene som forekom i mediet, inhiberer ikke tilfestingen av celler til TC-plast, men de forhindrer celle-celle-adhesjon og dannelse av intercellulære kontakter; (iv) behandling av immobilisert MN-protein og av aktive peptider med MAb M75 opphevet deres affinitet for cellene; og (v) bindingssetet for MN ble bestemt å være nært beslektet eller identisk med epitopen for MAb M75, hvorav minst to kopier er lokalisert i den 6 gangers tandemrepetisjon på 6 aminosyrer [aa 61-96 (SEKV. ID. NR:97)] i MN-proteinets proteoglykan-lignende domene.

Det ble trukket den konklusjon at ektopisk uttrykt MN-protein mest sannsynlig deltar i onkogenesen ved inngrep i normale celle-celle-kontakter. MN's bindingssete representerer et mulig mål som det kan konstrueres terapeutiske midler for.

Materialer og metoder

Affinitetskromatografi av MN/CA IX. MN/CA IX ble renset ved en enkelt adsorpsjonssyklus - eluering på sulfonamid-agarose, som beskrevet for andre CA [Falkbring *et al.*, **FEBS Letters**, **24**:229 (1972)]. Vi benyttet kolonner av p-aminoetyl-benzensulfonamid-agarose (Sigma). Kolonner med adsorbert MN/CA IX ble grundig vasket med PBS (NaCl 8,0 g/L, KCl 0,2 g/L, KH₂PO₄ 0,2 g/L, Na₂HPO₄ 1,15 g/L, pH=7,2) og eluert med 0,1 mM acetazolamid (Sigma). Alle rensetrinn ble foretatt ved 0-5°C, pH 7,2, ved fysiologiske saltkonsentrasjoner. Komplet MN/CA IX+ ble ekstrahert med 1% Triton X-100 i PBS fra Vero-celler infisert med vaccinia-virus som inneholdt et innskudd av fullstendig kodende region av MN/CA IX som beskrevet i Zavada *et al.*, **Int. J. Oncol.**, **10**:857 (1997). Før kromatografi ble ekstraktet fortynnet 1:6 med PBS og sentrifugert i 1 time ved 1500 x g. Trunkert MN/CA IX Δ TM Δ IC ble produsert fra en analog konstruksjon, bortsett fra at 3'-nedstrøms primeren for PCR var: 5' CGT CTA GAA GGA ATT CAG CTA GAC TGG CTC AGC A 3' [SEKV. ID. NR:117]. MN/CA IX Δ ble utskilt i mediet, hvorfra det ble affinitetsrenset etter sentrifugering som ovenfor. Alle rensetrinn ble overvåket ved dot-blotting.

Celler og medier. Følgende cellelinjer ble benyttet: HeLa, CGL1 = ikke-tumorigent hybrid HeLa x fibroblast, CGL3 = tumorigen segregant fra denne hybrid, NIH3T3-celler = muse-fibroblaster. Cellenes opprinnelse og vekstmedier er beskrevet i Zavada *et al.*, **Int. J. Cancer**, **54**:268 (1993) og Zavada *et al.*, **Int. J. Oncol.**, **10**:857 (1997). Dessuten benyttet vi også HT29, en cellelinje avledet fra kolorektalt karsinom (ATCC No. HBT-38).

Celleadhesjonsbestemmelse. Bestemmelsesbetingelsene er hovedsakelig som beskrevet i Eksempel 1. I korte trekk ble 1 μ g/mL renset MN/CA IX i 50 mM mono/bikarbonatbuffer, pH 9,2, adsorbert i 30 μ L dråper i bunnen av bakteriologiske 5 cm petriskåler i 1,5 timer. Deretter ble dråpene fjernet ved avsugning og skålene renset 3x med PBS og blokkert med 50% FCS i kulturmediet i 30 min. Det ble benyttet to varianter av testen. I den første ble hele bunnen av petriskålen blokkert med 50% FCS og skålene utsådd med 5 mL celleduspensjon (10^5 celler/mL). Etter inkubering over natten ble kulturene renset med PBS, fiksert og farvet. Etter den andre varianten ble bare området med adsorbert MN/CA IX blokkert, og på toppen av MN/CA IX-flekkene ble det tilsatt 30 μ L dråper celleduspensjon i vekstmedium, inneholdende tilsatte oligopeptider (eller kontroll uten peptider). Etter inkubering,

rensing og fiksering, ble kulturene farvet med 0,5% Trypanblått i 50 mM Tris-buffer, pH 8,5, i 1 time, rensset med vann og tørket. Farvede områder av tilheftede celler ble ekstrahert med 10% eddiksyre, ekstraktene overført til 96-brønns plater og absorbansen målt ved 630 nm med en mikroplateleser.

5 **ELISA.** Renset GST MN [Zavada *et al.*, (1993) *supra*] i konsentrasjonen 10 ng/mL i karbonatbuffer, pH 9,2, ble adsorbert i 3 timer i Maxisorb-strimler (NUNC). Etter vasking og blokkering (1 time) med 0,05% Tween 20 i PBS, ble 50 µL/brønn av antistoff + antigenblandingene tilsatt. Avsluttende fortynning av MAb 75 ascites-
væske var 10^{-6} ; peptidkonsentrasjonene varierte i henhold til deres affinitet for M75,
10 slik at bestemmelse av 50% slutt punktet var mulig. Disse blandinger ble adsorbert i 1,5 timer, etterfulgt av vasking med Tween-PBS. Bundet antistoff ble påvist med anti-mus IgG konjugat med peroksydase (SwAM-Px, SEVAC, Prag), fortynnet 1:1000. For farvereaksjonen ble OPD (o-fenylendiamin-dihydroklorid, Sigma) 1 mg/mL i 0,1M citratbuffer, pH 5,0, benyttet. Til dette ble det tilsatt H₂O₂ til en slutt konsentrasjon på
15 0,03%. Dette system balanseres slik at det muliggjør bestemmelse av antigen som konkurrerer om M75 så vel som om peptider som binder til epitopen av immobilisert GST-MN.

Peptider. Peptidene benyttet i denne undersøkelsen ble fremstillet ved hjelp av fastfase-metoden (Merrifield *et al.*; i: Gutte, B. (red.), **Peptides: Synthesis, Structures and Applications**, s. 93-169 (San Diego; Academic Press; 1995)] ved å
20 benytte Boc/Bzl-strategien. Peptidsyrene ble fremstillet på PAM-harpiks og peptid-amider på MeBHA-harpiks. Avbeskyttelse og avspaltning fra harpiksen ble foretatt med flytende hydrogenfluorid. Peptidene ble rensset ved C18 RP HPLC og karakterisert ved aminosyreanalyse og FAB MS spektroskopi.

25 **Western blott.** MN/CA IX antigener fra PAGE-gel ble overført til PVDF-membraner (Immobilon P, Millipore) og utviklet med M75, etterfulgt av SwAM-Px (se ovenfor) og diaminobenzidin (Sigma) med H₂O₂. For dot-blott benyttet vi nitro-cellulosemembraner.

30 **Fag-display.** Ph.D.-7 Phage Display Peptide Library kit ble benyttet for screening slik som anbefalt av produsenten (New England Biolabs.) En 96-brønns plate ble dekket med peptid SEKV. ID. NR:106. Biopanning ble foretatt ved inkubering av 2×10^{11} fag med målbelagt plate i 1 time. Ubundne fag ble vasket bort med TBST (50 mM Tris-HCl, pH 7,5, 150 mM NaCl, 0,1% Tween-20) og spesifikt

bundne fag ble eluert med M75 antistoff (2 µg i 100 µL TBS/brønn). Eluerte fag ble amplifisert og benyttet for ytterligere bindings- og amplifiseringsrunder for å anrike forrådet i favør av bindingssekvens. Etter 5 runder ble enkeltkloner plukket, amplifisert og sekvensert ved å benytte T7-sekvenseringssettet (Pharmacia).

5

Resultater

Affinitetskromatografi av MN/CA IX protein. For rensing av MN/CA IX protein bestemte vi å benytte affinitetskromatografi på sulfonamid-agarosekolonne, tidligere beskrevet for andre CA [Falkbring *et al.*, *supra*]. Fordelene ved denne metode er enkelhet og det at hele prosedyren foretas under ikke-denaturerende betingelser. Vaccinia-virusvektor med innskudd av det komplette MN/CA IX cDNA eller med trunkert cDNA (som mangler transmembrane og intracellulære domener) ble benyttet som kilde for MN/CA IX protein.

10

En enkelt syklus med adsorpsjon - eluering førte til relativt rene proteiner: MN/CA IX+ ga to bånd på 54 og 58 kDa, MN/CA IXΔ på 54,5 og 56 kDa. Disse proteinene reagerer sterkt med MAb M75 på western blott. I ekstrakter fra HeLa, CGL3 og HT29, viste blottet to bånd av samme størrelse som MN/CA IX+ renses fra vaccinia-viruskonstruksjon.

15

Adhesjon av celler til MN/CA IX protein. MN/CA IX immobilisert på hydrofob plast, muliggjorde tilfesting, utspredning og vekst av celler. Ekstremt lave konsentrasjoner av MN/CA IX tilsvarende 1 µg/mL renses protein i adsorpsjonsbuffer var tilstrekkelig til å gi denne effekt; andre celleadhesjonsmolekyler benyttes i 10-50x høyere konsentrasjoner. Bare komplett MN/CA IX protein var aktivt i celleadhesjonstesten, og trunkert MN/CA IX understøttet ikke celleadhesjon i det hele tatt eller oppviste kun en lav adhesjonsaktivitet, og i enkelte tilfeller virket endog som en celle «repellent».

20

25

Behandling av flekkene av immobilisert MN/CA IX med MAb M75 opphevet dets evne til å feste cellene, men kontrollen MAb M16, irrelevant for MN/CA IX, hadde ingen effekt. Blokkering av celletilfesting med M75 viser at epitopen er identisk med eller overlappende med bindingssettet av MN/CA IX for cellereseptorer.

30

Identifikasjon av epitopen som gjenkjennes av Mab M75. Forutgående kartlegging av M75 epitopen ved å benytte partielle sekvenser av ekstracellulære deler av MN/CA IX cDNA uttrykt fra bakterielle vektorer og testet på western blott,

lokaliserte det til PG-regionen. For nøyaktig kartlegging besto vår strategi i å syntetisere partielt overlappende oligopeptider på 15-25 aa som dekket PG-domenet og teste dem ved kompetitiv ELISA med M75. I henhold til resultatene ble dette etterfulgt av en rekke 6-12 aa oligopeptider. En vesentlig del av PG-domenet består av en 6-gangers tandemrepetisjon på 6 aa (aa 61-96) [SEKV. ID. NR:97]; 4 repetisjoner er identiske (GEEDLP) [SEKV. ID. NR:98] og 2 inneholder 2 aa utskiftet (SEEDSP [SEKV. ID. NR:141] og REEDPP [SEKV. ID. NR:142]).

Nedenfor følger resultatene av kompetitiv ELISA med rekombinant MN/CA IX og oligopeptider syntetisert i henhold til partialsekvenser av PG-regionen. MN/CA IX+ og Δ produsert i pattedyrceller, hadde en høyere serologisk aktivitet enn noe annet protein eller peptid inkludert i dette forsøk; fusjonsprotein GST-MN syntetisert i bakterier var mindre aktivt. Følgende peptider strekker seg over PG-regionen: GGSSGEDDPLGEEDLPSEEDSPC (aa 51-72) [SEKV. ID. NR:104]; GEEDLPSEEDSPREEDPPGEEDLPGEC (aa 61-85) [SEKV. ID. NR:105]; EDPPGEEDLPGEEDLPGEEDLPEVC (aa 75-98) [SEKV. ID. NR:106]; og EVKPKSEEEGSLKLE (aa 97-111) [SEKV. ID. NR:118]. SEKV. ID. NR:104 og 106 forårsaket 50% inhibering ved 1 ng/mL. Disse to oligopeptidene er gjensidig ikke-overlappende, og epitopen er således gjentatt i begge. SEKV. ID. NR:105 var 1000x mindre aktiv, sannsynligvis som følge av en annen konformasjon. SEKV. ID. NR:118 var inaktiv; den inneholder således ikke M75 epitopen.

Det neste trinn for å identifisere epitopen var å syntetisere oligopeptider som inneholdt alle sirkulære permutasjoner av motivet GEEDLP [SEKV. ID. NR:98] gjentatt to ganger. Alle 6 av de følgende dodekapeptidene [SEKV. ID. NR:119-124] var serologisk aktive (2 mer og 4 mindre aktive): GEEDLPGEEDLP [SEKV. ID. NR:119]; EEDLPGEEDLPG [SEKV. ID. NR:120]; EDLPGEEDLP [SEKV. ID. NR:121]; DLPGEEDLPGEE [SEKV. ID. NR:122]; LPGEEDLPGEED [SEKV. ID. NR:123] og PGEEDLPGEEDL [SEKV. ID. NR:124]. Den følgende rekke av 7 aa-sekvenser, flankert av alanin i begge ender, ble testet: APGEEDLPA [SEKV. ID. NR:125]; AGEEDLPGA [SEKV. ID. NR:126]; AEEDLPGEA [SEKV. ID. NR:127]; AEDLPGEEA [SEKV. ID. NR:128]; ADLPGEEDA [SEKV. ID. NR:129]; og ALPGEEDLA [SEKV. ID. NR:130]. Resultatene viste at den minste serologisk aktive sekvens er oligopeptidet APGEEDLPA [SEKV. ID. NR:125]. SEKV. ID. NR:127-130 viste seg negative i konkurransen med 100 μ g/mL. Dessuten konkurrerte ingen av de

følgende enda kortere oligopeptidene (6 + 2 aa) i ELISA om M75: AGEEDLPA [SEKV. ID. NR:131]; AEEDLPGA [SEKV. ID. NR:132]; AEDLPGEA [SEKV. ID. NR:133]; ADLPGEAA [SEKV. ID. NR:134]; ALPGEEDA [SEKV. ID. NR:135]; og APGEEDLA [SEKV. ID. NR:136].

I oligopeptidene av SEKV. ID. NR:104, 105, 106 og 118, forelå den C-terminale aminosyre som en syre, mens den C-terminale aminosyre i alle andre oligopeptider forelå som et amid. Det er klart at affiniteten mellom disse oligopeptidene og MAb M75 øker meget sterkt med størrelsen av peptidmolekylet.

Forsøk på demonstrering av adhesjon av celler til immobiliserte oligopeptider.

Vår opprinnelige plan var å følge pionérbetjentet til Piersbacher & Ruoslahti, **PNAS**, 81:5985 (1984). De koblet testet oligopeptider til adsorbert bovint serumalbumin med krysskoblende SPDP (N-succinimidyl-3[pyridylhydro]-propionat). Dette er årsaken til at vi til C-enden av oligopeptidene SEKV. ID. NR:104-106 adderte cystein, som ville muliggjøre orientert kobling til adsorbert albumin. Vi demonstrerte kobling av peptidene direkte i petriskåler ved immunperoksydase-farving med M75. Uheldigvis adhererte CGL1- eller CGL3-celler til kontrollalbumin behandlet med SPDP og blokkerte med etanolamin (i stedet for oligopeptider) så sterkt som til BSA-flekke med koblede oligopeptider. Vi klarte ikke å oppheve denne uspesifikke adhesjon. Oligopeptidene SEKV. ID. NR:104-106 adsorberer bare meget svakt til bakteriologiske petriskåler, og tillater derved ikke utførelse av celle-adhesjonsbestemmelsen.

Som et alternativ testet vi inhiberingen av celleadhesjon til MN/CA IX flekker av oligopeptider tilsatt til mediet sammen med celsuspensjonen, som beskrevet av Piersbacher & Ruoslahti, *supra*. Alle peptidene SEKV. ID. NR:104-106 og 118-136, ble testet i konsentrasjoner på 100 og 10 µg/mL. Ingen av dem inhiberte reproduserbart adhesjonen av CGL1-celler.

Oligopeptider med affinitet til M75 epitop som inhiberer celleadhesjon til MN/CA IX. Som et alternativ til monoklonale antistoffer startet vi selektering av oligopeptider som oppviste affinitet for M75 epitop så vel som for MN/CA IX reseptor-bindingssitet fra et fag-display-bibliotek av tilfeldige heptapeptider - Ph.D.-7. Vårt mål var å selektere fag som inneholdt de ønskede heptapeptidene ved panning på immobilisert peptid SEKV ID. NR:106 og påfølgende eluering med M75. Eluerte fag

ble formert i passende bakterier og underkastet 4 ytterligere runder panning og eluering. Fra den selekterte fag-populasjon ble det plukket 10 plakk, som ble amplifisert, og den heptapeptid-kodende region ble sekvensert. Bare 3 heptapeptider var representert. Disse tre heptapeptidene er etter addisjon av alanin på begge sider, følgende nonapeptider: AKKMKRRKA [SEKV. ID. NR:137]; AITFNAQYA [SEKV. ID. NR:138]; og ASASAPVSA [SEKV. ID. NR:139]. Det siste heptapeptidet, igjen syntetisert med adderte terminale alaniner som nonapeptid AGQTRSPLA [SEKV. ID. NR:140], ble identifisert ved panning på GST-MN og eluert med acetazolamid. Dette siste peptid har affinitet for det aktive setet av MN/CA IX karboanhydrase. Vi syntetiserte disse peptidene på 7+2 aa og testet dem ved kompetitiv ELISA og celle-adhesjonsinhibering. Begge testene ga i det vesentlige overensstemmende resultater: peptid SEKV. ID. NR:138 viste den høyeste aktivitet, peptid SEKV. ID. NR:137 var mindre aktiv, peptid SEKV. ID. NR:139 var marginalt positiv bare ved ELISA, og peptid SEKV. ID. NR:140 var inaktivt. I alle disse 4 nonapeptidene forelå det C-terminale amid som amid.

Diskusjon

Rensing av transmembrane proteiner som MN/CA IX fører ofte til tekniske problemer fordi de har tendens til å danne aggregater med andre membranproteiner som følge av deres hydrofobeTM-segmenter. For å unngå dette konstruerte vi trunkerte MN/CA IX Δ IC Δ TM, som utskilles i mediet. Trunkert MN/CA IX ble faktisk oppnådd i høyere renhet enn MN/CA IX⁺. Dessverre var dette protein av liten nytte for våre formål, siden det var inaktivt i celleadhesjonstesten. En slik situasjon har også vært beskrevet for andre celleadhesjonsmolekyler; idet deres utskilte, forkortede form enten antar en inaktiv konformasjon, eller adsorberes til hydrofob plast «upside down», mens komplette proteiner adsorberer med hydrofobeTM-segmenter i den «korrekte» posisjon.

MN/CA IX protein danner oligomerer på 150 kDa, koblet med disulfid-bindinger. Det var ikke kjent hvorvidt disse er homo- eller hetero-oligomerer, men PAGE og western blott analyse tyder på at disse sannsynligvis er homo-oligomerer, mest sannsynlig trimerer, siden det på gel farvet med Coomassieblått ikke fremkom noen ytterligere bånd av sammenlignbar intensitet som to spesifikke bånd for MN/CA IX. Det er også usannsynlig at det kunne eksistere et ytterligere protein som

migrerte sammen med én av de to hoved MN/CA IX båndene, siden intensiteten av deres farving på gelen og på western blott er godt sammenlignbare.

Det kan ikke være noen tvil om spesifisiteten av celle-tilfesting til rensset MN/CA IX+. Det oppheves av spesifikk MAb M75 ved en fortynning på 1:1000 av ascitesvæske. Dette er en korleksjon til vår tidligere rapport i Zavada *et al.*, *Int. J. Oncol.*, **10:857** (1997), hvor vi observerte at MN/CA IX produsert av vaccinia-virusvektor og fusjonsprotein GST-MN, støtter celleadhesjon, men vi innså ikke at GST-forankringen selv inneholder et annet bindingssete som ikke blokkeres av M75.

MAb M75 reagerer utmerket med MN/CA IX under alle omstendigheter - med nativt antigen på overflaten av levende celler, med denaturert protein på western blott og med antigen i parafinsnitt av biopsier fiksert med formaldehyd, hvilket tyder på at epitopen er liten og sammenhengende. Ved kompetitiv ELISA var den korteste sekvens som var reaktiv med M75, 7+2 aa, men affiniteten mellom M75 og undersøkte peptider avhang sterkt av deres molekylvekt. Komplet MN/CA IX var 100.000x mer aktiv enn det minste serologisk aktive peptid uttrykt i konsentrasjon i vekt/volum. Uttrykt i molar konsentrasjon ville denne forskjell være 150.000.000x. Oligopeptider av midlere størrelse oppviste også midlere aktiviteter. Det gjenstår å utforske hvorvidt slike forskjeller i aktivitet skyldes konformasjonen som avhenger av størrelsen av molekylet, eller av at det komplette MN/CA IX inneholder flere kopier av epitopen, men det minste molekyl bare én.

I betraktning av den mulighet at epitopen er identisk med celleadhesjonsstrukturen i MN/CA IX, kan vi forstå hvorfor vi ikke klarte å påvise inhibering av celleadhesjon med oligopeptidene. Bindingssetet er ganske enkelt ikke så likefrem som prototyp-peptidet, RGD [Winter, J., i Cleland & Craik (red.), **Protein Engineering. Principles and Practice**, s. 349-369 (N.Y.; Wiley-Liss; 1996)].

