(12) STANDARD PATENT (11) Application No. **AU 2014257267 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Veterinary decorin compositions and use thereof

(51) International Patent Classification(s)
C07K 14/47 (2006.01) ***C12P 21/02*** (2006.01)

(21) Application No: **2014257267** (22) Date of Filing: **2014.04.22**

(87) WIPO No: **WO14/176198**

(30) Priority Data

(31) Number	(32) Date	(33) Country
61/814,405	**2013.04.22**	**US**

(43) Publication Date: **2014.10.30**
(44) Accepted Journal Date: **2017.03.16**

(71) Applicant(s)
Catalent Pharma Solutions, LLC

(72) Inventor(s)
Bleck, Gregory T.

(74) Agent / Attorney
Spruson & Ferguson, L 35 St Martins Tower 31 Market St, Sydney, NSW, 2000, AU

(56) Related Art
US 2012/0238727 A1
WO 1993/009800 A1

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 October 2014 (30.10.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/176198 A1

(51) International Patent Classification:
C12P 21/02 (2006.01)

(21) International Application Number:
PCT/US2014/034877

(22) International Filing Date:
22 April 2014 (22.04.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/814,405 22 April 2013 (22.04.2013) US

(71) Applicant: **CATALENT PHARMA SOLUTIONS, LLC** [US/US]; 14 Schoolhouse Road, Somerset, New Jersey 08873 (US).

(72) Inventor: **BLECK, Gregory T.**; 1251 Gil's Way, Cross Plains, Wisconsin 53528 (US).

(74) Agent: **JONES, J. Mitchell**; Casimir Jones SC, 2275 Deming Way, Suite 310, Middleton, Wisconsin 53562 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

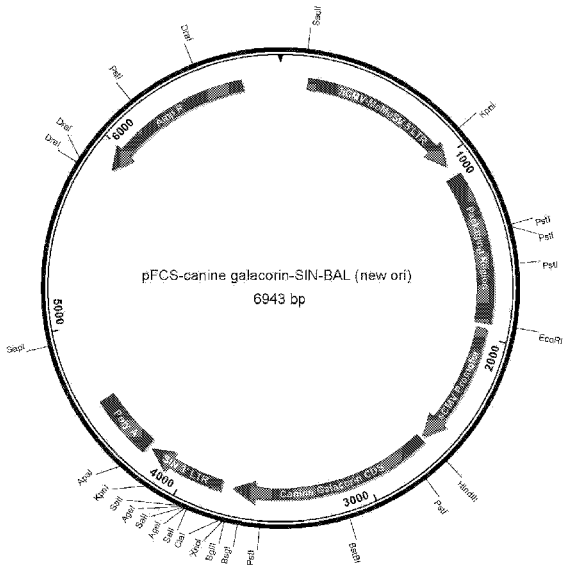
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: VETERINARY DECORIN COMPOSITIONS AND USE THEREOF

(57) Abstract: The present invention relates to veterinary decorin compositions and methods of their production.

FIG. 2



WO 2014/176198 A1

VETERINARY DECORIN COMPOSITIONS AND USE THEREOF

CROSS-REFERENCE TO RELATED APPLCIATIONS

This application claims the benefit of U.S. Prov. Appl. 61/814,405 filed April 22, 2013, the entire contents of which are incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates to veterinary decorin compositions and methods of their production and use.

BACKGROUND OF THE INVENTION

Proteoglycans carrying one or more glycosaminoglycan (GAG) chains form a large gene family that may be classifed into five groups according to the structural properties of the core protein. One of the groups is the small Leucine-rich proteoglycan family comprised of decorin (DCN), biglycan, fibromodulin and lumican. These are characterized by 40 kDa core proteins that contain Leucine-rich repeats of approximately 12 amino acids. DCN is a prototype of the group and is also referred to as PG-S2, PG40, proteodermatan sulphate and DS-PGII. It contains one dermatan chondroitin sulphate GAG chain covalently linked to a Serine of the mature core protein and is considered to be a multifunctional proteoglycan.

Decorin protein is present in most all animal tissues. The reduction or absence of decorin leads to problems within the body. The proposed functions of DCN include, but are not limited to, regulation of collagen fibrillogenesis, maintenance of tissue integrity via binding with fibronectin and thrombospondin, and a reservoir of transforming growth factor β (TGF-β). The latter function of DCN is achieved through its core protein sequestering the growth factor in the extracellular milieu from receptors expressed on the cell surface. Various conditions including all types of surgery, cuts, burns, eye injuries, spinal cord injury, head trauma, lung disease, kidney disease, liver disease and cancer can disrupt the balance of decorin in the effected tissue.

Therapeutic use of decorin in humans has been proposed. Examples of such uses include: suppression of cell proliferation by decorin (US 6,046,162); methods for inhibiting TGF-beta activity (US 6,277,812); methods of a pathology or a fibrotic condition by administering decorin (US 6,436,900); methods of preventing or reducing scarring with decorin or biglycan (US 6,509,314), treatment of glomerulonephritis with decorin (US 5,726,149); suppressing tumor cell growth by administering decorin (US 6,524,573); and

inhibiting proliferative diseases by inhibiting TGF-beta mediated angiogenesis (US 6,673,341). All of the patents referenced in this paragraph are incorporated herein by reference in their entirety.

What is needed in the art are decorin compositions for use in veterinary therapies.

SUMMARY OF THE INVENTION

The present invention relates to veterinary decorin compositions and methods of their production and use.

In a first aspect, the present invention provides an isolated veterinary decorin core protein molecule that is at least 95% identical to one of SEQ ID NOs: 4, 7, 10, 13, 16, 19, 22, and 27, with the proviso that the veterinary decorin core protein comprises a mutation at amino acid 4. In some embodiments, the mutation prevents gagylation of the molecule. In some embodiments, the mutation is a serine to alanine mutation. In some embodiments, the protein molecule is at least 98% identical to one of SEQ ID NOs: 4, 7, 10, 13, 16, 19, 22, and 27, with the proviso that the veterinary decorin core protein comprises a mutation at amino acid 4. In some embodiments, the protein molecule is at least 99% identical to one of SEQ ID NOs: 4, 7, 10, 13, 16, 19, 22, and 27, with the proviso that the veterinary decorin core protein comprises a mutation at amino acid 4. In some embodiments, the protein molecule is 100% identical to one of SEQ ID NOs: 4, 7, 10, 13, 16, 19, 22, and 27, with the proviso that the veterinary decorin core protein comprises a mutation at amino acid 4. In some embodiments, the core molecule is operably linked to an exogenous signal peptide. In some embodiments, the exogenous signal peptide is a bovine lactalbumin signal peptide.

In a second aspect, the present invention provides a composition comprising a veterinary decorin core protein molecule as described above in combination with a pharmaceutically acceptable carrier.

In some embodiments, the present invention provides a nucleic acid sequence encoding a veterinary decorin core molecule as described above. In some embodiments, the nucleic acid sequence encoding a veterinary decorin core molecule is operably associated with an exogenous signal peptide. In some embodiments, the exogenous signal peptide is a bovine lactalbumin signal peptide.

In a third aspect, the present invention provides an expression vector comprising a nucleic acid sequence encoding a veterinary decorin core molecule as described above.

In a fourth aspect, the present invention provides a host cell comprising the expression vector as described above.

In a fifth aspect, the present invention provides a method of producing a veterinary decorin protein comprising expressing the expression vector as described above in a host cell to produce the veterinary decorin protein and purifying the veterinary decorin protein.

In a sixth aspect, the present invention provides a veterinary decorin protein produced by the method described above.

In some embodiments, the present invention provides a pharmaceutical formulation comprising a veterinary decorin core protein molecule as described above, wherein the formulation is a liquid, powder, spray, gel, ointment, lotion, or eye drop.

In some embodiments, the present invention provides methods of treating a veterinary subject in need thereof comprising administering to the subject a veterinary decorin core protein according to claim 1 in an effective amount. In some embodiments, the veterinary decorin core protein is administered by a method selected from the group consisting of enteral, parenteral and topical administration. In some embodiments, the veterinary decorin core protein is administered by a method selected from the group consisting of oral administration, intravenous administration, intradermal administration, subcutaneous administration, transdermal administration, nasal administration, intramuscular administration, intrathecal administration, intraocular administration, intravitreal administration, intravaginal administration, and transmucosal administration.

In some embodiments, the veterinary subject is suffering from a wound or other injury to the skin and the veterinary decorin core protein is administered to inhibit scar formation. In some embodiments, the wound is the result of cosmetic or general surgery, injury to the skin, or injury causing proud flesh. In some embodiments, the scar is a keloid scar.

In some embodiments, the veterinary subject is suffering from an injury or disease to the eye. In some embodiments, the injury to the eye is the result of corneal surgery, an eye burn, an eye infection, and an abrasive injury.

In some embodiments, the veterinary subject is suffering from a lung disease. In some embodiments, the lung disease is selected from the group consisting of interstitial lung disease and pulmonary fibrosis. In some embodiments, the veterinary subject is suffering from kidney disease.

In some embodiments, the kidney disease is selected from the group consisting of diabetic nephropathy and renal fibrosis.

In some embodiments, the veterinary subject is suffering from liver disease. In some embodiments, the liver disease is selected from the group consisting of cirrhosis and hepatic fibrosis.

In some embodiments, the subject is suffering from a cancer. In some embodiments, cancer is an EGF Receptor or IGF-I receptor positive cancer.

In some embodiments, the veterinary subject is suffering from heart disease.

In some embodiments, the veterinary subject is suffereing from a neurological trauma. In some embodiments, the neurological trauma is selected from a brain injury and a spinal cord injury.

DESCRIPTION OF THE FIGURES

Figure 1. Canine decorin coding sequence and the flanking DNA cloning junctions in the final retrovector expression construct are shown. Kozak translation initiation sequence is highlighted. The cloning sites HindIII and XhoI are also shown.

Figure 2. Map of Canine Decorin in GPEx® Vector.

Figure 3. Feline Decorin coding sequence and the flanking DNA cloning junctions in the final retrovector expression construct are shown. The Kozak translation initiation sequence is highlighted. The cloning sites HindIII and XhoI are also shown.

Figure 4. Map of Feline Decorin in GPEx® Vector.

Figure 5. SDS-PAGE gel of media from the pooled CHO cell line expressing Canine decorin mature protein. Lane 1. Molecular weight standards. Lane 2. Non-reduced media sample. Lane 3. Reduced media sample. The decorin protein shows as a doublet in both conditions similar to the purified human decorin shown in Figure 7.

Figure 6. SDS-PAGE gel of media from the pooled CHO cell line expressing Feline decorin mature protein. Lane 1. Molecular weight standards. Lane 2. Non-reduced media sample. Lane 3. Reduced media sample. The decorin protein shows as a doublet in both conditions similar to the purified human decorin shown in Figure 7.

Figure 7. SDS-PAGE gel of purified human decorin mature protein. Lane 1. Molecular weight standards. Lane 2. 2 mg human decorin (reduced). Lane 3. 5 mg human decorin (reduced).

DEFINITIONS

To facilitate understanding of the invention, a number of terms are defined below.

As used herein, the term "veterinary decorin" refers to a decorin molecule from a companion or stock animal, e.g., poultry such as a chicken, cows, goats, sheep, pigs, horses, dogs and cats.

As used herein, the term "veterinary decorin core protein" refers to a veterinary decorin protein molecule that has a mutation at amino acid 4 of mature decorin and that substantially lacks modification with a glycosaminoglycan (GAG; i.e., is non-gagylated) at amino acid 4.

As used herein, the term “veterinary subject” encompasses stock and companion animals, including, but not limited to, cows, sheep, horses, pigs, goats, chickens, turkeys, dogs and cats.

As used herein, the term "host cell" refers to any eukaryotic cell (e.g., mammalian cells, avian cells, amphibian cells, plant cells, fish cells, and insect cells), whether located in vitro or in vivo.

As used herein, the term "cell culture" refers to any in vitro culture of cells. Included within this term are continuous cell lines (e.g., with an immortal phenotype), primary cell cultures, finite cell lines (e.g., non-transformed cells), and any other cell population maintained in vitro, including oocytes and embryos.

As used herein, the term "vector" refers to any genetic element, such as a plasmid, phage, transposon, cosmid, chromosome, virus, virion, etc., which is capable of replication when associated with the proper control elements and which can transfer gene sequences between cells. Thus, the term includes cloning and expression vehicles, as well as viral vectors.

As used herein, the terms "complementary" or "complementarity" are used in reference to polynucleotides (i.e., a sequence of nucleotides) related by the base-pairing rules. For example, the sequence "5'-A-G-T-3'," is complementary to the sequence "3'-T-C-A-5'." Complementarity may be "partial," in which only some of the nucleic acids' bases are matched according to the base pairing rules. Or, there may be "complete" or "total" complementarity between the nucleic acids. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, as well as detection methods that depend upon binding between nucleic acids.

The terms "homology" and "percent identity" when used in relation to nucleic acids refers to a degree of complementarity. There may be partial homology (i.e., partial identity) or complete homology (i.e., complete identity). A partially complementary sequence is one that at least partially inhibits a completely complementary sequence from hybridizing to a target nucleic acid sequence and is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or Northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe (i.e., an oligonucleotide which is capable of hybridizing to another oligonucleotide of interest) will compete for and inhibit the binding (i.e., the

hybridization) of a completely homologous sequence to a target sequence under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target which lacks even a partial degree of complementarity (e.g., less than about 30% identity); in the absence of non-specific binding the probe will not hybridize to the second non-complementary target.

The terms "in operable combination," "in operable order," and "operably linked" as used herein refer to the linkage of nucleic acid sequences in such a manner that a nucleic acid molecule capable of directing the transcription of a given gene and/or the synthesis of a desired protein molecule is produced. The term also refers to the linkage of amino acid sequences in such a manner so that a functional protein is produced.

As used herein, the term "signal sequence" refers to any DNA sequence which, when operably linked to a recombinant DNA sequence, encodes a signal peptide which is capable of causing the secretion of the recombinant polypeptide. In general, the signal peptides comprise a series of about 15 to 30 hydrophobic amino acid residues (See, e.g., Zwizinski et al., J. Biol. Chem. 255(16): 7973-77 [1980], Gray et al., Gene 39(2): 247-54 [1985], and Martial et al., Science 205: 602-607 [1979]). Such secretion signal sequences are preferably derived from genes encoding polypeptides secreted from the cell type targeted for tissue-specific expression (e.g., secreted milk proteins for expression in and secretion from mammary secretory cells). Secretory DNA sequences, however, are not limited to such sequences. Secretory DNA sequences from proteins secreted from many cell types and organisms may also be used (e.g., the secretion signals for t-PA, serum albumin, lactoferrin, and growth hormone, and secretion signals from microbial genes encoding secreted polypeptides such as from yeast, filamentous fungi, and bacteria).

As used herein, the term "purified" refers to molecules, either nucleic or amino acid sequences, that are removed from their normal environment, isolated or separated. An "isolated nucleic acid sequence" is therefore a purified nucleic acid sequence. "Substantially purified" molecules are at least 60% free, preferably at least 75% free, and more preferably at least 90% free from other components with which they are normally associated.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to veterinary decorin compositions and methods of their production and use.

Decorin

Native decorin is a glycoprotein with an attached glycosaminoglycan and an average molecular weight of 90-140 kD. The present invention contemplates the production and use recombinant veterinary decorin. In some preferred embodiments, the veterinary decorin is a veterinary decorin core protein, i.e., a substantially non-gagylated veterinary decorin. In some embodiments, the veterinary decorin core protein comprises a mutation at amino acid 4 (i.e., the 4th amino acid from the N-terminus) of the mature veterinary decorin core protein molecule. In some embodiments, the mutation is a serine to alanine mutation. In some embodiments, the veterinary decorin core protein molecule is a protein molecule having a sequence at least 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to one of SEQ ID NOS:4 (chicken), 7 (bovine), 10 (canine), 13 (caprine), 16 (equine), 19 (porcine), 22 (ovine), and 27 (feline) provided that that the decorin core protein comprises a mutation at amino acid 4 (i.e., the 4th amino acid from the N-terminus) of the mature decorin core protein molecule.

Decorin is commonly expressed as a pre-pro-protein. The present invention provides veterinary decorin fusion molecules comprising a heterologous signal sequence in operable association with the veterinary decorin pro-peptide and mature peptide sequences. In some embodiments, the heterologous signal polypeptide is an alpha-lactalbumin signal polypeptide. In some embodiments, the alpha-lactalbumin signal polypeptide is a bovine alpha-lactalbumin signal polypeptide. In some embodiments, the heterologous signal polypeptide is at least 80%, 90%, or 100% identical to MMSFVSLLLVGILFHATQA (SEQ ID NO:23). In some embodiments, the propeptide sequence is at least 80%, 90%, or 100% identical to the propeptide sequences identified in SEQ ID NOs:3, 6, 9, 12, 15, 18, 21 and 26. In some embodiments, the decorin core protein portion of the fusion polypeptide is at least 90%, 95% , 99% or 100% identical to SEQ ID NO:1 (mature decorin core protein), provided that that the decorin core protein comprises a mutation at amino acid 4 (i.e., the 4th amino acid from the N-terminus) of the mature decorin core protein molecule. In some embodiments, the fusion protein is 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to one of SEQ ID NOS:3 (chicken), 6 (bovine), 9 (canine), 12 (caprine), 15 (equine), 18 (porcine), 21 (ovine), and 26 (feline)(signal-propeptide veterinary decorin core protein), provided that that the decorin core protein comprises a mutation at amino acid 4 (i.e., the 4th amino acid from the N-terminus) of the mature decorin core protein molecule.

The present invention further provides nucleic acid sequences encoding the fusion proteins, as well as vectors comprising the nucleic acid sequences. In some embodiments,

the heterologous signal polypeptide is at least 80%, 90%, or 100% identical to ATGATGTCCTTTGTCTCTCTGCTCCTGGTAGGCATCCTATTCCATGCCACCCAGGCC (SEQ ID NO:24). In some embodiments, the propeptide nucleic acid sequence is at least 80%, 90%, or 100% identical to the propeptide sequences identified in SEQ ID NOs:1, 5, 8, 11, 14, 17, 20 and 25. In some embodiments, the veterinary decorin core protein portion of the fusion polypeptide is at least 90%, 95% , 99% or 100% identical to the veterinary decorin core protein portion identified in SEQ ID NOs:1, 5, 8, 11, 14, 17, 20, and 25 provided that that the decorin core protein nucleic acid sequence comprises a mutation at amino acid 4 (i.e., the 4th amino acid from the N-terminus) of the mature decorin core protein molecule. In some embodiments, the fusion protein is at least 90%, 95%, 99% or 100% identical to one of SEQ ID NOs: 1, 5, 8, 11, 14, 17, 20 and 25 (signal-propeptide veterinary decorin core protein nucleic acid sequences), provided that that the decorin core protein nucleic acid sequence comprises a mutation at amino acid 4 (i.e., the 4th amino acid from the N-terminus) of the mature decorin core protein molecule.

The veterinary decorin polynucleotides of the present invention may be employed for producing veterinary decorin polypeptides by recombinant techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing a polypeptide. In some embodiments of the present invention, vectors include, but are not limited to, retroviral vectors, chromosomal, nonchromosomal and synthetic DNA sequences (e.g., derivatives of SV40, bacterial plasmids, phage DNA; baculovirus, yeast plasmids, vectors derived from combinations of plasmids and phage DNA, and viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies). It is contemplated that any vector may be used as long as it is replicable and viable in the host. In some preferred embodiments, the vectors are retroviral vectors as described in US Pat. Nos. 6,852,510 and 7,332,333 and U.S. Pat. Publ. Nos. 200402335173 and 20030224415, all of which are incorporated herein by references in their entirety. In some especially preferred embodiments, the vectors are pseudotyped retroviral vectors.

In particular, some embodiments of the present invention provide recombinant constructs comprising one or more of the sequences as broadly described above (e.g., one of SEQ ID NOs: 1, 5, 8, 11, 14, 17, 20 and 25). In some embodiments of the present invention, the constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In still other embodiments, the heterologous structural sequence (e.g., one of SEQ ID NOs: 1, 5, 8, 11, 14, 17, 20 and 25) is assembled in appropriate phase with translation initiation and termination

sequences. In preferred embodiments of the present invention, the appropriate DNA sequence is inserted into the vector using any of a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art.

Large numbers of suitable vectors are known to those of skill in the art, and are commercially available. Such vectors include, but are not limited to, the following vectors: 1) Bacterial--pQE70, pQE60, pQE-9 (Qiagen), pBS, pD10, phagescript, psiX174, pbluescript SK, pBSKS, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); 2) Eukaryotic--pWLNEO, pSV2CAT, pOG44, PXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia); and 3) Baculovirus--pPbac and pMbac (Stratagene). Any other plasmid or vector may be used as long as they are replicable and viable in the host. In some preferred embodiments of the present invention, mammalian expression vectors comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation sites, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking non-transcribed sequences. In other embodiments, DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required non-transcribed genetic elements.

