

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 September 2007 (07.09.2007)

PCT

(10) International Publication Number
WO 2007/100357 A2

(51) International Patent Classification: Not classified

Alberta T6R-2K1 (CA). **FUNK, Walter** [CA/US]; 25432 Modoc Court, Hayward, CA 94542 (US).

(21) International Application Number:
PCT/US2006/039266

(74) Agents: **ROBINS, Roberta L.** et al.; Robins & Pasternak LLP, 1731 Embarcadero Road, Suite 230, Palo Alto, CA 94303 (US).

(22) International Filing Date: 6 October 2006 (06.10.2006)

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

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/724,908 7 October 2005 (07.10.2005) US

(71) Applicant (for all designated States except US): **NUVELO, INC.** [US/US]; 201 Industrial Road, Suite 310, San Carlos, CA 94070 (US).

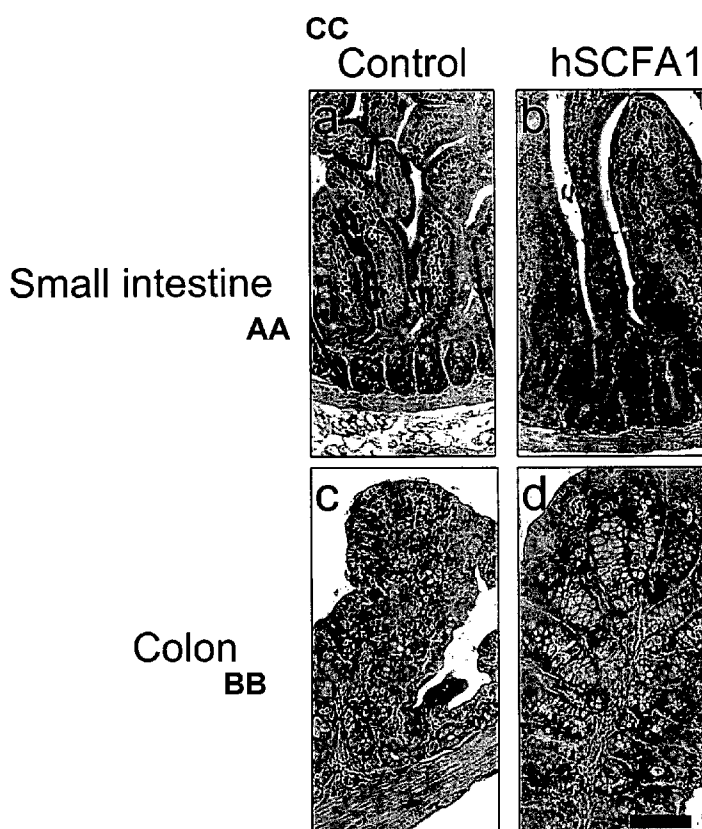
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

(72) Inventors; and
(75) Inventors/Applicants (for US only): **EMTAGE, Peter C.R.** [CA/CA]; 1129 Carter Crest Road, Edmonton,

[Continued on next page]

(54) Title: STEM CELL FACTOR-LIKE PROTEIN SCFA1 AND USES THEREOF

(57) Abstract: Methods for stimulating epithelial cell proliferation and for treating oral and gastrointestinal disorders are described. The methods use compositions comprising stem cell factor-like protein A1 (SCFA1) polypeptides and polynucleotides.





FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

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.

Published:

- *without international search report and to be republished upon receipt of that report*

STEM CELL FACTOR-LIKE PROTEIN SCFA1 AND USES THEREOF

CROSS-REFERENCE TO RELATED APPLICATION

5 This application claims the benefit under 35 U.S.C. §119(e)(1) of U.S. Provisional Application 60/724,908, filed October 7, 2005, which application is incorporated herein by reference in its entirety.

1. BACKGROUND

1.1 FIELD OF THE INVENTION

10 The present invention provides a method for augmenting proliferation of gastrointestinal epithelial cells. The invention further provides a method for treating or preventing mucositis in patients undergoing cancer treatment, to treat patients with inflammatory bowel disease, and to ameliorate digestion and nutritional
15 absorption problems in patients with short bowel syndrome.