Man kan selvsagt hevde at størrelsen av MN/CA IX er omtrent den samme som immunglobulinmolekylet og at binding av M75 til dens epitop sterisk kan hindre en forskjellig sekvens i et celle-tilfestingsete. Denne innvendig er gjort usannsynlig ved å blokkere både M75 epitop og cellebindingssetet med nonapeptider 7 + 2 aa. Det resultatet tyder sterkt på at epitopen og bindingssetet virkelig er identiske.

MN/CA IX og særlig dens PG-region, synes å være et potensielt mål molekyl for terapi av følgende grunner: (i) det eksponeres på celleoverflaten; (ii) det forekommer i en høy andel i visse humane karsinomer; (iii) det er normalt uttrykt

MN/CA IX i slimhinnen i alimentærtrakten, som i motsetning til tumorene ikke er tilgjengelig for sirkulerende antistoffer; (iv) det utskilles ikke (eller kun minimalt) i kroppsvæskene; (v) motivet GEEDLP [SEKV. ID. NR:98] gjentas 18x på overflaten av hvert MN/CA IX molekyl. Oligopeptid-display-bibliotek benyttes i de første trinn for å utvikle nye medikamenter [Winter, J. *supra*]. Utvalgte oligopeptider kan tjene som ledeforbindelser for datamaskinkonstruksjon av nye molekyler med ytterligere egenskaper som kreves av et medikament [DeCamp *et al.*, i, Cleland & Craik (red.) *supra*, s. 467-505].

Eksempel 3

Påvisning av peptider som binder til

MN-protein ved å benytte fag-display

(a) For å identifisere peptider som gjenkjennes av MN-protein ble et heptapeptid fag-display-bibliotek [Ph.D.®-7 Peptide 7-mer Library Kit (fag-display peptid-biblioteksett); New England Biolabs; Beverly, MA (USA)] underkastet screening. Ved screening av biblioteket ble det foretatt en seleksjonsprosess, dvs. biopanning [Parmley & Smith, *Gene*, **73**:308 (1988); Noren, C.J., *NEB Transcript*, **8**(1):1 (1996)] ved å inkubere de fag som koder for peptidene med en plate dekket med MN-protein, vaske bort ubundet fag, eluere og amplifisere den spesifikt bundne fag.

MN-målproteinet var i denne prosess et glutation-S-transferase (GST) MN-fusjonsprotein (GST-MN). GST-MN er et rekombinant produsert fusjonsprotein uttrykt fra pGEX-3X-MN inneholdende cDNA for MN-proteinet uten signalpeptidet. GST-MN ble produsert i bakterier under modifiserte dyrkningsbetingelser (nedsatt optisk tetthet, øket temperatur). Slik dyrkning forhindret for tidlig terminering av translasjon og resulterte i syntese av proteinmolekylene som i den alt overveiende del var av full lengde. GST-MN-proteinet ble benyttet for belegging av brønnene og binding av de relevante fag. De bundne fagene ble deretter eluert med acetazolamid, amplifisert og benyttet for to ytterligere screeningrunder.

Etter sekvensering av flere uavhengige fag-kloner oppnådd etter den tredje screeningrunde, ble følgende heptapeptider oppnådd:

- (1) GETRAPL (SEKV. ID. NR:107)
- (2) GETREPL (SEKV. ID. NR:108)
- (3) GQTRSPL (SEKV. ID. NR:109)
- (4) GQTRSPL (")
- 5 (5) GQTRSPL (")
- (6) GQTRSPL (")
- (7) GQTRSPL (")

Heptapeptidene oppviser meget like eller identiske sekvenser, hvilket indikerer at bindingen er spesifikk. Det at fag som bærer disse heptapeptidene ble eluert med acetazolamid, en inhibitor av karboanhydraseaktivitet, indikerer at peptidene binder til MN-proteinets CA-domene.

(b) Analog screening av heptapeptid-fag-display-biblioteket foretas ved å benytte kollagen I, vist å binde MN-protein, for eluering av fag. Forskjellige peptider som binder til ulike deler av MN-proteinmolekylet forventes å bli identifisert. Etter identifisering av slike MN-bindingspeptider skal de tilsvarende syntetiske peptidene deretter analyseres med henblikk på deres biologiske virkninger.

Eksempel 4

Tilgjengelighet *in vivo* av MN-protein uttrykt i tumorceller og i mavesekk

Lewisrotter (384 g) som var bærere av en BP6 subkutan tumor (ca. 1 cm i diameter) og uttrykte rotte MN-protein, ble injisert intraperitonealt (i.p.) med ^{125}I -M75 Mab ($2,5 \times 10^6$ cpm). 5 dager senere ble 0,5-1 g biter av tumoren og organer veid og deres radioaktivitet målt med en gammateller.

Tabell 2 sammenfatter resultatene. Den høyeste radioaktivitet forekom i tumoren. Relativt høy radioaktivitet ble funnet i leveren og nyre, hvilket formodentlig reflekterer utskillelsen av mus IgG fra blodet. Mavesekken hadde et vedvarende relativt lavt radioaktivitetsnivå, hvilket indikerer at M75 Mab bare hadde begrenset tilgang til MN-proteinet eksponert i maveslimhinnen.

TABELL 2

Fordeling av radioaktivitet av ^{125}I -M75 i rotteorganer og i tumoren

Organ		cpm/g			
Nyre	2153	2184			
Milt	653	555			
Lever	1993	1880			
Lunge	1183	1025			
Blod	1449				
Hjerte	568	477			
Mavesekk	1184	1170			
Testikkel	812	779			
Hale	647				
Tumor	3646	4058	3333	8653	3839

Eksempel 5

FACS-analyse av MN-proteinekspresjon i

CGL3-celler - apoptose

En FACS-undersøkelse ble utviklet for å bestemme de tilstander som påvirker syntesen av MN-protein og for å analysere celledyklusfordelingen av MN-positive *versus* MN-negative celler i en CGL3-populasjon stimulert til apoptose. Tidligere western blotting-analyser har vist at CGL3-celler uttrykker en relativt stor mengde MN-protein under forskjellige dyrkningsbetingelser. CGL3-celler ansees som en konstitutiv produsent av MN-proteiner. Western blotting gjenkjenner imidlertid ikke små forskjeller i proteinnivået. FACS derimot muliggjør påvisning av individuelle MN-positive celler, en beregning av deres andel av den analyserte populasjon, et estimat av nivået av MN-protein i cellene og en bestemmelse av celledyklusfordelingen.

For å studere effekten av dyrkningsbetingelsene på MN-ekspressjon i CGL3-celler, ble CGL3-celler utplatet i forskjellige relative tettheter og serum-konsentrasjoner. 3 dager etter utplating ble cellene oppsamlet, overflatemerket med M75 Mab etterfulgt av FITC-konjugert anti-mus IgG og umiddelbart analysert med FACS.

Analysen viste at i adherente celler er MN-ekspressjon, som i HeLa-celler, avhengig av celletetthet. Kulturer med lav tetthet produserte imidlertid fremdeles

påvisbare mengder MN-protein. I kulturer med lav tetthet synes serum-konsentrasjonen ikke å spille noen rolle. I kulturer med relativt høy tetthet resulterte en avtagende serumkonsentrasjon i svakt minsket MN-ekspresjon, sannsynligvis som følge av en lavere tetthet som cellene var i stand til å nå under dyrkingen i de 3 dagene.

Effekten av den aktuelle celletetthet er bemerkelsesverdig, og MN-ekspresjon (påvisbar i 15-90% av cellene) representerer en meget følsom overvåkningsfaktor. I alle forsøk var det en ca. 5% høyere prosentandel sykliserende celler i den MN-positive del av populasjonen sammenlignet med den MN-negative del. Dette faktum utløste omgående analyse av cellyklusfordelingen av MN-positive CGL3-celler under ugunstige vekstbetingelser, dvs. etter induksjon av apoptose.

Apoptose

CGL3-celler ble stimulert til apoptotisk død med diverse medikamenter, inklusivt cykloheksimid, aktinomycin D og deksametason. FACS-undersøkelsen viste at inntreden av apoptose forsinkes i MN-positive celler, hvilket tyder på en beskyttende rolle av MN i denne prosess. Det ble også observert at induksjonen av apoptose resulterte i nedregulering av MN-ekspresjon på en tidsavhengig måte. Det samme fenomen ble beskrevet for Bcl-2 anti-apoptotisk protein, og det er en eksisterende oppfatning at nedreguleringen av visse regulatorgener under apoptose sensibiliserer cellene til å undergå apoptotisk død. For å bevise rollen av MN ved apoptose må det foretas en lignende undersøkelse med celler transfektert med MN cDNA.

De foreløpige resultater indikerer den mulige medvirkning av MN ved undertrykkelsen av apoptose. Den seneste oppfatning at tumorer oppstår både som en følge av øket proliferasjon og minsket celledød, synes å være i overensstemmelse med assosiasjonen av MN-proteinet med tumorer *in vivo*.

ATCC deponeringer

Materialene angitt nedenfor ble deponert hos American Type Culture Collection (ATCC) nå ved 10810 University Blvd., Manassus, Virginia 20110-2209 (USA). Deponeringene ble foretatt i henhold til the Budapest Treaty on the International Recognition of Deposited Microorganisms for the Purposes of Patent

Procedure and Regulations thereunder (Budapest-avtalen). Oppretthold av en levedyktig kultur er sikret i 30 år fra deponeringsdato. Hybridomene og plasmidene vil bli gjort tilgjengelige av ATCC innenfor rammen av Budapest-avtalen, og forutsatt enighet mellom søkerne og ATCC som sikrer ubegrenset tilgjengelighet av de

5 deponerte hybridomer og plasmider for almenheten etter utstedelse av patent fra foreliggende søknad. Tilgang til den deponerte stamme er ikke å oppfatte som en lisens til å praktisere oppfinnelsen i strid med rettigheter som er gitt av en regjerings myndighet i henhold til dens patentlovgivning.

10	Hybridom	Deponeringsdato	ATCC nr.
	VU-M75	17. september, 1992	HB 11128
	MN 12.2.2	9. juni, 1994	HB 11647

	Plasmid	Deponeringsdato	ATCC nr.
15	A4a	6. juni, 1995	97199
	XE1	6. juni, 1995	97200
	XE3	6. juni, 1995	97198

Beskrivelsen av ovennevnte utførelsesformer av oppfinnelsen er presentert

20 med det formål å illustrere og beskrive.

25

30

Sekvensliste

<110> Zavada, Jan
Pastorekova, Silvia
Pastorek, Jaromir

<120> MN Gene and Protein

<130> D-0021.5 PCT

<140> PCT/US99/24879

<141> 1999-10-22

<150> 09/177,776

<151> 1998-10-23

<150> 09/178,115

<151> 1998-10-23

<160> 143

<170> PatentIn Ver. 2.0

<210> 1

<211> 1522

<212> DNA

<213> HUMAN

<220>

<221> CDS

<222> (13)..(1389)

<220>

<221> mat_peptide

<222> (124)..(1389)

<400> 1

acagtcagcc gc atg gct ccc ctg tgc ccc agc ccc tgg ctc cct ctg ttg 51
Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu
-35 -30 -25

atc ccg gcc cct gct cca ggc ctc act gtg caa ctg ctg ctg tca ctg 99
Ile Pro Ala Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu
-20 -15 -10

ctg ctt ctg atg cct gtc cat ccc cag agg ttg ccc cgg atg cag gag 147
Leu Leu Leu Met Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu
-5 -1 1 5

gat tcc ccc ttg gga gga ggc tct tct ggg gaa gat gac cca ctg ggc 195
Asp Ser Pro Leu Gly Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly
10 15 20

gag gag gat ctg ccc agt gaa gag gat tca ccc aga gag gag gat cca 243
Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro
25 30 35 40

ccc gga gag gag gat cta cct gga gag gag gat cta cct gga gag gag 291
Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu
45 50 55

gat cta cct gaa gtt aag cct aaa tca gaa gaa gag ggc tcc ctg aag 339
Asp Leu Pro Glu Val Lys Pro Lys Ser Glu Glu Glu Gly Ser Leu Lys
60 65 70

tta gag gat cta cct act gtt gag gct cct gga gat cct caa gaa ccc 387
Leu Glu Asp Leu Pro Thr Val Glu Ala Pro Gly Asp Pro Gln Glu Pro
75 80 85

Sekvensliste

cag Gln	aat Asn	aat Asn	gcc Ala	cac His	agg Arg	gac Asp	aaa Lys	gaa Glu	ggg Gly	gat Asp	gac Asp	cag Gln	agt Ser	cat His	tgg Trp	435
90						95					100					
cgc Arg	tat Tyr	gga Gly	ggc Gly	gac Asp	ccg Pro	ccc Pro	tgg Trp	ccc Pro	cgg Arg	gtg Val	tcc Ser	cca Pro	gcc Ala	tgc Cys	gcg Ala	483
105					110					115					120	
ggc Gly	cgc Arg	ttc Phe	cag Gln	tcc Ser	ccg Pro	gtg Val	gat Asp	atc Ile	cgc Arg	ccc Pro	cag Gln	ctc Leu	gcc Ala	gcc Ala	ttc Phe	531
				125					130					135		
tgc Cys	ccg Pro	gcc Ala	ctg Leu	cgc Arg	ccc Pro	ctg Leu	gaa Glu	ctc Leu	ctg Leu	ggc Gly	ttc Phe	cag Gln	ctc Leu	ccg Pro	ccg Pro	579
			140					145					150			
ctc Leu	cca Pro	gaa Glu	ctg Leu	cgc Arg	ctg Leu	cgc Arg	aac Asn	aat Asn	ggc Gly	cac His	agt Ser	gtg Val	caa Gln	ctg Leu	acc Thr	627
		155					160					165				
ctg Leu	cct Pro	cct Pro	ggg Gly	cta Leu	gag Glu	atg Met	gct Ala	ctg Leu	ggt Gly	ccc Pro	ggg Gly	cgg Arg	gag Glu	tac Tyr	cgg Arg	675
	170					175					180					
gct Ala	ctg Leu	cag Gln	ctg Leu	cat His	ctg Leu	cac His	tgg Trp	ggg Gly	gct Ala	gca Ala	ggt Gly	cgt Arg	ccg Pro	ggc Gly	tgc Ser	723
185					190					195					200	
gag Glu	cac His	act Thr	gtg Val	gaa Glu	ggc Gly	cac His	cgt Arg	ttc Phe	cct Pro	gcc Ala	gag Glu	atc Ile	cac His	gtg Val	gtt Val	771
				205					210					215		
cac His	ctc Leu	agc Ser	acc Thr	gcc Ala	ttt Phe	gcc Ala	aga Arg	gtt Val	gac Asp	gag Glu	gcc Ala	ttg Leu	ggg Gly	cgc Arg	ccg Pro	819
			220					225					230			
gga Gly	ggc Gly	ctg Leu	gcc Ala	gtg Val	ttg Leu	gcc Ala	gcc Ala	ttt Phe	ctg Leu	gag Glu	gag Glu	ggc Gly	ccg Pro	gaa Glu	gaa Glu	867
		235					240					245				
aac Asn	agt Ser	gcc Ala	tat Tyr	gag Glu	cag Gln	ttg Leu	ctg Leu	tct Ser	cgc Arg	ttg Leu	gaa Glu	gaa Glu	atc Ile	gct Ala	gag Glu	915
	250					255					260					
gaa Glu	ggc Gly	tca Ser	gag Glu	act Thr	cag Gln	gtc Val	cca Pro	gga Gly	ctg Leu	gac Asp	ata Ile	tct Ser	gca Ala	ctc Leu	ctg Leu	963
265					270					275					280	
ccc Pro	tct Ser	gac Asp	ttc Phe	agc Ser	cgc Arg	tac Tyr	ttc Phe	caa Gln	tat Tyr	gag Glu	ggg Gly	tct Ser	ctg Leu	act Thr	aca Thr	1011
				285					290					295		
ccg Pro	ccc Pro	tgt Cys	gcc Ala	cag Gln	ggt Gly	gtc Val	atc Ile	tgg Trp	act Thr	gtg Val	ttt Phe	aac Asn	cag Gln	aca Thr	gtg Val	1059
			300					305					310			
atg Met	ctg Leu	agt Ser	gct Ala	aag Lys	cag Gln	ctc Leu	cac His	acc Thr	ctc Leu	tct Ser	gac Asp	acc Thr	ctg Leu	tgg Trp	gga Gly	1107
		315					320					325				
cct Pro	ggt Gly	gac Asp	tct Ser	cgg Arg	cta Leu	cag Gln	ctg Leu	aac Asn	ttc Phe	cga Arg	gcg Ala	acg Thr	cag Gln	cct Pro	ttg Leu	1155
	330					335					340					
aat Asn	ggg Gly	cga Arg	gtg Val	att Ile	gag Glu	gcc Ala	tcc Ser	ttc Phe	cct Pro	gct Ala	gga Gly	gtg Val	gac Asp	agc Ser	agt Ser	1203
345					350					355					360	

Sekvensliste

cct cgg gct gct gag cca gtc cag ctg aat tcc tgc ctg gct gct ggt 1251
Pro Arg Ala Ala Glu Pro Val Gln Leu Asn Ser Cys Leu Ala Ala Gly
365 370 375

gac atc cta gcc ctg gtt ttt ggc ctc ctt ttt gct gtc acc agc gtc 1299
Asp Ile Leu Ala Leu Val Phe Gly Leu Leu Phe Ala Val Thr Ser Val
380 385 390

gcg ttc ctt gtg cag atg aga agg cag cac aga agg gga acc aaa ggg 1347
Ala Phe Leu Val Gln Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly
395 400 405

ggt gtg agc tac cgc cca gca gag gta gcc gag act gga gcc 1389
Gly Val Ser Tyr Arg Pro Ala Glu Val Ala Glu Thr Gly Ala
410 415 420

tagaggctgg atcttgagaga atgtgagaag ccagccagag gcatctgagg gggagccggt 1449

aactgtcctg tcctgctcat tatgccactt ccttttaact gccagaagaat tttttaaaat 1509

aaatatttat aat 1522

<210> 2
<211> 459
<212> PRT
<213> HUMAN

<400> 2
Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu Ile Pro Ala
-35 -30 -25

Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu Leu Leu Leu
-20 -15 -10

Met Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu Asp Ser Pro
-5 -1 1 5 10

Leu Gly Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp
15 20 25

Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu
30 35 40

Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro
45 50 55

Glu Val Lys Pro Lys Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu Asp
60 65 70 75

Leu Pro Thr Val Glu Ala Pro Gly Asp Pro Gln Glu Pro Gln Asn Asn
80 85 90

Ala His Arg Asp Lys Glu Gly Asp Asp Gln Ser His Trp Arg Tyr Gly
95 100 105

Gly Asp Pro Pro Trp Pro Arg Val Ser Pro Ala Cys Ala Gly Arg Phe
110 115 120

Gln Ser Pro Val Asp Ile Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala
125 130 135

Leu Arg Pro Leu Glu Leu Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu
140 145 150 155

Leu Arg Leu Arg Asn Asn Gly His Ser Val Gln Leu Thr Leu Pro Pro
160 165 170

Sekvensliste

Gly Leu Glu Met Ala Leu Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln
175 180 185
Leu His Leu His Trp Gly Ala Ala Gly Arg Pro Gly Ser Glu His Thr
190 195 200
Val Glu Gly His Arg Phe Pro Ala Glu Ile His Val Val His Leu Ser
205 210 215
Thr Ala Phe Ala Arg Val Asp Glu Ala Leu Gly Arg Pro Gly Gly Leu
220 225 230 235
Ala Val Leu Ala Ala Phe Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala
240 245 250
Tyr Glu Gln Leu Leu Ser Arg Leu Glu Glu Ile Ala Glu Glu Gly Ser
255 260 265
Glu Thr Gln Val Pro Gly Leu Asp Ile Ser Ala Leu Leu Pro Ser Asp
270 275 280
Phe Ser Arg Tyr Phe Gln Tyr Glu Gly Ser Leu Thr Thr Pro Pro Cys
285 290 295
Ala Gln Gly Val Ile Trp Thr Val Phe Asn Gln Thr Val Met Leu Ser
300 305 310 315
Ala Lys Gln Leu His Thr Leu Ser Asp Thr Leu Trp Gly Pro Gly Asp
320 325 330
Ser Arg Leu Gln Leu Asn Phe Arg Ala Thr Gln Pro Leu Asn Gly Arg
335 340 345
Val Ile Glu Ala Ser Phe Pro Ala Gly Val Asp Ser Ser Pro Arg Ala
350 355 360
Ala Glu Pro Val Gln Leu Asn Ser Cys Leu Ala Ala Gly Asp Ile Leu
365 370 375
Ala Leu Val Phe Gly Leu Leu Phe Ala Val Thr Ser Val Ala Phe Leu
380 385 390 395
Val Gln Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser
400 405 410
Tyr Arg Pro Ala Glu Val Ala Glu Thr Gly Ala
415 420

