



(12) PATENT

(19) NO

(11) 327754

(13) B1

NORGE

(51) Int Cl.

C07K 14/18 (2006.01)

A61K 39/12 (2006.01)

A61P 31/12 (2006.01)

G01N 33/569 (2006.01)

C12N 15/40 (2006.01)

Patentstyret

(21)	Søknadsnr	20005608	(86)	Int.inng.dag og søknadsnr	1999.05.06 PCT/EP99/03244
(22)	Inng.dag	2000.11.07	(85)	Videreføringssdag	2000.11.07
(24)	Løpedag	1999.05.06	(30)	Prioritet	1998.05.08, EP, 98201461
(41)	Alm.tilgj	2000.11.07			
(45)	Meddelt	2009.09.14			
(73)	Innehaver	Intervet International BV, Wim de Körverstraat 35 , 5831AN BOXMEER, NL			
(72)	Oppfinner	Jonathan Weston, Belfast, Nord-Irland, GB Daniel Todd, Belfast, Nord-Irland, GB			
(74)	Fullmektig	Bryn Aarflot AS, Postboks 449 Sentrum, 0104 OSLO			

(54)	Benevnelse	Strukturelle proteiner fra fisk-pankreatisk-sykdomsvirus, nukleotidsekvens, farmasøytisk preparat, DNA-vaksine, vektorvaksine, vaksine og diagnostisk sett.			
(56)	Anførte publikasjoner	EP 0712926 A2			
(57)	Sammendrag				

Foreliggende oppfinnelse angår de strukturelle proteiner fra midlet som forårsaker pankreatisk sykdom i fisk, nukleotidsekvenser som koder for nevnte proteiner, vaksiner omfattende nevnte proteiner eller nukleotidsekvenser og diagnostiske sett omfattende nevnte proteiner eller nukleotidsekvenser.

Foreliggende oppfinnelse vedrører strukturelle proteiner fra fisk-pankreatisk-sykdomsvirus, nukleotidsekvens, farmasøytisk preparat, DNA-vaksine, vektorvaksine, vaksine og diagnostisk sett.

Foreliggende oppfinnelse angår følgelig strukturelle proteiner fra det middel som forårsaker pankreatisk sykdom hos fisk, nukleotidsekvenser som koder for nevnte proteiner, vaksiner omfattende nevnte proteiner eller nukleotidsekvenser og diagnostiske sett som inneholder nevnte proteiner eller

5 nukleotidsekvenser.

Pankreatisk sykdom (PD) er en alvorlig sykdom som rammer fisk, spesielt laksefisk så som vill atlantisk laks, regnbueørret og lignende. Sykdommen forårsaker lesjoner i bukspyttkjertelen, inkludert tap av pankreatisk eksokrint vev og fibrose og hjerte- og skjelettmuskel-myopati.

10 Utbrudd av PD ble først beskrevet i 1984 av Munro et al, i Helgoland Meeresuntersuchungen 37:571-586 (1984). PD rammer typisk post-smolt fisk i løpet av det første året etter at de er overført til sjøanlegg og er angitt å spres raskt blant oppdrettsfisk som holdes i sjøinnhegninger. Kliniske tegn omfatter døsighet med en tendens til samlig i hjørnet av innhegningen og

15 mangel på evne til å opprettholde en horisontal stilling, opphør av matinntak (anoreksi) og betydelig dødelighet (Ferguson et al, J. Fish Disease 9:95-98, 1986). Murphy et al (i J. Fish Disease 15:401-408, 1992) bekreftet disse observasjoner i en senere undersøkelse, hvor det ble funnet at hjerte- og skjelett-myopati er forverret hos fisk som lider av PD.

20 Utbrudd av PD i et fiskebruk kan forårsake at veksten blir redusert og opptil 10 prosent av overlevende fisk kan vise seg å være misdannet. Ved irske fiskebruk forårsaker PD betydelige dødelighetsgrader på 10 til 60 prosent blant den unge fisken i løpet av det første året etter at de er overført til sjøanlegg (McLoughlin, M., Fish Farmer side 19, mars/april 1995). De

25 beregnede omkostninger for den irske industrien for tap av produksjon er for tiden antatt å være omkring £25 millioner pr. år. Følgelig er det et stort behov for en vaksine for forebygging og/eller behandling av PD hos fisk.

EP-A-712926 beskriver isolering av det forårsakende middel for PD fra vev fra PD-rammet fisk og identifikasjon av viruset som et toga-lignende virus.

For å forhindre PD-infeksjoner i fisk er følgelig anvendelse av svekket eller inaktivert PD for vaksiner av fisken foreslått. En ulempe ved fremstilling av inaktiverte vaksiner fra PD-viruset beskrevet i EP-A-712926 er den langsomme vekst av viruset, spesielt i cellekulturer, som gjør fremstilling av nevnte vaksiner til en relativt ineffektiv prosess. En ytterligere ulempe med de inaktiverte vaksiner er ustabiliteten til det inaktiverede virus i nærvær av andre inaktiverede patogener, hvilket resulterer i styrketap. Fiskevaksiner blir generelt produsert som multivalente vaksiner og betydelig høyere mengder av inaktiverede virus er nødvendig i en multivalent vaksine enn det som ville være nødvendig i en monovalent vaksine for å kompensere for styrketap.

Foreliggende oppfinnelse tilveiebringer vaksiner for å forhindre infeksjon hos fisk med PD, hvor ovennevnte vanskeligheter er overvunnet. Foreliggende oppfinnelse tilveiebringer nukleotidsekvensen av 3'-delen av genomisk RNA fra laks-PD-virus (SPDV). Denne sekvens av 5179 nukleotider er vist i SEKV ID NR 1 og inneholder mange åpne leserammer (ORF): I kodestrengen koder nukleotid 2 til 1186 for et ikke-strukturelt protein, og en annen overlappende ORF, som starter fra nukleotid 997 til 5076, koder for de strukturelle proteiner. Denne ORF ble betegnet som p130. Andre ikke-bestemte ORF'er ble funnet i den kodende strengen (3447 til 3767 og 4289 til 4612) og den ikke-kodende strengen (1207 til 890 og 1232-837).

ORF fra nukleotid 2 til 1186 koder for den C-terminale del av et ikke-strukturelt protein betegnet NSP4; og dens utledede aminosyre er vist i SEKV ID NR 2.

ORF p130 omfatter nukleotidsekvensene som koder for de strukturelle proteiner i PD-viruset. De strukturelle proteiner i PD-viruset består av et basisk kapsidprotein, tre kappeproteiner betegnet som E1, E2 og E3 og et protein betegnet som 6K-proteinet. Aminosyresekvensen til hele proteinet som kodes for av p130 ORF er vist i SEKV ID NR 3. Etter prosessering blir p130 proteinet spleiset inn i kapsidproteinet (aa 76-375 av p130), E3 (aa 358-428 av p130), E2 (aa429-866 av p130), 6K (aa 867-898 av p130) og E1 (aa 899-1359 av p130).

Nukleotidsekvensen som koder for kapsidproteinet i PD-viruset er lokalisert ved nukleotid 1222 til 2067 i SEKV ID NR 1. Den tilsvarende aminosyresekvens (totalt 282 aminosyrer) er vist i SEKV ID NR 4.

5 Nukleotidsekvensen som koder for kappeproteinene E3, E2 og E1 er lokalisert ved henholdsvis nukleotid 2068-2280, 2281-3594 og 3691-5076 i nukleotidsekvensen vist i SEKV ID NR 1. De tilsvarende aminosyresekvenser i E3-, E2- og E1-proteinene er vist i henholdsvis SEKV ID NR 5, 6 og 8.

10 Nukleotidsekvensen som koder for 6K-proteinet er lokalisert ved nukleotid 3595 til 3690 i nukleotidsekvensen vist i SEKV ID NR 1 og den tilsvarende aminosyresekvens i 6K-proteinet er vist i SEKV ID NR 7. Ytterligere sekvensanalyse av det virale RNA ekstrahert fra PD-infisert bukspyttkjertellev viste eksistensen av en lenger variant av 6K-proteinet som har en lengde på 68 aminosyrer sammenlignet med 6K-proteinet med 32 aminosyrer vist i SEKV ID NR 7. Nukleotidsekvensen (SEKV ID NR 14) som
15 koder for den lenger varianten av 6K-protein har en lengde på 204 nukleotider sammenlignet med de 96 nukleotider i nukleotidsekvensen som koder for det avkortede 6K-protein . Nukleotidsekvensen som koder for den lange varianten av 6K-proteinet og aminosyresekvensen utledet derav, er vist i Figur 2 og i henholdsvis SEKV ID NR 14 og SEKV ID NR 15.

20 Kloning og karakterisering av nukleotidsekvensene ifølge foreliggende oppfinnelse gir produksjon av de strukturelle proteiner i PD-virus ved anvendelse av standard rekombinant DNA-teknologi (Sambrooke et al., Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989). Kloningsteknikker og påfølgende protein-
25 ekspresjon ved anvendelse av in vitro ekspresjonssystemer er velkjent på området. På denne måten kan rekombinante strukturelle PDV-proteiner oppnås som er hovedsakelig fri for andre PDV-proteiner. Disse isolerte strukturelle proteiner kan anvendes for å fremstille underenhetsvaksiner for å beskytte mot infeksjon av PD hos fisk. Nevnte underenhetsvaksiner kan
30 anvendes som markørvaksiner i fisk for å skille vaksiner fra feltinfeksjoner med PD. Alternativt kan nukleotidsekvensene som koder for de strukturelle proteiner i PD-viruset anvendes for å fremstille DNA-vaksiner eller vektor-

vaksiner for å beskytte mot infeksjon av fisk med PD. Nukleotidsekvensene og rekombinante PD-proteiner kan videre anvendes for diagnostiske formål, for eksempel for å påvise tilstedeværelse av PD-virus i området eller anti-PD-antistoffer i fisk. I tillegg kan de rekombinante PD-proteiner ifølge foreliggende oppfinnelse anvendes for å produsere PD-spesifikke antistoffer. Disse antistoffer kan også anvendes for diagnostiske formål så som deteksjon av PD-virus hos fisk eller i området.

Foreliggende oppfinnelse vedrører strukturelt protein fra fisk-pankreatisk-sykdomsvirus, kjennetegnet ved at proteinet omfatter en aminosyresekvens angitt i SEKV ID NR 6.

Det er videre beskrevet nukleotidsekvens, kjennetegnet ved at den koder for det strukturelle proteinet fra fisk-pankreatisk-sykdomsvirus ifølge krav 1.

Foreliggende oppfinnelse vedrører videre farmasøytisk preparat, kjennetegnet ved at det omfatter et protein eller en nukleotidsekvens ifølge krav 1 eller 2.

Det er også beskrevet protein ifølge krav 1 eller en nukleotidsekvens ifølge krav 2, kjennetegnet ved at de er for anvendelse som et medikament.

Nukleotidsekvensene ifølge foreliggende oppfinnelse kan anvendes ved fremstilling av en DNA-vaksine for å vaksinere fisk mot PD-infeksjon.

Foreliggende oppfinnelse vedrører følgelig DNA-vaksine, kjennetegnet ved at den omfatter en farmasøytisk akseptabel bærer og et DNA-plasmid hvor en nukleotidsekvens ifølge krav 2 er operabelt bundet til en transkripsjonell regulatorisk sekvens.

DNA-vaksinering refererer til utløsning av en immunrespons på ett eller flere antigener som er uttrykt in vivo fra et gen innsatt i et DNA-plasmid som blir inokulert direkte i den vaksinerte fisken.

DNA-plasmider som er egnet for anvendelse i en DNA-vaksine ifølge oppfinnelsen er konvensjonelle klonings- eller ekspresjonsplasmider for bakteriell, eukaryote og gjær-vertsceller, som mange er kommersielt tilgjengelige. Velkjente eksempler på slike plasmider er pBR322 og pcDNA3 (Invitrogen). DNA-plasmidene må kunne fremkalle proteinekspresjon av

nukleotidsekvensene. DNA-plasmidet kan omfatte én eller flere nukleotidsekvenser ifølge oppfinnelsen. I tillegg kan DNA-plasmidet omfatte andre nukleotidsekvenser så som de immunstimulerende oligonukleotider som har umetylerte CpG-dinukleotider eller nukleotidsekvenser som koder for andre antigene proteiner eller hjelpe- cytokiner.

