



NORGE

(12) **PATENT**

(19) NO

(11) **317801**

(13) **B1**

(51) Int Cl⁷

C 07 K 14/51, 19/00 , C 12 N 15/12

Patentstyret

(21)	Søknadsnr	19962304	(86)	Innt.inng.dag og søknadsnr	1994.12.06 PCT/US94/14030
(22)	Inng.dag	1996.06.04	(85)	Videreføringsdag	1996.06.04
(24)	Løpedag	1994.12.06	(30)	Prioritet	1993.12.07, US, 164103 1994.03.25, US, 217780 1994.11.02, US, 333576
(41)	Alm.tilgj	1996.06.04			
(45)	Meddelt:	2004.12.13			
(71)	Søker	Genetics Institute Inc , 87 Cambridge Park Drive, Cambridge, MA 02140, US President and Fellows of Harvard College , Suite 727, Holyoke Center, 1350 Massachusetts Avenuc, MA02138 CAMBRIDGE, US			
(72)	Oppfinner	John M Wozney, 59 Old Bolton Road, Hudson, MA 01749, US Anthony J Celeste, 86 Packard Street, Hudson, MA 01749, US Neil M Wolfman, 5 Phillips Lane, MA02030 DOVER, US Vicki A. Rosen, , Brookline, MA, US Gerald H. Thomsen, , Port Jefferson, NY, US Douglas A. Melton, , Lexington, MA, US			
(74)	Fullmektig	Zacco Norway AS , Postboks 765 Sentrum, 0106 Oslo, NO			
(54)	Benevnelse	DNA-molekyl, kimaert DNA-molekyl eller vektor som koder for BMP12 eller BMP12-relaterte proteiner, og vertceller, fremgangsmåter, proteiner, sammensetninger og anvendelse derav.			
(56)	Anførte publikasjoner	WO 9316099 WO 9426893			
(57)	Sammendrag				

Benmorfogeniske proteiner BMP- 12 og BMP- 13 er blitt klonet. Sammensetningene av disse proteinene med tendon/ligament-lignende vevsinduserende aktivitet er beskrevet. Sammensetningene er nyttige for behandling av tendinitis og tendon eller ligament defekter og ved beslektet vevsreparering.

Foreliggende oppfinnelse vedrører en ny familie rensede proteiner, og sammensetninger som inneholder slike proteiner. Sammensetningene er nyttige for induksjon av tendon/-ligamentlignende vevsdannelse, leging av sår og ligament og annen vevsreparering. Disse proteinene kan også bli anvendt i sammensetninger for å øke aktiviteten til benmorfogeniske proteiner.

Søken for å oppnå molekylet eller molekylene som er ansvarlig for dannelse av ben, brusk, sene og andre vev tilstede i ben og andre vevsekstrakter har ført til oppdagelsen av et nytt sett molekyler betegnet benmorfogeniske proteiner (BMP). Strukturer av flere proteiner, betegnet BMP-1 til og med BMP-11 er tidligere blitt bestemt. De unike induserende aktivitetene til disse proteinene, sammen med deres tilstedeværelse i ben, tyder på at de er viktige regulatorer av benrepareringsprosesser, og kan være involvert i den normale opprettholdelsen av benvev. Det er nødvendig å identifisere ytterligere proteiner som spiller en rolle for dannelsen av andre vitale vev. Foreliggende oppfinnelse vedrører identifikasjon av en familie proteiner med sene/ligamentlignende vevsinduserende aktivitet, og som er nyttige i sammensetninger for induksjon av sene/ligamentlignende vevsdannelse og reparering.

I en utførelsesform omfatter foreliggende oppfinnelse DNA-molekyler kodende for et sene/ligamentlignende induserende protein som oppfinnerne har betegnet VL-1. Dette nye proteinet er nå betegnet BMP-12. Foreliggende oppfinnelse omfatter også DNA-molekyler kodende for BMP-12 relaterte proteiner.

BMP-12 relaterte proteiner er en undergruppe av BMP/TGF- β /Vg-1 familie av proteiner, inkludert BMP-12 og VL-1, som er definert som sene/ligamentlignende vevsinduserende proteiner som blir kodet av DNA-sekvenser som er klonet og identifisert, for eksempel ved anvendelse av PCR, ved anvendelse av BMP-12 spesifikke primere, så som primerene nr. 6 og nr. 7 beskrevet nedenfor, med reduserte stringensbetingelser. Det er foretrukket at DNA-sekvensene kodende for BMP-12 relaterte proteiner er minst omtrent 80% homologe på aminosyrenivå fra aminosyrer med aminosyrene nr. 3 til nr. 103 ifølge SEQ ID NO: 1.

DNA-molekylene har fortrinnsvis en DNA-sekvens kodende for BMP-12 proteinet, sekvensen som er gitt i SEQ ID NO: 1, eller et BMP-12 relatert protein som beskrevet heri. Både BMP-12 proteinet og BMP-12 relaterte proteiner er karakterisert ved deres evne til å indusere dannelsen av sene/ligamentlignende vev i analysen beskrevet i eksemplene.

DNA-molekylene ifølge oppfinnelsen omfatter fortrinnsvis en DNA-sekvens, som beskrevet i SEQ ID NO: 1; fortrinnsvis nukleotidene nr.496 til nr. 882, nr.571 til nr.882 eller nr.577 til nr.882 av SEQ ID NO:1; eller DNA-sekvenser som hybridiserer til ovennevnte under stringente hybridiseringsbetingelser og som koder for et protein som tilveiebringer evnen til å danne sene/ligamentlignende vev. DNA-molekylene ifølge oppfinnelsen kan også omfatte en DNA-sekvens som beskrevet i SEQ ID NO: 25; fortrinnsvis nukleotidene nr. 604 eller nr.658 til nr.964 til SEQ ID NO:25.

DNA-molekylene ifølge oppfinnelsen omfatter også DNA-molekyler omfattende en DNA-sekvens kodende for et BMP-12 relatert protein med aminosyresekvensen vist i SEQ ID NO: 2 eller SEQ ID NO:26, samt naturlig forekommende alleliske sekvenser og ekvivalente degenerative kodonsekvenser ifølge SEQ ID NO: 2 eller SEQ ID NO: 26. DNA-sekvensen ifølge foreliggende oppfinnelse koder fortrinnsvis for aminosyrene nr. 25 til nr. 104, nr 1 til nr 104 eller nr 3 til nr 103 til SEQ ID NO: 2; eller aminosyrene nr. 1 til nr 120 eller nr 19 til nr 120 til SEQ ID NO: 26. DNA-sekvensen kan omfatte, i en 5' til 3'-retning, nukleotider kodende for et propeptid, og nukleotider kodende for aminosyrene nr. 25 til nr. 104, nr 1 til nr 104 eller nr 3 til nr 103 til SEQ ID NO: 2; eller aminosyrene nr 1 til nr 120 eller nr 19 til nr 120 til SEQ ID NO: 26. Egenskapene som er nyttige i ovennevnte utførelsesform blir fortrinnsvis valgt fra gruppen bestående av nativt BMP-12 propeptid og et proteinpropeptid fra et annet medlem av TGF-B superfamilien eller BMP familien. Oppfinnelsen omfatter videre DNA-sekvenser som hybridiserer til ovennevnte DNA-sekvenser under stringente hybridiseringsbetingelser og koder for et BMP-12 relatert protein som utviser evne til å indusere dannelse av sene/ligamentlignende vev.

I en annen utførelsesform omfatter foreliggende oppfinnelse vertsceller og vektorer som omfatter et DNA-molekyl kodende for BMP-12 protein, eller et BMP-12 relatert protein. Vertscellene og vektorene kan videre omfatte den kodende sekvensen i operativ assosiasjon med en ekspresjonskontrollsekvens derav.

I en annen utførelsesform omfatter foreliggende oppfinnelse en fremgangsmåte for produksjon av et rensset BMP-12 relatert protein, i det metoden omfatter trinnene av (a) dyrke en vertscelle transformert med ovennevnte DNA-molekyl eller vektor omfattende en nukleotidsekvens kodende for et BMP-12 relatert protein; og (b) isolering og rensing av nevnte BMP-12 relaterte protein fra dyrkningsmediet. I en foretrukket utførelsesform omfatter fremgangsmåten

(a) dyrking av en celle transformert med et DNA-molekyl som omfatter nukleotidsekvensen fra nukleotid nr 496, nr 571 eller nr 577 til nr 879 eller nr 882 som vist i SEQ ID NO: 1; eller nukleotidsekvensen fra nr 604 eller nr 658 til nr 963 til SEQ ID NO: 25; og

(b) isolering og rensing fra nevnte kulturmedium av et protein som omfatter aminosyresekvensen fra aminosyre nr -25, nr 1 eller nr 3 til aminosyre nr 103 eller nr 104 som vist i SEQ ID NO: 2; eller fra aminosyre nr 1 eller nr 19 til aminosyre nr 120 som vist i SEQ ID NO: 26. Foreliggende oppfinnelse omfatter også et rensset protein produsert ifølge ovennevnte fremgangsmåter.

Foreliggende oppfinnelse omfatter videre rensset BMP-12 relatert protein kjennetegnet ved evnen til å indusere dannelsen av sene/ligamentlignende vev. BMP-12 relaterte polypeptider omfatter fortrinnsvis en aminosyresekvens som vist i SEQ ID NO: 2. Polypeptidet omfatter fortrinnsvis aminosyrene nr -25, nr 1 eller nr 3 til nr 103 eller 104 som angitt i SEQ ID NO: 2; eller aminosyrene nr 1 eller nr 19 til nr 120 som angitt i SEQ ID NO: 26. I en foretrukket utførelsesform kan det rensede polypeptidet være i form av en dimer som omfatter to subenheter som hver har aminosyresekvensen ifølge SEQ ID NO: 2.

I en annen utførelsesform omfatter foreliggende oppfinnelse sammensetninger omfattende en effektiv mengde av ovennevnte BMP-12 relaterte proteiner. I sammensetningene kan proteinet blandes sammen med en farmasøytisk akseptabel bærer.

Oppfinnelsen kan anvendes ved fremgangsmåter for sene/ligamentlignende vevsleging og vevsreparering, for behandling av tendinitt, eller andre sene- eller ligamentdefekter, og for indusering av sene/ligamentlignende vevsdannelse hos en pasient som trenger dette, omfattende administrering til nevnte pasient en effektiv mengde av ovennevnte sammensetning.

Andre utførelsesformer innbefatter kimære DNA-molekyler som omfatter en DNA-sekvens kodende for et propeptid fra et medlem av TGF- β superfamilien av proteiner koblet i riktig leseramme til en DNA-sekvens kodende for et BMP-12 relatert polypeptid. Et egnet propeptid er propeptidet fra BMP-2. Oppfinnelsen omfatter også heterodimeriske proteinmolekyler som omfatter en monomer som har aminosyresekvensen vist i SEQ ID NO: 2, og en monomer med aminosyresekvensen til et annet protein fra TGF- β superfamilien.

Foreliggende oppfinnelse kan også anvendes ved fremgangsmåter for å indusere sene/-ligamentlignende vevsdannelse i en pasient som trenger dette omfattende administrering til nevnte pasient en effektiv mengde av en sammensetning som omfatter et protein som utviser evne til å indusere dannelse av sene/ligamentlignende vev, i det nevnte protein har en aminosyresekvens vist i SEQ ID NO: 2 eller SEQ ID NO: 4 eller SEQ ID NO: 26. Aminosyresekvensene er fortrinnsvis en av følgende: (a) aminosyre nr -25, nr 1 eller nr 3 til nr 103 eller nr 104 til SEQ ID NO: 2; (b) aminosyre nr 1 eller nr 19 til nr 119 eller nr 120 til SEQ ID NO: 4; (c) aminosyre nr 1 eller nr 19 til nr 119 eller nr 120 til SEQ ID NO: 26; (d) mutanter og/eller varianter av (a), (b) eller (c) som utviser evne til å danne sene og/eller ligament. I andre utførelsesformer ifølge ovennevnte fremgangsmåte blir proteinet kodet av en DNA-sekvens ifølge SEQ ID NO: 1, SEQ ID NO: 3 eller SEQ ID NO: 25, mer foretrukket en av følgende: (a) nukleotidene nr 496, nr 571 eller nr 577 til nr 879 eller nr 882 av SEQ ID NO: 1; (b) nukleotidene nr 845 eller nr 899 til nr 1201 eller nr 1204 av SEQ ID NO: 3; (c) nukleotider nr 605 eller nr 659 til nr 961 eller nr 964 av SEQ ID NO: 25; og (d) sekvenser som hybridiserer til (a) eller (b) under stringente hybridiseringsbetingelser og som koder for et protein som utviser evne til å danne sene/ligamentlignende vev.

Beskrivelse av sekvensene

SEQ ID NO: 1 er nukleotidsekvensen som koder for humant BMP-12.

SEQ ID NO: 2 er aminosyresekvensen som omfatter modent humant BMP-12 polypeptid.

SEQ ID NO: 3 er nukleotidsekvensen som koder for proteinet MP52.

SEQ ID NO: 4 er aminosyresekvensen som omfatter modent MP52 polypeptid.

SEQ ID NO: 5 er nukleotidsekvensen til en spesifikk amplifisert del av human BMP-12 kodende sekvens.

SEQ ID NO: 6 er aminosyresekvensen som blir kodet av nukleotidsekvensen til SEQ ID NO: 5.

SEQ ID NO: 7 er nukleotidsekvensen til en spesifikk amplifisert del av human VL-1 kodende sekvens.

SEQ ID NO: 8 er aminosyresekvensen som blir kodet av nukleotidsekvensen til SEQ ID NO: 7.

SEQ ID NO: 9 er nukleotidsekvensen til plasmid pALV1-781, anvendt for ekspresjon av BMP-12 i E.coli.

SEQ ID NO: 10 er nukleotidsekvensen av et fragment av den murine klonen mV1.

SEQ ID NO: 11 er aminosyresekvensen til et fragment av murint protein kodet av mV1.

SEQ ID NO: 12 er nukleotidsekvensen til et fragment av murin klon mV2.

SEQ ID NO: 13 er aminosyresekvensen til et fragment av murin protein kodet av mV2.

SEQ ID NO: 14 er nukleotidsekvensen til et fragment av murin klon mV9.

SEQ ID NO: 15 er aminosyresekvensen til et fragment av murint protein kodet av mV9.

SEQ ID NO: 16 er aminosyresekvensen til en BMP/TGF- β /Vg-1 protein konsensus sekvens. Første Xaa representerer enten Gln eller Asn; andre Xaa representerer enten Val eller Ile.

SEQ ID NO: 17 er nukleotidsekvensen til oligonukleotid nr. 1.

SEQ ID NO: 18 er aminosyresekvensen til en BMP/TGF- β /Vg-1protein konsensus sekvens. Xaa representerer enten Val eller Leu.

SEQ ID NO: 19 er nukleotidsekvensen til oligonukleotid nr 2.

SEQ ID NO: 20 er nukleotidsekvensen til oligonukleotid nr 3.

SEQ ID NO: 21 er nukleotidsekvensen til oligonukleotid nr 4.

SEQ ID NO: 22 er nukleotidsekvensen til oligonukleotid nr 5.

SEQ ID NO: 23 er nukleotidsekvensen til oligonukleotid nr 6.

SEQ ID NO: 24 er nukleotidsekvensen til oligonukleotid nr 7.

SEQ ID NO: 25 er nukleotidsekvensen til human VL-1 (BMP-13) kodende sekvens.

SEQ ID NO: 26 er aminosyresekvensen som blir kodet av nukleotidsekvensen til SEQ ID NO: 25.

SEQ ID NO: 27 er nukleotidsekvensen kodende for en fusjon av BMP-2 propeptid og moden kodende sekvens til BMP-12.

SEQ ID NO: 28 er aminosyresekvensen som blir kodet av nukleotidsekvensen til SEQ ID NO: 27.

SEQ ID NO: 29 er nukleotidsekvensen kodende for murint mV1 protein . X01 er Val, Ala, Glu eller Gly; X02 er Ser, Pro Thr eller Ala; X03 er Ser eller Arg; X04 er Leu, Pro, Gln eller Arg; X05 er Cys eller Trp; X06 er Val, Ala, Asp eller Gly; C07 er Val, Ala, Glu eller Gly; X08 er Gln, Lys eller Glu.

SEQ ID NO: 30 er aminosyresekvensen som blir kodet av nukleotidsekvensen til SEQ ID NO: 29. X01 til og med X08 er de samme som i SEQ ID NO: 29.

SEQ ID NO: 31 er nukleotidsekvensen som blir kodet av murin mV2 protein. X01 er Pro eller Thr; X02 er Val.

SEQ ID NO: 32 er aminosyresekvensen som blir kodet av nukleotidsekvensen til SEQ ID NO: 31. X01 og X02 er de samme som i SEQ ID NO: 31.

SEQ ID NO: 33 er nukleotidsekvensen som blir kodet av humant BMP-12 protein.

SEQ ID NO: 34 er aminosyresekvensen som blir kodet av nukleotidsekvensen til SEQ ID NO: 33.

SEQ ID NO: 35: er nukleotidsekvensen til oligonukleotid nr. 8.

Kort beskrivelse av figurene

Figur 1 er en sammenligning av humane BMP-12 og humane MP52 sekvenser.

DNA-sekvensene ifølge foreliggende oppfinnelse er nyttige for fremstilling av proteiner som induserer dannelse av sene/ligamentlignende vev, som ytterligere beskrevet nedenfor. DNA-sekvensene ifølge foreliggende oppfinnelse er videre nyttige for isolering og kloning av ytterligere DNA-sekvenser kodende for BMP-12 relaterte proteiner med lignende aktivitet. Disse BMP-12 relaterte proteinene kan være homologe fra andre arter, eller kan være relaterte proteiner innenfor samme arter.

Et ytterligere aspekt av foreliggende oppfinnelse er DNA-sekvenser kodende for ekspressjon av et sene/ligamentlignende vevsinduserende protein. Slike sekvenser innbefatter sekvensen til nukleotidene i en 5' til 3' retning illustrert i SEQ ID NO:1 eller SEQ ID NO: 25, DNA-sekvenser som unntatt degenerasjonen til den genetiske koden, er identiske med DNA-sekvensen SEQ ID NO: 1 eller 25, og som koder for proteiner ifølge SEQ ID NO: 2 eller 26. Videre innbefatter foreliggende oppfinnelse DNA-sekvenser som hybridiserer under stringente betingelser med DNA-sekvensen til SEQ ID NO: 1 eller 25 og som koder for et protein som har evne til å indusere dannelse av sene eller ligament. Foretrukne DNA-sekvenser innbefatter de som hybridiserer under stringente betingelser som beskrevet i Maniatis et al, Molecular Cloning (A Laboratory Manual). Cold Spring Harbor Laboratory (1982) sidene 387 til 389. Alleliske eller andre variasjoner av sekvensene til SEQ ID NO: 1 eller 25, enten slike nukleotidforandringer resulterer i forandringer i peptidsekvensen eller ikke, men hvor peptidsekvensen fortsatt har sene/ligamentlignende vevsinduserende aktivitet, er også innbefattet i foreliggende oppfinnelse.

BMP-12 DNA-sekvensen (SEQ ID NO: 1) og aminosyresekvensen (SEQ ID NO: 2) er angitt i sekvenslisten. Et annet protein som er nyttig for sammensetningene og fremgangsmåtene ifølge foreliggende oppfinnelse er VL-1. VL-1 er et BMP-12 relatert protein som ble klonet ved anvendelse av sekvensene fra BMP-12. Oppfinnerne har nå betegnet VL-1 som BMP-13. En del DNA-sekvens til VL-1 (SEQ ID NO: 7) og kodende aminosyresekvens (SEQ ID NO: 8); samt en DNA-sekvens som koder for moden VL-1 (SEQ ID NO:25) og kodet aminosyresekvens (SEQ ID NO: 26) er angitt i sekvenslisten. Til tross for at ytterligere beskrivelser er utført med hensyn til BMP-12 sekvensen til SEQ ID NO: 1 og 2, er det å bemerke at oppfinnelsen innbefatter lignende modifikasjoner og forbedringer som kan bli dannet på andre BMP-12 relaterte sekvenser, så som VL-1 sekvensen vist i SEQ ID NO: 25 og 26.

Sekvensen til BMP-12 vist i SEQ ID NO: 1 innbefatter hele den modne sekvensen og omtrent 190 aminosyrer av propeptidet. Den kodende sekvensen til modent humant BMP-12 protein ser ut til å begynne ved nukleotid nr 496 eller nr 571 og fortsette til og med nukleotid nr 882 til SEQ ID NO: 1. Det første cysteinet i syv cystein strukturen som er karakteristisk for TGF- β proteinene begynner ved nukleotid nr 577. Det siste cysteinet ender ved nr 879. Det er følgelig ventet at DNA-sekvenser kodende for aktive BMP-12 arter vil omfatte nukleotidene nr 577 til nr 879 til SEQ ID NO: 1.

Det er ventet at BMP-12, som uttrykt fra pattedyrceller så som CHO celler, eksisterer som en heterogen populasjon av aktive BMP-12 protein arter med varierende N-terminale ender. Det er ventet at alle aktive arter vil inneholde aminosyresekvensen begynnende med cysteinresidiet ved aminosyre nr 3 til SEQ ID NO: 2 og fortsette til og med minst cysteinresidie ved aminosyre 103 eller helt til stoppkodonet etter aminosyre 104. Andre aktive arter inneholder ytterligere aminosyresekvens i den N-terminale retningen. Som ytterligere beskrevet heri er det ventet at den N-terminale enden til aktive arter produsert fra pattedyrceller begynner etter forekomst av et konsensus spaltningssete, kodende for en peptidsekvens Arg-X-X-Arg. Det er følgelig ventet at DNA-sekvensene kodende for aktive BMP-12 proteiner vil ha en nukleotidsekvens som omfatter nukleotidsekvensen begynnende ved hvilke som helst av nukleotidene nr 196, 199, 208, 217, 361, 388, 493, 496 eller 571 til nukleotid nr. 879 eller 882 til SEQ ID NO:1.

Den N-terminale enden til en aktiv art av human BMP-12 er blitt eksperimentelt bestemt ved ekspresjon i *E. coli* å være som følger: [M]SRXSRKPLHVDF hvor X angir et aminosyreresidie uten klart signal, som er i samsvar med et cysteinresidie ved denne beliggenheten. Det ser følgelig ut som om den N-terminale enden til denne arten av BMP-12 er ved aminosyre nr. 1 til SEQ ID NO: 1, og en DNA-sekvens kodende for nevnte arter av BMP-12 vil begynne ved nukleotid nr 571 til SEQ ID NO: 1. Den tilsynelatende molekylvekten til denne arten av human BMP-12 dimer ble bestemt ved SDS-PAGE å være omtrent 20-22 kd på en Novex 16% tricin gel. Humant BMP-12 protein eksisterer som en klar, farveløs oppløsning i 0,1% trifluoreddiksyre.

Som tidligere beskrevet er BMP-12 relaterte proteiner en undergruppe av BMP/TGF- β /Vg-1 familien av proteiner, inkludert BMP-12 og VL-1, som kan bli definert som sene/ligamentlignende vevsinduserende proteiner kodet av DNA-sekvenser som kan bli klonet og identifisert, for eksempel ved anvendelse av PCR, ved anvendelse av BMP-12 spesifikke primere, som primerne nr 6 og nr 7 beskrevet nedenfor, med reduserte strin-

gensbetingelser. Det er foretrukket at DNA-sekvensene ifølge foreliggende oppfinnelse har minst omtrent 80% homologi på aminosyrenivå i forhold til aminosyrene med DNA kodende aminosyre nr 3 til nr 103 til SEQ ID NO: 1. I henhold til foreliggende oppfinnelse angir betegnelsen BMP-12 relaterte proteiner ikke humant MP52 protein. Ved anvendelse av sekvensinformasjonen til SEQ ID NO: 1 og SEQ ID NO:3, og sammenligningen angitt i figur 1, hører det inn under dette fagområdet og konstruere primere til BMP-12 sekvensen som vil muliggjøre kloning av gener kodende for BMP-12 relaterte proteiner.

Et eksempel på BMP-12-relaterte proteiner ifølge foreliggende oppfinnelse er VL-1, fortrinnsvis referert til som BMP-13. Sekvensen til hele modne BMP-13 sekvensen og idet minst en del av propeptidet til BMP-13 er gitt i SEQ ID NO: 25. Som BMP-12 er det ventet at BMP-13, som uttrykt av pattedyrceller så som CHO-celler, eksisterer som en heterogen populasjon av aktive arter av BMP-13 proteinet med varierende N-terminale ender. Det er ventet at alle aktive arter vil inneholde aminosyresekvensen begynnende med cysteinresidiet ved aminosyre nr. 19 til SEQ ID NO:26 og fortsette forbi det minste cysteinresidiet ved aminosyre 119 eller helt til stoppkodonet etter aminosyre 120. Andre aktive arter inneholder ytterligere aminosyresekvens i den N-terminale retningen. Som ytterligere beskrevet heri vil den N-terminale enden til de aktive artene produsert av pattedyrceller begynne etter forekomsten av et konsensusspaltningssete kodende for en peptidsekvens Arg-X-X-Arg. Det er følgelig ventet at DNA-sekvensene kodende for aktive BMP-13 proteiner vil ha en nukleotidsekvens som omfatter nukleotidsekvensen begynnende ved hvilke som helst av nukleotidene nr 410, 458, 602, 605 eller 659 tilnukleotid nr 961 eller 964 av SEQ ID NO: 25.