<210> 3
<211> 29
<212> DNA
<213> HUMAN

<400> 3
cgcccagtgg gtcattctcc ccagaagag

29

<210> 4
<211> 19
<212> DNA
<213> HUMAN

<400> 4
ggaatcctcc tgcattccg

19

Sekvensliste

<210> 5
 <211> 10898
 <212> DNA
 <213> HUMAN
 <220>
 <221> gene
 <222> (1)..(10898)

```

<400> 5
ggatcctggt gactcgtgac cttaccccca accctgtgct ctctgaaaca tgagctgtgt 60
ccactcaggg ttaaatggat taagggcggt gcaagatgtg ctttggttaa ca gatgcttg 120
aaggcagcat gctcgttaag agtcatcacc aatcccta at ctcaagtaat caggacaca 180
aacactgcgg aaggccgcag ggtcctctgc ctaggaaaac cagagacctt tgttcacttg 240
tttacttgac cttccctcca ctattgtcca tgaccctgcc aaatccccct ctgtgagaaa 300
cacccaagaa ttatcaataa aaaaaataat ttaaaaaaa aatacaaaaa aaaaaaaaaa 360
aaaaaaaaaa gacttacgaa tagttattga taaatgaata gctattggta aagccaagta 420
aatgatcata ttcaaaacca gacggccatc atcacagctc aagtctacct gat ttgatct 480
ctttatcatt gtcattcttt ggattcacta gattagtc at catctcaaa attctcccc 540
aagttcta at tacgttccaa acatttaggg gtacatgaa gcttgaacct actaccttct 600
ttgcttttga gccatgagtt gtaggaatga tgagtttaca cttacatgc tggggatta 660
tttaaaactt acctctaagt cagttgggta gcctttggct tatttttgta gctaattttg 720
tagttaatgg atgcactgtg aatcttgcta tgatagtttt cctccacact ttgccactag 780
gggtaggtag gtactcagtt ttcagtaatt gcttaccta gaccctaagc cctatttctc 840
ttgtactggc ctttatctgt aatatgggca tatttaatac aatataattt ttggagtttt 900
tttgtttgtt tgtttgtttg tttttttgag acggagtctt gcatctgtca tgccaggct 960
ggagtagcag tgggtgccatc tcggctcact gcaagctcca cctcccgagt tcacgccatt 1020
ttcctgcctc agcctcccga gtagctggga ctacagggc ccgccaccat gccgggctaa 1080
ttttttgtat ttttggtaga gacgggggtt cacctgttta gccagaatgg tctcgatctc 1140
ctgacttcgt gatccaccgc cctcggcctc ccaaagttct gggattacag gtgtgagcca 1200
ccgcacctgg ccaatttttt gagtctttta aagtaaaaat atgtcttgta agctggtaac 1260
tatggtagat ttctttttat taatgtgggt ctgacggcta tatagtttct tttagtttg 1320
gcattcatat gctacttttt gcagtccttt cattacattt ttctctcttc atttgaagag 1380
catgttatat cttttagctt cacttggtt aaaaggttct ctcatagcc taacacagt 1440
tcattgtttg taccacttgg atcataagt gaaaaacagt caagaaattg cacagtaata 1500
cttgttttga agagggatga ttcaggtgaa tctgacacta agaaactccc ctacctgagg 1560
tctgagattc ctctgacata gctgtatata ggcttttctc ttgacagcct gtgactgcgg 1620
actatttttc ttaagcaaga tatgctaaag tttgtgagc ctttttccag agagaggtct 1680
catatctgca tcaagtgaga acatataatg tctgcatggt tccatatttc aggaatgttt 1740
gcttgtgttt tatgtctttt tatagacagg gaaacttgtt cctcagtgac ccaaaagagg 1800
tgggaattgt tttgttatat catcattggc caacgtttc tgaccttggg aacaattaa 1860
ggttcataat ctcaattctg tcagaattgg tacaagaaat agctgctatg tttcttgaca 1920
ttccacttgg taggaaataa gaatgtgaaa ctcttcagtt ggtgtgtgtc cctngttttt 1980
ttgcaatttc cttcttactg tgttaaaaaa aagtatgatc ttgtcttgag aggtgaggca 2040
ttcttaatca tgtattttaa agatcaataa tataatcctt tcaaggatta tgtctttatt 2100
ataataaaga taatttgtct ttaacagaat caataatata atcccttaaa ggattatatc 2160
tttgctgggc gcagtggctc acacctgtaa tcccagcact ttgggtggcc aaggtggaag 2220
gatcaaatgt gcctacttct atattatctt ctcaagcaga attcatctct cttccctcaa 2280
tatgatgata ttgacagggt ttgccctcac tcaactagatt gtgagctcct gctcagggca 2340
ggtagcgttt tttgtttttg tttttgtttt tcttttttga gacagggtct tgctctgtca 2400
cccaggccag agtgcaatgg tacagtctca gctcactgca gcctcaaccg cctcggtcca 2460
aaccatcatc ccatttcagc ctctgagta gctgggacta cagggtttgg ccatgttgcc cgggctggtc 2520
tggctaattt ttttgtattt ctagttaga cagggtttgg ccatgttgcc cgggctggtc 2580
tcgaactcct ggactcaagc aatccacca cctcagcctc ccaaaatgag ggaccgtgtc 2640
ttattcattt ccatgtccct agtccatagc ccagtgtgg acctatggta gtactaaata 2700
aatatttgtt gaatgcaata gtaaatagca tttcaggagg caagaactag attaaacaag 2760
gtggttaaa gtttggagaa aaaaataata gtttaatttg gctagagtat gagggagagt 2820
agtaggagac aagatggaaa ggtctcttgg gcaagggttt gaaggaggtt ggaagtcaga 2880
agtacacaat gtgcatatcg tggcaggcag tggggagcca atgaaggctt ttgagcagga 2940
gagtaatgtg ttgaaaaata aatataggtt aaacctatca gagccctct gacacataca 3000
cttgcctttt attcaagctt aagtttgtct cccacatacc cattacttaa ctcacctcg 3060
ggctccccct gcagcctgcc ctacctttt acctgcttcc tgggtggagt agggatgtat 3120
acatgagctg ctttccctct cagccagagg acatgggggg ccccgctcc cctgcctttc 3180
cccttctgtg cctggagctg ggaagcaggc cagggttagc tgaggctggc tggcaagcag 3240
ctgggtgggt ccaggagag cctgcatagt gccaggtggt gccttgggtt ccaagctagt 3300
ccatggcccc gataaccttc tgcctgtgca cacacctgcc cctcactcca ccccatcct 3360
agcttttgta tgggggagag ggcacagggc cagacaaacc tgtgagactt tggctccatc 3420

```

Sekvensliste

tctgcaaaaag	ggcgctctgt	gagtcagcct	gtccccctcc	aggcttgctc	ctccccacc	3480
cagctctcgt	ttccaatgca	cgtacagccc	gtacacaccg	tgtgctggga	cacccacag	3540
tcagccgcat	ggctccccctg	tgccccagcc	cctggctccc	tctgttgatc	ccggccccctg	3600
ctccaggcct	cactgtgcaa	ctgctgctgt	cactgctgct	tctgggtgct	gtccatcccc	3660
agagggttgc	ccggatgcag	gaggattccc	ccttgggagg	aggctcttct	ggggaagatg	3720
acccactggg	cgaggaggat	ctgcccagtg	aagaggattc	accagagag	gaggatccac	3780
ccggagagga	ggatctacct	ggagaggagg	atctacctgg	agaggaggat	ctacctgaag	3840
ttaagcctaa	atcagaagaa	gagggctccc	tgaagttaga	ggatctacct	actgttgagg	3900
ctcctggaga	tcctcaagaa	ccccagaata	atgcccacag	ggacaaagaa	ggtaagtgg	3960
catcaatctc	caaatccagg	ttccaggagg	ttcatgactc	ccctcccata	ccccagccta	4020
ggctctgttc	actcagggaa	ggagggggaga	ctgtactccc	cacagaagcc	cttcagagg	4080
tcccatacca	atatccccat	ccccactctc	ggaggtagaa	agggacagat	gtggagagaa	4140
aataaaaaag	gtgcaaaaag	agagaggtga	gtggatgag	atgggagaga	agggggaggc	4200
tggagaagag	aaagggatga	gaactgcaga	tgagagaaaa	aatgtgcaga	cagaggaaaa	4260
aaataggtgg	agaaggagag	tcagagagtt	tgaggggaag	agaaaaggaa	agcttgggag	4320
gtgaagtggg	taccagagac	aagcaagaag	agctggtaga	agtcattctca	tcttaggcta	4380
caggctggag	ttgcaatgag	ttgcaatgag	acacagcagg	tagagaaacg	tggcttcttg	4440
actcccaagc	caggaatttg	gggaaagggg	ttggagacca	tacaaggcag	agggatgagt	4500
ggggagaaga	aagaaggagg	aaaggaaaga	tggtgtactc	actcatttgg	gactcaggac	4560
tgaagtggcc	actcactttt	tttttttttt	tttttgagac	aaactttcac	ttttgttgcc	4620
caggctggag	tgcaatggcg	cgatctcggc	tcactgcaac	ctccacctcc	cgggttcaag	4680
tgattctcct	gcctcagcct	tagccaagt	agctgcgatt	acaggcatgc	gccaccacgc	4740
ccggctaatt	tttgatattt	tagtagagac	ggggtttcgc	catgttggtc	aggctggtct	4800
cgaactcctg	atctcaggtg	atccaaccac	cctggcctcc	caaagtgtcg	ggattatagg	4860
cgtgacagct	agcgcctggc	ctgaagcagc	cactcacttt	tacagaccct	aagacaatga	4920
ttgcaagctg	gtaggatttg	tggttgcccc	accagctgc	gggttgtagt	ttgggtgctg	4980
tcctctgtgc	tttgacacct	gcccgcctaa	ggcatttgtt	accgtaatg	ctcctgtaag	5040
gcatctgcgt	ttgtgacatc	gttttggtcg	ccaggaaggg	attggggctc	taagcttgag	5100
cggttcatcc	ttttcattta	tacaggggat	gaccagagtc	attggcgcta	tggaggtgag	5160
acacccaccc	gctgcacata	cccaatctgg	gaacccagct	ctgtggatct	cccctacagc	5220
cgctccctgaa	cactggctccc	ggcgctccca	cccgcgccc	accgtcccac	cccctcacct	5280
tttctaccgg	ggttccttaa	gttcctgacc	taggcgctcag	acttcctcac	tatactctcc	5340
caccccgagg	gacccgcctt	ggccccgggt	gtccccagcc	tgcgcgggcc	gcttccagtc	5400
cccgggtggat	atccgcccc	agctcgccgc	ctctcgccc	gcctcgccgc	ccctggaact	5460
cctgggcttc	cagctcccg	cgctcccgga	actgcgctg	cgcaacaatg	gccacagtgg	5520
tgaggggggtc	tccccgcgga	gacttgggga	tgggggcggg	cgaggggaag	ggaaccgtcg	5580
cgagtgctgc	gcccgggggt	tgggctggcc	ctaccgggct	gggcccggct	acttgcctct	5640
ccctacgcag	tgcaactgac	ctcgctcct	tggtctggga	tggtctggg	tcccggcgcg	5700
gagtaccggg	ctctgcagct	gcactgcac	tggggggctg	caggctcgcc	gggctcgag	5760
cacactgtgg	aaggccaccg	tttccctgcc	gaggtgagcg	cggactggcc	gagaaggggc	5820
aaaggagcgg	ggcggacggg	ggccagagac	gtggccctct	cctaccctcg	tgctcttttc	5880
agatccacgt	gggtccacct	agcacccgct	ttggcagagt	tgacgaggcc	ttggggcgcc	5940
cgggaggcct	ggccgtgttg	gccgcctttc	tgagggtacc	agatcctgga	cacccctac	6000
tccccgcttt	cccattcccat	gctcctcccg	gactctatcg	tgagccaga	gaccccatcc	6060
cagcaagctc	actcaggccc	ctggctgaca	aactcattca	cgcactgttt	gttcatttaa	6120
cacccactgt	gaaccagcca	ccagccccca	acaaggattc	tgaagctgta	ggctccttgc	6180
tctaaggagc	ccacagccag	tgggggaggc	tgacatgaca	gacacatagg	aaggacatag	6240
taaagatggt	ggtcacagag	gaggtgacac	ttaaagcctt	cactggtaga	aaagaaaagg	6300
aggtgttcat	tgacagggaa	acagaatgtg	caaagactca	gaatatggcc	tatttaggga	6360
atggctacat	acacagggag	tagagggagc	caggtaaaag	gaagggatgg	tgagatgcct	6420
gctagggttca	ctcactcact	tttattttatt	tattttatttt	tttgacagtc	tctctgtcgc	6480
ccaggctgga	gtgcagtggt	gtgatcttgg	gtcactgcaa	cttcgcctc	ccgggttcaa	6540
gggattctcc	tgctcagct	tcctgagtag	ctgggggttac	aggtgtgtgc	caccatgccc	6600
agctaatttt	tttttgattt	tttagtagac	aggggtttcac	catgttggtc	aggctggtct	6660
caaactcctg	gcctcaagtg	atccgcctga	ctcagcctac	caaagtgtcg	attacaagtg	6720
tgagccaccg	tgcccagcca	cactcactga	ttctttaatg	ccagccacac	agcacaaggt	6780
tcagagaaat	gcctccatca	tagcatgtca	atatgttcat	actcttaggt	tcatgatgtt	6840
cttaacatta	ggttctcata	caaaataaga	aaaaagaata	ataaaataaa	gaagtggcat	6900
gtcaggacct	cacctgaaaa	gccaaacaca	gaatcatgaa	ggtgaatgca	gaggtgacac	6960
caacacaaag	gtgtatatat	ggtttcctgt	ggggagtagt	tacggaggca	gcagtgagt	7020
agactgcaga	cgtcagaagg	ggcaggggtc	ctgagagcct	agtatcctag	taaagtgggc	7080
tctctccctt	ttcttccagc	ttgtcattga	aaaccagctc	accaagcttg	ttgggttcga	7140
cagcaagagt	acatatagagt	tgaaataata	cataggattt	taagaggag	acactgtctc	7200
taaaaaaaaa	aacaacagca	acaacaaaaa	gcaacaacca	ttacaatttt	atgttccctc	7260
agcatttctca	gagctgagga	atgggagagg	actatgggaa	cccccttcat	gttccggcct	7320
tcagccatgg	ccctggatac	atgcactcat	ctgtcttaca	atgtcattcc	cccaggagg	7380
ccgggaagaa	aacagtcct	atgagcagtt	gctgtctcgc	ttggaagaaa	tcgtgagga	7440
aggctcagttt	gttggtctgg	ccactaatct	ctgtggccta	gttcataaag	aatcaccctt	7500

Sekvensliste

tggagcttca	ggtctgaggg	tggagatggg	ctccctccag	tgcaggaggg	attgaagcat	7560
gagccagcgc	tcatcttgat	aataaccatg	aagctgacag	acacagttac	ccgcaaacgg	7620
ctgcctacag	attgaaaacc	aagcaaaaac	cgccggggcac	ggtgggtcac	gcctgtaatc	7680
ccagcacttt	gggaggccaa	ggcagggtga	tcacgagggtc	aagagatcaa	gaccatcctg	7740
gccaacatgg	tgaaacccca	tctctactaa	aaatacgaaa	aaatagccag	gcgtgggtggc	7800
gggtgcctgt	aatcccagct	actcggggagg	ctgaggcagg	agaatggcat	gaacccggga	7860
ggcagaagtt	gcagtgaagc	gagatcgtgc	cactgcactc	cagcctgggc	aacagagcga	7920
gactcttgct	tcaaaaaaaaa	aaaaaaaaaaa	gaaaaccaag	caaaaaccaa	aatgagacaa	7980
aaaaaacaag	accaaaaaat	ggtgttttga	aattgtcaag	gtcaagtctg	gagagctaaa	8040
ctttttctga	gaactgttta	tctttaataa	gcatcaaata	ttttaacttt	gtaaatactt	8100
ttgttggaag	tcgttctctt	cttagtcact	cttgggtcat	tttaaattctc	acttactcta	8160
ctagacttct	taggtttctg	ctagactagg	tagaactctg	cctttgcatt	tcttgtgtct	8220
gtttttgtata	gttatcataa	tctatattta	tttacaagtt	attcagatca	ttttttcttt	8280
tctttttttt	tttttttttt	ttttttacat	cttttagtaga	gacagggttt	caccatattg	8340
gccaggctgc	tctcaaaactc	ctgaccttgt	gatccaccag	cctcggcctc	ccaaagtgtc	8400
gggattcatt	ttttcttttt	aatttgctct	gggcttaaac	ttgtggccca	gcactttatg	8460
atggtagaca	gagtttaagag	gttagactca	gacgggtctt	cttctttcct	tcttctcctt	8520
cctcccttcc	ctccccactt	cccttctctc	cttcttttct	ttcttctctc	cttgcttctc	8580
caggcctctt	ccagttgtct	caaagccctg	tacttttttt	tgagttaacg	tcttatggga	8640
agggcctgca	cttagtgaag	aagtgggtctc	agagttgagt	taccttggct	tctgggaggt	8700
gaaactgtat	ccctataacc	tgaagcttta	agggggtgca	atgtagatga	gacccaaca	8760
tagatcctct	tcacaggctc	agagactcag	gtcccaggac	tgacatatac	tgactcctg	8820
ccctctgact	tcagccgcta	cttccaatat	gaggggtctc	tgactacacc	gccctgtgcc	8880
caggggtgtca	tctggactgt	gtttaaccag	acagtgtatg	tgagtgtctaa	gcagggtggc	8940
ctgggggtgtg	tgtggacaca	gtgggtgcgg	gggaaagagg	atgtaagatg	agatgagaaa	9000
caggagaaga	aagaaatcaa	ggctgggtct	tgtgggttac	gcctataatc	ccaccacgtt	9060
gggaggctga	ggtggggagaa	tgggtttgagc	ccaggagttc	aagacaaggc	ggggcaacat	9120
agtgtgaccc	catctctacc	aaaaaaaaacc	caacaaaacc	aaaaatagcc	gggcatgggtg	9180
gtatgcggcg	tagatccacc	tactcaagga	ggctgagggtg	ggaagatcgc	tgatttccag	9240
gagtttgaga	ctgcagtgaag	ctatgatccc	accactgcct	accatcttta	ggatacattt	9300
attttatttat	aaaagaaatc	aagaggctgg	atggggaata	caggagctgg	aggggtggagc	9360
cctgagggtgc	tggtttgtgag	ctggcctggg	acccttgttt	cctgtcatgc	catgaaccca	9420
cccacactgt	ccactgacct	ccctagctcc	acaccctctc	tgacaccctg	tggggacctg	9480
gtgactctctg	gctacagctg	aacttccgag	gcagcgagcc	tttgaatggg	cgagtgtattg	9540
aggcctcctt	ccctgctgga	gtggacagca	gtcctcgggc	tgtgtgagcca	ggtacagctt	9600
tgtctgtgtt	ccccccagcc	agtagtccct	tatcctccca	tgtgtgtgtcc	agtgtctgtc	9660
attggtgggtc	acagcccgcc	tctcacatct	cctttttctc	tccagtccag	ctgaattcct	9720
gcctggctgc	tggtgagtct	gcccctcctc	ttggctctga	tgccaggaga	ctcctcagca	9780
ccattcagcc	ccagggtctg	tcaggaccgc	ctctgtctcc	tctccttttc	tcgagaacag	9840
acccaaccc	caatattaga	gaggcagatc	atgggtggga	ttccccatt	gtccccagag	9900
gctaatttgat	tagaatgaag	cttgagaaat	ctcccagcat	ccctctcgca	aaagaatccc	9960
ccccctttt	tttaaagata	gggtctcact	ctgtttgccc	caggctgggg	tgttgtggca	10020
cgatcatagc	tcaactgcagc	ctcgaactcc	taggtctcagg	caatcctttc	accttagctt	10080
ctcaaagcac	tgggactgta	ggcatgagcc	actgtgcctg	gccccaaacg	gcccttttac	10140
ttggctttta	ggaagcaaaa	acgggtgctta	tcttaccctt	tctcgtgtat	ccaccctcat	10200
cccttgcctg	gcctcttctg	gagactgagg	cactatgggg	ctgcctgaga	actcggggca	10260
gggggtgggtg	agtgcactga	ggcagggtgtt	gaggaactct	gcagacccct	cttcttctcc	10320
aaagcagccc	tctctgtctt	ccatcgcagg	tgacatccta	gccctgtgtt	ttggcctcct	10380
ttttgtctgtc	accagcgtcg	cgttccttgt	gcagatgaga	aggcagcaca	ggtattacac	10440
tgaccctttc	ttcagggcga	agcttcccc	acccttgtgg	agtcacttca	tgcaaagcgc	10500
atgcaaataga	gctgtctctg	ggccagtttt	ctgattagcc	tttctgtgtg	tgtacacaca	10560
gaaggggaac	caaaggggggt	gtgagctacc	gcccagcaga	ggtagccgag	actggagcct	10620
agaggctgga	tcttggagaa	tgtgagaagc	cagccagagg	catctgaggg	ggagccggta	10680
actgtcctgt	cctgtctcatt	atgccacttc	cttttaactg	ccaagaaatt	ttttaaaata	10740
aatattttata	ataaaaatatg	tgttagtcac	ctttgttccc	caaatacaga	ggaggtattt	10800
gaattttccta	ttactgttat	tagcaccaat	ttagtgttaa	tgcattttatt	ctattacagt	10860
tcggcctcct	tccacacatc	actccaatgt	gttgctcc			10898

<210> 6
 <211> 37
 <212> PRT
 <213> HUMAN

<400> 6
 Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu Ile Pro Ala
 1 5 10 15

