

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
3 July 2003 (03.07.2003)

PCT

(10) International Publication Number
WO 03/054005 A2

- (51) International Patent Classification⁷: **C07K 14/00**
- (74) Agent: **ROBIC**; 55, St-Jacques, Montréal, Quebec H2Y 3X2 (CA).
- (21) International Application Number: PCT/IB02/05778
- (22) International Filing Date:
20 December 2002 (20.12.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/341,224 20 December 2001 (20.12.2001) US
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant (*for all designated States except US*): **PHENOGENE THERAPEUTIQUES INC.** [CA/CA]; 416, boulevard de Maisonneuve Ouest, Bureau 1020, Montréal, Québec H3A 1L2 (CA).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **LANCTOT, Christian** [CA/CA]; 4000 Kindersley, apt. 305, Montréal, Québec H4P 1K8 (CA). **SALOIS, Patrick** [CA/CA]; 10142 St-Hubert, Montréal, Québec H2C 2H5 (CA). **MOFFATT, Pierre** [CA/CA]; 2910 Terr. J.S. Aime Ashby, Lachine, Quebec H8S 4H9 (CA).
- Published:**
— *without international search report and to be republished upon receipt of that report*
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: BONE POLYPEPTIDE-1

(57) Abstract: The present invention relates to a bone polypeptide, particularly bone polypeptide-1 and nucleic acid molecules encoding the same. The present invention provides the human primary sequence of bone polypeptide-1 as well as sequences of vertebrate homologs. The invention also provides cell lines engineered to express the cDNA, antibodies to detect its translation products, and recombinant adenoviruses to deliver an expressible cDNA into a host. Bone polypeptide-1 can be utilized to treat diseases affecting renal and bone functions.



WO 03/054005 A2

BONE POLYPEPTIDE-1

RELATED APPLICATION

This application claims the benefit under 35 U.S.C. § 119(e) of United States provisional patent application number 60/341,224, filed December 20, 2001.

BACKGROUND OF THE INVENTION

a) Field of the invention

The present invention relates to a bone polypeptide and active fragments thereof. In particular, the present invention relates to bone polypeptide-1, active fragments thereof and nucleic acids encoding the same.

b) Brief description of the prior art

Bone is a specialized connective tissue that is constantly remodelled by the reciprocal action of bone-forming cells (osteoblasts) and bone-resorbing cells (osteoclasts) (Manolagas, 2000). Full citations for the references cited herein are provided before the claims. In the developing organism, the skeletal system is formed either through endochondral or intramembranous ossification (Baron, 1999). Endochondral bone formation, exemplified by the development of the long bones of the limbs, involves replacement of a preformed cartilage template whereas intramembranous ossification, a process typical of flat bones of the skull, does not rely on an intermediate cartilaginous step. In both cases, osteoblasts originate from the mesenchyme and initially differentiate from the inner layer of the periosteum. Periosteal osteoblasts deposit lamellar bone, progressively becoming enthumbed in the matrix they deposit and finally differentiating into osteocytes. Osteocytes elaborate cytoplasmic extensions through canaliculi, thereby constituting a network of interconnected cells within bone tissue (Aarder et al., 1994). It is thought that this network senses load on the bones and alters bone activity according to the demands of that load.

Osteoporosis is characterized by a loss of bone mass due to an imbalance between bone resorption and bone formation. This degenerative disease affects 20 million women in the United States. Current treatment mainly involves the inhibition of the activity of osteoclasts by inhibitors such as bisphosphonates (Russell, 1999). Such anti-resorptive therapies slow down the progression of the disease but do not really help in rebuilding lost bone. Effective bone anabolics are thus needed. Unfortunately, with the possible exception of a fragment of parathyroid hormone (PTH₁₋₃₄; Neer et al., 2001), very few molecules have been shown to notably increase bone mass in osteoporotic patients. A number of growth factors and related molecules are now being considered as therapeutic agents. Insulin-like growth factor 1 (IGF1) is a protein that displays bone-sparing activity in the ovariectomized rat, a model of postmenopausal osteoporosis (Bagi et al., 1995). Basic fibroblast growth factor (bFGF) is another protein that was shown to enhance bone formation *in vivo* (Mayahara et al., 1993). However IGF1, as the name implies, is also an hypoglycemic factor and bFGF severely disrupts hematopoiesis. Hence, the *in vivo* specificity of these growth factors is difficult to ascertain and their clinical potential is unproven. There is thus a need to identify molecules that can efficiently and specifically increase bone formation.

Regulation of bone mass by the central nervous system has recently been reported by Ducy and colleagues (Ducy et al., 2000). These authors have shown that the ob/ob mice lacking a functional leptin gene have a higher bone mass. Leptin is a 16kD hormone synthesized by the adipocytes and acting on hypothalamic neurons to regulate caloric intake (Unger, 2000). When injected intracerebroventricularly, leptin caused a loss of bone. It has been hypothesized that, in addition to adipocytes, osteoblasts *per se* might secrete a humoral factor that regulates bone mass through an hypothalamic relay. Thus, identification of such a factor might be useful to devise new therapies for the treatment of osteoporosis.

The inorganic component of bone is made up of hydroxyapatite crystals. Hence, in addition to its role as a supporting mechanical structure, the skeleton is a reservoir of calcium. Bone cells play a crucial role in the homeostasis of calcium and other ions such as phosphate. In particular, it has been proposed that bone

cells secrete a hormone, tentatively called phosphatonin, that regulates phosphate retention by kidney tubules. It has been reported that phosphatonin levels are elevated in conditioned medium of tumor cells derived from a rare disease called oncogenic hypophosphatemia osteomalacia (Kumar, 2000). Furthermore, it has been postulated that phosphatonin is a substrate for Phex, a metallopeptidase found at the cell surface of osteoblasts and osteocytes (Frota Ruchon et al., 2000). This hypothesis is supported by the fact that patients with X-linked hypophosphatemia harbor deleterious mutations in the Phex coding sequence (The Hyp consortium, 1995) and presumably have elevated levels of active phosphatonin. Low serum phosphate levels are associated with impaired bone quality. Hypophosphatemia could be remedied by injection of phosphatonin antagonists. In view of the above, it is clear that there is a need to identify the molecular nature of phosphatonin or other molecules that regulate phosphate metabolism. In particular, there is a tremendous need to identify bone polypeptides that may be useful as therapeutic agents.

SUMMARY OF THE INVENTION

In order to meet these needs, the present invention is directed to recombinant bone polypeptide-1 (BP-1), proteins sharing substantial homology to BP-1 and active fragments thereof. Bone polypeptide-1 is polypeptide expressed in bone. Bone polypeptide 1 can be synthesized chemically, recombinantly produced, isolated and/or purified from a recombinant host or it and it can be isolated and/or purified from its natural source. Sources of bone polypeptide-1 include all organisms containing bone since bone polypeptide-1 is predominantly expressed in bone cells. Preferred sources of bone polypeptide 1 include rat, mouse, python, cow and chicken. An especially preferred source of bone polypeptide-1 is a human.

The present invention is further directed to nucleic acids encoding bone polypeptide-1 and active fragments thereof; a vector containing the nucleic acids and a host cell carrying the vector. The present invention is further directed to processes to produce bone polypeptide-1 and active fragments thereof, processes to produce cells capable of producing the bone polypeptide-1 and active fragments

thereof and to produce a vector containing DNA or RNA encoding bone polypeptide-1 and active fragments thereof. The present invention is further directed to methods for treating bone and renal diseases and methods for using pharmacologic compositions comprising an effective amount of bone polypeptide-1 and active fragments thereof. The invention also encompasses monoclonal and polyclonal antibodies specifically recognizing bone polypeptide-1 and active fragments thereof.

The present invention is further directed to an isolated nucleic acid encoding bone polypeptide-1 and active fragments thereof. The nucleic acid may be isolated from an animal, preferably a human. Other sources of the nucleic acid include rats, cows, snakes, mice and chickens.

Bone polypeptide-1 may regulate bone cell proliferation and/or bone cell differentiation and/or osteoblast activity via an autocrine, paracrine or endocrine pathway. Active fragments of bone polypeptide-1 are fragments that have similar activity as the full length protein and therefore may regulate bone cell proliferation and/or bone cell differentiation and/or osteoblast activity via an autocrine, paracrine or endocrine pathway.

The present invention is further directed to nucleic acids having sequences selected from: SEQ ID NO: 2; SEQ ID NO: 3; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 6, SEQ ID NO: 7 and homologous sequences.

The present invention is further directed to vector including a nucleic acid encoding bone polypeptide-1 and active fragments thereof. In one format, the vector includes an expression control sequence operably linked to the nucleic acid. The expression control sequence may be a promoter such as a prokaryotic promoter or a eukaryotic promoter. The vector may be present in an isolated cell such as a prokaryotic cell or a eukaryotic cell.

The present invention is further directed to a method of producing bone polypeptide-1 or active fragments thereof by culturing a cell containing an vector containing a nucleic acid encoding bone polypeptide-1 under conditions permitting expression of the polypeptide and purifying the polypeptide from the cell or culture medium of the cell. The cell may be a prokaryotic cell or a eukaryotic cell.

The present invention is further directed to an isolated nucleic acid that hybridizes under high stringency conditions to a nucleic acid including nucleotide sequences such as SEQ ID NO: 2; SEQ ID NO: 3; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, and the complements thereof wherein the high stringency conditions includes hybridizing in 0.15 M NaCl at 72°C for about 15 minutes and washing the hybridized DNA in 0.2X SSC at 65°C for 15 minutes. The nucleic acid may be isolated from an animal, preferably a human. Other sources of the nucleic acid include rats, cows, snakes, mice and chickens.

The nucleic acids of the invention may be cloned into an vector. The vector may further include an expression control sequence operably linked to the nucleic acid. The expression control sequence may be a promoter such as a prokaryotic promoter or a eukaryotic promoter.

The present invention is further directed to isolated and/or purified and/or recombinant bone polypeptide-1 or an active fragment thereof in a pharmaceutical composition. The polypeptide may be isolated and/or purified from a human. Alternatively, the polypeptide may be recombinantly produced from the nucleic acid encoding the human polypeptide. The polypeptide may also be isolated from rat, cow, snake, mouse or chicken. Alternatively, the polypeptide may be recombinantly produced from the nucleic acid encoding the rat, cow, snake, mouse or chicken polypeptide.

The present invention is further directed to polypeptides having amino acid sequences selected from of SEQ ID NO: 9; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 12; SEQ ID NO: 13; SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 21 SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 24; SEQ ID NO: 25; SEQ ID NO: 26, SEQ ID NO: 27, and SEQ ID NO: 28.

The present invention is further directed to an expression vector containing a coding sequence the sequence of which encodes bone polypeptide-1 or an active fragment thereof, the vector being suitable for genetic therapy.

The present invention is further directed to a method to deliver a nucleic acid, the sequence of which encodes bone polypeptide-1 or a fragment thereof,

into a host for therapy. In one format of the method of the invention, the delivery is by adenovirus.

The present invention is further directed to a method of treating a bone disorder or osteoporosis comprising administering an effective amount of bone polypeptide-1 or a fragment thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A, B show the results of a Northern analysis comparing the expression of a mouse cDNA encoding bone polypeptide-1 (BP-1) in a variety of embryonic and adult mouse tissues. **A.** Five micrograms of total RNA was loaded in lanes 1 to 16. 1, brain e11.5; 2, adult brain; 3, gonads e14.5; 4, adult testis; 5, heart e12.5; 6, adult heart; 7, intestine e14.5; 8, adult intestine; 9, kidney e13.5; 10, adult kidney; 11, liver e11.5; 12, adult liver; 13, lung e12.5; 14, adult lung; 15, spleen e15.5; 16, adult spleen. **B.** One and a half microgram of polyA(+) RNA was loaded in lanes 21 to 23. 21, calvaria e15.5; 22, adult kidney; 23, adult liver. Bars indicate migration of 28S and 18S ribosomal RNA.

Figure 2 shows the nucleotide sequence of a full length mouse cDNA encoding bone protein-1 (SEQ ID NO:1). The open reading frame is written in capital letters, the polyadenylation signal is in bold characters, the sequence corresponding to the original ca155-cap1-c2-neg306 clone is underlined.

Figures 3A-F show brightfield (A,D) and darkfield (B,C,E,F) images of sections through the limb of a mouse embryo at 16 days of gestation (A-C) or through the head of a mouse embryo at 13 days of gestation (D-F) hybridized with a probe complementary to a novel cDNA encoding bone polypeptide-1 (B,E) or to Cbfa1, a transcription factor expressed mainly in the osteoblast lineage (C,F).

Figure 4 shows an alignment of the coding sequences from human (*Homo sapiens*) (SEQ ID NO:2), mouse (*Mus musculus*) (SEQ ID NO:3), rat (*Rattus norvegicus*) (SEQ ID NO:4), bovine (*Bos taurus*) (SEQ ID NO:5), chicken (*Gallus*

gallus) (SEQ ID NO:6), and python (*Python molurus bivittatus*) (SEQ ID NO: 7) BP-1 cDNAs. The consensus DNA sequence for BP-1 is provided as SEQ ID NO: 8.

Figure 5A shows an alignment of the primary sequences of BP deduced from cDNAs isolated from tissues of the human (*Homo sapiens*) (SEQ ID NO:9), mouse (*Mus musculus*) (SEQ ID NO:10), rat (*Rattus norvegicus*) (SEQ ID NO:11), bovine (*Bos taurus*) (SEQ ID NO:12), chicken (*Gallus gallus*) (SEQ ID NO:13) and python (*Python molurus bivittatus*) (SEQ ID NO:14). The consensus polypeptide sequence for BP-1 is provided as SEQ ID NO: 15. Sequence of the putative signal peptides are underlined. Putative processing sites are boxed.

Figure 5B shows an alignment of a region of homology between BP-1 and members of the natriuretic peptides family. Numbering of amino acids is according to the full length precursor protein (SEQ ID NO:9 for BP-1). The human sequence is SEQ ID NO: 16. The rat atrial natriuretic factor protein region is SEQ ID NO: 17. The rat brain natriuretic factor protein region is SEQ ID NO: 18. The rat C-type natriuretic factor protein region is SEQ ID NO: 19.

Figure 6A depicts the full length protein derived from the human BP-1 cDNA (SEQ ID NO: 2) and is presented as SEQ ID NO: 9. The signal peptide and corresponding signal peptidase cleavage site are shown. Also shown are the conserved dibasic cleavage sites, the putative products generated by processing of the protein (PROT 1 (SEQ ID NO: 20), PEPT 1 (SEQ ID NO: 21), PEPT 2A (SEQ ID NO: 22), PEPT 3A (SEQ ID NO: 23), PEPT 2B (SEQ ID NO: 24), PEPT 3B (SEQ ID NO: 25)), and the position of synthetic peptides used to raise antibodies (AG1 (SEQ ID NO: 26), AG2 (SEQ ID NO: 27), AG3 (SEQ ID NO: 28)). The residues at the N and C termini of peptides are indicated and numbered according to the initiator methionine of SEQ ID NO:9.

Figure 6B shows a picture of Western analysis on cell extracts (lanes 620) or precipitated medium (lanes 621) from cells transfected with a vector expressing

wild type BP-1 mouse cDNA (lanes 620, 621, 622), or a form of BP-1 mouse cDNA where one of the dibasic cleavage site was mutated (lanes 623). Migration of molecular weight markers is indicated.

Figure 7A depicts an expression vector for BP-1 (plasmid GD46) that can be used to generate cell lines secreting the BP-1 polypeptide or fragments thereof in the culture medium. The vector comprises components of a transcription unit for mouse BP-1 (700, 701, 702) as well as a transcription unit to confer resistance to puromycin (703).

Figure 7B shows the standard curve of an enzyme-linked immunoabsorbent assay (ELISA) performed on increasing quantities of synthetic peptide derived from the C terminal region of BP-1.

Figure 8A shows the levels of BP-1 immunoreactivity in fractions collected from a cation-exchange chromatography column after passage of conditioned medium from 293-GD46-7 cells overexpressing the mouse BP-1 coding sequence.

Figure 8B is a picture of a silver-stained polyacrylamide gel showing the proteins found in the conditioned medium from 293-GD46-7 cells (lane 810), in the flow through from a cation-exchange chromatography column (lane 811) and in BP-1-immunoreactive fractions from the same column (lane 812). Arrow 813 points to a BP-1 product. Arrowhead 814 points to aprotinin, a contaminating protease inhibitor included in the conditioned medium before the chromatographic step.

Figure 8C shows the result of a Western analysis performed on conditioned medium from 293-GD46-7 cells (lane 820), on the flow through from a cation-exchange chromatography column (lane 821) and on BP-1-immunoreactive fractions from the same column (lane 822).

Figure 9A depicts a vector (plasmid GD28b) that can be used to generate recombinant adenoviral particles expressing BP-1 coding sequence. The vector comprises a transcription unit for mouse BP-1 (901) flanked by two regions of the adenovirus serotype 5 genome (902, 903).

Figure 9B shows the result of an ELISA to detect BP-1 products in the medium of primary cultures of rat osteoblasts infected with recombinant adenoviral particles at day 11 of culture.

Figure 10 shows the result of a Northern analysis to detect osteocalcin (OCN) in total RNA extracted from primary cultures of osteoblasts treated chronically with control medium (1001) or medium containing BP-1 products (1002). The levels of a control messenger (GAPDH) are also shown.

DETAILED DESCRIPTION OF THE INVENTION

A) Definitions

Throughout the text, the word "kilobase" is generally abbreviated as "kb", the words "deoxyribonucleic acid" as "DNA", the words "ribonucleic acid" as "RNA", the words "complementary DNA" as "cDNA", the words "polymerase chain reaction" as "PCR", the words "expressed sequenced tag" as "EST", and the words "reverse transcription" as "RT". Nucleotide sequences are written in the 5' to 3' orientation unless stated otherwise. Amino acid sequences are written from the N terminus unless stated otherwise.

In order to provide an even clearer and more consistent understanding of the specification and the claims, including the scope given herein to such terms, the following definitions are provided:

Exogenous nucleic acid: A nucleic acid (such as cDNA, cDNA fragments, genomic DNA fragments, mRNA fragments, antisense RNA, oligonucleotide) which is not naturally part of another nucleic acid molecule. The "exogenous

nucleic acid" may be from any organism, purely synthetic, or any combination thereof.

Expressed sequenced tag: sequence information on a small nucleic acid (typically 200-500 bp) derived from a gene transcript.

Hormone: a molecule, generally polypeptidic in nature, secreted extracellularly by one cell type to modulate the activity of target cells.

Hormone precursor. Refers to a secreted protein that is processed in the secretory pathway such that one or more fragments are released extracellularly. Fragments can be further modified (e.g. amidation at the C terminus) before or after release in the extracellular space.

Host: A cell, tissue, organ or organism capable of providing cellular components for allowing the expression of an exogenous nucleic acid. The exogenous nucleic acid maybe cloned into a vector. This term is intended to also include hosts which have been modified in order to accomplish these functions. Bacteria, fungi, animal (cells, tissues, or organisms) and plant (cells, tissues, or organisms) are examples of a host.

Insertion: The process by which a nucleic acid is introduced into another nucleic acid. Methods for inserting a nucleic acid into another normally requires the use of restriction enzymes and such methods of insertion are well known in the art.

In silico: using computer and bioinformatics software and hardware.

Nucleic acid: Any DNA, RNA sequence or molecule having one nucleotide or more, including nucleotide sequences encoding a complete gene. The term is intended to encompass all nucleic acids whether occurring naturally or non-naturally in a particular cell, tissue or organism. This includes DNA and fragments thereof, RNA and fragments thereof, cDNAs and fragments thereof, expressed sequence tags, artificial sequences including randomized artificial sequences.

Open reading frame ("ORF"). The portion of a cDNA that is translated into a protein. Typically, an open reading frame starts with an initiator ATG codon and ends with a termination codon (TAA, TAG or TGA).

Protein products: refers to the various peptides or polypeptides generated after translation of a single mRNA. Polypeptides refer to amino acids linked to form a peptide or a full length protein.

Recombinant: The term "recombinant" in association with "vector" refers to a vector which has been modified to contain a non-native exogenous nucleic acid.

Secreted protein: any protein that enters the cellular secretory pathway. Typically, secreted proteins are synthesized bearing a signal peptide which is removed shortly after its synthesis.

Signal peptide: a peptide capable of directing a nascent protein to the cell secretory pathway. It is generally accepted that a signal peptide is composed of an initiating methionine, a highly hydrophobic stretch, typically 10 to 15 residues long, and a signal peptidase cleavage site.

Transcription unit: As used herein, a "transcription unit" refers to a nucleic acid which comprises an enhancer sequence, a promoter sequence, and a transcription termination sequence, all operably linked together. Preferably, the enhancer and promoter sequences are constitutively active in various hosts. Enhancer and promoter sequences can be derived for example from the cytomegalovirus (CMV) immediate-early genes or from the Rous sarcoma virus (RSV) long terminal repeat. Preferably, the transcription unit is comprised within a vector.

Transfection: the process of introducing nucleic acids in eukaryotic cells by any means such as electroporation, lipofection, precipitate uptake, micro-injection. A cell having incorporated an exogenous nucleic acid is said to be transfected.

Vector: A RNA or DNA molecule which can be used to transfer an RNA or DNA segment from one organism to another. Vectors are particularly useful for manipulating genetic constructs and different vectors may have properties particularly appropriate to express protein(s) in a recipient during cloning procedures and may comprise different selectable markers. Bacterial plasmids are commonly used vectors.