In certain embodiments of the present invention, the DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. Promoters useful in the present invention include, but are not limited to, the LTR or SV40 promoter, the E. coli lac or trp, the phage lambda P_L and P_R, T3 and T7 promoters, and the cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, and mouse metallothionein-I promoters and other promoters known to control expression of gene in prokaryotic or eukaryotic cells or their viruses. In other embodiments of the present invention, recombinant expression vectors include origins of replication and selectable markers permitting transformation of the host cell (e.g., dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or tetracycline or ampicillin resistance in E. coli).

In some embodiments of the present invention, transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Enhancers useful in the present invention include, but are not limited to, the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

In other embodiments, the expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. In still other embodiments of the present invention, the vector may also include appropriate sequences for amplifying expression.

In a further embodiment, the present invention provides host cells containing the above-described constructs. In some embodiments of the present invention, the host cell is a higher eukaryotic cell (e.g., a mammalian or insect cell). In other embodiments of the present invention, the host cell is a lower eukaryotic cell (e.g., a yeast cell). In still other embodiments of the present invention, the host cell can be a prokaryotic cell (e.g., a bacterial cell). Specific examples of host cells include, but are not limited to, *Escherichia coli*, *Salmonella typhimurium*, *Bacillus subtilis*, and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, as well as *Saccharomycees cerivisiae*, *Schizosaccharomycees pombe*, Drosophila S2 cells, *Spodoptera* Sf9 cells, Chinese hamster ovary (CHO) cells, COS-7 lines of monkey kidney fibroblasts, (Gluzman, Cell 23:175 [1981]), C127, 3T3, 293, 293T, HeLa and BHK cell lines.

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. In some embodiments, introduction of the construct into the host cell can be accomplished by retroviral transduction, calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (See e.g., Davis et al. [1986] Basic Methods in Molecular Biology). Alternatively, in some embodiments of the present invention, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y.

In some embodiments of the present invention, following transformation of a suitable host strain and growth of the host strain to an appropriate cell density in media, protein is secreted and cells are cultured for an additional period. In other embodiments of the present invention, cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. In still other embodiments of the present invention, microbial cells employed in expression of proteins can

be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

The present invention also provides methods for recovering and purifying decorin from recombinant cell cultures including, but not limited to, ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. In some embodiments, the present invention provides improved methods for the purification of decorin, especially decorin core protein. In some embodiments, the processes comprise two column chromatography steps and a polishing step.

In some preferred embodiments, cation exchange chromatography is used to capture decorin from medium containing decorin, preferably a clarified medium. In some embodiments, the cation exchange chromatography medium is SP-Sepharose FF. In some embodiments, the cation exchange medium is equilibrated at about 5 to 15 mM sodium phosphate, preferably 10 mM sodium phosphate, and 20 to 70 mM NaCL, preferably about 50 mM NaCL at a neutral pH. In some embodiments, after application of the decorin containing medium to the cation exchange medium, the cation exchange medium is washed. In some embodiments, the wash buffer comprises about 5 to 15 mM sodium phosphate, preferably 10 mM sodium phosphate, and 20 to 70 mM NaCL, preferably about 50 mM NaCL. In some embodiments, decorin is then eluted from the cation exchange medium. In some embodiments, the elution buffer comprises about 5 to 15 mM sodium phosphate, preferably 10 mM sodium phosphate, and 150 to 250 mM NaCL, preferably about 200 mM NaCL.

In some embodiments, the eluate containing decorin from the cation exchange chromatography step is applied to a hydroxyapatite medium. In some embodiments, the hydroxyapatite medium is CHT Type 1. In some embodiments, the hydroxyapatite medium is equilibrated at about 5 to 15 mM sodium phosphate, preferably 10 mM sodium phosphate, and 150 to 250 mM NaCL, preferably about 200 mM NaCL. In some embodiments, after application of the decorin containing medium to the hydroxyapatite medium, the hydroxyapatite medium is washed. In some embodiments, the wash buffer comprises about 5 to 15 mM sodium phosphate, preferably 10 mM sodium phosphate, and 150 to 250 mM NaCL, preferably about 200 mM NaCL. In some embodiments, decorin is then eluted from the hydroxyapatite medium. In some embodiments, the elution buffer comprises about 0.2 to

0.4 M sodium phosphate, preferably 0.3 M sodium phosphate, and 150 to 250 mM NaCL, preferably about 200 mM NaCL.

In some embodiments, the eluate containing decorin from the hydroxyapatite chromatography step is buffer exchanged and applied to an ion exchange membrane. In some embodiments, the ion exchange membrane is a Q ion exchange membrane, for examples a Mustang Q ion exchange medium. In some embodiments, the membrane is equilibrated with from about 30 mM to 70 mM Tris-HCl, preferably about 50 mM Tris-HCl. In some embodiments, after application of the decorin-containing solution to the membrane, the membrane is washed and the decorin passes through the membrane. In some embodiments, the wash buffer comprises from about 30 mM to 70 mM Tris-HCl, preferably about 50 mM Tris-HCl. Following purification, the decorin is preferably concentrated to a desired concentration, for example by flow filtration.

In other embodiments of the present invention, solutions of containing the decorin are treated to inactivate or remove viruses. In some embodiments, solutions comprising decorin are treated with a surfactant to inactivate viruses. In some embodiments, the surfactant is Triton X-100. In some embodiments, the surfactant treating step is performed after the cation exchange chromatography step. In some embodiments, the solutions comprising decorin are filtered to remove viruses. In some embodiments, the solutions are filtered through a a viral Filter (e.g., a Virosart viral filter). In some embodiments, the filtrations step is performed after the hydroxyapatite chromatography step.

The processes of the present invention preferably provide decorin compositions suitable for clinical use in human patients. In some embodiments, the compositions comprise a purified decorin core protein comprising a mutation at position 4 of the mature decorin core protein so that the decorin protein is substantially non-gagylated. In some embodiments, the compositions provide purified decorin proteins, and the compositions are characterized in comprising less than about 100, 50, 20, 10, 5 or 2 ng residual host cell protein/mg decorin core protein in the composition and/or less than about 20, 10, 5, or 2 pg residual host cell DNA/mg decorin core protein. In some embodiments, the decorin is provided in a aqueous solution. In some embodiments, the aqueous solution is phosphate buffered saline (e.g., 10mM sodium phosphate, 150mM sodium chloride), with a pH of from about 6.5 to 7.5, preferably about 7.0.

The present invention further provides pharmaceutical compositions comprising veterinary decorin purified as described above. Pharmaceutically acceptable carriers are well known in the art and include aqueous solutions such as physiologically buffered saline or

other solvents or vehicles such as glycols, glycerol, vegetable oils (e.g., olive oil) or injectable organic esters. Further pharmaceutically acceptable carriers include, for example, hyaluronic acid, and aqueous solutions such as bicarbonate buffers, phosphate buffers, Ringer's solution and physiological saline supplemented with 5% dextrose or human serum albumin, if desired. A pharmaceutically acceptable carrier can be used to administer the decorin polypeptide to a cell in vitro or to a subject in vivo. A pharmaceutically acceptable carrier can contain a physiologically acceptable compound that acts, for example, to stabilize the polypeptide or to increase or decrease the absorption of the agent. A physiologically acceptable compound can include, for example, carbohydrates, such as glucose, sucrose or dextrans, antioxidants, such as ascorbic acid or glutathione, chelating agents, low molecular weight proteins or other stabilizers or excipients. Other physiologically acceptable compounds include wetting agents, emulsifying agents, dispersing agents or preservatives, which are particularly useful for preventing the growth or action of microorganisms. Various preservatives are well known and include, for example, phenol and ascorbic acid. One skilled in the art would know that the choice of a pharmaceutically acceptable carrier, including a physiologically acceptable compound, depends, for example, on the route of administration of the polypeptide and on the particular physio-chemical characteristics of the specific polypeptide. For example, a physiologically acceptable compound such as aluminum monosterate or gelatin is particularly useful as a delaying agent, which prolongs the rate of absorption of a pharmaceutical composition administered to a subject.

A parmaceutical composition comprising an effective amount of veterinary decorin can be administered to a subject by various routes including, for example, topically, orally or parenterally, such as intravenously, intramuscularly, subcutaneously, intraperitoneally or by passive or facilitated absorption through the skin using, for example, a skin patch or transdermal iontophoresis, respectively. Topical administration can be passive, for example, by direct application of an ointment or powder, or active, for example, using a nasal spray or inhalant. Where the composition is administered as a topical spray, one component of the composition is an appropriate propellant. The pharmaceutical composition also can be incorporated, if desired, into liposomes, microspheres or other polymer matrices (Gregoriadis, Liposome Technology, Vol. 1 (CRC Press, Boca Raton, Fla. 1984), which is incorporated herein by reference). Liposomes, for example, which consist of phospholipids or other lipids, are nontoxic, physiologically acceptable and metabolizable carriers that are relatively simple to make and administer.

An effective amount of the pharmaceutical composition comprising veterinary decorin is generally in the range of about 0.01 to 100 mg/kg body weight. An effective amount can be determined using methods known to those in the art. The total effective amount can be administered to a subject as a single dose, either as a bolus or by infusion over a relatively short period of time, or can be administered using a fractionated treatment protocol, in which the multiple doses are administered over a more prolonged period of time. One skilled in the art would know that the amount of decorin required to obtain an effective dose in a subject depends on many factors including the age and general health of the subject as well as the route of administration and the number of treatments to be administered.

The veterinary decorin-containing compositions of the present invention can be combined with additional active agents. These agents include, but are not limited to, decorin-synthesis enhencers, collagen-synthesis enhencers, matrix metalloproteinases (MMP) inhibitors, antioxidants, collagen modulators, anti-wrinkle or anti-aging agents, antibiotics, depigmenting agents, analgesics, antimicrobials, anti-inflammatory agents, moisturers, skin lightening agents, corticosteroids, or sun-block agents.

The veterinary decorin-containing compositions can be combined with a cosmetically or pharmaceutically or dermatologically acceptable carrier. The carriers include, but are not limited to, water, mineral oil, ethylene glycol, propylene glycol, lanolin, glyceryl stearate, sorbitan stearate, isopropyl myristate, isopropyl palmitate, acetone, glycerol, phosphatidylcholine, sodium cholate, or ethanol.

The veterinary decorin-containing compositions can be combined with a skin penentration enhancer. The enhancers, helping to transport the active components through the normal intact skin, include, but are not limited to, liposomes, mixed lipid micelles, ethosomes, transfersomes, niosomes, ethanol, amides, ethers, glycols, hydrocarbon oils, sodium lauryl sulfate, oleic acid, hydroalcoholic solution, and soya phosphatidylcholine or their combinations. Other skin penetration enhancement includes different pH values, co-solvents, surfactants, cyclodextrins, and iontophoresis.

A suitable carrier or vehicle or enhancer will include the formulation of gels, creams, lotions, solutions, colloidal dispersions, emulsions (oil-in-water or water-in-oil), foams, sprays, suspensions, sunscreens, liquid and various skin care preparations for topical application to the skin. The decorin-containing compositions can be prepared in any formula for topical application to the skin.

The formulation mentioned above can also be combined with other ingredients, depending on the intended use of the formulation. These ingredients include, but are not

limited to, preservatives, vitamins, polymers, fragrances, water- or oil-soluble film formers, or flavoring agents.

The veterinary decorin compositions of the present invention find a variety of uses. While the methods utilizing veterinary decorin are generally applicable, specific examples of pathologies which can be treated include cancer, a fibrotic disease, and glomerulonephritis. In fibrotic cancer, for example, decorin can be used to bind TGF-β, destroying TGF-β's growth stimulating activity on the cancer cell. The verterinary decorin compositions further find use in treating tumors which are positive for the EGF rececptor and/or IGF-1 receptor. Other proliferative pathologies include rheumatoid arthritis, arteriosclerosis, adult respiratory distress syndrome, cirrhosis of the liver, fibrosis of the lungs, kidneys, or liver, post-myocardial infarction, cardiac fibrosis, post-angioplasty restenosis, renal interstitial fibrosis and certain dermal fibrotic conditions such as keloids and dermal scarring. In some particularly preferred embodiments, the compositions are used to treat or inhibit scarring in the proud flesh of horses. In further embodiments, the compositions are used to treat injuries to the eye, for example, injury to the eye resulting from corneal surgery, eye burns (chemical or thermal), eye infections, and abrasive injuries to the eye. In some embodiments, the decorin compositions are used to treat heart disease. In still further embodiments, the decorin compositions are used to treat neurological traumas, such as brain or spinal cord injuries.

EXPERIMENTAL

EXAMPLE 1

Expression constructs have been designed for a mutant form of veterinary decorin from a number of different species of companion animals and stock animals. A serine to alanine modification was made at amino acid 4 of mature decorin for each of the different species. The mutation prevents a GAG from being attached to the decorin molecule. The expression constructs use the bovine α-lactalbumin signal peptide instead of the endogenous signal peptide for protein production and secretion. The constructs are outlined below.

Chicken Decorin Expression Gene Sequence (SEQ ID NO:1):

ATGATGTCCTTTGTCTCTCTGCTCCTGGTAGGCATCCTATTCCATGCCACCC
AGGCCGGACCATTTCAACAGAAAGGCTTATTTGACTTTATGCTGGAA**GATGAGG**
CTGCAGGGATAGGCCCGGAAGAGCACTTTCCTGAAGTTCCTGAAATAGAGCCTA

TGGGCCCAGTCTGCCCCTTCCGCTGTCAGTGCCATCTGCGAGTTGTCCAGTGTTCT
GATCTGGGTCTGGAAAAAGTACCAAAAGACCTTCCTCCTGATACTGCGCTGCTGG
ACCTGCAAAACAACAAAATAACTGAGATCAAAGATGGAGACTTTAAGAACCTGA
AGAACCTTCATACACTGATTCTCATCAACAACAAAATTAGCAAAATCAGCCCTGG
GGCATTTGCTCCTTTGGTGAAATTGGAACGACTTTATCTTTCCAAGAATCAACTG
AAGGAATTGCCAGAGAAAATGCCCAAAACTCTTCAGGAGCTGCGTGTCCATGAG
AACGAGATCACCAAAGTGCGAAAGTCTGTGTTCAATGGATTGAACCAGATGATC
GTCGTAGAACTTGGCACCAACCCGCTGAAGAGCTCAGGCATTGAAAATGGAGCC
TTTCAGGGAATGAAGAAGCTCTCCTACATCCGCATTGCTGACACAAATATAACTA
CCATCCCTCAAGGTCTTCCTCCTTCCCTTACTGAATTACATCTCGATGGCAACAAA
ATCACCAAAGTTGATGCAGCTAGCCTGAAAGGACTGAATAATTTGGCTAAGTTG
GGACTGAGTTTCAACAGCATCTCTGCGGTTGACAATGGCTCTTTGGCCAACACTC
CTCATTTGAGGGAACTTCATTTGAACAACAACAAGCTTGTCAAAGTGCCCGGTGG
GCTGGCCGATCATAAGTACATCCAGGTTGTCTACCTTCACAACAACAATATCTCT
GCAATCGGCTCTAACGACTTCTGCCCACCCGGATACAACACCAAAAAGGCTTCTT
ATTCAGGAGTGAGCCTTTTCAGCAACCCAGTCCAGTACTGGGAGATCCAGCCATC
CACCTTCCGATGTGTCTATGTGCGTGCTGCCGTTCAGCTTGGAAACTACAAGTGA

SEQ ID NO:2:

ATGATGTCCTTTGTCTCTCTGCTCCTGGTAGGCATCCTATTCCATGCCACCC
AGGCCACGCGGTTCCACCAGAAGGGCCTCTTTGACTTTATGATAGAGGATGAAG
GGGCAGCCGACATGGCTCCAACAGATGATCCTGTCATATCTGGATTTGGGCCAGT
GTGCCCCTTCCGCTGCCAGTGTCATCTTCGCGTTGTGCAGTGCTCTGACCTAGGTC
TGGAAAGAGTGCCAAAAGACCTTCCCCCTGACACAACTCTGCTGGATTTACAGA
ACAACAAAATCACTGAAATTAAAGAAGGAGATTTCAAGAATTTGAAGAATCTTC
ATGCATTGATCCTTGTTAACAACAAAATCAGCAAAATAAGTCCGGCAGCTTTTGC
TCCTCTGAAGAAACTGGAAAGACTGTACCTATCCAAGAATAATTTGAAGGAACTT
CCAGAAAACATGCCAAAGTCTCTTCAGGAGATACGTGCTCATGAAAATGAGATC
TCCAAGTTGAGGAAGGCAGTTTTTAATGGACTGAATCAAGTGATTGTCTTAGAAC
TAGGCACCAATCCACTCAAGAGCTCAGGCATTGAAAATGGAGCTTTTCAAGGGA
TGAAGAGGCTTTCCTATATCCGCATCGCAGACACCAACATTACTAGCATCCCTAA
AGGTCTTCCTCCATCCCTTACTGAGCTTCACCTTGATGGCAACAAAATTAGCAAA
ATTGATGCGGAAGGTCTGTCTGGACTCACCAACTTGGCTAAATTGGGTCTCAGCT
TCAACAGTATTTCTTCTGTTGAAAATGGCTCTCTGAACAATGTACCTCATCTGAG

AGAACTTCATCTGAATAACAACGAACTTGTCAGAGTACCTAGTGGGTTGGGTGA
ACACAAATACATCCAGGTGGTCTATCTTCATAACAACAAGATTGCTTCAATTGGT
ATCAACGACTTTTGCCCTCTTGGCTACAACACCAAAAAGGCAACCTATTCTGGTG
TGAGTCTCTTCAGCAACCCCGTGCAGTACTGGGAAATCCAGCCCTCTGCTTTCCG
ATGTATCCATGAACGCTCTGCAGTACAGATCGGAAATTACAAATGA

The bovine α -lactalbumin signal peptide coding region is shown in boldface type
The chicken decorin pro-peptide coding region is underlined
The mutated mature decorin coding region is shown in standard type

Chicken Decorin Expression Protein Sequence (SEQ ID NO:3):

MMSFVSLLLVGILFHATQATRFHQKGLFDFMIEDEG**A**ADMAPTDDPVISGFGPVCP
FRCQCHLRVVQCSDLGLERVPKDLPPDTTLLDLQNNKITEIKEGDFKNLKNLHALILV
NNKISKISPAAFAPLKKLERLYLSKNNLKELPENMPKSLQEIRAHENEISKLRKAVFN
GLNQVIVLELGTNPLKSSGIENGAFQGMKRLSYIRIADTNITSIPKGLPPSLTELHLDGN
KISKIDAEGLSGLTNLAKLGLSFNSISSVENGSLNNVPHLRELHLNNNELVRVPSGLGE
HKYIQVVYLHNNKIASIGINDFCPLGYNTKKATYSGVSLFSNPVQYWEIQPSAFRCIH
ERSAVQIGNYK.

The bovine α -lactalbumin signal peptide is shown in boldface type
The chicken decorin pro-peptide is underlined

The mutated mature decorin coding region is shown in standard type (Amino acid #4 of mature decorin was mutated from a serine to an alanine to prevent addition of the GAG moiety to the mature protein. This amino acid is shown in boldface type as well as underlined).