Sekvensliste

Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu Leu Leu Leu
 20 25 30

Met Pro Val His Pro
 35

<210> 7
 <211> 25
 <212> DNA
 <213> HUMAN

<400> 7
 tggggttctt gaggatctcc aggag 25

<210> 8
 <211> 26
 <212> DNA
 <213> HUMAN

<400> 8
 ctctaacttc agggagccct cttctt 26

<210> 9
 <211> 48
 <212> DNA
 <213> HUMAN

<220>
 <221> primer_bind
 <222> (1)..(48)
 <400> 9
 cuacuacuac uagggcacgc gtcgactagt acgggnnggg nngggngng 48

<210> 10
 <211> 6
 <212> PRT
 <213> HUMAN

<400> 10
 Glu Glu Asp Leu Pro Ser
 1 5

<210> 11
 <211> 6
 <212> PRT
 <213> HUMAN

<400> 11
 Gly Glu Asp Asp Pro Leu
 1 5

<210> 12
 <211> 21
 <212> PRT
 <213> HUMAN

<400> 12
 Asn Asn Ala His Arg Asp Lys Glu Gly Asp Asp Gln Ser His Trp Arg
 1 5 10 15

Tyr Gly Gly Asp Pro

Sekvensliste

20

<210> 13
<211> 16
<212> PRT
<213> HUMAN

<400> 13
His Pro Gln Arg Leu Pro Arg Met Gln Glu Asp Ser Pro Leu Gly Gly
1 5 10 15

<210> 14
<211> 24
<212> PRT
<213> HUMAN

<400> 14
Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu Glu Asp Leu
1 5 10 15
Pro Gly Glu Glu Asp Leu Pro Gly
20

<210> 15
<211> 13
<212> PRT
<213> HUMAN

<400> 15
Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala Tyr Glu Gln
1 5 10

<210> 16
<211> 16
<212> PRT
<213> HUMAN

<400> 16
Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser Tyr Arg
1 5 10 15

<210> 17
<211> 45
<212> DNA
<213> HUMAN

<400> 17
gtcgtctagct ccatgggtca tatgcagagg ttgccccgga tgcag 45

<210> 18
<211> 43
<212> DNA
<213> HUMAN

<400> 18
gaagatctct tactcgagca ttctccaaga tccagcctct agg 43

<210> 19
<211> 10
<212> DNA
<213> HUMAN

Sekvensliste

<400> 19
 ctccatctct 10

<210> 20
 <211> 10
 <212> DNA
 <213> HUMAN

<400> 20
 ccacccccat 10

<210> 21
 <211> 205
 <212> DNA
 <213> HUMAN

<400> 21
 acctgcccct cactccaccc ccatacctagc ttgggtatgg gggagagggc acagggccag 60
 acaaacctgt gagacttttg ctccatctct gcaaaagggc gctctgtgag tcagcctgct 120
 cccctccagg ctgctcctc cccacccag ctctcgtttc caatgcacgt acagcccgt 180
 cacaccgtgt gctgggacac cccac 205

<210> 22
 <211> 8
 <212> PRT
 <213> HUMAN

<400> 22
 Leu Glu His His His His His His
 1 5

<210> 23
 <211> 10
 <212> DNA
 <213> HUMAN

<220>
 <221> misc_feature
 <222> (1)..(10)

<400> 23
 yyycayyyyy 10

<210> 24
 <211> 10
 <212> DNA
 <213> HUMAN

<300>
 <301> Locker and Buzard,
 <303> DNA Sequencing and Mapping
 <304> 1
 <306> 3-11
 <307> 1990

<400> 24
 tgtgagactt 10

<210> 25
 <211> 4
 <212> PRT

Sekvensliste

<213> HUMAN

<220>

<221> SITE

<222> (1)..(4)

<400> 25

Ser Pro Xaa Xaa

1

<210> 26

<211> 4

<212> PRT

<213> HUMAN

<220>

<221> SITE

<222> (1)..(4)

<400> 26

Thr Pro Xaa Xaa

1

<210> 27

<211> 540

<212> DNA

<213> HUMAN

<220>

<221> promoter

<222> (1)..(540)

<400> 27

cttgcttttc	attcaagctc	aagtttgtct	cccacatacc	cattacttaa	ctcaccctcg	60
ggctcccccta	gcagcctgcc	ctacctcttt	acctgcttcc	tggtggagtc	agggatgtat	120
acatgagctg	ctttccctct	cagccagagg	acatgggggg	ccccagctcc	cctgcctttc	180
cccttctgtg	cctggagctg	ggaagcaggc	cagggtttagc	tgaggctggc	tggaagcag	240
ctgggtgggtg	ccagggagag	cctgcatagt	gccaggtggt	gccttggggt	ccaagctagt	300
ccatggcccc	gataaccttc	tgctgtgca	cacacctgcc	cctcactcca	ccccatcct	360
agcttttgta	tgggggagag	ggcacagggc	cagacaaacc	tgtgagactt	tggtccatc	420
tctgcaaaag	ggcgctctgt	gagtcagcct	gctccccctc	aggcttgctc	ctccccacc	480
cagctctctgt	ttccaatgca	cgtacagccc	gtacacaccg	tgtgctggga	caccacacag	540

<210> 28

<211> 445

<212> DNA

<213> HUMAN

<220>

<221> exon

<222> (1)

<223> 1st MN exon

<400> 28

gcccgtacac	accgtgtgct	gggacacccc	acagtcagcc	gcatggctcc	cctgtgcccc	60
agccccctggc	tccctctggt	gatccccggc	cctgtctccag	gcctcactgt	gcaactgctg	120
ctgtcactgc	tgcttctggt	gcctgtccat	ccccagaggt	tgccccggat	gcaggaggat	180
tcccccttgg	gaggaggctc	ttctggggaa	gatgaccac	tgggagagga	ggatctgccc	240
agtgaagagg	attcaccag	agaggaggat	ccaccggag	aggaggatct	acctggagag	300
gaggatctac	ctggagagga	ggatctacct	gaagttaagc	ctaaatcaga	agaagagggc	360
tccctgaagt	tagaggatct	acctactggt	gaggctcctg	gagatcctca	agaaccccag	420
aataatgccc	acagggacaa	agaag				445

<210> 29

Sekvensliste

<211> 30
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 2nd MN exon

<400> 29
gggatgacca gagtcattgg cgctatggag 30

<210> 30
<211> 171
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 3rd MN exon

<400> 30
gcgacccgcc ctggccccgg gtgtccccag cctgcgcggg ccgcttccag tccccggtgg 60
atatccgccc ccagctcgcc gccttctgcc cggccctgcg cccctggaa ctcctgggct 120
tccagctccc gccgctcca gaactgcgcc tgcgcaacaa tggccacagt g 171

<210> 31
<211> 143
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 4th MN exon

<400> 31
tgcaactgac cctgcctcct gggctagaga tggctctggg tcccgggagg gagtaccggg 60
ctctgcagct gcatctgcac tgggggggctg caggctcgcc gggctcggag cacactgtgg 120
aaggccaccg tttccctgcc gag 143

<210> 32
<211> 93
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 5th MN exon

<400> 32
atccacgtgg ttcacctcag caccgccttt gccagagttg acgaggcctt ggggcgccc 60
ggaggcctgg ccgtgttggc cgcccttctg gag 93

<210> 33
<211> 67
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)

Sekvensliste

<223> 6th MN exon

<400> 33
gagggcccg aagaaaacag tgcctatgag cagttgctgt ctcgcttgga agaaatcgct 60
gaggaag 67

<210> 34
<211> 158
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 7th MN exon

<400> 34
gctcagagac tcagggtccca ggactggaca tatctgcact cctgccctct gacttcagcc 60
gctactttcca atatgagggg tctctgacta caccgccctg tgcccagggt gtcattctgga 120
ctgtgttttaa ccagacagtg atgctgagtg ctaagcag 158

<210> 35
<211> 145
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 8th MN exon

<400> 35
ctccacaccc tctctgacac cctgtgggga cctggtgact ctcggctaca gctgaacttc 60
cgagcgacgc agcctttgaa tgggcgagtg attgaggcct ccttccctgc tggagtggac 120
agcagtcctc gggctgctga gccag 145

<210> 36
<211> 27
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 9th MN exon

<400> 36
tccagctgaa ttctgcctg gctgctg 27

<210> 37
<211> 82
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 10th MN exon

<400> 37
gtgacatcct agccctgggt tttggcctcc tttttgctgt caccagcgtc gcgttccttg 60
tgcagatgag aaggcagcac ag 82

Sekvensliste

<210> 38
<211> 191
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 11th MN exon

<400> 38						
aaggggaacc	aaagggggtg	tgagctaccg	cccagcagag	gtagccgaga	ctggagccta	60
gaggctggat	cttggagaat	gtgagaagcc	agccagaggc	atctgagggg	gagccggtaa	120
ctgtcctgtc	ctgtctatta	tgccacttcc	ttttaactgc	caagaaattt	tttaaaataa	180
atatttataa	t					191

<210> 39
<211> 1174
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(1174)
<223> 1st MN intron

<400> 39						
gtaagtgggtc	atcaatctcc	aatccaggt	tccaggaggt	tcatgactcc	cctcccatat	60
cccagcctag	gctctgttca	ctcaggggaag	gaggggagac	tgtactcccc	acagaagccc	120
ttccagaggt	cccataccaa	tatccccatc	cccactctcg	gaggtagaaa	gggacagatg	180
tggagagaaa	ataaaaagg	tgcaaaaagga	gagaggtgag	ctggatgaga	tgggagagaa	240
gggggagggt	ggagaagaga	aagggatgag	aactgcagat	gagagaaaaa	atgtgcagac	300
agaggaaaaa	aataggtgga	gaaggagagt	cagagagt	gaggggaaga	gaaaaggaaa	360
gcttgggagg	tgaagtgggt	accagagaca	agcaagaaga	gctggtagaa	gtcatctcat	420
cttaggctac	aatgaggaat	tgagacctag	gaagaaggga	cacagcaggt	agagaaacgt	480
ggcttcttga	ctcccaagcc	aggaatttgg	ggaaaggggt	tggagaccat	acaaggcaga	540
gggatgagt	gggagaagaa	agaagggaga	aaggaaagat	ggtgtactca	ctcatttggg	600
actcaggact	gaagtgccca	ctcacttttt	tttttttttt	ttttgagaca	aacttttact	660
tttggtgccc	aggctggagt	gcaatggcgc	gatctcggct	actgcaacc	tccacctccc	720
gggttcaagt	gattctcctg	cctcagcctc	tagccaagta	gctgcgatta	caggcatgct	780
ccaccacgcc	cggctaattt	ttgtattttt	agtagagacg	gggttttcgc	atgttgggtca	840
ggctgggtctc	gaactcctga	tctcaggtga	tccaaccacc	ctggcctccc	aaagtgtctg	900
gattataggc	gtgagccaca	gcgcttgccc	tgaagcagcc	actcactttt	acagacccta	960
agacaatgat	tgcaagctgg	taggattgct	gtttggccca	cccagctgct	gtgttgagtt	1020
tgggtgctggt	ctcctgtgct	ttgcacctgg	cccgtttaag	gcatttgtta	cccgtaatgc	1080
tcctgtaagg	catctgcgtt	tgtgacatcg	ttttggtcgc	caggaaggga	ttggggctct	1140
aagcttgagc	ggttcacatc	tttcatttat	acag			1174

<210> 40
<211> 193
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(193)
<223> 2nd MN intron

<400> 40						
gtgagacacc	cacccgctgc	acagacccaa	tctgggaacc	cagctctgtg	gatctccctt	60
acagccgtcc	ctgaacactg	gtcccgggct	tcccacccgc	cgcccaccgt	cccacccctt	120
caccttttct	acccgggttc	cctaagttcc	tgacctaggt	gtcagacttc	ctcactatac	180
tctccacccc	cag					193

<210> 41

Sekvensliste

<211> 131
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(131)
<223> 3rd MN intron

<400> 41
gtgagggggt ctccccgccg agacttgggg atggggcggg gcgcagggaa gggaaccgtc 60
gcgcagtgcc tgccccgggg ttgggctggc cctaccgggc ggggccggct cacttgcctc 120
tccctacgca g 131

<210> 42
<211> 89
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(89)
<223> 4th MN intron

<400> 42
gtgagcgcgg actggccgag aaggggcaaa ggagcggggc ggacgggggc cagagacgtg 60
gccctctcct accctcgtgt ccttttcag 89

<210> 43
<211> 1400
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(1400)
<223> 5th MN intron

<400> 43
gtaccagatc ctggacaccc cctactcccc gctttcccat cccatgctcc tcccggactc 60
tatcgtggag ccagagaccc catcccagca agctcactca ggccccctggc tgacaaactc 120
attcacgcac tgtttgttca tttaacaccc actgtgaacc aggcaccagc cccaacaacg 180
gattctgaag ctgtagggtcc ttgcctctaa ggagcccaca gccagtgggg gaggttgaca 240
tgacagacac ataggaagga catagtaaaag atgggtgtca cagaggagggt gacacttaaa 300
gccttcactg gtagaaaaga aaaggagggtg ttcattgcag aggaaacaga atgtgcaaag 360
actcagaata tggcctatctt aggggaatggc tacatacacc atgattagag gagggccagt 420
aaaggggaagg gatgggtgaga tgcctgctag gttcactcac tcacttttat ttattttatt 480
atttttttga cagtctctct gtcgcccagg ctggagtgcg gtggtgtgat cttgggtcac 540
tgcaacttcc gcctcccggg ttcaagggat tctcctgcct cagcttcctg agtagctggg 600
gttacagggtg tgtgccacca tgcccagcta attttttttt gtatttttag tagacagggt 660
ttcaccatgt tgggtcaggct ggtctcaaac tcctggcctc aagtgatccg cctgactcag 720
cctaccaaag tgctgattac aagtgtgagc caccgtgccc agccacactc actgattctt 780
taatgccagc cacacagcac aaagttcaga gaaatgcctc catcatagca tgtcaatatg 840
ttcatactct taggttcatt atgttcttaa cattagggttc ataagcaaaa taagaaaaaa 900
gaataataaa taaaagaagt ggcattgtcag gacctcacct gaaaagccaa acacagaatc 960
atgaagggtg atgcagaggt gacaccaaca caaagggtga tatatggttt cctgtgggga 1020
gtatgtacgg aggcagcagt gagtgcagct gcaaacgtca gaagggcacg ggtcactgag 1080
agcctagtag cctagtaaa ggggctctct cctctctctc ccagcttgctc attgaaaacc 1140
agtccaccaa gcttgttgggt tcgcacagca agagtacata gagtttgaaa taatacatag 1200
gattttaaga gggagacact gtctctaaaa aaaaaaaca cagcaacaac aaaaagcaac 1260
aaccattaca attttatgtt cctcagagct tctcagagct gaggaatggg agaggactat 1320
gggaaccccc ttcatgttcc ggccttcagc catggccctg gatacatgca ctcatctgtc 1380
ttacaatgtc attccccag 1400

<210> 44

Sekvensliste

<211> 1334
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(1334)
<223> 6th MN intron

```
<400> 44
gtcagtttgt tggctctggcc actaatctct gtggcctagt tcataaagaa tcaccctttg 60
gagcttcagg tctgaggctg gagatgggct cctccagtg caggaggat tgaagcatga 120
gccagcgctc atcttgataa taaccatgaa gctgacagac acagttaccc gcaaacggct 180
gcctacagat tgaaaaccaa gcaaaaaccg ccgggcacgg tggctcacgc ctgtaatccc 240
agcacttttg gaggccaagg cagggtggatc acgagggtcaa gagatcaaga ccatacctggc 300
caacatggtg aaaccccatc tctactaaaa atacgaaaaa atagccaggc gtggtggcgg 360
gtgctgttaa tcccagctac tcgggaggct gaggcaggag aatggcatga acccgggagg 420
cagaagttgc agtgagccga gatcgtgccg ctgcactcca gcctgggcaa cagagcgaga 480
ctcttgtctc aaaaaaaaaa aaaaaaaaaa aaaccaagca aaaacaaaa tgagacaaaa 540
aaaacaagac caaaaaatgg tgtttggaat ttgtcaaggt caagtctgga gagctaaact 600
ttttctgaga actgtttatc ttaataaagc atcaaatatt ttaactttgt aaatactttt 660
gttggaatc gtctcttct tagtcactct tgggtcattt taaatctcac ttactctact 720
agacctttta ggtttctgct agactaggta gaactctgcc ttgcatcttc ttgtgtctgt 780
tttgatatag tatcaatatt catatttatt tacaagttat tcagatcatt ttttcttttc 840
tttttttttt tttttttttt ttttacctct ttagtagaga cagggtttca ccataattggc 900
caggctgtct tcaaacctct gaccttgtga tccaccagcc tcggcctccc aaagtgtctg 960
gattcatttt ttctttttaa ttgtctctgg gcttaaacct gtggcccagc actttatgat 1020
ggtacacaga gttaagagtg tagactcaga cggctcttct tctttccttc tcttccttcc 1080
tcccttccct cccaccttcc ctctctctct tcttttcttt ctctctctct tgcttcttca 1140
ggcctcttcc agttgctcca aagccctgta cttttttttg agttaacgtc ttatgggaag 1200
ggcctgcact tagtgaagaa gtggtctcag agttgagtta ccttggcttc tgggaggtga 1260
aactgtatcc ctataccctg aagctttaag ggggtgcaat gtagatgaga cccaacata 1320
gatcctcttc acag                                     1334
```

<210> 45
<211> 512
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(512)
<223> 7th MN intron

```
<400> 45
gtgggcctgg ggtgtgtgtg gacacagtgg gtgcggggga aagaggatgt aagatgagat 60
gagaaacagg agaagaaaga aatcaaggct gggctctgtg gcttacgcct ataatcccac 120
cacgttggga ggctgagggt ggagaatggg ttgagcccag gagttcaaga caaggcgggg 180
caacatagtg tgaccccatc tctaccaaaa aaaccccaac aaaacaaaa atagccgggc 240
atggtggtat gcggcctagt cccagctact caaggaggct gaggtgggaa gatcgcttga 300
ttccaggagt ttgagactgc agtgagctat gatcccacca ctgcctacca tctttaggat 360
acattttatt atttataaaa gaaatcaaga ggctggatgg ggaatacagg agctggaggg 420
tggagccctg aggtgctggg tgtgagctgg cctgggacct ttgtttcctg tcatgccatg 480
aaccaccca cactgtccac tgacctccct ag                                     512
```

<210> 46
<211> 114
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(114)
<223> 8th MN intron

<400> 46

Sekvensliste

gtacagcttt gtctggtttc cccccagcca gtagtccctt atcctcccat gtgtgtgcca 60
gtgtctgtca ttgggtgtca cagcccgctt ctcacatctc ctttttctct ccag 114

<210> 47
<211> 617
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(617)
<223> 9th MN intron

<400> 47
gtgagtctgc ccttcctctt ggtcctgatg ccaggagact cctcagcacc attcagcccc 60
agggctgctc aggaccgcct ctgctccctc tccttttctg cagaacagac cccaacccca 120
atattagaga ggcagatcat ggtggggatt cccccattgt cccagagagc taattgatta 180
gaatgaagct tgagaaatct cccagcatcc ctctcgcaaa agaatcccc cccctttttt 240
taaagatagg gtctcactct gtttgcccca ggctgggggtg ttgtggcacg atcatagctc 300
actgcagcct cgaactccta ggctcaggca atcctttcac cttagcttct caaagcactg 360
ggactgtagg catgagccac tgtgcctggc cccaaacggc ctttttactt ggcttttagg 420
aagcaaaaac ggtgcttata ttaccccttc tcgtgtatcc accctcatcc cttggctggc 480
ctcttctgga gactgaggca ctatggggct gcctgagaac tcggggcagg ggtgggtggag 540
tgcactgagg caggtgttga ggaactctgc agacccctct tccttcccaa agcagccctc 600
tctgctctcc atcgag 617

<210> 48
<211> 130
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(130)
<223> 10th MN intron

<400> 48
gtattacact gaccctttct tcaggcacia gcttccccca cccttgtgga gtcacttcat 60
gcaaagcgca tgcaaatgag ctgctcctgg gccagtttct tgattagcct ttcctgttgt 120
gtacacacag 130

<210> 49
<211> 1401
<212> DNA
<213> HUMAN

<400> 49
caaactttca cttttgttgc ccaggctgga gtgcaatggc gcgatctcgg ctcaactgcaa 60
cctccacctc ccgggttcaa gtgattctcc tgcctcagcc tctagccaag tagctgcat 120
tacaggcatg cgccaccag cccggctaata ttttgtattt ttagtagaga cggggtttcg 180
ccatgttggg caggctgggt tcgaactcct gatctcaggt gatccaacca ccttggcctc 240
ccaaagtgcg gggattatag gcgtgagcca cagcgcttgg cctgaagcag ccactcactt 300
ttacagacc taagacaatg attgcaagct ggtaggattg ctgtttggcc caccagctg 360
cgggtgttgg tttgggtgcg gtctcctgtg ctttgcacct ggcccgtta aggcatttgt 420
taccgttaat gctcctgtaa ggcatctgcg tttgtgacat cgttttggtc gccaggaagg 480
gattggggct ctaagcttga gcggttcatc cttttcattt atacagggga tgaccagagt 540
cattggcgct atggaggtga gacaccacc cgctgcacag acccaatctg ggaaccagc 600
tctgtggatc tcccctacag ccgtccctga acactggctc cgggctccc acccgccgc 660
caccgtcca cccctcacc ttttctacc gggttcccta agttcctgac ctaggcgta 720
gacttctca ctatactct ccacccagc cgacccgccc tggccccggg tgtccccagc 780
ctgcgcgggc cgcttccagt ccccggtgga tatccgccc cagctcgcg ctttctgccc 840
ggccctgcgc cccctggaac tcctgggctt ccagctccc cgcctcccag aactgcgcct 900
gcgcaacaat ggccacagt gtgaggggggt ctccccgcg agacttgggg atggggcggg 960
gcgaggggaa gggaaccgtc gcgcagtgc tgcccggggg ttgggctggc cctaccggc 1020
ggggccggct cacttgcctc tccctacgca gtgcaactga ccctgcctcc tgggctagag 1080