B) General overview of the invention

The invention is based on isolated nucleic acids encoding bone polypeptide-1. The invention encompasses isolated or substantially purified nucleic acid or protein compositions. In the context of the present invention, an "isolated" or "substantially purified" DNA molecule or an "isolated" or "substantially purified" polypeptide is a DNA molecule or polypeptide that, by the hand of man,

exists apart from its native environment and is therefore not a product of nature. An isolated DNA molecule or polypeptide may exist in a purified form or may exist in a non-native environment such as, for example, a transgenic host cell. An isolated or purified DNA or polypeptide may be synthesized chemically, may be produced using recombinant DNA techniques and then isolated or purified or may be isolated or purified from its natural host. An "isolated" or "substantially purified" nucleic acid molecule or protein, or biologically active portion thereof, is substantially free of other cellular material, or culture medium when produced by recombinant techniques and, in some circumstances, further purified, or substantially free of chemical precursors or other chemicals when chemically synthesized. Preferably, an "isolated" nucleic acid is free of sequences (preferably protein encoding sequences) that naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequences that naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. A protein that is substantially free of cellular material includes preparations of protein or polypeptide having less than about 30%, 20%, 10%, 5%, (by dry weight) of contaminating protein. When the protein of the invention, or biologically active portion thereof, is recombinantly produced, preferably culture medium represents less than about 30%, 20%, 10%, or 5% (by dry weight) of chemical precursors or non-protein of interest chemicals.

i) *Cloning of a cDNA fragment encoding bone polypeptide-1*

We have previously developed a screening system that allows the rapid identification of nucleic acids encoding signal peptides from complex libraries of cDNA fragments (International patent application No. PCT/CA01/01169; US patent application No. 09/641,931). A library enriched in 5' fragments of cDNAs derived from developing calvaria was obtained using the so-called "oligo-capping" method (Maruyama and Sugano, 1994). Briefly, the top halves of skulls of e15.5 mouse embryos were dissected free from nervous or skin tissue. Total RNA was extracted

using the guanidium/phenol method and mRNA was isolated using Oligotex mRNA kit (Qiagen, Mississauga, Canada). A synthetic oligoribonucleotide (RNA30-1; 5'-agcaucgagucggccuuguuggccuacugg-3') SEQ ID NO:29 was specifically ligated to the 5' extremities of mRNAs whose CAP moiety had been converted to a free phosphate group by the action of tobacco acid pyrophosphatase. The first strand of corresponding cDNA was synthesized using a primer-linker comprising a random sequence of 9 nucleotides (36-2V; 5'-gagagagagagagcgactcggatccannnnnnnnnc-3') SEQ ID NO 30. The resulting first strand cDNA was amplified by 30 cycles of PCR using primers complementary to both extremities (18-114V; 5'-agcatcgagtcggccttg-3') SEQ ID NO: 31 and (18-99V; 5'-gagagcgactcggatcca-3') SEQ ID NO: 32. Amplicons were purified and subjected to a further 10 cycles of semi-nested PCR using primers 18-99V and 32-4V (5'-ggacgagcggccgcgccttggtggcctactgg-3') SEQ ID NO:33. PCR products were digested with NotI and BamHI, size-selected by electrophoresis (500-1200bp) and directionally cloned in an expression/screening vector digested with NotI and BamHI. The library numbered approximately 3,000,000 primary clones. Sequencing of randomly chosen inserts revealed that the average insert size was 575 bp and that 60% of inserts corresponded to 5' fragments of cDNAs. The library was expression screened in BHK-21 fibroblasts. Selected fragments were amplified, cloned in the pBluescript KS II+ vector and sequenced using a CEQ2000TM automated sequencer.

One of the retrieved fragments (ca155-cap1-c2-neg306, 433 bp in length) encoded a putative translation product of 124 residues containing a signal peptide. Interestingly, this translation product harbored sequences reminiscent of dibasic cleavage sites found in peptide hormone precursor (KKKR at position 76-79 and KKR at position 110-112). This observation prompted us to determine the expression profile of the corresponding gene by Northern analysis of RNA extracted from 52 different tissues and cell lines. No signal was detected after an exposure of 7 days; a representative autoradiograph is shown in Figure 1A. In order to verify that the retrieved fragment derived from a cellular transcript, Northern analysis was performed with 1.5 µg of polyA⁺ RNA extracted from e15.5 mouse calvaria, adult liver and adult kidney (Figure 1B). A clear signal (arrow; 25),

corresponding to a mRNA of approximately 1.3 kb, was specifically detected in the embryonic calvaria. Taken together, these results indicate that ca155-cap1-c2-neg306 is derived from a mRNA that encodes a bone protein that is specifically expressed in bone tissue.

ii) Results from database mining; cloning of full length mouse cDNA containing c155-cap1-c2-neg306

Because of the interesting features of clone ca155-cap1-c2-neg306, we carried out extensive database searches to compare its sequence with those deposited in the public domain. We used the standard Blastn tool (v2.2.4, August 26, 2002) with the following parameter set : expect value of 10, low complexity filter. There was only one significant match to an entry (accession XM_155941.1) in the non redundant (nr) set of the GenbankTM database. This clone (designated LOC239790) encodes a putative protein whose N terminal sequence is identical to that of ca155-cap1-c2-neg306 but diverges at residues Gln¹⁰². Searches in dbEST revealed that ca155-cap1-c2-neg306 shows high homology to a bovine EST (77.5% homology, accession number BF045261), to an uncharacterized mouse EST (accession number BB638598.1), to 3 human ESTs from metastatic chondrosarcoma (accession numbers BQ021661, BQ001512, BQ000995) and that the last 102 nucleotides of ca155-cap1-c2-neg306 shows high homology (85.8%) to the last 105 nucleotides of a rat EST (accession number AI178209). It should be noted that a valid ORF can not be deduced from the sequences of either mouse, rat or human ESTs. The 526 bp bovine EST contained an ORF of 132 residues. *In silico* assembly of ca155-cap1-c2-neg306 and the rat EST yielded a 0.99 kb chimeric sequence, roughly corresponding to the size of the expected full length transcript minus polyA tail (see Northern analysis, Figure 1B).

Since we had initially cloned the 5' end of ca155-cap1-c2-neg306, we used a modified 3' RACE strategy to obtain a full length cDNA (see Materials and Methods in the Examples section). Sequences from at least five different clones were aligned together with ca155-cap1-c2-neg306 to reconstitute the full length

consensus cDNA sequence shown on Figure 2. The 1280 bp cDNA (SEQ ID NO:1) contains an ORF of 393 bp (uppercase) flanked by 61 bp and 811 bp of untranslated sequences at the 5' and 3' ends, respectively (lowercase). A polyadenylation signal is found 14 bp upstream of a polyA stretch. The putative initiator ATG codon is found in an adequate Kozak context (AAGATGC, SEQ ID NO: 25).

iii) Expression profiling on histological sections of mouse embryos

To determine which bone cells express the cDNA from which ca155-cap1-c2-neg306 derives, *in situ* hybridization was performed on sections of e13.5 and e16.5 mouse embryos using standard procedures. Consecutive sections were hybridized with a probe partially complementary to Cbfa1 transcripts. Cbfa1 is a transcription factor expressed in cells of the osteoblast lineage and whose activity is essential for proper development of the skeleton (Ducy, 2000). As shown in Figure 3B, the ca155-cap1-c2-neg306 probe specifically labels cells surrounding the bone collar in e16.5 limbs. These cells are located in the periosteum and are cells of the osteoblast lineage. This is revealed by the fact that they express Cbfa1 (Figure 3C). At e13.5, expression of ca155-cap1-c2-neg306 is seen in a few cells in the vicinity of cartilaginous formations in the skull (Figure 3E). These cells are also labeled by the Cbfa1 probe (Figure 3F). Thus, the cDNA from which ca155-cap1-c2-neg306 derives is expressed in cells of the osteoblast lineage.

iv) Cloning of vertebrate homologs

Because of its features, the cDNA from which ca155-cap1-c2-neg306 derives will be referred hereafter as mouse BP-1, mouse bone polypeptide-1. As discussed in more detail below, various homologs, including at least one human homolog to BP-1 exist. Bone polypeptide-1, as defined herein, refers to a polypeptide expressed in bone. As discussed in more detail below, bone polypeptide-1 may regulate bone cell proliferation and/or bone cell differentiation and/or osteoblast activity via an autocrine, paracrine or endocrine pathway.

BP-1 DNA sequences

The BP-1 DNA used in any embodiment of this invention can be BP-1 cDNA obtained as described herein, or alternatively, can be any oligonucleotide sequence having all or a portion of a sequence represented herein, or their functional equivalents. Such oligodeoxynucleotide sequences can be produced chemically or mechanically, using known techniques.

The following terms are used to describe the sequence relationships between two or more nucleic acids or polynucleotides: (a) "reference sequence", (b) "comparison window", (c) "sequence identity", (d) "percentage of sequence identity", and (e) "substantial identity".

(a) As used herein, "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset or the entirety of a specified sequence; for example, as a segment of a full length BP-1 cDNA or gene sequence, or the complete cDNA or gene sequence.

(b) As used herein, "comparison window" makes reference to a contiguous and specified segment of a polynucleotide sequence, wherein the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Generally, the comparison window is at least 20 contiguous nucleotides in length, and optionally can be 30, 40, 50, 100, or longer. Those of skill in the art understand that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence a gap penalty is typically introduced and is subtracted from the number of matches.

Methods of alignment of sequences for comparison are well known in the art. Thus, the determination of percent identity between any two sequences can be accomplished using a mathematical algorithm. Preferred, non-limiting examples of such mathematical algorithms are the algorithm of Myers and Miller, 1988; the local homology algorithm of Smith et al. 1981; the homology alignment algorithm of Needleman and Wunsch 1970; the search-for-similarity-method of

Pearson and Lipman 1988; the algorithm of Karlin and Altschul, 1990, modified as in Karlin and Altschul, 1993.

Computer implementations of these mathematical algorithms can be utilized for comparison of sequences to determine sequence identity. Such implementations include, but are not limited to: CLUSTAL in the PC/Gene program (available from Intelligenetics, Mountain View, California); the ALIGN program (Version 2.0) and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Version 8 (available from Genetics Computer Group (GCG), 575 Science Drive, Madison, Wisconsin, USA). Alignments using these programs can be performed using the default parameters. The CLUSTAL program is well described by Higgins et al. 1988; Higgins et al. 1989; Corpet et al. 1988; Huang et al. 1992; and Pearson et al. 1994. The ALIGN program is based on the algorithm of Myers and Miller, *supra*. The BLAST programs of Altschul et al., 1990, are based on the algorithm of Karlin and Altschul *supra*.

Software for performing BLAST analyses is publicly available through the web site for National Center for Biotechnology Information (NCBI). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al., 1990). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when the cumulative alignment score falls off by the quantity X from its maximum achieved value, the cumulative score goes to zero or below due to the accumulation of one or more negative-scoring residue alignments, or the end of either sequence is reached.

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul (1993). One measure of similarity provided by the BLAST algorithm is the smallest sum probability ($P(N)$), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a test nucleic acid sequence is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid sequence to the reference nucleic acid sequence is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

To obtain gapped alignments for comparison purposes, Gapped BLAST (in BLAST 2.0) can be utilized as described in Altschul et al. 1997. Alternatively, PSI-BLAST (in BLAST 2.0) can be used to perform an iterated search that detects distant relationships between molecules. See Altschul et al., *supra*. When utilizing BLAST, Gapped BLAST, PSI-BLAST, the default parameters of the respective programs (e.g. BLASTN for nucleotide sequences, BLASTX for proteins) can be used. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, $M=5$, $N=-4$, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, 1989). See the NCBI web site. Alignment may also be performed manually by inspection.

For purposes of the present invention, comparison of BP-1 nucleotide sequences for determination of percent sequence identity to the BP-1 sequences disclosed herein is preferably made using the BlastN program (version 1.4.7 or later) with its default parameters or any equivalent program. By "equivalent program" is intended any sequence comparison program that, for any two sequences in question, generates an alignment having identical nucleotide or amino acid residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by the preferred program.

(c) As used herein, "sequence identity" or "identity" in the context of two BP-1 nucleic acid or polypeptide sequences makes reference to the residues in the

two sequences that are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. When sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have "sequence similarity" or "similarity." Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California).

(d) As used herein, "percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

(e)(i) The term "substantial identity" of polynucleotide sequences means that a BP-1 polynucleotide comprises a sequence that has at least 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, or 79%, preferably at least 80%, 81%, 82%, 83%, 84%, 85%,

86%, 87%, 88%, or 89%, more preferably at least 90%, 91%, 92%, 93%, or 94%, and most preferably at least 95%, 96%, 97%, 98%, or 99% sequence identity, compared to a reference sequence using one of the alignment programs described using standard parameters. One of skill in the art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning, and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 70%, more preferably at least 80%, 90%, and most preferably at least 95%.

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions (see below). Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. However, stringent conditions encompass temperatures in the range of about 1°C to about 20°C, depending upon the desired degree of stringency as otherwise qualified herein. Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides they encode are substantially identical. This may occur, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. One indication that two nucleic acid sequences are substantially identical is when the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the polypeptide encoded by the second nucleic acid.

(e)(ii) The term "substantial identity" in the context of a BP-1 peptide indicates that a peptide comprises a sequence with at least 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, or 79%, preferably 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, or 89%, more preferably at least 90%, 91%, 92%, 93%, or 94%, or even more preferably, 95%, 96%, 97%, 98% or 99%, sequence identity to the reference sequence over a specified comparison window. Preferably, optimal alignment is conducted using the homology alignment algorithm of Needleman and Wunsch (1970). An indication that two peptide sequences are substantially identical is that

one peptide is immunologically reactive with antibodies raised against the second peptide. Thus, a peptide is substantially identical to a second peptide, for example, where the two peptides differ only by a conservative substitution.

For sequence comparison, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

As noted above, another indication that two BP-1 nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions. The phrase "hybridizing specifically to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (e.g., total cellular) DNA or RNA. "Bind(s) substantially" refers to complementary hybridization between a probe nucleic acid and a target nucleic acid and embraces minor mismatches that can be accommodated by reducing the stringency of the hybridization media to achieve the desired detection of the target nucleic acid sequence.

"Stringent hybridization conditions" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments such as Southern and Northern hybridization are sequence dependent, and are different under different environmental parameters. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl, 1984; $T_m = 81.5^\circ\text{C} + 16.6 (\log M) + 0.41 (\%GC) - 0.61 (\% \text{ form}) - 500/L$; where M is the molarity of monovalent cations, $\%GC$ is the percentage of guanosine and cytosine nucleotides in the DNA, $\% \text{ form}$ is the percentage of formamide in the hybridization solution, and L is the length of the

hybrid in base pairs. T_m is reduced by about 1°C for each 1% of mismatching; thus, T_m , hybridization, and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with >90% identity are sought, the T_m can be decreased 10°C. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point T_m for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4°C lower than the thermal melting point T_m ; moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10°C lower than the thermal melting point T_m ; low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20°C lower than the thermal melting point T_m . Using the equation, hybridization and wash compositions, and desired T_m , those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a T_m of less than 45°C (aqueous solution) or 32°C (formamide solution), it is preferred to increase the SSC concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen, 1993. Generally, highly stringent hybridization and wash conditions are selected to be about 5°C lower than the thermal melting point T_m for the specific sequence at a defined ionic strength and pH.

An example of highly stringent wash conditions is 0.15 M NaCl at 72°C for about 15 minutes. An example of stringent wash conditions is a 0.2X SSC wash at 65°C for 15 minutes (see, Sambrook, *infra*, for a description of SSC buffer). Often, a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example medium stringency wash for a duplex of, e.g., more than 100 nucleotides, is 1X SSC at 45°C for 15 minutes. An example low stringency wash for a duplex of, e.g., more than 100 nucleotides, is 4-6X SSC at 40°C for 15 minutes. For short probes (e.g., about 10 to 50 nucleotides), stringent conditions typically involve salt concentrations of less than about 1.5 M, more preferably about 0.01 to 1.0 M, Na ion concentration (or other salts) at pH 7.0 to 8.3, and the temperature is typically at least about 30°C and at least about 60°C for long probes (e.g., >50 nucleotides). Stringent conditions may also be

achieved with the addition of destabilizing agents such as formamide. In general, a signal to noise ratio of 2X (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization. Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the proteins that they encode are substantially identical. This occurs, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

Very stringent conditions are selected to be equal to the T_m for a particular probe. An example of stringent conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or Northern blot is 50% formamide, e.g., hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulphate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1.0 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C.

The following are examples of sets of hybridization/wash conditions that may be used to clone orthologous nucleotide sequences that are substantially identical to reference nucleotide sequences of the present invention: a reference nucleotide sequence preferably hybridizes to the reference nucleotide sequence in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 2X SSC, 0.1% SDS at 50°C, more desirably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 1X SSC, 0.1% SDS at 50°C, more desirably still in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.5X SSC, 0.1% SDS at 50°C, preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.1X SSC, 0.1% SDS at 50°C, more preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.1X SSC, 0.1% SDS at 65°C.

Protein sequences

The present invention includes peptides which are derivable from BP-1. A peptide is said to be "derivable from a naturally occurring BP-1 amino acid sequence" if it can be obtained by fragmenting a naturally occurring sequence, or if it can be synthesized based upon a knowledge of the sequence of the naturally occurring BP-1 amino acid sequence or of the genetic material (DNA or RNA) which encodes this sequence.

Included within the scope of the present invention are those molecules which can be said to be "derivatives" or "active fragments" of BP-1 referred herein as "BP-1 products". Such a "derivative" or "active fragment" has one or more of the following characteristics: (1) it shares substantial homology with BP-1 or a similarly sized fragment of BP-1; (2) it is capable of functioning as a bone hormone and (3) using at least one of the assays provided herein, the derivative has either (i) a bone hormone activity, or, (ii) an activity on the kidney. Bone hormone activity refers to an ability to regulate bone cell proliferation and/or bone cell differentiation and/or of osteoblast activity, via an autocrine, paracrine or endocrine pathways.

A derivative of BP-1 is said to share "substantial homology" with BP-1 if the amino acid sequences of the derivative is at least 60%, and more preferably at least 80%, and most preferably at least 90%, the same as that of BP-1.

The derivatives of the present invention include BP-1 fragments which, in addition to containing a sequence that is substantially homologous to that of a naturally occurring BP-1 peptide may contain one or more additional amino acids at their amino and/or their carboxy termini. Thus, the invention pertains to polypeptide fragments of BP-1 that may contain one or more amino acids that may not be present in a naturally occurring BP-1 sequence. The additional amino acids may be D-amino acids or L-amino acids or combinations thereof. Furthermore, the additional amino acids may be naturally occurring amino acids or non-naturally occurring amino acids such as L-tert-leucine; L-homophenylalanine; D-homophenylalanine; D-methionine; Halogenated D and L-phenylalanines, tyrosines, and tryptophans; D-2-aminopimelic acid and L-2-aminopimelic acid

The invention also includes BP-1 fragments which, although containing a

sequence that is substantially homologous to that of a naturally occurring BP-1 peptide may lack one or more additional amino acids at their amino and/or their carboxy termini that are naturally found on a BP-1 peptide. Thus, the invention pertains to polypeptide fragments of BP-1 that may lack one or more amino acids that are normally present in a naturally occurring BP-1.

The invention also encompasses the obvious or trivial variants of the above-described fragments which have inconsequential amino acid substitutions (and thus have amino acid sequences which differ from that of the natural sequence) provided that such variants have a bone hormone activity which is substantially identical to that of the above-described BP-1 derivatives. Examples of obvious or trivial substitutions include the substitution of one basic residue for another (i.e. Arg for Lys), the substitution of one hydrophobic residue for another (i.e. Leu for Ile), or the substitution of one aromatic residue for another (i.e. Phe for Tyr), etc.

Bone Polypeptide-1 Homologs

A cDNA encoding BP-1, or a portion thereof can be used to identify similar sequences in other vertebrate species and thus, to identify or "pull out" sequences which have sufficient homology to hybridize to BP-1 cDNA or mRNA or a portion thereof under conditions of low stringency. Those sequences which have sufficient homology (generally greater than 40%) can be selected for further assessment using the method described herein. Alternatively, high stringency conditions can be used as discussed above. In this manner, DNA of the present invention can be used to identify, in other species, sequences encoding polypeptides having amino acid sequences similar to that of BP-1 and thus to identify bone polypeptides in other species. Thus, the present invention includes not only BP-1, but also related proteins encoded by DNA which hybridizes, preferably under high stringency conditions, to DNA of the present invention.