Mature chicken decorin (SEQ ID NO:4):

DEG**A**ADMAPTDDPVISGFGPVCPFRCQCHLRVVQCSDLGLERVPKDLPPDTTLLDLQ
NNKITEIKEGDFKNLKNLHALILVNNKISKISPAAFAPLKKLERLYLSKNNLKELPEN
MPKSLQEIRAHENEISKLRKAVFNGLNQVIVLELGTNPLKSSGIENGAFQGMKRLSYI
RIADTNITSIPKGLPPSLTELHLDGNKISKIDAEGLSGLTNLAKLGLSFNSISSVENGSLN

NVPHLRELHLNNNELVRVPSGLGEHKYIQVVYLHNNKIASIGINDFCPLGYNTKKAT
YSGVSLFSNPVQYWEIQPSAFRCIHERSAVQIGNYK

Cow Decorin Expression Gene Sequence (SEQ ID NO:5):

ATGATGTCCTTTGTCTCTCTGCTCCTGGTAGGCATCCTATTCCATGCCACCC
AGGCCGGACCATTTCAACAGAAAGGCTTATTTGACTTTATGCTGGAAGATGAGG
CTGCAGGGATAGGCCCGGAAGAGCACTTTCCTGAAGTTCCTGAAATAGAGCCTA
TGGGCCCAGTCTGCCCCTTCCGCTGTCAGTGCCATCTGCGAGTTGTCCAGTGTTCT
GATCTGGGTCTGGAAAAAGTACCAAAAGACCTTCCTCCTGATACTGCGCTGCTGG
ACCTGCAAAACAACAAAATAACTGAGATCAAAGATGGAGACTTTAAGAACCTGA
AGAACCTTCATACACTGATTCTCATCAACAACAAAATTAGCAAAATCAGCCCTGG
GGCATTTGCTCCTTTGGTGAAATTGGAACGACTTTATCTTTCCAAGAATCAACTG
AAGGAATTGCCAGAGAAAATGCCCAAAACTCTTCAGGAGCTGCGTGTCCATGAG
AACGAGATCACCAAAGTGCGAAAGTCTGTGTTCAATGGATTGAACCAGATGATC
GTCGTAGAACTTGGCACCAACCCGCTGAAGAGCTCAGGCATTGAAAATGGAGCC
TTTCAGGGAATGAAGAAGCTCTCCTACATCCGCATTGCTGACACAAATATAACTA
CCATCCCTCAAGGTCTTCCTCCTTCCCTTACTGAATTACATCTCGATGGCAACAAA
ATCACCAAAGTTGATGCAGCTAGCCTGAAAGGACTGAATAATTTGGCTAAGTTG
GGACTGAGTTTCAACAGCATCTCTGCGGTTGACAATGGCTCTTTGGCCAACACTC
CTCATTTGAGGGAACTTCATTTGAACAACAACAAGCTTGTCAAAGTGCCCGGTGG
GCTGGCCGATCATAAGTACATCCAGGTTGTCTACCTTCACAACAACAATATCTCT
GCAATCGGCTCTAACGACTTCTGCCCACCCGGATACAACACCAAAAAGGCTTCTT
ATTCAGGAGTGAGCCTTTTCAGCAACCCAGTCCAGTACTGGGAGATCCAGCCATC
CACCTTCCGATGTGTCTATGTGCGTGCTGCCGTTCAGCTTGGAAACTACAAGTGA

The bovine α -lactalbumin signal peptide coding region is shown in boldface type
The bovine decorin pro-peptide coding region is underlined
The mutated mature decorin coding region is shown in standard type

Cow Decorin Expression Protein Sequence (SEQ ID NO:6):

MMSFVSLLLVGILFHATQAGPFQQKGLFDFMLEDEA**A**GIGPEEHFPEVPEIEPMGPV
CPFRCQCHLRVVQCSDLGLEKVPKDLPPDTALLDLQNNKITEIKDGDFKNLKNLHTLI

LINNKISKISPGAFAPLVKLERLYLSKNQLKELPEKMPKTLQELRVHENEITKVRKSVF
NGLNQMIVVELGTNPLKSSGIENGAFQGMKKLSYIRIADTNITTIPQGLPPSLTELHLD
GNKITKVDAASLKGLNNLAKLGLSFNSISAVDNGSLANTPHLRELHLNNNKLVKVPG
GLADHKYIQVVYLHNNNISAIGSNDFCPPGYNTKKASYSGVSLFSNPVQYWEIQPSTF
RCVYVRAAVQLGNYK.

The bovine α -lactalbumin signal peptide is shown in boldface type
The bovine decorin pro-peptide is underlined

The mutated mature decorin coding region is shown in standard type (Amino acid #4 of mature decorin was mutated from a serine to an alanine to prevent addition of the GAG moiety to the mature protein. This amino acid is shown in boldface type as well as underlined)

Mature cow decorin protein sequence (SEQ ID NO:7):

DEA**A**GIGPEEHFPEVPEIEPMGPVCPFRCQCHLRVVQCSDLGLEKVPKDLPPDTALLD
LQNNKITEIKDGDFKNLKNLHTLILINNKISKISPGAFAPLVKLERLYLSKNQLKELPE
KMPKTLQELRVHENEITKVRKSVFNGLNQMIVVELGTNPLKSSGIENGAFQGMKKLS
YIRIADTNITTIPQGLPPSLTELHLDGNKITKVDAASLKGLNNLAKLGLSFNSISAVDN
GSLANTPHLRELHLNNNKLVKVPGGLADHKYIQVVYLHNNNISAIGSNDFCPPGYNT
KKASYSGVSLFSNPVQYWEIQPSTFRCVYVRAAVQLGNYK

Dog Decorin Expression Gene Sequence (SEQ ID NO:8):

ATGATGTCCTTTGTCTCTCTGCTCCTGGTAGGCATCCTATTCCATGCCACCC
AGGCCGGGCCGTTCCAACAGAGAGGCTTATTTGACTTTATGCTAGAAGATGAGG
CTGCAGGGATAGGCCCGGAGGACCGTGCACCTGACATGCCTGACCTCGAGCTTC
TGGGACCTGTGTGTCCCTTCCGCTGTCAGTGCCATCTCCGAGTGGTCCAGTGTTCC
GACCTGGGTCTGGACAAAGTACCAAAAGATCTTCCCCCTGACACTACGCTGCTCG
ACTTGCAAAACAACAAAATCACCGAAATCAAAGATGGAGACTTCAAGAACCTCA
AGAACCTGCATACCTTGATTCTTGTAAACAACAAAATTAGCAAAATCAGCCCTGG
AGCATTTACACCTTTGTTGAAATTGGAACGACTTTATCTGTCCAAGAATCATCTG
AAGGAATTGCCAGAAAAAATGCCCAAAACTCTTCAGGAGCTGCGTGCCCATGAG

AATGAGATCACCAAAGTTCGAAAAGCTGTGTTCAATGGACTGAACCAGATGATC
GTCGTAGAGCTGGGCACCAATCCCCTGAAGAGTTCAGGGATTGAAAATGGAGCC
TTCCAGGGAATGAAGAAGCTCTCCTATATCCGCATTGCTGATACCAATATAACTA
CCATCCCTCAAGGTCTTCCTCCTTCCCTTACTGAATTACATCTTGAAGGCAACAAA
ATCACCAAGGTTGATGCATCTAGCCTGAAAGGACTGAATAATTTGGCTAAGTTGG
GACTGAGTTTTAACAGCATCTCCGCTGTTGACAATGGCACTCTAGCCAACACTCC
TCATCTGAGGGAGCTTCACTTGGACAACAATAAGCTCATCAGAGTACCCGGTGG
GCTGGCGGAGCATAAGTACATCCAGGTTGTCTACCTTCATAACAACAATATATCT
GCAGTCGGATCTAATGACTTCTGCCCACCTGGATACAACACCAAAAAGGCTTCTT
ATTCAGGTGTGAGCCTTTTCAGCAACCCAGTGCAGTACTGGGAGATCCAGCCATC
CACCTTCCGGTGTGTCTACGTGCGCTCTGCCATCCAGCTTGGAAATTATAAATGA

The bovine α -lactalbumin signal peptide coding region is shown in boldface type
The dog decorin pro-peptide coding region is underlined
The mutated mature decorin coding region is shown in standard type

Dog Decorin Expression Protein Sequence (SEQ ID NO:9):

MMSFVSLLLVGILFHATQAGPFQQRGLFDFMLEDEA**A**GIGPEDRAPDMPDLELLGP
VCPFRCQCHLRVVQCSDLGLDKVPKDLPPDTTLLDLQNNKITEIKDGDFKNLKNLHT
LILVNNKISKISPGAFTPLLKLERLYLSKNHLKELPEKMPKTLQELRAHENEITKVRKA
VFNGLNQMIVVELGTNPLKSSGIENGAFQGMKKLSYIRIADTNITTIPQGLPPSLTELH
LEGNKITKVDASSLKGLNNLAKLGLSFNSISAVDNGTLANTPHLRELHLDNNKLIRVP
GGLAEHKYIQVVYLHNNNISAVGSNDFCPPGYNTKKASYSGVSLFSNPVQYWEIQPS
TFRCVYVRSAIQLGNYK.

The bovine α -lactalbumin signal peptide is shown in boldface type
The dog decorin pro-peptide is underlined

The mutated mature decorin coding region is shown in standard type (Amino acid #4 of mature decorin was mutated from a serine to an alanine to prevent addition of the GAG moiety to the mature protein. This amino acid is shown in boldface type as well as underlined)

Mature dog decorin sequence (SEQ ID NO:10):

DEA**A**GIGPEDRAPDMPDLELLGPVCPFRCQCHLRVVQCSDLGLDKVPKDLPPDTTLL
DLQNNKITEIKDGDFKNLKNLHTLILVNNKISKISPGAFTPLLKLERLYLSKNHLKELP
EKMPKTLQELRAHENEITKVRKAVFNGLNQMIVVELGTNPLKSSGIENGAFQGMKK
LSYIRIADTNITTIPQGLPPSLTELHLEGNKITKVDASSLKGLNNLAKLGLSFNSISAVD
NGTLANTPHLRELHLDNNKLIRVPGGLAEHKYIQVVYLHNNNISAVGSNDFCPPGYN
TKKASYSGVSLFSNPVQYWEIQPSTFRCVYVRSAIQLGNYK

Goat Decorin Expression Gene Sequence (SEQ ID NO:11):

ATGATGTCCTTTGTCTCTCTGCTCCTGGTAGGCATCCTATTCCATGCCACCC
AGGCCGGACCGTTTCAACAGAAAGGCTTATTTGACTTTATGCTGGAAGATGAGG
CTGCAGGGATAGGCCCGGAAGAGCGCTTTCATGAAGTTCCTGAATTAGAGCCTAT
GGGCCCAGTCTGCCCCTTCCGCTGTCAGTGCCATCTGCGAGTTGTCCAGTGTTCTG
TTCTGGGTCTGGAAAAAGTGCCCAAAGACCTTCCTCCTGATACCGCGCTGCTGGA
CCTGCAAAACAACAAAATAACTGAGATCAAAGATGGAGACTTTAAGAACCTGAA
GAACCTTCATACACTGATTCTCATCAACAACAAAATTAGCAAAATCAGCCCTGGG
GCATTTGCTCCTCTGGTGAAATTGGAACGACTTTATCTTTCCAAGAATCAACTGA
AGGAATTGCCAGAGAAAATGCCCAAAACTCTTCAGGAGCTGCGTGTCCATGAGA
ACGAGATCACCAAAGTGCGAAAGTCTGTGTTCAATGGATTGAACCAGATGATCG
TCGTAGAACTTGGCACCAACCCACTGAAGAGCTCAGGCATTGAAAATGGAGCCT
TTCAGGGAATGAAGAAGCTCTCCTACATCCGCATTGCTGACACTAATATAACTAC
CATTCCTCAAGGTCTTCCTCCTTCCCTTACTGAATTACATCTCGATGGCAACAAAA
TCACCAAAGTTGATGCAGCTAGCCTGAAAGGACTGAATAATTTGGCTAAGTTGG
GACTGAGTTTCAACAGCATCTCTGCTGTTGACAATGGCTCTTTAGCCAACACTCC
TCATTTGAGGGAACTTCATTTGAACAACAACAAGCTTGTCAAAGTGCCCGGTGGG
CTGGCCGACCATAAGTACATCCAGGTTGTCTACCTTCACAACAACAATATCTCTG
CAATCGGCTCCAACGACTTCTGCCCACCCGGATACAACACCAAAAAGGCTTCTTA
TTCAGGAGTGAGCCTTTTCAGCAACCCAGTCCAGTACTGGGAGATCCAGCCATCC
ACCTTCCGATGTGTCTACGTGCGCGCTGCTGTTCAGCTTGGAAACTACAAGTGA

The bovine α -lactalbumin signal peptide coding region is shown in boldface type
The goat decorin pro-peptide coding region is underlined

The mutated mature decorin coding region is shown in standard type

Goat Decorin Expression Protein Sequence (SEQ ID NO:12):

MMSFVSLLLVGILFHATQAGPFQQKGLFDFMLEDEA**A**GIGPEERFHEVPELEPMGP
VCPFRCQCHLRVVQCSVLGLEKVPKDLPPDTALLDLQNNKITEIKDGDFKNLKNLHT
LILINNKISKISPGAFAPLVKLERLYLSKNQLKELPEKMPKTLQELRVHENEITKVRKS
VFNGLNQMIVVELGTNPLKSSGIENGAFQGMKKLSYIRIADTNITTIPQGLPPSLTELH
LDGNKITKVDAASLKGLNNLAKLGLSFNSISAVDNGSLANTPHLRELHLNNNKLVKV
PGGLADHKYIQVVYLHNNNISAIGSNDFCPPGYNTKKASYSGVSLFSNPVQYWEIQP
STFRCVYVRAAVQLGNYK.

The bovine α -lactalbumin signal peptide is shown in boldface type
The goat decorin pro-peptide is underlined

The mutated mature decorin coding region is shown in standard type (Amino acid #4 of mature decorin was mutated from a serine to an alanine to prevent addition of the GAG moiety to the mature protein. This amino acid is shown in boldface type as well as underlined)

Mature goat decorin protein sequence (SEQ ID NO:13):

DEA**A**GIGPEERFHEVPELEPMGPVCPFRCQCHLRVVQCSVLGLEKVPKDLPPDTALL
DLQNNKITEIKDGDFKNLKNLHTLILINNKISKISPGAFAPLVKLERLYLSKNQLKELP
EKMPKTLQELRVHENEITKVRKSVFNGLNQMIVVELGTNPLKSSGIENGAFQGMKKL
SYIRIADTNITTIPQGLPPSLTELHLDGNKITKVDAASLKGLNNLAKLGLSFNSISAVDN
GSLANTPHLRELHLNNNKLVKVPGGLADHKYIQVVYLHNNNISAIGSNDFCPPGYNT
KKASYSGVSLFSNPVQYWEIQPSTFRCVYVRAAVQLGNYK

Horse Decorin Expression Gene Sequence (SEQ ID NO:14):

ATGATGTCCTTTGTCTCTCTGCTCCTGGTAGGCATCCTATTCCATGCCACCC
AGGCCGGACCATTTCAACAGAGAGGCTTATTTGACTTCATGCTAGAAGATGAGG
CTGCAGGGATTGGCCCAGAAGATCGCATTCATGAAGTTCTAGACTTAGAGCCTCT

GGGACCAGTGTGTCCTTTCCGCTGTCAGTGCCATCTTCGAGTTGTCCAATGTTCTG
ATTTGGGTCTGGACAAAGTGCCCAAAGATCTTCCCCCTGACACCACGCTGCTGGA
CCTGCAAAACAACAAAATAACCGAAATCAAAGATGGAGACTTTAAGAACCTGAA
GAATCTTCATGCGTTGATTCTTGTCAACAACAAAATTAGCAAAATCAGCCCTGGA
GCATTTACACCTTTGGTGAAACTGGAACGACTTTATCTGTCCAAGAATCATTTGA
AGGAATTGCCAGAAAAAATGCCCAAAACTCTTCAGGAGCTGCGTGTCCATGAGA
ACGAGATCACCAAAGTGCGAAAAGCGGTGTTCAATGGACTGAACCAGATGATAG
TCGTAGAACTGGGCACCAACCCACTGAAGAGCTCAGGAATTGAAAATGGAGCCT
TCCAGGGGATGAAGAAGCTGTCCTACATCCGCATTGCTGACACCAACATAACCA
CCATCCCTCCAGGTCTTCCTCCTTCCCTTACTGAATTACATCTTGATGGCAACAAA
ATCACCAAAGTTGATGCAGCTAGCCTGAGAGGACTGAATAATTTGGCTAAATTG
GGACTGAGTTTCAACAGCATCTCTGCTGTTGACAATGGCTCTCTGGCCAACACTC
CTCATTTGAGGGAACTTCACTTGGACAACAACAAGCTTATCAAAGTGCCTGGTGG
GCTGGCGGATCATAAGTACATCCAGGTTGTCTACCTTCATAACAACAATATCTCT
GCAGTTGGATCTAATGACTTCTGCCCACCTGGATACAACACCAAAAAGGCTTCTT
ATTCGGGTGTGAGCCTTTTCAGCAACCCAGTCCAGTACTGGGAGATCCAGCCATC
CACCTTCCGATGTGTCTATGTGCGCTCTGCCATTCAGCTCGGAAACTACAAGTGA

The bovine α -lactalbumin signal peptide coding region is shown in boldface type
The horse decorin pro-peptide coding region is underlined
The mutated mature decorin coding region is shown in standard type

Horse Decorin Expression Protein Sequence (SEQ ID NO:15):

MMSFVSLLLVGILFHATQA<u>GPFQQRGLFDFMLE</u>DEA**<u>A</u>**GIGPEDRIHEVLDLEPLGP
VCPFRCQCHLRVVQCSDLGLDKVPKDLPPDTTLLDLQNNKITEIKDGDFKNLKNLHA
LILVNNKISKISPGAFTPLVKLERLYLSKNHLKELPEKMPKTLQELRVHENEITKVRKA
VFNGLNQMIVVELGTNPLKSSGIENGAFQGMKKLSYIRIADTNITTIPPGLPPSLTELH
LDGNKITKVDAASLRGLNNLAKLGLSFNSISAVDNGSLANTPHLRELHLDNNKLIKV
PGGLADHKYIQVVYLHNNNISAVGSNDFCPPGYNTKKASYSGVSLFSNPVQYWEIQP
STFRCVYVRSAIQLGNYK.

The bovine α -lactalbumin signal peptide is shown in boldface type

The horse decorin pro-peptide is underlined

The mutated mature decorin coding region is shown in standard type (Amino acid #4 of mature decorin was mutated from a serine to an alanine to prevent addition of the GAG moiety to the mature protein. This amino acid is shown in boldface type as well as underlined)

Mature horse decorin protein sequence (SEQ ID NO:16):

DEA**A**GIGPEDRIHEVLDLEPLGPVCPFRCQCHLRVVQCSDLGLDKVPKDLPPDTTLLD
LQNNKITEIKDGDFKNLKNLHALILVNNKISKISPGAFTPLVKLERLYLSKNHLKELPE
KMPKTLQELRVHENEITKVRKAVFNGLNQMIVVELGTNPLKSSGIENGAFQGMKKL
SYIRIADTNITTIPPGLPPSLTELHLDGNKITKVDAASLRGLNNLAKLGLSFNSISAVDN
GSLANTPHLRELHLDNNKLIKVPGGLADHKYIQVVYLHNNNISAVGSNDFCPPGYNT
KKASYSGVSLFSNPVQYWEIQPSTFRCVYVRSAIQLGNYK

Pig Decorin Expression Gene Sequence (SEQ ID NO:17):

ATGATGTCCTTTGTCTCTCTGCTCCTGGTAGGCATCCTATTCCATGCCACCC
AGGCCGGACCATTTCAACAGAAAGGCTTATTTGACTTTATGCTAGAAGATGAGG
CTGCAGGGATAGGCCCAGAAGACCGCTTTCCTGAAGTTCCTGAATTAGAGCCTCT
GGGACCCATGTGTCCCTTCCGCTGTCAATGCCATCTTCGAGTTGTCCAATGTTCTG
ATTTGGGTCTGGACAAAGTGCCCAAAGATCTTCCACCTGACACTGCCCTGCTGGA
TCTGCAAAACAACAAAATAACTGAAATCAAAGATGGAGACTTTAAGAACCTGAA
GAACCTTCATACACTGATTCTCATCAACAACAAAATTAGCAAAATCAGCCCTGGA
GCATTTGCACCTTTGGTGAAATTGGAACGACTTTATCTATCCAAGAATCAACTGA
AGGAATTGCCAGAGAAAATGCCCAAAACTCTTCAGGAGCTGCGTGTCCATGAGA
ATGAGATCACCAAAGTGCGAAAGGCTGTGTTCAATGGATTGAACCAGATGATCG
TCGTAGAACTTGGCACCAACCCGCTGAAGAGCTCAGGCATTGAAAACGGAGCTT
TCCAGGGAATGAAGAAGCTCTCCTACATCCGCATCGCTGACACCAACATTACCAC
CATCCCTCAAGGTCTTCCTCCTTCCCTTACTGAATTACATCTTGATGGCAACAAAA
TCAGCAAAGTTGATGCAGCTAGCCTAAAAGGACTGAATAATTTGGCTAAGTTGG
GACTGGGTTTCAATAGCATCTCAACTGTTGACAATGGCTCTCTGGCCAACACTCC
TCATTTGAGGGAACTTCATCTGAACAACAACAAGCTTAACAAAGTGCCTGGTGG

GCTGGCAGAGCATAAGTACATCCAGGTTGTCTACCTTCATAACAACAACATCTCT
GCAGTCGGCTCTAATGACTTCTGCCCGCCTGGATACAACACCAAAAAGGCTTCTT
ATTCGGGGGTGAGCCTTTTCAGCAACCCAGTCCAGTACTGGGAGATCCAGCCATC
CACCTTCCGATGTGTCTATGTGCGCTCTGCCATTCAGCTCGGAAACTACAAGTGA

The bovine α -lactalbumin signal peptide coding region is shown in boldface type
The pig decorin pro-peptide coding region is underlined
The mutated mature decorin coding region is shown in standard type

Pig Decorin Expression Protein Sequence (SEQ ID NO:18):

MMSFVSLLLVGILFHATQAGPFQQKGLFDFMLEDEA**A**GIGPEDRFPEVPELEPLGP
MCPFRCQCHLRVVQCSDLGLDKVPKDLPPDTALLDLQNNKITEIKDGDFKNLKNLHT
LILINNKISKISPGAFAPLVKLERLYLSKNQLKELPEKMPKTLQELRVHENEITKVRKA
VFNGLNQMIVVELGTNPLKSSGIENGAFQGMKKLSYIRIADTNITTIPQGLPPSLTELH
LDGNKISKVDAASLKGLNNLAKLGLGFNSISTVDNGSLANTPHLRELHLNNNKLNKV
PGGLAEHKYIQVVYLHNNNISAVGSNDFCPPGYNTKKASYSGVSLFSNPVQYWEIQP
STFRCVYVRSAIQLGNYK.