For å produsere rensede sene/ligamentlignende vevsinduserende proteiner nyttige i foreliggende oppfinnelse ble en fremgangsmåte anvendt som omfatter dyrking av en vertscelle transformert med en DNA-sekvens som omfatter en egnet kodende sekvens, spesielt DNA kodende sekvens fra nukleotid nr 496, nr 571 eller nr 577 til nr 879 eller nr 882 til SEQ ID NO: 1; og isolering og rensing fra kulturmediet av et protein som inneholder aminosyresekvensen eller en vesentlig homolog sekvens som representert ved aminosyrene nr -25, nr 1 eller nr 3 eller nr 104 til SEQ ID NO: 2. I en annen utførelsesform omfatter den anvendte fremgangsmåten dyrking av en vertscelle transformert med en DNA-sekvens som omfatter en egnet kodende sekvens, spesielt DNA-kodende sekvens fra nukleotid nr 605 eller nr 659 til nr 961 eller nr 964 til SEQ ID NO: 25; og isolering og rensing fra kulturmediet av et protein som inneholder aminosyresekvensen

eller en vesentlig homolog sekvens som representert ved aminosyrene nr 1 eller nr 19 til nr 119 eller nr 120 til SEQ ID NO: 26.

Humant MP52 DNA er beskrevet i WO93/16099 og beskrivelsen er inkorporert heri som referanse. Dette dokumentet beskriver ikke evnen som proteinet har til å danne sene/ligamentlignende vev, eller anvendelse derav i sammensetninger for induksjon av sene/ligamentlignende vev. Humant MP52 ble opprinnelig isolert ved anvendelse av RNA fra humant embryovev. Human MP52 nukleotidsekvens (SEQ ID NO:3) og kodende aminosyresekvenser (SEQ ID NO:4) er angitt i sekvenslisten heri. MP52 proteinet ser ut til å begynne ved nukleotid nr. 845 til SEQ ID NO: 3 og fortsetter til nukleotid nr. 1204 til SEQ ID NO: 3. Det første cysteinet av 7 cysteinstrukturen karakteristisk for TGF- β -proteinene begynner ved nukleotid nr. 899. Det siste cysteinet slutter ved nr. 1201. Andre aktive former av MP52 proteinet kan ha ytterligere nukleotider ved den N-terminale retningen fra nukleotid nr. 845 til SEQ ID NO:3.

Rensede humane MP52 proteiner ifølge foreliggende oppfinnelse kan bli produsert ved dyrking av en vertscelle transformert med en DNA-sekvens omfattende DNA-kodende sekvens til SEQ ID NO: 3 fra nukleotid nr 845 til nr 1204, og isolering og rensing fra kulturmediet et protein som inneholder aminosyresekvensen eller en vesentlig homolog sekvens som representert ved aminosyrene nr 1 til nr 120 til SEQ ID NO:4. Det er også ventet at aminosyresekvensen fra aminosyrene nr 17 eller nr 19 til nr 119 eller nr 120 til SEQ ID NO:4 vil beholde aktivitet. DNA-sekvensen fra nukleotidene nr 845, nr 893 eller nr 899 til nr 1201 eller nr 1204 er ventet å kode for aktive proteiner.

For ekspresjon av proteinet i pattedyrvertsceller blir vertscellen transformert med en kodende sekvens kodende for et propeptid egnet for sekresjon av proteinene fra vertscellen som er koblet i riktig leseramme til den kodende sekvensen for modent protein. Se for eksempel US-PS 5.168.050 og beskrivelsen er inkorporert heri som referanse, hvor et DNA kodende for en forløperdel av et pattedyrprotein forskjellig fra BMP-2 er koblet til DNA kodende for et modent BMP-2 protein. Foreliggende oppfinnelse innbefatter følgelig kimære DNA-molekyler som omfatter en DNA-sekvens kodende for et propeptid fra et medlem av TGF- β superfamilien til proteinene og som er koblet i riktig leseramme til en DNA-sekvens kodende for et sene/ligamentlignende vevsinduserende polypeptid. Betegnelsen "kimært" blir anvendt for å signifikere at propeptidet kommer fra et annet polypeptid enn det kodede modne polypeptidet. Vertscellen kan selvfølgelig bli transformert med en DNA-sekvens kodende for sekvensen som koder for det native propeptidet koblet i riktig leseramme til en kodende sekvens kodende for det modne

proteinet vist i SEQ ID NO: 2, SEQ ID NO:4 eller SEQ ID NO: 26. Hele sekvensen til det native propeptidet kan bli bestemt ved hjelp av fremgangsmåter kjent innenfor fagområdet ved anvendelse av sekvenser beskrevet i SEQ ID NO: 1, SEQ ID NO:3 eller SEQ ID NO:25 for å konstruere en egnet probe for identifisering og isolering av hele klonen.

Foreliggende oppfinnelse omfatter også nye DNA-sekvenser, fri for assosiasjon med DNA-sekvenser kodende for andre proteinholdige materialer, og kodende for ekspresjon av sene/ligamentlignende vevsinduserende proteiner. Disse DNA-sekvensene innbefatter de som er vist i SEQ ID NO:1 i en 5' til 3' retning og sekvenser som hybridiserer der-til under stringente hybridiseringsbetingelser (for eksempel 0,1X SSC, 0,1% SDS ved 65°C, se T. Maniatis et al, Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory (1982), sidene 387 til 389] og som koder for et protein med sene/-ligamentlignende vevsinduserende aktivitet.

DNA-sekvenser som koder for proteiner som blir kodet for av sekvensene ifølge SEQ ID NO: 1 eller SEQ ID NO:25, eller proteiner som omfatter aminosyresekvensen til SEQ ID NO: 2 eller SEQ ID NO: 26, men som har forskjellig kodonsekvens på grunn av degenerasjonen til den genetiske koden eller alleliske variasjoner (naturlig forekom-mende baseforandringer i arts populasjonen som kan eller som ikke behøver å resultere i en aminosyreforandring) også koder for sene/ligamentlignende vevsinduserende proteiner beskrevet heri. Variasjoner i DNA-sekvensene til SEQ ID NO: 1 eller SEQ ID NO:25 som blir forårsaket av punktmutasjoner eller ved induserte modifikasjoner (inkludert innskudd, delesjon og substitusjon) for å forsterke aktiviteten, halveringstiden eller produksjonen av polypeptidene som kodes for er også innbefattet i foreliggende oppfinnelse.

Et annet aspekt av foreliggende oppfinnelse tilveiebringer en ny fremgangsmåte for fremstilling av sene/ligamentlignende vevsinduserende proteiner. Fremgangsmåten ifølge foreliggende oppfinnelse innbefatter dyrking av en egnet cellelinje, som er blitt transformert med en DNA-sekvens kodende for et protein ifølge oppfinnelsen under kontroll av kjente regulatoriske sekvenser. Transformerte vertsceller blir dyrket og proteinene isolert og rensed fra kulturmediet. De rensede proteinene er vesentlig fri for andre proteiner som de blir produsert med samt andre kontaminanter.

Egnede celler eller cellelinjer kan være pattedyrceller, så som kinesisk hamstero-varie-celler (CHO). Som beskrevet ovenfor krever ekspresjon av protein i pattedyrcellene et

hensiktsmessig propeptid for å forsikre sekresjon av proteinet. Seleksjon av egnede pattedyrvertsceller og fremgangsmåte for transformasjon, dyrking, amplifikasjon, screening, produkt produksjon og rensing er kjent innenfor fagområdet. Se for eksempel Gething and Sambrook, *Nature*, 293:620-625 (1981), eller alternativt Kaufman et al, *Mol. Cell.Biol.*, 5(7):1750-1759 (1985) eller Howley et al. US-PS 4.419.446. En annen egnet pattedyrcellelinje, som er beskrevet i eksemplene, er ape COS-1 cellelinjen. Pattedyrcellen CV-1 kan også være egnet.

Bakterielle celler kan også være egnede verter. For eksempel er de forskjellige stammene av *E. coli* (for eksempel HB101, MC1061) velkjente som vertsceller innenfor området bioteknologi. Forskjellige stammer av *B.subtilis*, *Pseudomonas* eller andre baciller og lignende kan også bli anvendt i denne fremgangsmåten. For ekspresjon av proteinet i bakterielle celler er DNA kodende for et propeptid ikke nødvendig.

Bakteriell ekspresjon av pattedyrproteiner, inkludert medlemmer av TGF- β familien er kjent for å produsere proteinene i en ikke-glykosylert form, og i form av uoppløselige pelleter, kjent som inklusjonslegemer. Teknikker er blitt beskrevet innenfor fagområdet for oppløsning av disse inklusjonslegemene, denaturering av proteinet ved anvendelse av et chaotropisk middel og refolding av proteinet tilstrekkelig for å muliggjøre produksjon derav i en oppløselig form. Se for eksempel EP 0433225, og beskrivelsen er inkorporert heri som referanse.

Alternativt er fremgangsmåter blitt beskrevet som omgår inklusjonslegemedannelse, så som ekspresjon av genfusjonsproteiner, hvori det ønskede proteinet blir uttrykt som et fusjonsprotein med en fusjonspartner. Fusjonsproteinet blir senere utsatt for spaltning for å produsere det ønskede proteinet. Et eksempel på et slikt genfusjonsekspresjonssystem for *E. coli* er basert på anvendelse av *E. coli* tioredoksin genet som en fusjonspartner, LaVallie et al., *Bio/Technology*, 11:187-193 (1993), og beskrivelsen er herved inkorporert som referanse.

Mange stammer av gjærceller som er kjent for fagfolk innenfor dette området kan også bli anvendt som vertscelle for ekspresjon av polypeptidene ifølge foreliggende oppfinnelse. Alternativt kan insektceller anvendes som vertsceller i fremgangsmåten ifølge foreliggende oppfinnelse. Se for eksempel Miller et al, *Genetic Engineering*, 8:277-298 (Plenum Press 1986) og referansene sitert deri.

Et annet aspekt av foreliggende oppfinnelse tilveiebringer vektorer for anvendelse i fremgangsmåten for ekspresjon av disse sene/ligamentlignende vevsinduserende proteinene. Vektorene inneholder hele nye DNA-sekvenser som er beskrevet ovenfor og som koder for de nye faktorene ifølge oppfinnelsen. I tillegg inneholder vektorene hensiktsmessige ekspresjonskontrollsekvenser som muliggjør ekspresjon av proteinsekvensene. Alternativt er vektorer som inkorporerer modifiserte sekvenser som beskrevet ovenfor også utførelsesformer ifølge foreliggende oppfinnelse. I tillegg kan sekvensene ifølge SEQ ID NO: 1 eller SEQ ID NO: 3 eller SEQ ID NO: 25 manipuleres for å uttrykke et modent protein ved deletering av propeptidsekvensene og erstatning av disse med sekvenser som koder for fullstendig propeptid til BMP proteinene eller medlemmer av TGF- β superfamilien. Foreliggende oppfinnelse innbefatter følgelig kimære DNA-molekyler som koder for et propeptid fra et medlem av TGF- β superfamilien koblet i riktig leseramme til en DNA-sekvens kodende for et protein med aminosyresekvensen ifølge SEQ ID NO: 2 eller SEQ ID NO: 4 eller SEQ ID NO: 26. Vektorene kan bli anvendt i fremgangsmåten for transformering av celler og inneholder selekterte regulatoriske sekvenser i operativ assosiasjon med DNA kodende sekvenser ifølge oppfinnelsen som har evne til å lede replikasjon og ekspresjon derav i selekterte vertsceller. Regulatoriske sekvenser for slike vektorer er velkjente for fagfolk innenfor dette området og kan bli valgt avhengig av vertscellene. En slik seleksjon er rutinemessig og utgjør ikke en del av foreliggende oppfinnelse.

Et protein ifølge foreliggende oppfinnelse som induserer sene/ligamentlignende vev eller annen vevsdannelse i tilfeller hvor et slikt vev ikke normalt blir dannet kan anvendes for leging av sene eller ligament skader, deformiteter og andre sene eller ligament defekter i mennesker og andre dyr. En slik preparering som anvender et sene/ligamentlignende vevsinduserende protein kan ha profylaktisk anvendelse for å forhindre skade på sene eller ligamentvev, samt anvendelse ved forbedret fiksering av sene eller ligament til ben eller andre vev, og for reparering av defekter på sene eller ligament vev. De novo sene/ligamentlignende vevsdannelse indusert av en sammensetning ifølge foreliggende oppfinnelse bidrar til reparering av kongenital, traume indusert, eller andre sene eller ligament defekter med annen opprinnelse, og er også nyttig innen kosmetisk plastisk kirurgi for kobling eller reparering av sener eller ligamenter. Sammensetningene ifølge oppfinnelsen kan også være nyttige for behandling av tendinitt, carpal tunnelsyndrom og andre sene- eller ligamentdefekter. Sammensetningene ifølge foreliggende oppfinnelse kan også anvendes i andre indikasjoner hvor det er ønskelig å lege eller regenerere sene og/eller ligamentvev. Slike indikasjoner innbefatter, uten begrensning, regenerering eller reparering av skader på cementale fibre, så som ved tendonitt, og regenerering eller re-

parering av sene-til-ben kobling. Sammensetningene ifølge foreliggende oppfinnelse kan tilveiebringe et miljø for å tiltrekke sene- eller ligamentdannende celler, stimulere veksten av sene- eller ligamentdannende celler eller indusere differensiering av forløperne til sene- eller ligamentdannende celler.

BMP-12 relaterte proteiner kan bli isolert fra dyrkningsmediet og renses ved isolering av disse fra andre proteinholdige materialer som de blir produsert fra og fra andre kontaminanter som er tilstede. Proteinene ifølge foreliggende oppfinnelse har evnen til å inducere dannelsen av sene/ligamentlignende vev. Disse proteinene kan videre bli karakterisert ved evnen til å demonstrere sene/ligamentlignende vevsdannelsesaktivitet i rotte ektopisk implantatanalyse beskrevet nedenfor. Disse proteinene kan ha evnen til å inducere dannelsen av andre vevstyper, så som ligamenter.

Sene/ligamentlignende vevsinduserende proteiner gitt heri innbefatter også faktorer som blir kodet av sekvensene som er lik de til SEQ ID NO: 1 eller SEQ ID NO: 25, men hvor modifikasjoner er naturlig gitt (for eksempel alleliske variasjoner i nukleotidsekvensen som kan resultere i aminosyreforandringer i polypeptidet) eller med hensikt omkonstruert. For eksempel kan syntetiske polypeptider fullt eller delvis duplikere kontinuerlige sekvenser til aminosyreresidene ifølge SEQ ID NO: 2. Disse sekvensene i kraft av deres felles primære, sekundære eller tertiære strukturelle og konformasjonsmessige karaktertrekk med sene/ligamentlignende vevsvekstfaktor polypeptider ifølge SEQ ID NO: 2, kan ha sene/ligamentlignende eller andre vevsvekstfaktor biologiske egenskaper felles deri. De kan følgelig anvendes som biologiske aktive substituenten for naturlig forekommende sene/ligamentlignende vevsinduserende polypeptider i terapeutiske sammensetninger og prosesser.

Andre spesifikke mutasjoner i sekvensene til sene/ligamentlignende vevsinduserende proteiner beskrevet heri innbefatter modifikasjoner av glykosyleringsseter. Disse modifikasjonene kan innbefatte O-koblede eller N-koblede glykosyleringsseter. Blant annet resulterer fravær av glykosylering eller bare delglykosylering av aminosyresubstitusjon eller delesjon ved asparaginkoblede glykosyleringsgjenkjenningsseter. Asparaginkoblede glykosylerings gjenkjenningsseter omfatter tripeptidsekvenser som blir spesifikt oppdaget av hensiktsmessige cellulære glykosyleringsenzymene. Disse tripeptidsekvensene kan være asparagin-X-threonin, asparagin-X-serin eller asparagin-X-cystein, hvor X vanligvis er en hvilken som helst aminosyre unntatt prolin. Forskjellige aminosyre substitusjoner eller delesjoner ved en eller begge av første eller tredje aminosyreposisjonene til glykosyleringsgjenkjenningssete (og/eller aminosyredelesjon ved den andre posisjo-

nen) resulterer i ikke-glykosylering ved modifisert tripeptidsekvens. I tillegg vil bakteriell ekspresjon av proteinet også resultere i produksjonen av et ikke-glykosylert protein, selv hvis glykosyleringsssetene blir latt stå umodifiserte.

Sammensetningene ifølge foreliggende oppfinnelse omfatter et rensset BMP-12 relatert protein som kan bli produsert ved dyrking av en celle transformert med DNA-sekvensen ifølge SEQ ID NO: 1 eller SEQ ID NO: 25 og isolering og rensing av proteinet som har aminosyresekvensen til SEQ ID NO: 2 eller SEQ ID NO: 26 fra kulturmediet. Det rensede, uttrykte proteinet er vesentlig fri fra andre proteinholdige materialer som det blir produsert sammen med samt uten andre kontaminanter. Det isolerte rensede proteinet antas å utvise sene/ligamentlignende vevsdannelsesaktivitet, og annen vevs vekstaktivitet, så som ligament regenerering. Proteinene ifølge oppfinnelsen kan bli ytterligere karakterisert med evnen til å demonstrere sene/ligamentlignende vevsdannende aktivitet i rotteanalysen beskrevet nedenfor.

Sammensetninger for indusering av sene/ligamentlignende vevsdannelse ifølge oppfinnelsen kan omfatte en effektiv mengde av et sene/ligamentlignende vevsinduserende protein, hvor nevnte protein omfatter aminosyresekvensen ifølge SEQ ID NO: 2, fortrinnsvis aminosyrene nr -25, nr 1 eller nr 3 til nr 103 eller nr 104 til SEQ ID NO:2; eller aminosyrene nr 1 eller nr 19 til nr 120 til SEQ ID NO: 26; samt mutanter og/eller varianter ifølge SEQ ID NO: 2 eller SEQ ID NO: 26, som utviser evnen til å danne sene og/eller ligamentlignende vev.

Sammensetningene ifølge foreliggende oppfinnelse kan videre omfatte ytterligere proteiner, så som ytterligere medlemmer av TGF- β superfamilien av proteiner, så som aktiverer. Et annet aspekt ifølge oppfinnelsen tilveiebringer farmasøytiske sammensetninger inneholdende en terapeutisk effektiv mengde av et sene/ligamentinduserende protein, så som BMP-12 eller VL-1, i en farmasøytisk akseptabel bærer. Disse sammensetningene kan bli anvendt til å indusere dannelse av sene/ligamentlignende vev eller annet vev. Det antas at slike sammensetninger også kan bli anvendt for sene og ligament reparering, leging av sår og annen vevsreparering, så som hudreparering. Det blir videre betraktet at proteinene ifølge oppfinnelsen kan øke neuronaloverlevelsen og derfor være nyttig for transplantering og behandling av tilstander som utviser en reduksjon i neuronal overlevelse. Sammensetningene ifølge oppfinnelsen kan videre innbefatte minst et annet terapeutisk nyttig middel, så som BMP proteinene BMP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6 og BMP-7, blant annet beskrevet i US-PS 5.108.922; 5.013.649; 5.116.748; 5.106.748; 5.187.076; og 5.141.905; BMP-8, beskrevet i PCT

publikasjon WO91/18098; BMP-9, beskrevet i PCT publikasjon WO93/00432; og BMP-10 eller BMP-11, beskrevet i inngitte patentsøknader med serienr. 08/061.695 og 08/061/464, inngitt mai 12, 1993. Beskrivelsen av ovennevnte dokumenter er herved inkorporert som referanse.

Sammensetningene ifølge oppfinnelsen kan omfatte, i tillegg til et sene/ligamentinduserende protein så som BMP-12 eller VL-1 (BMP-13), andre terapeutiske nyttige midler som innbefatter MP52, epidermal vekstfaktor (EGF) fibroblast vekstfaktor (FGF), blodplateavledet vekstfaktor (PDGF), transformerende vekstfaktorer (TGF- α og TGF- β) og fibroblast vekstfaktor-4 (FGF-4), parathyroidhormon (PTH), leukemi inhibitorisk faktor (LIF/HILDA/DIA), insulinlignende vekstfaktorer (IGF-I og IGF-II). Deler av disse midlene kan også anvendes i sammensetningene ifølge oppfinnelsen. For eksempel gir en sammensetning som omfatter både BMP-2 og BMP-12 implantert sammen opphav til både ben og sene/ligamentlignende vev. En slik sammensetning kan være nyttig for behandling av defekter i embryonisk ledd hvor sene/ligamenter og ben dannes samtidig, ved siden av liggende anatomiske beliggenheter, og kan være nyttig for regenerering av vev ved stedet hvor senen blir festet til benet. Det er betraktet at sammensetningen ifølge oppfinnelsen kan være nyttig for leging av sår, så som leging av hud og beslektet vevsreparering. Sårtyper innbefatter, men er ikke begrenset til brannså, innsnitt og andre sår (se for eksempel PCT publikasjon WO84/01106 for en beskrivelse av leging av sår og beslektet vevsreparering).

Det antas at proteinene ifølge oppfinnelsen kan virke sammen med eller kanskje synergistisk med andre beslektede proteiner og vekstfaktorer. Ytterligere terapeutiske metoder og sammensetninger omfatter følgelig en terapeutisk mengde av minst et protein ifølge oppfinnelsen med en terapeutisk mengde av minst et av BMP-proteinene beskrevet ovenfor. Slike sammensetninger kan omfatte separate molekyler av BMP-proteiner eller heteromolekyler som omfatter forskjellige BMP deler. En fremgangsmåte og sammensetning ifølge oppfinnelsen kan for eksempel omfatte en disulfidkoblet dimer omfattende en BMP-12 relatert protein subenhet og i en enhet fra en av "BMP" proteinene beskrevet ovenfor. Foreliggende oppfinnelse omfatter følgelig sammensetninger som omfatter et renset BMP-12 relatert polypeptid som er en heterodimer hvor en subenhet omfatter aminosyresekvensen fra aminosyre nr. 1 til aminosyre nr. 104 til SEQ ID NO: 2, og en subenhet som omfatter en aminosyresekvens for et benmorfogent protein valgt fra gruppen omfattende BMP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6, BMP-7, BMP-8, BMP-9, BMP-10 og BMP-11. En ytterligere utførelsesform kan omfatte en heterodimer av disulfidbundet sene/ligamentlignende vevsinduserende deler så som

BMP-12, VL-1 (BMP-13) eller MP52. For eksempel kan heterodimeren omfatte en subenhet som omfatter en aminosyresekvens fra nr.1 til nr. 104 ifølge SEQ ID NO: 2 og den andre subenheten kan omfatte en aminosyresekvens fra nr. 1 til nr. 120 til SEQ ID NO: 4 eller nr. 1 til nr. 120 til SEQ ID NO: 26. Sammensetningene ifølge foreliggende oppfinnelse kan videre bli kombinert med andre midler som er nyttige for behandling av defekter, sår eller vev.

Preparering og formulering av slike fysiologiske akseptable proteinsammensetninger, men hensyn til pH, isotonisitet, stabilitet og lignende, hører inn under dette fagområdet. Terapeutiske sammensetninger er også verdifulle for veterinær anvendelser på grunn av mangel på artsspesifisitet i TGF- β proteinene. Husdyr og fullblodshester i tillegg til mennesker er ønskelige pasienter for slik behandling med sammensetningene ifølge foreliggende oppfinnelse.

Terapeutiske fremgangsmåter innbefatter administrering av sammensetningen topisk, systemisk eller lokalt som et implantat eller en innretning. Når administrert er den terapeutiske sammensetningen ifølge oppfinnelsen selvfølgelig i en pyrogenfri, fysiologisk akseptabel form. Sammensetningen kan fortrinnsvis bli innkapslet eller injisert i en viskøs form for levering til sted med vevsskade. Topisk administrering kan være egnet for leging av sår og reparering av vev. Terapeutiske nyttige midler forskjellige fra proteiner som også eventuelt kan bli innbefattet i sammensetningen som beskrevet ovenfor kan alternativt eller i tillegg, bli administrert samtidig eller sekvensielt med sammensetningen i fremgangsmåtene ifølge oppfinnelsen.

Sammensetningene kan også innbefatte en hensiktsmessig matrise og/eller sekvestreringsmiddel som en bærer. Matrisen kan blant annet understøtte sammensetningen eller utgjøre en overflate for sene ligamentlignende vevsdannelse og/eller annen vevsdannelse. Matrisen kan tilveiebringe sakte frigjøring av proteinet og/eller det hensiktsmessige miljøet for presentasjon derav. Sekvestreringsmiddelet kan være en forbindelse som hjelper administreringen ved injisering eller andre metoder, eller kan redusere migrering av proteinet fra applikasjonssetet.

Valg av et bærermateriale er basert på biokompatibilitet, biodegraderbarhet, mekaniske egenskaper, kosmetisk utseende og interfase egenskaper. Den bestemte applikasjonen av sammensetningene vil definere den hensiktsmessige formuleringen. Potensielle matriser for sammensetningene kan være biodegraderbare og kjemisk definerte. Ytterligere matriser består av rene proteiner eller ekstracellulære matrisekomponenter. Andre potensi-

elle matriser er ikke biodegraderbare og kjemisk definerte. Foretrukne matriser innbefatter kollagenbaserte materialer, så som Helistat svamp (Integra LifeSciences, Plainsboro, N.J.), eller collagen i en injiserbar form, samt sekvestreringsmidler, som også kan være biodegraderbare, og som kan innbefatte alkylcelluloseholdige materialer.

En annen foretrukket klasse av protein er porøse partikkelformige polymermatriser, inkludert polymerer av poly(melkesyre), poly(glykolsyre) og kopolymerer av melkesyre og glykolsyre. Disse matrisene kan også innbefatte et sekvestreringsmiddel. Egnede polymermatriser er for eksempel beskrevet i WO 93/00050, og beskrivelsen er inkorporert heri som referanse.