Transkripsjonelle regulatoriske sekvenser som er egnet for anvendelse i et DNA-plasmid omfatter promotere så som (human) cytomegalovirus immediate early promoter (Seed, B. et al., *Nature* 329, 840-842, 1987; Fynan, E.F. et al., *PNAS* 90, 11478-11482, 1993; Ulmer, J.B. et al., *Science* 259, 1745-1748, 1993), Rous sarcoma virus LTR (RSV, Gorman, C.M. et al., *PNAS* 79, 6777-6781, 1982; Fynan et al., *supra*; Ulmer et al., *supra*), MPSV LTR (Stacey et al., *J. Virology* 50, 725-732, 1984), SV40 immediate early promoter (Sprague J. et al., *J. Virology* 45, 773, 1983), the metallothionein-promoter (Brinster, R.L. et al., *Nature* 296, 39-42, 1982), the major late promoter of Ad2, the β -actin-promoter (Tang et al., *Nature* 356, 152-154, 1992). De regulatoriske sekvensene kan også omfatte terminator- og polyadenyleringssekvenser. Blant sekvensene som kan anvendes er den velkjente bovine veksthormon-polyadenyleringssekvens, SV40-polyadenyleringssekvensen, den humane cytomegalovirus- (hCMV) terminator og polyadenyleringssekvenser.

DNA-plasmidet omfattende en nukleotidsekvens ifølge foreliggende oppfinnelse som er operabelt bundet til en transkripsjonell regulatorisk sekvens for anvendelse i vaksinen ifølge oppfinnelsen kan være nakent eller kan pakkes i et leveringssystem. Egnede leveringssystemer er lipide vesikler, Iscomer, dendromerer, niosomer, polysakkarid-matrikser og lignende. Også meget egnet som leveringssystem er svekkede levende bakterier så som *Salmonella*.

Foreliggende oppfinnelse vedrører videre vektorvaksine, kjennetegnet ved at den omfatter levende, svekkede bakterier eller virus som er modifisert slik at de omfatter i deres genetiske materiale én eller flere av nukleotidsekvensene ifølge krav 2.

Det er videre beskrevet vaksine, kjennetegnet ved at den omfatter det strukturelle PD-proteinet ifølge krav 1 og en farmasøytisk akseptabel bærer.

Foreliggende oppfinnelse vedrører også diagnostisk sett.

Nukleotidsekvensene ifølge oppfinnelsen kan følgelig i tillegg
5 anvendes ved fremstilling av en vektorvaksine for å vaksinere fisk mot PD.
En vektorvaksine skal forstås å være en vaksine hvor levende, svekkede bakterier eller virus er modifisert slik at de inneholder én eller flere heterologe nukleotidsekvenser innsatt i deres genetiske materiale. Disse således betegnede vektorbakterier eller -virus kan meduttrykke de heterologe
10 proteiner kodet av de innsatte nukleotider. Meget egnet for anvendelse som en vaksinevektor er for eksempel vaccinia virus eller Semliki forest virus.

Nukleotidsekvensene ifølge oppfinnelsen kan også anvendes for rekombinant produksjon av strukturelle PD-proteiner, hovedsakelig fri for andre PD-proteiner.

15 Derivatproteiner skal forstås å være proteiner som har endringer i aminosyresekvensen(e) ifølge foreliggende oppfinnelse som ikke påvirker de antigene og/eller immunogene karakteristika til disse proteiner, dvs. disse derivatproteiner kan fortsatt fremkalle produksjon av antistoffer som gjenkjenner og (kryss)-reagerer med PD-virus og/eller fremkalle en
20 immunrespons hos fisk som beskytter mot PD-infeksjon. Antigene karakteristika skal forstås å være evnen til å fremkalle produksjon av antistoffer som gjenkjenner og (kryss)-reagerer med PD-viruset. Immunogene karakteristika skal forstås å være evnen til å fremkalle en immunrespons hos fisk som beskytter mot infeksjon med PD. Endringene som kan forekomme i
25 en sekvens ifølge foreliggende oppfinnelse kan for eksempel være et resultat av konservative aminosyresubstitusjoner, delesjoner, insersjoner, inversjoner eller addisjoner av aminosyre(r) i den totale sekvens. Aminosyre-substitusjoner som er forventet ikke å endre de immunologiske egenskaper er beskrevet. Aminosyreerstatninger mellom beslektede aminosyrer eller
30 erstatninger som ofte forekommer ved evolusjon er bl.a. Ser/Ala, Ser/Gly, Asp/Gly, Asp/Asn, Ile/Val (se Dayhof, M.D., Atlas of protein sequence and structure, Nat. Biomed. Res. Found., Washington D.C., 1978, vol. 5, suppl. 3).

Basert på denne informasjonen utviklet Lipman og Pearson en metode for rask og sensitiv proteinsammenligning (Science, 1985, vol. 227, 1435-1441) og bestemmelse av den funksjonelle similariteten mellom proteiner og peptider som har sekvenshomologi. Derivatproteinene er fortsatt i stand til å fremkalle produksjon av antistoffer som gjenkjenner og (kryss)-reagerer med PD-virus og/eller fremkalle en immunrespons hos fisken som beskytter mot PD-infeksjon. Kapsid-, E1-, E2-, E3- og 6K-proteiner avledet fra sovesyke- (SD) virus er slike derivatproteiner. Disse proteiner har en aminosyresekvens som er identisk eller nesten identisk med de til PD-viruset vist i SEKV ID NR 4 til 8 eller 15. Disse proteiner er i stand til å fremkalle antistoffer som gjenkjenner og kryssreagerer med PD-virus så vel som SD-virus. Andre derivater er proteinfragmenter som fortsatt er stand til å fremkalle produksjon av antistoffer som gjenkjenner og (kryss)-reagerer med PD-virus og/eller fremkalle en immunrespons hos fisken.

Proteinene ifølge oppfinnelsen kan fremstilles via standard rekombinante proteinkspresjonsteknikker. For dette formål blir en nukleotidsekvens som koder for ett eller flere av proteinene ifølge oppfinnelsen eller en multimer av nevnte protein, innsatt i en ekspresjonsvektor.

Egnede ekspresjonsvektorer er, blant andre, plasmider, kosmider, virus og YAC (Yeast Artificial Chromosomes) som omfatter de nødvendige kontrollregioner for replikasjon og ekspresjon. Ekspresjonsvektoren kan bringes til ekspresjon i en vertscelle. Egnede vertsceller er for eksempel bakterier, gjærceller og pattedyrceller. Slike ekspresjonsteknikker er velkjent på området (Sambrooke et al., Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989). De uttrykte proteiner kan isoleres og renses fra mediet. Ekspresjon av hele p130 ORF (nukleotidfragment 997 til 5076 i SEKV ID NR 1) kan føre til dannelse av viruslignende partikler på grunn av den spontane samling av de strukturelle proteiner.

Oppfinnelsen tilveiebringer følgelig videre en vaksine som omfatter ett eller flere av de strukturelle PD-proteiner og en farmasøytisk akseptabel

bærer. Mer spesifikt omfatter en vaksine ifølge oppfinnelsen, et E2-protein som har en aminosyresekvens vist i SEKV ID NR 6 eller et derivat derav.

Vaksiner ifølge oppfinnelsen er egnet for anvendelse som en markørvaksine for å skille mellom vaksinasjon og infeksjon av PD i området.

5 En vaksine ifølge oppfinnelsen kan fremstilles i henhold til teknikker velkjent for fagfolk på området. Generelle teknikker for fremstilling av DNA-vaksiner er omfattende beskrevet, for eksempel i EP patent 0 773 295 og US patent 5580859.

10 Vaksiner ifølge oppfinnelsen omfatter en effektiv mengde av de ovennevnte DNA-plasmider, vektorbakterier eller -virus eller proteiner og en farmasøytisk akseptabel bærer. Betegnelsen "effektiv" som anvendt her er definert som den mengde som er tilstrekkelig til å fremkalle en immunrespons hos målfisken. Mengden av plasmid, vektor eller protein vil avhenge av typen plasmid eller vektor, administreringsveien, tiden for administrering, fiskearten
15 så vel som alder, generell helse og diett.

Generelt kan en dose på 0,01 til 1000 µg protein pr. kg kroppsvekt, fortrinnsvis 0,5 til 500, mer foretrukket 0,1 til 100 µg protein anvendes. Med hensyn til DNA-vaksiner er generelt en minimumsdose på 10 pg opptil doser på 1000 µg plasmid angitt å være tilstrekkelig for en egnet ekspresjon av
20 antigenene in vivo.

Farmasøytisk akseptable bærere som er egnet for anvendelse i en vaksine ifølge oppfinnelsen er sterilt vann, saltvann, vandige buffere så som PBS og lignende. I tillegg kan en vaksine ifølge oppfinnelsen omfatte andre additiver så som tilsetningsmidler, stabiliseringsmidler, antioksydasjonsmidler
25 og annet.

Egnede tilsetningsmidler omfatter blant annet aluminiumhydroksyd, aluminiumfosfat, amfigen, tocofenoler, monofosfetyl lipid A, muramyldipeptid, oljeemulsjoner, glukaner, karbomerer, blokk-kopolymerer, cytokiner og saponiner så som Quil A. Mengden av tilsetningsmiddel som tilsettes
30 avhenger av typen tilsetningsmiddel.

Egnede stabiliseringsmidler for anvendelse i en vaksine ifølge oppfinnelsen er for eksempel karbohydrater omfattende sorbitol, mannitol,

stivelse, sukrose, dekstrin og glukose, proteiner så som albumin eller kasein og buffere så som alkaliske fosfater.

Vaksinene ifølge oppfinnelsen blir administrert til fisken via injeksjon, spray, immersjon eller peroralt. Administreringsprotokollen kan optimaliseres i henhold til standard vaksinasjonspraksis.

Nukleotidsekvensene og proteinene ifølge oppfinnelsen er også egnet for anvendelse innen diagnostikk. Nukleotidsekvensene eller fragmenter derav kan anvendes for å påvise tilstedeværelse av PD-virus i fisken. En primer som spenner over den C-terminale del av E2/6K/N-terminale del av E1 (se Fig. 2) ble anvendt i RT-PCR for med hell å påvise tilstedeværelse av PD-virus i en klinisk prøve ved PD-utbrudd. Proteinene kan anvendes for å påvise tilstedeværelse av antistoffer i fisken.

Proteinene ifølge oppfinnelsen kan i tillegg anvendes for fremstilling av antistoffer, ved anvendelse av de generelle teknikker som er tilgjengelige for fagfolk på området. Fortrinnsvis blir proteinene anvendt for å produsere spesifikke monoklonale antistoffer. De oppnådde antistoffer kan anvendes for diagnostikk, for å påvise PD-virus i området eller i fisken.

Ved et annet aspekt tilveiebringer således foreliggende oppfinnelse et diagnostisk sett omfattende én eller flere nukleotidsekvenser ifølge oppfinnelsen eller ett eller flere strukturelle proteiner ifølge oppfinnelsen eller antistoffer oppnådd med nevnte proteiner. Antistoffene kan fremstilles i henhold til standard teknikker. Prosedyrer for immunisering av dyr, f.eks. mus, med proteiner og seleksjon av hybridoma som produserer immunogenspesifikke monoklonale antistoffer er velkjent på området (se for eksempel Coligan et al. (ed.), *Current protocols in Immunology*, 1992; Kohler og Milstein, *Nature* 256:495-497, 1975; Steenbakkers et al., *Mol. Biol. Rep.* 19:125-134, 1994).

De følgende eksempler skal illustrere oppfinnelsen.

30 FORKLARING

FIGUR 1: strukturell organisering av de forskjellige klonede nukleotidsekvenser som koder for PD strukturelle proteiner.

FIGUR 2: Nukleotidsekvens fra C-terminus i E2-gen/ "lang" 6K-gen/N-terminus i E1-gen. De antatte spaltningssteder mellom E2/6K-protein og 6K/E1-protein er representert ved den vertikale linjen (|). Nukleotidsekvensen som koder for det "lange" 6K-protein er 204 nukleotider lang og koder for et protein med 68 aminosyrer. Nummereringen i parentes til høyre for sekvensen angir nukleotid- og aminosyrerester i henholdsvis 6K gen eller protein. I nukleotidstilling 44 i nukleotidsekvensen som koder for 6K-genet kan G-resten være erstattet med en A-rest, hvilket resulterer i et 6K-protein med en N-rest i aminosyrestilling 15 i aminosyresekvensen vist i figuren.