Sekvensliste

atggctctgg	gtcccggg	ggagtaccgg	gctctgcagc	tgcattctgca	ctggggggct	1140
gcaggtcgtc	cgggctcggg	gcacactgtg	gaaggccacc	gtttccctgc	cgaggtgagc	1200
gcggactggc	cgagaagggg	caaaggagcg	gggcggacgg	gggccagaga	cgaggccctc	1260
tcctaccctc	gtgtcctttt	cagatccacg	tggttcacct	cagcaccgcc	tttgccagag	1320
ttgacgaggc	cttggggcgc	ccgggaggcc	tggccgtgtt	ggccgccttt	ctggaggtac	1380
cagatcctgg	acaccccccta	c				1401

<210> 50
 <211> 59
 <212> PRT
 <213> HUMAN

<400> 50
 Ser Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp Leu Pro Ser Glu
 1 5 10 15
 Glu Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu Glu Asp Leu Pro
 20 25 30
 Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Glu Val Lys Pro
 35 40 45
 Lys Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu
 50 55

<210> 51
 <211> 257
 <212> PRT
 <213> HUMAN

<400> 51
 Gly Asp Asp Gln Ser His Trp Arg Tyr Gly Gly Asp Pro Pro Trp Pro
 1 5 10 15
 Arg Val Ser Pro Ala Cys Ala Gly Arg Phe Gln Ser Pro Val Asp Ile
 20 25 30
 Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala Leu Arg Pro Leu Glu Leu
 35 40 45
 Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu Leu Arg Leu Arg Asn Asn
 50 55 60
 Gly His Ser Val Gln Leu Thr Leu Pro Pro Gly Leu Glu Met Ala Leu
 65 70 75 80
 Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln Leu His Leu His Trp Gly
 85 90 95
 Ala Ala Gly Arg Pro Gly Ser Glu His Thr Val Glu Gly His Arg Phe
 100 105 110
 Pro Ala Glu Ile His Val Val His Leu Ser Thr Ala Phe Ala Arg Val
 115 120 125
 Asp Glu Ala Leu Gly Arg Pro Gly Gly Leu Ala Val Leu Ala Ala Phe
 130 135 140
 Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala Tyr Glu Gln Leu Leu Ser
 145 150 155 160
 Arg Leu Glu Glu Ile Ala Glu Glu Gly Ser Glu Thr Gln Val Pro Gly
 165 170 175
 Leu Asp Ile Ser Ala Leu Leu Pro Ser Asp Phe Ser Arg Tyr Phe Gln
 180 185 190

Sekvensliste

Tyr Glu Gly Ser Leu Thr Thr Pro Pro Cys Ala Gln Gly Val Ile Trp
195 200 205
Thr Val Phe Asn Gln Thr Val Met Leu Ser Ala Lys Gln Leu His Thr
210 215 220
Leu Ser Asp Thr Leu Trp Gly Pro Gly Asp Ser Arg Leu Gln Leu Asn
225 230 235 240
Phe Arg Ala Thr Gln Pro Leu Asn Gly Arg Val Ile Glu Ala Ser Phe
245 250 255

Pro

<210> 52
<211> 20
<212> PRT
<213> HUMAN

<400> 52
Ile Leu Ala Leu Val Phe Gly Leu Leu Phe Ala Val Thr Ser Val Ala
1 5 10 15

Phe Leu Val Gln
20

<210> 53
<211> 25
<212> PRT
<213> HUMAN

<400> 53
Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser Tyr Arg
1 5 10 15

Pro Ala Glu Val Ala Glu Thr Gly Ala
20 25

<210> 54
<211> 59
<212> PRT
<213> HUMAN

<400> 54
Ser Ala Ser Glu Glu Pro Ser Pro Ser Glu Val Pro Phe Pro Ser Glu
1 5 10 15

Glu Pro Ser Pro Ser Glu Glu Pro Phe Pro Ser Val Arg Pro Phe Pro
20 25 30

Ser Val Val Leu Phe Pro Ser Glu Glu Pro Phe Pro Ser Lys Glu Pro
35 40 45

Ser Pro Ser Glu Glu Pro Ser Ala Ser Glu Glu
50 55

<210> 55
<211> 470
<212> RNA
<213> HUMAN

<400> 55

Sekvensliste

cauggccccc	auaaccuucu	gccugugcac	acaccugccc	cucacuccac	ccccauccua	60
gcuuugguau	gggggagagg	gcacagggcc	agacaaaccu	gugagacuuu	ggcuuccaucu	120
cugcaaaagg	gcgucucugug	agucagccug	cucuccucca	ggcuugcucc	uccccacccc	180
agcucucguu	uccaaugcac	guacagcccc	uacacaccgu	gugcugggac	acccccacagu	240
cagccgcaug	gcucuccugug	gccccagccc	cuggcucccu	cuguugaucc	cgccccugc	300
uccagggcuc	acugugcaac	ugcugcuguc	acugcugcuu	cuggugccug	uccaucccca	360
gagguugccc	cggaugcagg	aggauucccc	cuugggagga	ggcucuucug	gggaagauga	420
cccacugggc	gaggaggau	ugcccaguga	agaggauuca	cccagagagg		470

<210> 56
 <211> 292
 <212> DNA
 <213> HUMAN

<400> 56						
gtttttttga	gacggagtct	tgcattctgtc	atgcccaggc	tggagtagca	gtggtgccat	60
ctcggctcac	tgaagctcc	acctcccag	ttcacgccat	tttcctgcct	cagcctccc	120
agtagctggg	actacaggcg	cccgccacca	tggccggcta	attttttgta	tttttggtag	180
agacggggtt	tcaccgtgtt	agccagaatg	gtctcgatct	cctgacttcg	tgatccaccc	240
gcctcggcct	cccaaagtcc	tgggattaca	ggtgtgagcc	accgcacctg	gc	292

<210> 57
 <211> 262
 <212> DNA
 <213> HUMAN

<400> 57						
tttctttttt	gagacaggg	cttgctctgt	cacccaggcc	agagtgcatt	ggtacagtct	60
cagctcactg	cagcctcaac	cgctctggct	caaaccatca	tccattttca	gcctcctgag	120
tagctgggac	tacaggcaca	tgccattaca	cctgggcta	ttttttgtat	ttctagtaga	180
gacaggggtt	ggccatgttg	cccgggctgg	tctcgaactc	ctggactcaa	gcaatccacc	240
cacctcagcc	tcccaaaatg	ag				262

<210> 58
 <211> 2501
 <212> DNA
 <213> HUMAN

<220>
 <221> misc_feature
 <222> (1)..(2501)

<400> 58						
tggtgactcg	tgaccttacc	cccaaccctg	tgctctctga	aacatgagct	gtgtccactc	60
aggggttaaat	ggattaaggg	cggtgcaaga	tggtgctttg	ttaaagagatg	cttgaaggca	120
gcatgtctct	taagagtcac	caccaatccc	taattctcaag	taatcaggga	cacaaacact	180
gcggaaggcc	gcagggtcct	ctgcctagga	aaaccagaga	cctttgttca	cttgtttatc	240
tgaccttccc	tccactattg	tccatgaccc	tgccaaatcc	ccctctgtga	gaaacaccca	300
agaattatca	ataaaaaaat	aaattttaaaa	aaaaaatata	aaaaaaaaaa	aaaaaaaaaa	360
aaaagactta	cgaatagtta	ttgataaatg	aatagctatt	ggtaaagcca	agtaaatgat	420
catattcaaa	accagacggc	catcatcaca	gctcaagtct	acctgatttg	atctctttat	480
cattgtcatt	ctttggattc	actagattag	tcatcatcct	caaaattctc	ccccagttc	540
taattacgtt	ccaaacattt	aggggttaca	tgaagcttga	acctactacc	ttctttgctt	600
ttgagccatg	agttgttagga	atgatgagtt	tacaccttac	atgctgggga	tttaatttaa	660
ctttacctct	aagtcagttg	ggtagccttt	ggcttatttt	tgtagcta	tttgtagtta	720
atggatgcac	tgtgaatctt	gctatgatag	ttttcctcca	cactttgcca	ctaggggtag	780
gtaggtactc	agttttcagt	aattgcttac	ctaagaccct	aagccctatt	tctctgtgac	840
tggcctttat	ctgtaaatag	ggcatattta	atacaatata	atttttggag	tttttttgtt	900
tggtttgttg	tttgtttttt	ttgagcggag	ttctgcatct	gtcatgccca	ggctggagta	960
gcagtgggtg	catctcggct	cactgcaagc	tccacctccc	gagttcacgc	cattttcctg	1020
cctcagcctc	ccgagtagct	gggactacag	gcgcccgcga	ccatgcccgg	ctaatttttt	1080
gtatttttgg	tagagacggg	gtttcaccgt	gttagccaga	atgggtctcg	tctcctgact	1140
tcgtgatcca	ccgcctcgg	cctcccaaa	ttctgggatt	acaggtgtga	gccaccgcac	1200
ctggccaatt	ttttgagctc	tttaaagtaa	aaatatgtct	tgtaagctgg	taactatggt	1260
acatttcctt	ttattaatgt	ggtgctgacg	gtcatatagg	ttcttttgag	tttggcatgc	1320

Sekvensliste

atatgctact	ttttgcagtc	ctttcattac	atTTTTtctct	cttcatttga	agagcatgtt	1380
atatctttta	gcttcacttg	gcttaaaagg	ttctctcatt	agcctaacac	agtgtcattg	1440
ttggtaccac	ttggatcata	agtggaaaaa	cagtcagaa	attgcacagt	aatacttggt	1500
tgtaagaggg	atgattcagg	tgaatctgac	actaagaaac	tcccttacct	gaggctctgag	1560
attcctctga	cattgctgta	tataggcctt	tcctttgaca	gcctgtgact	gcggactatt	1620
tttcttaagc	aagatatgct	aaagttttgt	gagccttttt	ccagagagag	gtctcatatc	1680
tgcatcaagt	gagaacatat	aatgtctgca	tgtttccata	tttcaggaat	gtttgcttgt	1740
gttttatgct	tttatataga	cagggaaaact	tgttcctcag	tgacccaaaa	gagggtggaa	1800
ttgttattgg	atatcatcat	tggcccacgc	tttctgacct	tggaacaat	taagggttca	1860
taatctcaat	tctgtcagaa	ttggtacaag	aaatagctgc	tatgtttctt	gacattccac	1920
ttggtaggaa	ataagaatgt	gaaactcttc	agttggtgtg	tgctccctngt	ttttttgcaa	1980
tttccttctt	actgtgttaa	aaaaaagtat	gatcttgctc	tgagaggtga	ggcattctta	2040
atcatgatca	ttaaagatca	ataatataat	attatgtctt	tattataata	tattataata	2100
aagataattt	gtctttaaca	gaatcaataa	tataatccct	taaaggatta	tatctttgct	2160
gggcgagtg	gctcacacct	gtaatcccag	cactttgggt	ggccaagggtg	gaaggatcaa	2220
atttgccctac	ttctatatta	tcttctaaag	cagaattcat	ctctcttccc	tcaatatgat	2280
gatattgaca	ggggtttgcc	tcactcacta	gattgtgagc	tcctgtcag	ggcaggtagc	2340
gttttttgtt	tttgtttttg	tttttctttt	ttgagacagg	gtcttgctct	gtcaccagc	2400
ccagagtgc	atggtacagt	ctcagctcac	tgagcctca	accgcctcgg	ctcaaaccat	2460
catcccat	cagcctcctg	agtagctggg	actacaggca	c		2501

<210> 59
 <211> 292
 <212> DNA
 <213> HUMAN

<220>
 <221> misc_feature
 <222> (1)

<400> 59						
ttttttttgag	acggagtcctt	gcattctgtca	tgcccaggct	ggagtagcag	tggtgccatc	60
tcggctcact	gcaagctcca	cctcccaggt	tcacgccatt	ttcctgcctc	agcctcccga	120
gtagctggga	ctacaggcgc	ccgccaccat	gcccggctaa	ttttttgtat	ttttggtaga	180
gacgggggtt	caccgtgtta	gccagaatgg	tctcgatctc	ctgacttcgt	gatccacccg	240
cctcggcctc	ccaaagttct	gggattacag	gtgtgagcca	ccgcacctgg	cc	292

<210> 60
 <211> 262
 <212> DNA
 <213> HUMAN

<400> 60						
ttctttttttg	agacagggtc	ttgctctgtc	acccaggcca	gagtgcattg	gtacagtctc	60
agctcactgc	agcctcaacc	gcctcggctc	aaaccatcat	cccatttcag	cctcctgagt	120
agctgggact	acaggcacat	gccattacac	ctgggctaatt	ttttttgtatt	tctagtagag	180
acagggtttg	gccatgttgc	ccgggctggg	ctcgaactcc	tggaactcaag	caatccaccc	240
acctcagcct	cccaaaatga	gg				262

<210> 61
 <211> 294
 <212> DNA
 <213> HUMAN

<400> 61						
ttttttttttg	agacaaactt	tcacttttgt	tgcccaggct	ggagtgcatt	ggcgcgatct	60
cggctcactg	caacctccac	ctcccgggtt	caagtgattc	tcctgcctca	gcctctagcc	120
aagtagctgc	gattacaggc	atgcgccacc	acgcccggct	aatttttcta	tttttagtag	180
agacgggggt	tcgccatgtt	ggtcaggctg	gtctcgaact	cctgatctca	ggtgatccaa	240
ccaccctggc	ctcccaaagt	gctgggatta	taggcgtgag	ccacagcgcc	tggc	294

<210> 62
 <211> 276
 <212> DNA

Sekvensliste

<213> HUMAN

<400> 62

tgacagtctc	tctgtcgccc	aggctggagt	gcagtgggtg	gatcttgggt	cactgcaact	60
tccgcctccc	gggttcaagg	gattctcctg	cctcagcttc	ctgagtagct	ggggttacag	120
gtgtgtgcc	ccatgcccag	ctaatttttt	tttgtatttt	tagtagacag	ggtttcacca	180
tgttgggtcag	gctgggtctca	aactcctggc	ctcaagtgat	ccgcctgact	cagcctacca	240
aagtgtctgat	tacaagtgtg	agccaccgtg	cccagc			276

<210> 63

<211> 289

<212> DNA

<213> HUMAN

<400> 63

cgccggggc	ggtggctcac	gcctgtaatc	ccagcacttt	gggaggccaa	ggcaggtgga	60
tcacgaggtc	aagagatcaa	gaccatcctg	gccaacatgg	tgaaacccca	tctctactaa	120
aaatacga	aaatagccag	gcgtgggtggc	gggtgcctgt	aatcccagct	actcgggagg	180
ctgaggcagg	agaatggcat	gaacccgga	ggcagaagtt	gcagttagcc	gagatcgtgc	240
cactgcactc	cagcctgggc	aacagagcga	gactcttgtc	tcaaaaaaa		289

<210> 64

<211> 298

<212> DNA

<213> HUMAN

<400> 64

aggctgggct	ctgtggctta	cgcttataat	cccaccacgt	tgggaggctg	agggtgggaga	60
atggttttgag	cccaggagtt	caagacaagg	cggggcaaca	tagtgtgacc	ccatctctac	120
caaaaaaac	ccaacaaaac	caaaaatagc	cgggcatggt	ggtatgcggc	ctagtcccag	180
ctactcaagg	aggctgaggt	gggaagatcg	cttgattcca	ggagtttgag	actgcagtga	240
gctatgatcc	caccactgcc	taccatcttt	aggatacatt	tatttattta	taaaagaa	298

<210> 65

<211> 105

<212> DNA

<213> HUMAN

<400> 65

ttttttacat	ctttagtaga	gacaggggtt	caccatattg	gccaggctgc	tctcaaactc	60
ctgacctgtg	gatccaccag	cctcggcctc	ccaaagtgtc	gggat		105

<210> 66

<211> 83

<212> DNA

<213> HUMAN

<400> 66

cctcgaactc	ctaggctcag	gcaatccttt	caccttagct	tctcaaagca	ctgggactgt	60
aggcatgagc	cactgtgcct	ggc				83

<210> 67

<211> 11

<212> DNA

<213> HUMAN

<400> 67

agaaggtg	t					11
----------	---	--	--	--	--	----

<210> 68

<211> 11

<212> DNA

Sekvensliste

<213> HUMAN	
<400> 68 tggaggtgag a	11
<210> 69 <211> 11 <212> DNA <213> HUMAN	
<400> 69 cagtcgtgag g	11
<210> 70 <211> 11 <212> DNA <213> HUMAN	
<400> 70 ccgaggtgag c	11
<210> 71 <211> 11 <212> DNA <213> HUMAN	
<400> 71 tggaggtacc a	11
<210> 72 <211> 11 <212> DNA <213> HUMAN	
<400> 72 ggaaggtcag t	11
<210> 73 <211> 11 <212> DNA <213> HUMAN	
<400> 73 agcaggtggg c	11
<210> 74 <211> 11 <212> DNA <213> HUMAN	
<400> 74 gccaggtaca g	11
<210> 75 <211> 11 <212> DNA <213> HUMAN	
<400> 75 tgctggtgag t	11

Sekvensliste

<210> 76	
<211> 11	
<212> DNA	
<213> HUMAN	
<400> 76	
cacaggtatt a	11
<210> 77	
<211> 11	
<212> DNA	
<213> HUMAN	
<400> 77	
atacagggga t	11
<210> 78	
<211> 11	
<212> DNA	
<213> HUMAN	
<400> 78	
ccccaggcga c	11
<210> 79	
<211> 11	
<212> DNA	
<213> HUMAN	
<400> 79	
acgcagtgca a	11
<210> 80	
<211> 11	
<212> DNA	
<213> HUMAN	
<400> 80	
tttcagatcc a	11
<210> 81	
<211> 11	
<212> DNA	
<213> HUMAN	
<400> 81	
ccccaggagg g	11
<210> 82	
<211> 11	
<212> DNA	
<213> HUMAN	
<400> 82	
tcacaggctc a	11
<210> 83	
<211> 11	
<212> DNA	
<213> HUMAN	

Sekvensliste

<400> 83
ccctagctcc a 11

<210> 84
<211> 11
<212> DNA
<213> HUMAN

<400> 84
ctccagtcca g 11

<210> 85
<211> 12
<212> DNA
<213> HUMAN

<400> 85
tcgcaggtga ca 12

<210> 86
<211> 11
<212> DNA
<213> HUMAN

<400> 86
acacagaagg g 11

<210> 87
<211> 377
<212> PRT
<213> HUMAN

<400> 87
Gln Arg Leu Pro Arg Met Gln Glu Asp Ser Pro Leu Gly Gly Gly Ser
1 5 10 15
Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp Leu Pro Ser Glu Glu
20 25 30
Asp Ser Pro Arg Gln Glu Asp Pro Pro Gly Glu Glu Asp Leu Pro Gly
35 40 45
Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Glu Val Lys Pro Lys
50 55 60
Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu Asp Leu Pro Thr Val Glu
65 70 75 80
Ala Pro Gly Asp Pro Gln Glu Pro Gln Asn Asn Ala His Arg Asp Lys
85 90 95
Glu Gly Asp Asp Gln Ser His Trp Arg Tyr Gly Gly Asp Pro Pro Trp
100 105 110
Pro Arg Val Ser Pro Ala Cys Ala Gly Arg Phe Gln Ser Pro Val Asp
115 120 125
Ile Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala Leu Arg Pro Leu Glu
130 135 140
Leu Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu Leu Arg Leu Arg Asn
145 150 155 160

Sekvensliste

Asn Gly His Ser Val Gln Leu Thr Leu Pro Pro Gly Leu Glu Met Ala
165 170 175

Leu Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln Leu His Leu His Trp
180 185 190

Gly Ala Ala Gly Arg Pro Gly Ser Glu His Thr Val Glu Gly His Arg
195 200 205

Phe Pro Ala Glu Ile His Val Val His Leu Ser Thr Ala Phe Ala Arg
210 215 220

Val Asp Glu Ala Leu Gly Arg Pro Gly Gly Leu Ala Val Leu Ala Ala
225 230 235 240

Phe Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala Tyr Glu Gln Leu Leu
245 250 255

Ser Arg Leu Glu Glu Ile Ala Glu Glu Gly Ser Glu Thr Gln Val Pro
260 265 270

Gly Leu Asp Ile Ser Ala Leu Leu Pro Ser Asp Phe Ser Arg Tyr Phe
275 280 285

Gln Tyr Glu Gly Ser Leu Thr Thr Pro Pro Cys Ala Gln Gly Val Ile
290 295 300

Trp Thr Val Phe Asn Gln Thr Val Met Leu Ser Ala Lys Gln Leu His
305 310 315 320

Thr Leu Ser Asp Thr Leu Trp Gly Pro Gly Asp Ser Arg Leu Gln Leu
325 330 335

Asn Phe Arg Ala Thr Gln Pro Leu Asn Gly Arg Val Ile Glu Ala Ser
340 345 350

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

Leu Asn Ser Cys Leu Ala Ala Gly Asp
370 375

<210> 88
<211> 34
<212> DNA
<213> HUMAN

<400> 88
tagacagatc tacgatggct cccctgtgcc ccag 34

<210> 89
<211> 34
<212> DNA
<213> HUMAN

<400> 89
attcctctag acagttaccg gctccccctc agat 34

<210> 90
<211> 3532
<212> DNA
<213> HUMAN

<220>
<221> misc_feature

Sekvensliste

<222> (1)..(3532)