The bovine α -lactalbumin signal peptide is shown in boldface type
The pig decorin pro-peptide is underlined

The mutated mature decorin coding region is shown in standard type (Amino acid #4 of mature decorin was mutated from a serine to an alanine to prevent addition of the GAG moiety to the mature protein. This amino acid is shown in boldface type as well as underlined)

Mature pig decorin sequence (SEQ ID NO:19):

DEA**A**GIGPEDRFPEVPELEPLGPMCPFRCQCHLRVVQCSDLGLDKVPKDLPPDTALL
DLQNNKITEIKDGDFKNLKNLHTLILINNKISKISPGAFAPLVKLERLYLSKNQLKELP
EKMPKTLQELRVHENEITKVRKAVFNGLNQMIVVELGTNPLKSSGIENGAFQGMKK
LSYIRIADTNITTIPQGLPPSLTELHLDGNKISKVDAASLKGLNNLAKLGLGFNSISTVD
NGSLANTPHLRELHLNNNKLNKVPGGLAEHKYIQVVYLHNNNISAVGSNDFCPPGY
NTKKASYSGVSLFSNPVQYWEIQPSTFRCVYVRSAIQLGNYK

Sheep Decorin Expression Gene Sequence (SEQ ID NO:20):

ATGATGTCCTTTGTCTCTCTGCTCCTGGTAGGCATCCTATTCCATGCCACCC
AGGCCGGACCGTTTCAACAGAAAGGCTTATTTGACTTTATGCTGGAAGATGAGG
CTGCAGGGATAGGCCCGGAAGAGCGCTTTCATGAGGTTCCTGAATTAGAGCCTAT
GGGCCCAGTCTGCCCCTTCCGTTGCCAGTGCCATCTGCGAGTTGTCCAGTGTTCTG
ATCTGGGTCTGGAAAAAGTGCCCAAAGACCTTCCTCCTGATACCGCGCTGCTGGA
CCTGCAAAACAACAAAATAACTGAGATCAAAGATGGAGACTTTAAAAACCTGAA
GAACCTTCATACACTGATTCTCATCAACAACAAAATTAGCAAAATTAGCCCTGGG
GCATTTGCTCCTCTGGTGAAATTGGAACGACTTTATCTTTCCAAGAATCAACTGA
AGGAATTGCCAGAGAAAATGCCCAAAACTCTTCAGGAGCTGCGTGTCCATGAGA
ACGAGATCACCAAAGTGCGAAAGTCTGTGTTCAATGGATTGAACCAGATGATCG
TCGTAGAACTTGGCACCAACCCACTGAAGAGCTCAGGCATTGAAAATGGAGCCT
TTCAGGGAATGAAGAAGCTCTCCTACATCCGCATTGCTGACACTAATATAACTAC
CATCCCTCAAGGTCTTCCTCCTTCCCTTACTGAATTACATCTCGACGGCAACAAA
ATCACCAAAGTTGATGCAGCTAGCCTGAAAGGACTGAATAATTTGGCTAAGTTG
GGACTGAGTTTCAACAGCATCTCTGCTGTTGACAATGGCTCTTTGGCCAACACTC
CTCATTTGAGGGAACTTCATTTGAACAACAACAAGCTTGTCAAAGTGCCCGGTGG
GCTGGCCGACCATAAGTACATCCAGGTTGTCTACCTTCACAACAACAATATCTCT
GCAATCGGCTCTAACGACTTCTGCCCACCTGGATACAACACCAAAAAGGCTTCTT
ATTCAGGAGTGAGCCTTTTCAGCAACCCAGTCCAGTACTGGGAGATCCAGCCATC
CACCTTCCGATGTGTCTACGTGCGCGCTGCTGTTCAGCTTGGAAACTACAAGTGA

The bovine α -lactalbumin signal peptide coding region is shown in boldface type
The sheep decorin pro-peptide coding region is underlined
The mutated mature decorin coding region is shown in standard type

Sheep Decorin Expression Protein Sequence (SEQ ID NO:21):

MMSFVSLLLVGILFHATQAGPFQQKGLFDFMLEDEA**A**GIGPEERFHEVPELEPMGP
VCPFRCQCHLRVVQCSDLGLEKVPKDLPPDTALLDLQNNKITEIKDGDFKNLKNLHT
LILINNKISKISPGAFAPLVKLERLYLSKNQLKELPEKMPKTLQELRVHENEITKVRKS
VFNGLNQMIVVELGTNPLKSSGIENGAFQGMKKLSYIRIADTNITTIPQGLPPSLTELH

LDGNKITKVDAASLKGLNNLAKLGLSFNSISAVDNGSLANTPHLRELHLNNNKLVKV
PGGLADHKYIQVVYLHNNNISAIGSNDFCPPGYNTKKASYSGVSLFSNPVQYWEIQP
STFRCVYVRAAVQLGNYK.

The bovine α -lactalbumin signal peptide is shown in boldface type
The sheep decorin pro-peptide is underlined

The mutated mature decorin coding region is shown in standard type (Amino acid #4 of mature decorin was mutated from a serine to an alanine to prevent addition of the GAG moiety to the mature protein. This amino acid is shown in boldface type as well as underlined)

Mature sheep decorin protein sequence (SEQ ID NO:22):

DEA**A**GIGPEERFHEVPELEPMGPVCPFRCQCHLRVVQCSDLGLEKVPKDLPPDTALL
DLQNNKITEIKDGDFKNLKNLHTLILINNKISKISPGAFAPLVKLERLYLSKNQLKELP
EKMPKTLQELRVHENEITKVRKSVFNGLNQMIVVELGTNPLKSSGIENGAFQGMKKL
SYIRIADTNITTIPQGLPPSLTELHLDGNKITKVDAASLKGLNNLAKLGLSFNSISAVDN
GSLANTPHLRELHLNNNKLVKVPGGLADHKYIQVVYLHNNNISAIGSNDFCPPGYNT
KKASYSGVSLFSNPVQYWEIQPSTFRCVYVRAAVQLGNYK

Feline Decorin Expression Gene Sequence (SEQ ID NO: 25)

ATGATGTCCTTTGTCTCTCTGCTCCTGGTAGGCATCCTATTCCATGCCACCC
AGGCCGGGCCGTTCCAACAGAGAGGCTTATTTGACTTTATGCTAGAAGATGAGG
CTGCAGGGATAGGCCCAGAAGAGCACGCTCCTGTTGATTCTGATTTAGAGCCTCT
GGGGCCAGTGTGTCCTTTCCGCTGTCAGTGCCACCTTCGAGTTGTGCAGTGTTCTG
ATTTGGGTTTGGAAAAAGTGCCAAAAGAGCTCCCTCCTGACACTACGCTGCTGGA
CTTGCAAAACAACAAAATAACCGAAATCAAAGATGGAGACTTCAAGAACCTGAA
GAACCTTCATACGTTGATCCTTGTCAACAACAAAATTAGCAAAATCAGCCCTGGA
GCATTTACACCTTTGTTGAAATTGGAACGACTTTATCTGTCCAAGAATCATCTGA
AGGAATTGCCAGAAAAAATGCCCAAAACTCTTCAGGAGCTGCGTGCTCACGAGA
ATGAGATCACCAAAGTGCGAAAAGCTGTGTTCAATGGCCTGAACCAGATGATCG
TCGTAGAACTGGGCACCAACCCGCTGAAGAGCTCGGGAATTGAAAATGGAGCCT
TCCAGGGAATGAAGAAGCTGTCCTACATCCGCATTGCCGACACCAATATAACCA

CCATCCCGCAAGGTCTTCCTCCTTCCCTTACTGAATTACATCTTGAAGGCAACAA
AATCTCCAAAGTTGATGCAGCTAGCCTGAAAGGACTGAATAATTTGGCTAAGTTG
GGACTGAGTTTTAACAGCATCTCTGCTATTGACAATGGCACTCTGGCCAACACTC
CTCATTTGAGGGAGCTTCACTTGGACAACAATAAGCTTATCAGAGTACCTGGTGG
GCTGGCGGAGCACAAATACATCCAGGTTGTCTACCTTCATAACAACAATATCTCT
GCAGTCGGGTCTAACGACTTCTGCCCACCTGGATACAACACCAAAAAGGCTTCTT
ATTCAGGTGTGAGCCTTTTCAGCAACCCAGTCCAGTACTGGGAGATCCAACCATC
CACCTTCCGATGTGTCTATGTGCGTTCCGCCATCCAGCTTGGAAATTATAAATGA

The bovine α –lactalbumin signal peptide coding region is shown in boldface type
The cat decorin pro-peptide coding region is underlined
The mutated mature decorin coding region is shown in standard type

Feline Decorin Expression Protein Sequence (SEQ ID NO: 26)

MMSFVSLLLVGILFHATQA<u>GPFQQRGLFDFMLE</u>DEA**<u>A</u>**GIGPEEHAPVDSDLEPLGP
VCPFRCQCHLRVVQCSDLGLEKVPKELPPDTTLLDLQNNKITEIKDGDFKNLKNLHT
LILVNNKISKISPGAFTPLLKLERLYLSKNHLKELPEKMPKTLQELRAHENEITKVRKA
VFNGLNQMIVVELGTNPLKSSGIENGAFQGMKKLSYIRIADTNITTIPQGLPPSLTELH
LEGNKISKVDAASLKGLNNLAKLGLSFNSISAIDNGTLANTPHLRELHLDNNKLIRVP
GGLAEHKYIQVVYLHNNNISAVGSNDFCPPGYNTKKASYSGVSLFSNPVQYWEIQPS
TFRCVYVRSAIQLGNYK
The bovine α –lactalbumin signal peptide is shown in boldface type
The cat decorin pro-peptide is underlined

The mutated mature decorin coding region is shown in standard type (Amino acid #4 of mature decorin was mutated from a serine to an alanine to prevent addition of the GAG moiety to the mature protein. This amino acid is shown in boldface type as well as underlined)

Mature feline decorin protein sequence (SEQ ID NO:27):

DEA**<u>A</u>**GIGPEEHAPVDSDLEPLGPVCPFRCQCHLRVVQCSDLGLEKVPKELPPDTTLLD
LQNNKITEIKDGDFKNLKNLHTLILVNNKISKISPGAFTPLLKLERLYLSKNHLKELPE
KMPKTLQELRAHENEITKVRKAVFNGLNQMIVVELGTNPLKSSGIENGAFQGMKKL
SYIRIADTNITTIPQGLPPSLTELHLEGNKISKVDAASLKGLNNLAKLGLSFNSISAIDN

GTLANTPHLRELHLDNNKLIRVPGGLAEHKYIQVVYLHNNNISAVGSNDFCPPGYNT
KKASYSGVSLFSNPVQYWEIQPSTFRCVYVRSAIQLGNYK

EXAMPLE 2

Gene Construct Development.

Canine and Feline Decorin DNA construction and cloning. The Canine and Feline gene sequences of decorin were synthesized with the bovine alpha-lactalbumin signal peptide attached. Both DNA sequences were cloned into Catalent's GPEx® expression vectors (Figures 1-4). See, e.g., Bleck, G.T. 2005 An alternative method for the rapid generation of stable, high-expressing mammalian cell lines (A Technical Review). Bioprocessing J. Setp/Oct. pp 1-7; Bleck, G.T., 2010. GPEx® A Flexible Method for the Rapid Generation of Stable, High Expressing, Antibody Producing Mammalian Cell Lines Chapter 4 In: Current Trends in Monoclonal Antibody Development and Manufacturing, Biotechnology: Pharmaceutical Aspects, Edited by: S.J. Shire et al. © 2010 American Association of Pharmaceutical Scientists, DOI 10.1007/978-0-387-76643-0_4.

CHO Cell Line Development

Retrovector Production. The expression constructs outlined above were introduced into a HEK 293 cell line that constitutively produces the MLV gag, pro, and pol proteins. An envelope containing expression plasmid was also co-tranfected with the decorin gene construct. The co-transfection resulted in the production of replication incompetent high titer retrovector that was concentrated by ultracentrifugation and used for cell transductions (1,2).

Transduction of GCHO Cells with Retrovector. The Canine and Feline Decorin pooled cell lines were made by performing multiple rounds of transduction of the GPEx® Chinese Hamster Ovary (GCHO) parental cell line with retrovector made from the gene constructs developed to express the two decorin molecules. Three independent transductions were performed to generate a pooled cell line for each of the two products.

Fed Batch Production of Canine and Feline from the Pooled Population of Cells. Post-transduction, the pooled cell lines for Canine and Feline decorin were scaled up for

productivity in a fed batch study in duplicate 250 mL shake flasks. Each shake flask was seeded with 300,000 viable cells per mL in a 60 mL working volume of PF CHO LS media (HyClone) and incubated in a humidified (70-80%) shaking incubator at 130 rpm with 5% CO2 and temperature of 37°C. Cultures were fed four times during the production run using two different feed supplements. Cultures were terminated when viabilities were ≤ 50% (Day 14). Confirmation of canine and feline decorin production was determined by SDS-PAGE gel analysis (Figures 5 and 6). The traditional doublet associated with the human form (see Figure 7) of this molecule was observed for both canine and feline.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention that are obvious to those skilled in the field of this invention are intended to be within the scope of the following claims.

CLAIMS

1. An isolated veterinary decorin core protein molecule that is at least 95% identical to one of SEQ ID NOs:4, 7, 10, 13, 16, 19, 22, and 27, with the proviso that the veterinary decorin core protein comprises a mutation at amino acid 4.

2. The veterinary decorin core protein molecule of claim 1, wherein said mutation prevents gagylation of said molecule.

3. The veterinary decorin core protein molecule of claim 1, wherein said mutation is a serine to alanine mutation.

4. The veterinary decorin core protein molecule of claim 1, wherein said protein molecule is at least 98% identical to one of SEQ ID NOs:4, 7, 10, 13, 16, 19, 22, and 27, with the proviso that the veterinary decorin core protein comprises a mutation at amino acid 4.

5. The veterinary decorin core protein molecule of claim 1, wherein said protein molecule is at least 99% identical to one of SEQ ID NOs:4, 7, 10, 13, 16, 19, 22, and 27, with the proviso that the veterinary decorin core protein comprises a mutation at amino acid 4.

6. The veterinary decorin core protein molecule of claim 1, wherein said protein molecule is 100% identical to one of SEQ ID NOs:4, 7, 10, 13, 16, 19, 22, and 27, with the proviso that the veterinary decorin core protein comprises a mutation at amino acid 4.

7. The veterinary decorin core protein molecule of claim 1, wherein said core molecule is operably linked to an exogenous signal peptide.

8. The veterinary decorin core protein molecule of claim 7, wherein said exogenous signal peptide is a bovine lactalbumin signal peptide.

9. A composition comprising a veterinary decorin core protein molecule according to claim 1 in combination with a pharmaceutically acceptable carrier.

10. An expression vector comprising a nucleic acid sequence encoding a veterinary decorin core molecule according to claim 1.

11. The expression vector of claim 10, wherein said nucleic acid sequence encoding a veterinary decorin core molecule is operably associated with an exogenous signal peptide.

12. The expression vector of claim 11, wherein said exogenous signal peptide is a bovine lactalbumin signal peptide.

13. A host cell comprising the expression vector of claim 10.

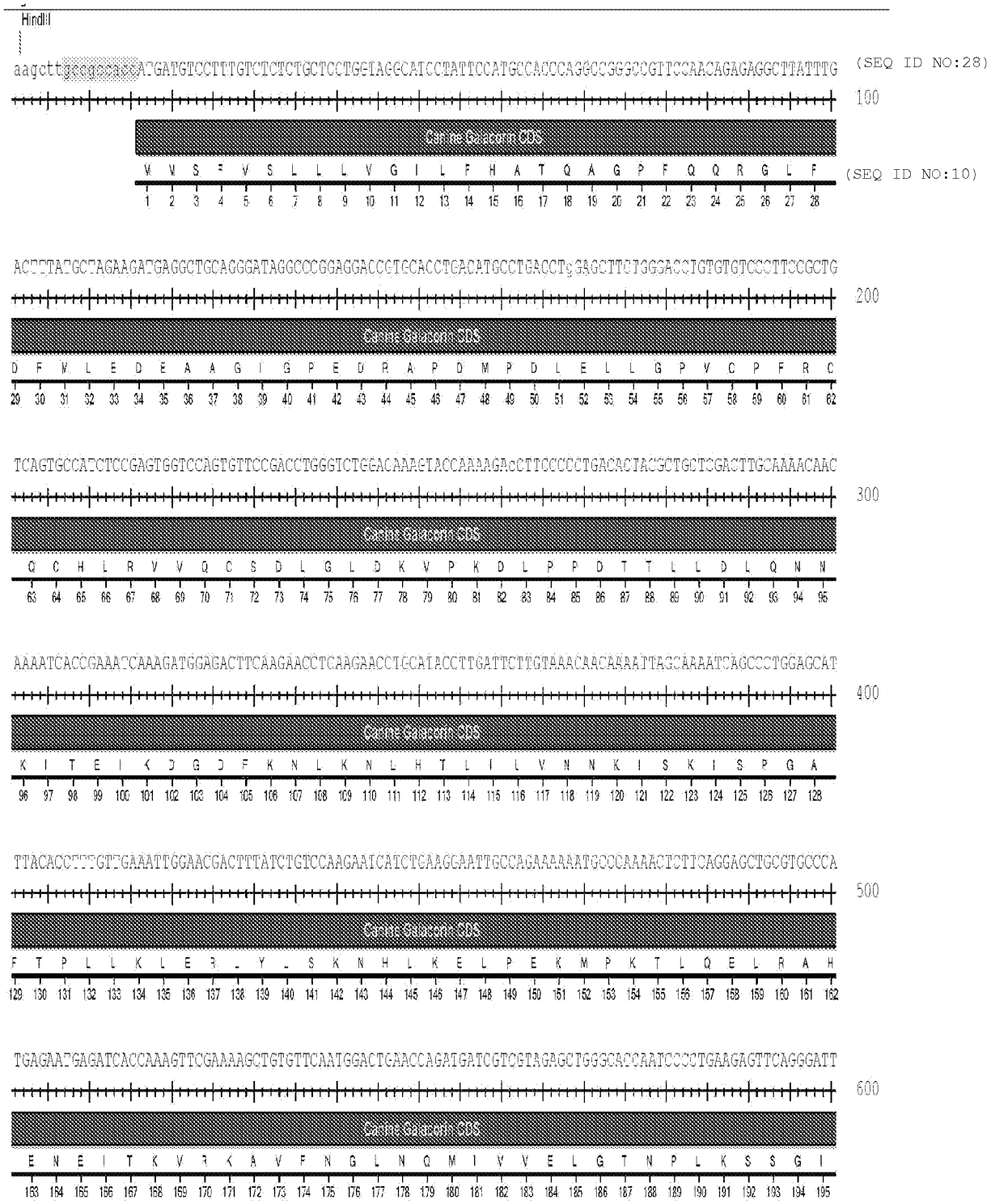
14. A method of producing a veterinary decorin protein comprising expressing the expression vector of claim 10 in a host cell to produce said veterinary decorin protein and purifying said veterinary decorin protein.

15. A pharmaceutical formulation comprising a veterinary decorin core protein molecule according to claim 1, wherein said formulation is a liquid, powder, spray, gel, ointment, lotion, or eye drop.