10

EKSEMPLER

Celler og virus

Isolering og kultivering av en laks-PD-virus- (SPDV) stamme ble utført generelt som beskrevet i EP-A-712926. F93125-isolat av SPDV ble dyrket i Chinook embryo- (CHSE-214) lakseceller som tidligere beskrevet (R.T. Nelson et al. (1995) Isolation of toga-like virus from farmed Atlantic salmon *Salmo salar* with pancreas disease. Diseases of Aquatic Organisms 22, s. 25-32). For virusrensingsformål ble monolagkulturer av CHSE-214 dyrket til ~80% sammenflyting i 75 cm² kolber infisert med 1 ml virus, for å gi en multiplisitet av infeksjon på ~1. Etter 1 times adsorpsjon ble ytterligere 14 ml supplert Eagle's minimale essensielle medium (MEM) innført i hver kolbe. De virusinfiserte kolber ble inkubert ved 15°C i 7 eller 8 dager, da virusfremkalt cytopatisk effekt var åpenbar og supernatanten ble oppsamlet.

25

Virusrensning

Supernatanten (typisk 500 ml fra virus-infiserte celler ble klaret ved 3000 g i 20 min. Polyetylen glykol (PEG) og NaCl ble tilsatt hvilket ga endelige konsentrasjoner på henholdsvis 6% og 2,2%. Etter inkubering natten over ved 4°C ble PEG-presipitatet oppsamlet ved sentrifugering i 2 timer ved 3000 g. Den resulterende pellet ble resuspendert i PBS (1-2 ml) og, etter klaring ved 1000 g i 5 minutter ble den urensede virussuspensjon fraksjonert ved

30

likevektsdensitetssentrifugering ved anvendelse av 11 ml gradienter (20-60% vekt/vekt i PBS) sukrose. Etter sentrifugering i 18 timer ved 75000 g ved 4°C, ble 1 ml fraksjoner oppsamlet fra bunnen av gradienten. Fraksjoner inneholdende virus ble identifisert ved immunoblotting ved anvendelse av
5 PD-spesifikt mus monoklonalt antistoff (Welsh et al., fremlagt 1999).

Produksjon av PD virus cDNA-kloner

Viralt RNA ble ekstrahert fra gradient-renset PD-virus og virusinfiserte celler ved anvendelse av RNA-isolator (Genosys) og lagret som etanol-presipitater.
10 Et cDNA-bibliotek ble laget ved random priming med RNA ekstrahert fra gradientrenset virus. Dette bibliotek besto av kloner inneholdende inserater (250-500bp) i vektoren pUC18 (Sureclone ligation Kit, Pharmacia). Kloner ble valgt tilfeldig fra biblioteket og ble etter sekvensering og analyse ved anvendelse av BLAST-program (University of Wisconsin, Genetics Computer
15 Group) kartlagt til alfavirusgenom. Sekvensene fra tre kloner, N11, N38 og N50, ble anvendt for å konstruere oligonukleotidprimere som ble anvendt i revers transkripsjon-polymerasekjedereaksjon (RT-PCR) for å amplifisere 3 overlappende fragmenter som omsluttet 5,2 kb regionen ved 3'-terminus i PD-genomet. Innføring av Not I seter i primerene lettet restriksjonsligering av to
20 av disse fragmenter i Not I setet av vektor pBluescript (Stratagene). PCR ble utført ved anvendelse av Expand Long Template PCR System (Boehringer Mannheim) ved 94°C i 30s, 60°C i 30s, 68°C i 2 min. En annen klon ble produsert ved anvendelse av 3'RACE (M.A. Frohmann et al., 1998; Rapid production of full-length cDNA's from rare transcripts using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci.USA. 85, s. 8998-9002).
25 Reaksjonen ble utført ved anvendelse av et 5'/3' RACE sett (Boehringer Mannheim) med noen modifikasjoner. Således ble RNA fra gradient-renset virus uavhengig underkastet førstestreng syntese og det resulterende cDNA ble amplifisert ved PCR ved 94°C i 30s, 60°C i 30s, 68°C i 1 min.

Sekvensering av PD-virus cDNA-kloner

Cyklus-sekvensering ble utført ved anvendelse av ABI PRISM dye terminator ready reaction sett på rensed plasmid DNA ved å følge produsentens protokoll (Perkin Elmer Cetus). Elektroferogrammer ble tolket ved anvendelse av Sequence Navigator software (Perkin Elmer Cetus). Den fullstendige nukleotidsekvens i den 3'-terminale 5,2 kb region i PD-virus RNA er presentert i SEKV ID NR1.

En RT-PCR og sekvensanalyse som anvender primere som flankerer C-terminus i E2 og N-terminus i E1 for viral RNA ekstrahert direkte fra PD-infisert bukspyttkjertelvev viste en lenger 6K-kodende nukleotidsekvens enn den vist ved nukleotider 3595-3690 i SEKV ID NR 1. Nukleinsyren som koder for fullengde 6K-protein så vel som den utledede aminosyresekvens er vist i Figur 2.

SPDV pFastBac1 og pcDNA3.1(+) konstruksjoner

Ved anvendelse av standard kloningsteknikker (Sambrooke et al., Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989) blir fire kloner som representerer den SPDV strukturelle region dannet i vektoren pFastBac1 (Gibco BRL) for ekspresjon i baculovirussystemet. Disse klonene blir også dannet i ekspresjonsvektoren pcDNA3.1 (Invitrogen) for monoklonal antistoffkarakterisering og anvendelse som en DNA-vaksine. Detaljer om hvorledes disse klonene er produsert er som følger:

Klon 1.

p130 koder for den fullstendige strukturelle genregion fra første ATG i kapsidproteinet til poly(A) trakt (3944nt). cDNA ble produsert fra viral RNA ved RT-PCR ved anvendelse av de følgende primere:

5' forlengs primer (5'130Not1): 5'-TGC ATG CGG CCG CAT GTT TCC CAT GCA ATT CAC CAA C-3' (SEKV ID NR 9)

3' invers primer (3'130Not1) (sekvens 5' til 3'): 5'-TGC ATG CGG CCG CTT
GTA TTG AAA ATT TTA AAA CCA A-3' (SEKV ID NR 10)

Disse primere inneholder et 5 nukleotidstrekk (sikrer
restriksjonsenzymgjenkjennelse) fulgt av et *Not1* sete og deretter den
5 passende SPDV-sekvens (uthevet i den vedlagte sekvens, fra 1222 til 1245 i
5'130Not1 og fra 5143 til 5166 i 3'130Not1). 3944nt cDNA-produktet ble
klonet inn i *Not1*-setet i både pFastBac1 og pcDNA3,1.

Klon 2.

10 p98 koder for E3, E2, 6K og E1 til poly(A) trakt (3098nt). cDNA ble
produsert fra viral RNA ved RT-PCR ved anvendelse av de følgende primere:
5' forlengs primer (5'E3Not1): 5'-TGC ATG CGG CCG CAT GAC ACG CGC
TCC GGC CCT CCT G-3' (SEKV ID NR 11)

3' invers primer (3'130Not1): 5'-TGC ATG CGG CCG CTT GTA TTG AAA
15 ATT TTA AAA CCA A -3' (SEKV ID NR 10)

Primeren 5'E3Not1 inneholder et 5 nukleotidstrekk (sikrer
restriksjonsenzymgjenkjennelse) fulgt av et *Not1* sete, en ATG (kunstig
startkodon) og deretter den passende SPDV-sekvens (fra 2067 til 2088).
Primeren 3'130Not1 er som beskrevet ovenfor under Klon 1. 3098nt cDNA-
20 produktet ble klonet inn i *Not1*-setet i både pFastBac1 og pcDNA3.1.

Klon 3.

pE2 som koder for E3- og E2-glykoproteiner (1527nt). cDNA ble
produsert fra viral RNA ved RT-PCR ved anvendelse av de følgende primere:
25 5' forlengs primer (5'E3Not1): 5'-TGC ATG CGG CCG CAT GAC ACG CGC
TCC GGC CCT CCT G -3'(SEKV ID NR 11)

3' invers primer (3'E2Not1): 5'-TGC ATG CGG CCG CTC ACG CGC GAG
CCC CTG GTA TGC AAC A-3' (SEKV ID NR 12)

Primeren 5'E3Not1 er som beskrevet ovenfor under Klon 2. Primeren
30 3'E2Not1 inneholder et 5 nukleotidstrekk (sikrer
restriksjonsenzymgjenkjennelse) fulgt av et *Not1* sete, TGA (kunstig
stoppkodon) og deretter den passende SPDV-sekvens (uthevet i den

vedlagte sekvens, fra 3571 til 3594). 1527nt cDNA-produktet ble klonet inn i *NotI* setet i både pFastBac1 og pcDNA3.1.

Klon 4.

5 E2 som koder for E2-glykoproteinet (1314nt). cDNA ble produsert fra viral RNA ved RT-PCR ved anvendelse av de følgende primere:

5' forlengs primer (5'E2Not1): 5'-TGC ATG CCG CCG CAT GGC TGT GTC TAC GTC GCC TGC C-3' (SEKV ID NR 13)

3' invers primer (3'E2Not1) : 5'-TGC ATG CCG CCG CTC ACG CGC GAG
10 CCC CTG GTA TGC AAC A-3' (SEKV ID NR 12).

Primeren 5'E2Not1 inneholder et 5 nukleotidstrekk (sikrer restriksjonsenzymgjenkjennelse) fulgt av et *NotI* sete, en ATG (kunstig startkodon) og deretter den passende SPDV-sekvens (fra 2281 til 2301).

15 Primeren 3'E2Not1 er som beskrevet ovenfor under Klon 3. 1314nt cDNA-produktet ble klonet inn i *NotI*-setet i både pFastBac1 og pcDNA3.1.

Insektceller (SF-9) ble infisert med de fire rekombinante baculovirus-konstruksjoner. Ved anvendelse av monoklonale elementer som var dannet mot fullstendig inaktiverede PD-virus, ble IFT-merking utført på disse rekombinante baculovirusinfiserte SF-9-celler. Alle de produserte proteiner
20 reagerte positivt med de monoklonale, hvilket indikerer at de rekombinante proteiner har villtype-epitoper.

Utfordringsforsøk

25 Proteinene produsert av alle fire konstruksjonene ble oppsamlet ved anvendelse av Triton-ekstraksjon. Proteinene ble BPL-inaktivert for å forhindre mulig spredning av overlevende rekombinante baculovirus til omgivelsene. Proteinene ble formulert i vann-i-olje-baserte vaksinepreparater og injisert i et 0,2 ml vaksinevolum.

30 ELISA-analyse ved anvendelse av anti-PD -E2 monoklonale elementer (2D9 opptak og 7A2) viste at mengden av reaktive epitoper pr. dose rekombinant vaksine var sammenlignbar eller til og med høyere enn mengden

av epitoper funnet i en dose av den konvensjonelle inaktiverede PD-virusvaksine.

Et standardisert utfordringsforsøk utført 8 uker etter vaksinerings av atlantisk laksefisk viste at beskyttelse mot utfordring med laks PD-virus kunne oppnås med disse rekombinante underenhetsvaksiner. I forsøket ble lesjoner i bukspyttkjertel, skjelettmuskel og hjertemuskel plassert i ordensrekkefølge. Signifikante nivåer ble beregnet fra Kruskal-Wallis enveis variansanalyse (ikke-parametrisk test). Vaksinepreparatet omfattende E2- eller E2-E3-proteiner ga lignende nivåer av beskyttelse som oppnådd ved den inaktiverede PD-virusvaksine, mens vaksiner inneholdende de rekombinante proteiner som er et resultat av henholdsvis p130- og p98-konstruksjonene var mindre beskyttende enn den inaktiverede PD-virusvaksine.

Produksjon av antistoffer .

DNA-vaksinering med proteiner oppnådd fra ekspresjon av p130-nukleotidkonstruksjonen ble utført i mus for å undersøke de antigene egenskapene til de rekombinante proteiner. Etter to intramuskulære inokuleringer med p130-pcDNA3.1 rekombinante ekspresjonsplasmider (se klon 1), viste sera fra mus en antistoffreaksjon med in vitro produserte PD-virus.