BLAST searches of the mouse BP-1 sequence using the Ensembl web site (containing the sequence of the human genome, v8.30a.1, August 2002 data) revealed that a human homolog of BP-1 might exist. Indeed, homologous sequences are found on human chromosome 3q28 in contigs AC019234.5.48926.64909 (nt 12659-12937, 75% homology),

AC019234.5.130697.209258 (nt 75405-75529, 79% homology), AC019234.5.89412.130596 (nt 39498-39560, 92% homology). The Genescan algorithm predicts 4 exons but those are disordered and not annotated as part of a single gene and, furthermore, the predicted translation product do not match the primary sequence of BP-1, presumably because Genescan has not identified the correct ORF. Thus, the ORF predicted from the BP-1 cDNA (Figure 2) can not be determined solely by *in silico* analysis of the sequence of the human genome. Therefore, in order to clone a human homolog of BP-1, we performed RT-PCR on human bone marrow mRNA using two primers encompassing the putative initiator and termination codons found in the human genome sequence. In addition, the rat homolog of BP was cloned by low-stringency RT-PCR using one primer encompassing the putative initiator codon and another encompassing a region in the 3' untranslated region of the mouse cDNA. To identify evolutionarily conserved domains within the BP-1 protein, we obtained sequence information on BP-1 from non mammalian species. The cDNA encoding chicken (*Gallus gallus*) BP-1 was retrieved by 'BLASTing' the human ORF against the BBSRC Chicken EST Project database (Blastn tool, all tissues, matrix BLOSUM62, expectation 10^{-2}). A portion of the cDNA encoding snake (*Python molurus bivittatus*) BP-1 was cloned by RT-PCR using degenerate oligonucleotide primers and starting from RNA extracted from the vertebrae of a young python (See Materials and Methods in the Examples section.) Figure 4 shows the alignment of the nucleic acid sequences encoding human, bovine, mouse, rat, chicken, and python BP-1. Allelic variations of the nucleic acid sequences described herein are also encompassed in the invention.

Figure 5A shows the alignment of the ORFs derived from the BP-1 cDNAs of mouse, human, rat, cow, chicken. The partial ORF deduced from the python sequence is also aligned. The presence of a cleavable signal peptide as well as the position of the putative processing sites are conserved across species (respectively underlined and boxed in Figure 5A). In silico analysis indicates that a) the region of the protein that lies between the 2 dibasic cleavage sites (residues 83-112 in the human sequence) is well conserved (overall identity of 60%, overall similarity of 76%); b) the C terminal region of the protein (residues 116-133 in the human sequence) is also well conserved (human and mouse share 94.4%

similarity and identity; human and chicken share 83.3% similarity and 61.1% identity). The observation that the dibasic cleavage sites are found in various species suggests that BP-1 is a prohormone precursor conserved in terrestrial vertebrates. Furthermore, analysis of the conserved domains suggests that the bioactive protein products derive from the C terminal half of BP-1.

In silico analysis indicates that full length human BP-1 has the following feature: it has a molecular weight of 14,832 g/mol and an isoelectric point of 9.62; it contains neither N-glycosylation sites nor disulfide bridges. Despite extensive searches in public databases (e.g. InterPro, ProDom, STN DGENE), BP-1 shows no significant homology to any known protein, protein motif or protein domain. We have found however that residues 116-124 of BP-1 (numbering from the human sequence) show some homology to members of the natriuretic peptides family. Figure 5B shows the alignment of the homologous regions between BP-1 and natriuretic peptides. The 2 cysteine residues conserved in all natriuretic peptides form an intramolecular disulfide bridge essential for their bioactivity (Inagami et al., 1985). Cysteine residues involved in cyclization are not found in BP-1. Interestingly however, it has been reported that a synthetic 'linear' analog of the atrial natriuretic peptide could displace the binding of atrial natriuretic factor to its receptor but failed to activate synthesis of second messengers (Olins et al., 1988). Thus, it is possible that a BP-1 protein product could be related both biochemically and functionally to the natriuretic peptides.

v) *Production of antibodies against portions of BP-1.*

In order to detect the BP-1 protein products, antibodies were raised against synthetic peptides derived from 3 regions of the full length protein AG-1: (DELVSLENDVIETK (SEQ ID NO:26), AG-2: (RLSAGSVDPHKGKQR) SEQ ID NO: 27, and AG-3: (MDRIGRNRLSNSRG) SEQ ID NO: 28). The chosen regions have a high antigenicity index as predicted by the algorithm of Hopp and Woods (Hopp and Woods, 1981). The positions of the antigenic peptides are indicated on the model of human BP-1 shown on Figure 6A. To increase immunogenicity, antigenic peptides comprise a N- or C-terminal cysteine to allow covalent coupling to a carrier protein (e.g. keyhole limpet hemocyanin, bovine serum albumin). The peptide/carrier complex are injected subcutaneously into animals (e.g. rabbits) and

antisera are obtained using standard protocols. (See Materials and Methods in Examples section.) Affinity chromatography was used to purify the fraction of immunoglobulins specific to the peptide used to elicit an immune response.

Both monoclonal and polyclonal antibodies are included within the scope of this invention as they can be produced by well established procedures known to those of skill in the art. Additionally, any secondary antibodies, either monoclonal or polyclonal, directed to the first antibodies would also be included within the scope of this invention.

BP-1 is Secreted Extracellularly and can be Processed

To begin to assess the biochemical and biological properties of BP-1, a cDNA encoding the mouse protein was inserted into a mammalian expression vector and this vector was transfected in HEK293A cells. Immunofluorescence analysis after cell permeabilization showed that BP-1 is mainly localized to the Golgi apparatus and in small cytoplasmic vesicles of transfected cells. No signal was detected at the cell surface. Western analysis of cell lysates and culture medium using antibodies raised against the C terminal region of human BP-1 (AG3) (SEQ ID NO: 28) revealed that the bulk of the protein products is secreted in the culture medium (Figure 6B, compare lanes 620 and 621). The major secreted form has an apparent molecular weight of 13kD, presumably corresponding to the BP-1 precursor after removal of the signal peptide. Upon longer exposure (Figure 6B, lane 622), a smaller product with an apparent molecular weight of ~6kD is reproducibly found in the culture medium of transfected HEK293A cells. This ~6kD band is absent when cells are transfected with a vector expressing a form of the mouse cDNA where the dibasic cleavage site at position 76-79 is mutated (KKKR -> AS, Figure 6B, lane 623).

Given these results, we predict that BP-1 is processed according to the following scheme. After translocation in the endoplasmic reticulum, the signal peptide is cleaved and the protein (PROT 1, SEQ ID NO: 20) may be further processed in the secretory pathway or extracellular space to yield bioactive products. Candidate processing enzymes include furin (Denault and Leduc, 1996; GenBank™ PID accession number g31478), a related convertase or, as discussed below, corin (GenBank™ PID accession number g5729989). Our results indicate

that the BP-1 precursor may be cleaved at arginine⁸² (numbering according to human sequence) in an heterologous expression system (e.g. HEK293A fibroblasts) to generate PEPT 1 (SEQ ID NO: 21). Additional processing may occur C terminal to arginine¹¹⁵ at a typical dibasic site to generate PEPT 2A (SEQ ID NO: 22) and PEPT 3A (SEQ ID NO: 23). Alternative processing may occur at a basic residue (lysine¹⁰⁴) upstream of the second dibasic site by an atypical convertase, similar to what has been reported for the processing of a natriuretic peptide precursor by the corin protease (Yan et al., 2000; Yan et al., 1999). This would generate PEPT 2B (SEQ ID NO: 24) and PEPT 3B (SEQ ID NO: 25). Interestingly in this latter case, both cleavage products of BP-1 end with a glycine residue. This observation raises the possibility that BP-1 products are amidated by peptidylglycine alpha-amidating monooxygenase (GenBankTM accession number no. BAC22594), an enzyme whose activity has been detected in mouse calvaria, a tissue that produces BP-1 (Birnbaum et al., 1989). Figure 6A illustrates the processing scenario for human BP-1 and depicts its putative protein products.

vi) *Methods of production of recombinant or synthetic BP-1 protein products.*

The present invention provides expression vectors and host cells transformed to express the nucleic acid sequences encoding BP-1 of the invention. Expression vectors of the invention comprise a nucleic acid sequence coding for at least one BP-1, or at least one antigenic fragment thereof, or derivative or homologue thereof, or the functional equivalent of such nucleic acid sequence. Nucleic acid sequences coding for BP-1, or at least one fragment thereof may be expressed in prokaryotic or eukaryotic host cells. Suitable host cells include bacterial cells such as *E. coli*, insect cells, yeast, or mammalian cells such as Chinese hamster ovary cells (CHO). Suitable expression vectors, promoters, enhancers, and other expression control elements may be found in Sambrook et al. *Molecular Cloning: A Laboratory Manual*, second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989.

Whatever the producing host, the BP-1 portion of the transcription unit may be engineered to increase or control the stability or translatability of the expressed

mRNA, e.g. through the insertion of exogenous *cis*-acting nucleic acid sequences, the replacement of poorly translated codons, the removal of untranslated sequences and/or the engineering of a proper context for the initiation of translation (Kozak, 1986). The BP-1 cDNA may also be engineered to increase the yield of secreted products, typically by replacing its signal peptide coding sequence by an exogenous nucleic acid encoding a more active signal peptide in the chosen host. Furthermore, minor modifications or variations may be introduced in the BP-1 coding sequence to enhance the stability and/or the activity of BP-1 product(s). These modifications may be made by site-directed mutagenesis or may be the result of spontaneous mutations.

Vertebrate cell lines

An aspect of this invention relates to modified cell lines and transgenic organisms used to produce BP-1 protein products. According to a preferred embodiment, there is provided an eukaryotic cell line having incorporated in its genome a DNA segment comprising a BP-1 coding sequence operably linked within a transcription unit. Enhancer and promoter sequences driving robust expression in a wide variety of cells are generally preferred. These include but are not limited to sequences derived from cytomegalovirus immediate-early genes (CMV; e.g. GenBank™ accession no. X03922) and Rous sarcoma virus long terminal repeat (RSV; e.g. GenBank™ accession no M83237) as well as sequences derived from widely expressed cellular genes such as chicken β -actin, human elongation factor 1 α , phosphoglycerate kinase, and metallothionein. Signals for the termination and polyadenylation of transcripts are well known in the art. Examples include part of the 3' untranslated region of the bovine growth hormone gene or of the SV40 virus.

Methods to incorporate DNA segments into the genome of a cell are well known in the art. According to a preferred embodiment of the invention, the transcription unit is inserted into a plasmid containing a gene conferring resistance to a selective agent (e.g. puromycin-N-acetyltransferase conferring resistance to puromycin). The resulting construct is electroporated into cells using standard protocols. Transfection by lipofection, calcium phosphate precipitate, and micro-

injection are other techniques available to introduce nucleic acids into eukaryotic cells. Selection is applied on the pool of transfected cells. Surviving and growing cells are thought to have incorporated the plasmid and are cloned. Individual clones are analyzed for the production of BP-1 protein products in the culture medium. Typically, the levels of BP-1 products are determined by Western analysis or ELISA. Preferred cellular hosts include, but are not limited to, human embryonic kidney 293 cells (HEK293; American Type Culture Collection no. CRL-1573) and chinese hamster ovary cells (CHO; American Type Culture Collection no. CCL-61). It is known in the prior art how to cultivate large quantities of these cells in order to obtain large amounts of recombinant proteins. It is understood that a BP-1-producing cell line can be further engineered to express a convertase involved in BP-1 processing, thereby allowing the release of bioactive BP-1 products into the culture medium.

Transgenic organisms

Alternatively, the transcription unit comprising the BP-1 coding sequence can be inserted into a fertilized egg (e.g. of a mouse), which is re-implanted into a pseudo-pregnant mother. DNA extracted from resulting organisms (e.g. embryos, pups or adults) is analyzed by Southern blotting to determine whether the organism is transgenic. Positive animals are bred and used to produce large quantities of BP-1 products. Ideally, the recombinant protein products should be produced in an easily collectable tissue or biological fluid (e.g. hair, milk). Furthermore, expression in tissues that are targets of BP-1 action (e.g. bone, kidney) should be limited. In order to achieve these goals, the enhancer and promoter elements of the transcription unit are chosen so that the BP-1 coding sequence is mainly expressed in the appropriate tissue (e.g. keratinocytes, mammary epithelium).

Fusion protein in bacteria

For some aspects of the present invention, it is desirable to produce a fusion protein comprising a BP-1 polypeptide or at least one fragment thereof or their derivatives and an amino acid sequence from another peptide or protein,

examples of the latter being enzymes such as beta-galactosidase, phosphatase, urease and fusion proteins incorporating purification moieties such as His-tags, FLAG-tags, myc-epitope tags and the like. Most fusion proteins are formed by the expression of a recombinant gene in which two coding sequences have been joined together such that their reading frames are in phase.

For expression in *E. coli*, suitable expression vectors include pTRC (Amann et al. (1988) Gene 69: 301-315); pET-11d (Novagen, Madison, Wis.); pGEX (Amrad Corp., Melbourne, Australia); pMAL (N.E. Biolabs, Beverly, Mass.); pRIT5 (Pharmacia, Piscataway, N.J.); PSEM (Knapp et al. (1990) BioTechniques 8: 280-281); pQE30 (Qiagen, Germany); and pTrxFus (Invitrogen, Carlsbad, Ca.). The use of pTRC and pET-11d will lead to the expression of unfused protein. The use of pGEX, pMAL, pRIT5, pSEM, pQE30, and pTrxFus will lead to the expression of BP-1 fused to glutathione S-transferase (pGEX), maltose E binding protein (pMAL), protein A (pRIT5), truncated β -galactosidase (PSEM), hexahistidine (His6), or thioredoxin. When a BP-1, fragment, or fragments thereof is expressed as a fusion protein, it is particularly advantageous to introduce an enzymatic cleavage site at the fusion junction between the carrier protein and the BP-1 or fragment thereof. A BP-1 or fragment thereof may then be recovered from the fusion protein through enzymatic cleavage at the enzymatic site and biochemical purification using conventional techniques for purification of proteins and peptides. Suitable enzymatic cleavage sites include those for blood clotting Factor Xa or thrombin for which the appropriate enzymes and protocols for cleavage are commercially available from for example Sigma Chemical Company, St. Louis, Mo. and N.E. Biolabs, Beverly, Mass.

Host cells can be transformed to express the nucleic acid sequences encoding the BP-1 of the invention using conventional techniques such as calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, or electroporation. Suitable methods for transforming the host cells may be found in Sambrook et al. supra, and other laboratory textbooks. The nucleic acid sequences of the invention may also be synthesized using standard techniques.

Chemical synthesis

Homologues and derivatives of BP-1 is meant to include synthetic derivatives thereof. The nucleotide sequences encoding BP-1 can be used to chemically synthesize the entire protein or generate any number of fragments (peptides) by chemical synthesis by well known methods (e.g., solid phase synthesis). All such chemically synthesized peptides are encompassed by the present invention. Alternatively, proteins or peptides can be linked in vitro by chemical means. All such fusion protein or hybrid genetic derivatives of BP-1 or its encoding nucleotide sequences are encompassed by the present invention. Accordingly, the present invention extends to isolated BP-1, fragments thereof and their derivatives, homologues and immunological relatives made by recombinant means or by chemical synthesis.

vii) *Methods of purification of recombinant BP-1 protein products.*

BP-1 and fragments (peptides) thereof can be purified from cell culture medium, host cells, or both using techniques known in the art for purifying peptides and proteins, including ion-exchange chromatography, hydrophobic chromatography, gel filtration chromatography, ultrafiltration, electrophoresis and immunopurification with antibodies specific for BP-1 or fragments thereof. The terms isolated and purified are used interchangeably herein and refer to peptides, protein, protein fragments, and nucleic acid sequences substantially free of cellular material or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when synthesized chemically.

Typically, recombinant proteins found in culture medium or biological fluids are purified by successive steps of chromatography according to methods known to an experimentator skilled in the art. According to a preferred embodiment of the invention, accumulation of secreted recombinant BP-1 is done by incubating engineered mammalian cells in serum-free medium, in order to limit the amount of contaminating proteins found in the medium. This medium is then diluted with column buffer (neutral HEPES-buffered saline) and passed on a cation-exchange column (e.g. SepharoseTM SP, Amersham Pharmacia Biosciences). Since the

isoelectric point of human BP-1 is 9.62, it binds strongly to the negatively-charged resin in column buffer. After extensive washes with column buffer, bound proteins are eluted with a salt gradient. Fractions are collected and assayed for BP-1 immunoreactivity by ELISA. Under these conditions, the BP-1 products elute at a salt concentration of 250-300 mM. Further purification is achieved by affinity chromatography. Monoclonal or polyclonal antibodies specific to BP-1 (e.g. raised against AG2, see section v) are covalently attached to a resin, either directly (e.g. CNBR-SepharoseTM) or through binding and crosslinking to immobilized protein A. In the latter case, protein A specifically binds the Fc portion of the antibodies, thereby avoiding crosslinking/inactivation of their F(ab)₂ portion to the resin and increasing the binding capacity of the affinity resin. BP-1-containing fractions from the ion-exchange column are pooled, diluted to a salt concentration of approximately 100 mM and passed over the affinity column at neutral pH. After washes, bound proteins are eluted under basic (pH 10) and acidic (pH 2.9) conditions. Fractions are collected, neutralized and assayed for BP-1 immunoreactivity by ELISA. BP-1-containing fractions are pooled and, if necessary, dialyzed against physiological saline and concentrated (e.g. by lyophilization).

viii) Recombinant adenoviruses expressing BP-1 cDNA for gene delivery

Since processing of BP-1 to bioactive products may occur uniquely in osteoblasts or within the bone microenvironment, it may be advantageous to express a BP-1 coding sequence in cultured osteoblasts or directly *in vivo*. Recombinant adenovirus particles are particularly useful to deliver a transcription unit into a whole organism or into cells that are difficult to transfect by conventional methods such as differentiated osteoblasts (Ragot et al., 1998). Standard methodology may be used to insert a transcription unit in an adenoviral genome and package the resulting recombinant adenoviral genome within an infectious viral particle. Briefly, the transcription unit comprising a BP-1 coding sequence is cloned in a plasmid between regions of the adenovirus serotype 5 genome. This plasmid is transfected, along with a replication-defective viral genome, in a cell line

that can complement the replication defect, usually HEK293 cells. Homologous recombination between a plasmid and a viral genome generates a recombinant viral genome having inserted a BP-1 transcription unit. This recombinant viral genome is subsequently packaged into infectious particles. Typically, the replication-defective viral genome minimally lacks the E1 protein such that it can not replicate in infected cells unless these are HEK293 cells that endogenously synthesize the E1 protein (Louis et al., 1997). The recombinant adenovirus can be propagated in HEK293 cells to very high titers ($>10^{10}$ pfu/ml). Stocks of recombinant adenoviral particles are concentrated and purified by centrifugation on cesium chloride gradient according to standard procedures.

The resulting recombinant adenovirus can be used to infect large numbers of various cell types, including primary osteoblasts, to produce recombinant BP-1. It can also be used as a delivery system for systemic or local gene therapy or other protocols that may require overexpression of BP-1 *in vivo*.

ix) Therapeutic utility of BP-1 products

Considering the fact that BP-1 is specifically expressed in the osteoblast lineage, BP-1 products may regulate bone cell proliferation and/or bone cell differentiation and/or of osteoblast activity, via an autocrine, paracrine or endocrine pathway. BP-1 products may also control osteoclast activity via a paracrine or endocrine pathway. In either cases, pharmaceutical compositions containing a therapeutically effective amount of BP-1 products may be useful in the treatment of bone diseases, particularly in those characterized by bone loss such as osteoporosis. The BP-1 compositions may be further employed in methods for treating bone fractures or defects. Drugs that augment, mimic, antagonize or blunt the activity of BP-1 products may also be beneficial for the treatment of bone diseases. Considering the *in vitro* effect of our BP-1 preparations (see Example 4), drugs that antagonize or blunt the activity of BP-1 products could be particularly useful to treat osteopenic or osteoporotic conditions.

Examples of drugs that augment the activity of BP-1 products include chemicals that activate proteins involved in the transcription of the BP-1 gene, thereby upregulating its expression. Drugs that inhibit the degradation of a given

BP-1 product should also lead to increased activity of this BP-1 product. Examples of drugs that mimick the activity of BP-1 products include peptidomimetics or peptides, modified or not, corresponding to fragments of BP-1 that regulate bone or kidney functions. In most cases, such drugs are agonists of the receptor that binds a given BP-1 product.

Examples of drugs that antagonize or blunt the activity of BP-1 products include a) one or a mixture of antisense oligonucleotides blocking the expression or translation of the BP-1 mRNA; b) one or a mixture of antibodies raised against a given BP-1 product quenching the activity of this product; c) an antagonist binding the cognate receptor of a given BP-1 product; d) a molecule inhibiting the enzymes responsible for the processing of the BP-1 precursor to a given product.

Phosphatonin is a putative hormone which regulates phosphate retention by kidney tubules. It has been hypothesized that phosphatonin is produced by cells of the osteoblast lineage and is a substrate for Phex, a metallopeptidase found at the cell surface of osteoblasts and osteocytes (Frota Ruchon et al., 2000; Ruchon et al., 1998; GenBankTM PID accession number g2499917). Our observations are consistent with the possibility that a given BP-1 product could have phosphatonin activity. If this is the case, then hypophosphatemia could be remedied by injection of a molecule that antagonize this activity. On the other hand, hyperphosphatemia could be remedied by injection of a molecule that stimulate or mimick phosphatonin activity. Alternatively, phosphatonin activity may be prolonged by using a drug that inhibits its degradation (e.g. inhibitors of the Phex endopeptidase activity). Such proteolysis inhibitors could be useful to treat hyperphosphatemic conditions. Thus, drugs that augment, mimick, antagonize or blunt the activity of BP-1 products may also be beneficial for the treatment of diseases characterized by abnormal serum phosphate levels.