16. A veterinary decorin protein produced by the method of claim 14.

Catalent Pharma Solutions, LLC
Patent Attorneys for the Applicant/Nominated Person
SPRUSON & FERGUSON

FIG. 1


HindIII
(SEQ ID NO:28)
100
Canine Galaconin CDS
(SEQ ID NO:10)
200
300
400
500
600

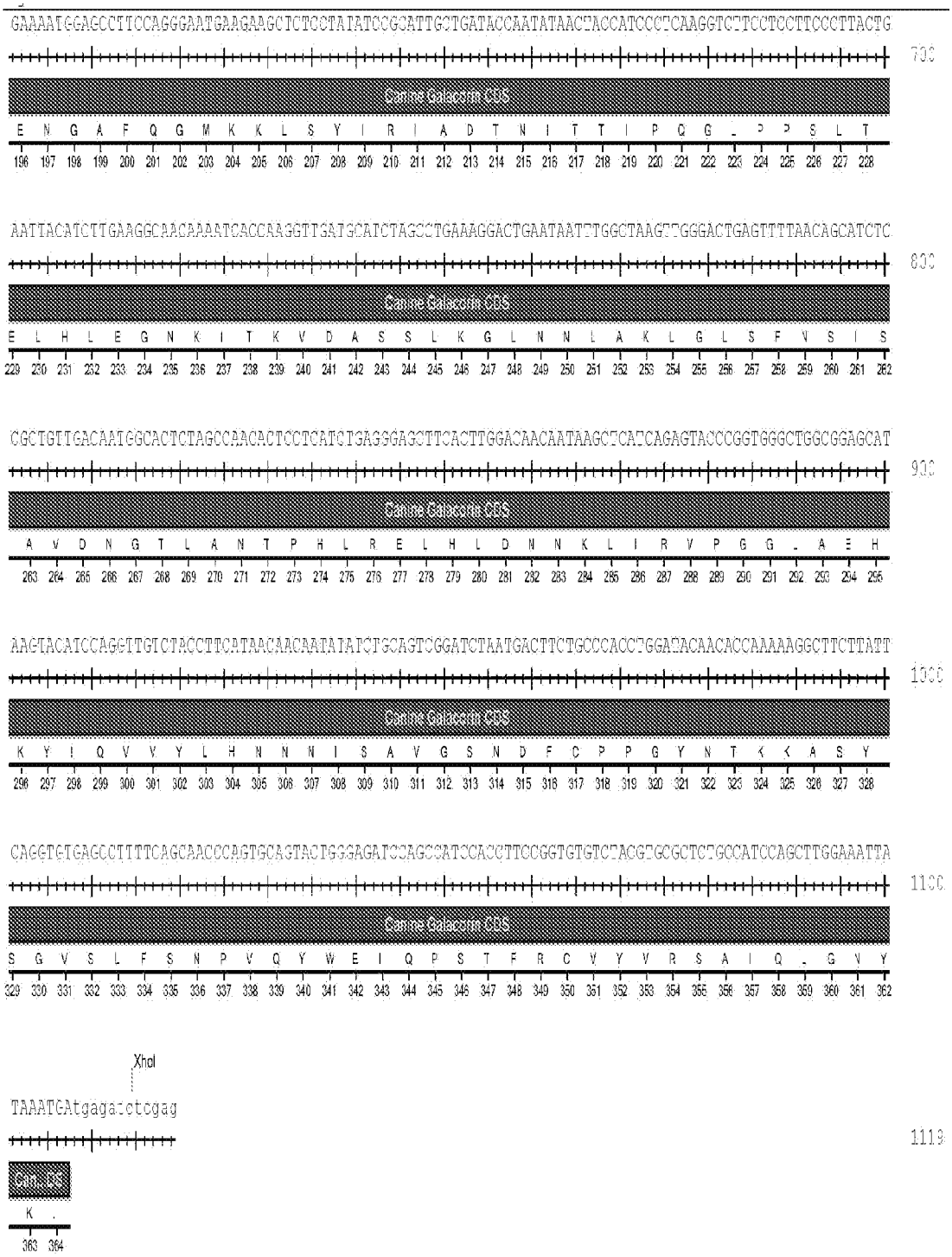

GAAAATGGAGCCTTCCAGGGAATGAAGAAGCTCTCCTATATCCGCATTGCTGATACCAATATAACCTACCATCCCTCAAGGTCTTCCTCCTTCCCTTACTG
700
Canine Calsarcin CDS
E N G A F Q G M K K L S Y I R I A D T N I T T I P Q G _ P P S L T
196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228
AATTACATCTTGAAGGCAACAAAATCACCAAGGTTGATGCATCTAGCCTGAAAGGACTGAATAATCTGGCTAAGCTCGGGACTGAGTTTTAACAGCATCTC
800
Canine Calsarcin CDS
E L H L E G N K I T K V D A S S L K G L N N L A K L G L S F N S I S
229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262
CGCTGTTGACAATGGCACTCTAGCCAACACTCCTCATCTGAGGGAGCTTCACTTGGACAACAATAAGCTCATCAGAGTACCCGGTGGGCTGGCGGAGCAT
900
Canine Calsarcin CDS
A V D N G T L A N T P H L R E L H L D N N K L I R V P G G _ A E H
263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295
AAGTACATCCAGGTTGTCTACCTTCATAACAACAATATATCTGCAGTCGGATCTAATGACTTCTGCCCACCTGGATACAACACCAAAAAGGCTTCTTATT
1000
Canine Calsarcin CDS
K Y I Q V V Y L H N N N I S A V G S N D F C P P G Y N T K K A S Y
296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328
CAGGTGTGAGCCTTTTCAGCAACCCAGTGCAGTACTGGGAGATCCAGCCATCCACCTTCCGGTGTGTCTACGTGCGCTCTGCCATCCAGCTTGGAAATTA
1100
Canine Calsarcin CDS
S G V S L F S N P V Q Y W E I Q P S T F R C V Y V R S A I Q _ G N Y
329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362
XhoI
TAAATGAtgagatctcgag
1119
Calm DS
K .
363 364

FIG. 1 (cont.)

FIG. 2

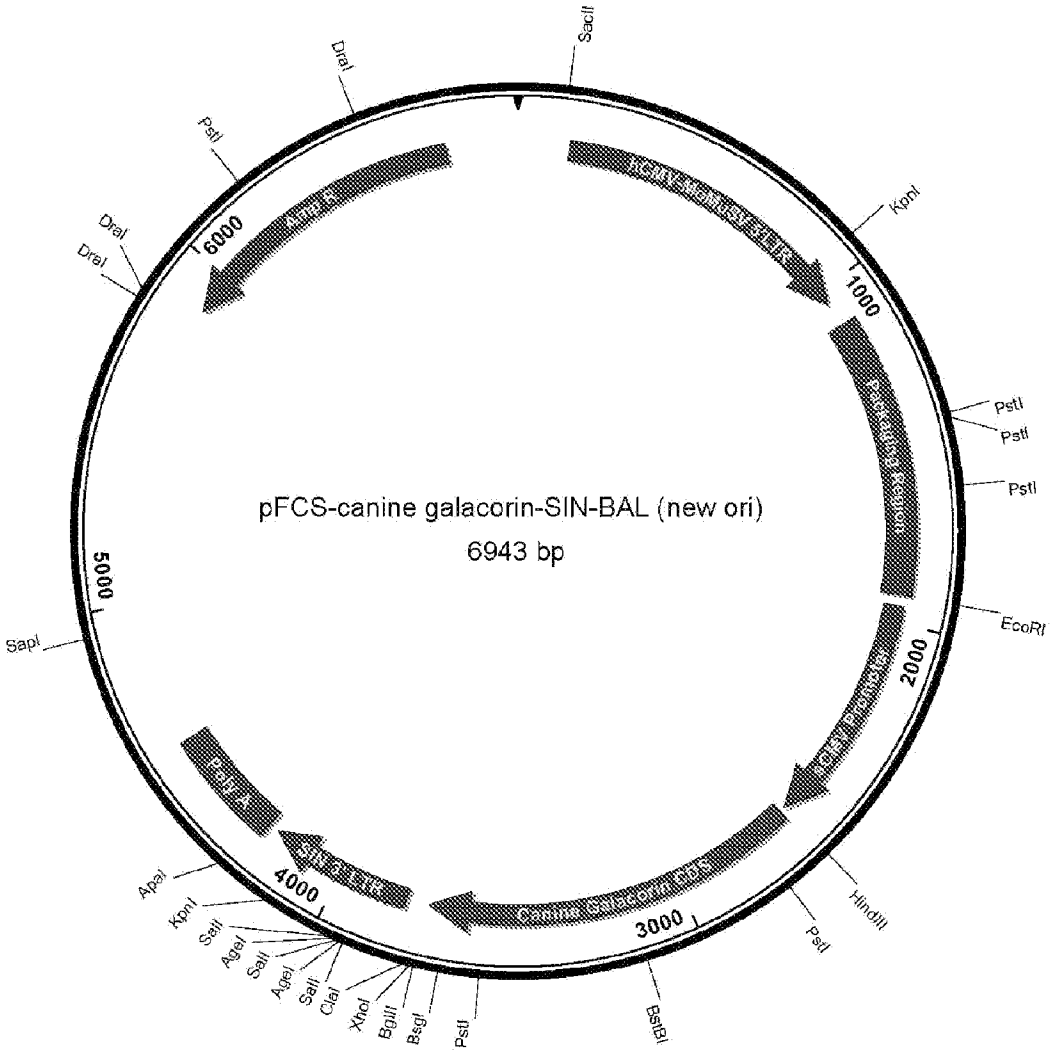
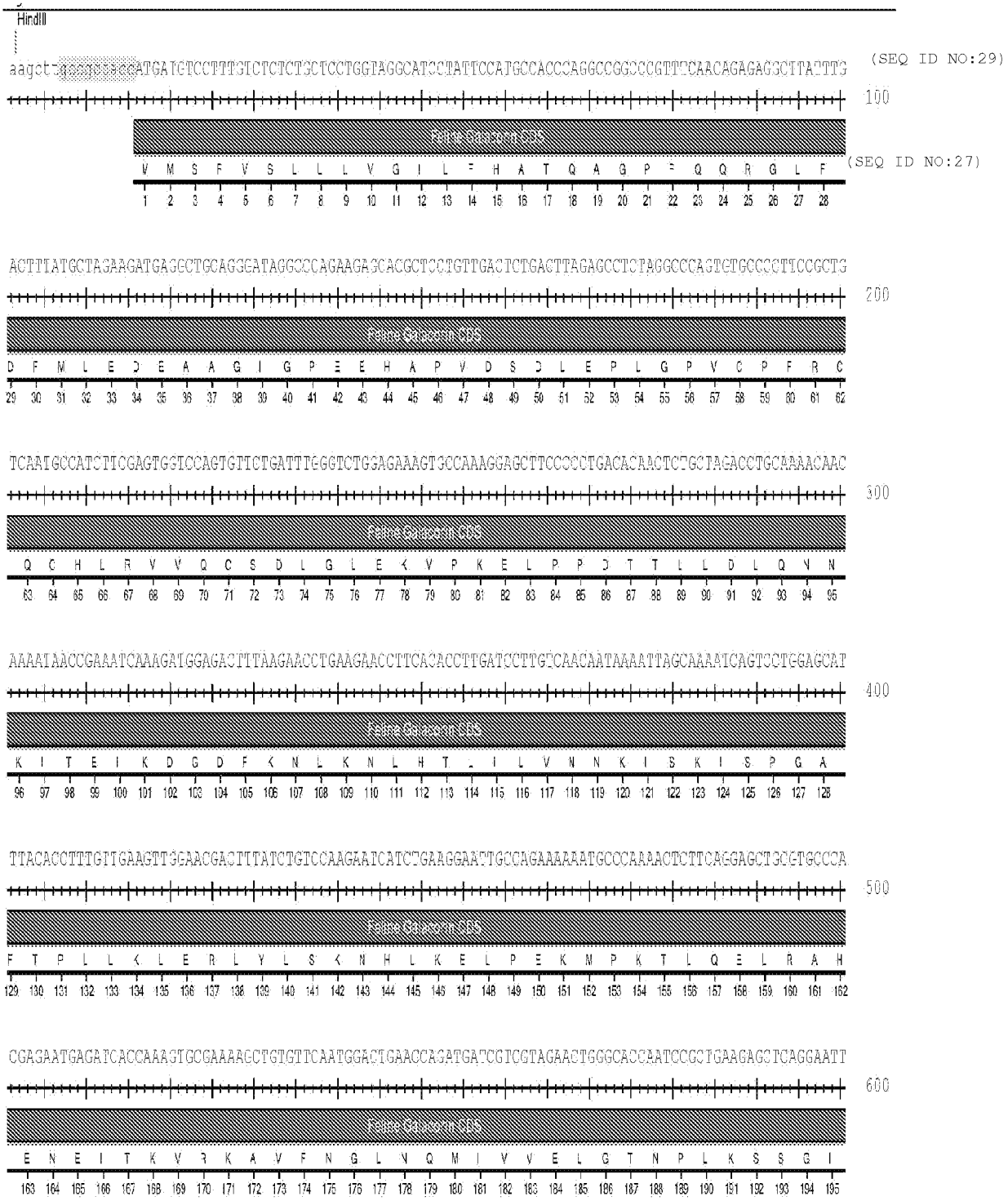

pFCS-canine galacorin-SIN-BAL (new ori)
6943 bp
SacII
KpnI
PstI
PstI
PstI
EcoRI
HindIII
PstI
BstBI
PstI
BglII
BsgI
XhoI
ClaI
SalI
AgeI
SalI
AgeI
SalI
KpnI
ApaI
SapI
DraI
DraI
DraI
PstI
1000
2000
3000
4000
5000
6000

FIG. 3



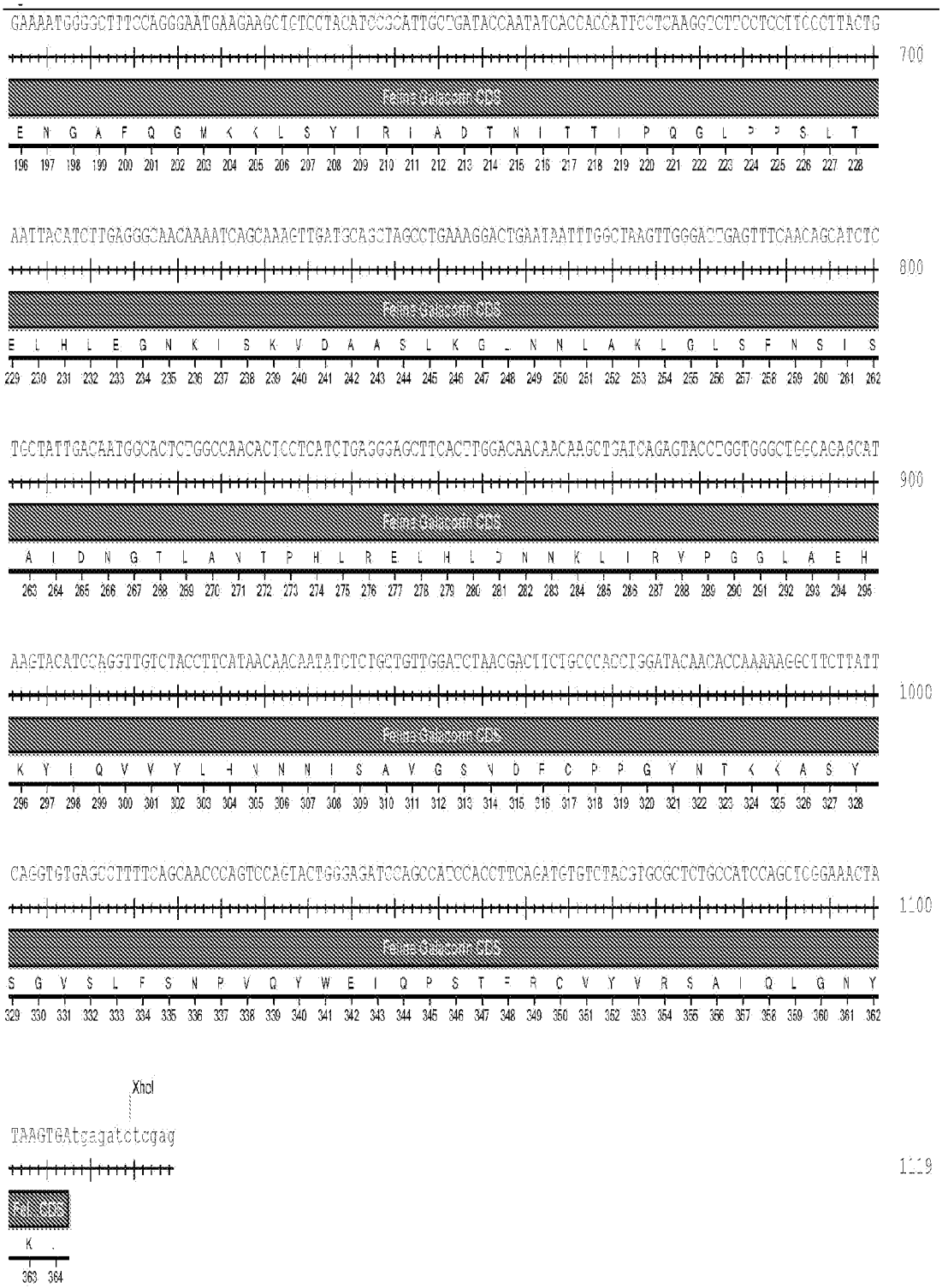

GAAAATGGGGCTTTCCAGGGAATGAAGAAGCTGTCCTACATCGGCATTGCTGATACCAATATCACCACCATTCCTCAAGGTCTTCCTCCTTCCCTTACTG
700
Penta-Stilbocarin CDS
E N G A F Q G M K K L S Y I R I A D T N I T T I P Q G L P P S L T
196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228
AATTACATCTTGAGGGCAACAAAATCAGCAAAGTTGATGCAGCTAGCCTGAAAGGACTGAATAATTTGGCTAAGTTGGGACTGAGTTTCAACAGCATCTC
800
Penta-Stilbocarin CDS
E L H L E G N K I S K V D A A S L K G L N N L A K L G L S F N S I S
229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262
TGCTATTGACAATGGCACTCTGGCCAACACTCCTCATCTGAGGGAGCTTCACTTGGACAACAACAAGCTGATCAGAGTACCTGGTGGGCTGGCAGAGCAT
900
Penta-Stilbocarin CDS
A I D N G T L A N T P H L R E L H L D N N K L I R V P G G L A E H
263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295
AAGTACATCCAGGTTGTCTACCTTCATAACAACAATATCTCTGCTGTTGGATCTAACGACTTCTGCCCACCTGGATACAACACCAAAAAGGCTTCTTATT
1000
Penta-Stilbocarin CDS
K Y I Q V V Y L H N N N I S A V G S N D F C P P G Y N T K K A S Y
296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328
CAGGTGTGAGCCTTTTCAGCAACCCAGTCCAGTACTGGGAGATCCAGCCATCCACCTTCAGATGTGTCTACGTGCGCTCTGCCATCCAGCTCGGAAACTA
1100
Penta-Stilbocarin CDS
S G V S L F S N P V Q Y W E I Q P S T F R C V Y V R S A I Q L G N Y
329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362
XhoI
TAAGTGAtgagatctcgag
1119
Pal cds
K .
363 364

FIG. 3 (cont.)