Sekvensliste

- (1) GENERAL INFORMATION:
- 5 (i) APPLICANT:
 (A) NAME: Akzo Nobel NV
 (B) STREET: Velperweg 76
 (C) CITY: Arnhem
 10 (E) COUNTRY: The netherlands
 (F) POSTAL CODE (ZIP): 6824 BM
 (G) TELEPHONE: 0412 666379
 (H) TELEFAX: 0412 650592
- 15 (ii) TITLE OF INVENTION: Structural Proteins of Fish Pancreatic
 Disease Virus and Uses Thereof
- (iii) NUMBER OF SEQUENCES: 15
- 20 (iv) COMPUTER READABLE FORM:
 (A) MEDIUM TYPE: Floppy disk
 (B) COMPUTER: IBM PC compatible
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
- (2) INFORMATION FOR SEQ ID NO: 1:
- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 5179 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- 35 (ii) MOLECULE TYPE: RNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- 40 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Salmon pancreatic disease virus
 (B) STRAIN: ..
 (C) INDIVIDUAL ISOLATE: ..
 45 (F) TISSUE TYPE: ..
- (vii) IMMEDIATE SOURCE:
 (A) LIBRARY: ..
 (B) KLONE: ..
- 50 (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 2..1186
 (D) OTHER INFORMATION: /partial
 55 /product= "NSP4 (C-terminal region)"
- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 997..5076
 60 (D) OTHER INFORMATION: /product= "p130"
- (ix) FEATURE:
 (A) NAME/KEY: mat_peptide

	GGC	TGC	CGC	CAC	ACG	CTA	TGG	CGT	GAA	CCA	CCT	GCC	GCT	GGC	CAC	AAT	1110
	Gly	Cys	Arg	His	Thr	Leu	Trp	Arg	Glu	Pro	Pro	Ala	Ala	Gly	His	Asn	
5			25					30					35				
	GGC	GAT	GGC	CAC	GCT	CGC	CCA	GGA	CTT	GAG	ATC	GTA	CCT	GGG	CGC	GCG	1158
	Gly	Asp	Gly	His	Ala	Arg	Pro	Gly	Leu	Glu	Ile	Val	Pro	Gly	Arg	Ala	
		40					45					50					
10	AGG	GGA	GTA	CGT	ATC	CCT	CTA	CGT	CTA	ACC	TTA	ATA	TTT	TCT	GCA	TCA	1206
	Arg	Gly	Val	Arg	Ile	Pro	Leu	Arg	Leu	Thr	Leu	Ile	Phe	Ser	Ala	Ser	
		55				60					65					70	
15	TAC	TTC	CAA	ACA	ATC	ATG	TTT	CCC	ATG	CAA	TTC	ACC	AAC	TCA	GCC	TAT	1254
	Tyr	Phe	Gln	Thr	Ile	Met	Phe	Pro	Met	Gln	Phe	Thr	Asn	Ser	Ala	Tyr	
				75						80					85		
20	CGC	CAG	ATG	GAG	CCC	ATG	TTT	GCA	CCG	GGT	TCC	CGA	GGA	CAA	GTA	CAG	1302
	Arg	Gln	Met	Glu	Pro	Met	Phe	Ala	Pro	Gly	Ser	Arg	Gly	Gln	Val	Gln	
				90						95				100			
25	CCG	TAC	CGG	CCG	CGC	ACT	AAG	CGC	CGC	CAG	GAG	CCG	CAA	GTC	GGC	AAC	1350
	Pro	Tyr	Arg	Pro	Arg	Thr	Lys	Arg	Arg	Gln	Glu	Pro	Gln	Val	Gly	Asn	
			105					110					115				
30	GCC	GCC	ATT	ACT	GCC	CTC	GCG	AAC	CAG	ATG	AGT	GCG	CTC	CAG	TTG	CAG	1398
	Ala	Ala	Ile	Thr	Ala	Leu	Ala	Asn	Gln	Met	Ser	Ala	Leu	Gln	Leu	Gln	
			120				125					130					
35	GTA	GCT	GGA	CTT	GCC	GGC	CAG	GCA	AGG	GTG	GAC	CGC	CGT	GGG	CCA	AGA	1446
	Val	Ala	Gly	Leu	Ala	Gly	Gln	Ala	Arg	Val	Asp	Arg	Arg	Gly	Pro	Arg	
						140					145					150	
40	CGT	GTT	CAG	AAG	AAC	AAG	CAG	AAG	AAG	AAG	AAC	TCT	TCC	AAC	GGA	GAA	1494
	Arg	Val	Gln	Lys	Asn	Lys	Gln	Lys	Lys	Lys	Asn	Ser	Ser	Asn	Gly	Glu	
					155						160					165	
45	AAA	CCC	AAA	GAG	AAG	AAG	AAG	AAG	CAA	AAA	CAA	CAG	GAG	AAG	AAG	GGA	1542
	Lys	Pro	Lys	Glu	Lys	Lys	Lys	Lys	Gln	Lys	Gln	Gln	Glu	Lys	Lys	Gly	
				170					175						180		
50	AGC	GGT	GGC	GAA	AAA	GTC	AAG	AAG	ACT	AGG	AAC	CGA	CCC	GGG	AAG	GAG	1590
	Ser	Gly	Gly	Glu	Lys	Val	Lys	Lys	Thr	Arg	Asn	Arg	Pro	Gly	Lys	Glu	
			185					190					195				
55	GTA	AGG	ATC	TCC	GTA	AAG	TGT	GCC	CGA	CAG	AGC	ACC	TTC	CCC	GTG	TAC	1638
	Val	Arg	Ile	Ser	Val	Lys	Cys	Ala	Arg	Gln	Ser	Thr	Phe	Pro	Val	Tyr	
							205					210					
60	CAC	GAA	GGT	GCT	ATA	TCC	GGC	TAC	GCT	GTG	CTG	ATT	GGA	TCT	CGC	GTA	1686
	His	Glu	Gly	Ala	Ile	Ser	Gly	Tyr	Ala	Val	Leu	Ile	Gly	Ser	Arg	Val	
						220					225					230	
65	TTC	AAG	CCG	GCA	CAC	GTG	AAG	GGT	AAG	ATC	GAC	CAC	CCT	GAA	CTG	GCA	1734
	Phe	Lys	Pro	Ala	His	Val	Lys	Gly	Lys	Ile	Asp	His	Pro	Glu	Leu	Ala	
					235					240					245		
70	GAC	ATC	AAG	TTC	CAG	GTC	GCC	GAG	GAC	ATG	GAC	CTC	GAA	GCA	GCT	GCG	1782
	Asp	Ile	Lys	Phe	Gln	Val	Ala	Glu	Asp	Met	Asp	Leu	Glu	Ala	Ala	Ala	
				250					255					260			
75	TAC	CCG	AAG	AGC	ATG	CGA	GAC	CAA	GCG	GCT	GAA	CCA	GCG	ACC	ATG	ATG	1830
	Tyr	Pro	Lys	Ser	Met	Arg	Asp	Gln	Ala	Ala	Glu	Pro	Ala	Thr	Met	Met	
			265					270					275				
80	GAC	AGA	GTG	TAC	AAC	TGG	GAG	TAT	GGC	ACT	ATC	AGA	GTG	GAG	GAT	AAT	1878

	Asp	Arg	Val	Tyr	Asn	Trp	Glu	Tyr	Gly	Thr	Ile	Arg	Val	Glu	Asp	Asn	
	280						285					290					
5	GTC	ATA	ATC	GAC	GCA	AGC	GGT	AGG	GGC	AAG	CCG	GGT	GAC	AGT	GGC	AGG	1926
	Val	Ile	Ile	Asp	Ala	Ser	Gly	Arg	Gly	Lys	Pro	Gly	Asp	Ser	Gly	Arg	
	295					300					305				310		
10	GCC	ATC	ACC	GAC	AAC	TCG	GGA	AAG	GTT	GTT	GGT	ATT	GTC	CTC	GGA	GGA	1974
	Ala	Ile	Thr	Asp	Asn	Ser	Gly	Lys	Val	Val	Gly	Ile	Val	Leu	Gly	Gly	
					315					320					325		
15	GGA	CCC	GAT	GGC	AGG	CGC	ACA	CGC	CTC	TCC	GTG	ATA	GGT	TTC	GAC	AAG	2022
	Gly	Pro	Asp	Gly	Arg	Arg	Thr	Arg	Leu	Ser	Val	Ile	Gly	Phe	Asp	Lys	
				330					335					340			
20	AAG	ATG	AAG	GCT	AGG	GAG	ATC	GCC	TAC	AGT	GAT	GCC	ATA	CCT	TGG	ACA	2070
	Lys	Met	Lys	Ala	Arg	Glu	Ile	Ala	Tyr	Ser	Asp	Ala	Ile	Pro	Trp	Thr	
			345					350					355				
25	CGC	GCT	CCG	GCC	CTC	CTG	CTG	CTG	CCT	ATG	GTT	ATT	GTC	TGC	ACC	TAC	2118
	Arg	Ala	Pro	Ala	Leu	Leu	Leu	Leu	Pro	Met	Val	Ile	Val	Cys	Thr	Tyr	
		360				365						370					
30	AAT	TCC	AAC	ACC	TTC	GAT	TGC	TCC	AAA	CCG	TCC	TGC	CAG	GAC	TGC	TGC	2166
	Asn	Ser	Asn	Thr	Phe	Asp	Cys	Ser	Lys	Pro	Ser	Cys	Gln	Asp	Cys	Cys	
	375					380					385				390		
35	ATT	ACT	GCT	GAA	CCA	GAG	AAG	GCC	ATG	ACC	ATG	CTG	AAG	GAC	AAT	CTG	2214
	Ile	Thr	Ala	Glu	Pro	Glu	Lys	Ala	Met	Thr	Met	Leu	Lys	Asp	Asn	Leu	
					395					400					405		
40	AAC	GAC	CCG	AAC	TAC	TGG	GAC	CTA	CTC	ATT	GCT	GTC	ACC	ACC	TGT	GGC	2262
	Asn	Asp	Pro	Asn	Tyr	Trp	Asp	Leu	Leu	Ile	Ala	Val	Thr	Thr	Cys	Gly	
				410					415					420			
45	TCC	GCC	CGG	AGA	AAG	AGG	GCT	GTG	TCT	ACG	TCG	CCT	GCC	GCC	TTT	TAC	2310
	Ser	Ala	Arg	Arg	Lys	Arg	Ala	Val	Ser	Thr	Ser	Pro	Ala	Ala	Phe	Tyr	
			425					430					435				
50	GAC	ACA	CAG	ATC	CTC	GCC	GCC	CAC	GCA	GCT	GCC	TCC	CCA	TAC	AGG	GCG	2358
	Asp	Thr	Gln	Ile	Leu	Ala	Ala	His	Ala	Ala	Ala	Ser	Pro	Tyr	Arg	Ala	
		440				445						450					
55	TAC	TGC	CCC	GAT	TGT	GAC	GGA	ACA	GCG	TGT	ATC	TCG	CCG	ATA	GCC	ATC	2406
	Tyr	Cys	Pro	Asp	Cys	Asp	Gly	Thr	Ala	Cys	Ile	Ser	Pro	Ile	Ala	Ile	
	455				460						465					470	
60	GAC	GAG	GTG	GTG	AGC	AGT	GGC	AGC	GAC	CAC	GTC	CTC	CGC	ATG	CGG	GTT	2454
	Asp	Glu	Val	Val	Ser	Ser	Gly	Ser	Asp	His	Val	Leu	Arg	Met	Arg	Val	
					475				480						485		
65	GGT	TCT	CAA	TCG	GGA	GTG	ACC	GCT	AAG	GGT	GGT	GCG	GCG	GGT	GAG	ACC	2502
	Gly	Ser	Gln	Ser	Gly	Val	Thr	Ala	Lys	Gly	Gly	Ala	Ala	Gly	Glu	Thr	
				490					495					500			
70	TCT	CTG	CGA	TAC	CTG	GGA	AGG	GAC	GGG	AAG	GTT	CAC	GCC	GCA	GAC	AAC	2550
	Ser	Leu	Arg	Tyr	Leu	Gly	Arg	Asp	Gly	Lys	Val	His	Ala	Ala	Asp	Asn	
			505					510					515				
75	ACG	CGA	CTC	GTG	GTG	CGC	ACG	ACT	GCA	AAG	TGC	GAC	GTG	CTG	CAG	GCC	2598
	Thr	Arg	Leu	Val	Val	Arg	Thr	Thr	Ala	Lys	Cys	Asp	Val	Leu	Gln	Ala	
			520				525					530					
80	ACT	GGC	CAC	TAC	ATC	CTG	GCC	AAC	TGC	CCA	GTG	GGG	CAG	AGC	CTA	ACC	2646
	Thr	Gly	His	Tyr	Ile	Leu	Ala	Asn	Cys	Pro	Val	Gly	Gln	Ser	Leu	Thr	
						540					545					550	