Pharmaceutical Compositions

The present invention, therefore, provides a pharmaceutical composition comprising a therapeutically effective amount of BP-1 or derivatives, homologues or immunological relatives thereof and one or more pharmaceutically acceptable carriers and/or diluents. The active ingredients of a pharmaceutical composition comprising BP-1 is contemplated to exhibit therapeutic activity when administered

in amount which depends on the particular case. Dosage regime may be adjusted to provide the optimum therapeutic response. For example, several divided doses may be administered daily or the dose may be proportionally reduced as indicated by the exigencies of the therapeutic situation. The active compound may be administered in a convenient manner such as by the oral, intravenous (where water soluble), intramuscular, subcutaneous, intranasal, intradermal or suppository routes or implanting (e.g., using slow release molecules). Depending on the route of administration, the active ingredients which comprise the pharmaceutical composition of the invention may be required to be coated in a material to protect the ingredients from the action of enzymes, acids and other natural conditions which may inactivate said ingredients

The active compounds may also be administered parenterally or intraperitoneally. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders of the extemporaneous dispersion. In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The preventions of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of

the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and the freeze-drying technique which yield a powder of the active ingredient plus any additional desired ingredient from previously sterile-filtered solution thereof.

When at least BP-1, or at least one fragment thereof is suitably protected as described above, the active compound may be orally administered, for example, with an inert diluent or with an assimilable edible carrier, or it may be enclosed in hard or soft shell gelatin capsule, or it may be compressed into tablets, or it may be incorporated directly with food of the diet. For oral therapeutic administration, the active compound may be incorporated with excipients and used in the form of ingestible tablets, buccal tablets, troches, capsules, elixirs, suspensions, syrups, wafers, and the like. Such compositions and preparations should contain at least 1% by weight of active compound. The percentage of the compositions and preparations may, of course, be carried and may conveniently be between about 5 to 80% of the weight of the unit. The amount of active compound in such therapeutically useful compositions is such that a suitable dosage will be obtained.

The tablets, troches, pills, capsules and the like may also contain the following: A binder such as gum tragacanth, acacia, corn starch or gelatin; excipients such as dicalcium phosphate; a disintegrating agent such as corn starch, potato starch, alginic acid and the like; a lubricant such as magnesium stearate; and a sweetening agent such as sucrose, lactose or saccharin may be added or a flavoring agent such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to

materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar or both. A syrup or elixir may contain the active compound, sucrose as a sweetening agent, methyl and propylparabens as preservatives, a dye and flavoring such as cherry or orange flavor of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compound may be incorporated into sustained-release preparations and formulations.

As used herein "pharmaceutically acceptable carrier and/or diluent" includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, use thereof in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

It is especially advantageous to formulate parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the mammalian subjects to be treated; each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the novel dosage unit forms of the invention are dictated by and directly dependent on (1) the unique characteristics of the active material and the particular therapeutic effect to be achieved, and (b) the limitations inherent in the art of compounding such an active material for the treatment of bone disease.

x) *Screening methods using BP-1 products*

It is known in the prior art that proteins or peptides such as those derived from BP-1 act through binding to cognate receptors. It is possible to use labelled BP-1 products to identify such receptors. This could be done as follows. BP-1 (PROT 1, SEQ ID NO: 20) can be synthesized in the presence of labeled amino

acids (e.g. tritiated L-leucine). The protein can be synthesized and labeled *in vivo* after transfection of a suitable host with a vector capable of expressing the BP-1 cDNA. Alternatively, transcripts can be synthesized, translated and labeled *in vitro* starting with a plasmid harboring the BP-1 cDNA operably linked to a DNA-dependent RNA polymerase promoter. Ideally, the *in vitro* translation step should be performed using rabbit reticulocyte lysates containing microsomal membranes to allow for co-translational and post-translational modifications to occur (Lingappa et al., 1979). In the case of peptides derived from BP-1 (PEPT 2A (SEQ ID NO: 22), PEPT 3A (SEQ ID NO: 23), PEPT 2B (SEQ ID NO: 24), PEPT 3B (SEQ ID NO: 25)), these can be labeled during chemical synthesis (e.g. by incorporating a biotinylated derivative of amino acids) or by adding a N-terminal tyrosine residue that can be iodinated according to standard methods (Tsomides and Eisen, 1993). The labeled BP-1 product(s) can be purified (e.g. by affinity chromatography). To screen for cognate receptors, a cDNA library is constructed according to standard protocols starting with mRNA extracted from a target organ (e.g. bone, kidney). cDNAs are cloned in an expression vector downstream of strongly active enhancer/promoter elements. The library of expression vectors is screened for BP-1 product(s) binding activity as follows. The initial step is to find a suitable host, i.e. one that does not bind the labeled BP-1 product(s) under basal conditions and in the absence of exogenous expression vectors, preferably a mammalian cell line. Examples of hosts include, but are not limited to, HEK293 cells, CHO cells, COS cells (American Type Culture Collection no. CRL-1650) and CV-1 cells (American Type Culture Collection no. CCL-70). The library is then divided into pools (e.g. 500 clones per pool), each pool is transfected into the host and incubated with the labeled BP-1 product(s). Expression vectors are extracted from cells transfected with a pool of clones that confers specific binding of the labeled BP-1 product(s). Rounds of selection are repeated until a single clone is identified. This clone encodes a putative receptor for BP-1 product(s). The receptor can be used to screen for molecules that bind to it and for molecules that modulate the binding of BP-1 product(s). Such assays can be performed on cells transfected with a vector expressing the receptor cDNA.

xi) Diagnostic assays

Based on the features described in the invention, it can be expected that methods to assay BP-1 products or nucleic acid sequences encoding BP-1 products will be useful as diagnostic reagents i.e to diagnose or monitor diseases such as osteoporosis and disorders of phosphate metabolism. Levels of BP-1 products or expression of the BP-1 gene may correlate with the occurrence of a bone disease. For example, the amount of BP-1 products may be elevated in bones of osteoporotic patients. Various immunological methods to assay soluble extracellular proteins are well known in the prior art (e.g. radioimmunoassays, ELISA, 'sandwich' ELISA). Such methods generally uses monoclonal or polyclonal antibodies specific to the protein of interest. In the case of the present invention, these antibodies could be raised against antigenic peptides and purified as described in section v. Example 1 describes an enzyme-linked immunoadsorbent assay to quantify BP-1 products found in cell culture medium. Other immunological methods to assay BP-1 products in tissues or biological fluids may be developed by an experimentator skilled in the art. Various hybridization-based methods to assay for specific nucleic acids are well known in the prior art. In the case of BP-1 for example, RNA extracted from cells of the bone marrow can be labeled during reverse transcription (e.g. radioactively or fluorescently) and single stranded cDNA can be hybridized to an immobilized nucleic acid probe comprising part of the BP-1 sequence (e.g. oligonucleotide, cDNA).

Proteins, peptides, or antibodies of the present invention can also be used for detecting and diagnosing bone or renal conditions that involve secretion of BP-1. For example, diagnosis can be accomplished by combining blood obtained from an individual to be tested with antibodies to BP-1 and determining the extent to which antibody is bound to the sample.

Futhermore, it is expected that there are sequence polymorphisms in the nucleic acid sequence coding for BP-1, and it will be appreciated by one skilled in the art that one or more nucleotides in the nucleic acid sequence coding for BP-1 may vary due to natural allelic variation. Any and all such nucleotide variations and resulting amino acid polymorphisms are within the scope of the invention. It may also be appreciated by one skilled in the art that BP-1 maybe a member of a family

of highly related genes. Nucleotide sequences and corresponding deduced amino acid sequences of any and all such related family members including BP-1 are within the scope of the invention.

EXAMPLES

As it will now be demonstrated by way of examples hereinafter, the invention provides methods to produce, partially purify and characterize the BP-1 protein products. Example 1 gives an example of a cell line secreting BP-1 protein products in the culture medium. Example 2 gives an example of a method to partially purify the BP-1 protein products from cell culture medium. Example 3 gives an example of recombinant adenovirus particles expressing a BP-1 coding sequence and uses thereof. Example 4 gives an example of treatment of primary cultures of osteoblasts with medium containing BP-1 products.

Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described.

A) Materials and Methods

The following are experimental procedures and materials that were used to describe the present invention.

Enzymes and reagents

Restriction enzymes and DNA-modifying enzymes were purchased from New England Biolabs (Cambridge, Ma.) unless otherwise stated. Synthetic oligonucleotides were obtained either from Hukabel Ltd. (Montreal, Quebec, Canada) or MWG Biotech Inc. (High Point, North Carolina). Cell culture reagents were from Life Technologies (subsidiary of Invitrogen, Carlsbad, Ca.) unless otherwise stated. Chemicals were obtained either from Roche Molecular Biochemicals (Laval, Canada), Calbiochem (San Diego, Ca.) or Sigma (St. Louis, Mo.). Radiochemicals were purchased from Amersham Pharmacia Biotech (Baie

d'Urfé, Canada). Sequencing of DNA was done using the Dye terminator cycle sequencing DTCS™ sequencing kit (Beckman Coulter, Fullerton, Ca.) and a Beckman CEQ 2000 automated sequencer.

RNA extraction and analysis

Total RNA was purified either by the guanidium isothiocyanate/acid phenol method (Chomczynski and Sacchi, 1987) or using the RNEASY™ kit according to the manufacturer's instructions (Qiagen, Mississauga, Canada). PolyA+ RNA was purified using the Oligotex™ kit according to the manufacturer's instructions (Qiagen). Northern analysis were performed as follows. RNA was electrophoresed on 1.2% agarose/1.2% formaldehyde gels in 1x MOPS buffer (10mM 3-(N-morpholino)propanesulfonic acid, 4mM sodium acetate, 0.5mM EDTA, pH 7) and transferred onto a nylon membrane (Osmonics, Westborough, Ma.) by capillarity in 20x SSC. After UV crosslinking, the blot was incubated in Church buffer (Church and Gilbert, 1984) for 4-8 hours at 63°C and hybridized overnight in Church buffer containing approximately 1 ng/ml of ca155-cap1-c2-neg306 fragment that had been labelled with [α^{32} P]dCTP by random priming.

Cloning of full length mouse cDNA corresponding to ca155-cap1-c2-neg306

First, an expression vector was generated as follows. A 5498 bp SacII-ApaI fragment of pQBI25fc3 (qBiogene, Montreal, Canada) was blunted and ligated unto itself to generate pCMVneo. Complementary oligonucleotides 30-11 (SEQ ID NO. 35) and 30-12 (SEQ ID NO. 36) were annealed and cloned in pCMVneo previously digested with BamHI and EcoRI and blunted using the Klenow fragment of DNA polymerase I. The resulting plasmid is pCMVneoXN. The first strand of cDNAs was reverse transcribed at 42°C from 5µg of e15.5 mouse calvaria total RNA using a dT₁₅ primer-linker (33-1V, SEQ ID NO: 37) and Superscript II™ (Invitrogen, Carlsbad, Ca.). After RNase H treatment (Roche Molecular Biochemicals, Laval, Canada), 3/10¹ of the first strand cDNA preparation was subjected to 25 cycles of PCR using DNA polymerases from the Titan™ RT-PCR kit, forward primer 25-10G (25-10G, 5'-aaacgctctgacttctcacaagatg-3', SEQ ID NO. 38) and reverse primer 18-30V (18-30V, 5'-gagatgaattcctcgagc-3', SEQ ID NO.

39). Cycling conditions were as follows: 94°C for 45 seconds, 54°C for 30 seconds, 68°C for 3 minutes. PCR products were cloned in pCMVneoXN to generate GD25. Bacterial colonies were hybridized with a ca155-cap1-c2-neg306 radioactive probe. Inserts from positive colonies were sequenced.

Cloning of vertebrate homologs of BP-1

PolyA+ RNA from human bone marrow was purchased from Clontech Inc. (Palo Alto, Ca.). RT-PCR was performed on 50 ng of polyA+ RNA using forward primer 18-131G (SEQ ID NO. 40) and reverse primer 19-9G (SEQ ID NO. 41) and Titan™ one tube RT-PCR system according to the manufacturer's instructions (Roche Molecular Biochemicals). Cycling conditions (40 cycles) were as follows: 94°C for 1 minute, 54°C for 1 minute, 68°C for 1 minute. PCR products (predominant band at approximately 400 bp) were cloned in pBluescript II KS (Stratagene, La Jolla, Ca.) and inserts from randomly chosen colonies were sequenced.

Total RNA was extracted from UMR-106 cells (CRL-1661, American Type Culture Collection, Manassas, Va.). This cell line is derived from a rat osteosarcoma and possesses characteristics of the osteoblast lineage (Partridge et al., 1983). The first strand of cDNAs was reverse transcribed at 42°C from 5 µg of total RNA using 33-1V and Superscript II™ (Invitrogen). After RNase H treatment (Roche Molecular Biochemicals) and purification using the minElute kit (Qiagen), 1/4 of the first strand cDNA preparation was subjected to 35 cycles of PCR using the DNA polymerases from the Titan™ RT-PCR kit, forward primer 18-131G and reverse primer 18-30V. Cycling conditions were as follows: 94°C for 1 minute, 48°C for 1 minute, 68°C for 1.5 minute. RT-PCR products were purified using the minElute kit. 1/6 of the initial PCR reaction was subjected to 35 cycles of semi-nested PCR using the DNA polymerases from the Titan™ RT-PCR kit, forward primer 18-131G and reverse primer 21-9G (SEQ ID NO. 42). PCR products (predominant band at approximately 500 bp) were cloned in pBluescript II KS (Stratagene, La Jolla, Ca.) and inserts from randomly chosen colonies were sequenced.

Total RNA was extracted from the vertebrae and surrounding muscles of a 2-month old birman python (~130g) using the Trizol™ reagent (Invitrogen). The first strand of cDNAs was synthesized at 42°C for 1 hour from 4µg of total RNA using 0.5µg of oligo-dT₁₈ and 200U Superscript II™ (Invitrogen) in a total volume of 20µl. After RNase H treatment (2U added, Roche Molecular Biochemicals) and purification using the QIAQuick kit (Qiagen) with 30µl of elution buffer, the first strand cDNA preparation (2µl) was subjected to 33 cycles of PCR using *Taq* DNA polymerase, forward primer 23-4V (SEQ ID NO. 44) and reverse primer 23-6V (SEQ ID NO. 45). Cycling conditions were as follows: 94°C for 30 seconds, 56°C for 30 seconds, 72°C for 40 seconds. RT-PCR products were purified using the minElute kit. PCR products (predominant band at approximately 130bp) were cloned in pBluescript II KS (Stratagene, La Jolla, Ca.) and inserts from randomly chosen colonies were sequenced.

In situ hybridization

In situ hybridization on paraformaldehyde-fixed paraffin-embedded tissue sections was performed according to standard procedures (for example, see Wilkinson and Nieto, 1993). Briefly, [α^{35} S]UTP-labeled cRNA probe were synthesized in vitro from 0.5 µg of purified linearized template DNA using DNA-dependent RNA polymerase (SP6, T7 or T3 RNA polymerase). After phenol/chloroform extraction and ethanol precipitation, quality of the probe was assessed by denaturing 4% polyacrylamide gel electrophoresis. Tissue sections were deparaffinized, rehydrated, treated with 5 µg/ml of proteinase K (Roche Molecular Biochemicals) for 15 minutes in 50mM Tris-HCl pH 8/5mM EDTA, refixed in 4% paraformaldehyde, acetylated with 0.25% (v/v) acetic anhydride in 0.1M triethanolamine pH 8, dehydrated and air-dried. Prehybridization was carried out for 2-4 hours at room temperature in 4x SET, 1x Denhardt's solution, 0.5mg/ml ssDNA, 0.6mg/ml yeast RNA, and 50% deionized formamide. 1x SET is 0.15M NaCl, 0.03M Tris-HCl pH 8 and 2mM EDTA. 1x Denhardt's solution is 0.02% (w/v) Ficoll, 0.02% (w/v) polyvinylpyrrolidone, 0.1% (w/v) BSA fraction V. ssDNA is salmon sperm DNA that has been sonicated to obtain fragments averaging 2kb, purified by phenol/chloroform extraction/ethanol precipitation/dialysis and

denatured by boiling. Hybridization was carried out for 12-16 hours at 54°C with 40,000 cpm/ μ l of cRNA probe in 4x SET, 1x Denhardt's solution, 0.1mg/ml ssDNA, 0.1mg/ml yeast RNA, 10% dextran sulfate, 10mM DTT, 0.1% SDS, and 50% deionized formamide. High stringency washes were performed as follows: formamide wash for 30 minutes at 55°C and 30 minutes at 62°C; 20 μ g/ml RNase A treatment in 3.5x SSC, 30 minutes, 37°C; formamide wash for 30 minutes at 62°C. Formamide wash is 50% deionized formamide/0.15M NaCl/0.15M sodium citrate/10mM DTT. After rinses, slides were dehydrated, dipped in Kodak NTB-2 liquid emulsion (Intersciences, Markham, Canada), developed after 3-4 weeks of exposition and counterstained by hematoxylin-eosin.

Cloning of mouse BP-1 into a mammalian expression vector

To optimize expression of BP-1 in heterologous systems, a vector expressing the coding sequence of mouse BP-1 was obtained as follows. A 411bp fragment was amplified from 100ng of plasmid GD25 using forward primer 25-10G (SEQ ID NO: 38) and reverse primer 24-10G (SEQ ID NO: 43). The PCR reaction contained 25 pmoles/ μ l of each primer, deoxynucleotides at 200 μ M, dithiotreitol at 5mM, the buffer and the enzyme mix from the *Titan*TM one-step RT-PCR kit (Roche Molecular Biochemicals). Cycling conditions were as follows: 94°C for 30 seconds, 56°C for 30 seconds, 68°C for 30 seconds. The PCR product was cloned in the unique PmeI site of pCMVneoXN to generate GD23. The identity of the insert was verified by DNA sequencing.

Production and purification of antisera

Peptides were synthesized, purified and coupled to activated keyhole limpet hemocyanin via sulfhydryl groups. The peptide/carrier complex was mixed with complete Freund's adjuvant and injected subcutaneously to rabbits on day 1. On days 21, 35 and 49, the peptide/carrier complex mixed with incomplete Freund's adjuvant was again injected. Blood samples were collected on days 44 and 59 to determine the antiserum titer by ELISA. Animals were exsanguinated on day 63. Immunoglobulins were precipitated from pooled antisera and purified by

peptide affinity chromatography according to standard protocols (Affinity Bioreagents, Golden, Co.).

Cell culture and transfection

HEK293A cells (purchased from Quantum Biotechnologies, Montreal, Canada) are grown in Dulbecco modified essential medium supplemented with 10% (v/v) fetal bovine serum, 100 U/ml penicillin, 100 mg/ml streptomycin (referred hereafter as complete medium). Cells are passaged when reaching 80-95% confluence by incubating with 0.05% (v/v) trypsin/0.5mM EDTA (Wisent Inc., St-Bruno, Canada). Lipofection is performed as follows. Lipid-DNA complexes (containing typically 1 μ g of DNA) are formed using the Effectene™ reagent (Qiagen) according to the manufacturer's instructions. Cells are transfected the day after plating (typically 18,700 cells/cm²) by adding the lipid-DNA complex to the culture medium. After a 6 hour incubation, the medium is changed and cells are usually processed after 48 hours.

Immunofluorescence

All incubations and washes are done at room temperature. Cells are rinsed with PBS and fixed with 2% (w/v) paraformaldehyde in PBS. Cells are washed with PBS. For detection of intracellular protein, cells are permeabilized by incubation in 0.1% TRITON X-100™ in PBS for 4 minutes and washed twice with PBS. Cells are then incubated in 50mM NH₄Cl in PBS for 10 minutes at room temperature and washed with PBS. Cells are incubated in PBS supplemented with 0.1% (w/v) bovine serum albumin fraction V and 2% (w/v) dried milk for 1 hour. Cells are then incubated for 1 hour with a 1/100 dilution (v/v) of antiserum raised against BP-1 in PBS supplemented with 0.1% (w/v) bovine serum albumin and 0.5% (w/v) dried milk. Cells are washed twice with PBS and incubated for 1 hour in 1/200 dilution of goat anti-mouse coupled to fluorescein isothiocyanate (Sigma) in PBS supplemented with 0.1% (w/v) bovine serum albumin. Cells are washed twice and observed by fluorescence microscopy.

Western analysis

Cell extracts. Cells are rinsed with PBS. Membrane proteins are solubilized in 0.1ml of Lysis buffer (50 mM Tris-Cl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% IGEPAL-630™ and 1% (v/v) protease inhibitor cocktail (Sigma)). Cell debris and insoluble material are pelleted by centrifugation at 12,000 g for 5 minutes at 4°C. Protein concentration in the supernatant is determined using the Bradford assay according to the manufacturer's instructions (Bio-Rad, Hercules, Ca.).

Medium. 24 hours post-transfection, medium is replaced by medium containing 0.5% (v/v) fetal bovine serum. The following day, this medium is collected and centrifuged at 12,000g for 5 minutes at 4°C to remove cell debris. Proteins in the supernatant are precipitated by adding trichloroacetic acid at a final concentration of 10% (v/v) and incubating on ice for 1 hour. After centrifugation at 12,000g for 15 minutes at 4°C, the pellet is washed once with cold acetone, dried briefly and resuspended in 24µl of Lysis buffer.