FIG. 4

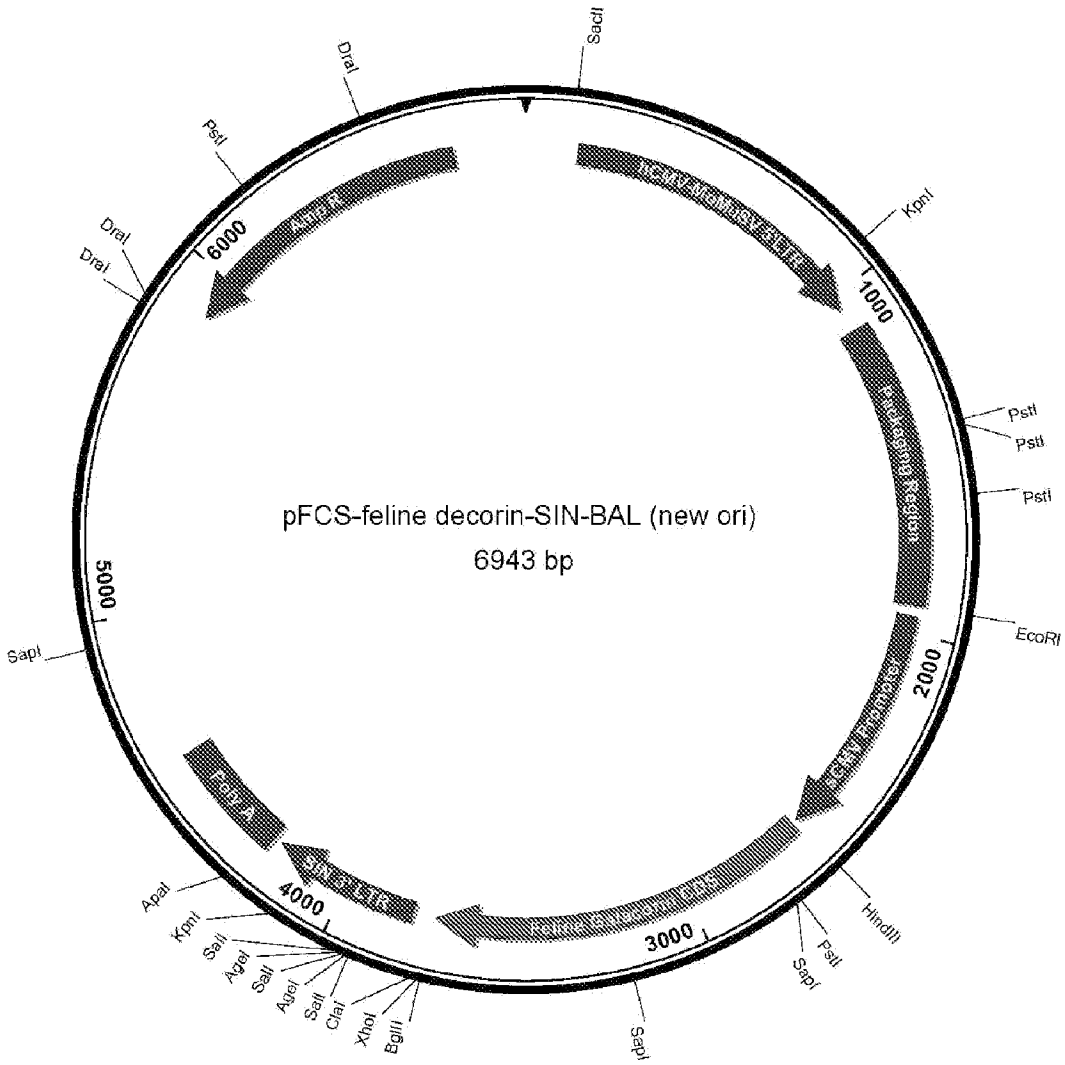

pFCS-feline decorin-SIN-BAL (new ori)
6943 bp
SacII
KpnI
PstI
PstI
PstI
EcoRI
HindIII
PstI
SapI
SapI
BglII
XhoI
ClaI
SalI
AgeI
SalI
AgeI
SalI
KpnI
ApaI
SapI
DraI
DraI
DraI
PstI
1000
2000
3000
4000
5000
6000

FIG. 5

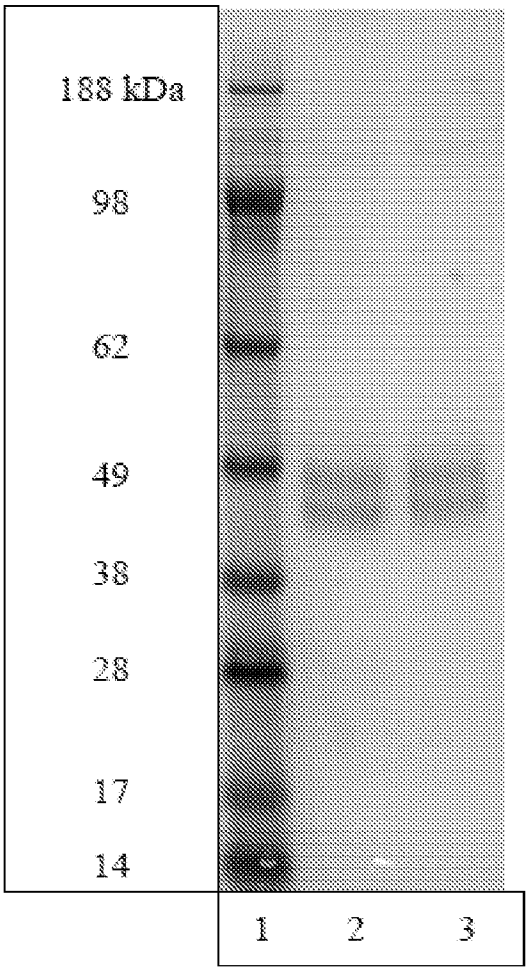


FIG. 6

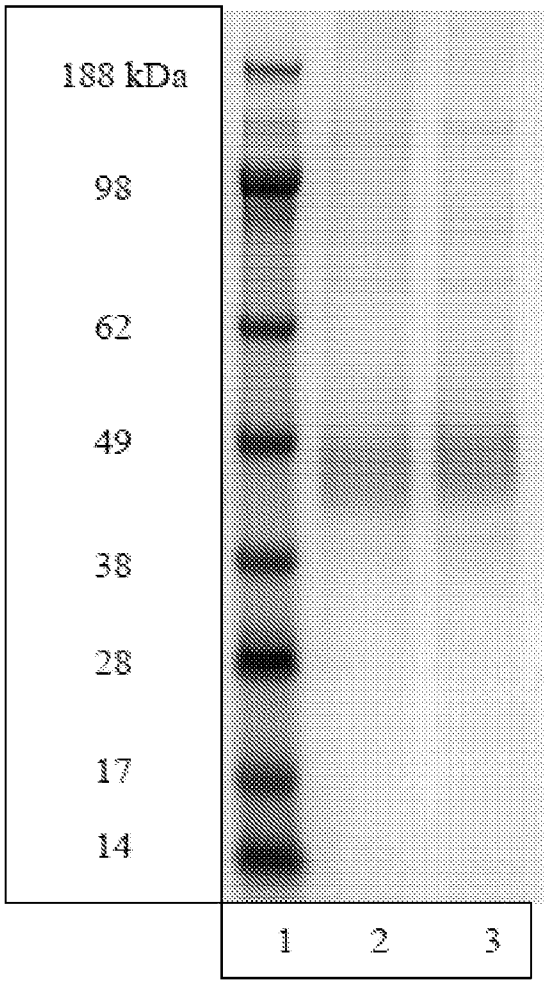
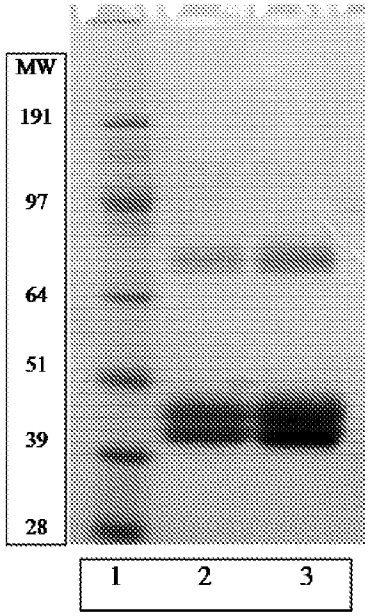


FIG. 7



SEQUENCE LISTING

<110> Catalent Pharma Solutions, LLC

<120> VETERINARY DECORIN COMPOSITIONS AND USE THEREOF

<130> GALA-33129/WO-1/ORD

<150> US 61/814,405
<151> 2013-04-22

<160> 29

<170> PatentIn version 3.5

<210> 1
<211> 1092
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 1
atgatgtcct ttgtctctct gctcctggta ggcatcctat tccatgccac ccaggccgga 60

ccatttcaac agaaaggctt atttgacttt atgctggaag atgaggctgc agggataggc 120

ccggaagagc actttcctga agttcctgaa atagagccta tgggcccagt ctgccccttc 180

cgctgtcagt gccatctgcg agttgtccag tgttctgatc tgggtctgga aaaagtacca 240

aaagaccttc ctcctgatac tgcgctgctg gacctgcaaa acaacaaaat aactgagatc 300

aaagatggag actttaagaa cctgaagaac cttcatacac tgattctcat caacaacaaa 360

attagcaaaa tcagccctgg ggcatttgct cctttggtga aattggaacg actttatctt 420

tccaagaatc aactgaagga attgccagag aaaatgccca aaactcttca ggagctgcgt 480

gtccatgaga acgagatcac caaagtgcga aagtctgtgt tcaatggatt gaaccagatg 540

atcgtcgtag aacttggcac caacccgctg aagagctcag gcattgaaaa tggagccttt 600

cagggaatga agaagctctc ctacatccgc attgctgaca caaatataac taccatccct 660

caaggtcttc ctccttccct tactgaatta catctcgatg gcaacaaaat caccaaagtt 720

gatgcagcta gcctgaaagg actgaataat ttggctaagt tgggactgag tttcaacagc 780

atctctgcgg ttgacaatgg ctctttggcc aacactcctc atttgaggga acttcatttg 840

aacaacaaca agcttgtcaa agtgcccggt gggctggccg atcataagta catccaggtt 900

gtctaccttc acaacaacaa tatctctgca atcggctcta acgacttctg cccacccgga 960

tacaacacca aaaaggcttc ttattcagga gtgagccttt tcagcaaccc agtccagtac 1020

tgggagatcc agccatccac cttccgatgt gtctatgtgc gtgctgccgt tcagcttgga 1080

aactacaagt ga 1092

<210> 2
<211> 1083
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 2
atgatgtcct ttgtctctct gctcctggta ggcatcctat tccatgccac ccaggccacg 60

cggttccacc agaagggcct ctttgacttt atgatagagg atgaaggggc agccgacatg 120

gctccaacag atgatcctgt catatctgga tttgggccag tgtgcccctt ccgctgccag 180

tgtcatcttc gcgttgtgca gtgctctgac ctaggtctgg aaagagtgcc aaaagacctt 240

ccccctgaca caactctgct ggatttacag aacaacaaaa tcactgaaat taaagaagga 300

gatttcaaga atttgaagaa tcttcatgca ttgatccttg ttaacaacaa aatcagcaaa 360

ataagtccgg cagcttttgc tcctctgaag aaactggaaa gactgtacct atccaagaat 420

aatttgaagg aacttccaga aaacatgcca aagtctcttc aggagatacg tgctcatgaa 480

aatgagatct ccaagttgag gaaggcagtt tttaatggac tgaatcaagt gattgtctta 540

gaactaggca ccaatccact caagagctca ggcattgaaa atggagcttt tcaagggatg 600

aagaggcttt cctatatccg catcgcagac accaacatta ctagcatccc taaaggtctt 660

cctccatccc ttactgagct tcaccttgat ggcaacaaaa ttagcaaaat tgatgcggaa 720

ggtctgtctg gactcaccaa cttggctaaa ttgggtctca gcttcaacag tatttcttct 780

gttgaaaatg gctctctgaa caatgtacct catctgagag aacttcatct gaataacaac 840

gaacttgtca gagtacctag tgggttgggt gaacacaaat acatccaggt ggtctatctt 900

cataacaaca agattgcttc aattggtatc aacgactttt gccctcttgg ctacaacacc 960

aaaaaggcaa cctattctgg tgtgagtctc ttcagcaacc ccgtgcagta ctgggaaatc 1020

cagccctctg ctttccgatg tatccatgaa cgctctgcag tacagatcgg aaattacaaa 1080

tga 1083

<210> 3
<211> 360
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 3

Met Met Ser Phe Val Ser Leu Leu Leu Val Gly Ile Leu Phe His Ala
1 5 10 15

Thr Gln Ala Thr Arg Phe His Gln Lys Gly Leu Phe Asp Phe Met Ile
20 25 30

Glu Asp Glu Gly Ala Ala Asp Met Ala Pro Thr Asp Asp Pro Val Ile
35 40 45

Ser Gly Phe Gly Pro Val Cys Pro Phe Arg Cys Gln Cys His Leu Arg
50 55 60

```
Val Val Gln Cys Ser Asp Leu Gly Leu Glu Arg Val Pro Lys Asp Leu
65                  70                  75                  80

Pro Pro Asp Thr Thr Leu Leu Asp Leu Gln Asn Asn Lys Ile Thr Glu
                85                  90                  95

Ile Lys Glu Gly Asp Phe Lys Asn Leu Lys Asn Leu His Ala Leu Ile
            100                 105                 110

Leu Val Asn Asn Lys Ile Ser Lys Ile Ser Pro Ala Ala Phe Ala Pro
        115                 120                 125

Leu Lys Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn Asn Leu Lys Glu
    130                 135                 140

Leu Pro Glu Asn Met Pro Lys Ser Leu Gln Glu Ile Arg Ala His Glu
145                 150                 155                 160

Asn Glu Ile Ser Lys Leu Arg Lys Ala Val Phe Asn Gly Leu Asn Gln
                165                 170                 175

Val Ile Val Leu Glu Leu Gly Thr Asn Pro Leu Lys Ser Ser Gly Ile
            180                 185                 190

Glu Asn Gly Ala Phe Gln Gly Met Lys Arg Leu Ser Tyr Ile Arg Ile
        195                 200                 205

Ala Asp Thr Asn Ile Thr Ser Ile Pro Lys Gly Leu Pro Pro Ser Leu
    210                 215                 220

Thr Glu Leu His Leu Asp Gly Asn Lys Ile Ser Lys Ile Asp Ala Glu
225                 230                 235                 240

Gly Leu Ser Gly Leu Thr Asn Leu Ala Lys Leu Gly Leu Ser Phe Asn
                245                 250                 255

Ser Ile Ser Ser Val Glu Asn Gly Ser Leu Asn Asn Val Pro His Leu
            260                 265                 270

Arg Glu Leu His Leu Asn Asn Asn Glu Leu Val Arg Val Pro Ser Gly
        275                 280                 285

Leu Gly Glu His Lys Tyr Ile Gln Val Val Tyr Leu His Asn Asn Lys
    290                 295                 300

Ile Ala Ser Ile Gly Ile Asn Asp Phe Cys Pro Leu Gly Tyr Asn Thr
305                 310                 315                 320

Lys Lys Ala Thr Tyr Ser Gly Val Ser Leu Phe Ser Asn Pro Val Gln
                325                 330                 335
```

```
Tyr Trp Glu Ile Gln Pro Ser Ala Phe Arg Cys Ile His Glu Arg Ser
            340                 345                 350


Ala Val Gln Ile Gly Asn Tyr Lys
        355                 360


<210>  4
<211>  327
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  4

Asp Glu Gly Ala Ala Asp Met Ala Pro Thr Asp Asp Pro Val Ile Ser
1               5                   10                  15


Gly Phe Gly Pro Val Cys Pro Phe Arg Cys Gln Cys His Leu Arg Val
            20                  25                  30


Val Gln Cys Ser Asp Leu Gly Leu Glu Arg Val Pro Lys Asp Leu Pro
        35                  40                  45


Pro Asp Thr Thr Leu Leu Asp Leu Gln Asn Asn Lys Ile Thr Glu Ile
    50                  55                  60


Lys Glu Gly Asp Phe Lys Asn Leu Lys Asn Leu His Ala Leu Ile Leu
65                  70                  75                  80


Val Asn Asn Lys Ile Ser Lys Ile Ser Pro Ala Ala Phe Ala Pro Leu
                85                  90                  95


Lys Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn Asn Leu Lys Glu Leu
            100                 105                 110


Pro Glu Asn Met Pro Lys Ser Leu Gln Glu Ile Arg Ala His Glu Asn
        115                 120                 125


Glu Ile Ser Lys Leu Arg Lys Ala Val Phe Asn Gly Leu Asn Gln Val
    130                 135                 140


Ile Val Leu Glu Leu Gly Thr Asn Pro Leu Lys Ser Ser Gly Ile Glu
145                 150                 155                 160


Asn Gly Ala Phe Gln Gly Met Lys Arg Leu Ser Tyr Ile Arg Ile Ala
                165                 170                 175


Asp Thr Asn Ile Thr Ser Ile Pro Lys Gly Leu Pro Pro Ser Leu Thr
            180                 185                 190


Glu Leu His Leu Asp Gly Asn Lys Ile Ser Lys Ile Asp Ala Glu Gly
```

195 200 205

Leu Ser Gly Leu Thr Asn Leu Ala Lys Leu Gly Leu Ser Phe Asn Ser
210 215 220

Ile Ser Ser Val Glu Asn Gly Ser Leu Asn Asn Val Pro His Leu Arg
225 230 235 240

Glu Leu His Leu Asn Asn Asn Glu Leu Val Arg Val Pro Ser Gly Leu
245 250 255

Gly Glu His Lys Tyr Ile Gln Val Val Tyr Leu His Asn Asn Lys Ile
260 265 270

Ala Ser Ile Gly Ile Asn Asp Phe Cys Pro Leu Gly Tyr Asn Thr Lys
275 280 285

Lys Ala Thr Tyr Ser Gly Val Ser Leu Phe Ser Asn Pro Val Gln Tyr
290 295 300

Trp Glu Ile Gln Pro Ser Ala Phe Arg Cys Ile His Glu Arg Ser Ala
305 310 315 320

Val Gln Ile Gly Asn Tyr Lys
325

<210> 5
<211> 1092
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 5
atgatgtcct ttgtctctct gctcctggta ggcatcctat tccatgccac ccaggccgga 60

ccatttcaac agaaaggctt atttgacttt atgctggaag atgaggctgc agggataggc 120

ccggaagagc actttcctga agttcctgaa atagagccta tgggcccagt ctgccccttc 180

cgctgtcagt gccatctgcg agttgtccag tgttctgatc tgggtctgga aaaagtacca 240

aaagaccttc ctcctgatac tgcgctgctg gacctgcaaa acaacaaaat aactgagatc 300

aaagatggag actttaagaa cctgaagaac cttcatacac tgattctcat caacaacaaa 360

attagcaaaa tcagccctgg ggcatttgct cctttggtga aattggaacg actttatctt 420

tccaagaatc aactgaagga attgccagag aaaatgccca aaactcttca ggagctgcgt 480

gtccatgaga acgagatcac caaagtgcga aagtctgtgt tcaatggatt gaaccagatg 540

atcgtcgtag aacttggcac caacccgctg aagagctcag gcattgaaaa tggagccttt 600

cagggaatga agaagctctc ctacatccgc attgctgaca caaatataac taccatccct 660

caaggtcttc ctccttccct tactgaatta catctcgatg gcaacaaaat caccaaagtt 720

gatgcagcta gcctgaaagg actgaataat ttggctaagt tgggactgag tttcaacagc 780

atctctgcgg ttgacaatgg ctctttggcc aacactcctc atttgaggga acttcatttg 840

aacaacaaca agcttgtcaa agtgcccggt gggctggccg atcataagta catccaggtt 900

gtctaccttc acaacaacaa tatctctgca atcggctcta acgacttctg cccacccgga 960

tacaacacca aaaaggcttc ttattcagga gtgagccttt tcagcaaccc agtccagtac 1020

tgggagatcc agccatccac cttccgatgt gtctatgtgc gtgctgccgt tcagcttgga 1080

aactacaagt ga 1092

<210> 6
<211> 363
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 6

Met Met Ser Phe Val Ser Leu Leu Leu Val Gly Ile Leu Phe His Ala
1 5 10 15

Thr Gln Ala Gly Pro Phe Gln Gln Lys Gly Leu Phe Asp Phe Met Leu
20 25 30

Glu Asp Glu Ala Ala Gly Ile Gly Pro Glu Glu His Phe Pro Glu Val
35 40 45

Pro Glu Ile Glu Pro Met Gly Pro Val Cys Pro Phe Arg Cys Gln Cys
50 55 60

His Leu Arg Val Val Gln Cys Ser Asp Leu Gly Leu Glu Lys Val Pro
65 70 75 80

Lys Asp Leu Pro Pro Asp Thr Ala Leu Leu Asp Leu Gln Asn Asn Lys
85 90 95

Ile Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His
100 105 110

Thr Leu Ile Leu Ile Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala
115 120 125

Phe Ala Pro Leu Val Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn Gln
130 135 140

Leu Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg
145 150 155 160

Val His Glu Asn Glu Ile Thr Lys Val Arg Lys Ser Val Phe Asn Gly
165 170 175

```
Leu Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser
            180                 185                 190

Ser Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr
        195                 200                 205

Ile Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro
    210                 215                 220

Pro Ser Leu Thr Glu Leu His Leu Asp Gly Asn Lys Ile Thr Lys Val
225                 230                 235                 240

Asp Ala Ala Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu
                245                 250                 255

Ser Phe Asn Ser Ile Ser Ala Val Asp Asn Gly Ser Leu Ala Asn Thr
            260                 265                 270

Pro His Leu Arg Glu Leu His Leu Asn Asn Asn Lys Leu Val Lys Val
        275                 280                 285

Pro Gly Gly Leu Ala Asp His Lys Tyr Ile Gln Val Val Tyr Leu His
    290                 295                 300

Asn Asn Asn Ile Ser Ala Ile Gly Ser Asn Asp Phe Cys Pro Pro Gly
305                 310                 315                 320

Tyr Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn
                325                 330                 335

Pro Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr
            340                 345                 350

Val Arg Ala Ala Val Gln Leu Gly Asn Tyr Lys
        355                 360


<210>  7
<211>  329
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  7

Glu Ala Ala Gly Ile Gly Pro Glu Glu His Phe Pro Glu Val Pro Glu
1               5                   10                  15

Ile Glu Pro Met Gly Pro Val Cys Pro Phe Arg Cys Gln Cys His Leu
            20                  25                  30

Arg Val Val Gln Cys Ser Asp Leu Gly Leu Glu Lys Val Pro Lys Asp
```

```
        35                  40                  45

Leu Pro Pro Asp Thr Ala Leu Leu Asp Leu Gln Asn Asn Lys Ile Thr
    50                  55                  60

Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His Thr Leu
65                  70                  75                  80

Ile Leu Ile Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala Phe Ala
                85                  90                  95

Pro Leu Val Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn Gln Leu Lys
            100                 105                 110

Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg Val His
        115                 120                 125

Glu Asn Glu Ile Thr Lys Val Arg Lys Ser Val Phe Asn Gly Leu Asn
    130                 135                 140

Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser Ser Gly
145                 150                 155                 160

Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr Ile Arg
                165                 170                 175

Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro Pro Ser
            180                 185                 190

Leu Thr Glu Leu His Leu Asp Gly Asn Lys Ile Thr Lys Val Asp Ala
        195                 200                 205

Ala Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu Ser Phe
    210                 215                 220

Asn Ser Ile Ser Ala Val Asp Asn Gly Ser Leu Ala Asn Thr Pro His
225                 230                 235                 240

Leu Arg Glu Leu His Leu Asn Asn Asn Lys Leu Val Lys Val Pro Gly
                245                 250                 255

Gly Leu Ala Asp His Lys Tyr Ile Gln Val Val Tyr Leu His Asn Asn
            260                 265                 270

Asn Ile Ser Ala Ile Gly Ser Asn Asp Phe Cys Pro Pro Gly Tyr Asn
        275                 280                 285

Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn Pro Val
    290                 295                 300

Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr Val Arg
```