En foretrukket familie av sekvestreringsmidler er celluloseholdige materialer så som alkylcellulose (inkludert hydroksyalkylcellulose), inkludert metylcellulose, etylcellulose, hydroksyetylcellulose, hydroksypropylcellulose, hydroksypropylmetylcellulose og karboksymetylcellulose, og den mest foretrukne er kationiske salter av (CMC). Andre foretrukne sekvestreringsmidler innbefatter hyaluronsyre, natriumalginat, poly(etylen-glykol), polyoksyetylenoksyd, karboksyvinylpolymer og poly(vinylalkohol). Mengden av sekvestreringsmiddel som er nyttig heri er 0,5-20 vektprosent, fortrinnsvis 1-10 vektprosent basert på total formuleringsvekt, som representerer mengden som er nødvendig for å forhindre desorpsjon av proteinet fra polymermatrisen og tilveiebringe hensiktsmessig håndtering av sammensetningen, men ikke så mye at forløpercellene blir forhindret fra infiltrering av matrisen og derved gir proteinet muligheten av å assistere aktiviteten til forløpercellene.

Ytterligere valgfrie komponenter som er nyttige ved anvendelse av applikasjonen innbefatter for eksempel kryogene beskyttelsesmidler så som mannitol, sukrose, laktose, glukose eller glycin (for å beskytte proteinet fra degradering i løpet av lyofilisering), antimikrobielle konserveringsmidler så som metyl eller propylparabener og benzylalkohol; antioksyderingsmidler så som EDTA, sitrat og BHT (butylert hydroksytoluen); og overflateaktive midler så som poly(sorbater) og poly(oksetylener); osv.

Som beskrevet ovenfor kan sammensetningen ifølge oppfinnelsen bli anvendt i fremgangsmåter for behandling av et antall sendefekter, så som regenerering av sene/ligamentlignende vev i områder med sene eller ligamentskade, for å assistere i reparering av revning av senevev, ligamenter og forskjellige andre typer av vevsdefekter eller sår. Disse fremgangsmåtene omfatter administrering til en pasient som trenger et slik sene-/ligamentlignende vev eller annen vevsreparering av en sammensetning som omfatter en

effektiv mengde av et sene/ligamentlignende vevsinduserende protein, som beskrevet i SEQ ID NO: 2, SEQ ID NO: 4 og/eller SEQ ID NO: 26. Disse fremgangsmåtene kan også innbefatte administrering av et sene/ligamentlignende vevsinduserende protein sammen med minst et av BMP proteinene beskrevet ovenfor.

I en annen utførelsesform kan fremgangsmåtene innbefatte administrering av et heterodimerisk protein hvor en av monomerene er et sene/ligamentlignende vevsinduserende polypeptid så som BMP-12, VL-1 (BMP-13) eller MP52, og den andre monomeren er et medlem av TGF- β superfamilien av vekstfaktorer. I tillegg kan disse fremgangsmåtene også omfatte administrering av et sene/ligamentlignende vevsinduserende protein med andre vekstfaktorer inkludert EGF, FGF, TGF- α , TGF- β og IGF.

Et ytterligere aspekt av oppfinnelsen er en sammensetning for reparering av sene/ligamentlignende vev, for reparering av sene eller ligament samt behandling av tendinitt og andre tilstander relatert til sene eller ligament defekter. Slike sammensetninger omfatter en terapeutisk effektiv mengde av en eller flere sene/ligamentlignende vevsinduserende proteiner, så som BMP-12, et BMP-12 relatert protein, eller MP52, blandet sammen med en farmasøytisk akseptabel bærer eller matrise.

Doseringsregimet vil bestemmes av behandlende lege ved å betrakte forskjellige faktorer som modifierer virkningen til sammensetningen, for eksempel mengden av sene eller ligamentvev som det er ønskelig at skal bli dannet. Sted for sene eller ligament-skade, tilstanden til skadet sene eller ligament, størrelsen på såret, type skadet vev, pasientens alder, kjønn og diett, alvorligheten av eventuell infeksjon, administreringstid og andre kliniske faktorer. Doseringen kan variere med matrisetype som blir anvendt ved rekonstitusjon og type av ytterligere proteiner i sammensetningen. Tilsetning av andre kjente faktorer, så som IGF-1 (insulin lignende vekstfaktor I), til den endelige sammensetningen kan også påvirke doseringen.

Progresjonen kan bli registrert ved periodisk vurdering av sene/ligamentlignende vevsdannelse, eller sene eller ligamentvekst og/eller reparering. Progresjonen kan bli registrert ifølge fremgangsmåter kjent innenfor fagområdet for eksempel røntgenstråler, artroskopi, histomorfometriske bestemmelser og tetracyklin merking.

Foreliggende oppfinnelse omfatter et DNA-molekyl kjennetegnet ved at det har en sekvens som koder for et benmorfogenisk protein med evne til å indusere dannelse av sene/ligamentlignende vev, idet DNA-molekylet omfatter:

- (a) nukleotid nr 196, nr 199, nr 208, nr 217, nr 361, nr 388, nr 493, nr 496, nr 571 eller nr 577 til nr 879 eller nr 882 av SEQ ID NO: 1;
- (b) nukleotider kodende for aminosyrene nr. 25, nr 1 eller nr 3 til nr 103 eller nr 104 av SEQ ID NO: 2;
- (c) nukleotid nr 410, nr 458, nr 602, nr 605 eller nr 659 til nr 961 eller nr 964 av SEQ ID NO: 25;
- (d) nukleotider kodende for aminosyrene nr 1 eller nr 19 til nr 119 eller nr 120 av SEQ ID NO: 26;
- (e) naturlig forekommende aleliske sekvenser eller ekvivalente degenererte kodonsekvenser av hvilke som helst ifølge (a) til (d); eller
- (f) sekvenser som hybridiserer til hvilke som helst av (a) til (e) under standard hybridiseringsbetingelser, eller et komplement av disse.

Foreliggende oppfinnelse omfatter videre et kimært DNA-molekyl kjennetegnet ved at det omfatter en DNA-sekvens som koder for et propeptid fra et medlem av TGF- β superfamilien av proteiner koblet i riktig leseramme til en DNA-sekvens ifølge krav 1.

Foreliggende oppfinnelsen omfatter også en vektor kjennetegnet ved at den omfatter DNA-molekylet ifølge det ovenstående i operativ assosiasjon med en ekspresjonskontrollsekvens derav.

Videre omfatter foreliggende oppfinnelse en vertscelle kjennetegnet ved at den er transformert med DNA-molekyler ifølge det ovenstående.

Videre omfatter foreliggende oppfinnelse en fremgangsmåte for fremstilling av et BMP-12 eller BMP-12-relatert protein som har evnen til å indusere dannelse av sene/ligamentlignende vev kjennetegnet ved at den omfatter:

- (a) dyrking av vertscellen ifølge det ovenstående og
- (b) isolering og, eventuelt, rensing av nevnte BMP-12 eller BMP-12 relaterte protein fra kulturmediet.

Videre omfatter foreliggende oppfinnelse et BMP-12 eller BMP-12-relatert protein med evnen til å indusere dannelse av sene/ligamentlignende vev kjennetegnet ved at det kodes av DNA-molekylet ifølge det ovenstående eller produseres ifølge fremgangsmåten beskrevet over.

Foreliggende oppfinnelse omfatter videre et polypeptid kjennetegnet ved at det er i form av en dimer omfattende to subenheter, hver med aminosyresekvensen kodet av DNA-molekylet ifølge ovenstående.

Oppfinnelsen omfatter også et heterodimerisk proteinmolekyl som er kjennetegnet ved at det omfatter en monomer med aminosyresekvensen til polypeptidet som beskrevet ovenfor, og en monomer som har aminosyresekvensen til et protein fra TGF- β superfamilien.

Videre omfatter foreliggende oppfinnelse en farmasøytisk sammensetning kjennetegnet ved at den omfatter polypeptidet eller proteinet som beskrevet ovenfor, eventuelt blandet sammen med en farmasøytisk akseptabel bærer.

Videre omfatter oppfinnelsen anvendelse av polypeptidet eller proteinet ifølge det ovenstående for fremstilling av en sammensetning for indusering av sene/ligamentlignende vevsdannelse hos en pasient som trenger dette.

Foreliggende oppfinnelse vedrører til slutt anvendelse av polypeptidet eller proteinet som beskrevet ovenfor for fremstilling av en sammensetning for behandling av tendinitt, eller annen sene- eller ligamentdefekt, hos en pasient som trenger dette.

Følgende eksempler illustrerer utførelse av foreliggende oppfinnelse for isolering og karakterisering av humant sene/ligamentlignende vevsinduserende protein og anvendelse av disse for å isolere andre sene/ligamentlignende vevsinduserende proteiner, oppnåelse av humane proteiner, uttrykking av proteinene ved hjelp av rekombinante teknikker og demonstrering av evnen som sammensetningene ifølge foreliggende oppfinnelse har til å danne sene/ligamentlignende vev i en in vivo modell. Til tross for at eksemplene demonstrerer oppfinnelsen med hensyn på BMP-12 kan det med mindre modifikasjoner innenfor dette fagområdet oppnås samme resultater med MP52 og VL-1.

Eksempel 1

Isolering av DNA

DNA-sekvenser kodende for BMP-12 og BMP-12 relaterte proteiner kan bli isolert ved hjelp av forskjellige teknikker kjent innenfor fagområdet. Som beskrevet nedenfor kan oligonukleotidprimerene bli konstruert på grunnlag av aminosyresekvensene tilstede i andre BMP-proteiner, Vg-1 relaterte proteiner og andre proteiner fra TGF- β

superfamilien. Regioner inneholdende aminosyresekvenser som er sterkt konserverte med BMP familien av proteiner og innenfor andre medlemmer av TGF- β 1 superfamilien av proteiner kan bli identifisert og konsensus aminosyresekvensene til disse sterkt konserverte regionene kan bli konstruert basert på likheten til tilsvarende regioner til individuelle BMP/TGF- β /Vg-1 proteiner. Et eksempel på en slik konsensusaminosyresekvens er angitt nedenfor.

Konsensus aminosyresekvens (1):

Trp-Gln/Asn-Asp-Trp-Ile-Val/Ile-Ala (SEQ ID NO: 16)

hvor X/Y indikerer at hvilke som helst aminosyreresidie kan fremkomme ved den posisjonen.

Følgende oligonukleotid er konstruert på grunnlag av ovennevnte identifiserte konsensusaminosyresekvens (1):

Nr. 1 CGGATCCTGGVANGAYTGGATHRTNGC (SEQ ID NO:17)

Denne oligonukleotidsekvensen blir syntetisert på en automatisert DNA synthesizer. Standardnukleotidsymboler i ovennevnte identifiserte oligonukleotidprimere er som følger: A, adenosin; C, cytosin; G, guanin; T, tymidin, N, adenosin eller cytosin eller guanin eller tymin; R, adenosin eller cytosin; Y, cytosin eller tymin; H, adenosin eller cytosin eller tymin, V, adenosin eller cytosin eller guanin; D, adenosin eller guanin eller tymin.

De første 7 nukleotidene til oligonukleotid nr.1 (understreket) inneholder gjenkjenningssekvensen for restriksjonsendonuklease BamHI for å lette manipuleringen av en spesifikt amplifisert DNA-sekvens kodende for BMP-12 proteinet og blir følgelig ikke avledet fra konsensusaminosyresekvensen (1) presentert ovenfor.

En annen konsensus aminosyresekvens er avledet fra en annen sterkt konservert region av BMP/TGF- β /Vg-1 proteinene som beskrevet nedenfor:

His-Ala-Ile-Val/Leu-Gln-Thr (SEQ ID NO: 18)

Følgende oligonukleotid blir konstruert på grunnlag av ovenfor identifiserte konsensus aminosyresekvens (2):

nr.2 TTTCTAGAARNGTYTGNACDATNGCRTG (SEQ ID NO: 19)

Denne oligonukleotidsekvensen blir syntetisert på en automatisert DNA synthesizer. De samme nukleotidsymbolene blir anvendt som beskrevet ovenfor.

De første 7 nukleotidene til oligonukleotid nr. 1 (understreket) inneholder gjenkjenningssekvensen til restriksjonsendonuklease XbaI for å lette manipuleringen av en spesifikt amplifisert DNA-sekvens kodende for BMP-12 proteinet og er følgelig ikke avledet fra konsensus aminosyresekvensen (2) presentert ovenfor.

Det blir betraktet at BMP-12 proteinet ifølge oppfinnelsen og andre BMP/TGF- β /Vg-1 relaterte proteiner kan inneholde aminosyresekvenser som ligner konsensus aminosyresekvensene beskrevet ovenfor og beliggenheten av de sekvensene innenfor et BMP-12 protein eller andre nye relaterte proteiner vil tilsvare de relative beliggenhetene i proteinene som de er blitt avledet fra. Det er videre betraktet at denne posisjonelle informasjonen oppnådd fra strukturen til andre BMP/TGF- β /Vg-1 proteiner og oligonukleotidsekvensene nr. 1 og nr. 2 som er blitt avledet fra konsensus aminosyresekvensene (1) og (2) kan bli anvendt for spesifikk amplifisering av DNA-sekvensene som koder for de tilsvarende aminosyrene til BMP-12 proteinet eller andre BMP/TGF- β /Vg-1 relaterte proteiner.

Basert på kunnskapen av genstrukturene til BMP/TGF- β /Vg-1 proteinene er det videre betraktet at humant genomisk DNA kan bli anvendt som et templat for å utføre spesifikke amplifikasjonsreaksjoner som vil resultere i identifikasjon av BMP-12 BMP/TGF- β /Vg-1 (BMP-12 relatert protein) kodende sekvenser. Slike spesifikke amplifikasjonsreaksjoner til et humant genomisk DNA templat kan bli initiert ved anvendelse av oligonukleotidprimerene nr 1 og nr 2 beskrevet tidligere. Oligonukleotidene nr 1 og nr 2 identifisert ovenfor blir anvendt som primere for å muliggjøre spesifikk amplifikasjon av en spesifikk nukleotidsekvens fra humant genomisk DNA. Amplifikasjonsreaksjonen blir utført som følger:

humant genomisk DNA (kilde:perifere blodlymfocytter), oppnådd fra Ken Jacobs, Genetics Institute, blir spaltet ved gjentatt gjennomkjøring gjennom en 25 gauge nål, denaturert 100°C i 5 minutter og deretter avkjølt på is før det ble tilsatt til en reaksjonsblanding inneholdende 200 μ m av hver av deoksynukleotidtrifosfatene (dATP, dGTP, dCTP og dTTP), 10 mM tris-HCl pH 8,3, 50 mM KCl, 1,5 mM MgCl₂, 0,001%v gelatin, 1,25 enheter taq DNA polymerase, 100 pM oligonukleotid nr 1 og 100 pM oligonukleotid nr 2. Denne reaksjonsblandingen blir inkubert ved 94°C i 2 minutter og deretter utsatt for termisk cyklisering på følgende måter: 1 minutt ved 94°C, 1 minutt ved 40°C, 1 minutt ved 72°C i tre sykluser; og deretter 1 minutt ved 94°C, 1 minutt ved 55°C, 1 minutt ved 72°C i 37 sykluser, etterfulgt av 10 minutter inkubasjon ved 72°C.

DNA som blir spesifikt amplifisert ifølge denne reaksjonen blir etanolpresipitert, spaltet med restriksjonsendonukleasene BamHI og XbaI og deretter utsatt for agarosegelelektroforese. En region av gelen som tilsvarer den antatte størrelsen til BMP-12 eller annet BMP/TGF- β /Vg-1 kodende DNA-fragment, blir spaltet ut og spesifikt amplifiserte DNA-fragmenter innbefattet deri blir elektroeluert og subklonet inn i plasmidvektor pGEM-3 mellom XbaI og BamHI setene til polylinkeren. DNA-sekvens analysene av en av de resulterende BMP-12 relaterte subklonene indikerer spesifikt amplifisert DNA-sekvens produkt innbefattet deri som koder for en del av BMP-12 proteinet ifølge oppfinnelsen.

DNA-sekvensen (SEQ ID NO: 5) og oppnådd aminosyresekvens (SEQ ID NO:6) til dette spesifikt amplifiserte DNA-fragmentet til BMP-12 er vist i sekvenslisten.

Nukleotidene nr.1-nr.26 til SEQ ID NO: 5 omfatter en del av oligonukleotid nr. 1 og nukleotidene nr 103 - nr 128 omfatter en del av omvendte komplement til oligonukleotid nr. 2 anvendt for å utføre den spesifikke amplifikasjonsreaksjonen. På grunn av funksjonen til oligonukleotidene nr 1 og nr 2 for initiering av amplifikasjonsreaksjonen behøver ikke å korrespondere nøyaktig med den virkelige sekvensen kodende for et BMP-12 protein og blir følgelig ikke translatert i den tilsvarende aminosyrederivasjonen (SEQ ID NO: 6).

DNA-sekvensanalyser av en annen subklon indikerer at spesifikt amplifisert DNA produkt innbefattet deri koder for en del av et annet BMP/TGF- β /Vg-1 (BMP-12 relatert) protein ifølge oppfinnelsen betegnet VL-1.

DNA-sekvensen (SEQ ID NO:7) og avledet aminosyresekvens (SEQ ID NO: 8) ifølge dette spesifikt amplifiserte DNA-fragmentet er vist i sekvenslisten.

Nukleotidene nr 1 - nr 26 til SEQ ID NO: 7 omfatter en del av oligonukleotid nr 1 og nukleotidene nr 103 - nr 128 omfatter en del av omvendte komplement til oligonukleotid nr 2 anvendt for å utføre den spesifikke amplifikasjonsreaksjonen. På grunn av funksjonen til oligonukleotidene nr 1 og nr 2 når det gjelder initiering av amplifikasjonsreaksjonen behøver det ikke å tilsvare nøyaktig til den sanne sekvensen kodende for et VL-1 protein ifølge oppfinnelsen og blir ikke translatert til den tilsvarende aminosyrederivasjonen (SEQ ID NO:8).

Følgende oligonukleotidprobe blir konstruert på grunnlag av den spesifikt amplifiserte BMP-12 humane DNA-sekvensen som angitt ovenfor (SEQ ID NO: 5) og syntetisert på en automatisert DNA synthesizer:

nr 3: CCACTGCGAGGGCCTTTGCGACTTCCCTTTGCGTTCGCAC (SEQ ID NO: 20).

Denne oligonukleotidproben blir radioaktivt merket med ^{32}P og anvendt for å screene et humant genomisk bibliotek konstruert i vektor λFIX (Stratagene katalog nr. 944291). 500.000 rekombinanter av humant genomisk bibliotek blir sådd ut i en tetthet på omtrent 10.000 rekombinanter pr skål på 50 skåler. Duplikate nitrocellulose replikaer av rekombinante bakteriofag plakk og hybridisert til oligonukleotidprobe nr 3 i standard-hybridiseringsbuffer (SHB = 5X SSC, 0,1% SDS, 5X Denhardt's, 100 $\mu\text{g/ml}$ laksesperm DNA) ved 65°C over natt. Dagen deretter blir radioaktivt merket oligonukleotid inneholdende hybridiseringsoppløsning fjernet og filtrere blir vasket med 0,2X SSC, 0,1% SDS ved 65°C. En enkelt positivt hybridiseringsrekombinant blir identifisert og plakkrenset. Denne plakkrensede rekombinante bakteriofagklonen som hybridiserer til BMP-12 oligonukleotidprobe nr 3 blir betegnet $\lambda\text{HuG-48}$. En bakteriofag platestokk blir dannet og bakteriofag DNA blir isolert fra $\lambda\text{HuG-48}$ human genomisk klon. Bakteriofagen $\lambda\text{HuG-48}$ er blitt deponert til American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD "ATCC" under aksjesjonsnr. 75625 desember 7, 1993. Denne deponeringen fyller kravene ifølge Budapest Treaty of the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and Regulations. Oligonukleotidhybridiserende region til denne rekombinanten $\lambda\text{HuG-48}$ er lokalisert til et 3,2 kb BamHI fragment. Dette fragmentet blir subklonet inn i plasmidvektor (pGEM-3) og DNA-sekvensanalyser blir utført. Denne plasmidsubklonen blir betegnet PCR1-1nr2 og er blitt deponert til American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD "ATCC" under aksjesjon nr 69517, desember 7, 1993. Denne deponeringen oppfyller kravene ifølge Budapest Treaty of the International Recognition og the Deposit of Microorganisms for the Purpose of Patent Procedure and Regulations. Del DNA-sekvensen (SEQ ID NO: 1) og avledet aminosyresekvens (SEQ ID NO: 2) til 3,2 kb DNA innskuddet til plasmid subklon PCR1-1nr2, avledet fra klon $\lambda\text{HuG-48}$, er vist i sekvenslistene.

Det er å bemerke at nukleotidene nr 639 - nr 714 til SEQ ID NO: 1 tilsvarer nukleotidene nr 27 - nr 102 ifølge spesifikt amplifisert BMP-12 kodende DNA-fragment angitt i SEQ ID NO: 5 og bekrefter følgelig at den humane genomiske bakteriofagklonen $\lambda\text{HuG-48}$ og avledede subklon PCR1-1nr2 koder for minst en del av BMP-12 proteinet

ifølge oppfinnelsen. Nukleotidsekvensen til en del av 3,2 kb BamHI innskuddet til plasmid PCR1-1nr2 inneholder en åpen leseramme med minst 882 basepar, som definert ved nukleotidene nr 1- nr 882 til SEQ ID NO: 1. Denne åpne leserammen koder for minst 294 aminosyrer av humant BMP-12 protein ifølge oppfinnelsen. Kodet 294 aminosyre humant BMP-12 protein innbefatter et fullstendig modent humant BMP-12 protein (aminosyrene nr 1 - nr 104 til SEQ ID NO: 2) samt den C-terminale delen av propeptidregionen til det primære translasjonsproduktet (aminosyre nr -190 til nr -1 til SEQ ID NO: 2).

Ytterligere DNA-sekvens til 3,2 kb BamHI innskuddet til plasmid PCR1- 1nr2 angitt i SEQ ID NO: 33 demonstrerer tilstedeværelse av en 1164 bp åpen leseramme, som definert ved nukleotidene nr 138 til og med nr 1301 ifølge SEQ ID NO: 33 (det er å bemerke at hele sekvensen beskrevet i SEQ ID NO: 11 er innbefattet i SEQ ID NO: 33). Fordi denne sekvensen er avledet fra en genomisk klon er det vanskelig å bestemme grensen mellom 5' omfanget av den kodende sekvensen og 3' grensen til mellomliggende sekvens (intron/ikke-kodende sekvens).

Basert på kunnskapen om andre BMP proteiner og andre proteiner innenfor TGF- β familien antas det at forløperpolypeptidet kan bli spaltet ved flerbasisisk sekvens Arg-Arg-Gly-Arg i samsvar med en foreslått konsensus proteolytisk prosesserende sekvens Arg-X-X-Arg. Spaltning av BMP-12 forløperpolypeptidet er antatt å danne et 104 aminosyre modent peptid begynnende med aminosyren Ser i posisjon nr 1 ifølge SEQ ID NO: 2. Prosessering av BMP-12 til den modne formen antas å innbefatte dimerisering og fjerning av den N-terminale regionen på en måte som er analog med prosessering av beslektet protein TGF- β [Gentry et al., *Molec & Cell.Biol.*, 8:4162 (1988); Derynck et al., *Nature*, 316:701 (1985)].

Det antas derfor at modne aktive former av BMP-12 omfatter en homodimer med to polypeptidsubenheter hvor hver subenhet omfatter aminosyrene nr 1 til nr 104 ifølge SEQ ID NO: 2 med en antatt molekylvekt på omtrent 12.000 dalton. Ytterligere aktive former antas å omfatte minst aminosyrene nr 3 til nr 103 ifølge SEQ ID NO: 2 og følgende innbefatter første og siste konserverte cysteinresidie. Som andre medlemmer av TGF- β /BMP familien av proteinene utviser den karboksy-terminale delen av BMP-12 proteinet høyere sekvenskonservering enn den mer amino-terminale delen. Prosent aminosyreidentitet til humant BMP-12 protein i det cystein-rike C-terminale domenet (aminosyrene nr 3 - nr 104) med tilsvarende region av humane BMP-proteiner og andre proteiner innenfor TGF- β familien er som følger: BMP-2, 55%; BMP-3, 43%, BMP-4,

53%; BMP-5, 49%; BMP-6, 49%; BMP-7, 50%; BMP-8, 57%; BMP-9, 48%; BMP-10, 57%; aktivin WC (BMP-11), 38%; Vg1, 46%; GDF-1, 47%; TGF- β 1; 36%; TGF- β 2, 36%; TGF- β 3, 39%; inhibin β (B), 36%; inhibin β (A), 41%.

Human BMP-12 DNA-sekvens (SEQ ID NO: 1), eller en del derav, kan anvendes som en probe for å identifisere en human cellelinje eller et vev som syntetiserer BMP-12 mRNA. Kort forklart blir RNA ekstrahert fra valgte celle eller vevskilde og enten elektroforet på en formaldehydagarosegel og overført til nitrocellulose, eller omsatt med formaldehyd og applisert direkte på nitrocellulose. Nitrocellulose blir deretter hybridisert til en probe avledet fra den kodende sekvensen til human BMP-12.

Human BMP-12 sekvens blir alternativt anvendt for å konstruere oligonukleotidprimere som vil spesifikt amplifisere en del av BMP-12 kodende sekvens beliggende i regionen mellom primerene anvendt for å utføre den spesifikke amplifikasjonsreaksjonen. Det antas at disse humane BMP-12 avledede primerene vil muliggjøre spesifikk amplifisering av tilsvarende BMP-12 kodende sekvenser fra mRNA, cDNA eller genomiske DNA-templater. Når en positiv kilde er blitt identifisert ifølge en av de ovennevnte metodene blir mRNA selektert ved oligo (dT) cellulosekromatografi og cDNA blir syntetisert og klonet i λ gt10 eller λ bakteriofag vektorer kjent for fagfolk innenfor dette området, for eksempel, λ ZAP ifølge kjente teknikker (Toole et al., supra). Det er også mulig å utføre oligonukleotidprimerrettet amplifikasjonsreaksjon, beskrevet ovenfor, direkte på forhånd etablerte humane cDNA eller genomiske bibliotek som er blitt klonet inn i en λ bakteriofag vektor. I slike tilfeller vil et bibliotek som tilveiebringer et spesifikt amplifisert DNA produkt kodende for en del av humant BMP-12 protein bli screenet direkte ved anvendelse av fragmentet av amplifisert BMP-12 kodende DNA som en probe.