Proteins are boiled for 5 minutes in the following Laemmli 1x solution: 50mM Tris-HCl, pH 6.8, 100mM dithiothreitol, 2% sodium dodecyl sulfate, 0.1% bromophenol blue, 10% glycerol. Proteins are electrophoresed on denaturing polyacrylamide gel and transferred to 0.2µm nitrocellulose (Protran®, Schleicher&Schnell, Keene, NH) according to standard Western protocols. The nitrocellulose membrane is incubated overnight in Tris-buffered saline (TBS; 25mM Tris-HCl, pH 7.4, 137mM NaCl, 2.7mM KCl) supplemented with 5% (w/v) dried milk and 0.1% (v/v) TWEEN-20™. It is then incubated for 1 hour at room temperature with a 1/800 dilution (v/v) of antiserum raised against BP-1 in TBS supplemented with 2.5% (w/v) dried milk and 0.1% (v/v) TWEEN-20™. The membrane is washed twice with TBS supplemented with 0.1% (v/v) TWEEN-20™. It is then incubated for 1 hour at room temperature with goat anti-mouse coupled to horseradish peroxidase (Sigma) diluted 1/30,000 in TBS supplemented with 2.5% (w/v) dried milk and 0.1% (v/v) TWEEN-20™. The membrane is washed twice with TBS supplemented with 0.1% (v/v) TWEEN-20™. Detection of the protein bound to the antibody complex is performed with the ECL™ reagent according to the manufacturer's instructions (Amersham Pharmacia Biotech, Baie d'Urfé, Canada).

Primary culture of rat calvarial osteoblasts

Calvariae were dissected from embryonic rats from days 19-21 of gestation. Osteoblasts were isolated by 5 successive digestions of the tissues with a mix of collagenase (2mg/ml, Worthington, NJ.), trypsin (0.05% w/v, Roche Molecular Biochemicals, IN.) and EDTA (0.02% w/v) for 20 minutes with agitation at 37°C. Cells from digests 2 to 5 were seeded at 10,000 cells/cm² and cultured in α MEM (Invitrogen, Carlsbad, Ca.) supplemented with 10% (v/v) FBS, 20mM HEPES pH 7.4, 2mM L-glutamine, 5mM β -glycerophosphate, 50 μ g/ml ascorbic acid, 100U/ml penicillin, 100 μ g/ml streptomycin. Medium was changed every 2-3 days.

Incorporation of ⁴⁵Ca into primary cultures of rat calvarial osteoblasts

Cells were incubated in medium containing 1 μ Ci/ml ⁴⁵CaCl₂ to assay calcium uptake into mineralized matrix. Medium was collected 48 hours later from each well and added to 4ml of scintillation cocktail, and counted on a scintillation counter. The cells were washed twice in PBS then scraped into 500 μ l of lysis buffer (20mM Tris, 0.5mM MgCl₂, 0.1% (v/v) Triton X-100TM). Cellular debris was spun out, the lysis buffer removed and stored and the pellet resuspended in 1ml 0.5M EDTA. The pellet in EDTA was placed overnight on a rocking platform then the debris spun out and the EDTA solution removed and stored. The pellet was then resuspended in 1ml 0.5M NaOH and placed overnight on a rocking platform. The remaining cell debris was spun out and the NaOH removed and stored. To ascertain ⁴⁵Ca incorporation, 100 μ l of the lysis buffer and 500 μ l of the EDTA and NaOH fractions were counted in 4ml of scintillation fluid. The proportion of ⁴⁵Ca taken up into the cells in each well was then calculated as a percentage of the total counts in each well.

B) Example 1. Cell lines secreting BP-1 protein products in the culture medium

The expression vector used to obtain stable cell lines producing BP-1 is schematized on Figure 7A. Plasmid RC33 (see provisional patent application

docket number 29313-0002 and international patent application PCT/CA02/00997) was digested with KpnI and EcoRI and plasmid GD23 (see Materials and Methods) was digested with KpnI and MluI. Both digests were treated using the Klenow fragment of DNA polymerase I to blunt the 5' protruding extremities generated by EcoRI and MluI. The 5194 bp fragment of the RC33 digest and the 473 bp fragment of the GD23 digest were purified and ligated. The resulting vector is designated GD46. A similar vector comprising the coding sequence for human BP-1 was obtained and is designated GD45. One of the transcription unit of GD46 comprises the cytomegalovirus immediate early gene enhancer/promoter regions (700) followed by the coding sequence for mouse BP-1 (701) and a bovine growth hormone polyadenylation signal (702). GD46 also comprises a transcription unit to confer resistance to puromycin (703).

To obtain cell lines having integrated one or more copies of GD46 in its genome, 3.5 million of HEK293 cells were electroporated at 600V/cm with a mixture of 3.5µg of Scal-linearized expression vector and 15µg of denatured salmon sperm DNA. Cells were plated and selection (2.5 µg/ml puromycin) was applied 24 hours later. The concentration of puromycin was reduced to 0.2µg/ml on day 4 and colonies were picked on day 11. Clones were grown until confluence in 24-well plates. At this time, cells were incubated in 0.5ml of medium without serum for 48 hours. One-fifth (100µl) was analyzed by ELISA as follows. One volume (100µl) of 2X carbonate buffer (13mM Na₂CO₃/35 mM NaHCO₃) was added to the sample. Samples were transferred in wells of a high-binding capacity polystyrene plate (Corning, USA) and incubated 1 hour at 37°C. The medium was then aspirated and the well was washed once with TBST (Tris-buffered saline pH 7.4 [TBS; 100mM Tris], 0.1% (v/v) Tween-20TM). Non-specific binding sites were blocked for 1 hour at 37°C with 300µl of TBS supplemented with 5% (w/v) of non fat milk. After 5 washes with TBST, incubation proceeded for 1 hour at 37°C with 200µl of a 1/2500 dilution of the affinity-purified antibody against the C-terminal portion of BP-1 in TBST supplemented with 2.5% (w/v) of non fat milk. The plate was then washed 5 times with TBST. Samples were incubated for 1 hour at 37°C with 200µl of a 1/30,000 dilution of anti-rabbit IgG coupled to horseradish peroxidase (Sigma, St.Louis, Mo.). After 5 washes with TBST, the antibody

complex was revealed with Sigma Fast™ o-phenylenediamine dihydrochloride according to the manufacturer's recommendations. Reactions were stopped by adding 25µl of 5N sulfuric acid per well. Absorbance at 490nm was determined using a plate reader. Known quantities of antigenic peptide are also assayed to determine a standard curve (Figure 7B). The clones of GD46 transfectants secreting immunoreactive BP-1 products in the culture medium were expanded. The amount of BP-1 products released by each clone was calculated by regression analysis after assaying varying volumes of culture medium. The highest expressor (clone 293-GD46-7) was found to secrete 8pg of BP-1 products per cell in the culture medium in 48 hours. This corresponds to a production of approximately 10mg/l for a confluent monolayer in a 175cm² flask. We also obtained a cell line overexpressing the human BP-1 coding sequence (293-GD45-58B). Confluent 175cm² monolayers of these cells secrete approximately 0.6mg of BP-1 products per liter of culture medium in 48 hours.

To ascertain that the cell lines we derived were of monoclonal origin and would not be diluted by non BP-1-expressing cells upon passaging, immunofluorescence analysis was performed on confluent monolayers of 293A-GD46-7 using an antibody directed against the C terminal portion of BP-1. BP-1 immunoreactivity was seen in the secretory apparatus in over 98% of the cells.

C) Example 2. Method to partially purify the BP-1 protein products from cell culture medium

Six million cells stably overexpressing mouse BP-1 (clone 293-GD46-7) were seeded in a 175cm² flask and grown for 48 hours in DMEM supplemented with 10% FBS, at which stage the monolayer was approximately 90% confluent. The cells were then washed 2 times with pre-warmed phosphate buffered saline and once with serum-free DMEM. The cells were incubated for 48 hours in 20ml of serum-free DMEM. The medium was then collected and spun down at 800g for 2 minutes to remove floating cells and debris. All subsequent procedures were performed at 4°C using a BioRad BioLogic™ LP Chromatography system. The conditioned medium containing BP-1 products was first diluted 2-fold with

equilibration buffer (Buffer A: 25 mM HEPES pH 7.8 and 100mM NaCl) and filtered through a 0.45µm FiltropurTM S membrane. The diluted medium was then loaded at a flow rate of 2 ml/min on a column (1.5cm x 20cm) packed with 3.5ml gel bed of a SepharoseTM SP cation exchange resin (Amersham-Pharmacia Biosciences). The flow-through was collected until the absorbance at 280 nm returned to baseline. The resin was washed with 5 bed volume of Buffer A at 2 ml/min. The bound-proteins were eluted at 1 ml/min with 50 ml of a linear gradient from 0-100% of Buffer B (25 mM HEPES pH 7.8 and 1 M NaCl). Fractions of 1 ml were collected and monitored for the presence of BP-1 by ELISA (see Example 1). The fractions containing BP-1 were pooled and concentrated 10-fold on an Amicon CentriplusTM YM-3 (3 kDa cut-off membrane) by centrifugation at 3400g at 4°C. No significant loss of BP-1 immunoreactivity occurred during this concentration procedure.

Figure 8 shows the results of the initial steps of the purification scheme described above. The fractions collected from the cation-exchange column were individually assayed for BP-1 immunoreactivity by ELISA (Figure 8A). The input (conditioned medium 0.9% of total), the flow through (0.9% of total) and BP-1-containing fractions (0.8% of total) were analyzed by SDS-PAGE 10-20% (Figure 8B, lanes 810, 811, 812, respectively). After electrophoresis, the gel was stained by PlusOneTM silver staining kit (Amersham Pharmacia Biosciences). Lane 800 shows the migration of molecular weight markers. Arrow 813 points to a BP-1 product. Arrowhead 814 points to aprotinin, a contaminating protease inhibitor included in the conditioned medium before the chromatographic step. A duplicate gel was transferred onto nitrocellulose membrane and BP-1 products were revealed by immunostaining (Figure 8C, lane 820:input, lane 821:flow through, lane 822: BP-1-containing fractions). Results indicate that the bulk of the proteins in the culture medium do not bind the cation exchange column (compare lanes 810 and 811). By Coomassie and silver staining, we estimate that these initial steps of purification allowed a 25-fold enrichment in BP-1 products. As can be seen on the SepharoseTM SP elution profile (Figure 8A), BP-1 products eluted at a salt concentration of approximately 300mM NaCl. Importantly, the immunoreactive fractions contain products of lower apparent molecular weight (lane 822, band at ~6kD) in addition to the ~13kD form of BP-1. This indicates that the partial

purification scheme described above can enrich for all BP-1 products secreted by an overexpressing cell line.

D) Example 3. Recombinant adenovirus particles expressing the BP-1 coding sequence and uses thereof

This example illustrates the various functionalities of an adenovirus-based expression vector designed according to the present invention. To obtain a 'shuttle' vector for inserting a transcription unit into an adenoviral genome, a 1689 bp *Stu*I-*Nae*I fragment from GD23 (see Materials and Methods) was blunted and cloned in the *Eco*RV site of pQBI-AdBN (qBiogene, Montreal, Canada) which had been previously engineered to replace the unique *Cla*I site with a unique *Pme*I site. The resulting vector is designated GD28b and comprises a transcription unit for mouse BP-1 coding sequence (901) flanked by nucleotides 1-102 (902) and nucleotides 3334-5779 (903) of the Adenovirus serotype 5 genome (GenBankTM accession number 9626187). A map of GD28b is given in Figure 9A.

The transcription unit for BP-1 comprised in GD28b was incorporated in an adenoviral genome by *in vivo* homologous recombination. This was done by co-transfecting 5 µg of *Pme*I-linearized GD28b with 5 µg of AdCMVlacZΔE1/ΔE3, a replication-defective genome obtained commercially (qBiogene, Montreal, Canada). Co-transfection of DNA molecules in HEK293 cells was carried out by means of a calcium phosphate precipitate using standard protocols. Two days post-transfection, cells were overlaid with medium containing 1.25% (w/v) low melting agarose. Recombinant viral genome resulting from homologous recombination between GD28b and the replication-defective adenoviral genome can be propagated in HEK293 cells, as indicated by the appearance of viral plaques starting at day 8 post-infection. Plaques were picked at day 14 post-transfection. After elution, viral particles (1/10 of the elutate) were used to infect 3×10^5 HEK293 cells to assay for BP-1 expression. This was done by measuring BP-1 immunoreactivity by ELISA in 1/10 of the culture medium 2 days post-infection. Expression of BP-1 was also confirmed by Northern and Western analysis. A stock of recombinant adenoviral particles (Ad5GD28b) was obtained

after 2 successive rounds of plaque-purification according to standard protocols. This stock was amplified on HEK293 cells. At the time of full cytopathic effect, the infected cells were harvested and the viral particles were purified on a discontinuous cesium chloride gradient and on a continuous cesium chloride gradient according to standard protocols (O.D.260 Inc., Boise, Idaho, USA).

The following experiment was performed to assess the efficiency of BP-1 gene transfer in osteoblasts by recombinant adenoviruses. Primary cultures of osteoblasts were grown in vitro (see Materials and Methods). Cells reached confluence 5 days after seeding, at which time they started depositing an extracellular collagen matrix. Mineralization of this matrix was routinely visible around days 13-15. At day 11, cells were infected with CsCl-purified Ad5GD28b at a multiplicity of infection of 100 pfu/cell for 3 hours. Relative levels of BP-1 products were determined by ELISA on aliquots (1/40) of culture medium taken 2, 4, 7 and 10 days post-infection (Figure 9B). Results show that BP-1 immunoreactivity is detected as early as 2 days post-infection, peaks at day 4 and is still detectable 10 days post-infection. These results indicate that a preparation of Ad5GD28b can be used to express high levels of BP-1 in osteoblasts.

E) Example 4. Treatment of osteoblasts with medium containing BP-1 products

This example demonstrates the effects of treating primary cultures of osteoblasts with medium containing BP-1 products. The BP-1-containing medium was obtained by transient transfection of 1.5 million HEK293A cells with 2 μ g of GD23, an expression vector for mouse BP-1 cDNA. Control medium was obtained by similarly transfecting pCMVneo, a control expression vector. At 24 hours post-transfection, cells were washed twice with serum-free α MED. The transfected cells were then incubated for 48 hours in primary culture medium. BP-1 products were characterized by Western analysis using an antiserum raised against a C-terminal portion of BP-1 (SEQ AG3). The concentration of BP-1 products was estimated at 5-10 μ g/ml with the bulk of immunoreactive protein having an apparent molecular weight of ~13 kD and a minor fraction having an apparent molecular weight of ~6 kD after SDS-PAGE (e.g. Figure 6B).

Primary cultures of osteoblasts prepared as described in Materials and Methods were cultivated in a 1/6 (v/v) dilution of either BP-1-containing or control medium. β -glycerophosphate (10mM) was added to the culture medium from day 12. The effect of the treatment was determined by measuring osteocalcin expression and matrix mineralization, both markers of the mature osteoblastic phenotype. As shown in Figure 10, treatment with BP-1-containing medium (1002) for 18 days completely suppressed osteocalcin (OCN) expression compared to treatment of cells with control medium (1001). Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) expression in the two cell populations was similar, demonstrating no general effects on the cell populations. Incorporation of ^{45}Ca was also markedly reduced in the cells treated with BP-1-containing medium compared to cells treated with control medium (-60%, $p < 0.01$). These results suggest that BP-1 products restrict the progression from early to late stage osteoblasts.

REFERENCES

Throughout this paper, reference is made to a number of articles of scientific literature which are listed below.

Aarder, E.M., Burger, E.H., Nijweide, P.J. (1994) Function of osteocytes in bone, *J. Cell Biochem.*, 55, 287-299

Bagi, C.M., DeLeon, E., Brommage, R., Adams, S., Rosen, D., Sommer, A. (1995) Systemic administration of rhIGF-I or rhIGF-I/IGFBP-3 increases cortical bone and lean body mass in ovariectomized rats, *Bone*, 16(Suppl.), 263S-269S

Baron, R. (1999) Anatomy and ultrastructure of bone, in *Primer on the Metabolic Bone Diseases and Disorders of Mineral Metabolism*, ed. Lippincott Williams & Wilkins, Philadelphia, pp.3-10

Birnbaum, R.S., Howard, G.A., Roos, B.A. (1989) Ontogeny of peptidylglycine α -amidating monooxygenase activity in rapidly mineralizing bone from neonatal mouse, *Endocrinology*, 124, 3134-3136

Chomczynski P, Sacchi N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction, *Anal Biochem.*, 162, 156-9.

Church GM, Gilbert W. (1984) Genomic Sequencing, *Proc Natl Acad Sci U S A*, 81, 1991-5

Denault, J.-B., Leduc, R. (1996) Furin/PACE/SPC1: a convertase involved in exocytic and endocytic processing of precursor proteins, *FEBS Lett.*, 379, 113-116

Ducy, P. (2000) Cbfa1: a molecular switch in osteoblast biology, *Dev. Dyn.* 219, 461-471

Ducy, P., Amling, M., Takeda, S., Priemel, M., Schilling, A.F., Biel, F.T., Shen, J., Vinson, C., Rueger, J.M., Karsenty, G. (2000) Leptin inhibits bone formation through a hypothalamic relay: a central control bone mass, *Cell*, 100, 197-207

Frota Ruchon, A., Tenenhouse, T.S., Marcinkiewicz, M., Siegfried, G., Aubin, J.E., Desgroseillers, L., Crine, P., Boileau, G. (2000) Developmental expression and tissue distribution of pthex protein: effect of the *hyp* mutation and relationship to bone markers, *J. Bone Miner. Res.*, 15, 1440-1450.

Henikoff S., Henikoff J.G.; RT "Automated assembly of protein blocks for database searching."; *RL Nucleic Acids Res.* 19:6565-6572(1991).

Hopp, T.P., Woods, K.R. (1981) Prediction of protein antigenic determinants from amino acid sequences, *Proc Natl Acad Sci USA*, 86, 152-156

The Hyp consortium (1995) A gene (PEX) with homologies to endopeptidases is mutated in patients with X-linked hypophosphatemic rickets, *Nat. Genet.* 11, 130-6

Inagami T, Misono KS, Grammer RT, Fukumi H, Maki M, Tanaka I, McKenzie JC, Takayanagi R, Pandey KN, Parmentier M. (1985) Biochemical studies of rat atrial natriuretic factor, *Clin Exp Hypertens A*, 7, 851-867

Karlin S, Altschul SF, *Proc Natl Acad Sci U S A* 1990 March, 87:2264-8;

Karlin S, Altschul SF, *Proc Natl Acad Sci U S A* 1993 Jun 15;90(12):5873-7;
Altschul, SF (1993), *J. Mol. Evol.* 36:290-300

Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, *Cell*, 44, 283-292

Kumar, R. (2000) Tumor-induced osteomalacia and the regulation of phosphate homeostasis, *Bone*, 27, :333-8

Lingappa V.R., Cunningham B.A., Jazwinski S.M., Hopp T.P., Blobel G., Edelman G.M. (1979) Cell-free synthesis and segregation of beta 2-microglobulin, *Proc Natl Acad Sci U S A*, 76, 3651-3655

Louis, N., Eveleigh, C., Graham, F.L. (1997) Cloning and sequencing of the cellular-viral junctions from the human adenovirus type 5 transformed 293 cell line, *Virology*, 233, 423-429

Mayahara H, Ito T, Nagai H, Miyajima H, Tsukuda R, Taketomi S, Mizoguchi J, kato K (1993) In vivo stimulation of endosteal bone formation by basic fibroblast growth factor in rats, *Growth Factors*, 9, 73-80

Manolagas, S.C. (2000) Birth and death of bone cells: basic regulatory mechanisms and implications for the pathogenesis and treatment of osteoporosis, *Endocr. Rev.*, 21, 115-137

Maruyama K, Sugano S (1994) Oligo-capping: a simple method to replace cap structure of eukaryotic mRNAs with oligoribonucleotides, *Gene*, 138, 171-174

Myers., W. Technical Report 29, Department of Computer Science, University of Arizona, Tucson, 1991.

Needleman, S. B. & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48, 443-453.

Neer R.M., Arnaud C.D., Zanchetta J.R., Prince R., Gaich G.A., Reginster J.Y., Hodsman A.B., Eriksen E.F., Ish-Shalom S., Genant H.K., Wang O., Mitlak B.H. (2001) Effect of parathyroid hormone (1-34) on fractures and bone mineral density in postmenopausal women with osteoporosis, *N Engl J Med*, 344, 1434-41

Olins, G.M., Patton, D.R., Bovy, P.R., Mehta, P.P. (1988) A linear analog of atrial natriuretic peptide (ANP) discriminates guanylate cyclase-coupled ANP receptors from non-coupled receptors, *J. Biol. Chem.*, 263, 10989-10993

Partridge, N.C., Alcorn, D., Michelangeli, V.P., Ryan, G., Martin, T.J. (1983) Morphological and biochemical characterization of four clonal osteogenic sarcoma cell lines of rat origin, *Cancer Res.*, 43, 4308-4314

Pearson, W. R. & D. J. Lipman., Improved Tools for Biological Sequence Analysis., 1988, *Proc. Natl. Acad. Sci.*, 85, 2444-2448

Ragot, T., *et al.* (1998), *Meth. Cell Biol.*, 52, 229-260.

Ruchon A.F., Marcinkiewicz M., Siegfried G., Tenenhouse H.S., DesGroseillers L., Crine P., Boileau G. (1998) Pex mRNA is localized in developing mouse osteoblasts and odontoblasts, *J Histochem Cytochem*, 46, 459-68

Russell, R.G.G, Rogers, M.J. (1999) Bisphosphonates: from the laboratory to the clinic and back again, *Bone*, 25, 97-106

Smith, T. F. & Waterman, M. S. (1981). Identification of common molecular subsequences. *J. Mol. Biol.* 147, 195-197.