	GTT	GCG	GCC	ACA	CTG	GAT	GGC	ACC	CGG	CAT	CAA	TGC	ACC	ACG	GTT	TTC	2694
	Val	Ala	Ala	Thr	Leu	Asp	Gly	Thr	Arg	His	Gln	Cys	Thr	Thr	Val	Phe	
5					555					560					565		
	GAA	CAC	CAA	GTA	ACG	GAG	AAG	TTC	ACC	AGA	GAA	CGC	AGC	AAG	GGC	CAC	2742
	Glu	His	Gln	Val	Thr	Glu	Lys	Phe	Thr	Arg	Glu	Arg	Ser	Lys	Gly	His	
				570					575					580			
10	CAT	CTG	TCC	GAC	ATG	ACC	AAG	AAA	TGC	ACC	AGA	TTT	TCC	ACT	ACA	CCA	2790
	His	Leu	Ser	Asp	Met	Thr	Lys	Lys	Cys	Thr	Arg	Phe	Ser	Thr	Thr	Pro	
			585					590					595				
15	AAA	AAG	TCC	GCC	CTC	TAC	CTC	GTT	GAT	GTG	TAT	GAC	GCT	CTG	CCG	ATT	2838
	Lys	Lys	Ser	Ala	Leu	Tyr	Leu	Val	Asp	Val	Tyr	Asp	Ala	Leu	Pro	Ile	
		600					605					610					
20	TCT	GTA	GAG	ATT	AGC	ACC	GTC	GTA	ACA	TGC	AGC	GAC	AGC	CAG	TGC	ACA	2886
	Ser	Val	Glu	Ile	Ser	Thr	Val	Val	Thr	Cys	Ser	Asp	Ser	Gln	Cys	Thr	
	615				620						625					630	
25	GTG	AGG	GTG	CCA	CCT	GGT	ACC	ACA	GTG	AAA	TTC	GAC	AAG	AAA	TGC	AAG	2934
	Val	Arg	Val	Pro	Pro	Gly	Thr	Thr	Val	Lys	Phe	Asp	Lys	Lys	Cys	Lys	
				635					640						645		
	AGC	GCT	GAC	TCG	GCA	ACC	GTC	ACT	TTC	ACC	AGC	GAC	TCC	CAG	ACG	TTT	2982
	Ser	Ala	Asp	Ser	Ala	Thr	Val	Thr	Phe	Thr	Ser	Asp	Ser	Gln	Thr	Phe	
			650						655					660			
30	ACG	TGT	GAG	GAG	CCA	GTC	CTA	ACG	GCT	GCC	AGT	ATC	ACC	CAG	GGC	AAG	3030
	Thr	Cys	Glu	Glu	Pro	Val	Leu	Thr	Ala	Ala	Ser	Ile	Thr	Gln	Gly	Lys	
			665				670						675				
35	CCA	CAC	CTC	AGA	TCG	GCA	ATG	TTG	CCT	AGC	GGA	GGC	AAG	GAA	GTG	AAA	3078
	Pro	His	Leu	Arg	Ser	Ala	Met	Leu	Pro	Ser	Gly	Gly	Lys	Glu	Val	Lys	
		680					685					690					
40	GCA	AGG	ATC	CCG	TTC	CCG	TTC	CCG	CCG	GAA	ACC	GCA	ACT	TGC	AGA	GTG	3126
	Ala	Arg	Ile	Pro	Phe	Pro	Phe	Pro	Pro	Glu	Thr	Ala	Thr	Cys	Arg	Val	
	695					700				705						710	
45	AGT	GTA	GCC	CCA	CTG	CCG	TCG	ATC	ACC	TAC	GAG	GAA	AGC	GAT	GTC	CTG	3174
	Ser	Val	Ala	Pro	Leu	Pro	Ser	Ile	Thr	Tyr	Glu	Glu	Ser	Asp	Val	Leu	
				715					720						725		
50	CTA	GCC	GGT	ACC	GCA	AAA	TAC	CCT	GTG	CTG	CTA	ACC	ACA	CGG	AAC	CTT	3222
	Leu	Ala	Gly	Thr	Ala	Lys	Tyr	Pro	Val	Leu	Leu	Thr	Thr	Arg	Asn	Leu	
				730					735					740			
55	GGT	TTC	CAT	AGC	AAC	GCC	ACA	TCC	GAA	TGG	ATC	CAG	GGC	AAG	TAC	CTG	3270
	Gly	Phe	His	Ser	Asn	Ala	Thr	Ser	Glu	Trp	Ile	Gln	Gly	Lys	Tyr	Leu	
			745					750					755				
60	CGC	CGC	ATC	CCG	GTC	ACG	CCT	CAA	GGG	ATC	GAG	CTA	ACA	TGG	GGA	AAC	3318
	Arg	Arg	Ile	Pro	Val	Thr	Pro	Gln	Gly	Ile	Glu	Leu	Thr	Trp	Gly	Asn	
		760					765					770					
65	AAC	GCG	CCG	ATG	CAC	TTT	TGG	TCA	TCC	GTC	AGG	TAC	GCA	TCC	GGG	GAC	3366
	Asn	Ala	Pro	Met	His	Phe	Trp	Ser	Ser	Val	Arg	Tyr	Ala	Ser	Gly	Asp	
	775					780					785					790	
70	GCT	GAT	GCG	TAC	CCC	TGG	GAA	CTT	CTG	GTG	TAC	CAC	ACC	AAG	CAC	CAT	3414
	Ala	Asp	Ala	Tyr	Pro	Trp	Glu	Leu	Leu	Val	Tyr	His	Thr	Lys	His	His	
				795					800						805		
75	CCA	GAG	TAC	GCG	TGG	GCG	TTT	GTA	GGA	GTT	GCA	TGC	GGC	CTG	CTG	GCT	3462

	Pro	Glu	Tyr	Ala	Trp	Ala	Phe	Val	Gly	Val	Ala	Cys	Gly	Leu	Leu	Ala	
				810					815					820			
5	ATC	GCA	GCG	TGC	ATG	TTT	GCG	TGC	GCA	TGC	AGC	AGG	GTG	CGG	TAC	TCT	3510
	Ile	Ala	Ala	Cys	Met	Phe	Ala	Cys	Ala	Cys	Ser	Arg	Val	Arg	Tyr	Ser	
			825					830					835				
10	CTG	GTC	GCC	AAC	ACG	TTC	AAC	TCG	AAC	CCA	CCA	CCA	TTG	ACC	GCA	CTG	3558
	Leu	Val	Ala	Asn	Thr	Phe	Asn	Ser	Asn	Pro	Pro	Pro	Leu	Thr	Ala	Leu	
		840					845					850					
15	ACT	GCA	GCA	CTG	TGT	TGC	ATA	CCA	GGG	GCT	CGC	GCG	GAC	CAA	CCC	TAC	3606
	Thr	Ala	Ala	Leu	Cys	Cys	Ile	Pro	Gly	Ala	Arg	Ala	Asp	Gln	Pro	Tyr	
		855			860						865					870	
20	TTG	GAC	ATC	ATT	GCC	TAC	TTT	TTA	GGG	GTA	AGA	GGG	TGG	TCA	GCC	CTG	3654
	Leu	Asp	Ile	Ile	Ala	Tyr	Phe	Leu	Gly	Val	Arg	Gly	Trp	Ser	Ala	Leu	
					875					880					885		
25	CTG	GTC	ATC	CTT	GCG	TAT	GTA	CAG	AGC	TGC	AAG	AGC	TAC	GAA	CAC	ACC	3702
	Leu	Val	Ile	Leu	Ala	Tyr	Val	Gln	Ser	Cys	Lys	Ser	Tyr	Glu	His	Thr	
				890						895				900			
30	GTG	GTG	GTC	CCA	ATG	GAT	CCA	AGA	GCC	CCG	TCG	TAC	GAA	GCA	GTG	ATA	3750
	Val	Val	Val	Pro	Met	Asp	Pro	Arg	Ala	Pro	Ser	Tyr	Glu	Ala	Val	Ile	
			905					910					915				
35	AAC	CGG	AAT	GGG	TAT	GAT	CCA	TTG	AAG	CTG	ACC	ATC	TCA	GTG	AAT	TTC	3798
	Asn	Arg	Asn	Gly	Tyr	Asp	Pro	Leu	Lys	Leu	Thr	Ile	Ser	Val	Asn	Phe	
		920					925					930					
40	ACC	GTC	ATC	TCA	CCA	ACT	ACG	GCT	CTG	GAA	TAT	TGG	ACC	TGC	GCA	GGA	3846
	Thr	Val	Ile	Ser	Pro	Thr	Thr	Ala	Leu	Glu	Tyr	Trp	Thr	Cys	Ala	Gly	
						940					945					950	
45	GTC	CCC	ATC	GTC	GAG	CCG	CCC	CAT	GTG	GGC	TGC	TGC	ACG	TCG	GTG	TCC	3894
	Val	Pro	Ile	Val	Glu	Pro	Pro	His	Val	Gly	Cys	Cys	Thr	Ser	Val	Ser	
					955					960					965		
50	TGC	CCC	TCT	GAC	CTC	TCT	ACG	CTG	CAT	GCG	TTT	ACT	GGC	AAA	GCT	GTC	3942
	Cys	Pro	Ser	Asp	Leu	Ser	Thr	Leu	His	Ala	Phe	Thr	Gly	Lys	Ala	Val	
				970					975					980			
55	TCC	GAC	GTG	CAC	TGC	GAT	GTG	CAC	ACA	AAC	GTG	TAC	CCC	TTG	TTG	TGG	3990
	Ser	Asp	Val	His	Cys	Asp	Val	His	Thr	Asn	Val	Tyr	Pro	Leu	Leu	Trp	
			985					990					995				
60	GGC	GCG	GCT	CAC	TGC	TTC	TGT	TCC	ACC	GAG	AAT	ACA	CAG	GTC	AGC	GCT	4038
	Gly	Ala	Ala	His	Cys	Phe	Cys	Ser	Thr	Glu	Asn	Thr	Gln	Val	Ser	Ala	
			1000				1005					1010					
65	GTG	GCA	GCC	ACC	GTT	TCT	GAG	TTC	TGT	GCC	CAG	GAC	TCA	GAG	CGT	GCC	4086
	Val	Ala	Ala	Thr	Val	Ser	Glu	Phe	Cys	Ala	Gln	Asp	Ser	Glu	Arg	Ala	
			1015			1020					1025					1030	
70	GAA	GCG	TTC	AGC	GTA	CAC	AGC	AGC	TCA	GTC	ACC	GCT	GAG	GTC	CTG	GTG	4134
	Glu	Ala	Phe	Ser	Val	His	Ser	Ser	Ser	Val	Thr	Ala	Glu	Val	Leu	Val	
					1035					1040					1045		
75	ACG	CTT	GGT	GAA	GTG	GTG	ACG	GCA	GTC	CAC	GTT	TAC	GTG	GAC	GGG	GTA	4182
	Thr	Leu	Gly	Glu	Val	Val	Thr	Ala	Val	His	Val	Tyr	Val	Asp	Gly	Val	
				1050					1055					1060			
80	ACA	TCA	GCC	AGG	GGC	ACT	GAC	CTC	AAG	ATC	GTG	GCT	GGA	CCA	ATA	ACA	4230
	Thr	Ser	Ala	Arg	Gly	Thr	Asp	Leu	Lys	Ile	Val	Ala	Gly	Pro	Ile	Thr	
			1065					1070					1075				

	ACC	GAC	TAC	TCC	CCA	TTC	GAT	CGC	AAA	GTA	GTC	CGC	ATC	GGC	GAA	GAG	4278
	Thr	Asp	Tyr	Ser	Pro	Phe	Asp	Arg	Lys	Val	Val	Arg	Ile	Gly	Glu	Glu	
		1080					1085					1090					
5	GTC	TAT	AAC	TAT	GAC	TGG	CCT	CCT	TAC	GGG	GCT	GGC	CGA	CCA	GGC	ACA	4326
	Val	Tyr	Asn	Tyr	Asp	Trp	Pro	Pro	Tyr	Gly	Ala	Gly	Arg	Pro	Gly	Thr	
	1095					1100					1105					1110	
10	TTC	GGA	GAC	ATT	CAA	GCT	AGG	TCA	ACC	AAC	TAT	GTC	AAA	CCC	AAC	GAT	4374
	Phe	Gly	Asp	Ile	Gln	Ala	Arg	Ser	Thr	Asn	Tyr	Val	Lys	Pro	Asn	Asp	
					1115					1120					1125		
15	CTG	TAT	GGG	GAC	ATC	GGA	ATT	GAA	GTA	CTG	CAG	CCG	ACT	AAC	GAC	CAC	4422
	Leu	Tyr	Gly	Asp	Ile	Gly	Ile	Glu	Val	Leu	Gln	Pro	Thr	Asn	Asp	His	
				1130					1135					1140			
20	GTA	CAT	GTG	GCT	TAC	ACG	TAT	ACG	ACC	TCT	GGG	TTA	CTG	CGT	TGG	CTG	4470
	Val	His	Val	Ala	Tyr	Thr	Tyr	Thr	Thr	Ser	Gly	Leu	Leu	Arg	Trp	Leu	
			1145					1150					1155				
25	CAG	GAC	GCT	CCG	AAA	CCA	CTC	AGT	GTC	ACA	GCA	CCG	CAC	GGT	TGT	AAG	4518
	Gln	Asp	Ala	Pro	Lys	Pro	Leu	Ser	Val	Thr	Ala	Pro	His	Gly	Cys	Lys	
		1160					1165					1170					
25	ATC	AGT	GCC	AAT	CCG	CTC	CTG	GCC	CTC	GAT	TGT	GGG	GTT	GGT	GCC	GTC	4566
	Ile	Ser	Ala	Asn	Pro	Leu	Leu	Ala	Leu	Asp	Cys	Gly	Val	Gly	Ala	Val	
	1175					1180					1185					1190	