<400> 90

tggtgactcg	tgaccttacc	cccaaccctg	tgctctctga	aacatgagct	gtgtccactc	60
agggttaaat	ggattaaggg	cggtgcaaga	tggtgcttgt	taaacagatg	cttgaaggca	120
gcatgctcgt	taagagtcac	caccaatccc	taatctcaag	taatcaggga	cacaaacact	180
gcggaaggcc	gcagggtcct	ctgcctagga	aaaccagaga	cctttgttca	cttgtttatc	240
tgaccttccc	tccactattg	tccatgacct	tgccaaatcc	ccctctgtga	gaaacaccca	300
agaattatca	ataaaaaaat	aaattttaaa	aaaaaatata	aaaaaaaaaa	aaaaaaaaaa	360
aaaagactta	cgaatagtta	ttgataaatg	aatagctatt	ggtaaagcca	agtaaagtat	420
catattcaaa	accagacggc	catcatcaca	gctcaagtct	acctgatttg	atctctttat	480
cattgtcatt	ctttggattc	actagattag	tcatcatcct	caaaattctc	ccccaagttc	540
taattacggt	ccaaacattt	aggggttaca	tgaaagcttg	acctactacc	ttctttgctt	600
ttgagccatg	agttgttaga	atgatgagtt	tacaccttac	atgctgggga	ttaatttaaa	660
ctttacctct	aagtcagttg	ggtagccttt	ggcttatttt	tgtagctaata	ttttagtgta	720
atggaatgcac	tgtagaatctt	gctatgatag	ttttcctcca	cactttgcca	ctaggggtag	780
gtaggtatca	agttttcagt	ataagacctt	tgaaagacct	aagccctatt	tctctgtgac	840
tgccctttat	ctgtaatatg	ggcatattta	atacaatata	atttttggag	tttttttgtt	900
tgtttggttg	tttggttttt	tgagacggag	tcttgcatct	gtcatgcccc	ggctggagta	960
gcagtgggtg	catctcggct	cactgcaagc	tccacctccc	gagttcacgc	cattttcctg	1020
cctcagcctc	ccgagtagct	gggactacag	gcgcccgcga	ccatgcccgg	ctaatttttt	1080
gtatttttgg	tagagacggg	gtttcaccgt	gttagccaga	atgggtctga	tctcctgact	1140
tcgtgatcca	cccgcctcgg	cctcccaaag	tcttgggatt	acaggtgtga	gccaccgcac	1200
ctggccaatt	ttttgagctt	tttaaagtaa	aaatatgtct	tgtaagctgg	taactatggt	1260
acattttcctt	ttattaatgt	gggtgctgac	gtcataatag	ttcttttgag	tttggcatgc	1320
atatgtctact	ttttgcagtc	ctttcattac	atttttctct	cttcatttga	agagcatgtt	1380
atatctttta	gcttcacttg	gcttaaaagg	ttctctcatt	agcctaacac	agtgtcattg	1440
ttggtaccac	ttggatcata	agtggaaaaa	cagtcaagaa	attgcacagt	aatacttgtt	1500
tgtaagaggg	atgattcagg	tgaatctgac	actaagaaac	tcccctacct	gaggtctgag	1560
attcctctga	cattgtctga	tataggtctt	tcctttgaca	gcctgtgact	gcggactatt	1620
tttcttaagc	aagatatgct	aaagttttgt	gagccttttt	ccagagagag	gtctcatatc	1680
tgcatcaagt	gagaacatat	aatgtctgca	tgtttccata	tttcaggaat	gtttgcttgt	1740
gttttatggt	tttatataga	cagggaacct	tgttcctcag	tgacccaaaa	gaggtgggaa	1800
ttgttatgtg	atatcatcat	tggcccacgc	ttcttgacct	tggaacaat	taagggttca	1860
taatctcaat	tctgtcagaa	ttggtacaag	aaatagctgc	tatgtttctt	gacattccac	1920
ttggtagga	ataagaatgt	gaaactcttc	agttgggtgt	tgctccctngt	ttttttgcaa	1980
tttccttctt	actgtgttaa	aaaaaagtat	gatcttgctc	tgagaggtga	ggcattctta	2040
atcatgatct	tttaagatca	ataatataat	cctttcaagg	attatgtctt	tattataata	2100
aagataattt	gtctttaaca	gaatcaataa	tataatccct	taaaggatta	tatctttgct	2160
gggcgagtg	gctcacacct	gtaatcccag	cactttgggt	ggccaagggt	gaaggatcaa	2220
atttgccctac	ttctatatta	tccttctaaag	cagaattcat	ctctcttccc	tcaatatgat	2280
gattttgaca	gggtttgccc	tcactcacta	gattgtgagc	tcctgtctcag	ggcaggtagc	2340
gttttttgtt	tttggttttt	tttttctttt	ttgagacagg	gtcttgctct	gtcaccacag	2400
ccagagtgc	atggtacagt	ctcagctcac	tgacgcctca	accgcctcgg	ctcaaaccat	2460
catcccatct	cagcctcctg	agtagctggg	actacaggca	catgccatta	cacctggcta	2520
atttttttgt	atttctagta	gagacagggt	ttggccatgt	tgcccgggct	ggtctcgaac	2580
tcctggactc	aagcaatcca	cccacctcag	cctcccaaaa	tgagggaccg	tgtcttattc	2640
atttccatgt	ccctagtcga	tagcccagtg	ctggaccctat	ggtagtacta	aataaatatt	2700
tggtgaatgc	aatagttaa	agcattttcag	ggagcaagaa	ctagattaac	aaagggtgga	2760
aaagggttgg	agaaaaaaat	aatagtttaa	tttggtctaga	gtatgaggga	gagtagtagg	2820
agacaagatg	gaaaggctct	ttgggcaagg	ttttgaagga	agttggaagt	cagaagtaca	2880
caatgtgcat	atcgtggcag	gcagtgggga	gccaatgaag	gcttttgagc	aggagagtaa	2940
tgtgttgaaa	aataaatata	ggttaaacct	atcagagccc	ctctgacaca	tacacttgct	3000
tttcattcaa	gctcaagttt	gtctcccaca	taccattac	ttactcacc	ctcgggctcc	3060
cctagcagcc	tgccctacct	ctttacctgc	ttctgggtgg	agtcagggat	gtatacatga	3120
gctgctttcc	ctctcagcca	gaggacatgg	ggggccccag	ctcccctgcc	tttccccttc	3180
tggtgcttga	gctgggaagc	aggccagggt	tagctgaggc	tggtgggcaa	gcagctgggt	3240
gggtgccagg	agagcctgca	tagtgccagg	tggtgccttg	ggttccaagc	tagtccatgg	3300
ccccgataac	cttctgcctg	tgacacaccc	tgcccctcac	tccaccccca	tcctagcttt	3360
gggatggggg	agagggcaca	gggcccagaca	aacctgtgag	actttggctc	catctctgca	3420
aaagggcgtc	ctgtgagtca	gcctgctccc	ctccaggctt	gtcctctccc	caccagctc	3480
tcgtttccaa	tgcacgtaca	gcccgtacac	accgtgtgct	gggacacccc	ac	3532

<210> 91

<211> 204

<212> DNA

<213> HUMAN

Sekvensliste

<400> 91
 cctgcccctc actccacccc catcctagct ttggtatggg ggagagggca cagggccaga 60
 caaacctgtg agactttggc tccatctctg caaaagggcg ctctgtgagt cagcctgctc 120
 ccctccaggc ttgctcctcc cccacccagc tctcgtttcc aatgcacgta cagcccgtac 180
 acaccgtgtg ctgggacacc ccac 204

<210> 92
 <211> 132
 <212> DNA
 <213> HUMAN

<400> 92
 ggatcctgtt gactcgtgac cttaccccca accctgtgct ctctgaaaca tgagctgtgt 60
 ccactcaggg ttaaattggat taagggcggg gcaagatgtg ctttgttaaa cagatgcttg 120
 aaggcagcat gc 132

<210> 93
 <211> 275
 <212> DNA
 <213> HUMAN

<400> 93
 gcatagtgcc aggtgggtgcc ttgggttcca agctagtcca tggccccgat aaccttctgc 60
 ctgtgcacac acctgcccct cactccaccc ccatactagc tttggtatgg gggagagggc 120
 acagggccag acaaacctgt gagactttgg ctccatctct gcaaaagggc gctctgtgag 180
 tcagcctgct cccctccagg cttgctcctc cccacccag ctctcgtttc caatgcacgt 240
 acagcccgtg cacaccgtgt gctgggacac cccac 275

<210> 94
 <211> 89
 <212> DNA
 <213> HUMAN

<400> 94
 ctgctcccct ccaggcttgc tcttcccca cccagctctc gtttccaatg cacgtacagc 60
 ccgtacacac cgtgtgctgg gacaccca 89

<210> 95
 <211> 61
 <212> DNA
 <213> HUMAN

<400> 95
 caccagctc tcgtttccaa tgcacgtaca gcccgtagc accgtgtgct gggacacccc 60
 a 61

<210> 96
 <211> 116
 <212> DNA
 <213> HUMAN

<400> 96
 acctgcccct cactccaccc ccatactagc tttggtatgg gggagagggc acagggccag 60
 acaaacctgt gagactttgg ctccatctct gcaaaagggc gctctgtgag tcagcc 116

<210> 97
 <211> 36
 <212> PRT
 <213> HUMAN

<400> 97

Sekvensliste

Gly Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp
 1 5 10 15
 Pro Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu
 20 25 30
 Glu Asp Leu Pro
 35

<210> 98
 <211> 6
 <212> PRT
 <213> HUMAN

<400> 98
 Gly Glu Glu Asp Leu Pro
 1 5

<210> 99
 <211> 4
 <212> PRT
 <213> HUMAN

<400> 99
 Glu Glu Asp Leu
 1

<210> 100
 <211> 5
 <212> PRT
 <213> HUMAN

<400> 100
 Glu Glu Asp Leu Pro
 1 5

<210> 101
 <211> 6
 <212> PRT
 <213> HUMAN

<400> 101
 Glu Asp Leu Pro Ser Glu
 1 5

<210> 102
 <211> 7
 <212> PRT
 <213> HUMAN

<400> 102
 Glu Glu Asp Leu Pro Ser Glu
 1 5

<210> 103
 <211> 6
 <212> PRT
 <213> HUMAN

<400> 103
 Asp Leu Pro Gly Glu Glu
 1 5

Sekvensliste

<210> 104
 <211> 22
 <212> PRT
 <213> HUMAN

<400> 104
 Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp Leu Pro
 1 5 10 15
 Ser Glu Glu Asp Ser Pro
 20

<210> 105
 <211> 25
 <212> PRT
 <213> HUMAN

<400> 105
 Gly Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp
 1 5 10 15
 Pro Pro Gly Glu Glu Asp Leu Pro Gly
 20 25

<210> 106
 <211> 24
 <212> PRT
 <213> HUMAN

<400> 106
 Glu Asp Pro Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro
 1 5 10 15
 Gly Glu Glu Asp Leu Pro Glu Val
 20

<210> 107
 <211> 7
 <212> PRT
 <213> HUMAN

<400> 107
 Gly Glu Thr Arg Ala Pro Leu
 1 5

<210> 108
 <211> 7
 <212> PRT
 <213> HUMAN

<400> 108
 Gly Glu Thr Arg Glu Pro Leu
 1 5

<210> 109
 <211> 7
 <212> PRT
 <213> HUMAN

<400> 109
 Gly Gln Thr Arg Ser Pro Leu

Sekvensliste

1

5

<210> 110
<211> 1247
<212> DNA
<213> HUMAN

<220>
<221> misc_feature
<222> (1)..(1247)

```
<400> 110
tatgctactt tttgcagtc tttcattaca tttttctctc ttcatttgaa gagcatgtta 60
tatcttttag cttcacttgg cttaaaaagg tctctcatta gcctaacaca gtgtcattgt 120
tggtagcact tggatcataa gtggaaaaac agtcaagaaa ttgcacagta atacttgttt 180
gtaagaggga tgattcaggt gaatctgaca ctaagaaaact cccctacctg aggtctgaga 240
ttcctctgac attgctgtat ataggccttt cctttgacag cctgtgactg cggactatct 300
ttcctaagca agatatgcta aagttttgtg agcctttttc cagagagagg tctcatatct 360
gcatcaagtg agaacaataa atgtctgcat gtttccatat ttcaggaatg tttgcttggt 420
ttttatgctt ttatatagac agggaaaact gttcctcagt gacccaaaag aggtgggaat 480
tggtattgga tatcatcatt ggcccacgct ttctgacctt ggaaacaatt aaggggtcat 540
aatctcaatt ctgtcagaat tggtagaaga aatagctgct atgtttcttg acattccact 600
tggtaggaaa taagaatgtg aaactcttca gttgggtgtg gtccctngtt tttttgcaat 660
ttccttctta ctgtgttaaa aaaaagtatg atcttgctct gagagggtgag gcattcttaa 720
tcgatctctt taaagatcaa taatataatc ctttcaagga ttatgtcttt attataataa 780
agataatttg tctttaacag aatcaataat ataatccctt aaaggattat atctttgctg 840
ggcgtagtgg ctcacacctg taatcccagc actttgggtg gccagggtgg aaggatcaaa 900
tttgccctact tctatattat cttctaaagc agaattcatc tctcttccct caatatgatg 960
atattgacag gggtttgccct cactcactag attgtgagct cctgctcagg gcaggtagcg 1020
ttttttgttt ttgtttttgt ttttcttttt tgagacaggg tcttgctctg tcaaccaggg 1080
cagagtgcaa tggtagagtc tcagctcact gcagcctcaa ccgcctcggc tcaaaccatc 1140
atccccattc agcctcctga gtagctggga ctacaggcac atgccattac acctggctaa 1200
tttttttgta tttctagtag agacagggtt tggccatggt gcccggg 1247
```

<210> 111
<211> 17
<212> DNA
<213> HUMAN

```
<400> 111
ctctgtgagt cagcctg 17
```

<210> 112
<211> 23
<212> DNA
<213> HUMAN

```
<400> 112
aggcttgctc ctccccacc cag 23
```

<210> 113
<211> 18
<212> DNA
<213> HUMAN

```
<400> 113
agactttggc tccatctc 18
```

<210> 114
<211> 20
<212> DNA
<213> HUMAN

Sekvensliste

<400> 114
cactccaccc ccatacctagc 20

<210> 115
<211> 26
<212> DNA
<213> HUMAN

<400> 115
gggagagggc acagggccag acaaac 26

<210> 116
<211> 15
<212> PRT
<213> HUMAN

<400> 116
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10 15

<210> 117
<211> 34
<212> DNA
<213> HUMAN

<400> 117
cgtctagaag gaattcagct agactggctc agca 34

<210> 118
<211> 15
<212> PRT
<213> HUMAN

<400> 118
Glu Val Lys Pro Lys Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu
1 5 10 15

<210> 119
<211> 12
<212> PRT
<213> HUMAN

<400> 119
Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro
1 5 10

<210> 120
<211> 12
<212> PRT
<213> HUMAN

<400> 120
Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly
1 5 10

<210> 121
<211> 10
<212> PRT
<213> HUMAN

Sekvensliste

<400> 121

Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro
1 5 10

<210> 122

<211> 12

<212> PRT

<213> HUMAN

<400> 122

Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu
1 5 10

<210> 123

<211> 12

<212> PRT

<213> HUMAN

<400> 123

Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp
1 5 10

<210> 124

<211> 12

<212> PRT

<213> HUMAN

<400> 124

Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu
1 5 10

<210> 125

<211> 9

<212> PRT

<213> HUMAN

<400> 125

Ala Pro Gly Glu Glu Asp Leu Pro Ala
1 5

<210> 126

<211> 9

<212> PRT

<213> HUMAN

<400> 126

Ala Gly Glu Glu Asp Leu Pro Gly Ala
1 5

<210> 127

<211> 9

<212> PRT

<213> HUMAN

<400> 127

Ala Glu Glu Asp Leu Pro Gly Glu Ala
1 5

<210> 128

<211> 9

<212> PRT

Sekvensliste

<213> HUMAN

<400> 128

Ala Glu Asp Leu Pro Gly Glu Glu Ala
1 5

<210> 129

<211> 9

<212> PRT

<213> HUMAN

<400> 129

Ala Asp Leu Pro Gly Glu Glu Asp Ala
1 5

<210> 130

<211> 9

<212> PRT

<213> HUMAN

<400> 130

Ala Leu Pro Gly Glu Glu Asp Leu Ala
1 5

<210> 131

<211> 8

<212> PRT

<213> HUMAN

<400> 131

Ala Gly Glu Glu Asp Leu Pro Ala
1 5

<210> 132

<211> 8

<212> PRT

<213> HUMAN

<400> 132

Ala Glu Glu Asp Leu Pro Gly Ala
1 5

<210> 133

<211> 8

<212> PRT

<213> HUMAN

<400> 133

Ala Glu Asp Leu Pro Gly Glu Ala
1 5

<210> 134

<211> 8

<212> PRT

<213> HUMAN

<400> 134

Ala Asp Leu Pro Gly Glu Glu Ala
1 5

<210> 135

Sekvensliste

<211> 8
 <212> PRT
 <213> HUMAN

<400> 135
 Ala Leu Pro Gly Glu Glu Asp Ala
 1 5

<210> 136
 <211> 8
 <212> PRT
 <213> HUMAN

<400> 136
 Ala Pro Gly Glu Glu Asp Leu Ala
 1 5

<210> 137
 <211> 9
 <212> PRT
 <213> HUMAN

<400> 137
 Ala Lys Lys Met Lys Arg Arg Lys Ala
 1 5

<210> 138
 <211> 9
 <212> PRT
 <213> HUMAN

<400> 138
 Ala Ile Thr Phe Asn Ala Gln Tyr Ala
 1 5

<210> 139
 <211> 9
 <212> PRT
 <213> HUMAN

<400> 139
 Ala Ser Ala Ser Ala Pro Val Ser Ala
 1 5

<210> 140
 <211> 9
 <212> PRT
 <213> HUMAN

<400> 140
 Ala Gly Gln Thr Arg Ser Pro Leu Ala
 1 5

<210> 141
 <211> 6
 <212> PRT
 <213> HUMAN

<400> 141
 Ser Glu Glu Asp Ser Pro
 1 5

Sekvensliste

<210> 142
<211> 6
<212> PRT
<213> HUMAN

<400> 142
Arg Glu Glu Asp Pro Pro
1 5

<210> 143
<211> 12
<212> DNA
<213> HUMAN

<400> 143
agggcacagg gc

12

PATENTKRAV

1. Polypeptid, k a r a k t e r i s e r t v e d a t det er valgt fra SEKV ID NR: 137
5 og 138.
2. Polypeptid ifølge krav 1, k a r a k t e r i s e r t v e d a t det er for bruk innen medisin.
- 10 3. Anvendelse av polypeptid ifølge krav 1, for fremstilling av et medikament for behandling av preneoplastisk og/eller neoplastisk sykdom som er kjennetegnet ved unormal ekspresjon av MN protein, ved inhibering av veksten av vertebrate preneoplastiske og/eller neoplastiske celler som uttrykker MN protein unormalt.
- 15 4. Peptidkompleks, k a r a k t e r i s e r t v e d a t det omfatter et polypeptid ifølge krav 1, kovalentlig bundet til polylysin hvortil det er bundet en nukleinsyre som koder for et cytotoxisk protein eller cytotoxisk polypeptid operativt koblet til MN-genpromoter, hvor nevnte peptidkompleks når administrert inn i en vertebrat preneoplastisk eller neoplastisk celle som unormalt uttrykker MN-protein, inhiberer
20 veksten av nevnte celle.
5. Peptidkompleks ifølge krav 4, k a r a k t e r i s e r t v e d a t nevnte cytotoxiske protein er HSV tymidinkinaser.
- 25 6. Peptidkompleks ifølge krav 5, k a r a k t e r i s e r t v e d a t nevnte nukleinsyre ytterligere omfatter en nukleinsyre som koder for et cytokin operativt koblet til nevnte MN-genpromoter.
- 30 7. Fremgangsmåte for identifisering av et organisk eller et uorganisk molekyl som spesifikt bindes til et sete på et MN protein, hvortil vertebrat celler adhererer i en celleadhesjonsanalyse, idet nevnte sete består av en aminosyresekvens valgt fra SEKV ID NR: 10 og 97-106, k a r a k t e r i s e r t v e d a t den omfatter testing av organiske og uorganiske molekyler i nevnte celleadhesjonsanalyse og identifisering av molekyler som inhiberer adhesjonen av vertebratcellene til nevnte MN protein

som spesifikt bindende til nevnte sete, og nevnte molekyl er nyttig for behandling av preneoplastisk og/eller neoplastisk sykdom, som er kjennetegnet ved unormal ekspresjon av MN protein, idet nevnte celleadhesjonsanalyse omfatter:

5 (a) muliggjøre at MN proteinet bindes til et substrat hvortil nevnte vertebratceller ikke binder;

 (b) skylle ubundet MN protein fra nevnte substrat;

10 (c) inkubering av nevnte bundete MN protein med nevnte molekyl og med nevnte vertebratceller;

 (d) skyling av ubundede vertebratceller fra nevnte MN protein; og

15 (e) bestemme om nevnte molekyl inhiberer adhesjonen av nevnte vertebratceller til nevnte MN protein.