1.2 SEQUENCE LISTING

20 The sequences of the polynucleotides and polypeptides of the invention are listed in the Sequence Listing and are submitted on a disc containing the file labeled "NUVO-27PCT.ST25.txt"— 48.6 KB (49,771 bytes) which was created on an IBM PC, Windows 2000 operating system on September 25, 2006 at 11:01:04 AM. The Sequence Listing entitled "NUVO-27PCT.ST25.txt" is herein incorporated by reference in its entirety. A computer readable format ("CRF") and a paper copy of
25 the Sequence Listing "NUVO-27PCT.ST25.txt" are submitted herein. Applicants hereby state that the content of the CRF and the paper copy of the Sequence Listing, submitted in accordance with 37 CFR §1.821(c) and (e), respectively, are the same.

1.3 BACKGROUND

30 Ionizing radiation therapy and cytotoxic chemotherapy produce injuries to the oral and gastrointestinal mucosa, which remain significant problems for patients undergoing anti-neoplastic treatments. Mucositis is the inflammation of the mucous

membranes and is a particularly common problem in this patient population due to the use of chemotherapy and radiation therapy used with curative or palliative intent. The mucosal injuries to the gastrointestinal tract seen with radiation and chemotherapy (to these areas) include the destruction of crypt cells, a decrease in villous height and ulceration and necrosis of the gastrointestinal epithelium (Berthrong M, *World J Surg* 10:155-170 (1986)), which underlie disorders including gastrointestinal mucositis and enterocolitis. To the patients this can mean abdominal pain, bloody diarrhea, malabsorption and in some cases bacterial translocation (Guzman et al., *J Surg Res* 46:104-107 (1989)). In addition, chemotherapy and ionizing radiation can affect other mucous membranes including those of the oropharynx and lips, and those of the esophagus. It is well known that combined modality therapy of concurrent radiation and chemotherapy can produce highly symptomatic stomatitis in patients with head and neck cancer, and esophagitis in patients with small cell lung cancer.

Chemotherapy and radiation therapy cause injury to the oral and gastrointestinal mucosa through direct and indirect toxicity. The mechanism for direct mucositis is nonspecific cell killing of rapidly dividing basal epithelial cells that results in epithelial thinning, inflammation, decreased cell renewal, and ultimately ulceration. These painful lesions also produce an increased risk for local and systemic infection. Indirect mucotoxicity is a byproduct of chemotherapy-induced myelosuppression, which permits bacterial and viral infections at the site of direct mucosal injury. The severity of these effects may preclude dose escalation, delay treatment, and warrant dose reductions, thus limiting the effectiveness of cancer therapy.

A well-established prophylaxis or therapy for chemotherapy and radiation therapy-induced (mucosal) gastrointestinal injuries (mucositis) is unavailable, other than a prescription of suboptimal doses of chemotherapy or radiotherapy, a downward dose modification in subsequent treatment courses following toxicity, or the use of specific antidotes such as leucovorin after moderate-dose or high-dose methotrexate (Allegra CJ. Antifolates. *In*: Chabner and Collins, eds. *Cancer Chemotherapy: Principles and Practice*. Philadelphia, Pa. JP Lippincott Co; 1990:110-153.)

Injury to the gastrointestinal mucosa is also associated with chronic inflammatory disorders of the gastrointestinal tract, which are collectively referred to as inflammatory bowel disease. While cytokine-based therapies are available for the treatment of inflammatory bowel disease, none can be considered as a permanent cure (Bouma and Strober, *Nature Rev* 3:521-533 (2003)). Often resection of the small intestine is indicated in patients with inflammatory bowel disease such as Crohn's disease. Surgical resection of the small intestine may also be necessary following traumatic injury, vascular accidents, and cancer. Surgical resection that leaves less than 200 cm of viable small bowel places a patient at risk for developing short-bowel syndrome (SBS). SBS is a disorder that is clinically defined by malabsorption, diarrhea, fluid and electrolyte disturbances, and malnutrition. The management of patients with SBS frequently requires long-term, if not life long use of parenteral nutrition DiBaise et al., *Am J Gastroenterol* 99:1823-1832 (2004)).

Thus, there is a need to find agents that may be used prophylactically or therapeutically to increase the tolerance to anti-neoplastic treatments, to advance current therapies for treating inflammatory bowel disease, and to restore the digestive and absorptive processes that are compromised following surgical resection of the intestine.