CCC ATG TCC ATC AAC ATT CCG GAC GCG AAG TTT ACC CGC AAA TTA AAG 4614
 Pro Met Ser Ile Asn Ile Pro Asp Ala Lys Phe Thr Arg Lys Leu Lys
 1195 1200 1205
 5 GAT CCG AAA CCA TCG GCC CTG AAA TGC GTG GTG GAC AGC TGC GAG TAC 4662
 Asp Pro Lys Pro Ser Ala Leu Lys Cys Val Val Asp Ser Cys Glu Tyr
 1210 1215 1220
 10 GGG GTG GAC TAC GGG GGC GCC GCC ACG ATC ACC TAC GAG GGC CAC GAG 4710
 Gly Val Asp Tyr Gly Gly Ala Ala Thr Ile Thr Tyr Glu Gly His Glu
 1225 1230 1235
 15 GCC GGG AAG TGC GGG ATT CAT TCC CTG ACA CCA GGA GTC CCC CTG AGA 4758
 Ala Gly Lys Cys Gly Ile His Ser Leu Thr Pro Gly Val Pro Leu Arg
 1240 1245 1250
 20 ACA TCG GTG GTT GAA GTG GTT GCT GGC GCC AAT ACC GTC AAA ACG ACC 4806
 Thr Ser Val Val Glu Val Val Ala Gly Ala Asn Thr Val Lys Thr Thr
 1255 1260 1265 1270
 TTC TCC TCA CCC ACG CCC GAG GTT GCA CTC GAG GTA GAG ATC TGT TCG 4854
 Phe Ser Ser Pro Thr Pro Glu Val Ala Leu Glu Val Glu Ile Cys Ser
 1275 1280 1285
 25 GCA ATA GTG AAG TGC GCT GGT GAG TGC ACT CCA CCG AAG GAA CAT GTG 4902
 Ala Ile Val Lys Cys Ala Gly Glu Cys Thr Pro Pro Lys Glu His Val
 1290 1295 1300
 30 GTC GCA ACC AGG CCT CGC CAT GGC AGC GAC CCT GGA GGC TAC ATC TCC 4950
 Val Ala Thr Arg Pro Arg His Gly Ser Asp Pro Gly Gly Tyr Ile Ser
 1305 1310 1315
 35 GGG CCC GCA ATG CGC TGG GCC GGA GGG ATT GTA GGG ACC CTA GTG GTC 4998
 Gly Pro Ala Met Arg Trp Ala Gly Gly Ile Val Gly Thr Leu Val Val
 1320 1325 1330
 40 CTG TTC CTT ATC CTT GCC GTC ATC TAC TGC GTG GTG AAG AAG TGC CGC 5046
 Leu Phe Leu Ile Leu Ala Val Ile Tyr Cys Val Val Lys Lys Cys Arg
 1335 1340 1345 1350
 TCC AAA AGA ATC CGG ATA GTC AAG AGC TAAATTCCGG TATACAAATT 5093
 Ser Lys Arg Ile Arg Ile Val Lys Ser
 1355
 45 GCTCACTAGG AGCCCATCCG ATCCCACAGG GAGTAGGATG AGTCATCTAT TGTTTTTAAA5153
 ATTTTCAATA CAAAAA AAAA 5179

50

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

55 (A) LENGTH: 394 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

60

(ix) FEATURE:

(D) OTHER INFORMATION: /label= "NSP4"

65

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Thr Met Asp Ser Ala Ala Met Asn Val Glu Ala Phe Lys Ser Phe Ala

	1			5				10					15			
	Cys	Lys	Asp	Thr	Asp	Leu	Trp	Thr	Glu	Phe	Ala	Glu	Lys	Pro	Val	Arg
				20					25					30		
5	Leu	Ser	Pro	Gly	Gln	Ile	Glu	Glu	Tyr	Val	Phe	His	Leu	Gln	Gly	Ala
			35					40					45			
	Lys	Ala	Asn	Val	Met	His	Ser	Arg	Val	Glu	Ala	Val	Cys	Pro	Asp	Leu
10		50					55					60				
	Ser	Glu	Val	Ala	Met	Asp	Arg	Phe	Thr	Leu	Asp	Met	Lys	Arg	Asp	Val
	65				70						75				80	
	Lys	Val	Thr	Pro	Gly	Thr	Lys	His	Val	Glu	Glu	Arg	Pro	Lys	Val	Gln
15					85					90					95	
	Glu	Ile	Gln	Ala	Ala	Asp	Pro	Met	Ala	Thr	Ala	Tyr	Leu	Cys	Ala	Ile
20			100						105					110		
	His	Arg	Glu	Leu	Val	Arg	Arg	Leu	Lys	Ala	Val	Leu	Lys	Pro	Ser	Ile
			115					120					125			
	His	Val	Leu	Phe	Asp	Met	Ser	Ser	Glu	Asp	Phe	Asp	Ala	Ile	Val	Gly
25		130				135						140				
	His	Gly	Met	Lys	Leu	Gly	Asp	Lys	Val	Leu	Glu	Thr	Asp	Ile	Ser	Ser
	145					150					155				160	
	Phe	Asp	Lys	Ser	Gln	Asp	Gln	Ala	Met	Ala	Val	Thr	Ala	Leu	Met	Leu
30					165					170					175	
	Leu	Arg	Asp	Leu	Gly	Val	Glu	Glu	Asp	Leu	Leu	Thr	Leu	Ile	Glu	Ala
35				180					185					190		
	Ser	Phe	Gly	Asp	Ile	Thr	Ser	Ala	His	Leu	Pro	Thr	Gly	Thr	Arg	Phe
			195					200					205			
	Gln	Phe	Gly	Ser	Met	Met	Lys	Ser	Gly	Leu	Phe	Leu	Thr	Leu	Phe	Val
40		210					215						220			
	Asn	Thr	Leu	Leu	Asn	Ile	Thr	Ile	Ala	Ala	Arg	Val	Leu	Arg	Glu	Gln
	225					230					235				240	
	Leu	Ala	Asp	Thr	Arg	Cys	Ala	Ala	Phe	Ile	Gly	Asp	Asp	Asn	Val	Ile
45					245					250					255	
	Thr	Gly	Val	Val	Ser	Asp	Asp	Met	Met	Val	Ala	Arg	Cys	Ala	Ser	Trp
50				260					265					270		
	Leu	Asn	Met	Glu	Val	Lys	Ile	Met	Asp	Met	Glu	Ile	Gly	Asn	Met	Ser
			275					280					285			
	Pro	Tyr	Phe	Cys	Gly	Gly	Phe	Leu	Leu	Leu	Asp	Thr	Val	Thr	Gly	Thr
55		290					295					300				
	Val	Ser	Arg	Val	Ser	Asp	Pro	Val	Lys	Arg	Leu	Met	Lys	Met	Gly	Lys
	305					310					315				320	
	Pro	Ala	Leu	Asn	Asp	Pro	Glu	Thr	Asp	Val	Asp	Arg	Cys	Arg	Ala	Leu
60					325					330					335	
	Arg	Glu	Glu	Val	Glu	Ser	Trp	Tyr	Arg	Val	Gly	Ile	Gln	Trp	Pro	Leu
				340					345					350		
65																
	Gln	Val	Ala	Ala	Ala	Thr	Arg	Tyr	Gly	Val	Asn	His	Leu	Pro	Leu	Ala

	Leu	Ile	Gly	Ser	Arg	Val	Phe	Lys	Pro	Ala	His	Val	Lys	Gly	Lys	Ile
	225					230					235					240
5	Asp	His	Pro	Glu	Leu	Ala	Asp	Ile	Lys	Phe	Gln	Val	Ala	Glu	Asp	Met
					245					250					255	
	Asp	Leu	Glu	Ala	Ala	Ala	Tyr	Pro	Lys	Ser	Met	Arg	Asp	Gln	Ala	Ala
				260					265					270		
10	Glu	Pro	Ala	Thr	Met	Met	Asp	Arg	Val	Tyr	Asn	Trp	Glu	Tyr	Gly	Thr
			275					280					285			
15	Ile	Arg	Val	Glu	Asp	Asn	Val	Ile	Ile	Asp	Ala	Ser	Gly	Arg	Gly	Lys
		290					295					300				
	Pro	Gly	Asp	Ser	Gly	Arg	Ala	Ile	Thr	Asp	Asn	Ser	Gly	Lys	Val	Val
	305					310					315					320
20	Gly	Ile	Val	Leu	Gly	Gly	Gly	Pro	Asp	Gly	Arg	Arg	Thr	Arg	Leu	Ser
					325					330					335	
	Val	Ile	Gly	Phe	Asp	Lys	Lys	Met	Lys	Ala	Arg	Glu	Ile	Ala	Tyr	Ser
				340					345					350		
25	Asp	Ala	Ile	Pro	Trp	Thr	Arg	Ala	Pro	Ala	Leu	Leu	Leu	Leu	Pro	Met
			355					360					365			
30	Val	Ile	Val	Cys	Thr	Tyr	Asn	Ser	Asn	Thr	Phe	Asp	Cys	Ser	Lys	Pro
		370					375					380				
	Ser	Cys	Gln	Asp	Cys	Cys	Ile	Thr	Ala	Glu	Pro	Glu	Lys	Ala	Met	Thr
	385					390					395					400
35	Met	Leu	Lys	Asp	Asn	Leu	Asn	Asp	Pro	Asn	Tyr	Trp	Asp	Leu	Leu	Ile
					405					410					415	
	Ala	Val	Thr	Thr	Cys	Gly	Ser	Ala	Arg	Arg	Lys	Arg	Ala	Val	Ser	Thr
				420					425					430		
40	Ser	Pro	Ala	Ala	Phe	Tyr	Asp	Thr	Gln	Ile	Leu	Ala	Ala	His	Ala	Ala
			435					440					445			
45	Ala	Ser	Pro	Tyr	Arg	Ala	Tyr	Cys	Pro	Asp	Cys	Asp	Gly	Thr	Ala	Cys
	450						455					460				
	Ile	Ser	Pro	Ile	Ala	Ile	Asp	Glu	Val	Val	Ser	Ser	Gly	Ser	Asp	His
	465					470					475					480
50	Val	Leu	Arg	Met	Arg	Val	Gly	Ser	Gln	Ser	Gly	Val	Thr	Ala	Lys	Gly
					485					490					495	
	Gly	Ala	Ala	Gly	Glu	Thr	Ser	Leu	Arg	Tyr	Leu	Gly	Arg	Asp	Gly	Lys
				500					505					510		
55	Val	His	Ala	Ala	Asp	Asn	Thr	Arg	Leu	Val	Val	Arg	Thr	Thr	Ala	Lys
			515					520					525			
60	Cys	Asp	Val	Leu	Gln	Ala	Thr	Gly	His	Tyr	Ile	Leu	Ala	Asn	Cys	Pro
	530							535					540			
	Val	Gly	Gln	Ser	Leu	Thr	Val	Ala	Ala	Thr	Leu	Asp	Gly	Thr	Arg	His
	545					550					555					560
65	Gln	Cys	Thr	Thr	Val	Phe	Glu	His	Gln	Val	Thr	Glu	Lys	Phe	Thr	Arg
					565					570					575	