Tsomides, T.J., Eisen, H.N. (1993) Stoichiometric labeling of peptides by iodination on tyrosyl or histidyl residues, *Anal. Biochem.*, 210, 129-135

Unger, R.H. (2000) Leptin physiology: a second look, *Reg. Pept.*, 92, 87-95

Wilkinson DG, Nieto MA. (1993) Detection of messenger RNA by in situ hybridization to tissue sections and whole-mounts, *Methods Enzymol*, 225, 361-73

Yan, W., Sheng, N., Seto, M., Morser, J., Wu, Q. (1999) Corin, a mosaic transmembrane serine protease encoded by a novel cDNA from human heart, *J. Biol. Chem.*, 274, 14926-14935

Yan, W., Wu, F., Morser, J., Wu, Q. (2000) Corin, a transmembrane cardiac serine protease, acts as a pro-atrial natriuretic peptide-converting enzyme, *Proc Natl Acad Sci U S A*, 97, 8525-8529

We Claim:

1. An isolated nucleic acid encoding bone polypeptide-1.
2. The nucleic acid of claim 1 wherein said polypeptide is human bone polypeptide-1.
3. The nucleic acid of claim 2 wherein said protein comprises the polypeptide sequence of SEQ ID NO:9.
4. The nucleic acid of claim 1 wherein said protein is selected from the group consisting of rat, cow, snake, mouse and chicken bone polypeptide-1 proteins.
5. The nucleic acid of claim 1 wherein said polypeptide comprises a polypeptide sequence selected from the group consisting of SEQ ID NO:10; SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 and SEQ ID NO:14.
6. The nucleic acid of claim 1 wherein said protein comprises a polypeptide selected from the group consisting of SEQ ID NO: 20 SEQ ID NO:21; SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 and SEQ ID NO:25.
7. The nucleic acid of claim 1 wherein said nucleic acid comprises the nucleotide sequence of SEQ ID NO: 2.
8. The nucleic acid of claim 1 wherein said nucleic acid comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO: 6 and SEQ ID NO:7.
9. An vector comprising the nucleic acid of claim 1.
10. The vector of claim 9 comprising an expression control sequence operably linked to said nucleic acid.
11. The vector of claim 10 wherein said expression control sequence is a promoter.
12. The vector of claim 11 wherein the promoter is a prokaryotic promoter.
13. The vector of claim 11 wherein the promoter is a eukaryotic promoter.

14. An isolated cell comprising the vector of claim 9.
15. The cell of claim 14 wherein the isolated cell is a prokaryotic cell.
16. The cell of claim 14 wherein the isolated cell is a eukaryotic cell.
17. A method of producing bone polypeptide-1, comprising culturing the cell of claim 14 under conditions permitting expression of the polypeptide and purifying the polypeptide from the cell or culture medium of the cell.
18. The method of claim 17 wherein the cell is a prokaryotic cell.
19. The method of claim 17 wherein the cell is a eukaryotic cell.
20. An isolated nucleic acid wherein said nucleic acid hybridizes under high stringency conditions to the full length nucleic acid sequence of a nucleic acid selected from the group consisting of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO: 6 and SEQ ID NO:7 or the complement thereof wherein said high stringency conditions includes hybridizing in 0.15 M NaCl at 72°C for about 15 minutes and washing said hybridized DNA in 0.2X SSC at 65°C for 15 minutes.
21. The nucleic acid of claim 20 wherein said nucleic acid comprises the nucleotide sequence of SEQ ID NO:2.
22. A vector comprising the nucleic acid of claim 20.
23. The vector of claim 22 comprising an expression control sequence operably linked to said nucleic acid.
24. The vector of claim 23 wherein said expression control sequence is a promoter.
25. The vector of claim 24 wherein the promoter is a prokaryotic promoter.
26. The vector of claim 24 wherein the promoter is a eukaryotic promoter.

27. An isolated cell comprising the vector of claim 22.
28. The cell of claim 27 wherein the isolated cell is a prokaryotic cell.
29. The cell of claim 27 wherein the isolated cell is a eukaryotic cell.
30. A method of producing bone polypeptide-1, comprising culturing the cell of claim 27 under conditions permitting expression of bone polypeptide-1 and purifying bone polypeptide-1 from the cell or culture medium of the cell.
31. The method of claim 30 wherein the cell is a prokaryotic cell.
32. The method of claim 30 wherein the cell is a eukaryotic cell.
33. An isolated or recombinant bone polypeptide-1 or an active fragment thereof.
34. The polypeptide of claim 33 wherein said bone polypeptide-1 is human bone polypeptide-1.
35. The polypeptide of claim 33 wherein said bone polypeptide-1 is rat, bovine, snake, mouse or chicken bone polypeptide-1.
36. The polypeptide of claim 34 wherein said polypeptide comprises the amino acid sequence of SEQ ID NO: 9.
37. The polypeptide of claim 33 wherein said fragment comprises one or more peptides selected from the group consisting of SEQ ID NO:20, SEQ ID NO: 21, SEQ ID NO:22, SEQ ID NO:22; SEQ ID NO:23, SEQ ID NO:24 and SEQ ID NO:25.
38. The polypeptide of claim 33 wherein said fragment comprises one or more peptides selected from the group consisting of SEQ ID NO:20, SEQ ID NO: 21, and SEQ ID NO:23.
39. A pharmaceutical composition comprising the bone polypeptide-1 or active fragment thereof of claim 33.
40. An antibody that binds to an epitope in the polypeptide encoded by the nucleic acid of SEQ ID NO:2.
41. The antibody of claim 40 wherein the epitope comprises a peptide selected from SEQ ID NO:26, SEQ ID NO: 27, or SEQ ID NO:28.
42. An expression vector comprising a nucleic acid sequence encoding bone polypeptide-1, the vector being suitable for genetic therapy.

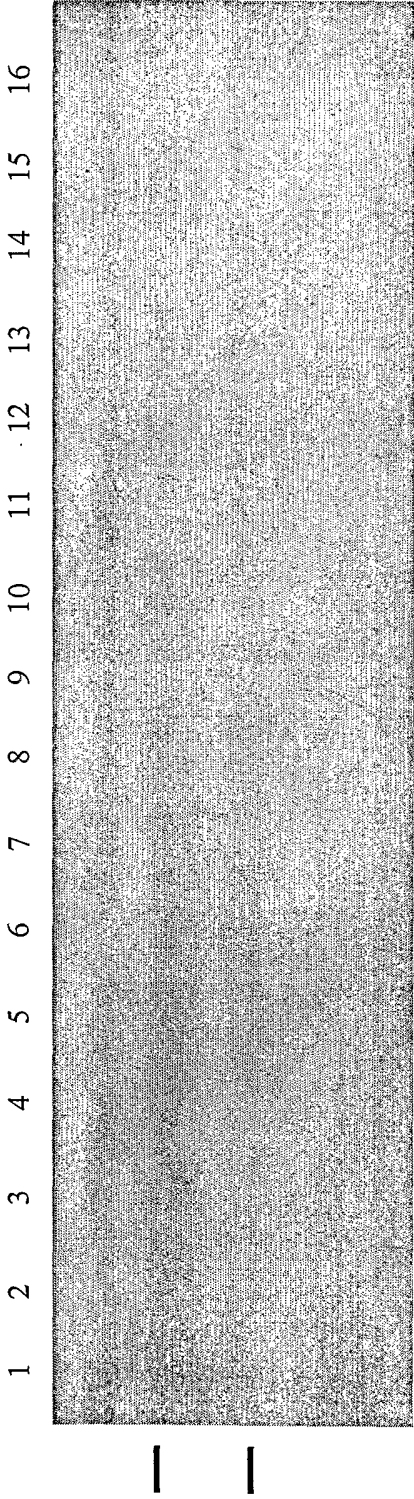
43. A method to deliver a nucleic acid, the sequence of which encodes a nucleic acid encoding bone polypeptide-1 or a fragment thereof, into a host for therapy.

44. The method of claim 43 wherein delivery is by adenovirus.

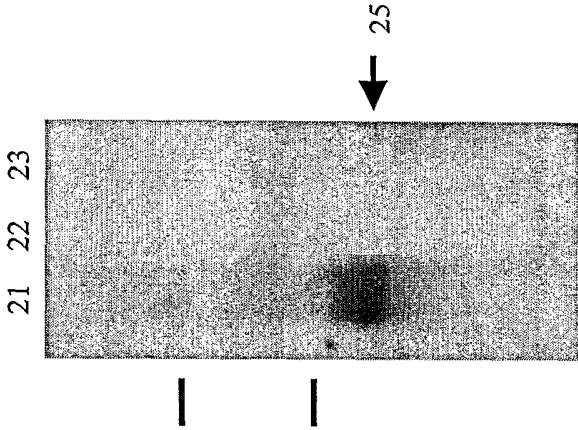
45. A method of treating a bone disorder or osteoporosis of a patient comprising administering an effective amount of bone polypeptide-1 or a fragment thereof to said patient.

Figure 1

A



B



2/15

Figure 2

1 gctaagtttg ggataagctg cagggcgggac tccaaagtta ggagctctga ctctcacaaa
 61 gATGCTGGAC TGGAGATTGG CAAGTACACA CTTCAATCCTG GCTATGATTG TGATGCTGTG
 121 GGGCTCAGGA AAGGCATTCT CTGTGGACTT AGCATCACAG GAGTTTGGAA CAGCAAGCTT
 181 GCAGTCTCCA CCCACAGCCA GAGAAAGAGAA GTCAGCCACT GAGCTTTTCGG CTAAGCTCCT
 241 GGTCTTGAT GATCTGGTGT CCTTAGAGAA TGACGTATTT GAGACCAAGA AAAAGAGAAAG
 301 CTTCTCTGGC TTTGGGTCTC CCCTTGACAG ACTCTCAGCT GGGTCTGTAG AGCATAGAGG
 361 GAAACAAAGG AAAGCAGTAG ATCATTCAAA AAAGCGGTTT GGTATTCCCA TGGATCGGAT
 421 TGGTAGAAAC CGGCTCTCCA GTTCCAGAGG CTGAtggatt cttattgtgc gacttacttg
 481 tgtgagatgg cacagaacta tagaagacac ttcagtgaag ttcactaccc cttttgtcaa
 541 ggaattggcc tttcgcaaac cttcccaaaag cttgatcctc cccagaccat caggtcatag
 601 tgttgctgtg gttttagttg agttgtgcag atcatttcag tgcattgata tctctgaaaag
 661 tatttttcaa tgattcccaa attgtaacgt gggccctgaa cctacttttt taaacagcag
 721 accaatataa tgcattctct tggcattaat attttcacat ttcagttaat caatgtgctt
 781 tctagaaacc tagtgtctga agatctgatg atctaaagaa atcagaaaatg agcacatggt
 841 gatttatata ggtttcttta gttttcttga ggtttgtcga attgttgtaa acttcaaact
 901 caagcttaga aaaaagacat tacatgagtg tttgcttcaa ctgtgtcaga gggcaataaa
 961 attttgagaa acctgagcaa ttgtgttctt taggaactaa taaaggatag tataattggc
 1021 ccataatgtaa tattctgaca aactctgaat gtaaaagact catttgaaaa gaagtactg
 1081 cctggcttgt ttacttctac cagcctaggg gtgaattgtt caaatgtttc ctatgttagc
 1141 agctttttctt cttctttttt tttcttctat tttacttttt ttcttcattc aatgtttata
 1201 agctaaaaat ccaaccaa atgtgctttgt gctttaaaag ggggtattaa aatcaacatt
 1261 aatctaaaaa aaaaaaaaaa

Figure 3

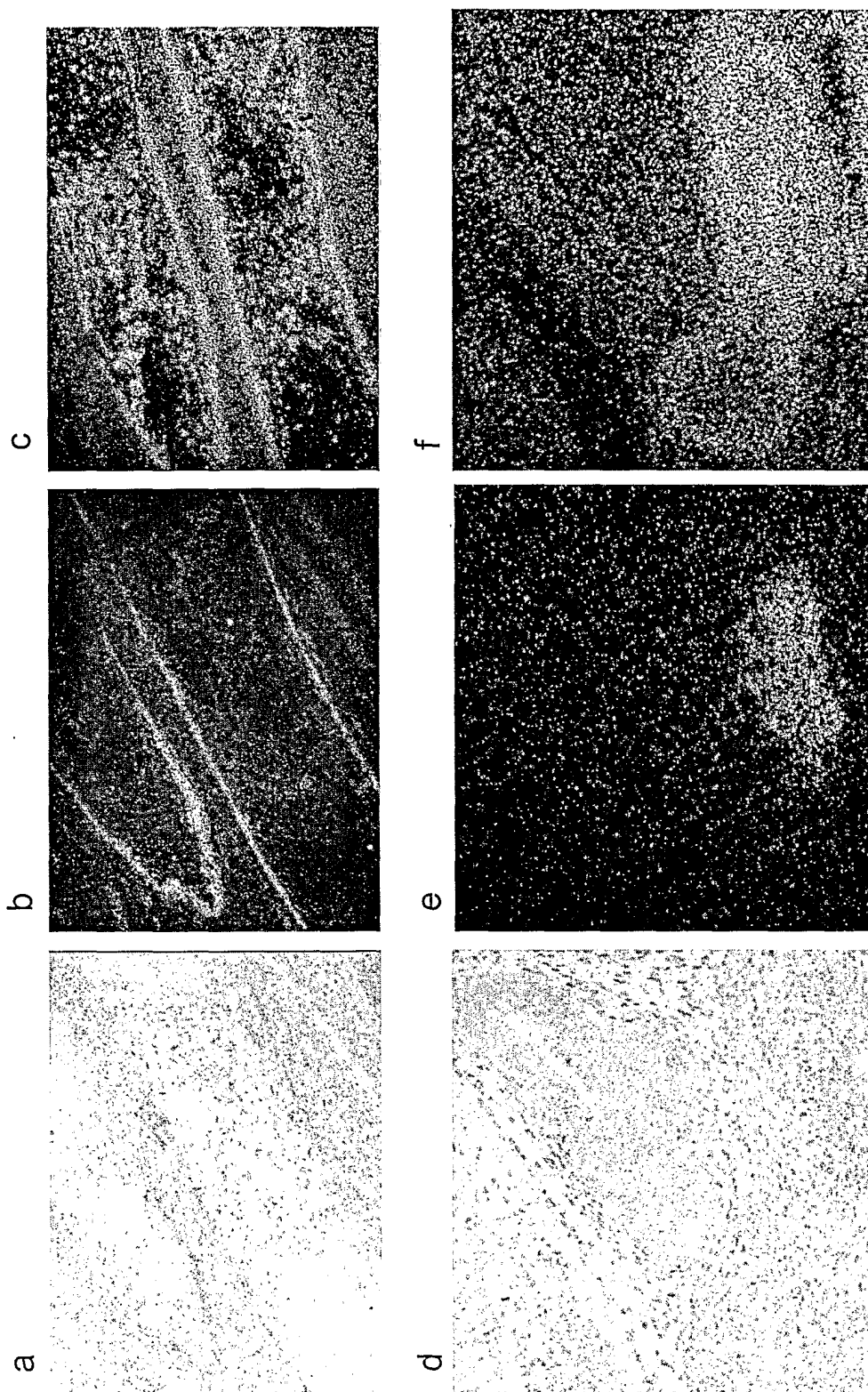


Figure 4

	1	60
BP-1 ORF Human	(1)	ATGCTGGACTGGAGATTGGCAAGTGCACATTTCATCCTGGCTGTGACACTGACACTGTGG
BP-1 ORF Bovine	(1)	ATGCTGGACTGGAGATTAGCAAGTGCACATTTTATCCTGGCTATGACACTGATGCTCTGG
BP-1 ORF Mouse	(1)	ATGCTGGACTGGAGATTGGCAAGTACACACTTCATCCTGGCTATGATTGTGATGCTGTGG
BP-1 ORF Rat	(1)	ATGCTGGACTGGAGATTGGCAAGTGCACACTTCCTCCTGGCTATGATCCTGATGCTGTGG
BP-1 ORF Chicken	(1)	ATGCTGCAGTTCACGCTTGTGTGTGTCATCTGGCCCTTGTGATCACCCCTGCTGCAGTGG
Consensus	(1)	ATGCTGGACTGGAGATTGGCAAGTGCACATTTCATCCTGGCTATGAC CTGATGCTGTGG
	61	120
BP-1 ORF Human	(61)	AGCTCAGGAAAAGTCCCTCTCAGTAGATGTAAACAACAGAGGCCCTTTGAT---TCTGGA
BP-1 ORF Bovine	(61)	AGCTCAGGAAAAGTGTTCCTCAGTGGGTGTCAACA---GAGGCCCTTTGAT---TCTGGA
BP-1 ORF Mouse	(61)	GGCTCAGGAAAAGGCATTCTCTGTGGACTTAGCATCA---CAGG-----AG---TTTGGA
BP-1 ORF Rat	(61)	GGCTCAGGAAAAGGCATTCTCCGTGGACTTAGCATCA---GAGGCCCTCCGAG---TTTGGA
BP-1 ORF Chicken	(61)	CATTCTAGTTCAGTGTCTCCTTGCAGAGGCAGCTCCA---GAGCCCTTTGGAGCCTTCTGCT
Consensus	(61)	GCTCAGGAAAAGT TTCTC GTGGA GTAGCA CA GAGGCCCTT GAG TCTGGA

Figure 4 (continued)

	121		180
BP-1 ORF Human	(118)	GTCA TAGATGTGCAGTCAACACCCACAGTCAGGGAAGAGAAATCAGCCACTGACCTGACA	
BP-1 ORF Bovine	(115)	GTCTTAGGTGTTCAAGTCATCACCCACAGTCAGAGAACGAAAGTCGGCCACTGACCTGGCA	
BP-1 ORF Mouse	(109)	ACAGCAAGCTTGCAGTCTCCACCCACAGCCAGAGAGAGAAAGTCAGCCACTGAGCTTTTCG	
BP-1 ORF Rat	(115)	GCAGAAAGCTTGCAGTCCCCACCCACAAACAGAGAAAGAGAAAGTCAGCCACGGAGCTTGCA	
BP-1 ORF Chicken	(118)	GCTCTGGGCATGGCAGCACATCCTACTGCCAGCGAGGAGAAAGTCAGCCTCCAGCCTGGCA	
BP-1 ORF Python	(1)	TACGGCGTCGGAGGAGAAAGTCGGCTACTGACCTGGTG	
Consensus	(121) GC	TAGGC TGCAGTCACCCACCCACAGCCAGAGAAAGTCAGCCACTGACCTGGCA	
	181		240
BP-1 ORF Human	(178)	GCAAAACTCTTGCTTCTTGATGAATTGGTGTCCCTAGAAAATGATGTGATTGAGACAAAAG	
BP-1 ORF Bovine	(175)	GCAAAACTCTTACTTCTTGATGAACCTTGTGTCTCTGGAGAAATGACGTGATTGAAAACAAAAG	
BP-1 ORF Mouse	(169)	GCTAAGCTCCTGCGTCTTGATGATCTGGTGTCCCTTAGAGAAATGACGTATTTGAGACCAAAG	
BP-1 ORF Rat	(175)	GCTAAGCTCCTGCTTCTTGATGATCTGTGGTGTCCCTTGGAGAAATGATGTGTTTGGAGACCAAAG	
BP-1 ORF Chicken	(178)	GCCAAACTGCTCCTTCTTGATGAGTTGGTGTCTCTGGAGAAATGAGGTAACCTGAGACAAAAG	
BP-1 ORF Python	(38)	GCCAAAATTTTGCTCCTCAACGAATTGGTGTCCCTTGAAAACGATGTCTTTGAGACCAAAG	
Consensus	(181) GC	AAACTCTTGCTTCTTGATGAATTGGTGTCCCTTGGAGAAATGATGTGTTTGGAGACCAAAG	

Figure 4 (continued)

	241		300
BP-1 ORF Human	(238)	AAGAAAAGGAGTTTCTCTGGTTTGGGTCTCCCTTGACAGACTCTCAGCTGGCTCTGTA	
BP-1 ORF Bovine	(235)	AAGAAAAGAAAGCTTCTCTGGTTTGGTTTCTCCCTGGACAGACTCTCAGCTGGCTCTGTA	
BP-1 ORF Mouse	(229)	AAAAAGAGAAAGCTTCTCTGGCTTTGGGTCTCCCTTGACAGACTCTCAGCTGGGTCTGTA	
BP-1 ORF Rat	(235)	AAGAAGAGAAAGCTTCTCTGGCTTCGGGTCTCCCTTGACAGACTCTCGGCTGGGTCTGTA	
BP-1 ORF Chicken	(238)	AAGAAAAGAAAGTTTCCAGGATTTGGCTCCCGATCGACAGAATTTCTGCGACATCTGTG	
BP-1 ORF Python	(98)	AAGAAGAGGAGCTTCTCCGGGTTTGGCTCCCACTTGACAGACTTTCGG-TGGGCCCTG--	
Consensus	(241)	AAGAAGAGAAAGCTTCTCTGG TTTGGGTCTCCCTTGACAGACTCTCAGCTGGGTCTGTA	
	301		360
BP-1 ORF Human	(298)	GATCACAAGGTAAACAGAGGAAAGTAGATCATCCAAAAGGCGATTTGGTATCCCC	
BP-1 ORF Bovine	(295)	AGTCATAAAGGTAAACAGAGGAAAGTAGATCATCCAAAAGGCGATTTGGTATCCCCT	
BP-1 ORF Mouse	(289)	GAGCATAGAGGGAAACAAAGGAAAGCAGTAGATCATTCAAAAAGCGGTTTGGTATTCCC	
BP-1 ORF Rat	(295)	GAGCATAGAGGGAAACAAAGGAGAGTAGTTGATCATTCAAAAAGCGATTTGGTATTCCC	
BP-1 ORF Chicken	(298)	GATGCTAAAGGCAACACAGAGGAAAGTGGTTGAGCTGCCTAAGAGACGGTTTGGAGTTCCT	
BP-1 ORF Python	(155)	-----AAAGCCAAGCAGAGGAAAGCTGTGGAGCTGCCAAAGAGCGGTTTGGGATTCCCT	
Consensus	(301)	GATCATAAAGG AAACAGAGGAAAGTAGATCATCCAAAAGGCGGTTTGGTATTCCCT	