To this end, the inventors herein have discovered an agent that induces the proliferation of gastrointestinal epithelial cells, and which may be useful for treating conditions in which proliferation of epithelial cells may be desired.

2. SUMMARY OF THE INVENTION

The present invention is based on the discovery that stem cell factor-like protein A1, SCFA1, induces the proliferation of epithelial cells of the gastrointestinal tract. Thus, compositions comprising SCFA1, fragments or analogs thereof, may be used for the treatment of conditions where epithelialization is required, such as for the treatment of oral and gastrointestinal disorders including chemotherapy and radiation therapy-induced mucositis, mucositis of the oropharynx, lips and esophagus, inflammatory bowel disease, and other conditions including wounds, burns, ophthalmic disorders, and any disorder where stimulation of epithelial cell proliferation or regeneration is desired.

The compositions for use with the present invention include isolated polynucleotides (SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 and 15) encoding SCFA1 polypeptides, including recombinant DNA molecules, and cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants. The compositions also include vectors such as expression vectors containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The compositions for use with the invention comprise isolated polynucleotides that include, but are not limited to, a SCFA1 polynucleotide, a fragment, or variant thereof; a polynucleotide comprising the full length protein coding sequence of SEQ ID NO: 1 (for example, SEQ ID NO: 2); a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of SEQ ID NO: 9 (for example SEQ ID NO: 10); a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of SEQ ID NO: 3, 5, or 9 (for example SEQ ID NO: 4, 6, or 10); a polynucleotide comprising the nucleotide sequence of the thrombospondin domain of SEQ ID NO: 15 (for example SEQ ID NO: 16); a polynucleotide of SEQ ID NO: 11 or 13 comprising the nucleotide sequence that encodes furin-like cysteine-rich domains (for example SEQ ID NO: 12 or 14). The polynucleotide compositions of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any of the nucleotide sequences set forth in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, or 15 (b) a nucleotide sequence encoding any of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, or 16; a polynucleotide which is an allelic variant of any polynucleotides recited above having at least 70% polynucleotide sequence identity to the polynucleotides; a polynucleotide which encodes a species homolog (e.g. ortholog) of any of the peptides recited above; or a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptide of SEQ ID NO: 2, 4, 6, 8 or 10.

This invention further provides cloning or expression vectors comprising at least a fragment of the polynucleotides set forth above and host cells or organisms transformed with these expression vectors. Useful vectors include plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the

invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

An embodiment of the invention is directed to the use of a composition comprising a pharmaceutically effective amount of SCFA1 polypeptide and a pharmaceutically acceptable carrier.

The pharmaceutical compositions for use with the present invention include polypeptides comprising, but not limited to, an isolated polypeptide selected from the group comprising the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14 or 16. Polypeptides for use with the present invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active analogs of any of the protein sequences listed as SEQ ID NO: 2, 4, 6, 8, 10, 12, 14 or 16, and substantial equivalents thereof that retain biological activity are also contemplated. The polypeptides may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The polypeptides can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantifying the polypeptide in tissue.

In further embodiments, the subject invention is directed to a method of stimulating epithelial cell proliferation. The method comprises contacting epithelial cells with a composition that includes a therapeutically effective amount of a SCFA1 polypeptide, fragment or analog thereof, and a pharmaceutically acceptable carrier. Specifically, a subject in need of stimulation (including cytoprotection, proliferation and/or differentiation) of epithelial cells will be administered therapeutically-effective

or prophylactically-effective amounts of an SCFA1 polypeptide, such as any of the above SCFA1 proteins, fragments or analogs thereof.

In all the methods described, epithelial cells may be contacted with the SCFA1 polypeptides *in vitro* or *in vivo*.

5 Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a peptide of the present invention and a pharmaceutically acceptable carrier.

10 In particular, the SCFA1 polypeptides described herein may be used to induce the proliferation and/or differentiation of gastrointestinal crypt cells to regenerate the epithelial layer of the alimentary tract. Thus, the SCFA1 polypeptides and polynucleotides may be used in the treatment of chemotherapy or radiation therapy-induced mucositis and enterocolitis. SCFA1 polypeptides may also be used to increase absorptive surface area of unresected intestinal tissue
15 thereby ameliorating symptoms associated with short bowel syndrome.