	Glu	Arg	Ser	Lys	Gly	His	His	Leu	Ser	Asp	Met	Thr	Lys	Lys	Cys	Thr	
				580					585					590			
5	Arg	Phe	Ser	Thr	Thr	Pro	Lys	Lys	Ser	Ala	Leu	Tyr	Leu	Val	Asp	Val	
			595					600					605				
	Tyr	Asp	Ala	Leu	Pro	Ile	Ser	Val	Glu	Ile	Ser	Thr	Val	Val	Thr	Cys	
		610					615					620					
10	Ser	Asp	Ser	Gln	Cys	Thr	Val	Arg	Val	Pro	Pro	Gly	Thr	Thr	Val	Lys	
		625				630					635					640	
	Phe	Asp	Lys	Lys	Cys	Lys	Ser	Ala	Asp	Ser	Ala	Thr	Val	Thr	Phe	Thr	
15					645					650					655		
	Ser	Asp	Ser	Gln	Thr	Phe	Thr	Cys	Glu	Glu	Pro	Val	Leu	Thr	Ala	Ala	
				660					665					670			
20	Ser	Ile	Thr	Gln	Gly	Lys	Pro	His	Leu	Arg	Ser	Ala	Met	Leu	Pro	Ser	
			675					680					685				
	Gly	Gly	Lys	Glu	Val	Lys	Ala	Arg	Ile	Pro	Phe	Pro	Phe	Pro	Pro	Glu	
25		690					695				700						
	Thr	Ala	Thr	Cys	Arg	Val	Ser	Val	Ala	Pro	Leu	Pro	Ser	Ile	Thr	Tyr	
		705				710					715					720	
	Glu	Glu	Ser	Asp	Val	Leu	Leu	Ala	Gly	Thr	Ala	Lys	Tyr	Pro	Val	Leu	
30					725					730					735		
	Leu	Thr	Thr	Arg	Asn	Leu	Gly	Phe	His	Ser	Asn	Ala	Thr	Ser	Glu	Trp	
				740					745					750			
35	Ile	Gln	Gly	Lys	Tyr	Leu	Arg	Arg	Ile	Pro	Val	Thr	Pro	Gln	Gly	Ile	
			755					760					765				
	Glu	Leu	Thr	Trp	Gly	Asn	Asn	Ala	Pro	Met	His	Phe	Trp	Ser	Ser	Val	
40		770					775					780					
	Arg	Tyr	Ala	Ser	Gly	Asp	Ala	Asp	Ala	Tyr	Pro	Trp	Glu	Leu	Leu	Val	
		785				790					795					800	
	Tyr	His	Thr	Lys	His	His	Pro	Glu	Tyr	Ala	Trp	Ala	Phe	Val	Gly	Val	
45					805					810					815		
	Ala	Cys	Gly	Leu	Leu	Ala	Ile	Ala	Ala	Cys	Met	Phe	Ala	Cys	Ala	Cys	
				820					825					830			
50	Ser	Arg	Val	Arg	Tyr	Ser	Leu	Val	Ala	Asn	Thr	Phe	Asn	Ser	Asn	Pro	
			835					840					845				
	Pro	Pro	Leu	Thr	Ala	Leu	Thr	Ala	Ala	Leu	Cys	Cys	Ile	Pro	Gly	Ala	
55							855					860					
	Arg	Ala	Asp	Gln	Pro	Tyr	Leu	Asp	Ile	Ile	Ala	Tyr	Phe	Leu	Gly	Val	
		865				870					875					880	
	Arg	Gly	Trp	Ser	Ala	Leu	Leu	Val	Ile	Leu	Ala	Tyr	Val	Gln	Ser	Cys	
60					885					890					895		
	Lys	Ser	Tyr	Glu	His	Thr	Val	Val	Val	Pro	Met	Asp	Pro	Arg	Ala	Pro	
				900					905					910			
65	Ser	Tyr	Glu	Ala	Val	Ile	Asn	Arg	Asn	Gly	Tyr	Asp	Pro	Leu	Lys	Leu	
			915					920					925				

Thr Ile Ser Val Asn Phe Thr Val Ile Ser Pro Thr Thr Ala Leu Glu
 930 935 940
 5 Tyr Trp Thr Cys Ala Gly Val Pro Ile Val Glu Pro Pro His Val Gly
 945 950 955 960
 Cys Cys Thr Ser Val Ser Cys Pro Ser Asp Leu Ser Thr Leu His Ala
 965 970 975
 10 Phe Thr Gly Lys Ala Val Ser Asp Val His Cys Asp Val His Thr Asn
 980 985 990
 Val Tyr Pro Leu Leu Trp Gly Ala Ala His Cys Phe Cys Ser Thr Glu
 995 1000 1005
 15 Asn Thr Gln Val Ser Ala Val Ala Ala Thr Val Ser Glu Phe Cys Ala
 1010 1015 1020
 20 Gln Asp Ser Glu Arg Ala Glu Ala Phe Ser Val His Ser Ser Ser Val
 1025 1030 1035 1040
 Thr Ala Glu Val Leu Val Thr Leu Gly Glu Val Val Thr Ala Val His
 1045 1050 1055
 25 Val Tyr Val Asp Gly Val Thr Ser Ala Arg Gly Thr Asp Leu Lys Ile
 1060 1065 1070
 Val Ala Gly Pro Ile Thr Thr Asp Tyr Ser Pro Phe Asp Arg Lys Val
 1075 1080 1085
 30 Val Arg Ile Gly Glu Glu Val Tyr Asn Tyr Asp Trp Pro Pro Tyr Gly
 1090 1095 1100
 35 Ala Gly Arg Pro Gly Thr Phe Gly Asp Ile Gln Ala Arg Ser Thr Asn
 1105 1110 1115 1120
 Tyr Val Lys Pro Asn Asp Leu Tyr Gly Asp Ile Gly Ile Glu Val Leu
 1125 1130 1135
 40 Gln Pro Thr Asn Asp His Val His Val Ala Tyr Thr Tyr Thr Ser
 1140 1145 1150
 Gly Leu Leu Arg Trp Leu Gln Asp Ala Pro Lys Pro Leu Ser Val Thr
 1155 1160 1165
 45 Ala Pro His Gly Cys Lys Ile Ser Ala Asn Pro Leu Leu Ala Leu Asp
 1170 1175 1180
 50 Cys Gly Val Gly Ala Val Pro Met Ser Ile Asn Ile Pro Asp Ala Lys
 1185 1190 1195 1200
 Phe Thr Arg Lys Leu Lys Asp Pro Lys Pro Ser Ala Leu Lys Cys Val
 1205 1210 1215
 55 Val Asp Ser Cys Glu Tyr Gly Val Asp Tyr Gly Gly Ala Ala Thr Ile
 1220 1225 1230
 Thr Tyr Glu Gly His Glu Ala Gly Lys Cys Gly Ile His Ser Leu Thr
 1235 1240 1245
 60 Pro Gly Val Pro Leu Arg Thr Ser Val Val Glu Val Val Ala Gly Ala
 1250 1255 1260
 65 Asn Thr Val Lys Thr Thr Phe Ser Ser Pro Thr Pro Glu Val Ala Leu
 1265 1270 1275 1280

Glu Val Glu Ile Cys Ser Ala Ile Val Lys Cys Ala Gly Glu Cys Thr
 1285 1290 1295
 5 Pro Pro Lys Glu His Val Val Ala Thr Arg Pro Arg His Gly Ser Asp
 1300 1305 1310
 Pro Gly Gly Tyr Ile Ser Gly Pro Ala Met Arg Trp Ala Gly Gly Ile
 1315 1320 1325
 10 Val Gly Thr Leu Val Val Leu Phe Leu Ile Leu Ala Val Ile Tyr Cys
 1330 1335 1340
 15 Val Val Lys Lys Cys Arg Ser Lys Arg Ile Arg Ile Val Lys Ser
 1345 1350 1355

20 (2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 282 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
 (D) OTHER INFORMATION: / label= "Capsid"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

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

Val Ala Glu Asp Met Asp Leu Glu Ala Ala Ala Tyr Pro Lys Ser Met
 180 185 190
 5 Arg Asp Gln Ala Ala Glu Pro Ala Thr Met Met Asp Arg Val Tyr Asn
 195 200 205
 Trp Glu Tyr Gly Thr Ile Arg Val Glu Asp Asn Val Ile Ile Asp Ala
 210 215 220
 10 Ser Gly Arg Gly Lys Pro Gly Asp Ser Gly Arg Ala Ile Thr Asp Asn
 225 230 235 240
 Ser Gly Lys Val Val Gly Ile Val Leu Gly Gly Gly Pro Asp Gly Arg
 15 245 250 255
 Arg Thr Arg Leu Ser Val Ile Gly Phe Asp Lys Lys Met Lys Ala Arg
 260 265 270
 20 Glu Ile Ala Tyr Ser Asp Ala Ile Pro Trp
 275 280

(2) INFORMATION FOR SEQ ID NO: 5:

25

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 71 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(D) OTHER INFORMATION: / label= "E3"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

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

55 (2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 438 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

60

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(D) OTHER INFORMATION: / label= "E2"

65

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

5	Ala	Val	Ser	Thr	Ser	Pro	Ala	Ala	Phe	Tyr	Asp	Thr	Gln	Ile	Leu	Ala	1	5	10	15
	Ala	His	Ala	Ala	Ala	Ser	Pro	Tyr	Arg	Ala	Tyr	Cys	Pro	Asp	Cys	Asp	20	25	30	
10	Gly	Thr	Ala	Cys	Ile	Ser	Pro	Ile	Ala	Ile	Asp	Glu	Val	Val	Ser	Ser	35	40	45	
	Gly	Ser	Asp	His	Val	Leu	Arg	Met	Arg	Val	Gly	Ser	Gln	Ser	Gly	Val	50	55	60	
15	Thr	Ala	Lys	Gly	Gly	Ala	Ala	Gly	Glu	Thr	Ser	Leu	Arg	Tyr	Leu	Gly	65	70	75	80
	Arg	Asp	Gly	Lys	Val	His	Ala	Ala	Asp	Asn	Thr	Arg	Leu	Val	Val	Arg	85	90	95	
20	Thr	Thr	Ala	Lys	Cys	Asp	Val	Leu	Gln	Ala	Thr	Gly	His	Tyr	Ile	Leu	100	105	110	
25	Ala	Asn	Cys	Pro	Val	Gly	Gln	Ser	Leu	Thr	Val	Ala	Ala	Thr	Leu	Asp	115	120	125	
	Gly	Thr	Arg	His	Gln	Cys	Thr	Thr	Val	Phe	Glu	His	Gln	Val	Thr	Glu	130	135	140	
30	Lys	Phe	Thr	Arg	Glu	Arg	Ser	Lys	Gly	His	His	Leu	Ser	Asp	Met	Thr	145	150	155	160
	Lys	Lys	Cys	Thr	Arg	Phe	Ser	Thr	Thr	Pro	Lys	Lys	Ser	Ala	Leu	Tyr	165	170	175	
35	Leu	Val	Asp	Val	Tyr	Asp	Ala	Leu	Pro	Ile	Ser	Val	Glu	Ile	Ser	Thr	180	185	190	
40	Val	Val	Thr	Cys	Ser	Asp	Ser	Gln	Cys	Thr	Val	Arg	Val	Pro	Pro	Gly	195	200	205	
	Thr	Thr	Val	Lys	Phe	Asp	Lys	Lys	Cys	Lys	Ser	Ala	Asp	Ser	Ala	Thr	210	215	220	
45	Val	Thr	Phe	Thr	Ser	Asp	Ser	Gln	Thr	Phe	Thr	Cys	Glu	Glu	Pro	Val	225	230	235	240
	Leu	Thr	Ala	Ala	Ser	Ile	Thr	Gln	Gly	Lys	Pro	His	Leu	Arg	Ser	Ala	245	250	255	
50	Met	Leu	Pro	Ser	Gly	Gly	Lys	Glu	Val	Lys	Ala	Arg	Ile	Pro	Phe	Pro	260	265	270	
55	Phe	Pro	Pro	Glu	Thr	Ala	Thr	Cys	Arg	Val	Ser	Val	Ala	Pro	Leu	Pro	275	280	285	
	Ser	Ile	Thr	Tyr	Glu	Glu	Ser	Asp	Val	Leu	Leu	Ala	Gly	Thr	Ala	Lys	290	295	300	
60	Tyr	Pro	Val	Leu	Leu	Thr	Thr	Arg	Asn	Leu	Gly	Phe	His	Ser	Asn	Ala	305	310	315	320
	Thr	Ser	Glu	Trp	Ile	Gln	Gly	Lys	Tyr	Leu	Arg	Arg	Ile	Pro	Val	Thr	325	330	335	
65																				

Pro Gln Gly Ile Glu Leu Thr Trp Gly Asn Asn Ala Pro Met His Phe
 340 345 350
 5 Trp Ser Ser Val Arg Tyr Ala Ser Gly Asp Ala Asp Ala Tyr Pro Trp
 355 360 365
 Glu Leu Leu Val Tyr His Thr Lys His His Pro Glu Tyr Ala Trp Ala
 370 375 380
 10 Phe Val Gly Val Ala Cys Gly Leu Leu Ala Ile Ala Ala Cys Met Phe
 385 390 395 400
 Ala Cys Ala Cys Ser Arg Val Arg Tyr Ser Leu Val Ala Asn Thr Phe
 405 410 415
 15 Asn Ser Asn Pro Pro Pro Leu Thr Ala Leu Thr Ala Ala Leu Cys Cys
 420 425 430
 20 Ile Pro Gly Ala Arg Ala
 435