Oligonukleotidprimerene konstruert på grunnlag av DNA-sekvensen til human BMP-12 genomisk klon λ HuG-48 antas å muliggjøre spesifikk amplifikasjon av humane BMP-12 kodende DNA-sekvenser fra på forhånd etablerte humane cDNA bibliotek som er kommersielt tilgjengelige (Stratagene, La Jolla, CA eller Clontech Laboratories, Inc., Palo Alto, CA). Følgende oligonukleotidprimer blir konstruert på grunnlag av nukleotidene nr 571 til nr 590 til DNA-sekvensen angitt i SEQ ID NO: 1 og syntetisert på en automatisert DNA synthesizer:

nr 4 TGCGGATCCAGCCGCTGCAGCCGCAAGCC (SEQ ID NO: 21).

De første ni nukleotidene til primer nr 4 (understreket) omfatter gjenkjenningssekvensen for restriksjonsendonuklease BamHI som kan bli anvendt for å lette manipuleringen av en spesifikt amplifisert DNA-sekvens kodende for humant BMP-12 protein ifølge oppfinnelsen og er følgelig ikke avledet fra DNA-sekvensen presentert i SEQ ID NO: 1.

Følgende oligonukleotidprimer blir konstruert på grunnlag av nukleotidene nr 866 - nr 885 til DNA-sekvensen angitt i SEQ ID NO: 1 og syntetisert på en automatisert DNA synthesizer.

nr 5 GACTCTAGACTACCTGCAGCCGCAGGCCT (SEQ ID NO: 22)

De første ni nukleotidene til primer nr 5 (understreket) omfatter gjenkjenningssekvensen for restriksjonsendonuklease XbaI som kan bli anvendt for å lette manipuleringen av en spesifikt amplifisert DNA-sekvens kodende for humant BMP-12 protein ifølge oppfinnelsen og er følgelig ikke avledet fra DNA-sekvensen presentert i SEQ ID NO: 1.

Standardnukleotidsymboler i ovenfor identifiserte primere er som følger: A, adenin; C, cytosin; G, guanin; T, thymin.

Primerne nr 4 og nr 5 identifisert ovenfor blir anvendt som primere for å muliggjøre amplifikasjon av en spesifikk BMP-12 kodende nukleotidsekvens fra på forhånd etablerte cDNA-bibliotek som kan innbefatte følgende: human føtal hjerne cDNA/ λ ZAPII (Stratagene katalog nr. 936206), human lever/ λ UNI-ZAP XR (Stratagene katalog nr. 937200), human lunge/ λ UNI-ZAP XR (Stratagene katalog nr. 937206) og human føtal milt/ λ UNI-ZAP XR (Stratagene katalog nr. 937205).

Omtrent 1×10^8 pfu (plakkdannende enheter) av λ bakteriofag-bibliotek som inneholder humane cDNA innskudd som de som er angitt ovenfor blir denaturert ved 95°C i 5 minutter før tilsetning til en reaksjonsblanding inneholdende 200 μ M av hver av deoksynukleotidtrifosfatene (dATP, dGTP, dCTP og dTTP) 10 mM tris-HCl pH 8,3, 50 mM KCl, 1,5 mM $MgCl_2$, 0,001% gelatin, 1,25 enheter Taq DNA polymerase, 100 pM oligonukleotidprimer nr 4 og 100 pM oligonukleotidprimer nr 5. Reaksjonsblandingen blir deretter utsatt for termisk cyklisering på følgende måte: 1 minutt ved 94°C, 1 minutt ved 50°C, 1 minutt ved 72°C i 39 sykluser etterfulgt av 10 minutter ved 72°C.

DNA som blir spesifikt amplifisert ifølge denne reaksjonen ventes å danne et BMP-12 kodende produkt på omtrent 333 basepar, hvor indre 315 bp tilsvarer nukleotidene nr 571 til nr 885 til SEQ ID NO: 1 og også innbefatter 9 bp ved hver ende av BMP-12

spesifikt fragment som tilsvarer restriksjonssetene definert ved nukleotidene nr 1 - nr 9 til primerene nr 4 og nr 5. Resulterende 333 bp DNA produkt blir spaltet med restriksjonsendonukleasene BamHI og XbaI, fenolekstrahert, kloroformekstrahert og etanolpresipitert.

I forhold til etanolpresipitering blir bufferskift og fjerning av små fragmenter av DNA resulterende fra BamHI/XbaI restriksjonsspaltningen oppnådd ved fortykning av spaltet DNA produkt i 10 mM tris-HCl pH 8,0, 1 mM EDTA etterfulgt av sentrifugering gjennom en CentriconTM 30 mikrokonsentrator (W.R.Grace & Co., Beverly, MA; Product nr 4209). Resulterende BamHI/XbaI spaltet amplifisert DNA produkt blir subklonet inn i en plasmidvektor (dvs. pBluescript, pGEM-3 etc) mellom BamHI og XbaI setene til polylinkerregionen. DNA-sekvensanalyse av resulterende subkloner vil være nødvendig for å bekrefte integriteten til BMP-12 kodende innskudd. Når en positiv cDNA kilde er blitt identifisert på denne måten ble tilsvarende cDNA bibliotek hvorfra en 333 bp BMP-12 spesifikk sekvens ble amplifisert screenet direkte med 333 bp innskuddet eller andre BMP-12 spesifikke prober for å identifisere og isolere cDNA kloner kodende for full-lengde BMP-12 proteinet ifølge oppfinnelsen.

Ytterligere fremgangsmåter kjent for fagfolk innenfor dette området kan bli anvendt for å isolere andre full-lengde cDNA kodende humane BMP-12 relaterte proteiner, eller full-lengde cDNA kloner kodende for BMP-12 relaterte proteiner ifølge oppfinnelsen fra arter som er forskjellige fra mennesker, spesielt andre pattedyrarter.

Følgende eksempler demonstrerer anvendelse av human BMP-12 sekvens for å isolere homologer fra BMP-12 relaterte proteiner i et murint genomisk DNA bibliotek.

DNA-sekvensen som koder for humant BMP-12 protein ifølge oppfinnelsen antas å være betydelig homolog med BMP-12 og BMP-12-relaterte sekvenser fra arter som er forskjellige fra mennesker slik at det kan bli anvendt for spesifikk amplifisering av DNA-sekvenser fra andre arter som koder for tilsvarende BMP-12 relaterte proteiner. Følgende oligonukleotider blir konstruert på grunnlag av human BMP-12 sekvens (SEQ ID NO: 1) og blir syntetisert på en automatisert DNA synthesizer:

nr 6: GCGGATCCAAAGGAGCTCGGCTGGGACGA (SEQ ID NO: 23)

nr.7: GGAATTCCCCACCACCATGTCCTCGTAT (SEQ ID NO: 24)

De første 8 nukleotidene til oligonukleotidprimerene nr 6 og nr 7 (understreket) omfatter gjenkjenningsssekvensen til restriksjonsendonukleasene BamHI og EcoRI. Disse sekven-

sene blir anvendt for å lette manipulering av en spesifikk amplifisert DNA-sekvens kodende for et BMP-12 eller BMP-12-relatert protein fra en art som er forskjellig fra mennesker og er følgelig ikke avledet fra DNA-sekvensen presentert i SEQ ID NO: 1. Oligonukleotidprimer nr 6 blir konstruert på grunnlag av nukleotidene nr 607 - nr 626 til SEQ ID NO: 1. Oligonukleotidprimer nr 7 blir konstruert på grunnlag av omvendt komplement til nukleotidene nr 846 - nr 865 av DNA-sekvensen angitt i SEQ ID NO: 1.

Oligonukleotidprimerene nr 6 og nr 7 identifisert ovenfor blir anvendt som primere for å muliggjøre amplifikasjon av spesifikke BMP-12 relaterte sekvenser fra genomisk DNA avledet fra arter forskjellige fra mennesker. Amplifikasjonsreaksjonen blir utført som følger:

murin genomisk DNA (kilde: stamme Balb c) blir spaltet ved gjentatt føring gjennom en 25 gauge nål, denaturert ved 100°C i 5 minutter og deretter avkjølt på is før tilsetning til en reaksjonsblanding inneholdende 200 µm av hver av deoksynukleotidtrifosfatene (dATP, dGTP, dCTP og dTTP) 10 mM tris-HCl pH 8,3, 50 mM KCl, 1,5 mM MgCl₂, 0,001% gelatin, 1,25 enheter Taq DNA polymerase, 100 pM oligonukleotidprimer nr 6 og 100 pM oligonukleotidprimer nr 7. Reaksjonsblandingen blir deretter utsatt for termisk cyklisering på følgende måte: 1 min. ved 95°C, 1 min. ved 55°C, 1 min. ved 72°C i 40 sykluser etterfulgt av 10 min. ved 72°C.

DNA som blir spesifikt amplifisert ifølge denne reaksjonen blir etanolpresipitert, spaltet med restriksjonsendonukleasene BamHI og EcoRI og utsatt for agarose gelelektroforese. En region av gelen, tilsvarende antatt størrelse av murin BMP-12 eller BMP-1-relatert kodende DNA-fragment, blir spaltet ut og spesifikt amplifiserte DNA-fragmenter innbefattet deri blir ekstrahert (ved elektroeluering eller ved hjelp av andre kjente metoder) og subklonet inn i en plasmidvektor, så som pGEM-3 eller pBluescript mellom BamHI og EcoRI setene til polylinkeren. DNA-sekvensanalyser av en av de resulterende subklonene betegnet mV1, indikerer at den spesifikt amplifiserte DNA-sekvensen innbefattet deri koder for en del av et protein som ser ut til å være murin homolog med enten BMP-12 eller VL-1 sekvensen ifølge oppfinnelsen. DNA-sekvensen (SEQ ID NO: 10) og avledet aminosyresekvens (SEQ ID NO: 11) til dette spesifikke amplifiserte murine DNA-fragmentet er vist i sekvenslisten.

Nukleotidene nr 1 - nr 26 til SEQ ID NO: 10 omfatter en del av oligonukleotid nr 6 og nukleotidene nr 246 - nr 272 omfatter en del av omvendt komplement ("reverse complement") av oligonukleotid nr. 7 anvendt for å utføre den spesifikke amplifika-

sjonsreaksjonen. Nukleotid nr. 27 til SEQ ID NO: 10 ser ut til å være det siste nukleotidet fra en kodontriplt, og nukleotidene nr 244 - nr 245 til SEQ ID NO: 10 ser ut til å være de første 2 nukleotidene til en kodontriplt. Nukleotidene nr 28 til nr 243 til SEQ ID NO: 10 tilsvarer følgelig en delkodende sekvens av mV1. På grunn av funksjonen til oligonukleotidene nr 6 og nr 7 for initiering av amplifikasjonsreaksjonen behøver disse ikke korrespondere nøyaktig med den aktuelle sekvensen som koder for murin homolog til human BMP-12 eller VL-1 protein ifølge oppfinnelsen og blir følgelig ikke translatert i tilsvarende aminosyrederivasjon (SEQ ID NO: 11).

Oligonukleotidprober konstruert på grunnlag av spesifikk amplifisert murin BMP-12 eller VL-1 DNA-sekvens er angitt i SEQ ID NO: 10 kan bli anvendt av fagfolk innenfor dette området for å identifisere full-lengde murin BMP-12 eller VL-1 kodende klon (enten cDNA eller genomisk).

DNA-sekvensanalyser av andre resulterende subkloner betegnet mV2 indikerer at spesifikk amplifisert DNA-sekvens innbefattet deri koder for en del av en murin BMP-12 relatert sekvens ifølge oppfinnelsen. DNA-sekvensen (SEQ ID NO: 12) og avledede aminosyresekvens (SEQ ID NO: 13) ifølge det spesifikt amplifiserte murine DNA-fragmentet er vist i sekvenslistene.

Nukleotidene nr 1 - nr 26 til SEQ ID NO: 12 omfatter en del av oligonukleotid nr 6 og nukleotidene nr 246 - nr 272 omfatter en del av omvendte komplement til oligonukleotid nr 7 anvendt for å utføre den spesifikke amplifikasjonsreaksjonen. Nukleotid nr 27 til SEQ ID NO: 12 ser ut til å være det siste nukleotidet av en kodontriplt, og nukleotidene nr 244 - nr 245 til SEQ ID NO: 12 der ut til å være de første to nukleotidene til en kodontriplt. Nukleotidene nr 28 - nr 243 ifølge SEQ ID NO: 12 tilsvarer følgelig en delkodende sekvens av mV2. På grunn av funksjonen til oligonukleotidene nr 6 og nr 7 når det gjelder initiering av amplifikasjonsreaksjonen så behøver disse ikke nøyaktig tilsvare den virkelige sekvensen som koder for muring BMP-12 relatert protein ifølge oppfinnelsen og blir følgelig ikke translatert i tilsvarende aminosyresekvensderivasjon (SEQ ID NO: 13).

Oligonukleotidprobene konstruert på grunnlag av spesifikk amplifisert murin BMP-12 relatert DNA-sekvens angitt i SEQ ID NO: 12 kan bli anvendt av fagfolk innenfor dette området for å identifisere full-lengde murin BMP-12 relatert kodende kloner (enten cDNA eller genomisk).

DNA-sekvensanalyse av en annen av resulterende subkloner betegnet mV9 indikerer at den spesifikke amplifiserte DNA-sekvensen innbefattet deri koder for en del av en murin BMP-12 relatert sekvens ifølge oppfinnelsen. Denne sekvensen ser ut til å være den murine homologen til human MP52 DNA-sekvens beskrevet i SEQ ID NO: 3. DNA-sekvensen (SEQ ID NO:14) og avledet aminosyresekvens (SEQ ID NO: 15) ifølge dette spesifikke amplifiserte amplifiserte murine DNA-fragmentet er vist i sekvenslisten.

Nukleotidene nr 1 - nr 26 til SEQ ID NO: 14 omfatter en del av oligonukleotid nr 6 og nukleotidene nr 246 - nr 245 til SEQ ID NO: 14 omfatter en del av revers komplement til oligonukleotid nr 7 anvendt for å utføre den spesifikke amplifikasjonsreaksjonen. Nukleotid nr 27 til SEQ ID NO:14 ser ut til å være det siste nukleotidet til en kodontriplett, og nukleotidene nr 244 - nr 245 til SEQ ID NO: 14 ser ut til å være de første to nukleotidene til en kodontriplett. Nukleotidene nr 28 til nr 243 til SEQ ID NO:14 tilsvarer en delvis kodende sekvens til mV9. På grunn av funksjonen til oligonukleotidene nr 6 og nr 7 når det gjelder initiering av amplifikasjonsreaksjonen er det ikke sikkert at de tilsvarer nøyaktig den virkelige sekvensen kodende for murin BMP-12 relatert protein ifølge oppfinnelsen og blir følgelig ikke translatert i den tilsvarende aminosyresekvensderivasjonen (SEQ ID NO: 15).

Oligonukleotidprober konstruert på grunnlag av spesifikt amplifisert murin BMP-12 relatert DNA-sekvens angitt i SEQ ID NO: 14 kan bli anvendt av fagfolk innenfor dette området for å identifisere full-lengde murin BMP-12 relaterte kodende kloner (enten cDNA eller genomisk).

Alternativt blir oligonukleotidprimerene nr 6 og nr 7 identifisert ovenfor anvendt som primere for å muliggjøre spesifikk amplifikasjon av en 275 basepar DNA probe hvor indre 259 bp tilsvarer nukleotidene nr 607 til nr 865 av SEQ ID NO: 1, fra BMP-12 kodende plasmid subklon PCR1-1 nr 2. Denne 275 bp DNA proben ble radioaktivt merket med ^{32}P og anvendt for å screene et murint genomisk bibliotek konstruert i vektor λ FIX II (Stratagene katalog nr. 946306). 1 million rekombinanter av murin genomisk bibliotek ble sådd ut i en tetthet på omtrent 20.000 rekombinanter pr skål på 50 skåler. Duplikate nitrocellulosereplikaer av rekombinante bakteriofagplakk ble hybridisert, under reduserte stringensbetingelser, til spesifikk amplifisert 33 bp probe i standard hybridiseringsbuffer (SHB = 5X SSC, 0,1% SDS, 5X Denhardt's, 100 $\mu\text{g/ml}$ laksesperm DNA) ved 60°C over natten. Neste dag ble radioaktivt merket oligonukleotid inneholdende hybridiseringsoppløsning fjernet og filtrere vasket under reduserte stringensbetingelser med 2X SSC, 0,1% SDS ved 60°C. Flere positivt hybridiseringsrekombinanter

blir identifisert og plakkrenset. Fragmenter av positivt hybridiserende murine genomiske rekombinant kloner blir subklonet inn i standardplasmidvektorer (dvs. pGEM-3) og ut-satt for DNA-sekvensanalyser.

DNA-sekvensanalyser av en av disse subklonene betegnet MVR3 indikerer at det koder for en del av musegenet som tilsvare PCR produktet mV1 (murin homolog av human BMP-12 sekvens angitt i SEQ ID NO: 1) beskrevet ovenfor. Del DNA-sekvensen av denne subklonen og tilsvarende aminosyretranslasjon er angitt i SEQ ID NO: 29 og SEQ ID NO: 30.

DNA-sekvensanalyser av en annen av disse subklonene betegnet MVR32 indikerer at det koder for en del av musegenet som tilsvare PCR produktet mV2 (murin homolog av human VL-1 sekvensen angitt i SEQ ID NO: 7) beskrevet ovenfor. Del DNA-sekvensen av denne subklonen og tilsvarende aminosyretranslasjon er angitt i SEQ ID NO: 31 og SEQ ID NO: 32.

DNA-sekvensanalyser av en annen av disse subklonene betegnet MVR23 indikerer at den koder for en del av musegenet som tilsvare PCR produktet mV9 (murin homolog av MP-52 sekvensen angitt i SEQ ID NO: 3) beskrevet ovenfor.

På lignende måte som beskrevet ovenfor for identifisering og isolering av humane genomiske kloner kodende for BMP-12 proteinet ifølge oppfinnelsen kan oligonukleotidprober tilsvarende VL-1 kodende sekvens angitt i SEQ ID NO: 7 bli konstruert og anvendt for å identifisere humane genomiske eller cDNA-sekvenser kodende for VL-1 (BMP-13) protein. Disse oligonukleotidene kan bli konstruert til regioner spesifikke for VL-1 kodende sekvenser og vil derfor sannsynligvis være avledet fra regioner av lavest grad av nukleotidsekvensidentitet mellom spesifikke amplifiserte VL-1 kodende sekvens (SEQ ID NO: 7) og spesifikke amplifisert BMP-12 kodende sekvens (SEQ ID NO: 5).

Alternativt blir oligonukleotidprimerene nr 4 og nr 5 identifisert ovenfor anvendt som primere for å muliggjøre spesifikke amplifikasjon av en 333 basepar DNA probe hvor indre 315 bp tilsvare nukleotidene nr 571 til nr 885 til SEQ ID NO: 1, fra BMP-12 kodende plasmidsubklon PCR1-1 nr 2. Denne 333 bp DNA proben ble radioaktivt merket med ³²P og anvendt for å screene et humant genomisk bibliotek konstruert i vektoren λ DASH II (Stratagene katalog nr. 945203). 1 million rekombinanter av humant genomisk bibliotek blir sådd ut i en tetthet på omtrent 20.000 rekombinanter pr skål på 50 skåler. Duplikate nitrocellulose replikaer av rekombinant bakteriofage plakk blir hybridisert,

under reduserte stringensbetingelser, til spesifikk amplifisert 333 bp probe i standard-hybridiseringsbuffer (SHB = 5X SSC, 0,1% SDS, 5X Denhardt's, 100 µg/ml laksesperm DNA) ved 60°C over natt. Neste dag blir radioaktivt merket oligonukleotid inneholdende hybridiseringsoppløsning fjernet og filtrene vasket under reduserte stringensbetingelser med 2X SSC, 0,1% SDS ved 60°C. Flere (omtrent 15) positivt hybridiserende rekombinante blir identifisert og plakkrenset.

For å skjelne positivt hybridiserende rekombinanter kodende for VL-1 proteinet ifølge oppfinnelsen fra BMP-12 og andre BMP-12 relaterte kodende rekombinanter som antas å hybridisere positivt til 333 bp DNA proben dannet fra BMP-12 kodende plasmid PCR1-1 nr 2 anvendt i denne screeningsprosedyren, etter oligonukleotidproben, basert på VL-1 sekvensen angitt i SEQ ID NO: 7, er konstruert og syntetisert på en automatisert DNA synthesizer:

nr. 8 TGTATGCGACTTCCCGC (SEQ ID NO: 35)

Et oligonukleotid som tilsvare oligonukleotidene nr 60 til nr 76 til SEQ ID NO:7 som inneholder 5 nukleotidforskjeller i forhold til tilsvarende region av BMP-12 kodende sekvens angitt i SEQ ID NO: 1 (nukleotidene nr 672 til 689). En av rekombinante bakteriofagkloner som hybridiserer til VL-1 oligonukleotidprobe nr 8 er betegnet λJLDc31. Denne rekombinante bakteriofagklonen blir plakkrenset, en bakteriofagstokk blir dannet og bakteriofag DNA blir isolert fra λJLDc31 human genomisk klon. Bakteriofagen λJLDc31 er blitt deponert til American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD "ATCC" under aksesjonsnr. 75922 oktober 20, 1994. Denne deponeringen oppfyller kravene ifølge Budapest Treaty of the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and Regulations. Oligonukleotidhybridiseringsregionen til denne rekombinanten λJLDc31 er lokalisert til et 2,5 kb EcoRI fragment. Dette fragmentet blir subklonet inn i en plasmidvektor (pGEM-3) og DNA-sekvensanalyser blir utført. Denne plasmid subklonen er betegnet pGEMJLDc31/2,5 og er blitt deponert til American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD "ATCC" under aksesjon nr. 69710 oktober 20, 1994. Denne deponeringen oppfyller kravene ifølge Budapest Treaty of the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and Regulations.

Del DNA-sekvensen (SEQ ID NO:25) og avledet aminosyresekvens (SEQ ID NO: 26) av en del av 2,5 kb DNA innskuddet til plasmidsubklon pGEMJLDc31/2,5 avledet fra klon λJLDc31, er vist i sekvenslisten.

DNA-sekvensen til en del av 2,5 kb EcoRI innskuddet til plasmid pGEMJLDC31/2,5 er angitt i SEQ ID NO: 25 og inneholder en 912 bp åpen leseramme, som definert ved nukleotidene nr. 52 til og med nr. 963 ifølge SEQ ID NO: 25. I det denne sekvensen er avledet fra en genomisk klon er det vanskelig å bestemme grensen mellom 5' omfanget av den kodende sekvensen og 3' grensen til mellomliggende sekvens (intron/ikke-kodende sekvens). Hele den åpne leserammen (nukleotidene nr 52 til og med nr 963 ifølge SEQ ID NO: 25) koder for en del av VL-1 proteinet ifølge oppfinnelsen opp til 304 aminosyrer.

Basert på kunnskapen om andre BMP proteiner og andre proteiner innenfor TGF- β familien antas det at forløper polypeptidet kan bli spaltet ved flerbasesekvensen Arg-Arg-Arg i samsvar med en foreslått konsensus proteolytisk prosesseringssekvens til Arg-X-X-Arg. Spaltning av VL-1 forløperpolypeptidet er ventet å danne et 120 aminosyre modent peptid begynnende med aminosyren Thr i posisjon nr 1 ifølge SEQ ID NO: 26. Prosessering av VL-1 til den modne formen antas å involvere dimerisering og fjerning av den N-terminale regionen på en måte som er analog med prosessering av beslektet protein TGF- β [Gentry et al., *Molec & Cell. Biol.*, 8:4162 (1988); Derynck et al. *Nature*, 316:701 (1985)].

Det antas derfor at modne aktive former av VL-1 omfatter en homodimer med 2 polypeptidsubenheter og hver subenhet omfatter aminosyrene nr 1 til nr 120 ifølge SEQ ID NO: 26 med en antatt molekylvekt på omtrent 12.000 dalton. Ytterligere aktive arter antas å omfatte minst aminosyrene nr 19 til nr 119 eller nr 120 ifølge SEQ ID NO: 26, og inkluderer følgende første og siste konserverte cysteinresidie.

Ved anvendelse av en slik fremgangsmåte ble en klon kodende for moden human VL-1 (BMP-13) oppnådd. Nukleotidsekvensen og tilsvarende aminosyresekvens som blir kodet av denne klonen er ført i sekvenslisten ved SEQ ID NO: 25 og 26.

EKSEMPEL 2

Ekspresjon av BMP-12

For å produsere humane BMP-12 proteiner blir DNA kodende derav overført til en hensiktsmessig ekspresjonsvektor og ført inn i pattedyrceller eller andre foretrukne eukaryote eller prokaryote verter ved hjelp av konvensjonelle genetiske konstruksjonsteknikker.

For å produsere humant BMP-12 protein i bakterielle celler blir følgende prosedyre anvendt.