Figure 4 (continued)

	361		405
BP-1 ORF Human	(358)	ATGGATCGGATTGGTAGAAACCGGCTTTCAAATTCCAGAGGCTAA	
BP-1 ORF Bovine	(355)	ATGGATCGGATTGGAGAAACCGGCTTTCAAATTCCAGAGGCTAA	
BP-1 ORF Mouse	(349)	ATGGATCGGATTGGTAGAAACCGGCTCTCCAGTTCCAGAGGCTGA	
BP-1 ORF Rat	(355)	ATGGATCGAATTGGTAGAAACCGTCTCTCCAGTTCCAGGGGCTGA	
BP-1 ORF Chicken	(358)	CTTGACCGGATCGGAGTGAGTCGTCTTGGCAACACCAAGGTTAG	
BP-1 ORF Python	(209)	CTAGATCGGATTGGCGTGAAATCGTTTGAGCGGCTCCAGAGGTTAG	
Consensus	(361)	ATGGATCGGATTGGTAGAAACCGTCTTTCAGTTCCAGAGGCTAA	

Figure 5A

Homo sapiens	(1)	<u>MLDWRLASAHFILAVTLTLWSSGKVLSDVTTTEAFDSGVIDVQSTPTVR</u>	
Bos taurus	(1)	<u>MLDWRLASAHFILAMTLMWSSGKVFSGVTT-EAFDSGVLGVQSSPTVR</u>	
Mus musculus	(1)	<u>MLDWRLASTHFILAMIVMLWGSGKAFSVDLAS-QEFGTASLQS--PPTAR</u>	
Rattus norvegicus	(1)	<u>MLDWRLASAHFILLAMILMLWGSGKAFSVDLAS-EASEFGAESLQSPPTTR</u>	
Gallus gallus	(1)	<u>MLQFQLVVVHLALVITLLIQWHSSVLLAEAAPEPLEPSAALGMAAHPAS</u>	
Python molurus bivittatus			
Consensus	(1)	MLDWRLASAHFILAMTLMW-SGKVFSVDLASE--DSG-L-LQS-PT-R	
Homo sapiens	(51)	EEKSATDLTAKLLLLLDELVSLENDVIET	KKKR SFGFGSPLDRLSAGSVD
Bos taurus	(50)	EAKSATDLAAKLLLLLDELVSLENDVIET	KKKR SFGFGSPLDRLSAGSVS
Mus musculus	(48)	EEKSATELSAKLLRLDDLVSLENDVFET	KKKR SFGFGSPLDRLSAGSVE
Rattus norvegicus	(50)	EEKSATELAAKLLLLLDDLVSLENDVFET	KKKR SFGFGSPLDRLSAGSVE
Gallus gallus	(51)	EEKSASSLAAKLLLLLDELVSLENEVTET	KKKR SFPGFGSPIDRISATSVD
Python molurus bivittatus		TDLVAKILLNLNELVSLENDVFET	KKKR SFGFGSPLDRLSVG-L-
Consensus	(51)	EEKSATDLAAKLLLLLDELVSLENDVFET	KKKR SFGFGSPLDRLSAGSVD
Homo sapiens	(101)	HKG KQK VVDHP	KRR FGIPMDRIGRNRLNSRG-
Bos taurus	(100)	HKG KQK VVDHP	KRR FGIPMDRIGRNRLNSRG-
Mus musculus	(98)	HRG KQK AVDHS	KRR FGIPMDRIGRNRLSSSRG-
Rattus norvegicus	(100)	HRG KQK VVDHS	KRR FGIPMDRIGRNRLSSSRG-
Gallus gallus	(101)	AKG KQK VVELP	KRR FGVPLDRIGVSRLGNTKG-
Python molurus bivittatus		-KA KQK AVELP	KRR FGIPLDRIGVNRLSGSRG
Consensus	(101)	HKG KQK VVDHP	KRR FGIPMDRIGRNRLNSRG

Figure 5B

Human	BP-1		F ¹¹⁶ G	I P M	D R I G ¹²⁴	
Rat	Atrial natriuretic factor	C ¹²⁹	F G G R	I	D R I G	A Q S G L G C ¹⁴⁵
Rat	Brain natriuretic factor	C ⁹⁹	F G Q K	I	D R I G	A V S R L G C ¹¹⁵
Rat	C-type natriuretic factor	C ¹¹⁰	F G L K	L	D R I G	S M S G L G C ¹²⁶

Figure 6A

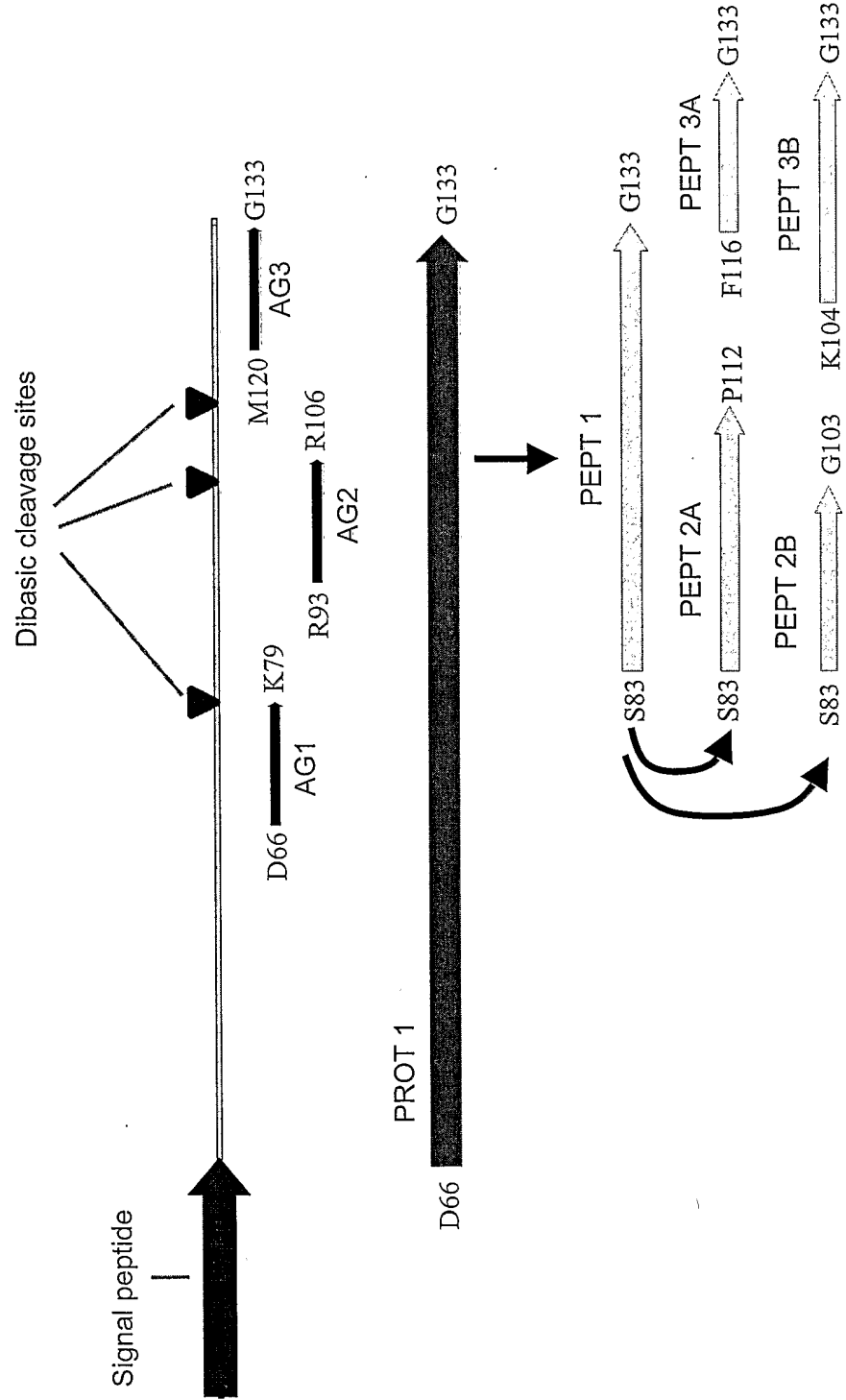


Figure 6B

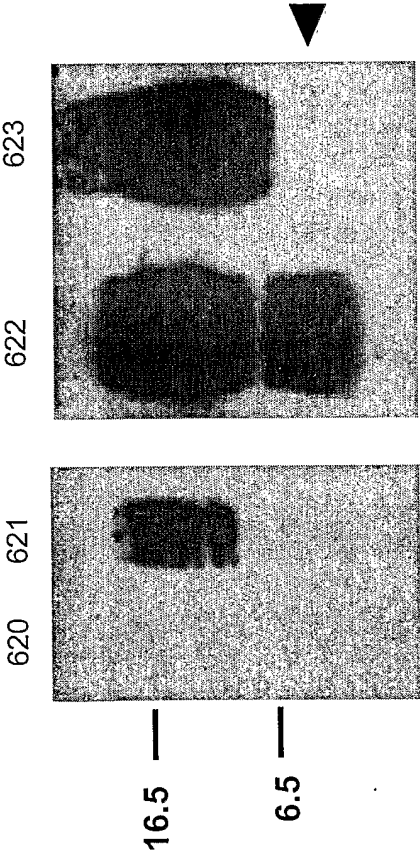
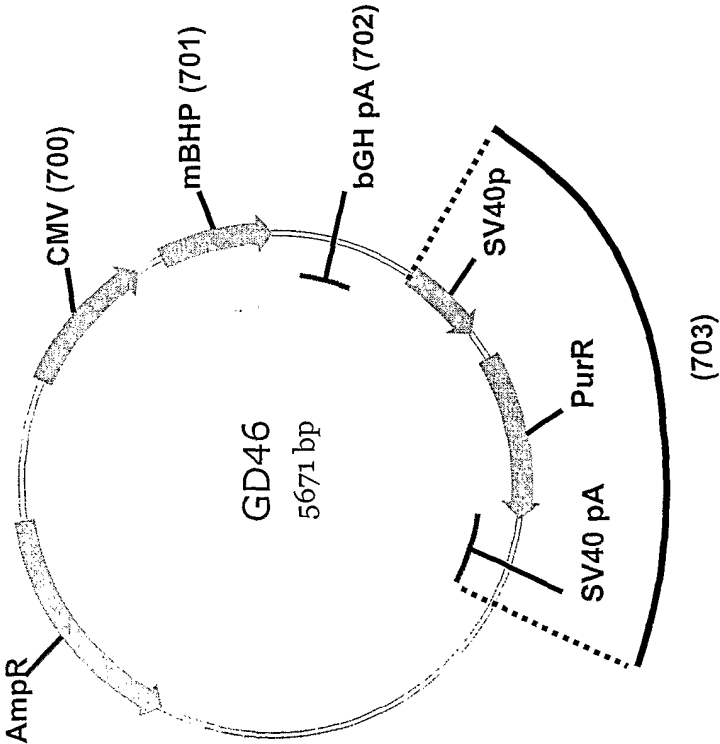


Figure 7

A



B

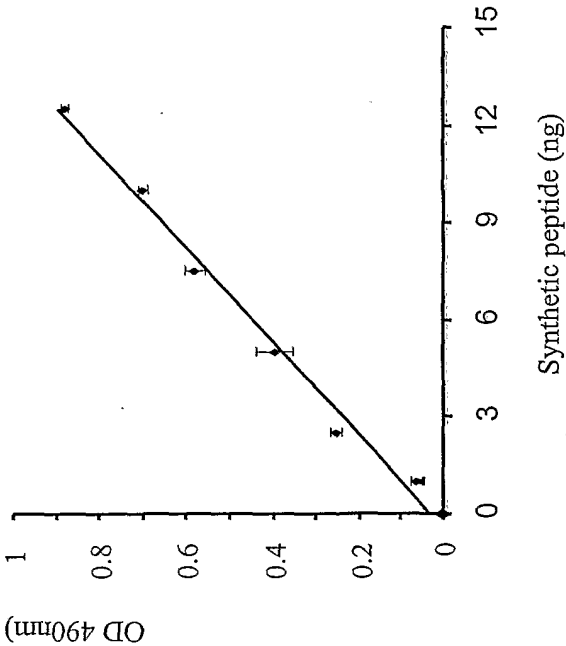


Figure 8

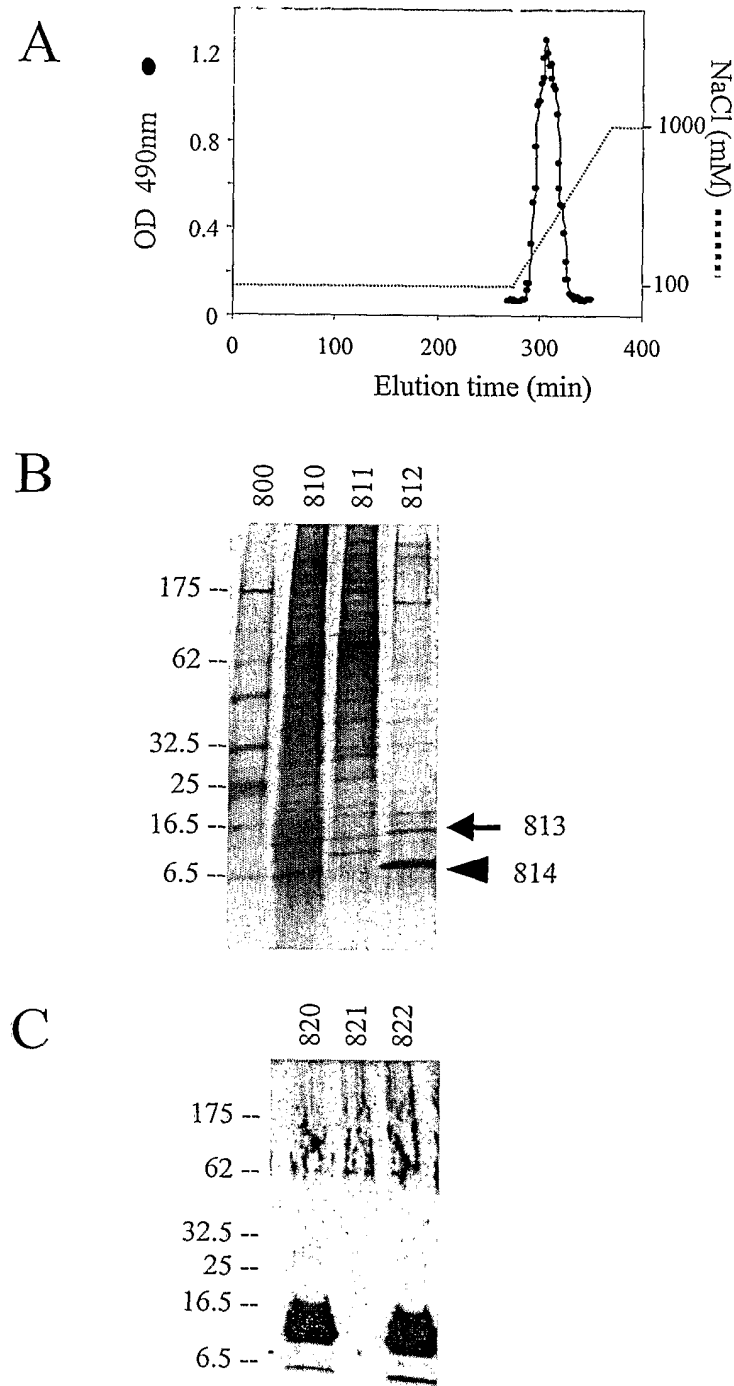
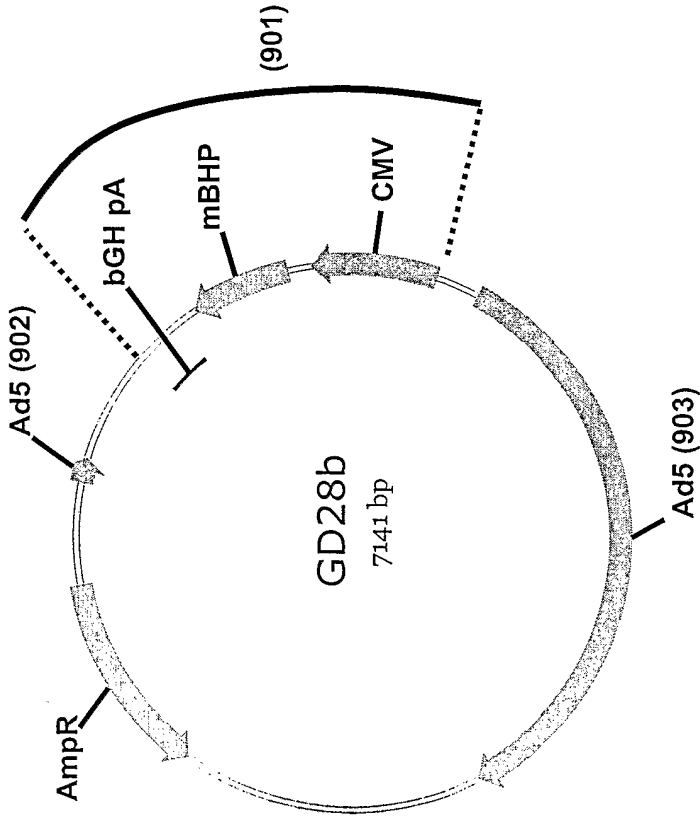


Figure 9

A



B

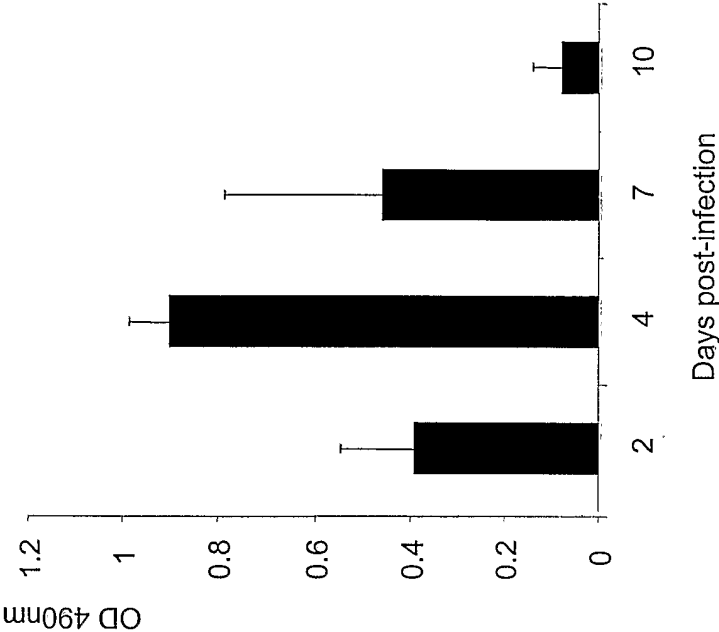
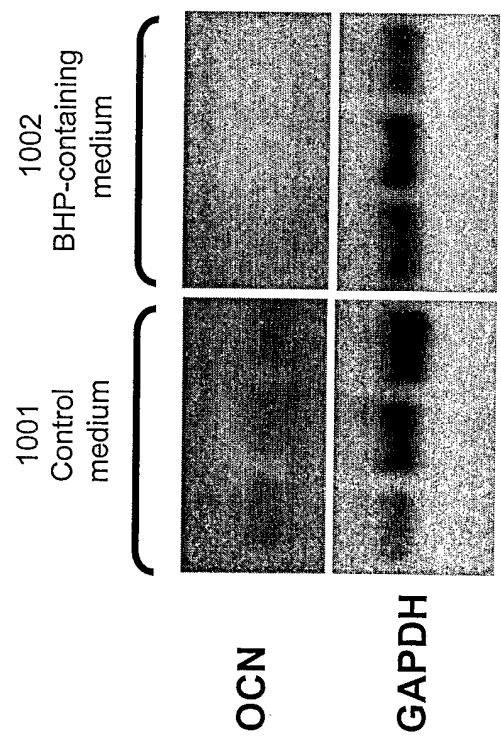


Figure 10



SEQUENCE LISTING

<110> PHENOGENE THERAPEUTIQUES INC.
 LANCTOT, Christian
 SALOIS, Patrick
 MOFFATT, Pierre