(2) INFORMATION FOR SEQ ID NO: 7:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 30 (ii) MOLECULE TYPE: protein
 (ix) FEATURE:
 (D) OTHER INFORMATION: / label= "6K"
 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Asp Gln Pro Tyr Leu Asp Ile Ile Ala Tyr Phe Leu Gly Val Arg Gly
 1 5 10 15
 40 Trp Ser Ala Leu Leu Val Ile Leu Ala Tyr Val Gln Ser Cys Lys Ser
 20 25 30

45 (2) INFORMATION FOR SEQ ID NO: 8:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 461 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: protein
 (ix) FEATURE:
 (D) OTHER INFORMATION: / label= "E1"
 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Tyr Glu His Thr Val Val Val Pro Met Asp Pro Arg Ala Pro Ser Tyr
 1 5 10 15
 65 Glu Ala Val Ile Asn Arg Asn Gly Tyr Asp Pro Leu Lys Leu Thr Ile
 20 25 30

	Ser	Val	Asn	Phe	Thr	Val	Ile	Ser	Pro	Thr	Thr	Ala	Leu	Glu	Tyr	Trp
			35					40					45			
5	Thr	Cys	Ala	Gly	Val	Pro	Ile	Val	Glu	Pro	Pro	His	Val	Gly	Cys	Cys
		50					55					60				
	Thr	Ser	Val	Ser	Cys	Pro	Ser	Asp	Leu	Ser	Thr	Leu	His	Ala	Phe	Thr
	65				70					75						80
10	Gly	Lys	Ala	Val	Ser	Asp	Val	His	Cys	Asp	Val	His	Thr	Asn	Val	Tyr
					85					90					95	
	Pro	Leu	Leu	Trp	Gly	Ala	Ala	His	Cys	Phe	Cys	Ser	Thr	Glu	Asn	Thr
			100						105					110		
15	Gln	Val	Ser	Ala	Val	Ala	Ala	Thr	Val	Ser	Glu	Phe	Cys	Ala	Gln	Asp
			115					120					125			
	Ser	Glu	Arg	Ala	Glu	Ala	Phe	Ser	Val	His	Ser	Ser	Ser	Val	Thr	Ala
20		130					135					140				
	Glu	Val	Leu	Val	Thr	Leu	Gly	Glu	Val	Val	Thr	Ala	Val	His	Val	Tyr
	145					150					155					160
25	Val	Asp	Gly	Val	Thr	Ser	Ala	Arg	Gly	Thr	Asp	Leu	Lys	Ile	Val	Ala
					165					170					175	
	Gly	Pro	Ile	Thr	Thr	Asp	Tyr	Ser	Pro	Phe	Asp	Arg	Lys	Val	Val	Arg
			180						185					190		
30	Ile	Gly	Glu	Glu	Val	Tyr	Asn	Tyr	Asp	Trp	Pro	Pro	Tyr	Gly	Ala	Gly
		195					200						205			
	Arg	Pro	Gly	Thr	Phe	Gly	Asp	Ile	Gln	Ala	Arg	Ser	Thr	Asn	Tyr	Val
35		210					215					220				
	Lys	Pro	Asn	Asp	Leu	Tyr	Gly	Asp	Ile	Gly	Ile	Glu	Val	Leu	Gln	Pro
	225					230					235					240
40	Thr	Asn	Asp	His	Val	His	Val	Ala	Tyr	Thr	Tyr	Thr	Thr	Ser	Gly	Leu
				245						250					255	
	Leu	Arg	Trp	Leu	Gln	Asp	Ala	Pro	Lys	Pro	Leu	Ser	Val	Thr	Ala	Pro
			260						265					270		
45	His	Gly	Cys	Lys	Ile	Ser	Ala	Asn	Pro	Leu	Leu	Ala	Leu	Asp	Cys	Gly
			275					280					285			
	Val	Gly	Ala	Val	Pro	Met	Ser	Ile	Asn	Ile	Pro	Asp	Ala	Lys	Phe	Thr
50		290					295					300				
	Arg	Lys	Leu	Lys	Asp	Pro	Lys	Pro	Ser	Ala	Leu	Lys	Cys	Val	Val	Asp
	305					310					315					320
55	Ser	Cys	Glu	Tyr	Gly	Val	Asp	Tyr	Gly	Gly	Ala	Ala	Thr	Ile	Thr	Tyr
					325					330					335	
	Glu	Gly	His	Glu	Ala	Gly	Lys	Cys	Gly	Ile	His	Ser	Leu	Thr	Pro	Gly
			340						345					350		
60	Val	Pro	Leu	Arg	Thr	Ser	Val	Val	Glu	Val	Val	Ala	Gly	Ala	Asn	Thr
			355					360					365			
	Val	Lys	Thr	Thr	Phe	Ser	Ser	Pro	Thr	Pro	Glu	Val	Ala	Leu	Glu	Val
65		370					375					380				

Glu Ile Cys Ser Ala Ile Val Lys Cys Ala Gly Glu Cys Thr Pro Pro
 385 390 395 400
 5 Lys Glu His Val Val Ala Thr Arg Pro Arg His Gly Ser Asp Pro Gly
 405 410 415
 Gly Tyr Ile Ser Gly Pro Ala Met Arg Trp Ala Gly Gly Ile Val Gly
 420 425 430
 10 Thr Leu Val Val Leu Phe Leu Ile Leu Ala Val Ile Tyr Cys Val Val
 435 440 445
 15 Lys Lys Cys Arg Ser Lys Arg Ile Arg Ile Val Lys Ser
 450 455 460

(2) INFORMATION FOR SEQ ID NO: 9:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 37 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

25 (ii) MOLECULE TYPE: cDNA

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TGCATGCGGC CGCATGTTTC CCATGCAATT CACCAAC

37

(2) INFORMATION FOR SEQ ID NO: 10:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 37 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 40 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TGCATGCGGC CGCTTGTATT GAAAATTTTA AAACCAA

37

50

(2) INFORMATION FOR SEQ ID NO: 11:

55 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 37 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11

65

TGCATGCGGC CGCATGACAC GCGCTCCGGC CCTCCTG

37

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 40 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

5

(ii) MOLECULE TYPE: cDNA

10

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

TGCATGCGGC CGCTCACGCG CGAGCCCTG GTATGCAACA 40

15

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 37 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

20

(ii) MOLECULE TYPE: cDNA

25

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TGCATGCGGC CGCATGGCTG TGTCTACGTC GCCTGCC 37

35

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 204 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

40

(ii) MOLECULE TYPE: cDNA

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GACCAACCCT ACTTGGACAT CATTGCCTAC TTGTGGACCA ACAGCAAAGT GGCCTTCGGG 60

50

CTACAATTTG CGGCGCCCGT GGCCTGTGTG CTCATCATTA CATACGCCCT TAGGCACTGC 120

55

AGATTGTGCT GCAAGTCTTT TTTAGGGGTA AGAGGGTGGT CAGCCCTGCT GGTCATCCTT 180

GCGTATGTAC AGAGCTGCAA GAGC 204

60

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 68 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

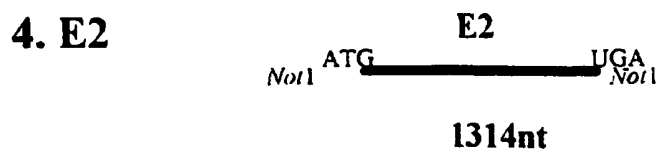
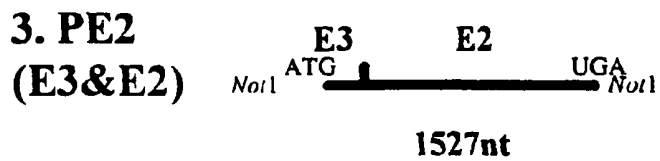
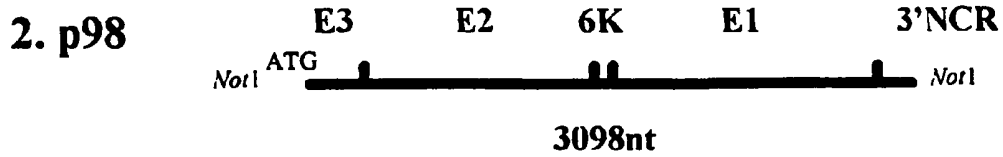
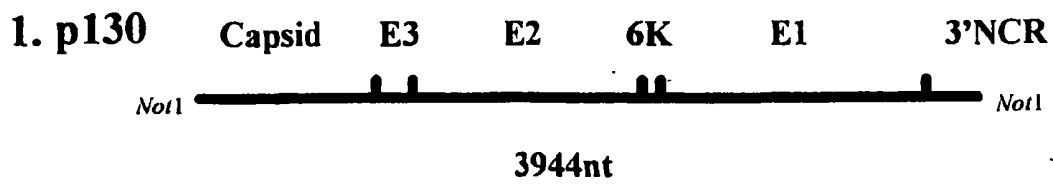
65

(ii) MOLECULE TYPE: protein

P A T E N T K R A V

1. Strukturelt protein fra fisk-pankreatisk-sykdomsvirus,
karakterisert ved at proteinet omfatter en aminosyresekvens angitt i
5 SEKV ID NR 6.
2. Nukleotidsekvens, karakterisert ved at den koder for det
strukturelle proteinet fra fisk-pankreatisk-sykdomsvirus ifølge krav 1.
- 10 3. Farmasøytisk preparat, karakterisert ved at det omfatter et
protein eller en nukleotidsekvens ifølge krav 1 eller 2.
4. Protein ifølge krav 1 eller en nukleotidsekvens ifølge krav 2,
karakterisert ved at de er for anvendelse som et medikament.
15
5. DNA-vaksine, karakterisert ved at den omfatter en
farmasøytisk akseptabel bærer og et DNA-plasmid hvor en nukleotidsekvens
ifølge krav 2 er operabelt bundet til en transkripsjonell regulatorisk sekvens.
- 20 6. Vektorvaksine, karakterisert ved at den omfatter levende,
svakkede bakterier eller virus som er modifisert slik at de omfatter i deres
genetiske materiale én eller flere av nukleotidsekvensene ifølge krav 2.
7. Vaksine, karakterisert ved at den omfatter det strukturelle
25 PD-proteinet ifølge krav 1 og en farmasøytisk akseptabel bærer.
8. Diagnostisk sett, karakterisert ved at det omfatter et protein
ifølge krav 1, eller én eller flere nukleotidsekvenser eller fragmenter derav
ifølge krav 2.

pFastBac1 SPDV Constructs



FIGUR 1

TGC AGC AGG GTG CGG TAC TCT CTG GTC GCC AAC
 C S R V R Y S L V A N

ACG TTC AAC CCG AAC CCA CCA CCA TTG ACC GCA
 T F N P N P P P L T A

E2 C-TERMINUS

CTG ACT GCA GCA CTG TGT TGC ATA CCA GGG GCT
 L T A A L C C I P G A

| 6K PROTEIN

CGC GCG **GAC CAA CCC TAC TTG GAC ATC ATT GCC** (27)
 R A D Q P Y L D I I A (9)

TAC TTG TGG ACC AAC AGC AAA GTG GCC TTC GGG (60)
Y L W T N S K V A F G (20)

CTA CAA TTT GCG GCG CCC GTG GCC TGT GTG CTC (93)
L Q F A A P V A C V L (31)

ATC ATT ACA TAC GCC CTT AGG CAC TGC AGA TTG (126)
I I T Y A L R H C R L (42)

TGC TGC AAG TCT TTT TTA GGG GTA AGA GGG TGG (159)
C C K S F L G V R G W (53)

TCA GCC CTG CTG GTC ATC CTT GCG TAT GTA CAG (192)
S A L L V I L A Y V Q (64)

| E1 N-TERMINUS

AGC TGC AAG AGC TAC GAA CAC ACC GTG GTG GTC (204)
S C K S Y E H T V V V (68)

CCA ATG GAT CCA AGA GCC CCG TCG TAC GAA GCA
 P M D P R A P S Y E A

GTG ATA AAC CGG AAT GGG TAT GAT CCA TTG AAG
 V I N R N G Y D P L K

CTG ACC ATC TCA GTG AAT TTC ACC GTC ATC TCA
 L T I S V N F T V I S

CCA ACT ACG GCT CTG GAA T 3'
 P T T A L E

FIGUR 2