8. Fremgangsmåte ifølge krav 7, k a r a k t e r i s e r t v e d a t nevnte molekyl er uorganisk.

20 9. Fremgangsmåte ifølge krav 7, k a r a k t e r i s e r t v e d a t nevnte molekyl er organisk.

10. Fremgangsmåte ifølge krav 9, k a r a k t e r i s e r t v e d a t nevnte molekyl er et protein eller et polypeptid.

25 11. Fremgangsmåte ifølge krav 10, k a r a k t e r i s e r t v e d a t nevnte protein eller polypeptid omfatter en aminosyresekvens valgt fra SEKV ID NR: 137 og 138.

30 12. Fremgangsmåte ifølge krav 10, k a r a k t e r i s e r t v e d a t nevnte polypeptid er valgt fra SEKV ID NR: 137 og 138.

13. Fremgangsmåte ifølge krav 7, k a r a k t e r i s e r t v e d a t nevnte organiske eller uorganiske molekyl, når i kontakt med en vertebrat preneoplastisk og/eller neoplastisk celle som uttrykker MN protein unormalt, hemmer veksten av nevnte celle.

5

14. Fremgangsmåte ifølge krav 7, k a r a k t e r i s e r t v e d a t setet på MN proteinet hvortil nevnte vertebratceller adhererer i nevnte celleadhesjonsanalyse er innenfor den proteoglykan-lignende domenen til MN proteinet.

10

15. Fremgangsmåte ifølge krav 7, k a r a k t e r i s e r t v e d a t vertebratcellene er pattedyrceller.

16. Fremgangsmåte ifølge krav 7, k a r a k t e r i s e r t v e d a t nevnte vertebratceller er humane celler.

15

1	ACA	GTC	AGC	CGC	ATG	GCT	CCC	CTG	TGC	CCC	P	S	P	W	L	P	L	12
1																		48
13	L	I	P	A	P	A	P	G	L	T	V	Q	L	L	L	L	S	28
49	TTG	ATC	CCG	GCC	CCT	GCT	CCA	GGC	CTC	ACT	GTG	CAA	CTG	CTG	CTG	CTG	TCA	96
29	L	L	L	L	M	P	P	H	P	Q	R	L	P	P	R	M	Q	44
97	CTG	CTG	CTT	CTG	ATG	CCT	GTC	CAT	CCC	CAG	AGG	TTG	CCC	CGG	ATG	CAG		144
45	E	D	S	P	L	G	G	G	S	S	G	E	D	D	P	L		60
145	GAG	GAT	TCC	CCC	TTG	GGA	GGA	GGC	TCT	TCT	GGG	GAA	GAT	GAC	CCA	CTG		192
61	G	E	E	D	L	P	S	E	E	D	S	P	R	E	E	D		76
193	GGC	GAG	GAG	GAT	CTG	CCC	AGT	GAA	GAG	GAT	TCA	CCC	AGA	GAG	GAG	GAT		240
77	P	P	G	E	E	D	L	P	G	E	E	D	L	P	G	E		92
241	CCA	CCC	GGA	GAG	GAG	GAT	CTA	CCT	GGA	GAG	GAG	GAT	CTA	CCT	GGA	GAG		288
93	E	D	L	P	E	V	K	P	K	S	E	E	E	G	S	L		108
289	GAG	GAT	CTA	CCT	GAA	GTT	AAG	CCT	AAA	TCA	GAA	GAA	GAG	GGC	TCC	CTG		336
109	K	L	E	D	L	P	T	V	E	A	P	G	D	P	Q	E		124
337	AAG	TTA	GAG	GAT	CTA	CCT	ACT	GTT	GAG	GCT	CCT	GGA	GAT	CCT	CAA	GAA		384
125	P	Q	N	A	H	R	D	K	E	G	D	D	Q	Q	S	H		140
385	CCC	CAG	AAT	AAT	GCC	CAC	AGG	GAC	AAA	GAA	GGG	GAT	GAC	CAG	AGT	CAT		432
141	W	R	Y	G	G	D	P	P	W	P	R	V	S	P	A	C		156
433	TGG	CGC	TAT	GGA	GGC	GAC	CCG	CCC	TGG	CCC	CGG	GTG	TCC	CCA	GCC	TGC		480
157	A	G	R	F	Q	S	P	V	D	I	R	P	Q	L	A	A		172
481	GCG	GGC	CGC	TTC	CAG	TCC	CCG	GTG	GAT	ATC	CGC	CCC	CAG	CTC	GCC	GCC		528

FIG.-1A

173	F	C	P	A	L	R	P	L	E	L	L	G	F	Q	L	P	188
529	TTC	TGC	CCG	GCC	CTG	CGC	CCC	CTG	GAA	CTC	CTG	GGC	TTC	CAG	CTC	CCG	576
189	P	L	P	E	L	R	L	R	N	N	G	H	S	V	Q	L	204
577	CCG	CTC	CCA	GAA	CTG	CGC	CTG	CGC	AAC	AAT	GGC	CAC	AGT	GTG	CAA	CTG	624
205	T	L	P	P	G	L	E	M	A	L	G	P	G	R	E	Y	220
625	ACC	CTG	CCT	CCT	GGG	CTA	GAG	ATG	GCT	CTG	GGT	CCC	GGG	CGG	GAG	TAC	672
221	R	A	L	Q	L	H	L	H	W	G	A	A	G	R	P	G	236
673	CGG	GCT	CTG	CAG	CTG	CAT	CTG	CAC	TGG	GGG	GCT	GCA	GGT	CGT	CCG	GGC	720
237	S	E	H	T	V	E	G	H	R	F	P	A	E	I	H	V	252
721	TCG	GAG	CAC	ACT	GTG	GAA	GGC	CAC	CGT	TTC	CCT	GCC	GAG	ATC	CAC	GTG	768
253	V	H	L	S	T	A	F	A	R	V	D	E	A	L	G	R	268
769	GTT	CAC	CTC	AGC	ACC	GCC	TTT	GCC	AGA	GTT	GAC	GAG	GCC	TTG	GGG	CGC	816
269	P	G	G	L	A	V	L	A	A	F	L	E	E	G	P	E	284
817	CCG	GGA	GGC	CTG	GCC	GTG	TTG	GCC	GCC	TTT	CTG	GAG	GAG	GGC	CCG	GAA	864
285	E	N	S	A	Y	E	Q	L	L	S	R	L	E	E	I	A	300
865	GAA	AAC	AGT	GCC	TAT	GAG	CAG	TTG	CTG	TCT	CGC	TTG	GAA	GAA	ATC	GCT	912
301	E	E	G	S	E	T	Q	V	P	G	L	D	I	S	A	L	316
913	GAG	GAA	GGC	TCA	GAG	ACT	CAG	GTC	CCA	GGA	CTG	GAC	ATA	TCT	GCA	CTC	960
317	L	P	S	D	F	S	R	Y	F	Q	Y	E	G	S	L	T	332
961	CTG	CCC	TCT	GAC	TTC	AGC	CGC	TAC	TTC	CAA	TAT	GAG	GGG	TCT	CTG	ACT	1008
333	T	P	P	C	A	Q	G	V	I	W	T	V	F	N	Q	T	348
1009	ACA	CCG	CCC	TGT	GCC	CAG	GGT	GTC	ATC	TGG	ACT	GTG	TTT	AAC	CAG	ACA	1056

FIG. 1B

349	V	M	L	S	A	K	Q	L	H	T	L	S	D	T	L	W	364
1057	GTG	ATG	CTG	AGT	GCT	AAG	CAG	CTC	CAC	ACC	CTC	TCT	GAC	ACC	CTG	TGG	1104
365	G	P	G	D	S	R	L	Q	L	N	F	R	A	T	Q	P	380
1105	GGA	CCT	GGT	GAC	TCT	CGG	CTA	CAG	CTG	AAC	TTC	CGA	GCG	ACG	CAG	CCT	1152
381	L	N	G	R	V	I	E	A	S	F	P	A	G	V	D	S	396
1153	TTG	AAT	GGG	CGA	GTG	ATT	GAG	GCC	TCC	TTT	CCT	GCT	GGA	GTG	GAC	AGC	1200
397	S	P	R	A	A	E	P	V	Q	L	N	S	C	L	A	A	412
1201	AGT	CCT	CGG	GCT	GCT	GAG	CCA	GTC	CAG	CTG	AAT	TCC	TGC	CTG	GCT	GCT	1248
413	G	D	I	L	A	L	V	F	G	L	L	F	A	V	T	S	428
1249	GGT	GAC	ATC	CTA	GCC	CTG	GTT	TTT	GGC	CTC	CTT	TTT	GCT	GTC	ACC	AGC	1296
429	V	A	F	L	V	Q	M	R	R	Q	H	R	R	G	T	K	444
1297	GTC	GCG	TTC	CTT	GTG	CAG	ATG	AGA	AGG	CAG	CAC	AGA	AGG	GGA	ACC	AAA	1344
445	G	G	V	S	Y	R	P	A	E	V	A	E	T	G	A	*	460
1345	GGG	GGT	GTG	AGC	TAC	CGC	CCA	GCA	GAG	GTA	GCC	GAG	ACT	GGA	GCC	TAG	1392
1393	AGG	CTG	GAT	CTT	GGA	GAA	TGT	GAG	AAG	CCA	GCC	AGA	GGC	ATC	TGA	GGG	1440
1441	GGA	GCC	GGT	AAC	TGT	CCT	GTC	CTG	CTC	ATT	ATG	CCA	CTT	CCT	TTT	AAC	1488
1489	TGC	CAA	GAA	ATT	TTT	TAA	AAT	AAA	TAT	TTA	TAA	T					1522

FIG._1C

FIG._1A

FIG._1B

FIG._1C

FIG._1

```

1  ggatcctgtt gactcgtgac cttaccccca accctgtgct ctctgaaaca tgagctgtgt
61  ccactcaggg ttaaatggat taagggcggg gcaagatgtg cttgtttaa cagatgcttg
121  aaggcagcat gctcgttaag agtcatcacc aatcccta atcaagta at cagggacaca
181  aacactgagg aaggccgcag ggtcctctgc ctaggaaaac cagagacctt tgttcacttg
241  ttatctgac cttccctcca ctattgtcca tgaccctgcc aatccccc at cgtgagaaa
301  cacccaagaa ttatcaataa aaaaataaat ttaaaaaaaa aatacaaaa aaaaaaaa
361  aaaaaaaa gacttacgaa tagttattga taaatgaata gctattggt aagccaagta
421  aatgatcata ttcaaaaacca gacggccatc atcacagctc aagtcctacc gatcttgatct
481  ctttatcat gtcatctctt gatttcacta gattagtc atccctcaa atctctcccc
541  aagttctaat tacgttccaa acatttaggg gttacatgaa gcttgaacct actaccttct
601  ttgcttttga gccatgagtt gtaggaatga tgagtttaca cttacatgc tggggattaa
661  ttaaaactt accctaaagt cagttgggta gccttggct tatcttcta gctaatcttg
721  tagttaatgg atgcactgtg aatcttgcta tgatagtctt cctccacac ttgccactag
781  gggtaggtag gtactcagtt ttcagtaatt gcttacctaa gaccctaag cctatttctc
841  ttgtactggc ctttatctgt aatatgggca tatttaatac aatataattt ttggagtctt
901  ttgtgtgtt tgttgtgtt ttttttgag acggagtctt gcatctgtca tggccaggct
961  ggagtagcag tggtgccatc tcggctcact gcaagctcca cctcccgagt tcacgccatt
1021  ttcctgcctc agcctccga gtagctggga ctacaggcgc ccgccaccat gccgggctaa
1081  ttttttgtat ttttggtaga gacggggtt caccgtgta gccagaatgg tctcgatctc
1141  ctgacttcgt gatccacccg cctcggcctc ccaaagtctt gggattacag gtgtgagcca
1201  ccgcacctgg ccaattttt gagtctttta agtaaaaa atgtcttcta agctggtaac
1261  tatggtagat ttccctttat taatgtgtgt ctgacgggtca tataggttct tttgagtctg
1321  gcatgcata gctactttt gcagtccttt cattacattt ttctctcttc atttgaagag
1381  catgttatat cttttagctt cacttggctt aaaagggtct ctcattagcc taacacagtg
1441  tcattgttgg taccacttgg atcataagt atcataagtg gaaaaacagt caagaaatg cacagtaata
1501  cttgttttga agagggatga ttcagggtgaa tctgacacta agaaactccc ctacctgagg
1561  tctgagattc ccttgacatt gctgtatata ggcttttctt ttgacagcct tgactgagg
1621  actatttttc ttaagcaaga tatgtctaaag ttttgtgagc ctttttccag agagaggtct
1681  catatctgca tcaagtga gaacataaatg tctgcatgtt tccatatctc aggaatgttt
1741  gcttgtgttt tatgtttta tatagacagg gaaactgtt gaaactgtt cctcagtgac ccaaaagagg
1801  tgggaattgt tattggatat catcatcttc ccacgcttc tgaccttgg aacaattaag
1861  ggttcataat ctcaattctg tcagaattgg tacaagaa at agctgctatg tttcttgaca
1921  tccacttgg taggaaataa gaatgtgaaa ctcttcagtt ggtgtgtgtc cct?gttttt

```

FIG._2A

```

1981 ttgcaatttc cttcttactg tgtaaaaaa aagtatgac ttgctctgag agtgaggca
2041 ttcttaataca tgatctttaa agatcaataa tataatcctt tcaaggatta tgtctttatt
2101 ataataaaga taatttgtct ttaacagaat caataatata atcccttaaa ggattatatc
2161 ttctgctgggc gcagtggctc acacctgtaa tcccagcact ttgggtggcc aaggtggaag
2221 gatcaaatct gcctactctc atatatctt cttaagcaga attcatctct ctccctcaa
2281 tatgatgata ttgacaggtt ttgctctcac tccactagatt gtgagctcct gctcagggca
2341 ggtagcgttt ttgtttttg tttttgttt tactttttga gacagggtct tgcctgttca
2401 ccaggccag agtgcaatgg tacagtctca gctcactgca gcctcaaccg cctcggctca
2461 aaccatcatc ccatttcagg ctcttgagta gctgggacta caggcacatg ccattacacc
2521 tggctaattt ttttgtattt ctagttaga cagggtttgg ccattgttgc cgggctggtc
2581 tcgaactcct ggactcaagc aatccacca cctcagcctc ccaaatgag ggaccgtgtc
2641 ttattcatct ccatgtccct agtccatagc ccagtgtgg acctatggtg gtactaaata
2701 aatatctgtt gaatgcaata gtaatatagca tttcaggag caagaactag attaacaaag
2761 gtggtaaaag gtttggagaa aaaaataata gtttaatttg gctagagtat gagggagagt
2821 agtaggagac aagatggaaa ggtctcttgg gcaagggtttt gaaggaaagt ggaagtcaga
2881 agtacacaat gtgcatactg tggcaggcag tggggagcca atgaaggctt ttgagcagga
2941 gagtaatgtg ttgaaaaata aatataggtt aaacctatca gagccctct gacacataca
3001 cttgcttttc attcaagctc aagtttgtct ccacataacc cattaactaa ctccccctg
3061 ggctccccta gcagcctgcc ctacctctt acctgcttcc tggaggagtc agggatgtat
3121 acatgagctg ctttccctct cagccagagg acatgggggg cccagctcc cctgctttc
3181 cccttctgtg cctggagctg ggaagcaggc cagggttagc tgaggctggc tggcaagcag
3241 ctgggtggtg ccagggagag cctgcatagt gccaggtgtt gccttgggtt ccaagctagt
3301 ccatggcccc gataaccttc tgcctgtgca cacacctgcc cctcactcca ccccatcct
3361 agcttttggt agggggagag ggcacagggc cagacaaacc tgtgagactt tggctccatc
3421 tctgcaaaag ggcgctctgt gagtcaacct gctccccctc aggttggctc ctccccacc
3481 cagctctcgt ttccaatgca cgtacagccc gtacacaccc tgtgctggga caccacACAG
3541 TCAGCCGCAT GGCTCCCCTG TGCCCCAGCC CCTGGCTCCC TCTGTTGATC CCGGCCCTG
3601 CTCCAGGCCT CACTGTGCAA CTGCTGCTGT CACTGCTGCT TCTGGTGCCT GTCCATCCCC
3661 AGAGGTTGCC CCGGATGCAG GAGGATTCCC CCTTGGGAGG AGGCTCTTCT GGGGAAGATG
3721 ACCCACTGGG CGAGGAGGAT CTGCCAGTG AGAGGATTC ACCCAGAGAG GAGGATCCAC
3781 CCGGAGAGGA GGATCTACCT GGAGAGGAGG ATCTACCTGG AGAGGAGGAT CTACCTGAAG
3841 TTAAGCCCTAA ATCAGAAGAA GAGGGCTCCC TGAAGTTAGA GGATCTACCT ACTGTTGAGG
3901 CTCTGGGAGA TCCCTCAAGAA CCCCAGAATA ATGCCACACG GGACAAAGAA Ggtaagtgg

```

FIG..2B

3961 catcaatctc caaatccagg ttccaggagg ttcatgactc cctcccata cccagccta
 4021 ggctctgttc actcaggga ggaggggaga ctgtactccc cacagaagcc ctccagagg
 4081 tcccatacca atatcccat cccactctc ggaggtagaa agggacagat gtggagagaa
 4141 aataaaaagg gtgcaaaagg agagaggtga gctggatgag atggagaga agggggaggc
 4201 tggagaagag aaagggatga gaactgcaga tgagagaaaa atgtgcaga cagaggaaaa
 4261 aaataggtgg agaaggagag tcagagagtt tgaggggaag agaaaaggaa agcttgggag
 4321 gtgaagtggg taccagagag aagcaagaag agctggtaga agtcatctca tcttaggcta
 4381 caatgaggaa ttgagacctg ggaagaaggg acacagcagg tagagaaacg tggcttcttg
 4441 actcccaagc caggaatttg gggaaggagg ttggagacca tacaaggcag agggatgagt
 4501 ggggagaaga aagaagggag aaaggaaaaga tggtgtactc actcatttgg gactcaggac
 4561 tgaagtggcc actcactttt tttttttttt tttttgagac aaactttcac tttgtttgcc
 4621 caggctggag tgcaatggcg cgatctcggc tcaactgcaac ctccacctcc cgggttcaag
 4681 tgattctcct gcctcagcct ctagccaagt agctgcgatt acaggcatgc gccaccacgc
 4741 ccggctaatt tttgtatttt tagtagagac ggggtttcgc catgttggtc aggttggtct
 4801 cgaactcctg atctcaggty atccaaccac cctggcctcc caaagtgtg ggattatagg
 4861 cgtgagccac agcgctggc ctgaagcagc cactcacttt tacagacctt aagacaaatga
 4921 ttgcaagctg gtaggattgc tgtttggccc acccagctgc ggtgtttagt ttgggtgcgg
 4981 tctcctgtgc ttgacacctg gcccgcttaa ggcatttgtt acccgtaatg ctccgtgaag
 5041 gcatctgcgt ttgtgacatc gtttttggtcg ccaggaaggg attggggctc taagcttgag
 5101 cggttcatcc ttttcattta tacaggggat gaccagagtc attggcgcta TGGAGgtgag
 5161 acacccaccc gctgcacaga cccaatctgg gaacccagct ctgtggatct cccctacagc
 5221 cgtccctgaa cactgggtccc gggcgtccca ccgcccgcct accgtccac cccctcacct
 5281 ttctaccgg ggttccctaa gttcctgacc taggcgtcag acttccctcac tatactctcc
 5341 caccccaggg GACCCGCCCT GGCCCCGGGT GTCCCCAGCC TGCGCGGGCC GCTTCCAGTC
 5401 CCCGGTGGAT ATCCGCCCCC AGCTCGCCCG CTCTGCCCG GCCCTGCGCC CCTTGGAACT
 5461 CCTGGGCTTC CAGCTCCCGC CGCTCCCGA ACTGCGCCTG CGCAACAATG GCCACAGTGG
 5521 tgaggggggt tcccgcgca gacttgggga tggggcgggg cgaggggaag ggaaccgtcg
 5581 cgcagtgcct gcccggggt tgggctggcc ctaccgggcy gggccggctc acttgccctc
 5641 ccctacgcag TGCAACTGAC CTGCTCCTT GGGCTAGAGA GGGCTCGTCC CAGGTCGGAG
 5701 GAGTACCGGG CTCTGCAGCT GCATCTGCAC TGGGGGGCTG GAGGTgagcg gagaaagggc
 5761 CACACTGTGG AAGGCCACCG TTTCCCTGCC GAGgtgagcg cggactggcc gagaagggc
 5821 aaaggagcgg ggcggacggg gtggccctct cctaccctcg tgtccttttc
 5881 agATCCACGT GGTTCACCTC AGCACCGCCT TTGCCAGAGT TGACGAGGCC TTGGGGCGCC