305 310 315 320

Ala Ala Val Gln Leu Gly Asn Tyr Lys
325

<210> 8
<211> 1092
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 8
atgatgtcct ttgtctctct gctcctggta ggcatcctat tccatgccac ccaggccggg 60

ccgttccaac agagaggctt atttgacttt atgctagaag atgaggctgc agggataggc 120

ccggaggacc gtgcacctga catgcctgac ctcgagcttc tgggacctgt gtgtcccttc 180

cgctgtcagt gccatctccg agtggtccag tgttccgacc tgggtctgga caaagtacca 240

aaagatcttc cccctgacac tacgctgctc gacttgcaaa acaacaaaat caccgaaatc 300

aaagatggag acttcaagaa cctcaagaac ctgcatacct tgattcttgt aaacaacaaa 360

attagcaaaa tcagccctgg agcatttaca cctttgttga aattggaacg actttatctg 420

tccaagaatc atctgaagga attgccagaa aaaatgccca aaactcttca ggagctgcgt 480

gcccatgaga atgagatcac caaagttcga aaagctgtgt tcaatggact gaaccagatg 540

atcgtcgtag agctgggcac caatcccctg aagagttcag ggattgaaaa tggagccttc 600

cagggaatga agaagctctc ctatatccgc attgctgata ccaatataac taccatccct 660

caaggtcttc ctccttccct tactgaatta catcttgaag gcaacaaaat caccaaggtt 720

gatgcatcta gcctgaaagg actgaataat ttggctaagt tgggactgag ttttaacagc 780

atctccgctg ttgacaatgg cactctagcc aacactcctc atctgaggga gcttcacttg 840

gacaacaata agctcatcag agtacccggt gggctggcgg agcataagta catccaggtt 900

gtctaccttc ataacaacaa tatatctgca gtcggatcta atgacttctg cccacctgga 960

tacaacacca aaaaggcttc ttattcaggt gtgagccttt tcagcaaccc agtgcagtac 1020

tgggagatcc agccatccac cttccggtgt gtctacgtgc gctctgccat ccagcttgga 1080

aattataaat ga 1092

<210> 9
<211> 363
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 9

Met Met Ser Phe Val Ser Leu Leu Leu Val Gly Ile Leu Phe His Ala
1 5 10 15

```
Thr Gln Ala Gly Pro Phe Gln Gln Arg Gly Leu Phe Asp Phe Met Leu
            20                  25                  30

Glu Asp Glu Ala Ala Gly Ile Gly Pro Glu Asp Arg Ala Pro Asp Met
        35                  40                  45

Pro Asp Leu Glu Leu Leu Gly Pro Val Cys Pro Phe Arg Cys Gln Cys
    50                  55                  60

His Leu Arg Val Val Gln Cys Ser Asp Leu Gly Leu Asp Lys Val Pro
65                  70                  75                  80

Lys Asp Leu Pro Pro Asp Thr Thr Leu Leu Asp Leu Gln Asn Asn Lys
                85                  90                  95

Ile Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His
            100                 105                 110

Thr Leu Ile Leu Val Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala
        115                 120                 125

Phe Thr Pro Leu Leu Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn His
    130                 135                 140

Leu Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg
145                 150                 155                 160

Ala His Glu Asn Glu Ile Thr Lys Val Arg Lys Ala Val Phe Asn Gly
                165                 170                 175

Leu Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser
            180                 185                 190

Ser Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr
        195                 200                 205

Ile Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro
    210                 215                 220

Pro Ser Leu Thr Glu Leu His Leu Glu Gly Asn Lys Ile Thr Lys Val
225                 230                 235                 240

Asp Ala Ser Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu
                245                 250                 255

Ser Phe Asn Ser Ile Ser Ala Val Asp Asn Gly Thr Leu Ala Asn Thr
            260                 265                 270

Pro His Leu Arg Glu Leu His Leu Asp Asn Asn Lys Leu Ile Arg Val
        275                 280                 285
```

```
Pro Gly Gly Leu Ala Glu His Lys Tyr Ile Gln Val Val Tyr Leu His
    290                 295                 300

Asn Asn Asn Ile Ser Ala Val Gly Ser Asn Asp Phe Cys Pro Pro Gly
305                 310                 315                 320

Tyr Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn
                325                 330                 335

Pro Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr
            340                 345                 350

Val Arg Ser Ala Ile Gln Leu Gly Asn Tyr Lys
        355                 360


<210>  10
<211>  330
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  10

Asp Glu Ala Ala Gly Ile Gly Pro Glu Asp Arg Ala Pro Asp Met Pro
1               5                   10                  15

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

Leu Arg Val Val Gln Cys Ser Asp Leu Gly Leu Asp Lys Val Pro Lys
        35                  40                  45

Asp Leu Pro Pro Asp Thr Thr Leu Leu Asp Leu Gln Asn Asn Lys Ile
    50                  55                  60

Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His Thr
65                  70                  75                  80

Leu Ile Leu Val Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala Phe
                85                  90                  95

Thr Pro Leu Leu Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn His Leu
            100                 105                 110

Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg Ala
        115                 120                 125

His Glu Asn Glu Ile Thr Lys Val Arg Lys Ala Val Phe Asn Gly Leu
    130                 135                 140

Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser Ser
```

145 150 155 160

Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr Ile
165 170 175

Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro Pro
180 185 190

Ser Leu Thr Glu Leu His Leu Glu Gly Asn Lys Ile Thr Lys Val Asp
195 200 205

Ala Ser Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu Ser
210 215 220

Phe Asn Ser Ile Ser Ala Val Asp Asn Gly Thr Leu Ala Asn Thr Pro
225 230 235 240

His Leu Arg Glu Leu His Leu Asp Asn Asn Lys Leu Ile Arg Val Pro
245 250 255

Gly Gly Leu Ala Glu His Lys Tyr Ile Gln Val Val Tyr Leu His Asn
260 265 270

Asn Asn Ile Ser Ala Val Gly Ser Asn Asp Phe Cys Pro Pro Gly Tyr
275 280 285

Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn Pro
290 295 300

Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr Val
305 310 315 320

Arg Ser Ala Ile Gln Leu Gly Asn Tyr Lys
325 330

<210> 11
<211> 1092
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 11
atgatgtcct ttgtctctct gctcctggta ggcatcctat tccatgccac ccaggccgga 60

ccgtttcaac agaaaggctt atttgacttt atgctggaag atgaggctgc agggataggc 120

ccggaagagc gctttcatga agttcctgaa ttagagccta tgggcccagt ctgccccttc 180

cgctgtcagt gccatctgcg agttgtccag tgttctgttc tgggtctgga aaaagtgccc 240

aaagaccttc ctcctgatac cgcgctgctg gacctgcaaa acaacaaaat aactgagatc 300

aaagatggag actttaagaa cctgaagaac cttcatacac tgattctcat caacaacaaa 360

attagcaaaa tcagccctgg ggcatttgct cctctggtga aattggaacg actttatctt 420

tccaagaatc aactgaagga attgccagag aaaatgccca aaactcttca ggagctgcgt 480

gtccatgaga acgagatcac caaagtgcga aagtctgtgt tcaatggatt gaaccagatg 540

atcgtcgtag aacttggcac caacccactg aagagctcag gcattgaaaa tggagccttt 600

cagggaatga agaagctctc ctacatccgc attgctgaca ctaatataac taccattcct 660

caaggtcttc ctccttccct tactgaatta catctcgatg gcaacaaaat caccaaagtt 720

gatgcagcta gcctgaaagg actgaataat ttggctaagt tgggactgag tttcaacagc 780

atctctgctg ttgacaatgg ctctttagcc aacactcctc atttgaggga acttcatttg 840

aacaacaaca agcttgtcaa agtgcccggt gggctggccg accataagta catccaggtt 900

gtctaccttc acaacaacaa tatctctgca atcggctcca acgacttctg cccacccgga 960

tacaacacca aaaaggcttc ttattcagga gtgagccttt tcagcaaccc agtccagtac 1020

tgggagatcc agccatccac cttccgatgt gtctacgtgc gcgctgctgt tcagcttgga 1080

aactacaagt ga 1092

<210> 12
<211> 363
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 12

Met Met Ser Phe Val Ser Leu Leu Leu Val Gly Ile Leu Phe His Ala
1 5 10 15

Thr Gln Ala Gly Pro Phe Gln Gln Lys Gly Leu Phe Asp Phe Met Leu
20 25 30

Glu Asp Glu Ala Ala Gly Ile Gly Pro Glu Glu Arg Phe His Glu Val
35 40 45

Pro Glu Leu Glu Pro Met Gly Pro Val Cys Pro Phe Arg Cys Gln Cys
50 55 60

His Leu Arg Val Val Gln Cys Ser Val Leu Gly Leu Glu Lys Val Pro
65 70 75 80

Lys Asp Leu Pro Pro Asp Thr Ala Leu Leu Asp Leu Gln Asn Asn Lys
85 90 95

Ile Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His
100 105 110

Thr Leu Ile Leu Ile Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala
115 120 125

```
Phe Ala Pro Leu Val Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn Gln
    130                 135                 140

Leu Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg
145                 150                 155                 160

Val His Glu Asn Glu Ile Thr Lys Val Arg Lys Ser Val Phe Asn Gly
                165                 170                 175

Leu Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser
            180                 185                 190

Ser Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr
        195                 200                 205

Ile Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro
    210                 215                 220

Pro Ser Leu Thr Glu Leu His Leu Asp Gly Asn Lys Ile Thr Lys Val
225                 230                 235                 240

Asp Ala Ala Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu
                245                 250                 255

Ser Phe Asn Ser Ile Ser Ala Val Asp Asn Gly Ser Leu Ala Asn Thr
            260                 265                 270

Pro His Leu Arg Glu Leu His Leu Asn Asn Asn Lys Leu Val Lys Val
        275                 280                 285

Pro Gly Gly Leu Ala Asp His Lys Tyr Ile Gln Val Val Tyr Leu His
    290                 295                 300

Asn Asn Asn Ile Ser Ala Ile Gly Ser Asn Asp Phe Cys Pro Pro Gly
305                 310                 315                 320

Tyr Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn
                325                 330                 335

Pro Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr
            340                 345                 350

Val Arg Ala Ala Val Gln Leu Gly Asn Tyr Lys
        355                 360
```

<210> 13
<211> 330
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 13

```
Asp Glu Ala Ala Gly Ile Gly Pro Glu Glu Arg Phe His Glu Val Pro
1               5                  10                  15

Glu Leu Glu Pro Met Gly Pro Val Cys Pro Phe Arg Cys Gln Cys His
            20                  25                  30

Leu Arg Val Val Gln Cys Ser Val Leu Gly Leu Glu Lys Val Pro Lys
        35                  40                  45

Asp Leu Pro Pro Asp Thr Ala Leu Leu Asp Leu Gln Asn Asn Lys Ile
    50                  55                  60

Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His Thr
65                  70                  75                  80

Leu Ile Leu Ile Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala Phe
                85                  90                  95

Ala Pro Leu Val Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn Gln Leu
            100                 105                 110

Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg Val
        115                 120                 125

His Glu Asn Glu Ile Thr Lys Val Arg Lys Ser Val Phe Asn Gly Leu
    130                 135                 140

Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser Ser
145                 150                 155                 160

Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr Ile
                165                 170                 175

Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro Pro
            180                 185                 190

Ser Leu Thr Glu Leu His Leu Asp Gly Asn Lys Ile Thr Lys Val Asp
        195                 200                 205

Ala Ala Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu Ser
    210                 215                 220

Phe Asn Ser Ile Ser Ala Val Asp Asn Gly Ser Leu Ala Asn Thr Pro
225                 230                 235                 240

His Leu Arg Glu Leu His Leu Asn Asn Asn Lys Leu Val Lys Val Pro
                245                 250                 255

Gly Gly Leu Ala Asp His Lys Tyr Ile Gln Val Val Tyr Leu His Asn
```

260 265 270

Asn Asn Ile Ser Ala Ile Gly Ser Asn Asp Phe Cys Pro Pro Gly Tyr
275 280 285

Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn Pro
290 295 300

Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr Val
305 310 315 320

Arg Ala Ala Val Gln Leu Gly Asn Tyr Lys
325 330

<210> 14
<211> 1092
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 14
atgatgtcct ttgtctctct gctcctggta ggcatcctat tccatgccac ccaggccgga 60

ccatttcaac agagaggctt atttgacttc atgctagaag atgaggctgc agggattggc 120

ccagaagatc gcattcatga agttctagac ttagagcctc tgggaccagt gtgtcctttc 180

cgctgtcagt gccatcttcg agttgtccaa tgttctgatt tgggtctgga caaagtgccc 240

aaagatcttc cccctgacac cacgctgctg gacctgcaaa acaacaaaat aaccgaaatc 300

aaagatggag actttaagaa cctgaagaat cttcatgcgt tgattcttgt caacaacaaa 360

attagcaaaa tcagccctgg agcatttaca cctttggtga aactggaacg actttatctg 420

tccaagaatc atttgaagga attgccagaa aaaatgccca aaactcttca ggagctgcgt 480

gtccatgaga acgagatcac caaagtgcga aaagcggtgt tcaatggact gaaccagatg 540

atagtcgtag aactgggcac caacccactg aagagctcag gaattgaaaa tggagccttc 600

caggggatga agaagctgtc ctacatccgc attgctgaca ccaacataac caccatccct 660

ccaggtcttc ctccttccct tactgaatta catcttgatg gcaacaaaat caccaaagtt 720

gatgcagcta gcctgagagg actgaataat ttggctaaat tgggactgag tttcaacagc 780

atctctgctg ttgacaatgg ctctctggcc aacactcctc atttgaggga acttcacttg 840

gacaacaaca agcttatcaa agtgcctggt gggctggcgg atcataagta catccaggtt 900

gtctaccttc ataacaacaa tatctctgca gttggatcta atgacttctg cccacctgga 960

tacaacacca aaaaggcttc ttattcgggt gtgagccttt tcagcaaccc agtccagtac 1020

tgggagatcc agccatccac cttccgatgt gtctatgtgc gctctgccat tcagctcgga 1080

aactacaagt ga 1092

<210> 15

<211> 363
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 15

```
Met Met Ser Phe Val Ser Leu Leu Leu Val Gly Ile Leu Phe His Ala
1               5                  10                  15

Thr Gln Ala Gly Pro Phe Gln Gln Arg Gly Leu Phe Asp Phe Met Leu
            20                  25                  30

Glu Asp Glu Ala Ala Gly Ile Gly Pro Glu Asp Arg Ile His Glu Val
        35                  40                  45

Leu Asp Leu Glu Pro Leu Gly Pro Val Cys Pro Phe Arg Cys Gln Cys
    50                  55                  60

His Leu Arg Val Val Gln Cys Ser Asp Leu Gly Leu Asp Lys Val Pro
65                  70                  75                  80

Lys Asp Leu Pro Pro Asp Thr Thr Leu Leu Asp Leu Gln Asn Asn Lys
                85                  90                  95

Ile Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His
            100                 105                 110

Ala Leu Ile Leu Val Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala
        115                 120                 125

Phe Thr Pro Leu Val Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn His
    130                 135                 140

Leu Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg
145                 150                 155                 160

Val His Glu Asn Glu Ile Thr Lys Val Arg Lys Ala Val Phe Asn Gly
                165                 170                 175

Leu Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser
            180                 185                 190

Ser Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr
        195                 200                 205

Ile Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Pro Gly Leu Pro
    210                 215                 220

Pro Ser Leu Thr Glu Leu His Leu Asp Gly Asn Lys Ile Thr Lys Val
225                 230                 235                 240
```

```
Asp Ala Ala Ser Leu Arg Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu
                245                 250                 255

Ser Phe Asn Ser Ile Ser Ala Val Asp Asn Gly Ser Leu Ala Asn Thr
            260                 265                 270

Pro His Leu Arg Glu Leu His Leu Asp Asn Asn Lys Leu Ile Lys Val
        275                 280                 285

Pro Gly Gly Leu Ala Asp His Lys Tyr Ile Gln Val Val Tyr Leu His
    290                 295                 300

Asn Asn Asn Ile Ser Ala Val Gly Ser Asn Asp Phe Cys Pro Pro Gly
305                 310                 315                 320

Tyr Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn
                325                 330                 335

Pro Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr
            340                 345                 350

Val Arg Ser Ala Ile Gln Leu Gly Asn Tyr Lys
        355                 360


<210>  16
<211>  330
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  16

Asp Glu Ala Ala Gly Ile Gly Pro Glu Asp Arg Ile His Glu Val Leu
1               5                   10                  15

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

Leu Arg Val Val Gln Cys Ser Asp Leu Gly Leu Asp Lys Val Pro Lys
        35                  40                  45

Asp Leu Pro Pro Asp Thr Thr Leu Leu Asp Leu Gln Asn Asn Lys Ile
    50                  55                  60

Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His Ala
65                  70                  75                  80

Leu Ile Leu Val Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala Phe
                85                  90                  95

Thr Pro Leu Val Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn His Leu
```

100 105 110

Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg Val
115 120 125

His Glu Asn Glu Ile Thr Lys Val Arg Lys Ala Val Phe Asn Gly Leu
130 135 140

Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser Ser
145 150 155 160

Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr Ile
165 170 175

Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Pro Gly Leu Pro Pro
180 185 190

Ser Leu Thr Glu Leu His Leu Asp Gly Asn Lys Ile Thr Lys Val Asp
195 200 205

Ala Ala Ser Leu Arg Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu Ser
210 215 220

Phe Asn Ser Ile Ser Ala Val Asp Asn Gly Ser Leu Ala Asn Thr Pro
225 230 235 240

His Leu Arg Glu Leu His Leu Asp Asn Asn Lys Leu Ile Lys Val Pro
245 250 255

Gly Gly Leu Ala Asp His Lys Tyr Ile Gln Val Val Tyr Leu His Asn
260 265 270

Asn Asn Ile Ser Ala Val Gly Ser Asn Asp Phe Cys Pro Pro Gly Tyr
275 280 285

Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn Pro
290 295 300

Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr Val
305 310 315 320

Arg Ser Ala Ile Gln Leu Gly Asn Tyr Lys
325 330

<210> 17
<211> 1092
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 17

atgatgtcct ttgtctctct gctcctggta ggcatcctat tccatgccac ccaggccgga 60

ccatttcaac agaaaggctt atttgacttt atgctagaag atgaggctgc agggataggc 120

ccagaagacc gctttcctga agttcctgaa ttagagcctc tgggacccat gtgtcccttc 180

cgctgtcaat gccatcttcg agttgtccaa tgttctgatt tgggtctgga caaagtgccc 240

aaagatcttc cacctgacac tgccctgctg gatctgcaaa acaacaaaat aactgaaatc 300

aaagatggag actttaagaa cctgaagaac cttcatacac tgattctcat caacaacaaa 360

attagcaaaa tcagccctgg agcatttgca cctttggtga aattggaacg actttatcta 420

tccaagaatc aactgaagga attgccagag aaaatgccca aaactcttca ggagctgcgt 480

gtccatgaga atgagatcac caaagtgcga aaggctgtgt tcaatggatt gaaccagatg 540

atcgtcgtag aacttggcac caacccgctg aagagctcag gcattgaaaa cggagctttc 600

cagggaatga agaagctctc ctacatccgc atcgctgaca ccaacattac caccatccct 660

caaggtcttc ctccttccct tactgaatta catcttgatg gcaacaaaat cagcaaagtt 720

gatgcagcta gcctaaaagg actgaataat ttggctaagt tgggactggg tttcaatagc 780

atctcaactg ttgacaatgg ctctctggcc aacactcctc atttgaggga acttcatctg 840

aacaacaaca agcttaacaa agtgcctggt gggctggcag agcataagta catccaggtt 900

gtctaccttc ataacaacaa catctctgca gtcggctcta atgacttctg cccgcctgga 960

tacaacacca aaaaggcttc ttattcgggg gtgagccttt tcagcaaccc agtccagtac 1020

tgggagatcc agccatccac cttccgatgt gtctatgtgc gctctgccat tcagctcgga 1080

aactacaagt ga 1092

<210> 18
<211> 363
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 18

Met Met Ser Phe Val Ser Leu Leu Leu Val Gly Ile Leu Phe His Ala
1 5 10 15

Thr Gln Ala Gly Pro Phe Gln Gln Lys Gly Leu Phe Asp Phe Met Leu
20 25 30

Glu Asp Glu Ala Ala Gly Ile Gly Pro Glu Asp Arg Phe Pro Glu Val
35 40 45

Pro Glu Leu Glu Pro Leu Gly Pro Met Cys Pro Phe Arg Cys Gln Cys
50 55 60

His Leu Arg Val Val Gln Cys Ser Asp Leu Gly Leu Asp Lys Val Pro
65 70 75 80

```
Lys Asp Leu Pro Pro Asp Thr Ala Leu Leu Asp Leu Gln Asn Asn Lys
                85                  90                  95

Ile Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His
            100                 105                 110

Thr Leu Ile Leu Ile Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala
        115                 120                 125

Phe Ala Pro Leu Val Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn Gln
    130                 135                 140

Leu Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg
145                 150                 155                 160

Val His Glu Asn Glu Ile Thr Lys Val Arg Lys Ala Val Phe Asn Gly
                165                 170                 175

Leu Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser
            180                 185                 190

Ser Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr
        195                 200                 205

Ile Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro
    210                 215                 220

Pro Ser Leu Thr Glu Leu His Leu Asp Gly Asn Lys Ile Ser Lys Val
225                 230                 235                 240

Asp Ala Ala Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu
                245                 250                 255

Gly Phe Asn Ser Ile Ser Thr Val Asp Asn Gly Ser Leu Ala Asn Thr
            260                 265                 270

Pro His Leu Arg Glu Leu His Leu Asn Asn Asn Lys Leu Asn Lys Val
        275                 280                 285

Pro Gly Gly Leu Ala Glu His Lys Tyr Ile Gln Val Val Tyr Leu His
    290                 295                 300

Asn Asn Asn Ile Ser Ala Val Gly Ser Asn Asp Phe Cys Pro Pro Gly
305                 310                 315                 320

Tyr Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn
                325                 330                 335

Pro Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr
            340                 345                 350
```