They may also be used in the treatment of diseases, and other conditions including wounds, burns, ophthalmic disorders, and any disorder where stimulation of epithelial cell proliferation or regeneration is desired.

20 Polynucleotides and polypeptides described herein may also be used as markers of differentiation and development of gastrointestinal epithelium.

25 Thus, in certain embodiments, the invention is directed to a method of stimulating epithelial cell proliferation in a subject in need thereof, the method comprising administering to the subject a therapeutically effective amount of a composition comprising a SCFA1 polypeptide and a pharmaceutically acceptable carrier. In certain embodiments, the SCFA1 polypeptide comprises the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10, or a sequence of amino acids with at least 90% sequence identity to the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8 or 10. In certain embodiments of the above method,
30 the method comprises stimulating epithelial cell proliferation in the oral or gastrointestinal tract, such as stimulating epithelial cell proliferation in the esophagus, in the small intestine, in the large intestine, in the stomach and/or in the oral cavity.

In additional embodiments, the invention is directed to a method of treating a gastrointestinal disorder or oral mucosal disorder in a mammalian subject in need thereof, the method comprising administering to the subject a therapeutically effective amount of a composition comprising a SCFA1 polypeptide and a pharmaceutically acceptable carrier. In certain embodiments, the SCFA1 polypeptide comprises the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10, or a sequence of amino acids with at least 90% sequence identity to the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8 or 10. In additional embodiments, the disorder is mucositis, inflammatory bowel disease, or short bowel syndrome.

In yet further embodiments, the invention is directed to a method for treating a mammalian subject at risk for damage to the epithelial cells lining at least a portion of the oral or gastrointestinal tract, the method comprising administering to the subject a therapeutically effective amount of a composition comprising a SCFA1 polypeptide and a pharmaceutically acceptable carrier. In certain embodiments, the SCFA1 polypeptide comprises the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10, or a sequence of amino acids with at least 90% sequence identity to the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8 or 10. In additional embodiments, the subject has undergone or will undergo radiation therapy and/or chemotherapy.

In additional embodiments, the invention is directed to an adenoviral vector comprising a polynucleotide encoding a SCFA1 polypeptide operably associated with an expression control sequence. In certain embodiments, the SCFA1 polypeptide comprises the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10, or a sequence of amino acids with at least 90% sequence identity to the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8 or 10.

In further embodiments, the invention is directed to a pharmaceutical composition comprising the adenoviral vector described above and a pharmaceutically acceptable carrier.

In additional embodiments, the invention is directed to a method of stimulating epithelial cell proliferation in a subject in need thereof, the method comprising administering to the subject a therapeutically effective amount of the pharmaceutical composition comprising the adenoviral vector described above.

In yet further embodiments, the invention is directed to a method of stimulating epithelial cell proliferation in a subject in need thereof, the method comprising administering to the subject a therapeutically effective amount of a pharmaceutical composition comprising a vector encoding a SCFA1 polypeptide operably associated with an expression control sequence and a pharmaceutically acceptable carrier.

The methods of the invention also provide methods for the treatment of disorders as recited herein which comprise the administration of a therapeutically effective amount of a composition comprising a polynucleotide or polypeptide of the invention and a pharmaceutically acceptable carrier to a mammalian subject exhibiting symptoms or tendencies related to disorders as recited herein. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising the step of administering a composition comprising compounds and other substances that modulate the overall activity of the target gene products and a pharmaceutically acceptable carrier. Compounds and other substances can effect such modulation either on the level of target gene/protein expression or target protein activity. Specifically, methods are provided for preventing, treating or ameliorating a medical condition, including mucositis, inflammatory bowel disease, and wounds, which comprises administering to a mammalian subject, including but not limited to humans, a therapeutically effective amount of a composition comprising a polypeptide of the invention or a therapeutically effective amount of a composition comprising a binding partner of a SCFA1 polypeptide. The mechanics of the particular condition or pathology will dictate whether the polypeptides or binding partners thereof are beneficial to the individual in need of treatment.

The invention further provides methods for manufacturing medicaments useful in the above-described methods.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides as described herein in a sample (e.g., tissue or sample). Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions.

The invention provides a method for detecting a polypeptide as described herein in a sample comprising contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting formation of the complex, so that if a
5 complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in
10 clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can enhance the therapeutic
15 activity of the SCFA1 polypeptides, and ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention.