Ekspresjon av BMP-12 i E.coli

Et ekspresjonsplasmid pALV1-781, for produksjon av BMP-12 i E. coli ble konstruert og inneholder følgende hovedtrekk: nukleotidene 1-2060 inneholder DNA-sekvenser med opphav fra plasmid pUC-18 [Norrande et al., Gene 26:101-106 (1983)] inkludert sekvenser inneholdende genet for β -laktamase som utøver resistens overfor antibiotika ampicillin i E. coli vertsstammer, og et colE1-avledet replikasjonsorigo. Nukleotidene 2061-2221 inneholder DNA-sekvenser for hoved "leftward" promoter (pL) til bakteriofag λ [Sanger et al., J. Mol. Biol. 162:729-773 (1982)], inkludert 3 operatorsekvenser O_L1, O_L2 og O_L3. Operatorene er bindingssteder for λ C1 repressorproteinet og intracellulære nivåer kontrollerer mengden av transkripsjonsinitiering fra pL. Nukleotidene 2222-2723 inneholder en sterk ribosom bindende sekvens inkludert på en sekvens avledet fra nukleotidene 35566 til 35472 og 38137 til 38361 fra bakteriofag lambda beskrevet i Sanger et al., J. Mol. Biol. 162:729-773 (1982). Nukleotidene 2724-3041 inneholder en DNA-sekvens kodende for modent BMP-12 protein hvor alle 3' utranslaterte sekvenser er fjernet. BMP-12 DNA-sekvenser introdusert i pALV1-781 ekspresjonsvektoren ble modifisert i 5' enden for å øke A+T innholdet uten å endre den kodende kapasiteten. Disse forandringene ble utført for å øke translasjonseffektiviteten initiert på BMP-12 mRNA i E.coli. Nukleotidene 3042-3058 tilveiebringer en "linker" DNA-sekvens inneholdende restriksjonsendonuklease seter. Nukleotidene 3059-3127 tilveiebringer en transkripsjonstermineringssekvens basert på den til E. coli Asp A genet [Takagi et al., Nucl. Acids Res. 13:2063-2074 (1985)]. Nukleotidene 3128-3532 er DNA-sekvenser avledet fra pUC-18.

Plasmid pALV1-781 ble transformert inn i E. coli vertsstamme G1724 (F, lacI^q, lacp^{L8}, ampC:: λ C1⁺) ifølge prosedyren til Dagert og Ehrlich, Gene 6:23 (1979). G1724 (ATCC aksesjonsnr. 55151) inneholder en kopi av vill-type λ C1 repressorgenet stabilt integrert inn i kromosomet ved ampC lokuset, hvor det er blitt plassert under transkripsjonell kontroll av Salmonella typhimurium trp promoter/operatorsekvenser. I G1724 blir λ C1 proteinet bare dannet i løpet av vekst i tryptofan-fritt medium, så som minimalmedium eller et minimalmedium supplementert med casaminosyrer så som IMC, beskrevet ovenfor. Tilførsel av tryptofan til en kultur av G1724 vil undertrykke trp promoteren og stenge syntesen av λ C1, og gradvis forårsake induksjon av transkripsjon fra pL promotorene dersom de er tilstede i cellen.

Transformantene ble selektert på 1,5% w/v agar skåler inneholdende IMC medium som består av M9 medium [Miller, "Experiments in Molecular Genetics", Cold Spring Harbor Laboratory, New York (1972)] inneholdende 1 mM MgSO_4 og supplementert med 0,5% w/v glukose, 0,2% w/v casaminosyrer og 100 $\mu\text{g/ml}$ ampicillin. GI274 transformert med pALV1-781 ble dyrket ved 37°C til en A_{550} 0,5 i IMC medium inneholdende 100 $\mu\text{g/ml}$ ampicillin. Tryptofan ble deretter tilsatt til en sluttkonsentrasjon på 100 $\mu\text{g/ml}$ og kulturen inkubert i ytterligere 4 timer. I løpet av denne tiden akkumulerer BMP-12 protein innenfor "inklusjonslegeme" fraksjonen.

Preparering av proteinmonomer

18 g frosne celler ble veid og resuspendert i 60 ml 100 mM tris, 10 mM EDTA, 1 mM fenylmetylsulfonylfluorid [PMSF], pH 8,3. Cellene ble lysert ved å bli sendt tre ganger gjennom en MicrofluidizerTM [modell nr MCF 100 T]. Inklusjonslegeme pelleten ble oppnådd ved sentrifugering ved 15.000 g ved 4°C i 20 minutter. Supernatanten ble dekantert og pelleten vasket med 100 ml 100 mM tris, 1,0 M NaCl, 10 mM EDTA, 1 mM PMSF, pH 8,3. Suspensjonen ble på ny sentrifugert ved 15.000 g ved 4°C i 10 minutter og supernatanten dekantert. Pelleten ble deretter vasket med 100 ml 100 mM tris, 10 mM EDTA, 1% triton X-100, 1 mM PMSF, pH 8,3. Suspensjonen ble på ny sentrifugert ved 15.000 g ved 4°C i 10 minutter og supernatanten dekantert. Pelleten ble resuspendert med 50 ml 20 mM tris, 1 mM EDTA, 1 mM PMSF, pH 8,3, inneholdende 1% DTT i en glassvevshomogenisator. Monomerisk BMP-12 ble deretter oppløst ved surgjøring til pH 2,5 med iseddiksyre. Oppløselig fraksjon ble isolert ved sentrifugering ved 15.000 g i 20 minutter ved 4°C.

Supernatanten fra denne sentrifugeringen ble samlet og kromatografert over en Sephacryl S-100TM størrelsesseksklusjonskolonne (83 cm x 2,6 cm; $\approx$ 440 ml sjikt) i 20 ml deler. Sephacryl S-100TM kolonnen ble kjørt med en mobil fase bestående av 1% eddiksyre ved en strømningsrate på 1,4 ml/min. Fraksjoner som tilsvarer BMP-12 monomer ble detektert ved absorbanse ved 280 nm ved anvendelse av en databeregnet ekstinksjonskoeffisient på $18200\text{M}^{-1}\text{cm}^{-1}$ og molekylvekt (11667 dalton). Dette størrelsesseksklusjonskolonne materialet ble anvendt som utgangsmateriale for refoldingsreaksjoner.

Som et alternativ til ovennevnte blir 1,0 g celler lagret ved -80°C og målt. Oppløsningen (3,4 ml 100 mM tris, 10 mM EDTA, pH 8,5) blir tilsatt. Oppløsningen blir vortexbehandlet helt til cellene er suspenderte. 40 μl 100 mM PMSF i isopropanol blir tilsatt. Cellene blir lysert ved 1000 psi i en French presscelle. Inklusjonslegemene blir sentrifu-

gert ved 4°C i 20 minutter i Eppendorf mikrofuge for å danne pelleter. Supernatantene blir dekantert. Til en pelleten (av 4 totalt) blir 1,0 ml avgasset 8.0 M guanidinhydroklorid, 0,5 M tris, 5 mM EDTA, pH 8,5, inneholdende 250 mM DTT tilsatt. Pelleten blir løst opp og argon blir blåst over væsken i 30 sekunder. Deretter blir oppløsningen inkubert ved 37°C i 1 time. Uoppløselig materiale blir pelletert i 2-3 minutter i en Eppendorf mikrofuge ved 23°C. 0,5-1,0 ml supernatant blir injisert inn i en Supelco 2 cm guard cartridge (LC-304) og eluert med en acetonitrilgradient i 0,1% TFA fra 1-70% over 35 minutter. BMP-12 eluerer mellom 29 og 31 minutter. Fraksjonene blir slått sammen og proteinkonsentrasjonen bestemt ved adsorbanse ved 280 nanometer mot 0,1% TFA ved anvendelse av teoretisk ekstinksjonskoeffisient basert på aminosyreinnholdet.

Som en annen alternativ metode i forhold til den ovennevnte blir frosne cellepelleter oppnådd fra E. coli transformanter som beskrevet ovenfor og tint i 30 ml TE8.3(100:10) buffer (100 mM tris-HCl pH 8,3, 10 mM Na₂EDTA, 1 mM PMSF). Cellene blir lysert ved å sende dem tre ganger gjennom en MicrofluidizerTM [modell nr MCF 100 T]. Opprinnelig inklusjonslegeme materialpelleten blir løst opp i 8 M guanidin-HCl, TE8.5 (100:10) buffer (100 mM tris-HCl pH 8,5, 10 mM Na₂EDTA som inneholdt 100 mM DTT) og inkubert ved 37°C i 1 time. Dette materialet blir sentrifugert ved 12000 x g i 15 minutter ved romtemperatur.

Refolding av BMP-12 protein ved anvendelse av CHAPS-system

Et tilstrekkelig volum av BMP-12-blandingen blir lyofilisert for å oppnå 10 µg protein. 5 µl av glassdestillert vann blir tilsatt for å gjenoppløse resten og deretter blir 100 µl av refoldet blanding (50 mM tris, 1,0 M NaCl, 2% 3-(2-chloramidopropyl)dimetylammonio-1-propansulfat (CHAPS), 5 mM EDTA, 2 mM glutation (redusert) 1 mM glutation (oksidert); ved pH på omtrent 8,5). Oppløsningen blir forsiktig blandet og lagret ved 23°C i 1-4 dager. Dimerdannelsen blir vurdert ved å kjøre en alikvot på en Novex 16% tricin gel ved 125 volt i 2,5 timer, etterfulgt av Coomassie blå farving og avfarving.

BMP-12 dimeren ble rensset ved anvendelse av en C4 analytisk RP-HPLC (revers fase-høy ytelse væskechromatografi) kolonne (Vydac 214TP54) som ble ekvilibrert til 1% B-buffer (fortynnet inn i A buffer) og ble kjørt i løpet av 35 minutter, hvor proteinet eluerer ved anvendelse av følgende gradient (A buffer = 0,1% trifluor eddiksyre, B buffer = 95% acetonitril, 0,1% trifluor eddiksyre [TFA], med en strømningsrate på 1 ml/min:

1-5 minutter 20% B buffer

5-10 minutter 20-30% B buffer

10-30 minutter 30-50% B buffer

30-35 minutter 50-100% B buffer.

Proteinet ble registrert ved absorbanse ved 280 nm. Topp BMP-12 fraksjoner (eluerende mellom 29 og 31 minutter) ble slått sammen. Renheten ble vurdert ved SDS-PAGE. Konsentrasjonen ble bestemt ved absorbanse ved 280 nm og anvendelse av computerberegnet ekstinksjonskoeffisient og molekylvekt som angitt ovenfor.

Ekspresjon av BMP-12 i pattedyrceller:

Et annet betraktet foretrukket ekspresjonssystem for biologisk aktiv rekombinant human BMP-12 er stabilt transformerte pattedyrceller.

Fagfolk innenfor dette området kan konstruere pattedyr ekspresjonsvektorer ved anvendelse av sekvensen til SEQ ID NO: 1, eller andre DNA-sekvenser kodende for BMP-12 proteiner eller andre modifiserte sekvenser og kjente vektorer, så som pCD [Okayama et al., Mol. Cell. Biol., 1:161-170 (1982)], pJL3, pJL4 [Gough et al., EMBO J., 4:645-653 (1985)] og pMT2 CXM.

Pattedyr ekspresjonsvektor pMT2 CXM er et derivat av p91023(b) (Wong et al., Science 228:810-815, 1985) forskjellig fra sistnevnte ved at det inneholder ampicillin resistensgenet i stedet for tetracyklinresistensgenet og videre inneholder et XhoI sete for innskudd av cDNA kloner. De funksjonelle elementene til pMT2 CXM er blitt beskrevet (Kaufman, R.J., 1985, Proc. Natl. Acad. Sci. USA 82:689-693) og innbefatter adenovirus VA gener, SV40 replikasjonsorigo inkludert 72 bp enhancer, adenovirus major late promoter inkludert et 5' spleisesete og det meste av adenovirus tripartit ledersekvensen tilstede på adenovirus late mRNA, et 3' spleise akseptor sete, et DHFR innskudd, SV40 early polyadenyleringssete (SV40) og pBR322 sekvenser nødvendig for propagering i E. coli.

Plasmid pMT2 CXM blir oppnådd ved EcoRI spaltning av pMT2-VWF som er blitt deponert til American Type Culture Collection (ATCC), Rockville, MD (USA) under aksjonsnr. ATCC 67122. EcoRI spaltningen spalter ut cDNA innskuddet til stede i pMT2-VWF, og tilveiebringe pMT2 i lineær form som kan bli ligert og anvendt for å transformere E. coli HB 101 eller DH-5 til ampicillin resistens. Plasmid pMT2 DNA kan bli dannet ved hjelp av konvensjonelle metoder. pMT2 CXM blir deretter konstruert ved anvendelse av "loopout/i mutagenese" [Moringa, et al., Biotechnology 84: 636 (1984). Dette fjerner basen 1075 til 1145 i forhold til HindIII setet nær ved SV40 replikasjonsorigo og enhancer sekvenser til pMT2. I tillegg blir det innskutt en sekvens inne-

holdende gjenkjenningssete for restriksjonsendonuklease XhoI. Et derivat pMT2CXM betegnet pMT23, inneholder gjenkjenningsseter for restriksjonsendonukleasene PstI, EcoRI, Sall og XhoI. Plasmid pMT2 CXM og pMT23 DNA kan bli dannet ved hjelp av vanlige metoder.

pEMC2B1 avledet fra pMT21 kan også være egnet for utførelse av oppfinnelse. pMT21 blir avledet fra pMT2 som er avledet fra pMT2-VWF. Som beskrevet ovenfor spalter EcoRI spaltningen cDNA innskudd tilstede i pMT-VWF, og tilveiebringer pMT2 i lineær form som kan bli ligert og anvendt for å transformere *E. coli* HB101 og DH-5 til ampicillinresistens. Plasmid pMT2 DNA kan bli dannet ved hjelp av konvensjonelle metoder.

pMT21 blir avledet fra pMT2 ved hjelp av følgende to modifikasjoner. For det første blir 76 bp av 5' utranslatert region til DHFR cDNA inkludert et område med 19 G residier fra G/C halen for cDNA kloning deletert. I denne prosessen blir et XhoI sete skutt inn for å oppnå følgende sekvens rett oppstrøms fra DHFR. Deretter blir et unikt ClaI sete introdusert ved spaltning med EcoRV og XbaI, behandling med Klenow fragment DNA polymerase I og ligering til en ClaI linker (CATCGATG). Dette deleterer et 250 bp segment fra adenovirus assosiert RNA (VAI) region, men interfererer ikke med VAI RNA genekspressjonen eller funksjonen. pMT21 blir spaltet med EcoRI og XhoI og anvendt for å oppnå vektor pEMC2B1.

En del av EMCV lederen blir oppnådd fra pMT2-ECAT1 [S.K. Jung, et al, J.Virol 63: 1651-1660 (1989)] ved spaltning med EcoRI og PstI og resulterer i et 2752 bp fragment. Dette fragmentet blir spaltet med TaqI og tilveiebringer et EcoRI-TaqI fragment på 508 bp som blir rensset ved elektroforese på lavt smeltende agarosegel. En 68 bp adapter og dets komplementære tråd blir syntetisert med en 5' TaqI utragende ende og en 3' XhoI utragende ende som har en sekvens som samsvarer med EMC virus ledersekvensen fra nukleotid 763 til 827. Den forandrer også ATG ved posisjon 10 innenfor EMC viruslederen til en ATT og blir etterfulgt av et XhoI sete. En treveis ligering av pMT21 Eco RI-XhoI fragmentet, EMC virus EcoRI-TaqI fragmentet og 68 bp oligonukleotidadapter TaqI-XhoI adapter resulterer i vektoren pEMC2B1.

Denne vektoren inneholder SV40 replikasjonsorigo og enhancer, adenovirus major late promoter, en cDNA-kopi av det meste av adenovirus tripartit ledersekvensen, en liten hybrid intervenserende sekvens, et SV40 polyadenyleringssignal og adenovirus VA I

genet, DHFR og β -laktamase markører og EMC sekvens, i hensiktsmessige forhold for å lede ekspresjon i høyt nivå av ønsket cDNA i pattedyrcellene.

Konstruksjon av vektorene kan innbefatte modifikasjon av BMP-12 DNA-sekvenser. Blant annet kan BMP-12 cDNA modifiseres ved fjerning av ikke-kodende nukleotider på 5' og 3' endene til den kodende regionen. Deleterte ikke-kodende nukleotider kan erstattes av andre sekvenser kjent for å være fordelaktige for ekspresjon. Disse vektorene blir transdormert inn i hensiktsmessige vertsceller for ekspresjon av BMP-12 proteiner. I tillegg kan sekvensen til SEQ ID NO: 1 eller andre sekvenser kodende for BMP-12 proteinene bli manipulert for å uttrykke BMP-12 proteinet ved isolering av den modne kodende sekvensen til nukleotidene 571 til 882 ifølge SEQ ID NO: 1 og tilsetning ved 5'-enden sekvenser som koder for fullstendige propeptider til andre BMP proteiner.

For eksempel kan fagfolk innenfor dette området danne et fusjonsprotein hvor propeptidet til BMP-2 er koblet på en operabel måte til modent BMP-12 peptid ved å danne en DNA-vektor hvor DNA-sekvensen kodende for BMP-2 propeptidet er koblet i riktig leseramme til DNA-sekvensen kodende for modent BMP-12 peptid. DNA-sekvensen til et slikt fusjonsprotein er vist i SEQ ID NO: 27.

Fagfolk innenfor dette området kan manipulere sekvensene ifølge SEQ ID NO: 1 ved å eliminere eller erstatte pattedyr regulatoriske sekvenser som flankerer den kodende sekvensen med bakterielle sekvenser for å danne bakterielle vektorer for intracellulær eller ekstracellulær ekspresjon av bakterielle celler, som beskrevet ovenfor. Som et annet eksempel kan kodende sekvenser bli ytterligere manipulert (for eksempel ligert til andre kjente linkere eller modifisert ved deletering av ikke-kodende sekvenser derfra eller endring av nukleotider deri ved hjelp av andre kjente teknikker). Modifisert BMP-12 kodende sekvens kan deretter bli innskutt i en kjent bakteriell vektor ved anvendelse av prosedyrene som beskrevet i T. Taniguchi et al., Proc. Natl. Acad. Sci. USA, 77:5230-5233 (1980). Denne eksempelvis bakterielle vektoren kan deretter bli transformert inn i bakterielle vertsceller og et BMP-12 protein dermed uttrykt. For en strategi for produksjon av ekstracellulær ekspresjon av BMP-12 proteiner i bakterielle celler se for eksempel europeisk patentsøknad 177.343.

Lignende manipuleringer kan bli utført for konstruksjon av en innskuddsvektor (se for eksempel prosedyrer beskrevet i publisert europeisk patentsøknad 155.476) for ekspresjon i insektsceller. En gjærvektor kan også bli konstruert ved anvendelse av regulato-

riske sekvenser fra gjær for intracellulær eller ekstracellulær ekspresjon av faktorene ifølge foreliggende oppfinnelse fra gjærceller (se for eksempel prosedyrer beskrevet i publisert PCT søknad WO86/00639 og europeiske patentsøknad EPA 123.289).

En fremgangsmåte for produsering av høye nivåer av et BMP-12 protein ifølge oppfinnelsen i pattedyrceller kan bli anvendt for konstruksjon av celler inneholdende mange kopier av heterologt BMP-12 gen. Heterologt gen blir koblet til en amplifiserbar markør, for eksempel dihydrofolat reduktase (DHFR) gen hvor celler inneholdende økte genkopier kan bli selektert for propagering av økende konsentrasjoner av metotreksat (MTX) ifølge prosedyren til Kaufman og Sharp, J.Mol. Biol., 159:601-629 (1982). Denne metoden kan anvendes med mange forskjellige celletyper.

Et plasmid inneholdende en DNA-sekvens for en BMP-12 ifølge oppfinnelsen i operativ assosiasjon med andre plasmidsekvenser som etablerer ekspresjonen derav og DHFR ekspresjonsplasmid pAdA26SV(A)3 [Kaufman og Sharp, Mol. Cell. Biol., 2:1304 (1982)] kan sammen innføres i DHFR-manglende CHO celler, DUKX-BII, ved hjelp av forskjellige fremgangsmåter som innbefatter kalsiumfosfat kopresipitasjon og transfeksjon, elektroporasjon eller protoplastfusjon. DHFR ekspresjonstransformanter blir valgt for vekst i alfa medium med dialysert føtalt kalveserum, og deretter selektert for amplifikasjon ved vekst i økende konsentrasjoner MTX (for eksempel sekvensielle trinn i 0,02, 0,2, 1,0 og 5 µm MTX) som beskrevet i Kaufman et al., Mol. Cell.Biol., 5:1750 (1983). Transformanter blir klonet og biologisk aktiv BMP-12 ekspresjon blir registrert i Rosen-modifisert Sampath-Reddi rotteanalyse beskrevet nedenfor i eksempel 5. BMP-12 ekspresjonen øker med økende nivåer av MTX resistens. BMP-12 polypeptider er karakterisert ved hjelp av standardteknikker kjent innenfor fagområdet så som pulsmarking med [35S] metionin eller cystein og polyakrylamid gelelektroforese. Lignende prosedyrer kan bli fulgt for å produsere andre beslektede BMP-12 proteiner.

Eksempel 3

Preparering av BMP-2 propeptid/BMP-12 moden peptidfusjon

For å konstruere en vektor som koder for BMP-2 propeptid/BMP-12 moden peptidfusjon ble følgende kloningsprosedyre anvendt for å fusjonere de to sekvensene sammen.

Først blir et DNA restriksjonsenzymfragment omfattende propeptidet til humant BMP-2 protein, omfattende nukleotidene 1 til og med 843 ifølge SEQ ID NO: 27 spaltet fra pBMP2ΔEMC. pBMP2ΔEMC er et plasmid avledet fra lambda U20S-39 (ATCC nr.

40345) som omfatter hele den kodende sekvensen for humant BMP-2 protein med ikke-translaterte 5' og 3'sekvenser til BMP-2 deletert fra vektoren. 5' restriksjonsenzymet anvendt var BglII og det spalter pBMP2ΔEMC i vektoren ved nukleotid 979. 3' restriksjonsenzymet var Mae II og det spalter pBMP2ΔEMC i BMP-2 propeptidet ved nukleotid 1925 ved karboksyterminusen. Det resulterende 954 baseparproduktet ble deretter gelisolert og genrenset. Deretter blir et DNA restriksjonsenzymfragment omfattende 5' delen av human BMP-12 moden peptid DNA-sekvens spaltet fra pPCR1-1 nr 2 V1-1 (ATCC nr 69517). 5' restriksjonsenzymet anvendt var Eae I og spalter pPCR1-1 nr 2 V1-1 like 3' for den N-terminale enden til human BMP-12 moden peptidsekvens. Resulterende 259 baseparprodukt ble gelisolert og genrenset. Deretter ble to DNA oligoer konstruert og syntetisert slik at når sammensmeltet blir det dannet et lite DNA-fragment som omfatter fusjonssekvensen til ekstrem 3' ende av humant BMP-2 propeptid og 5' ende til BMP-12 moden peptid. DNA-fragmentet har en 5' Mae II komplementær klebrig ende som blir sammensmeltet til 3' restriksjonsenzymfragmentet som omfatter humant BMP-2 propeptid. Sammensmeltet oligo DNA-fragment har en 3' Eae I komplementær klebrig ende som blir sammensmeltet til 5' av restriksjonsenzymfragmentet som omfatter det modne peptidet til human BMP-12. Kodende oligotråd blir betegnet B2/12 og har en lengde på 13 basepar. Deretter ble et DNA-fragment kodende for 123 baseparene ved 3' enden av BMP-12 moden peptidfragment oppnådd som følger. Først blir et DNA-fragment som omfatter propeptidet til humant BMP-2 protein, omfattende nukleotidene nr 1 til og med 846 PCR amplifisert fra pBMP2ΔEMC. 5' primerene (oligo 655a) sammensmelter like 5' for polylinkeren. 3' primeren (BMPpro3) blir sammensmeltet til BMP-2 propeptid 3' enden og innfører et Bgl II restriksjonsenzymsete ved stille sekvensmutasjoner. Resulterende PCR produkt ble spaltet med Sal I som blir spaltet i polylinkeren, og Bgl II. 850 basepar restriksjonsenzym fragment (som ender i aminosyresekvens REKR) ble gelisolert og genrenset. BMP-12 moden peptid ble PCR amplifisert ved anvendelse av 5' primer (oligo 5-1) kodende for Bgl II restriksjonsenzymsetet ved stille sekvensmutasjoner, og sammensmeltning til 5' enden av et mulig moden spaltningsprodukt, begynnende med aminosyresekvens SRCS. 3' primeren (V1-1 3) blir sammensmeltet til BMP-12 moden peptid 3' ende og innfører et XbaI restriksjonsenzymsete etter stoppkodonet. Det resulterende PCR produktet ble spaltet med Bgl II og XbaI. 321 basepar restriksjonsenzym fragmentet ble gelisolert og genet ble renset.

To restriksjonsfragmenter ble 3-veis ligert inn i en på forhånd SalI og XbaI spaltet vektor. Den resulterende konstruksjonen ble sekvensert for å undersøke PCR induserte feil og en stille C til T mutasjon ble observert ved basepar 185 i propeptidet. Dette plasmidet ble betegnet pREKRSRC. Deretter ble pREKRSRC spaltet med BglII og NgoMI, og

vektorfragmentet som omfatter siste 123 basepar av BMP-12 moden sekvens ble derved isolert. De tre restriksjonsfragmentene og sammensmeltet oligolinker ble 4-veis ligert for å tilveiebringe pREKR-TAL med BMP-2 propeptid med modent spaltningssete ved 3' enden koblet til (TAL) 5' enden av BMP-12 modent peptid. Den kodende sekvensen til den resulterende ligerte vektoren er vist i SEQ ID NO: 27.

Eksempel 4

Biologisk aktivitet til uttrykt BMP-12

For å måle den biologiske aktiviteten til uttrykte BMP-12 proteiner oppnådd i eksempel 2 ovenfor ble proteinene isolert fra cellekulturen og renset ved isolering av BMP-12 proteinene fra andre proteinholdige materialer som de blir produsert sammen med samt fra andre kontaminanter. Renset protein kan bli analysert i henhold til rotteanalysen beskrevet nedenfor i eksempel 5.