Val Arg Ser Ala Ile Gln Leu Gly Asn Tyr Lys
355 360

<210> 19
<211> 330
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 19

Asp Glu Ala Ala Gly Ile Gly Pro Glu Asp Arg Phe Pro Glu Val Pro
1 5 10 15

Glu Leu Glu Pro Leu Gly Pro Met Cys Pro Phe Arg Cys Gln Cys His
20 25 30

Leu Arg Val Val Gln Cys Ser Asp Leu Gly Leu Asp Lys Val Pro Lys
35 40 45

Asp Leu Pro Pro Asp Thr Ala Leu Leu Asp Leu Gln Asn Asn Lys Ile
50 55 60

Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His Thr
65 70 75 80

Leu Ile Leu Ile Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala Phe
85 90 95

Ala Pro Leu Val Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn Gln Leu
100 105 110

Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg Val
115 120 125

His Glu Asn Glu Ile Thr Lys Val Arg Lys Ala Val Phe Asn Gly Leu
130 135 140

Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser Ser
145 150 155 160

Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr Ile
165 170 175

Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro Pro
180 185 190

Ser Leu Thr Glu Leu His Leu Asp Gly Asn Lys Ile Ser Lys Val Asp
195 200 205

Ala Ala Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu Gly

210 215 220

Phe Asn Ser Ile Ser Thr Val Asp Asn Gly Ser Leu Ala Asn Thr Pro
225 230 235 240

His Leu Arg Glu Leu His Leu Asn Asn Asn Lys Leu Asn Lys Val Pro
245 250 255

Gly Gly Leu Ala Glu His Lys Tyr Ile Gln Val Val Tyr Leu His Asn
260 265 270

Asn Asn Ile Ser Ala Val Gly Ser Asn Asp Phe Cys Pro Pro Gly Tyr
275 280 285

Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn Pro
290 295 300

Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr Val
305 310 315 320

Arg Ser Ala Ile Gln Leu Gly Asn Tyr Lys
325 330

<210> 20
<211> 1092
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 20
atgatgtcct ttgtctctct gctcctggta ggcatcctat tccatgccac ccaggccgga 60

ccgtttcaac agaaaggctt atttgacttt atgctggaag atgaggctgc agggataggc 120

ccggaagagc gctttcatga ggttcctgaa ttagagccta tgggcccagt ctgccccttc 180

cgttgccagt gccatctgcg agttgtccag tgttctgatc tgggtctgga aaaagtgccc 240

aaagaccttc ctcctgatac cgcgctgctg gacctgcaaa acaacaaaat aactgagatc 300

aaagatggag actttaaaaa cctgaagaac cttcatacac tgattctcat caacaacaaa 360

attagcaaaa ttagccctgg ggcatttgct cctctggtga aattggaacg actttatctt 420

tccaagaatc aactgaagga attgccagag aaaatgccca aaactcttca ggagctgcgt 480

gtccatgaga acgagatcac caaagtgcga aagtctgtgt tcaatggatt gaaccagatg 540

atcgtcgtag aacttggcac caacccactg aagagctcag gcattgaaaa tggagccttt 600

cagggaatga agaagctctc ctacatccgc attgctgaca ctaatataac taccatccct 660

caaggtcttc ctccttccct tactgaatta catctcgacg gcaacaaaat caccaaagtt 720

gatgcagcta gcctgaaagg actgaataat ttggctaagt tgggactgag tttcaacagc 780

atctctgctg ttgacaatgg ctctttggcc aacactcctc atttgaggga acttcatttg 840

aacaacaaca agcttgtcaa agtgcccggt gggctggccg accataagta catccaggtt 900

gtctaccttc acaacaacaa tatctctgca atcggctcta acgacttctg cccacctgga 960

tacaacacca aaaaggcttc ttattcagga gtgagccttt tcagcaaccc agtccagtac 1020

tgggagatcc agccatccac cttccgatgt gtctacgtgc gcgctgctgt tcagcttgga 1080

aactacaagt ga 1092

<210> 21
<211> 363
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 21

Met Met Ser Phe Val Ser Leu Leu Leu Val Gly Ile Leu Phe His Ala
1 5 10 15

Thr Gln Ala Gly Pro Phe Gln Gln Lys Gly Leu Phe Asp Phe Met Leu
20 25 30

Glu Asp Glu Ala Ala Gly Ile Gly Pro Glu Glu Arg Phe His Glu Val
35 40 45

Pro Glu Leu Glu Pro Met Gly Pro Val Cys Pro Phe Arg Cys Gln Cys
50 55 60

His Leu Arg Val Val Gln Cys Ser Asp Leu Gly Leu Glu Lys Val Pro
65 70 75 80

Lys Asp Leu Pro Pro Asp Thr Ala Leu Leu Asp Leu Gln Asn Asn Lys
85 90 95

Ile Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His
100 105 110

Thr Leu Ile Leu Ile Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala
115 120 125

Phe Ala Pro Leu Val Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn Gln
130 135 140

Leu Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg
145 150 155 160

Val His Glu Asn Glu Ile Thr Lys Val Arg Lys Ser Val Phe Asn Gly
165 170 175

Leu Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser
180 185 190

```
Ser Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr
        195                 200                 205

Ile Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro
    210                 215                 220

Pro Ser Leu Thr Glu Leu His Leu Asp Gly Asn Lys Ile Thr Lys Val
225                 230                 235                 240

Asp Ala Ala Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu
                245                 250                 255

Ser Phe Asn Ser Ile Ser Ala Val Asp Asn Gly Ser Leu Ala Asn Thr
            260                 265                 270

Pro His Leu Arg Glu Leu His Leu Asn Asn Asn Lys Leu Val Lys Val
        275                 280                 285

Pro Gly Gly Leu Ala Asp His Lys Tyr Ile Gln Val Val Tyr Leu His
    290                 295                 300

Asn Asn Asn Ile Ser Ala Ile Gly Ser Asn Asp Phe Cys Pro Pro Gly
305                 310                 315                 320

Tyr Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn
                325                 330                 335

Pro Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr
            340                 345                 350

Val Arg Ala Ala Val Gln Leu Gly Asn Tyr Lys
        355                 360


<210>  22
<211>  330
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  22

Asp Glu Ala Ala Gly Ile Gly Pro Glu Glu Arg Phe His Glu Val Pro
1               5                   10                  15

Glu Leu Glu Pro Met Gly Pro Val Cys Pro Phe Arg Cys Gln Cys His
            20                  25                  30

Leu Arg Val Val Gln Cys Ser Asp Leu Gly Leu Glu Lys Val Pro Lys
        35                  40                  45

Asp Leu Pro Pro Asp Thr Ala Leu Leu Asp Leu Gln Asn Asn Lys Ile
```

50 55 60

Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His Thr
65 70 75 80

Leu Ile Leu Ile Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala Phe
85 90 95

Ala Pro Leu Val Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn Gln Leu
100 105 110

Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg Val
115 120 125

His Glu Asn Glu Ile Thr Lys Val Arg Lys Ser Val Phe Asn Gly Leu
130 135 140

Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser Ser
145 150 155 160

Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr Ile
165 170 175

Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro Pro
180 185 190

Ser Leu Thr Glu Leu His Leu Asp Gly Asn Lys Ile Thr Lys Val Asp
195 200 205

Ala Ala Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu Ser
210 215 220

Phe Asn Ser Ile Ser Ala Val Asp Asn Gly Ser Leu Ala Asn Thr Pro
225 230 235 240

His Leu Arg Glu Leu His Leu Asn Asn Asn Lys Leu Val Lys Val Pro
245 250 255

Gly Gly Leu Ala Asp His Lys Tyr Ile Gln Val Val Tyr Leu His Asn
260 265 270

Asn Asn Ile Ser Ala Ile Gly Ser Asn Asp Phe Cys Pro Pro Gly Tyr
275 280 285

Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn Pro
290 295 300

Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr Val
305 310 315 320

Arg Ala Ala Val Gln Leu Gly Asn Tyr Lys

325 330

<210> 23
<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 23

Met Met Ser Phe Val Ser Leu Leu Leu Val Gly Ile Leu Phe His Ala
1 5 10 15

Thr Gln Ala

<210> 24
<211> 57
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 24
atgatgtcct ttgtctctct gctcctggta ggcatcctat tccatgccac ccaggcc 57

<210> 25
<211> 1092
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 25
atgatgtcct ttgtctctct gctcctggta ggcatcctat tccatgccac ccaggccggg 60

ccgttccaac agagaggctt atttgacttt atgctagaag atgaggctgc agggataggc 120

ccagaagagc acgctcctgt tgattctgat ttagagcctc tggggccagt gtgtcctttc 180

cgctgtcagt gccaccttcg agttgtgcag tgttctgatt tgggtttgga aaaagtgcca 240

aaagagctcc ctcctgacac tacgctgctg gacttgcaaa acaacaaaat aaccgaaatc 300

aaagatggag acttcaagaa cctgaagaac cttcatacgt tgatccttgt caacaacaaa 360

attagcaaaa tcagccctgg agcatttaca cctttgttga aattggaacg actttatctg 420

tccaagaatc atctgaagga attgccagaa aaaatgccca aaactcttca ggagctgcgt 480

gctcacgaga atgagatcac caaagtgcga aaagctgtgt tcaatggcct gaaccagatg 540

atcgtcgtag aactgggcac caacccgctg aagagctcgg gaattgaaaa tggagccttc 600

cagggaatga agaagctgtc ctacatccgc attgccgaca ccaatataac caccatcccg 660

caaggtcttc ctccttccct tactgaatta catcttgaag gcaacaaaat ctccaaagtt 720

gatgcagcta gcctgaaagg actgaataat ttggctaagt tgggactgag ttttaacagc 780

atctctgcta ttgacaatgg cactctggcc aacactcctc atttgaggga gcttcacttg 840

gacaacaata agcttatcag agtacctggt gggctggcgg agcacaaata catccaggtt 900

gtctaccttc ataacaacaa tatctctgca gtcgggtcta acgacttctg cccacctgga 960

tacaacacca aaaaggcttc ttattcaggt gtgagccttt tcagcaaccc agtccagtac 1020

tgggagatcc aaccatccac cttccgatgt gtctatgtgc gttccgccat ccagcttgga 1080

aattataaat ga 1092

<210> 26
<211> 363
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 26

Met Met Ser Phe Val Ser Leu Leu Leu Val Gly Ile Leu Phe His Ala
1 5 10 15

Thr Gln Ala Gly Pro Phe Gln Gln Arg Gly Leu Phe Asp Phe Met Leu
20 25 30

Glu Asp Glu Ala Ala Gly Ile Gly Pro Glu Glu His Ala Pro Val Asp
35 40 45

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

His Leu Arg Val Val Gln Cys Ser Asp Leu Gly Leu Glu Lys Val Pro
65 70 75 80

Lys Glu Leu Pro Pro Asp Thr Thr Leu Leu Asp Leu Gln Asn Asn Lys
85 90 95

Ile Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His
100 105 110

Thr Leu Ile Leu Val Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala
115 120 125

Phe Thr Pro Leu Leu Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn His
130 135 140

Leu Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg
145 150 155 160

Ala His Glu Asn Glu Ile Thr Lys Val Arg Lys Ala Val Phe Asn Gly
165 170 175

Leu Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser

180 185 190

Ser Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr
195 200 205

Ile Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro
210 215 220

Pro Ser Leu Thr Glu Leu His Leu Glu Gly Asn Lys Ile Ser Lys Val
225 230 235 240

Asp Ala Ala Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu
245 250 255

Ser Phe Asn Ser Ile Ser Ala Ile Asp Asn Gly Thr Leu Ala Asn Thr
260 265 270

Pro His Leu Arg Glu Leu His Leu Asp Asn Asn Lys Leu Ile Arg Val
275 280 285

Pro Gly Gly Leu Ala Glu His Lys Tyr Ile Gln Val Val Tyr Leu His
290 295 300

Asn Asn Asn Ile Ser Ala Val Gly Ser Asn Asp Phe Cys Pro Pro Gly
305 310 315 320

Tyr Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn
325 330 335

Pro Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr
340 345 350

Val Arg Ser Ala Ile Gln Leu Gly Asn Tyr Lys
355 360

<210> 27
<211> 330
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 27

Asp Glu Ala Ala Gly Ile Gly Pro Glu Glu His Ala Pro Val Asp Ser
1 5 10 15

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

Leu Arg Val Val Gln Cys Ser Asp Leu Gly Leu Glu Lys Val Pro Lys
35 40 45

```
Glu Leu Pro Pro Asp Thr Thr Leu Leu Asp Leu Gln Asn Asn Lys Ile
    50                  55                  60

Thr Glu Ile Lys Asp Gly Asp Phe Lys Asn Leu Lys Asn Leu His Thr
65                  70                  75                  80

Leu Ile Leu Val Asn Asn Lys Ile Ser Lys Ile Ser Pro Gly Ala Phe
                85                  90                  95

Thr Pro Leu Leu Lys Leu Glu Arg Leu Tyr Leu Ser Lys Asn His Leu
            100                 105                 110

Lys Glu Leu Pro Glu Lys Met Pro Lys Thr Leu Gln Glu Leu Arg Ala
        115                 120                 125

His Glu Asn Glu Ile Thr Lys Val Arg Lys Ala Val Phe Asn Gly Leu
    130                 135                 140

Asn Gln Met Ile Val Val Glu Leu Gly Thr Asn Pro Leu Lys Ser Ser
145                 150                 155                 160

Gly Ile Glu Asn Gly Ala Phe Gln Gly Met Lys Lys Leu Ser Tyr Ile
                165                 170                 175

Arg Ile Ala Asp Thr Asn Ile Thr Thr Ile Pro Gln Gly Leu Pro Pro
            180                 185                 190

Ser Leu Thr Glu Leu His Leu Glu Gly Asn Lys Ile Ser Lys Val Asp
        195                 200                 205

Ala Ala Ser Leu Lys Gly Leu Asn Asn Leu Ala Lys Leu Gly Leu Ser
    210                 215                 220

Phe Asn Ser Ile Ser Ala Ile Asp Asn Gly Thr Leu Ala Asn Thr Pro
225                 230                 235                 240

His Leu Arg Glu Leu His Leu Asp Asn Asn Lys Leu Ile Arg Val Pro
                245                 250                 255

Gly Gly Leu Ala Glu His Lys Tyr Ile Gln Val Val Tyr Leu His Asn
            260                 265                 270

Asn Asn Ile Ser Ala Val Gly Ser Asn Asp Phe Cys Pro Pro Gly Tyr
        275                 280                 285

Asn Thr Lys Lys Ala Ser Tyr Ser Gly Val Ser Leu Phe Ser Asn Pro
    290                 295                 300

Val Gln Tyr Trp Glu Ile Gln Pro Ser Thr Phe Arg Cys Val Tyr Val
305                 310                 315                 320
```

Arg Ser Ala Ile Gln Leu Gly Asn Tyr Lys
325 330

<210> 28
<211> 1119
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 28
aagcttgccg ccaccatgat gtcctttgtc tctctgctcc tggtaggcat cctattccat 60

gccacccagg ccgggccgtt ccaacagaga ggcttatttg actttatgct agaagatgag 120

gctgcaggga taggcccgga ggaccgtgca cctgacatgc ctgacctcga gcttctggga 180

cctgtgtgtc ccttccgctg tcagtgccat ctccgagtgg tccagtgttc cgacctgggt 240

ctggacaaag taccaaaaga tcttcccccc gacactacgc tgctcgactt gcaaaacaac 300

aaaatcaccg aaatcaaaga tggagacttc aagaacctca agaacctgca taccttgatt 360

cttgtaaaca acaaaattag caaaatcagc cctggagcat ttacaccttt gttgaaattg 420

gaacgacttt atctgtccaa gaatcatctg aaggaattgc cagaaaaaat gcccaaaact 480

cttcaggagc tgcgtgccca tgagaatgag atcaccaaag ttcgaaaagc tgtgttcaat 540

ggactgaacc agatgatcgt cgtagagctg ggcaccaatc ccctgaagag ttcagggatt 600

gaaaatggag ccttccaggg aatgaagaag ctctcctata tccgcattgc tgataccaat 660

ataactacca tccctcaagg tcttcctcct tcccttactg aattacatct tgaaggcaac 720

aaaatcacca aggttgatgc atctagcctg aaaggactga ataatttggc taagttggga 780

ctgagtttta acagcatctc cgctgttgac aatggcactc tagccaacac tcctcatctg 840

agggagcttc acttggacaa caataagctc atcagagtac ccggtgggct ggcggagcat 900

aagtacatcc aggttgtcta ccttcataac aacaatatat ctgcagtcgg atctaatgac 960

ttctgcccac ctggatacaa caccaaaaag gcttcttatt caggtgtgag ccttttcagc 1020

aacccagtgc agtactggga gatccagcca tccaccttcc ggtgtgtcta cgtgcgctct 1080

gccatccagc ttggaaatta taaatgatga gatctcgag 1119

<210> 29
<211> 1119
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 29
aagcttgccg ccaccatgat gtcctttgtc tctctgctcc tggtaggcat cctattccat 60

gccacccagg ccgggccgtt ccaacagaga ggcttatttg actttatgct agaagatgag 120

gctgcaggga taggcccaga agagcacgct cctgttgatt ctgatttaga gcctctgggg 180

```
ccagtgtgtc ctttccgctg tcagtgccac cttcgagttg tgcagtgttc tgatttgggt      240

ttggaaaaag tgccaaaaga gctccctcct gacactacgc tgctggactt gcaaaacaac      300

aaaataaccg aaatcaaaga tggagacttc aagaacctga agaaccttca tacgttgatc      360

cttgtcaaca acaaaattag caaaatcagc cctggagcat ttacaccttt gttgaaattg      420

gaacgacttt atctgtccaa gaatcatctg aaggaattgc cagaaaaaat gcccaaaact      480

cttcaggagc tgcgtgctca cgagaatgag atcaccaaag tgcgaaaagc tgtgttcaat      540

ggcctgaacc agatgatcgt cgtagaactg ggcaccaacc cgctgaagag ctcgggaatt      600

gaaaatggag ccttccaggg aatgaagaag ctgtcctaca tccgcattgc cgacaccaat      660

ataaccacca tcccgcaagg tcttcctcct tcccttactg aattacatct tgaaggcaac      720

aaaatctcca aagttgatgc agctagcctg aaaggactga ataatttggc taagttggga      780

ctgagtttta acagcatctc tgctattgac aatggcactc tggccaacac tcctcatttg      840

agggagcttc acttggacaa caataagctt atcagagtac ctggtgggct ggcggagcac      900

aaatacatcc aggttgtcta ccttcataac aacaatatct ctgcagtcgg gtctaacgac      960

ttctgcccac ctggatacaa caccaaaaag gcttcttatt caggtgtgag ccttttcagc     1020

aacccagtcc agtactggga gatccaacca tccaccttcc gatgtgtcta tgtgcgttcc     1080

gccatccagc ttggaaatta taaatgatga gatctcgag                            1119
```