FIG.-2C

5941 CGGGAGGCCT GGGCGTGTG GCCGCTTTC TGGAGgtacc agatcctgga cacccttac
 6001 tccccgctt cccatcccat gctcctccc gactctatcg tggagccaga gaccctatcc
 6061 cagcaagctc actcaggccc ctggctgaca aactcattca cgcactgttt gttcatttaa
 6121 caccactgt gaaccaggca ccagccccc ccagtaaaag gaaggatgg tgagatgcct
 6181 tctaaggagc ccacagccag tgggggaggg tgacatgaca gacacatagg aaggacatag
 6241 taaagatggt ggtcacagag gaggtgacac ttaaagcctt cactggtaga aaagaaaagg
 6301 aggtgttcat tgcagaggaa acagaatgtg caaagactca gaatatggcc tatttaggga
 6361 atggctacat acaccatgat tagaggaggg ccagtaaaag gaaggatgg tgagatgcct
 6421 gctaggttca ctcaactcact tttatttatt ttttattttt ttgacagtc tctctgtcgc
 6481 ccaggctgga gtgcagtgggt gtgatcttgg gtcaactgcaa ctccgcctc ccgggttcaa
 6541 gggattctcc tgcctcagct tcctgagtag ctgggggttac aggtgtgtgc caccatgccc
 6601 agctaatttt ttttgtatt tttagtagac aggttttcac catgttggtc aggtgtgtct
 6661 caaactcctg gcctcaagtg atccgcctga ctacagcctac caaagtgtg attacaagtg
 6721 tgagccaccg tggccagcca cactcactga ttctttaatg ccagccacac agcacaaaagt
 6781 tcagagaaat gcctccatca tagcatgtca atatgttcat actcttaggt tcatgatgtt
 6841 cttaacatta ggttcataag caaaataaga aaaaagaata ataaataaaa gaagtggcat
 6901 gtcaggacct cacctgaaa gccaacaca gaatcatgaa ggtgaatgca gaggtgacac
 6961 caacacaaag gtgtatatat ggtttcctgt ggggagtatg tacggaggca gcagtgagtg
 7021 agactgcaaa cgtcagaagg gcacgggtca ctgagagcct agtatcctag taaagtgggc
 7081 tctctccctc tctctcagc ttgtcattga aaaccagtcc accaagcttg ttggttcgca
 7141 cagcaagagt acatagagtt tgaataata cataggattt taagaggagg acactgtctc
 7201 taacaaaaaa aacaacagca acaacaaaa actatgggaa ccccttcat gttccggcct
 7261 agcattctca gagctgagga atgggagagg actatgggaa ccccttcat gttccggcct
 7321 tcagccatgg ccctggatc acctggaatc atgtcttaca atgtcattcc ccagGAGGG
 7381 CCGGAAGAA AACAGTGCTT ATGAGCAGTT GCTGTCTCGC TTGGAAGAAA TCGCTGAGGA
 7441 AGgtcagttt gttggtctgg ccaactaatc ctgtggccta gttcataaag aatcacccct
 7501 tggagcttca ggtctgaggc tggagatggg ctccctccag tgcaggaggg attgaagcat
 7561 gagccagcgc tcatcttgat aataaccatg aagctgacag acacagttac ccgcaaacgg
 7621 ctgcctacag attgaaaaac aagcaaaaac cgcggggcac ggtggctcac gcctgtaatc
 7681 ccagcacttt gggaggccaa ggcagggtgga tcacgaggtc aagagatcaa gaccatcctg
 7741 gccaacatgg tgaaacccca tctctactaa aaatagccag aaatagccag gcgtggtggc
 7801 ggtgtgcctgt aatccagct actcgggagg ctgaggcagg agaattggcat gaaccggga
 7861 ggcagaagtt gcagtgagcc gagatcgtgc cactgcactc cagcctgggc aacagagcga

FIG._2D

7921 gactcttgtc tcaaaaaaa ggtgtttgga gaaaccaag caaaaccaa aatgagacaa
 7981 aaaaaaag accaaaaaat ggtgtttgga aatgtcaag gtcaagtctg gagagctaaa
 8041 ctttttctga gaactgttta tctttaataa gcatcaata ttttaacttt gtaaatactt
 8101 ttgttggaaa tcgttctctt cttagtcact cttagttaga gacagggttt caccatattg
 8161 ctgacctttt taggtttctg ctgacctagg ttacaagt tttagtaga gacagggttt caccatattg
 8221 gttttgtata gttatcaata ttcatattta ttacaagt tttagtaga gacagggttt caccatattg
 8281 tctttttttt tttttttttt ttttttttat ctgacctgt gatccaccag cctcgccctc ccaaagtgtc
 8341 gccagggtgc tctcaaaact ttttcttttt aatttgctct gggcttaaac ttgtggccca gcactttatg
 8401 gggattcat ttttcttttt aatttgctct gggcttaaac ttgtggccca gcactttatg
 8461 atggtacaca ggttaagag tgtagactca gacgggtctt ctctcttctt tctcttctt
 8521 cctcccttcc ctccacctt ccttctctc caaagccctg tacttttttt tgagttaacg tcttatggga
 8581 caggcctctt cttagtgaag aagtggctc aagtggtctc agagttgagt taccttggct tctgggaggt
 8641 agggcctgca cctatatacc tgaagcttta aggggtgca atgtagatga gaccccaaca
 8701 gaaactgtat ccctataccc tcacaggctc agagactcag gtcccaggac tggacatatc TGCACTCCTG
 8761 tagatcctct TCAGCCGCTA CTTCCAATAT GAGGGGTCTC TGACTACACC GCCCTGTGCC
 8821 CCTCTGACT TCTGGACTGT GTTTAACCA GAGGGGTCTC TGACTACACC GCCCTGTGCC
 8881 CAGGGTGTCA TCTGGACTGT GTTTAACCA GAGGGGTCTC TGACTACACC GCCCTGTGCC
 8941 ctgggggtgtg tgtggacaca gtgggtgctg gggaaagagg atgtaagatg agatgagaaa
 9001 caggagaaga aagaaatcaa ggctgggctc tgtggcttac gcctataatc ccaccacgtt
 9061 gggaggctga ggtggagaa tggtttgagc ccaggagttc aagacaaggc ggggcaacat
 9121 agtgtgacct catctctacc tactcaagga ggctgagggt ggaagatcgc ttgattccag
 9181 gtatgcggcc tagtcccagc ctatgatccc accactgcct accatcttta ggatacatat
 9241 gaggttgaga ctgcagtgag ctatgatccc accactgcct accatcttta ggatacatat
 9301 attatttat aaaagaaatc aagaggctgg atggggaata caggagctgg aggggtggagc
 9361 cctgagggtg tggttgtgag ctggcctggg acccttgttt cctgtcatgc catgaaccca
 9421 ccacactgt ccactgacct ccctagCTCC ACACCTCTC TGACACCTCTG TGGGACCTG
 9481 GTGACTCTCG GCTACAGCTG AACTTCCGAG CGACGCAGCC TTTGAATGGG CGAGTGATTG
 9541 AGGCCTCCTT CCTGCTGGA GTGGACAGCA GTCTCGGGC TGCTGAGCCA Ggtacagctt
 9601 tgtctgggtt ccccccagcc agtagtccct tctctccca tgtgtgtgct agtgtctgtc
 9661 attggtggtc acagcccgc tctcacatct cctttttctc tccagTCCAG CTGAATTCTT
 9721 GCCTGGCTGC TGgtgagttt gcccctctc ttggtcctga tggcaggaga ctctcagca
 9781 ccattcagcc ccagggtgc tcaggaccgc ctctgctccc tctcttctt tgcagaacag
 9841 accccaacc caatataga gaggcagatc atgggtggga tccccatt gtccccagag

9901 gctaattgat tagaatgaag cttgagaaaat ctccagcat ccctctcgca aaagaatccc
 9961 cccccctttt tttaagata gggctcact ctgtttgccc caggctggg tggtgtggca
 10021 cgatcatagc tcaactgagc ctcgaaactcc taggctcagg actgtgcctg gcccacaaag gcccttttac
 10081 ctcaagcac tgggactgta ggcactgagcc tcttaccctt tctcgtgtat ccaccctcat
 10141 ttggctttta ggaagcaaaa acggtgctta cactatgggg ctccttgaga actcggggca
 10201 cccttggctg gcctcttctg gagactgagg gaggaactct gcagacccct ctcccttccc
 10261 ggggtggtgg agtgcaactga ggcaggtgtt TGACATCCTA GCCCTGGTTT TTGGCCTCCT
 10321 aaagcagccc tctctgctct ccatcgcagg GCAGATGAGA AGGCAGCAC Ggtattacac
 10381 TTTTGCTGTC ACCAGCGTCG CGTTCCTTGT acccttgtgg agtcacttca tgcaaaagcg
 10441 tgaccctttc ttcaggcaca agcttccccc ctgattagcc ttccctgttg tgtacacaca
 10501 atgcaaatga gctgctcctg ggccagtttt GCCCAGCAGA GGTAGCCGAG ACTGGAGCCT
 10561 gAAGGGGAAC CAAAGGGGGT GTGAGCTACC CAGCCAGAGG CATCTGAGGG GGAGCCGGTA
 10621 AGAGGCTGGA TCTTGGAGAA TGTGAGAAGC CTTTAACTG CCAAGAAATT TTTTAAATA
 10681 ACTGTCCTGT CCTGCTCATT ATGCCACTTC ctttgttccc caaatcagaa ggaggatttt
 10741 AATATTATA ATaaaatatg tgttagtcac ttagtggttaa tgcatttatt ctattacagt
 10801 gaatttccta ttactgttat tagcaccaat gttgctcc
 10861 tcggcctcct tccacacatc actccaatgt gttgctcc

FIG._2F

FIG._2A

FIG._2B

FIG._2C

FIG._2D

FIG._2E

FIG._2F

FIG._2

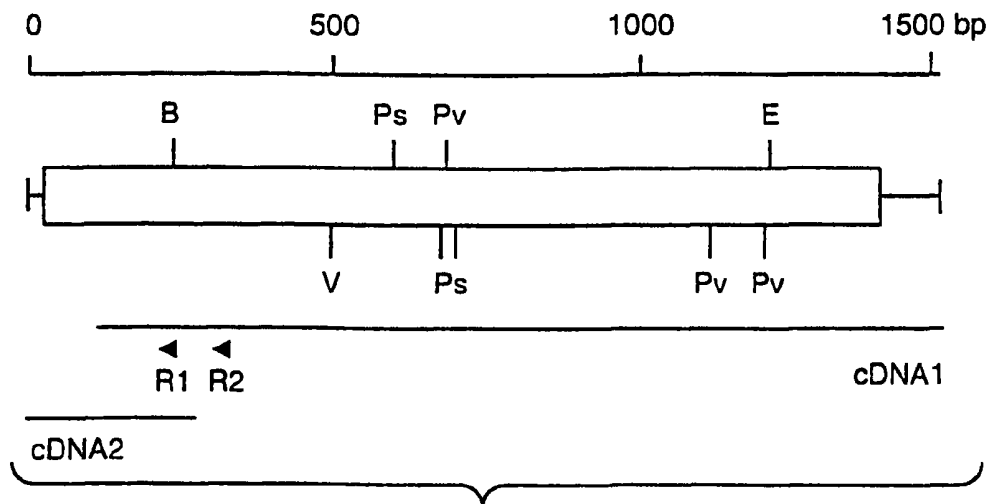


FIG._3

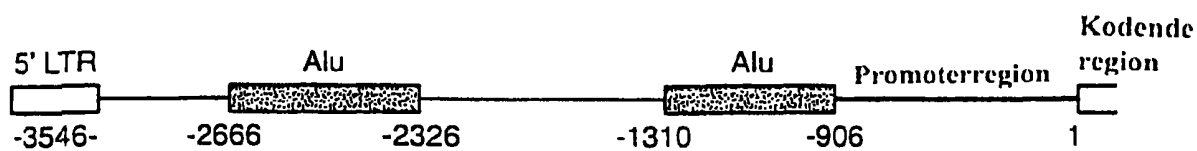


FIG._4

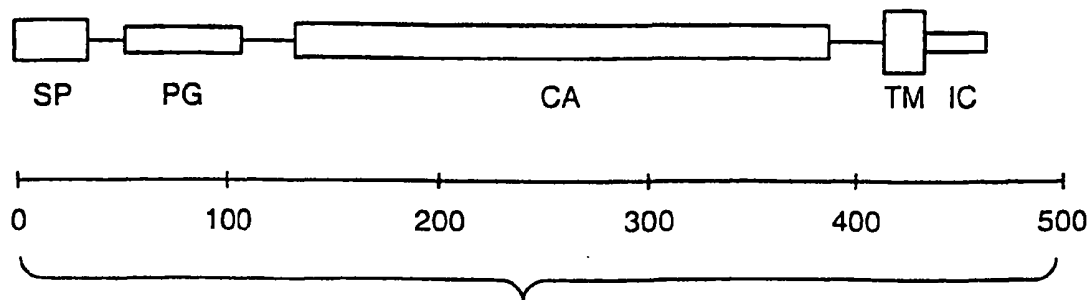


FIG._8

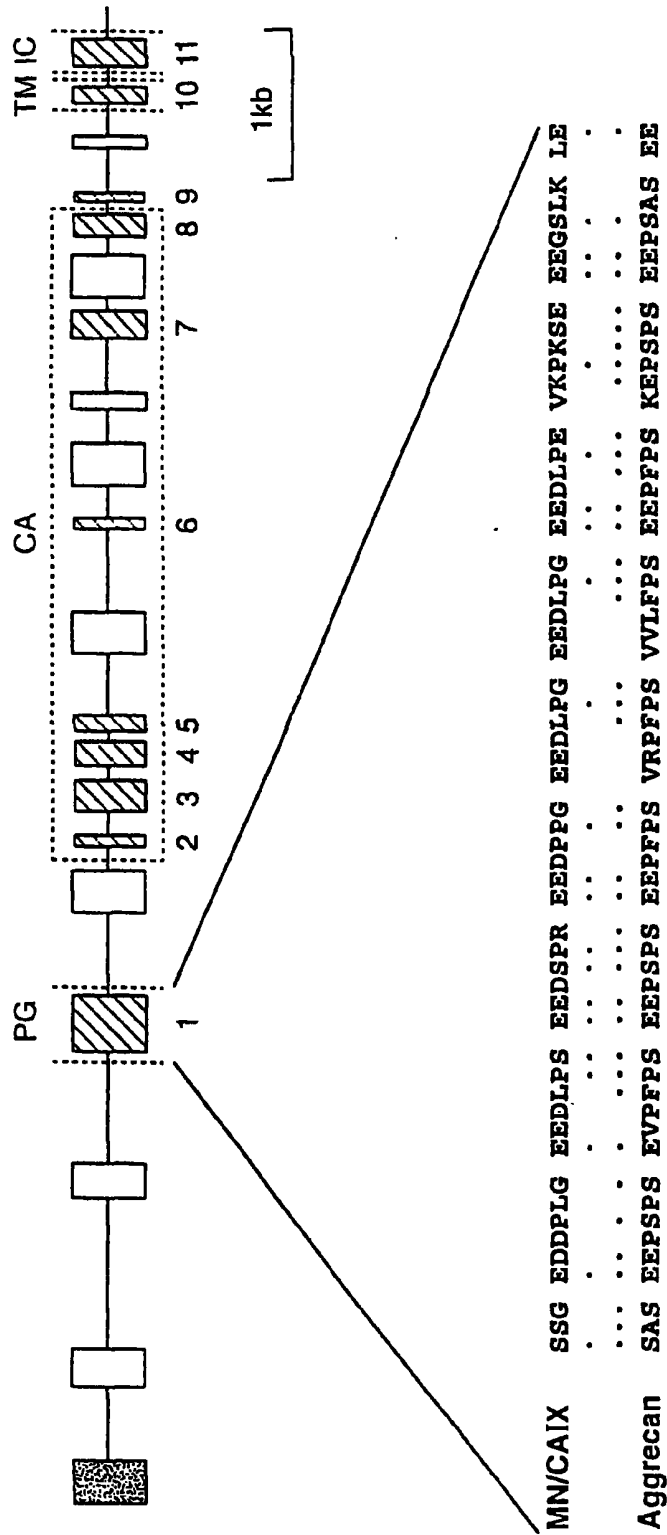


FIG._5

```

-506 CTTGCTTTTC ATTCAAGCTC AAGTTTGCTT CCCACATACC CATTACTTAA CTCACCCCTCG

-446 GGCTCCCCCTA GCAGCCTGCC CTACCTCTTT ACCTGCTTCC TGGTGGAGTC AGGATGTAT
      AP2

-386 ACATGAGCTG CTTTCCCTCT CAGCCAGAGG ACATGGGGGG CCCCAGCTCC CCTGCCTTTC

-326 CCCTTCTGTG CCTGGAGCTG GGAAGCAGGC CAGGGTTAGC TGAGGCTGGC TGGCAAGCAG

-266 CTGGGTGGTG CCAGGGAGAG CCTGCATAGT GCCAGGTGGT GCCTTGGGTT CCAAGCTAGT
      VII p53

-206 CCATGSCCCC GATAACCTTC TGCCTGTGCA CACACCTGCC CCTCACTCCA CCCCCATCCT
      VI Inr V

-146 AGCTTTGGTA TGGGGGAGAG GGCACAGGGC CAGACAAACC TGTGAGACTT TGGCTCCATC
      IV AP1 III Inr

-86 TCTGCAAAAG GCGGCTCTGT GAGTCAGCCT GCTCCCCCTCC AGGCTTGCTC CTCCTCCACCC
      II AP1 p53 I AP2 ***

-26 CAGCTCTCGT TTCCAATGCA CGTACAGCCC GTACACACCG TGTGCTGGGA CACCCACAG
      ...

```

FIG..6

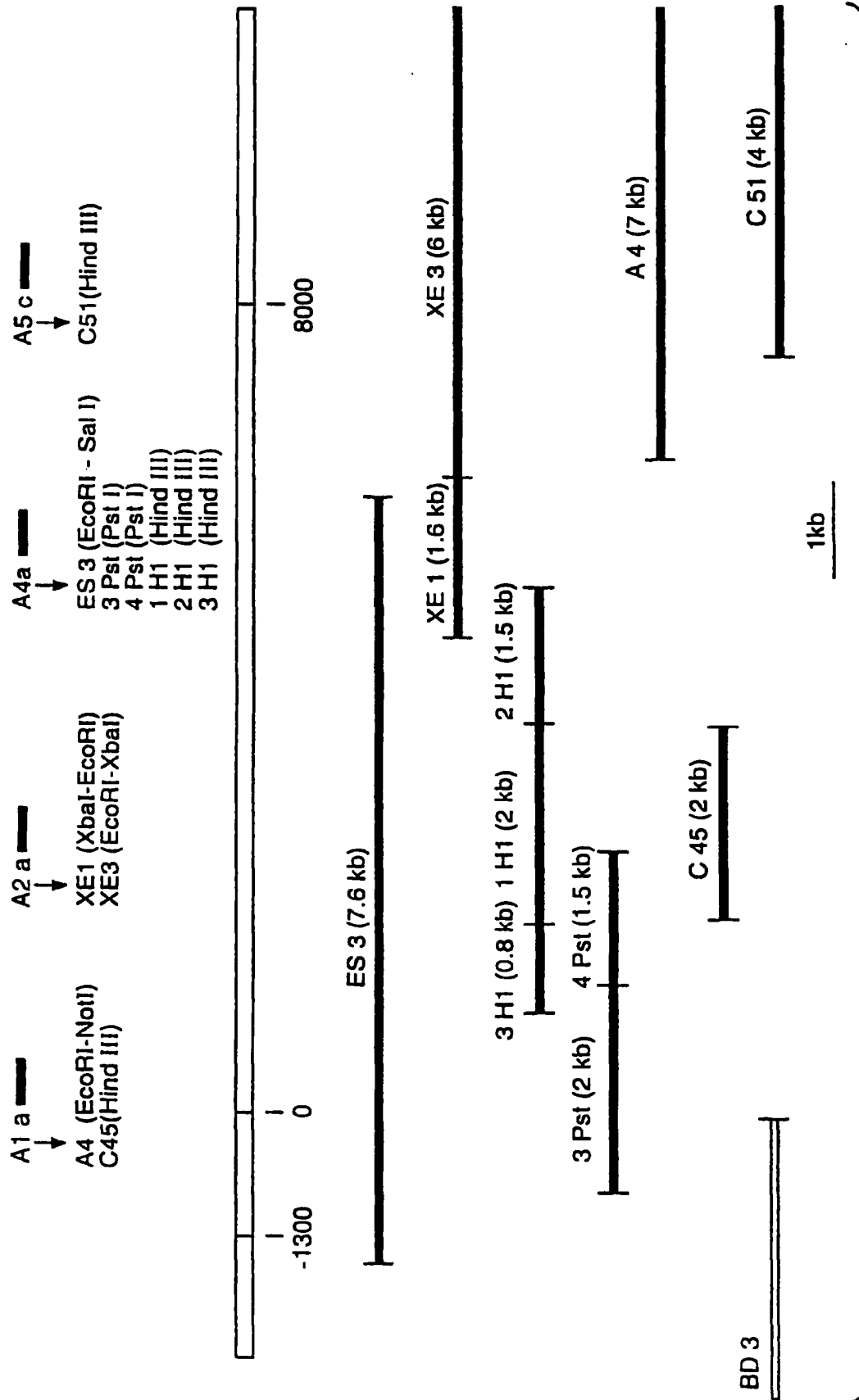


FIG. 7