<120> BONE POLYPEPTIDE-1

<130> 544172000140

<150> US 60/341,224

<151> 2001-12-20

<160> 45

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1280

<212> DNA

<213> Mus musculus

<400> 1

```

gctaagtttg ggataagctg caggcgggac tccaaagtta ggagctctga cttctcacia 60
gatgctggac tggagatttg caagtacaca cttcatcctg gctatgattg tgatgctgtg 120
gggctcagga aaggcattct ctgtggactt agcatcacag gagtttggaa cagcaagctt 180
gcagtctcca cccacagcca gagaagagaa gtcagccact gagctttcgg ctaagctcct 240
gcgtcttgat gatctgggtg ccttagagaa tgacgtatct gagaccaaga aaaagagaag 300
cttctctggc tttgggtctc cccttgacag actctcagct gggctctgtg agcatagagg 360
gaaacaaagg aaagcagtag atcattcaaa aaagcgggtt ggtattccca tggatcggat 420
tggtagaaac cggctctcca gttccagagg ctgatggatt cttattgtgc gacttacttg 480
tgtgagatgg cacagaacta tagaagacac ttcagtgaag ttcactaccc cttttgtcaa 540
ggaattggcc tttcgcaaac cttcccaaag cttgatcctc cccagaccat cacgtcatag 600
tggtgctgtg gttttagttg agttgtgcag atcatttcag tgcattggata tctctgaaag 660
tatttttcaa tgattcccaa attgtaacgt ggcccctgaa cctacttttt ttaaacagca 720
gaccaatata atgcattctc ttgccattaa tattttcaca tttcagttaa tcaatgtgct 780
ttctagaaac ctagtgtctg aagatctgat gatctaaaga aatcagaaat gagcacatgg 840
tgatttatat aggtttcttt agtttttctg aggtttgtcg aattgttgta aacttcaact 900
tcaagcttag aaaaaagaca ttacatgagt gtttgcttca actgtgtcag agggcacaata 960
aattttgaga aacctgagca attgtgttct ttaggaacta ataaaggata gtataattgg 1020
cccatatgta atattctgac aaactctgaa tgtaaaagac tcatattgaaa agaagttact 1080
gcctggcctg tttacttcta ccagcctagg ggtgaattgt tcaaatgttt cctatgttag 1140
cagcttttct tcttcttttt tttcttcta ttttactttt tttcttcatt caatgtttat 1200
aagctaaaaa tccaaccaa tagtgctttg tgctttaaaa gggggtatta aaatcaacat 1260
taatctaaaa aaaaaaaaaa

```

<210> 2

<211> 402

<212> DNA

<213> Homo sapiens

<400> 2

```

atgctggact ggagattggc aagtgcacat ttcacatcctg ctgtgacact gacactgtgg 60
agctcaggaa aagtcctctc agtagatgta acaacaacag aggcctttga ttctggagtc 120
atagatgtgc agtcaacacc cacagtcagg gaagagaaat cagccactga cctgacagca 180
aaaactcttg tcttgatga attggtgtcc ctagaaaatg atgtgattga gacaaagaag 240
aaaaggagtt tctctgggtt tgggtctccc cttgacagac tctcagctgg ctctgtagat 300
cacaaaggta aacagaggaa agtagtagat catccaaaaa ggcgatttgg tatccccatg 360

```

gatcggattg gtagaaaccg gctttcaa at tccagaggct aa

402

<210> 3

<211> 393

<212> DNA

<213> Mus musculus

<400> 3

atgctggact ggagattggc aagtacacac ttcacctctg ctatgattgt gatgctgtgg 60
 ggctcaggaa aggcatcttc tgtggactta gcatcacagg agtttggaa agcaagcttg 120
 cagtctccac ccacagccag agaagagaag tcagccactg agctttcggc taagctcctg 180
 cgtcttgatg atctgggtgtc cttagagaat gacgtatttg agaccaagaa aaagagaagc 240
 ttctctggct ttgggtctcc ccttgacaga ctctcagctg ggtctgtaga gcatagaggg 300
 aaacaaagga aagcagtaga tcattcaaaa aagcggtttg gtattcccat ggatcggatt 360
 ggtagaaacc ggctctccag ttccagaggc tga 393

<210> 4

<211> 399

<212> DNA

<213> Rattus norvegicus

<400> 4

atgctggact ggagattggc aagtgcacac ttctctcttg ctatgacct gatgctgtgg 60
 ggctcaggaa aggcatcttc cgtggactta gcatcagagg cctccgagtt tggagcagaa 120
 agcttgcagt cccacccac aaccagagaa gagaagtcag ccacggagct tgcagctaag 180
 ctctgcttc ttgatgatct ggtgtccttg gagaatgatg tgtttgagac caagaagaag 240
 agaagcttct ctggcttcgg gtctccctt gacagactct cggctgggtc tgtagagcat 300
 agagggaaac aaaggagagt agttgatcat tcaaaaaagc gatttggtat tcccatggat 360
 cgaattggta gaaaccgtct ctccagttcc aggggctga 399

<210> 5

<211> 399

<212> DNA

<213> Bos taurus

<400> 5

atgctggact ggagattagc aagtgcacat tttatctctg ctatgacact gatgctctgg 60
 agctcaggaa aagtgttctc agtgggtgtc acaacagagg cctttgattc tggagtctta 120
 ggtgttcagt catcaccac agtcagagaa gcgaagtcgg cactgacct ggcagcaaaa 180
 ctcttacttc ttgatgaact tgtgtctctg gagaatgacg tgattgaaac aaagaagaaa 240
 agaagcttct ctgggttttg ttctcccttg gacagactct cagctggctc tgtaagtcac 300
 aaaggtaaac agaggaaagt agtagatcat ccaaaaaggc gatttggtat ccctatggat 360
 cggattggaa gaaaccggct ttcaaattcc agaggctaa 399

<210> 6

<211> 402

<212> DNA

<213> Gallus gallus

<400> 6

atgctgcagt tccagcttgt tgtggtccat ctggcccttg tgatcaccct gctgcagtgg 60
 cattctagtt cagtgtctct tgcagaggca gctccagagc ctttggagcc ttctgctgct 120
 ctgggcatgg cagcacatcc tactgccagc gaggagaagt cagcctccag cctggcagcc 180
 aaactgctcc ttcttgatga gttggtgtct ctggagaatg aggttaactga gacaaagaag 240
 aaaagaagtt ttccaggatt tggctccccg atcgacagaa tttctgagc atctgtggat 300
 gctaaaggca aacagaggaa agtggttgag ctgcctaaga gacgggtttg agttcctctt 360
 gaccggatcg gagtgagtcg tcttggcaac accaagggtt ag 402

<210> 7

<211> 253

<212> DNA

<213> Python molurus bivittatus

<400> 7

```
tacggcgtcg gaggagaagt cggctactga cctgggtggcc aaaattttgc tcctcaacga 60
attggtgtcc cttgaaaacg atgtctttga gaccaagaag aagaggagct tctccgggtt 120
tggctcccca cttgacagac tttcgggtggg cctgaaagcc aagcagagga aagctgtgga 180
gctgccaaag aagcgggttg ggattcctct agatcggatt ggcgtgaatc gtttgagcgg 240
ctccagaggt tag 253
```

<210> 8

<211> 405

<212> DNA

<213> Artificial Sequence

<220>

<223> Consensus DNA sequence for BP-1

<221> misc_feature

<222> 48, 61, 75, 81, 87, 94, 97, 98, 99, 108, 112, 113, 114, 123, 124, 130, 183, 261, 312

<223> n = A,T,C or G

<400> 8

```
atgctggact ggagattggc aagtgcacat ttcacacctg ctatgacnct gatgctgtgg 60
ngctcaggaa aagtnttctc ngtggangta gcancannng aggccttnga gnnntctgga 120
gcnntaggcn tgcagtcacc acccacagcc agagaagaga agtcagccac tgacctggca 180
gcnaaactct tgcttcttga tgaattggtg tccctggaga atgatgtgtt tgagaccaag 240
aagaagagaa gcttctctgg ntttgggtct ccccttgaca gactctcagc tgggtctgta 300
gatcataaag gnaaacagag gaaagtagta gatcatccaa aaaggcgggt tggtattcct 360
atggatcgga ttggtagaaa ccgtctttcc agttccagag gctaa 405
```

<210> 9

<211> 133

<212> PRT

<213> Homo sapiens

<400> 9

```
Met Leu Asp Trp Arg Leu Ala Ser Ala His Phe Ile Leu Ala Val Thr
1           5           10          15
Leu Thr Leu Trp Ser Ser Gly Lys Val Leu Ser Val Asp Val Thr Thr
20          25          30
Thr Glu Ala Phe Asp Ser Gly Val Ile Asp Val Gln Ser Thr Pro Thr
35          40          45
Val Arg Glu Glu Lys Ser Ala Thr Asp Leu Thr Ala Lys Leu Leu Leu
50          55          60
Leu Asp Glu Leu Val Ser Leu Glu Asn Asp Val Ile Glu Thr Lys Lys
65          70          75          80
Lys Arg Ser Phe Ser Gly Phe Gly Ser Pro Leu Asp Arg Leu Ser Ala
85          90          95
Gly Ser Val Asp His Lys Gly Lys Gln Arg Lys Val Val Asp His Pro
100         105         110
Lys Arg Arg Phe Gly Ile Pro Met Asp Arg Ile Gly Arg Asn Arg Leu
115         120         125
Ser Asn Ser Arg Gly
130
```

<210> 10

<211> 130

<212> PRT

<213> Mus musculus

<400> 10

```

Met Leu Asp Trp Arg Leu Ala Ser Thr His Phe Ile Leu Ala Met Ile
 1           5           10           15
Val Met Leu Trp Gly Ser Gly Lys Ala Phe Ser Val Asp Leu Ala Ser
          20           25           30
Gln Glu Phe Gly Thr Ala Ser Leu Gln Ser Pro Pro Thr Ala Arg Glu
          35           40           45
Glu Lys Ser Ala Thr Glu Leu Ser Ala Lys Leu Leu Arg Leu Asp Asp
          50           55           60
Leu Val Ser Leu Glu Asn Asp Val Phe Glu Thr Lys Lys Lys Arg Ser
65           70           75           80
Phe Ser Gly Phe Gly Ser Pro Leu Asp Arg Leu Ser Ala Gly Ser Val
          85           90           95
Glu His Arg Gly Lys Gln Arg Lys Ala Val Asp His Ser Lys Lys Arg
          100          105          110
Phe Gly Ile Pro Met Asp Arg Ile Gly Arg Asn Arg Leu Ser Ser Ser
          115          120          125
Arg Gly
          130

```

<210> 11

<211> 132

<212> PRT

<213> Rattus norvegicus

<400> 11

```

Met Leu Asp Trp Arg Leu Ala Ser Ala His Phe Leu Leu Ala Met Ile
 1           5           10           15
Leu Met Leu Trp Gly Ser Gly Lys Ala Phe Ser Val Asp Leu Ala Ser
          20           25           30
Glu Ala Ser Glu Phe Gly Ala Glu Ser Leu Gln Ser Pro Pro Thr Thr
          35           40           45
Arg Glu Glu Lys Ser Ala Thr Glu Leu Ala Ala Lys Leu Leu Leu Leu
          50           55           60
Asp Asp Leu Val Ser Leu Glu Asn Asp Val Phe Glu Thr Lys Lys Lys
65           70           75           80
Arg Ser Phe Ser Gly Phe Gly Ser Pro Leu Asp Arg Leu Ser Ala Gly
          85           90           95
Ser Val Glu His Arg Gly Lys Gln Arg Arg Val Val Asp His Ser Lys
          100          105          110
Lys Arg Phe Gly Ile Pro Met Asp Arg Ile Gly Arg Asn Arg Leu Ser
          115          120          125
Ser Ser Arg Gly
          130

```

<210> 12

<211> 132

<212> PRT

<213> Bos taurus

<400> 12

```

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

```

```

      50              55              60
Asp Glu Leu Val Ser Leu Glu Asn Asp Val Ile Glu Thr Lys Lys Lys
65              70              75              80
Arg Ser Phe Ser Gly Phe Gly Ser Pro Leu Asp Arg Leu Ser Ala Gly
      85              90              95
Ser Val Ser His Lys Gly Lys Gln Arg Lys Val Val Asp His Pro Lys
      100              105              110
Arg Arg Phe Gly Ile Pro Met Asp Arg Ile Gly Arg Asn Arg Leu Ser
      115              120              125
Asn Ser Arg Gly
      130

```

<210> 13
 <211> 133
 <212> PRT
 <213> Gallus gallus

```

<400> 13
Met Leu Gln Phe Gln Leu Val Val Val His Leu Ala Leu Val Ile Thr
 1              5              10              15
Leu Leu Gln Trp His Ser Ser Ser Val Leu Leu Ala Glu Ala Ala Pro
      20              25              30
Glu Pro Leu Glu Pro Ser Ala Ala Leu Gly Met Ala Ala His Pro Thr
      35              40              45
Ala Ser Glu Glu Lys Ser Ala Ser Ser Leu Ala Ala Lys Leu Leu Leu
      50              55              60
Leu Asp Glu Leu Val Ser Leu Glu Asn Glu Val Thr Glu Thr Lys Lys
65              70              75              80
Lys Arg Ser Phe Pro Gly Phe Gly Ser Pro Ile Asp Arg Ile Ser Ala
      85              90              95
Thr Ser Val Asp Ala Lys Gly Lys Gln Arg Lys Val Val Glu Leu Pro
      100              105              110
Lys Arg Arg Phe Gly Val Pro Leu Asp Arg Ile Gly Val Ser Arg Leu
      115              120              125
Gly Asn Thr Lys Gly
      130

```

<210> 14
 <211> 75
 <212> PRT
 <213> Python molurus bivittatus

```

<400> 14
Thr Asp Leu Val Ala Lys Ile Leu Leu Leu Asn Glu Leu Val Ser Leu
 1              5              10              15
Glu Asn Asp Val Phe Glu Thr Lys Lys Lys Arg Ser Phe Ser Gly Phe
      20              25              30
Gly Ser Pro Leu Asp Arg Leu Ser Val Gly Leu Lys Ala Lys Gln Arg
      35              40              45
Lys Ala Val Glu Leu Pro Lys Lys Arg Phe Gly Ile Pro Leu Asp Arg
      50              55              60
Ile Gly Val Asn Arg Leu Ser Gly Ser Arg Gly
65              70              75

```

<210> 15
 <211> 132
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Consensus polypeptide sequence for BP-1

<221> VARIANT

<222> 21, 34, 35, 36, 40, 42, 46, 49

<223> Xaa = Any Amino Acid

<400> 15

```

Met Leu Asp Trp Arg Leu Ala Ser Ala His Phe Ile Leu Ala Met Thr
 1           5           10           15
Leu Met Leu Trp Xaa Ser Gly Lys Val Phe Ser Val Asp Leu Ala Ser
 20           25           30
Glu Xaa Xaa Xaa Asp Ser Gly Xaa Leu Xaa Leu Gln Ser Xaa Pro Thr
 35           40           45
Xaa Glu Glu Lys Ser Ala Thr Asp Leu Ala Ala Lys Leu Leu Leu Leu
 50           55           60
Asp Glu Leu Val Ser Leu Glu Asn Asp Val Phe Glu Thr Lys Lys Lys
 65           70           75           80
Arg Ser Phe Ser Gly Phe Gly Ser Pro Leu Asp Arg Leu Ser Ala Gly
 85           90           95
Ser Val Asp His Lys Gly Lys Gln Arg Lys Val Val Asp His Pro Lys
 100          105          110
Lys Arg Phe Gly Ile Pro Met Asp Arg Ile Gly Arg Asn Arg Leu Ser
 115          120          125
Asn Ser Arg Gly
 130

```

<210> 16

<211> 9

<212> PRT

<213> Homo sapiens

<400> 16

```

Phe Gly Ile Pro Met Asp Arg Ile Gly
 1           5

```

<210> 17

<211> 17

<212> PRT

<213> Rattus norvegicus

<400> 17

```

Cys Phe Gly Gly Arg Ile Asp Arg Ile Gly Ala Gln Ser Gly Leu Gly
 1           5           10           15
Cys

```

<210> 18

<211> 17

<212> PRT

<213> Rattus norvegicus

<400> 18

```

Cys Phe Gly Gln Lys Ile Asp Arg Ile Gly Ala Val Ser Arg Leu Gly
 1           5           10           15
Cys

```

<210> 19
 <211> 17
 <212> PRT
 <213> Rattus norvegicus

<400> 19
 Cys Phe Gly Leu Lys Leu Asp Arg Ile Gly Ser Met Ser Gly Leu Gly
 1 5 10 15
 Cys

<210> 20
 <211> 106
 <212> PRT
 <213> Homo sapiens

<400> 20
 Val Asp Val Thr Thr Thr Glu Ala Phe Asp Ser Gly Val Ile Asp Val
 1 5 10 15
 Gln Ser Thr Pro Thr Val Arg Glu Glu Lys Ser Ala Thr Asp Leu Thr
 20 25 30
 Ala Lys Leu Leu Leu Asp Glu Leu Val Ser Leu Glu Asn Asp Val
 35 40 45
 Ile Glu Thr Lys Lys Lys Arg Ser Phe Ser Gly Phe Gly Ser Pro Leu
 50 55 60
 Asp Arg Leu Ser Ala Gly Ser Val Asp His Lys Gly Lys Gln Arg Lys
 65 70 75 80
 Val Val Asp His Pro Lys Arg Arg Phe Gly Ile Pro Met Asp Arg Ile
 85 90 95
 Gly Arg Asn Arg Leu Ser Asn Ser Arg Gly
 100 105

<210> 21
 <211> 51
 <212> PRT
 <213> Homo sapiens

<400> 21
 Ser Phe Ser Gly Phe Gly Ser Pro Leu Asp Arg Leu Ser Ala Gly Ser
 1 5 10 15
 Val Asp His Lys Gly Lys Gln Arg Lys Val Val Asp His Pro Lys Arg
 20 25 30
 Arg Phe Gly Ile Pro Met Asp Arg Ile Gly Arg Asn Arg Leu Ser Asn
 35 40 45
 Ser Arg Gly
 50

<210> 22
 <211> 30
 <212> PRT
 <213> Homo sapiens

<400> 22
 Ser Phe Ser Gly Phe Gly Ser Pro Leu Asp Arg Leu Ser Ala Gly Ser
 1 5 10 15
 Val Asp His Lys Gly Lys Gln Arg Lys Val Val Asp His Pro

20

25

30

<210> 23

<211> 18

<212> PRT

<213> Homo sapiens

<400> 23

Phe Gly Ile Pro Met Asp Arg Ile Gly Arg Asn Arg Leu Ser Asn Ser

1

5

10

15

Arg Gly

<210> 24

<211> 21

<212> PRT

<213> Homo sapiens

<400> 24

Ser Phe Ser Gly Phe Gly Ser Pro Leu Asp Arg Leu Ser Ala Gly Ser

1

5

10

15

Val Asp His Lys Gly

20

<210> 25

<211> 26

<212> PRT

<213> Homo sapiens

<400> 25

Val Val Asp His Pro Lys Arg Arg Phe Gly Ile Pro Met Asp Arg Ile

1

5

10

15

Gly Arg Asn Arg Leu Ser Asn Ser Arg Gly

20

25

<210> 26

<211> 14

<212> PRT

<213> Homo sapiens

<400> 26

Asp Glu Leu Val Ser Leu Glu Asn Asp Val Ile Glu Thr Lys

1

5

10

<210> 27

<211> 14

<212> PRT

<213> Homo sapiens

<400> 27

Arg Leu Ser Ala Gly Ser Val Asp His Lys Gly Lys Gln Arg

1

5

10

<210> 28

<211> 14

<212> PRT

<213> Homo sapiens

<400> 28

Met Asp Arg Ile Gly Arg Asn Arg Leu Ser Asn Ser Arg Gly
1 5 10

<210> 29

<211> 30

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic oligoribonucleotide

<400> 29

agcaucgagu cggccuuguu ggccuacugg 30

<210> 30

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic primer-linker

<221> misc_feature

<222> 27, 28, 29, 30, 31, 32, 33, 34, 35

<223> n = A,T,C or G

<400> 30

gagagagaga gagcgactcg gatccannnn nnnnnc 36

<210> 31

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic primer

<400> 31

agcatcgagt cggccttg 18

<210> 32

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic primer

<400> 32

gagagcgact cggatcca 18

<210> 33

<211> 32

<212> DNA

<213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 33
 ggacgagcgg ccgcgcccttg ttggcctact gg 32

<210> 34
 <211> 7
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic construct

<400> 34
 aagatgc 7

<210> 35
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic oligonucleotide

<400> 35
 gtgcggcggc gtttaaacgg tcacctcgag 30

<210> 36
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic oligonucleotide

<400> 36
 ctcgaggtga ccgtttaaac gcggccgcac 30

<210> 37
 <211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer-linker

<400> 37
 gagatgaatt cctcgagctt tttttttttt ttt 33

<210> 38
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 38
 aaacgctctg acttctcaca agatg 25

<210> 39
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 39
 gagatgaatt cctcgagc 18

<210> 40
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 40
 agatgctgga ctggagat 18

<210> 41
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 41
 gaatcaatta gcctctgga 19

<210> 42
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 42
 ggaaagaact tcactgaagt g 21

<210> 43
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 43
 ccatcagcct ctggaactgg agag 24

<210> 44
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Synthetic primer

<221> modified_base

<222> 3, 6, 9

<223> n= I

<400> 44

acnrynhsng argmgaartc rgc

23

<210> 45

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic primer

<221> modified_base

<222> 6, 15, 21

<223> n= I

<400> 45

ctrctnadsg grswnccraa ncc

23