Rensingen blir utført ved anvendelse av standardteknikker kjent for fagfolk innenfor dette området.

Proteinanalyser blir utført ved anvendelse av standardteknikker så som SDS-PAGE akrylamid [Laemmli, Nature 227:680 (1970)] farvet med Coomassie blå eller sølv [Oakley, et al. Anal. Biochem. 105:361 (1980)] og ved immunoblotting [Towbin, et al. Proc. Natl. Acad. Sci. USA 76:4350 (1979)].

Eksempel 5

Rosen-modifisert Sympath-Reddi analyse

En modifisert versjon av rotte ektopisk implanteringsanalyse beskrevet i Sampath og Reddi, Proc. Natl. Acad. Sci. USA, 80:6591-6595 (1983) blir anvendt for å vurdere aktiviteten til BMP-12 proteinene. Denne modifiserte analysen er heri betegnet Rosen-modifisert Sympath-Reddi analyse. Analysen er blitt anvendt for å vurdere ben og brusk-induserende aktivitet til BMP. Etanolpresipiteringstrinnet til Sympath-Reddi prosedyren er erstattet med dialysering (dersom sammensetningen er en oppløsning) eller diafiltrering (dersom sammensetningen er en suspensjon) av fraksjonen som skal bli analysert mot vann. Oppløsningen eller suspensjonen blir deretter ekvilibret til 0,1% TFA. Den resulterende oppløsningen blir tilsatt til 20 mg rottematrise. En uekte rottematriseprøve ikke behandlet med protein virker som en kontroll. Dette materialet blir frosset og lyofilisert og resulterende pulver innesluttet i nr 5 gelatinkapsler. Kapslene blir implantert subkutan i abdominal thoraks område til 21-49 dager gamle hann Long Evans rotter. Implanatene blir fjernet etter 10 dager. Et snitt av hvert implantat blir fiksert og bearbeidet for

histologisk analyse. 1 μm glykolmetakrylat deler blir farvet med Von Kossa og sur fuschin for å registrere mengden av indusert sene/ligamentlignende vevsdannelse tilstede i hvert implantat.

BMP-12 ble implantert i rotter i doser på 1, 5, 25 og 50 μg pr implantat i 10 dager. BMP-2 i en dose på 5 μg ble innbefattet som en positiv kontroll. For alle testede doser av BMP-12 ble ingen ben eller bruskdannelse observert i implantatene etter 10 dager. Implantatene ble fylt med vevslignende embryonisk sene som lett kan gjenkjennes ved tilstedeværelse av tette knuter av fibroblaster orientert i samme plan og pakket tett sammen (sene/ligamentlignende vev er for eksempel beskrevet i Ham og Cormack, Histology (JB Lippincott Co. (1979), s. 367-369, og beskrivelsen er inkorporert heri som referanse). Disse funnene ble reproduert i et annet sett av analyser hvor sene/ligamentlignende vev var tilstede i alle BMP-12 inneholdende implantater. I kontrast til dette viste BMP-2 implantater, som ventet, brusk og bendannelse, men ikke sene/ligamentlignende vev.

BMP-12 proteinene og relaterte proteiner ifølge oppfinnelsen kan bli vurdert for aktivitet i denne analysen.

Eksempel 6

Anvendelse av fremgangsmåter i henhold til ovennevnte eksempler, med mindre modifikasjoner innenfor fagområdet, ble humant MP-52 protein og murin homolog av en BMP-13 protein uttrykt og analysert for sene/ligamentlignende vevsinduserende aktivitet. Alle proteinene viste sammenlignbare resultater som ligner de som er beskrevet ovenfor for human BMP-12.

SEKVENSLISTE

(1) GENERELL INFORMASJON:

- 5 (i) APPLICANT: GENETICS INSTITUTE, INC.
PRESIDENT AND FELLOWS OF HARVARD COLLEGE
- (ii) TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
- (iii) NUMBER OF SEQUENCES: 35
- (iv) CORRESPONDENCE ADDRESS:
10 (A) ADDRESSEE: GENETICS INSTITUTE, INC.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: Massachusetts
(E) COUNTRY: USA
(F) ZIP: 02140
- (v) COMPUTER READABLE FORM:
15 (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: Herewith
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
20 (A) APPLICATION NUMBER: US 08/164,103
(B) FILING DATE: 07-DEC-1993
(C) APPLICATION NUMBER: US 08/217,780
(D) FILING DATE: 25-MAR-1994
(E) APPLICATION NUMBER: US 08/333,576
(F) FILING DATE: 02-NOV-1994
- (viii) ATTORNEY/AGENT INFORMATION:
25 (A) NAME: Lazar, Steven R.
(B) REGISTRATION NUMBER: 32,618
(C) REFERENCE/DOCKET NUMBER: 5202D-PCT
- (ix) TELECOMMUNICATION INFORMATION:
30 (A) TELEPHONE: 617 498-8260
(B) TELEFAX: 617 876-5851

(2) INFORMASJON FOR SEQ ID NO: 1:

(i) SEKVENSKARAKTERTREKK:

- (A) LENGDE: 926 basepar
- (B) TYPE: nukleinsyre
- 35 (C) TRÅDTYPE: enkel
- (D) TOPOLOGI: lineær

(ii) MOLEKYLTYP: DNA (genomisk)

(iv) OPPRINNELIG KILDE:

(A) ORGANISME: Homo sapiens

(vii) ØYEBLIKKELEG KILDE:

5 (B) KLON: v1-1

(ix) TREKK:

(A) NAVN/NØKKEL: mat_peptid

(B) BELIGGENHET: 571..882

(ix) TREKK:

10 (A) NAVN/NØKKEL: CDS

(B) BELIGGENHET: 1..882

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 1:

15	GCG CGT AAT ACG ACT CAC TAT AGG GCG AAT TGG GTA CCG GGC CCA GGC Ala Arg Asn Thr Thr His Tyr Arg Ala Asn Trp Val Arg Gly Pro Gly -190 -185 -180 -175	48
	AGC TGG ACT TCT CCG CCG TTG CTG CTG CTG TCC ACG TGC CCG GGC GCC Ser Trp Thr Ser Pro Pro Leu Leu Leu Leu Ser Thr Cys Pro Gly Ala -170 -165 -160	96
20	GCC CGA GCG CCA CGC CTG CTG TAC TCG CCG GCA GCT GAG CCC CTA GTC Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val -155 -150 -145	144
	GGT CAG CGC TGG GAG GCG TTC GAC GTG GCG GAC GCC ATG AGG CGC CAC Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His -140 -135 -130	192
25	CGT CGT GAA CCG CGC CCC CCC CGC GCG TTC TGC CTC TTG CTG CGC GCA Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu Leu Arg Ala -125 -120 -115	240
	GTG GCA GGC CCG GTG CCG AGC CCG TTG GCA CTG CCG CGA CTG GGC TTC Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg Leu Gly Phe -110 -105 -100 -95	288
	GGC TGG CCG GGC GGA GGG GGC TCT GCG GCA GAG GAG CGC GCG GTG CTA Gly Trp Pro Gly Gly Gly Gly Ser Ala Ala Glu Glu Arg Ala Val Leu -90 -85 -80	336
30	GTC GTC TCC TCC CGC ACG CAG AGG AAA GAG AGC TTA TTC CCG GAG ATC Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile -75 -70 -65	384
	CGC GCC CAG GCC CGC GCG CTC GGG GCC GCT CTG GCC TCA GAG CCG CTG Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser Glu Pro Leu -60 -55 -50	432

	CCC GAC CCA GGA ACC GGC ACC GCG TCG CCA AGG GCA GTC ATT GGC GGC	480
	Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val Ile Gly Gly	
	-45 -40 -35	
	CGC AGA CGG AGG AGG ACG GCG TTG GCC GGG ACG CGG ACA GCG CAG GGC	528
	Arg Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr Ala Gln Gly	
	-30 -25 -20 -15	
5	AGC GGC GGG GGC GCG GGC CGG GGC CAC GGG CGC AGG GGC CGG AGC CGC	576
	Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg	
	-10 -5 1	
	TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG GAG CTC GGC TGG GAC	624
	Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp	
	5 10 15	
10	GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG TAC CAC TGC GAG GGC	672
	Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly	
	20 25 30	
	CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG CCC ACC AAC CAT GCC	720
	Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala	
	35 40 45 50	
15	ATC ATT CAG ACG CTG CTC AAC TCC ATG GCA CCA GAC GCG GCG CCG GCC	768
	Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala	
	55 60 65	
	TCC TGC TGT GTG CCA GCG CGC CTC AGC CCC ATC AGC ATC CTC TAC ATC	816
	Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile	
	70 75 80	
	GAC GCC GCC AAC AAC GTT GTC TAC AAG CAA TAC GAG GAC ATG GTG GTG	864
	Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val	
20	85 90 95	
	GAG GCC TGC GGC TGC AGG TAGCGCGCGG GCCGGGGAGG GGGCAGCCAC	912
	Glu Ala Cys Gly Cys Arg	
	100	
	GCGGCCGAGG ATCC	926

25

30

(2) INFORMASJON FOR SEQ ID NO: 2:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 294 aminosyrer

(B) TYPE: aminosyre

(C) TOPOLOGI: lineær

35

(ii) MOLEKYLTTYPE: protein

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 2:

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

 Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp
 5 10 15
 Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly
 20 25 30
 Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
 35 40 45 50
 Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala
 55 60 65
 Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile
 70 75 80
 Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
 85 90 95
 Glu Ala Cys Gly Cys Arg
 100

(2) INFORMASJON FOR SEQ ID NO: 3:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 1207 basepar

(B) TYPE: nukleinsyre

5 (C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

10 (vi) OPPRINNELIG KILDE:

(A) ORGANISME: Homo sapiens

(vii) ØYEBLIKKELEG KILDE:

(B) KLON: MP52

(ix) TREKK:

15 (A) NAVN/NØKKEL: CDS

(B) BELIGGENHET: 845..1204

(xi) SEKVENSBESKRIVELSE FOR SEQ ID NO: 3:

20 ACCGGGCGGC CCTGAACCCA AGCCAGGACA CCCTCCCCAA ACAAGGCAGG CTACAGCCCG 60
 GACTGTGACC CAAAAGGAC AGCTTCCCGG AGGCAAGGCA CCCCCAAAAG CAGGATCTGT 120
 CCCAGCTCC TTCCTGCTGA AGAAGGCCAG GGAGCCCGGG CCCCACGAG AGCCCAAGGA 180
 GCCGTTTCGC CCACCCCCCA TCACACCCA CGAGTACATG CTCTCGCTGT ACAGGACGCT 240
 GTCCGATGCT GACAGAAAGG GAGGCAACAG CAGCGTGAAG TTGGAGGCTG GCCTGGCCAA 300
 25 CACCATCACC AGCTTTATTG ACAAAGGGCA AGATGACCGA GGTCCCGTGG TCAGGAAGCA 360
 GAGGTACGTG TTTGACATTA GTGCCCTGGA GAAGGATGGG CTGCTGGGGG CCGAGCTCCG 420
 GATCTTGCGG AAGAAGCCCT CGGACACGGC CAAGCCAGCG GCCCCCGAG GCGGGCGGGC 480
 TGCCAGCTG AAGCTGTCCA GCTGCCCCAG CGGCCGGCAG CGGCCCTCCT TGCTGGATGT 540
 GCGCTCCGTG CCAGGCCTGG ACGGATCTGG CTGGGAGGTG TTCGACATCT GGAAGCTCTT 600
 30 CCGAAACTTT AAGAACTCGG CCCAGCTGTG CCTGGAGCTG GAGGCCTGGG AACGGGCGAG 660
 GGCCGTGGAC CTCCGTGGCC TGGGCTTCGA CCGCGCCGCC CGGCAGGTCC ACGAGAAGGC 720

	CCTGTTCTCTG GTGTTTGGCC GCACCAAGAA ACGGGACCTG TTCITTAATG AGATTAAGGC	780
	CCGCTCTGGC CAGGACGATA AGACCGTGTA TGAGTACCTG TTCAGCCAGC GGCGAABACG	840
5	GCGG GCC CCA CTG GCC ACT CGC CAG GGC AAG CGA CCC AGC AAG AAC CTT Ala Pro Leu Ala Thr Arg Gln Gly Lys Arg Pro Ser Lys Asn Leu 1 5 10 15	889
	AAG GCT CGC TGC AGT CGG AAG GCA CTG CAT GTC AAC TTC AAG GAC ATG Lys Ala Arg Cys Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met 20 25 30	937
10	GGC TGG GAC GAC TGG ATC ATC GCA CCC CTT GAG TAC GAG GCT TTC CAC Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His 35 40 45	985
	TGC GAG GGC CTG TGC GAG TTC CCA TTG CGC TCC CAC CTG GAG CCC ACG Cys Glu Gly Leu Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr 50 55 60	1033
	AAT CAT GCA GTC ATC CAG ACC CTG ATG AAC TCC ATG GAC CCC GAG TCC Asn His Ala Val Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser 65 70 75	1081
15	ACA CCA CCC ACC TGC TGT GTG CCC ACG CGG CTG AGT CCC ATC AGC ATC Thr Pro Pro Thr Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile 80 85 90 95	1129
	CTC TTC ATT GAC TCT GCC AAC AAC GTG GTG TAT AAG CAG TAT GAG GAC Leu Phe Ile Asp Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp 100 105 110	1177
20	ATG GTC GTG GAG TCG TGT GGC TGC AGG TAG Met Val Val Glu Ser Cys Gly Cys Arg 115 120	1207

25 (2) INFORMASJON FOR SEQ ID NO: 4:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 120 aminosyrer

(B) TYPE: aminosyre

(D) TOPOLOGI: linear

30 (ii) MOLEKYLTTYPE: protein

(xi) SEKVENSBEKRIVELSE: SEQ ID NO: 4:

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

51

His Ala Val Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr
 65 70 75 80
 Pro Pro Thr Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu
 85 90 95
 Phe Ile Asp Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met
 100 105 110
 Val Val Glu Ser Cys Gly Cys Arg
 115 120

(2) INFORMASJON FOR SEQ ID NO: 5:

- (i) SEKVENSKARAKTERTREKK:
 (A) LENGDE: 128 basepar
 (B) TYPE: nukleinsyre
 (C) TRÅDTYPE: enkel
 (D) TOPOLOGI: lineær
 (ii) MOLEKYLTTYPE: DNA (genomisk)
 (vi) ØPRINNELIG KILDE:
 (A) ORGANISME:Homo Sapiens
 (vii) ØYEBLIKKELIG KILDE:
 (B) KLON: VI-1 fragment
 (ix) TREKK:
 (A) NAVN/NØKKEL:CDS
 (B) BELIGGENHET: 28..102
 (xi) SEKVENSBESKRIVELSE: SEQ ID NO: 5:

GGATCCTGGA AGGATTGGAT CATTGCG CCG CTG GAC TAC GAG GCG TAC CAC 51
 Pro Leu Asp Tyr Glu Ala Tyr His
 1 5
 TGC GAG GGC CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG CCC ACC 99
 Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr
 10 15 20
 AAC CACGCTATAG TCCAAACCTT TCTAGA 128
 Asn
 25

35

(2) INFORMASJON FOR SEQ ID NO: 6:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 25 aminosyrer

(B) TYPE: aminosyre

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: protein

(xi) SEKVENSBEKRIVELSE: SEQ ID NO: 6:

Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Leu Cys Asp Phe Pro
 1 5 10 15
 Leu Arg Ser His Leu Glu Pro Thr Asn
 20 25

(2) INFORMASJON FOR SEQ ID NO: 7:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 128 basepar

(B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

(vi) OPPRINNELIG KILDE: Homo Sapiens

(vii) ØYEBLIKKELIG KILDE:

(B) KLON: VL-1

(ix) TREKK:

(A) NAVN/NØKKEL: CDS

(B) BELIGGENHET: 28..102

(xi) SEKVENSBEKRIVELSE: SEQ ID NO: 7:

GGATCCTGGG ATGACTGGAT TATGGCG CCG CTG GAC TAC GAG GCG TAC CAC 51
 Pro Leu Asp Tyr Glu Ala Tyr His
 1 5
 TGC GAG GGT GTA TGC GAC TTC CCG CTG CGC TCG CAC CTG GAG CCC ACC 99
 Cys Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr
 10 15 20
 AAC CACGCCATGC TACAAACGCT TCTAGA 128
 Asn
 25

(2) INFORMASJON FOR SEQ ID NO: 8:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 25 aminosyrer

(B) TYPE: aminosyre

(D) TOPOLOGI: lineær

(ii) MOLEKYLTYPE: protein

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 8:

Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro
 1 5 10 15
 Leu Arg Ser His Leu Glu Pro Thr Asn
 20 25

(2) INFORMASJON FOR SEQ ID NO: 9:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 3585 basepar

(B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTYPE: DNA (genomisk)

(vii) ØYEBLIKKELIG KILDE:

(B) KLON:pALV1-781

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 9:

CTAACTACCC AACTCAAAAA AAAAAAAAAA AAAAACCCCC TCTAACCCCC ATTGACGAAA 60
 GGGCCTCGTG ATACGCCTAT TTTTATAGGT TAATGTCATG ATAATAATGG TTTCTTAGAC 120
 GTCAGGTGGC ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTTCTAAT 180

ACATTCAAT ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTO	240
AAAAAGGAAG AGTATGAGTA TTCAACATTT CCGTGTGCGC CTTATTCCTT TTTTGGCGC	300
ATTTTGCTT CCGTTTTTG CTCACCCAGA AACGCTGGTG AAAGTAAAG ATGCTGAAGA	360
TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA	420
GAGTTTTGCG CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTAAAGTTC TGCTATGTGG	480
CGCGGTATTA TCCCGTATTG ACGCCGGGCA AGAGCAATC GGTGCGCGCA TACACTATTG	540
TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC	600
AGTAAGAGAA TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT	660
TCTGACAACG ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA	720
TGTAATCGC CTTGATCGTT GGGAAACCGA GCTGAATGAA GCCATACCA ACGACGAGCG	780
TGACACCAG ATGCGGTAG CAATGGCAAC AACGTTGCGC AAATATTAA CTGGCGAACT	840
ACTTACTCTA GCTTCCCGC AACAAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG	900
ACCACTTCTG CGCTCGGCCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG	960
TGAGCGTGGG TCTCGCGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCGTAT	1020
CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC	1080
TGAGATAGGT GCCTCACTGA TTAAGCATTG GTAACGTCA GACCAAGTTT ACTCATATAT	1140
ACTTTAGATT GATTTAAAC TTCATTTTAA ATTTAAAGG ATCTAGGTGA AGATCCTTTT	1200
TGATAATCTC ATGACCAAAA TCCCTTAACG TGAGTTTTCG TTCCACTGAG CGTCAGACCC	1260
CGTAGAAAAG ATCAAAGGAT CTCCTTGAGA TCCTTTTTT CTGCGGTAA TCTGCTGCTT	1320
GCAAAACAAA AAACCACCGC TACCAGCGGT GGTGTGTTG CCGGATCAAG AGCTACCAAC	1380
TCTTTTTCCG AAGGTAAC TGCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT	1440
GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCTACAT ACCTCGCTCT	1500
GCTAATCCTG TTACCACTGG CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA	1560
CTCAAGACGA TAGTTACCG ATAAGCGCA GCGGTCGGC TGAACGGGG GTTCGTGCAC	1620
ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCATTG	1680
AGAAAGCGCC ACGCTTCCG AAGGGAGAAA GCGGACAGG TATCCGTAA GCGGCAGGGT	1740
CGGAACAGGA GAGCGCACGA GGGAGCTTC AGGGGAAAC GCCTGGTATC TTTATAGTCC	1800
TGTCGGGTTT CGCCACCTCT GACTTGAGCG TCGATTTTG TGATGCTCGT CAGGGGGCG	1860
GAGCCTATGG AAAACGCCA GCAACGCGC CTTTTACGG TTCTGGCCT TTTGCTGGCC	1920
TTTTGCTCAC ATGTTCTTC CTGCGTTATC CCCTGATTCT GTGGATAACC GTATTACCGC	1980
CTTTGAGTGA GCTGATACCG CTCGCCGAG CCGAACGACC GAGCGCAGCG AGTCAGTGAG	2040
CGAGGAAGCG GAAGAGCGC CAATACGCA ACCGCTCTC CCCGCGGTT GCGCGATTCA	2100
TTAATGCAGA ATTGATCTCT CACCTACCA ACAATGCCCC CCTGCAAAA ATAAATTCAT	2160
ATAAAAAACA TACAGATAAC CATCTGCGGT GATTAATTAT CTCTGGCGGT GTTGACATAA	2220

	ATACCACTGG	CGGTGATACT	GAGCACATCA	GCAGGACGCA	CTGACCACCA	TGAAGGTGAC	2280
	GCTCTTAAA	ATTAAGCCCT	GAAGAAGGGC	AGCATTCAAA	GCAGAAGGCT	TTGGGGTGTG	2340
5	TGATACGAAA	CGAAGCATTG	GCCGTAAGTG	CGATTCCGGA	TTAGCTGCCA	ATGTGCCAAT	2400
	CGCGGGGGGT	TTTCGTTTCA	GACTACAACT	GCCACACACC	ACCAAAGCTA	ACTGACAGGA	2460
	GAATCCAGAT	GGATGCACAA	ACACGCCGCC	GCGAACGTCG	CGCAGAGAAA	CAGGCTCAAT	2520
	GGAAAGCAGC	AAATCCCCTG	TTGGTTGGGG	TAAGCGCAAA	ACCAGTTCGG	AAAGATTTTT	2580
	TTAACTATAA	ACGCTGATGG	AAGCGTTTAT	GCGGAAGAGG	TAAAGCCCTT	CCCGAGTAAC	2640
10	AAAAAAACAA	CAGCATAAAT	AACCCCGCTC	TTACACATTC	CAGCCCTGAA	AAAGGGCATC	2700
	AAATTAAACC	ACACCTATGG	TGTATGCATT	TATTTGCATA	CATTCAATCA	ATTGTTATCT	2760
	AAGGAAATAC	TTACATATGT	CTCGTTGTTT	TCGTAAACCA	CTGCATGTAG	ATTTTAAAGA	2820
	GCTCGGCTGG	GACGACTGGA	TCATCGCGCC	GCTGGACTAC	GAGGCGTACC	ACTGCGAGGG	2880
	CCTTTGCGAC	TTCCCTTTGC	GTTGCGACCT	CGAGCCCAAC	AACCATGCCA	TCAITTCAGAC	2940
15	GCTGCTCAAC	TCCATGGCAC	CAGACGCGGC	GCCGGCCTCC	TGCTGTGTGC	CAGCGCGCCT	3000
	CAGCCCCATC	AGCATCCTCT	ACATCGACGC	CGCCAACAAC	GTGTGTCTAC	AGCAATACGA	3060
	GGACATGGTG	GTGGAGGCCT	GCGGCTGCAG	GTAGTCTAGA	GTCGACCTGC	AGTAATCGTA	3120
	CAGGGTAGTA	CAATATAAAA	AGGCACGTCA	GATGACGTGC	CTTTTCTCTT	GTGAGCAGTA	3180
	AGCTTGGCAC	TGGCCGTCGT	TTTACAACGT	COTGACTGGG	AAAACCCTGG	CGTTACCCAA	3240
20	CTTAATCGCC	TTGCAGCACA	TCCCCCTTTC	GCCAGCTGGC	GTAATAGCGA	AGAGGCCCGC	3300
	ACCGATCGCC	CTTCCCAACA	GTTGCGCAGC	CTGAATGGCG	AATGGCGCCT	GATGCGGTAT	3360
	TTTCTCCTTA	CGCATCTGTG	CGGTATTTCA	CACCGCATAT	ATGGTGCACT	CTCAGTACAA	3420
	TCTGCTCTGA	TGCCGCATAG	TTAAGCCAGC	CCCGACACCC	GCCAACACCC	GCTGACGCGC	3480
25	CCTGACGGGC	TTGTCTGCTC	CCGGCATCCG	CTTACAGACA	AGCTGTGACC	GTCTCCGGGA	3540
	GCTGCATGTG	TCAGAGGTTT	TCACCGTCAT	CACCGAAACG	CGCGA		3585

30

(2) INFORMASJON FOR SEQ ID NO: 10:

(i) SEKVENNS KARAKTERTREKK:

(A) LENGDE: 272 basepar

35

(B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

(vi) OPPRINNELIG KILDE:

(A) ORGANISEM: mus

(vii) ØYEBLIKKELIG KILDE:

5 (B) KLON: mV1

(ix) TREKK:

(A) NAVN/NØKKEL:CDS

(B) BELIGGENHET: 28..243

10 (xi) SEKVENSBESKRIVELSE: SEQ ID NO: 10:

	GGATCCAAGG AGCTCGGCTG GGACGAC TGG ATC ATC GCG CCA TTA GAC, TAC	51
	Trp Ile Ile Ala Pro Leu Asp Tyr	
	1 5	
15	GAG GCA TAC CAC TGC GAG GGC GTT TGC GAC TTT CCT CTG CGC TCG CAC	99
	Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His	
	10 15 20	
	CTG GAG CCT ACC AAC CAC GCC ATC ATT CAG ACG CTG CTC AAC TCC ATG	147
	Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met	
	25 30 35 40	
20	GCG CCC GAC GCT GCG CCA GCC TCC TGC TGC GTG CCC GCA AGG CTC AGT	195
	Ala Pro Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser	
	45 50 55	
	CCC ATC AGC ATT CTC TAC ATC GAT GCC GCC AAC AAC GTG GTC TAC AAG	243
	Pro Ile Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr Lys	
	60 65 70	
25	CAATACGAGG ACATGGTGGT GGGGAATTC	272
30		

(2) INFORMASJON FOR SEQ ID NO: 11:

35 (i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 72 aminosyrer

(B) TYPE: aminosyre

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: protein

(xi) SEKVENSBESKRIVELSE: SEQ ID NO:11:

5

```

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

```

15

20 (2) INFORMASJON FOR SEQ ID NO: 12:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 272 basepar

(B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

25

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

(vi) OPPRINNELIG KILDE:

(A) ORGANISME: mus

(vii) ØYEBLIKKELIG KILDE:

30

(B) KLON: mV2

(ix) TREKK:

(A) NAVN/NØKKEL: CDS

(B) BELIGGENHET: 28..243

35

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 12:

	GGATCCAAGG AGCTCGGCTG GGACGAC TGG ATT ATC GCG CCC CTA GAG TAC	51
	Trp Ile Ile Ala Pro Leu Glu Tyr	
	1 5	
5	GAG GCC TAT CAC TGC GAG GGC GTG TGC GAC TTT CCG CTG CGC TCG CAC	99
	Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His	
	10 15 20	
	CTT GAG CCC ACT AAC CAT GCC ATC ATT CAG ACG CTG ATG AAC TCC ATG	147
	Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu Met Asn Ser Met	
	25 30 35 40	
10	GAC CCG GGC TCC ACC CCG CCT AGC TGC TGC GTT CCC ACC AAA CTG ACT	195
	Asp Pro Gly Ser Thr Pro Pro Ser Cys Cys Val Pro Thr Lys Leu Thr	
	45 50 55	
	CCC ATT AGC ATC CTG TAC ATC GAC GCG GGC AAT AAT GTA GTC TAC AAG	243
	Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly Asn Asn Val Val Tyr Lys	
	60 65 70	
	CAATACGAGG ACATGGTGGT GGGGAATTC	272

15

20

25 (2) INFORMASJON FOR SEQ ID NO:13:

(i) SEKVENNS KARAKTERTREKK:

(A) LENGDE: 72 aminosyrer

(B) TYPE: aminosyre

(D) TOPOLOGI: lineær

30 (ii) MOLEKYLTTYPE: protein

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 13:

35

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

(2) INFORMASJON FOR SEQ ID NO:14:

15 (i) SEKVENSKARAKTERTREKK:

- (A) LENGDE: 272 basepar
- (B) TYPE: nukleinsyre
- (C) TRPDTYPE: enkel
- (D) TOPOLOGI: lineær

20 (ii) MOLEKYLTTYPE: DNA (genomisk)

(vi) OPPRINNELIG KILDE:

- (A) ORGANISME: mus

(vii) ØYEBLIKKELIG KILDE:

- (B) KLON: mV9

25 (ix) TREKK:

- (A) NAVN/NØKKEL: CDS
- (B) BELIGGENHET: 28..243

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 14:

30

35

	GGATCCAAGG AGCTCGGCTG GGACGAC TGG ATC ATC GCA CCT CTT GAG TAT	51
	Trp Ile Ile Ala Pro Leu Glu Tyr	
	1 5	
5	GAG GCC TTC CAC TGC GAA GGA CTG TGT GAG TTC CCC TTG CGC TCC CAC	99
	Glu Ala Phe His Cys Glu Gly Leu Cys Glu Phe Pro Leu Arg Ser His	
	10 15 20	
	TTG GAG CCC ACA AAC CAC GCA GTC ATT CAG ACC CTA ATG AAC TCT ATG	147
	Leu Glu Pro Thr Asn His Ala Val Ile Gln Thr Leu Met Asn Ser Met	
	25 30 35 40	
10	GAC CCT GAA TCC ACA CCA CCC ACT TGT TGT GTG CCT ACA CGG CTG AGT	195
	Asp Pro Glu Ser Thr Pro Pro Thr Cys Cys Val Pro Thr Arg Leu Ser	
	45 50 55	
	CCT ATT AGC ATC CTC TTC ATC GAC TCT GCC AAC AAC GTG GTG TAT AAA	243
	Pro Ile Ser Ile Leu Phe Ile Asp Ser Ala Asn Asn Val Val Tyr Lys	
	60 65 70	
	CAATACGAGG ACATGGTGCT GGGGAATTC	272
15		

(2) INFORMASJON FOR SEQ ID NO: 15:

(i) SEKVENS KARAKTERTREKK:

(A) LENGDE: 72 aminosyrer

(B) TYPE: aminosyre

(D) TOPOLOGI: lineær

(ii) MOLEKYLTYPE: protein

(xi) SEKVENSBEKRIVELSE: SEQ ID NO:15:

30

Trp 1 Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys Glu Gly Leu
5 10 15

Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Val
20 25 30

Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr Pro Pro Thr
35 40 45

Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp
50 55 60

35

Ser Ala Asn Asn Val Val Tyr Lys
65 70

(2) INFORMASJON FOR SEQ ID NO: 16:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 7 aminosyrer

(B) TYPE: aminosyre

5 (C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: peptid

(vi) OPPRINNELIG KILDE:

(A) ORGANISME: BMP/TGF-beta konsensus sekvens

10

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 16:

Trp Xaa Asp Trp Ile Xaa Ala

1

5

15 (2) INFORMASJON FOR SEQ ID NO: 17:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 27 basepar

(B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

20 (D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

(vi) OPPRINNELIG KILDE:

(B) KLON: oligonukleotid nr. 1

25 (xi) SEKVENSBESKRIVELSE: SEQ ID NO: 17:

CGGATCCTGG VANGAYTGGA THRTNGC

27

(2) INFORMASJON FOR SEQ ID NO: 18:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 6 aminosyrer

(B) TYPE: aminosyre

5 (C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: peptid

(vi) OPPRINNELIG KILDE:

(B) KLON: BMP/TGF-beta konsensus sekvens

10

(xi) SEKVENSBESKRIVELSE: SEQ ID NO:18:

His Ala Ile Gln Thr

1 5

15

(2) INFORMASJON FOR SEQ ID NO: 19:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 28 basepar

(B) TYPE: nukleinsyre

20 (C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

25

(vii) ØYEBLIKKELIG KILDE:

(B) KLON:oligonukleotid nr 2

(xi) SEKVENSBESKRIVELSE: SEQ ID NO:19:

TTTCTAGAAR NGTYTGNACD ATNGCRTG

28

30

(2) INFORMASJON FOR SEQ ID NO: 20:

SEKVENSKARAKTERTREKK:

(A) LENGDE: 40 basepar

(B) TYPE: nukleinsyre

35 (C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

(vii) ØYEBLIKKELIG KILDE:

(B) KLON: oligonukleotid nr. 3

5

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 20:

CCACTGCGAG GGCCTTTGCG ACTTCCCTTT GCGTTCGCAC

40

10 (2) INFORMASJON FOR SEQ ID NO: 21:

SEKVENNS KARAKTERTREKK:

(A) LENGDE: 29 basepar

(B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

15

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

(vii) ØYEBLIKKELIG KILDE:

20

(A) BIBLIOTEK: oligonukleotid nr. 4

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 21:

TGCGGATCCA GCCGCTGCAG CCGCAAGCC

29

25

(2) INFORMASJON FOR SEQ ID NO: 22:

SEKVENNS KARAKTERTREKK:

(A) LENGDE: 29 basepar

(B) TYPE: nukleinsyre

30

(C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

35

(vii) ØYEBLIKKELIG KILDE:

(B)KLON: oligonukleotid nr. 5

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 22:

GACTCTAGAC TACCTGCAGC CGCAGGCCT

29

5 (2) INFORMASJON FOR SEQ ID NO: 23:

SEKVENNS KARAKTERTREKK:

(A) LENGDE: 28 basepar

(B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

10 (D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

(vii) ØYEBLIKKELIG KILDE:

15 (A) BIBLIOTEK: oligonukleotid nr. 6

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 23:

GCGGATCCAA GGAGCTCGGC TGGGACGA

28

20

(2) INFORMASJON FOR SEQ ID NO: 24:

SEKVENNS KARAKTERTREKK:

(A) LENGDE: 28 basepar

25 (B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

30

(vii) ØYEBLIKKELIG KILDE:

(B) KLON: oligonukleotid nr. 7

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 24:

35

GGAATTCCCC ACCACCATGT CCTCGTAT

28

(2) INFORMASJON FOR SEQ ID NO: 25:

SEKVENSKARAKTERTREKK:

(A) LENGDE: 1171 basepar

(B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

(vii) ØYEBLIKKELIG KILDE:

(B) KLON: human VL-1 protein

(ix) TREKK:

(A) NAVN/NØKKEL: CDS

(B) BELIGGENHET: 2..964

(ix) TREKK:

(A) NAVN/NØKKEL: mat_peptid

(B) BELIGGENHET: 605..964

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 25:

G AAT TCG GAT CTC TCG CAC ACT CCT CTC CGG AGA CAG AAG TAT TTG	46
Asn Ser Asp Leu Ser His Thr Pro Leu Arg Arg Gln Lys Tyr Leu	
-201-200 -195 -190	
TTT GAT GTG TCC ATG CTC TCA GAC AAA GAA GAG CTG GTG GGC GCG GAG	94
Phe Asp Val Ser Met Leu Ser Asp Lys Glu Glu Leu Val Gly Ala Glu	
-185 -180 -175	
CTG CGG CTC TTT CGC CAG GCG CCC TCA GCG CCC TGG GGG CCA CCA GCC	142
Leu Arg Leu Phe Arg Gln Ala Pro Ser Ala Pro Trp Gly Pro Pro Ala	
-170 -165 -160 -155	
GGG CCG CTC CAC GTG CAG CTC TTC CCT TGC CTT TCG CCC CTA CTG CTG	190
Gly Pro Leu His Val Gln Leu Phe Pro Cys Leu Ser Pro Leu Leu Leu	
-150 -145 -140	
GAC GCG CGG ACC CTG GAC CCG CAG GGG GCG CCG CCG GCC GGC TGG GAA	238
Asp Ala Arg Thr Leu Asp Pro Gln Gly Ala Pro Pro Ala Gly Trp Glu	

-135	-130	-125	
TTC GAC GTG TGG CAG GGC CTG CGC CAC CAG CCC TGG AAG CAG CTG Val Phe Asp Val Trp Gln Gly Leu Arg His Gln Pro Trp Lys Gln Leu -120 -115 -110			286
TGC TTG GAG CTG CGG GCC GCA TGG GGC GAG CTG GAC GCC GGG GAG GCC Cys Leu Glu Leu Arg Ala Ala Trp Gly Glu Leu Asp Ala Gly Glu Ala -105 -100 -95			334
GAG GCG CGC GCG CGG GGA CCC CAG CAA CCG CCG CCC CCG GAC CTG CGG Glu Ala Arg Ala Arg Gly Pro Gln Gln Pro Pro Pro Asp Leu Arg -90 -85 -80 -75			382
AGT CTG GGC TTC GGC CGG AGG GTG CGG CCT CCC CAG GAG CGG GCC CTG Ser Leu Gly Phe Gly Arg Arg Val Arg Pro Gln Glu Arg Ala Leu -70 -65 -60			430
CTG GTG GTA TTC ACC AGA TCC CAG CGC AAG AAC CTG TTC GCA GAG ATG Leu Val Val Phe Thr Arg Ser Gln Arg Lys Asn Leu Phe Ala Glu Met -55 -50 -45			478
CGC GAG CAG CTG GGC TCG GCC GAG GCT GCG GGC CCG GGC GCG GGC GCC Arg Glu Gln Leu Gly Ser Ala Glu Ala Ala Gly Pro Gly Ala Gly Ala -40 -35 -30			526
GAG GGG TCG TGG CCG CCG CCG TCG GGC GCC CCG GAT GCC AGG CCT TGG Glu Gly Ser Trp Pro Pro Ser Gly Ala Pro Asp Ala Arg Pro Trp -25 -20 -15			574
CTG CCC TCG CCC GGC CGC CGG CGG CGC ACG GCC TTC GCC AGT CGC Leu Pro Ser Pro Gly Arg Arg Arg Arg Thr Ala Phe Ala Ser Arg -10 -5 1 5			622
CAT GGC AAG CCG CAC GGC AAG AAG TCC AGG CTA CGC TGC AGC AAG AAG His Gly Lys Arg His Gly Lys Lys Ser Arg Leu Arg Cys Ser Lys Lys 10 15 20			670
CCC CTG CAC GTG AAC TTC AAG GAG CTG GGC TGG GAC GAC TGG ATT ATC Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile 25 30 35			718
GCG CCC CTG GAG TAC GAG GCC TAT CAC TGC GAG GGT GTA TGC GAC TTC Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe 40 45 50			766
CCG CTG CCG TCG CAC CTG GAG CCC ACC AAC CAC GCC ATC ATC CAG ACG Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr 55 60 65 70			814
CTG ATG AAC TCC ATG GAC CCC GGC TCC ACC CCG CCC AGC TGC TGC GTG Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro Ser Cys Cys Val 75 80 85			862
CCC ACC AAA TTG ACT CCC ATC AGC ATT CTA TAC ATC GAC GCG GGC AAT Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly Asn 90 95 100			910
AAT GTG GTC TAC AAG CAG TAC GAG GAC ATG GTG GTG GAG TCG TGC GGC Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys Gly 105 110 115			958
TGC AGG TAGCGGTGCC TTTCGCGCCG CCTTGGCCCG GAACCAAGGT GGGCCAAGGT Cys Arg 120			1014
CCGCCTTGCA GGGGAGGCCT GGCTGCAGAG AGGCGGAGGA GGAAGCTGGC GCTGGGGGAG			1074

GCTGAGGGTG AGGGAACAGC CTGGATGTGA GAGCCGCTGG GAGAGAAGGG AGCGCACCTT 1134
 CCCAGTAACT TCTACCTGCC AGCCCAGAGG GAAATAT 1171

5

(2) INFORMASJON FOR SEQ ID NO: 26:

- 10 (i) SEKVENSTREKK:
 (A) LENGDE: 321 aminosyrer
 (B) TYPE: aminosyre
 (D) TOPOLOGI: lineær

- 15 (ii) MOLEKYLTYP: protein

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 26:

Asn Ser Asp Leu Ser His Thr Pro Leu Arg Arg Gln Lys Tyr Leu Phe
 -201 -200 -195 -190
 Asp Val Ser Met Leu Ser Asp Lys Glu Glu Leu Val Gly Ala Glu Leu
 -185 -180 -175 -170
 Arg Leu Phe Arg Gln Ala Pro Ser Ala Pro Trp Gly Pro Pro Ala Gly
 -165 -160 -155
 Pro Leu His Val Gln Leu Phe Pro Cys Leu Ser Pro Leu Leu Asp
 -150 -145 -140
 Ala Arg Thr Leu Asp Pro Gln Gly Ala Pro Pro Ala Gly Trp Glu Val
 -135 -130 -125
 Phe Asp Val Trp Gln Gly Leu Arg His Gln Pro Trp Lys Gln Leu Cys
 -120 -115 -110
 Leu Glu Leu Arg Ala Ala Trp Gly Glu Leu Asp Ala Gly Glu Ala Glu
 -105 -100 -95 -90
 Ala Arg Ala Arg Gly Pro Gln Gln Pro Pro Pro Pro Asp Leu Arg Ser
 -85 -80 -75
 Leu Gly Phe Gly Arg Arg Val Arg Pro Pro Gln Glu Arg Ala Leu Leu
 -70 -65 -60
 Val Val Phe Thr Arg Ser Gln Arg Lys Asn Leu Phe Ala Glu Met Arg
 -55 -50 -45

35

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

20

25

(2) INFORMASJON FOR SEQ ID NO: 27:

(i) SEKVENNS KARAKTERTREKK:

- (A) LENGDE: 1233 basepar
 (B) TYPE: nukleinsyre
 (C) TRÅDTYPE: enkel
 (D) TOPOLOGI: lineær

30

(ii) MOLEKYLTTYPE: DNA (genomisk)

35

(vii) ØYEBLIKKELIG KILDE:

(B) KLON: DNA kodende for BMP2 propeptid/BMP-12 modent peptid

(ix) TREKK:

(A) NAVN/NØKKEL: CDS

(B) BELIGGENHET: 1..1233

(ix) TREKK:

5 (A) NAVN/NØKKEL: mat_peptid

(B) BELIGGENHET: 847..1233

(xi) SEKVENSBEKRIVELSE: SEQ ID NO: 27:

10	ATG GTG GCC GGG ACC CGC TGT CTT CTA GCG TTG CTG CTT CCC CAG GTC Met Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Pro Gln Val -282 -280 -275 -270	48
	CTC CTG GGC GGC GCG GCT GGC CTC GTT CCG GAG CTG GGC CGC AGG AAG Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys -265 -260 -255	96
15	TTC GCG GCG GCG TCG TCG GGC CGC CCC TCA TCC CAG CCC TCT GAC GAG Phe Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu -250 -245 -240 -235	144
	GTC CTG AGC GAG TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA Val Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys -230 -225 -220	192
	CAG AGA CCC ACC CCC AGC AGG GAC GCC GTG GTG CCC CCC TAC ATG CTA Gln Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu -215 -210 -205	240
20	GAC CTG TAT CGC AGG CAC TCA GGT CAG CCG GGC TCA CCC GCC CCA GAC Asp Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp -200 -195 -190	288
	CAC CGG TTG GAG AGG GCA GCC AGC CGA GCC AAC ACT GTG CGC AGC TTC His Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe -185 -180 -175	336
25	CAC CAT GAA GAA TCT TTG GAA GAA CTA CCA GAA ACG AGT GGG AAA ACA His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr -170 -165 -160 -155	384
	ACC CGG AGA TTC TTC TTT AAT TTA AGT TCT ATC CCC ACG GAG GAG TTT Thr Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe -150 -145 -140	432
30	ATC ACC TCA GCA GAG CTT CAG GTT TTC CGA GAA CAG ATG CAA GAT GCT Ile Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala -135 -130 -125	480
	TTA GGA AAC AAT AGC AGT TTC CAT CAC CGA ATT AAT ATT TAT GAA ATC Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile -120 -115 -110	528
	ATA AAA CCT GCA ACA GCC AAC TCG AAA TTC CCC GTG ACC AGA CTT TTG Ile Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu -105 -100 -95	576
	GAC ACC AGG TTG GTG AAT CAG AAT GCA AGC AGG TGG GAA AGT TTT GAT Asp Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp -90 -85 -80 -75	624

	GTC ACC CCC GCT GTG ATG CGG TGG ACT GCA CAG GGA CAC GCC AAC CAT Val Thr Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His -70 -65 -60	672
5	GGA TTC GTG GTG GAA GTG GCC CAC TTG GAG GAG AAA CAA GGT GTC TCC Gly Phe Val Val Glu Val Ala His Leu Glu Glu Lys Gln Gly Val Ser -55 -50 -45	720
	AAG AGA CAT GTT AGG ATA AGC AGG TCT TTG CAC CAA GAT GAA CAC AGC Lys Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His Ser -40 -35 -30	768
	TGG TCA CAG ATA AGG CCA TTG CTA GTA ACT TTT GGC CAT GAT GGA AAA Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp Gly Lys -25 -20 -15	816
10	GGG CAT CCT CTC CAC AAA AGA GAA AAA CGT ACG GCG TTG GCC GGG ACG Gly His Pro Leu His Lys Arg Glu Lys Arg Thr Ala Leu Ala Gly Thr -10 -5 1 5	864
	CGG ACA GCG CAG GGC AGC GGC GGG GGC GCG GGC CGG GGC CAC GGG CGC Arg Thr Ala Gln Gly Ser Gly Gly Ala Gly Arg Gly His Gly Arg 10 15 20	912
15	AGG GGC CGG AGC CGC TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG Arg Gly Arg Ser Arg Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys 25 30 35	960
	GAG CTC GGC TGG GAC GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala 40 45 50	1008
20	TAC CAC TGC GAG GGC CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG Tyr His Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu 55 60 65 70	1056
	CCC ACC AAC CAT GCC ATC ATT CAG ACG CTG CTC AAC TCC ATG GCA CCA Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro 75 80 85	1104
25	GAC GCG GCG CCG GCC TCC TGC TGT GTG CCA GCG CGC CTC AGC CCC ATC Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile 90 95 100	1152
	AGC ATC CTC TAC ATC GAC GCC GCC AAC AAC GTT GTC TAC AAG CAA TAC Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr 105 110 115	1200
30	GAG GAC ATG GTG GTG GAG GCC TGC GGC TGC AGG Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg 120 125	1233

(i) SEKVENNS KARAKTERTREKK:

(A) LENGDE: 411 aminosyrer

(B) TYPE: aminosyre

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: protein

(xi) SEKVENNS BESKRIVELSE: SEQ ID NO: 28:

Met Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Pro Gln Val
 -282 -280 -275 -270
 Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys
 -265 -260 -255
 Phe Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu
 -250 -245 -240 -235
 Val Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys
 -230 -225 -220
 Gln Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu
 -215 -210 -205
 Asp Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp
 -200 -195 -190
 His Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe
 -185 -180 -175
 His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr
 -170 -165 -160 -155
 Thr Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe
 -150 -145 -140
 Ile Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala
 -135 -130 -125
 Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile
 -120 -115 -110
 Ile Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu
 -105 -100 -95
 Asp Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp
 -90 -85 -80 -75

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

25

(2) INFORMASJON FOR SEQ ID NO: 29:

30 (i) SEKVENNS KARAKTERTREKK:

(A) LENGDE: 1203 basepar

(B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

35

(ii) MOLEKYLTTYPE: DNA (genomisk)

(vii) ØYEBLIKKELIG KILDE:

(B) KLON: murin MV1

(ix) TREKK:

5 (A) NAVN/NØKKEL:CDS

(B) BELIGGENHET: 2..721

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 29:

10 A AAG TTC TGC CTG GTG CTG GNG ECG GTG ACG GCC TCG GAG AGC AGN 46
 Lys Phe Cys Leu Val Leu X01 X02 Val Thr Ala Ser Glu Ser X03
 1 5 10 15

CNG CTG GCC CTG AGA CGA CTG GGC TTC GGC TGN CCG GGC GGT GGC GAC 94
 X04 Leu Ala Leu Arg Arg Leu Gly Phe Gly X05 Pro Gly Gly Gly Asp
 20 25 30

15 GGC GGC GGC ACT GCG GNC GAG GAG CGC GCG CTG TTG GTG ATC TCC TCC 142
 Gly Gly Gly Thr Ala X06 Glu Glu Arg Ala Leu Leu Val Ile Ser Ser
 35 40 45

CGT ACG CAA AGG AAA GAG AGT CTG TTC CCG GAG ATC CGA GCC CAG GCC 190
 Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile Arg Ala Gln Ala
 50 55 60

20 CGT GCT CTC CCG GCC GCT GCA GAG CCG CCA CCG GAT CCA GGA CCA GGC 238
 Arg Ala Leu Arg Ala Ala Glu Pro Pro Pro Asp Pro Gly Pro Gly
 65 70 75

GCT GGG TCA CGC AAA GCC AAC CTG GGC GGT CGC AGG CCG CAG CCG ACT 286
 Ala Gly Ser Arg Lys Ala Asn Leu Gly Gly Arg Arg Arg Gln Arg Thr
 80 85 90 95

25 GCG CTG GCT GGG ACT CCG GGA GNG NAG GGA AGC GGT GGT GGC GGC GGT 334
 Ala Leu Ala Gly Thr Arg Gly X07 X08 Gly Ser Gly Gly Gly Gly Gly
 100 105 110

GGC GGT GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GCA 382
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ala
 115 120 125

GGC AGG GGC CAC GGG CGC AGA GGC CCG AGC CGC TGC GGT CGC AAG TCA 430
 Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg Cys Gly Arg Lys Ser
 130 135 140

30 CTG CAC GTG GAC TTT AAG GAG CTG GGC TGG GAC GAC TGG ATC ATC GCG 478
 Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala
 145 150 155

5	CCA TTA GAC TAC GAG GCA TAC CAC TGC GAG GGC GTT TGC GAC TTT CCT Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro 160 165 170 175	526
	CTG CGC TCG CAC CTG GAG CCT ACC AAC CAC GCC ATC ATT CAG ACG CTG Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu 180 185 190	574
	CTC AAC TCC ATG GCG CCC GAC GCT GCG CCA GCC TCC TGC TGC GTG CCC Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala Ser Cys Cys Val Pro 195 200 205	622
10	GCA AGG CTC AGT CCC ATC AGC ATT CTC TAC ATC GAT GCC GCC AAC AAC Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn 210 215 220	670
	GTG GTC TAC AAG CAG TAC GAA GAC ATG GTG GTG GAG GCC TGC GGC TGC Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ala Cys Gly Cys 225 230 235	718
15	AGG TAGCATGCGG TCTGGGGAGG GTCTGGCCGC CCAGGACCCT AGCTCAAGAG Arg 240	771
	CAGGTGTCAT CAGGCCCGAG GGACGGCGGA CTATGGCCTC TGCCAGCACA GAGGAGAGCA	831
	CACAGTTAAC ACTCACATTT ACACACTCCT TCACTCAGGC ACATGTTTAC CGTGGACGGC	891
	AGGCGCTAAA AGCCTTGCTT ATTGCTACC ATTGATACAA ACCTCTGTCC TTTTCGGGAG	951
20	AGGGAAGGGC ATCTGTGTTT ATGTTGCACT AATTGGCACT AATCCAAGT AGAAATGGGT	1011
	TAGCATTGGA TTCTCCTTTT AGTTGGAGGC GGTGTGGCTG GATTCTGRC GTTGGATATG	1071
	GRGTGCACTG CAGGGCTGGG ATACCCAGAT TCTCTGGAGT GGGCATTGGG AACCTTCAAA	1131
	AGTAAGGAGC CACTGGGGCT TGGGAGGGAG CACCCGGTTC CTAACAAGT CTGATGTGTA	1191
25	CTGCTCAGTT TG	1203

30

(2) INFORMASJON FOR SEQ ID NO: 30:

(i) SEKVENNS KARAKTERTREKK:

(A) LENGDE: 240 aminosyrer

(B) TYPE: aminosyre

35

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: protein

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 30:

```

Lys Phe Cys Leu Val Leu X01 X02 Val Thr Ala Ser Glu Ser X03 X04
 1           5           10           15
Leu Ala Leu Arg Arg Leu Gly Phe Gly X05 Pro Gly Gly Gly Asp Gly
 20           25           30
Gly Gly Thr Ala X06 Glu Glu Arg Ala Leu Leu Val Ile Ser Ser Arg
 35           40           45
Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile Arg Ala Gln Ala Arg
 50           55           60
Ala Leu Arg Ala Ala Ala Glu Pro Pro Pro Asp Pro Gly Pro Gly Ala
 65           70           75           80
Gly Ser Arg Lys Ala Asn Leu Gly Gly Arg Arg Arg Gln Arg Thr Ala
 85           90           95
Leu Ala Gly Thr Arg Gly X07 X08 Gly Ser Gly Gly Gly Gly Gly Gly
100           105           110
Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ala Gly
115           120           125
Arg Gly His Gly Arg Arg Gly Arg Ser Arg Cys Gly Arg Lys Ser Leu
130           135           140
His Val Asp Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala Pro
145           150           155           160
Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro Leu
165           170           175
Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu
180           185           190
Asn Ser Met Ala Pro Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala
195           200           205
Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val
210           215           220
Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg
225           230           235           240

```

(2) INFORMASJON FOR SEQ ID NO: 31:

(i) SEKVENNS KARAKTERTREKK:

(A) LENGDE: 1046 basepar

(B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

(iii) HYPOTETISK: nei

(iv) ANTI-SENS: nei

(vii) ØYEBLIKKELIG KILDE:

(B) KLON: MURIN MV2

(ix) TREKK:

(A) NAVN/NØKKEL: CDS

(B) BELIGGENHET: 2..790

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 31:

A	AGA	AAA	CAA	GCT	TGC	ATT	CCT	GCA	GGT	CCG	ACT	CTA	AGA	GGA	TCC	46
	Arg	Lys	Gln	Ala	Cys	Ile	Pro	Ala	Gly	Pro	Thr	Leu	Arg	Gly	Ser	
	1				5					10					15	
TCA	GGG	ACC	CAA	CCC	AGG	CCG	GCT	GGG	AAG	TCT	TTC	GAC	GTG	TGG	CAG	94
Ser	Gly	Thr	Gln	Pro	Arg	Pro	Ala	Gly	Lys	Ser	Phe	Asp	Val	Trp	Gln	
				20					25						30	
GGC	CTG	CGC	CCT	CAG	CCT	TGG	AAG	CAG	CTG	TGC	CTG	GAG	TTC	CGG	GCA	142
Gly	Leu	Arg	Pro	Gln	Pro	Trp	Lys	Gln	Leu	Cys	Leu	Glu	Leu	Arg	Ala	
			35					40						45		
GCC	TGG	GGT	GAG	CTG	GAC	RCC	GGG	GAT	ACG	GGG	GCG	CGC	GCG	AGG	GGT	190
Ala	Trp	Gly	Glu	Leu	Asp	X01	Gly	Asp	Thr	Gly	Ala	Arg	Ala	Arg	Gly	
		50					55						60			
CCC	CAG	CAG	CCA	CCG	CCT	CTG	GAC	CTG	CGG	AGT	CTG	GGC	TTC	GGT	CGG	238
Pro	Gln	Gln	Pro	Pro	Pro	Leu	Asp	Leu	Arg	Ser	Leu	Gly	Phe	Gly	Arg	
	65					70					75					

	AGG GTG AGA CCG CCC CAG GAG CGC GCC CTG CTT GTA GTG TTC ACC AGA	286
	Arg Val Arg Pro Pro Gln Glu Arg Ala Leu Leu Val Val Phe Thr Arg	
	80 85 90 95	
5	TCG CAG CGC AAG AAC CTG TTC ACT GAG ATG CAT GAG CAG CTG GGC TCT	334
	Ser Gln Arg Lys Asn Leu Phe Thr Glu Met His Glu Gln Leu Gly Ser	
	100 105 110	
	GCA GAG GCT GCG GGA GCC GAG GGG TCA TGT CCA GCG CCG TCG GGC TCC	382
	Ala Glu Ala Ala Gly Ala Glu Gly Ser Cys Pro Ala Pro Ser Gly Ser	
	115 120 125	
10	CCA GAC ACC GGG TCT TGG CTG CCC TCG CCC GGC CGC CCG CGG CGA CGC	430
	Pro Asp Thr Gly Ser Trp Leu Pro Ser Pro Gly Arg Arg Arg Arg	
	130 135 140	
	ACC GCC TTC GCC AGC CGT CAC GGC AAG CGA CAT GGC AAG AAG TCC AGG	478
	Thr Ala Phe Ala Ser Arg His Gly Lys Arg His Gly Lys Lys Ser Arg	
	145 150 155	
	CTG CGC TGC AGC AGA AAG CCT CTG CAC GTG AAT TTT AAG GAG TTA GGC	526
	Leu Arg Cys Ser Arg Lys Pro Leu His Val Asn Phe Lys Glu Leu Gly	
	160 165 170 175	
15	TGG GAC GAC TGG ATT ATC GCG CCC CTA GAG TAC GAG GCC TAT CAC TGC	574
	Trp Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys	
	180 185 190	
	GAG GGC GTG TGC GAC TTT CCG CTG CGC TCG CAC CTT GAG CCC ACT AAC	622
	Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn	
	195 200 205	
20	CAT GCC ATC ATT CAG ACG CTG ATG AAC TCC ATG GAC CCG GGC TCC ACC	670
	His Ala Ile Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr	
	210 215 220	
	CCG CCT AGC TGC TGC GTT CCC ACC AAA CTG ACT CCC ATT AGC ATC CTG	718
	Pro Pro Ser Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu	
	225 230 235	
25	TAC ATC GAC GCG GGC AAT AAT GTN GTC TAC AAG CAG TAT GAG GAC ATG	766
	Tyr Ile Asp Ala Gly Asn Asn X02 Val Tyr Lys Gln Tyr Glu Asp Met	
	240 245 250 255	
	GTG GTG GAG TCC TGC GGC TGT AGG TAGCGGTGCT GTCCCGCCAC CTGGGCCAGG	820
	Val Val Glu Ser Cys Gly Cys Arg	
	260	
	GACCATGGAG GGAGGCCTGA CTGCCGAGAA AGGAGCAGGA GCTGGCCTTG GAAGAGGCCA	880
30	CAGGTGGGGG ACAGCCTGAA AGTAGGAGCA CAGTAAGAAG CAGCCCAGCC TTCCAGAAC	940
	CTTCCAATCC CCCAACCCAG AAGCAGCTAA GGGGTTTCAC AACTTTTGGC CTGCCCAGCC	1000
	TGGAAAGACT AGACAAGAGG GATTCTTCTC TTTTATTAT GGCTTG	1046

(2) INFORMASJON FOR SEQ ID NO: 32:

(i) SEKVENSKARAKTERTREKK:

(A) LENGDE: 263 aminosyrer

(B) TYPE: aminosyrer

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: protein

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 32:

Arg Lys Gln Ala Cys Ile Pro Ala Gly Pro Thr Leu Arg Gly Ser Ser
 1 5 10 15
 Gly Thr Gln Pro Arg Pro Ala Gly Lys Ser Phe Asp Val Trp Gln Gly
 20 25 30
 Leu Arg Pro Gln Pro Trp Lys Gln Leu Cys Leu Glu Leu Arg Ala Ala
 35 40 45
 Trp Gly Glu Leu Asp X01 Gly Asp Thr Gly Ala Arg Ala Arg Gly Pro
 50 55 60
 Gln Gln Pro Pro Pro Leu Asp Leu Arg Ser Leu Gly Phe Gly Arg Arg
 65 70 75 80
 Val Arg Pro Pro Gln Glu Arg Ala Leu Leu Val Val Phe Thr Arg Ser
 85 90 95
 Gln Arg Lys Asn Leu Phe Thr Glu Met His Glu Gln Leu Gly Ser Ala
 100 105 110
 Glu Ala Ala Gly Ala Glu Gly Ser Cys Pro Ala Pro Ser Gly Ser Pro
 115 120 125
 Asp Thr Gly Ser Trp Leu Pro Ser Pro Gly Arg Arg Arg Arg Thr
 130 135 140
 Ala Phe Ala Ser Arg His Gly Lys Arg His Gly Lys Lys Ser Arg Leu
 145 150 155 160
 Arg Cys Ser Arg Lys Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp
 165 170 175
 Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu
 180 185 190
 Gly Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His
 195 200 205
 Ala Ile Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro
 210 215 220
 Pro Ser Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr
 225 230 235 240
 Ile Asp Ala Gly Asn Asn X02 Val Tyr Lys Gln Tyr Glu Asp Met Val
 245 250 255
 Val Glu Ser Cys Gly Cys Arg
 260

(A) LENGDE: 1345 basepar

(C) TRADTYPE: enkel

(D) TOPOLOGI: lineær

(iii) **HYPOTETISK:** nei

(iv) **ANTI-SENS: nei**

(B) KLON: HUMAN V1-1

(A) NAVN/NØKKEL: CDS

(B) BELIGGENHET: 138..1301

(A) NAVN/NØKKEL: mat peptid

(B) BELIGGENHET: 990..1301

(xi) SEKVENSS BESKRIVELSE: SEQ ID NO: 33:

AACTATAGCA	CCTGCAGTCC	CTGGTCTTGG	GTGTAGGGGT	GCCTCCTGG	TCCCGCGGCT	60
CAGGGATATG	CAGTGACCAA	TGGGTGTGTTG	GCCTGATGGG	ACTTTTGGCT	TGCTAAACCA	120
AAGCTCGGTT	CGGATAG	CCC GGG CGA AGA CGT CCG CTG CTC TGG GCC AGG				170
		Pro Gly Arg Arg Arg Pro Leu Leu Trp Ala Arg				
		-284	-280		-275	
CTG GCA GCG TTC AGG CTG GGG CAG AGA CGC GGA GTC GGG CGC TGG CTC						218
Leu Ala Ala Phe Arg Leu Gly Gln Arg Arg Gly Val Gly Arg Trp Leu						
		-270	-265		-260	

-CAA CAG GCC TGG CTC CCA CAT CGA AGA CAG CTG GGC CAT TTG CTG TTA Gln Gln Ala Trp Leu Pro His Arg Arg Gln Leu Gly His Leu Leu Leu -255 -250 -245	266
GGA GGC CCC GCG CTG ACA GTG TGC AGG ATT TGC TCT TAC ACA GCT CTT Gly Gly Pro Ala Leu Thr Val Cys Arg Ile Cys Ser Tyr Thr Ala Leu -240 -235 -230	314
TCT CTC TGT CCC TGC CGG TCC CCC GCA GAC GAA TCG GCA GCC GAA ACA Ser Leu Cys Pro Cys Arg Ser Pro Ala Asp Glu Ser Ala Ala Glu Thr -225 -220 -215 -210	362
GGC CAG AGC TTC CTG TTC GAC GTG TCC AGC CTT AAC GAC GCA GAC GAG Gly Gln Ser Phe Leu Phe Asp Val Ser Ser Leu Asn Asp Ala Asp Glu -205 -200 -195	410
GTG GTG GGT GCC GAG CTG CGC GTG CTG CGC CGG GGA TCT CCA GAG TCG Val Val Gly Ala Glu Leu Arg Val Leu Arg Arg Gly Ser Pro Glu Ser -190 -185 -180	458
GGC CCA GGC AGC TGG ACT TCT CCG CCG TTG CTG CTG CTG TCC ACG TGC Gly Pro Gly Ser Trp Thr Ser Pro Pro Leu Leu Leu Ser Thr Cys -175 -170 -165	506
CCG GGC GCC GCC CGA GCG CCA CGC CTG CTG TAC TCG CGG GCA GCT GAG Pro Gly Ala Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu -160 -155 -150	554
CCC CTA GTC GGT CAG CGC TGG GAG GCG TTC GAC GTG GCG GAC GCC ATG Pro Leu Val Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met -145 -140 -135 -130	602
AGG CGC CAC CGT CGT GAA CCG CGC CCC CCC CGC GCG TTC TGC CTC TTG Arg Arg His Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu -125 -120 -115	650
CTG CGC GCA GTG GCA GGC CCG GTG CCG AGC CCG TTG GCA CTG CGG CGA Leu Arg Ala Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg -110 -105 -100	698
CTG GGC TTC GGC TGG CCG GGC GGA GGG GGC TCT GCG GCA GAG GAG CGC Leu Gly Phe Gly Trp Pro Gly Gly Gly Ser Ala Ala Glu Glu Arg -95 -90 -85	746
GCG GTG CTA GTC GTC TCC TCC CGC ACG CAG AGG AAA GAG AGC TTA TTC Ala Val Leu Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe -80 -75 -70	794
CGG GAG ATC CGC GCC CAG GCC CGC GCG CTC GGG GCC GCT CTG GCC TCA Arg Glu Ile Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser -65 -60 -55 -50	842
GAG CCG CTG CCC GAC CCA GGA ACC GGC ACC GCG TCG CCA AGG GCA GTC Glu Pro Leu Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val -45 -40 -35	890
ATT GGC GGC CGC AGA CGG AGG AGG ACG GCG TTG GCC GGG ACG CGG ACA Ile Gly Gly Arg Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr -30 -25 -20	938
GCG CAG GGC AGC GGC GGG GGC GCG GGC CGG GGC CAC GGG CGC AGG GGC Ala Gln Gly Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly -15 -10 -5	986
CGG AGC CGC TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG GAG CTC Arg Ser Arg Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu 1 5 10 15	1034

	GGC TGG GAC GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG TAC CAC	1082
	Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His	
	20 25 30	
5	TGC GAG GGC CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG CCC ACC	1130
	Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr	
	35 40 45	
	AAC CAT GCC ATC ATT CAG ACG CTG CTC AAC TCC ATG GCA CCA GAC GCG	1178
	Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala	
	50 55 60	
	GCG CCG GCC TCC TGC TGT GTG CCA GCG CGC CTC AGC CCC ATC AGC ATC	1226
	Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile	
	65 70 75	
10	CTC TAC ATC GAC GCC GCC AAC AAC GTT GTC TAC AAG CAA TAC GAG GAC	1274
	Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp	
	80 85 90 95	
	ATG GTG GTG GAG GCC TGC GGC TGC AGG TAGCGCGCGG GCCGGGGAGG	1321
	Met Val Val Glu Ala Cys Gly Cys Arg	
	100	
15	GGGCAGCCAC GCGGCCGAGG ATCC	1345

20

(2) INFORMASJON FOR SEQ ID NO: 34:

(i) SEKVENSKARAKTERTREKK:

- (A) LENGDE: 388 aminosyrer
- (B) TYPE: aminosyre
- (D) TOPOLOGI: lineær

25

(ii) MOLEKYLTYP: protein

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 34:

Pro Gly Arg Arg Arg Pro Leu Leu Trp Ala Arg Leu Ala Ala Phe Arg
 -284 -280 -275 -270
 Leu Gly Gln Arg Arg Gly Val Gly Arg Trp Leu Gln Gln Ala Trp Leu
 -265 -260 -255
 5 Pro His Arg Arg Gln Leu Gly His Leu Leu Leu Gly Gly Pro Ala Leu
 -250 -245 -240
 Thr Val Cys Arg Ile Cys Ser Tyr Thr Ala Leu Ser Leu Cys Pro Cys
 -235 -230 -225
 Arg Ser Pro Ala Asp Glu Ser Ala Ala Glu Thr Gly Gln Ser Phe Leu
 -220 -215 -210 -205
 10 Phe Asp Val Ser Ser Leu Asn Asp Ala Asp Glu Val Val Gly Ala Glu
 -200 -195 -190
 Leu Arg Val Leu Arg Arg Gly Ser Pro Glu Ser Gly Pro Gly Ser Trp
 -185 -180 -175
 Thr Ser Pro Pro Leu Leu Leu Leu Ser Thr Cys Pro Gly Ala Ala Arg
 -170 -165 -160
 15 Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val Gly Gln
 -155 -150 -145
 Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His Arg Arg
 -140 -135 -130 -125
 Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu Leu Arg Ala Val Ala
 -120 -115 -110
 Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg Leu Gly Phe Gly Trp
 -105 -100 -95
 Pro Gly Gly Gly Gly Ser Ala Ala Glu Glu Arg Ala Val Leu Val Val
 -90 -85 -80
 Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile Arg Ala
 -75 -70 -65
 Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser Glu Pro Leu Pro Asp
 -60 -55 -50 -45
 Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val Ile Gly Gly Arg Arg
 -40 -35 -30
 Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr Ala Gln Gly Ser Gly
 -25 -20 -15
 Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg Cys Ser
 -10 -5 1
 Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp Asp Trp
 5 10 15 20
 Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Leu Cys
 25 30 35
 Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile
 40 45 50
 Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala Ser Cys
 55 60 65
 Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile Asp Ala
 70 75 80
 Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ala
 85 90 95 100
 Cys Gly Cys Arg

(2) INFORMASJON FOR SEQ ID NO: 35:

(i) SEKVENNS KARAKTERTREKK:

(A) LENGDE: 17 basepar

5 (B) TYPE: nukleinsyre

(C) TRÅDTYPE: enkel

(D) TOPOLOGI: lineær

(ii) MOLEKYLTTYPE: DNA (genomisk)

10

(iii) HYPOTETISK: nei

(iv) ANTI-SENS: nei

15 (vi) OPPRINNELIG KILDE:

(C) INDIVIDUELT ISOLAT: primer nr. 8

(xi) SEKVENSBESKRIVELSE: SEQ ID NO: 35:

TGTATGCGAC TTCCCGC

20

P a t . e n t k r a v

1.

DNA-molekyl, k a r a k t e r i s e r t v e d at det har en sekvens som koder for et benmorfogenisk protein med evne til å indusere dannelse av sene/ligamentlignende vev, idet DNA-molekylet omfatter:

- (a) nukleotid nr 196, nr 199, nr 208, nr 217, nr 361, nr 388, nr 493, nr 496, nr 571 eller nr 577 til nr 879 eller nr 882 av SEQ ID NO: 1;
- (b) nukleotider kodende for aminosyrene nr. 25, nr 1 eller nr 3 til nr 103 eller nr 104 av SEQ ID NO: 2;
- (c) nukleotid nr 410, nr 458, nr 602, nr 605 eller nr 659 til nr 961 eller nr 964 av SEQ ID NO: 25;
- (d) nukleotider kodende for aminosyrene nr 1 eller nr 19 til nr 119 eller nr 120 av SEQ ID NO: 26;
- (e) naturlig forekommende aleliske sekvenser eller ekvivalente degenererte kodonsekvenser av hvilke som helst ifølge (a) til (d); eller
- (f) sekvenser som hybridiserer til hvilke som helst av (a) til (e) under standard hybridiseringsbetingelser, eller et komplement av disse.

2.

Kimært DNA-molekyl, k a r a k t e r i s e r t v e d at det omfatter en DNA-sekvens som koder for et propeptid fra et medlem av TGF- β superfamilien av proteiner koblet i riktig leseramme til en DNA-sekvens ifølge krav 1.

3.

Kimært DNA-molekyl ifølge krav 2, k a r a k t e r i s e r t v e d at propeptidet er fra BMP-2.

4.

Vektor, k a r a k t e r i s e r t v e d at den omfatter DNA-molekylet ifølge et hvilket som helst av kravene 1 til 3 i operativ assosiasjon med en ekspresjonskontrollsekvens derav.

5.

Vertscelle, k a r a k t e r i s e r t v e d at den er transformert med DNA-molekyler ifølge et hvilket som helst av kravene 1 til 3 eller vektoren ifølge krav 5.

6.

Fremgangsmåte for fremstilling av et BMP-12 eller BMP-12-relatert protein som har evnen til å indusere dannelse av sene/ligamentlignende vev, k a r a k -
t e r i s e r t v e d at den omfatter:

- (a) dyrking av vertscellen ifølge krav 5 og
- (b) isolering og, eventuelt, rensing av nevnte BMP-12 eller BMP-12 relaterte protein fra kulturmediet.

7.

BMP-12 eller BMP-12-relatert protein med evnen til å indusere dannelse av sene/ligamentlignende vev, k a r a k t e r i s e r t v e d at det kodes av DNA-molekylet ifølge et hvilket som helst av kravene 1 til 4 eller produseres ifølge fremgangsmåten i krav 6.

8.

Polypeptid, k a r a k t e r i s e r t v e d at det er i form av en dimer omfattende to subenheter, hver med aminosyresekvensen kodet av DNA-molekylet ifølge et hvilket som helst av kravene 1 til 3.

9.

Heterodimerisk proteinmolekyl, k a r a k t e r i s e r t v e d at det omfatter en monomer med aminosyresekvensen til polypeptidet ifølge krav 7, og en monomer som har aminosyresekvensen til et protein fra TGF- β superfamilien.

10.

Farmasøytisk sammensetning, k a r a k t e r i s e r t v e d at den omfatter polypeptidet eller proteinet ifølge et hvilket som helst av kravene 7 til 9, eventuelt blandet sammen med en farmasøytisk akseptabel bærer.

11.

Anvendelse av polypeptidet eller proteinet ifølge et hvilket som helst av kravene 7 til 9 for fremstilling av en sammensetning for indusering av sene/ligamentlignende vevsdannelse hos en pasient som trenger dette.

12.

Anvendelse av polypeptidet eller proteinet ifølge et hvilket som helst av kravene 7 til 9 for fremstilling av en sammensetning for behandling av tendinitt, eller annen sene- eller ligamentdefekt, hos en pasient som trenger dette.

FIG 1/1

SAMMENLIGNING AV HUMAN V1-1 VS. HUMAN MP-52

V1-1	Ser	Arg	Cys	Ser	Arg	Lys	Pro	Leu	His	Val	Asp	Phe	Lys	Glu	Leu
1	AGC	CGC	TGC	AGC	CGC	AAG	CCG	TTG	CAC	GTG	GAC	TTC	AAG	GAG	CTC
MP52	GCT	CGC	TGC	AGT	CGG	AAG	GCA	CTG	CAT	GTG	AAC	TTC	AAG	GAC	ATG
1	Ala	Arg	Cys	Ser	Arg	Lys	Ala	Leu	His	Val	Asn	Phe	Lys	Asp	Met
16	Gly	Trp	Asp	Asp	Trp	Ile	Ile	Ala	Pro	Leu	Asp	Tyr	Glu	Ala	Tyr
46	GGC	TGG	GAC	GAC	TGG	ATC	ATC	GCG	CCG	CTG	GAC	TAC	GAG	GCG	TAC
46	GGC	TGG	GAC	GAC	TGG	ATC	ATC	GCA	CCC	CTT	GAG	TAC	GAG	GCT	TTC
16	Gly	Trp	Asp	Asp	Trp	Ile	Ile	Ala	Pro	Leu	Glu	Tyr	Glu	Ala	Phe
31	His	Cys	Glu	Gly	Leu	Cys	Asp	Phe	Pro	Leu	Arg	Ser	His	Leu	Glu
91	CAC	TGC	GAG	GGC	CTT	TGC	GAC	TTC	CCT	TTG	CGT	TCG	CAC	CTC	GAG
91	CAC	TGC	GAG	GGG	CTG	TGC	GAG	TTC	CCA	TTG	CGC	TCC	CAC	CTG	GAG
31	His	Cys	Glu	Gly	Leu	Cys	Glu	Phe	Pro	Leu	Arg	Ser	His	Leu	Glu
46	Pro	Thr	Asn	His	Ala	Ile	Ile	Gln	Thr	Leu	Leu	Asn	Ser	Met	Ala
121	CCC	ACC	AAC	CAT	GCC	ATC	ATT	CAG	ACG	CTG	CTC	AAC	TCC	ATG	GCA
121	CCC	ACG	AAT	CAT	GCA	GTC	ATC	CAG	ACC	CTG	ATG	AAC	TCC	ATG	GAC
46	Pro	Thr	Asn	His	Ala	Val	Ile	Gln	Thr	Leu	Met	Asn	Ser	Met	Asp
61	Pro	Asp	Ala	Ala	Pro	Ala	Ser	Cys	Cys	Val	Pro	Ala	Arg	Leu	Ser
181	CCA	GAC	GCG	GCG	CCG	GCC	TCC	TGC	TGT	GTG	CCA	GCG	CGC	CTC	AGC
181	CCC	GAC	TCC	ACA	CCA	CCC	ACC	TGC	TGT	GTG	CCC	ACG	CGG	CTG	AGT
61	Pro	Glu	Ser	Thr	Pro	Pro	Thr	Cys	Cys	Val	Pro	Thr	Arg	Leu	Ser
76	Pro	Ile	Ser	Ile	Leu	Tyr	Ile	Asp	Ala	Ala	Asn	Asn	Val	Val	Tyr
226	CCC	ATC	AGC	ATC	CTC	TAC	ATC	GAC	GCC	GCC	AAC	AAC	GTT	GTC	TAC
226	CCC	ATC	AGC	ATC	CTC	TTC	ATT	GAC	TCT	GCC	AAC	AAC	GTG	GTG	TAT
76	Pro	Ile	Ser	Ile	Leu	Phe	Ile	Asp	Ser	Ala	Asn	Asn	Val	Val	Tyr
91	Lys	Gln	Tyr	Glu	Asp	Met	Val	Val	Glu	Ala	Cys	Gly	Cys	Arg	
271	AAG	CAA	TAC	GAG	GAC	ATG	GTG	GTG	GAG	GCC	TGC	GGC	TGC	AGG	
271	AAG	CAG	TAT	GAG	GAC	ATG	GTC	GTG	GAG	TGC	TGT	GGC	TGC	AGG	
91	Lys	Gln	Tyr	Glu	Asp	Met	Val	Val	Glu	Ser	Cys	Gly	Cys	Arg	

Homologi på nukleotidnivå: 249/312 = 79,8%

Homologi på aminosyrenivå: 84/104 = 80,8%

1/1