Another embodiment of the invention provides gene therapy by delivery of
20 SCFA1 polynucleotides encoding SCFA1 polypeptides for the treatment of conditions or disorders recited herein.

In a related embodiment, the invention is directed to the use of a vector comprising a gene encoding a SCFA1 polypeptide operably associated with an expression control sequence that provides for expression of the SCFA1 polypeptide
25 in the manufacture of a medicament for treating disorders as recited herein. The invention provides a virus vector comprising a gene encoding a SCFA1 polypeptide operably associated with an expression control sequence. In a preferred embodiment, the virus vector is an adenovirus vector. The virus vectors of the invention can provide a gene encoding any of the SCFA1 polypeptides, as set forth
30 above. More particularly, the invention provides for use of an adenoviral vector of the invention, in the manufacture of a medicament for treating mucositis, inflammatory bowel disease or short bowel syndrome. The related pharmaceutical

composition comprises the virus vector of the invention and a pharmaceutically acceptable carrier.

Additional aspects and advantages of the invention will be apparent to those skilled in the art upon consideration of the following description, which details the practice of the invention.

3. BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the H&E staining of cross-sections derived from the small intestine and colon derived from a control mouse (PBS) (Figure 1 a and c), and from a mouse treated with 1×10^{10} viral particles of adenovirus expressing SCFA1-V5His6 (SEQ ID NO: 6) (Figure 1 b and d), as described in Example 3.

Figure 2 shows the H&E staining of cross-sections derived from the small intestine and colon derived from a control mouse (PBS) (Figure 2 a and c), and from a mouse treated with purified recombinant SCFA1 protein (Figure 2 b and d).

Figure 3 shows the incorporation of BrdU into proliferating crypt cells of the small intestine and colon of control mice (PBS) (Figure 3 a and c) and of mice that had received purified recombinant SCFA1 protein (Figure 3 b and d). The sections shown were obtained from the same mice for which the H&E staining is described in Figure 2.

Figure 4 shows the dose-dependent stabilization of β -catenin by SCFA1 in HEK293 cells.

Figure 5 shows the effect of SCFA1 on β -catenin/TCF-mediated transcription.

4. DETAILED DESCRIPTION OF THE INVENTION

Stem cell factor-like protein A1, SCFA1, was previously described and shown to promote the proliferation and maintain the survival of hematopoietic stem cells (U.S. Patent 6,824,973). SCFA1 is a member of the thrombospondin type 1 repeat (TSR) superfamily (Chen et al., *Molecular Biol Report* 29: 287-292 (2002); Kamata et al., *Biochim Biophys Acta* 1676: 51-62 (2004); herein incorporated by reference), and is also known as hPWTSR (Chen et al., (2002)) and R-spondin 3 (Kazanskaya et al., *Developmental Cell* 7: 525-534 (2004); herein incorporated by reference).

The SCFA1 polypeptide of SEQ ID NO: 2 is an approximately 272 amino acid protein with a predicted molecular mass of approximately 30 kDa unglycosylated.

The initial methionine starts at position 291 of SEQ ID NO: 1 and the putative stop codon begins at position 1107 of SEQ ID NO: 1. Protein database searches with the BLAST algorithm (Altschul S.F. et al., *J. Mol. Evol.* 36:290-300 (1993) and Altschul S.F. et al., *J. Mol. Biol.* 21:403-10 (1990), herein incorporated by reference) indicate that an ortholog of SEQ ID NO: 1 is present in mouse (NCBI Accession No. NM028351; gi30794143), cow (NCBI Accession No. AV615874; gi9751544) and chicken (NCBI Accession No. AL587653; gi13192687), and a paralog of SEQ ID NO: 2 exists in zebrafish (NCBI Accession No. AI545208; gi4462581) and rat (NCBI Accession No. BE11437; gi8503542).

A predicted approximately twenty-one amino acid residue signal peptide is encoded from approximately residue 1 to residue 21 of SEQ ID NO: 2. The extracellular portion is useful on its own. The signal peptide region was predicted using the Neural Network Signal P VI.I program (Nielsen et al., *Int. J. Neural Syst.* 8:581-599 (1997)), incorporated herein by reference) and/or using Neural Network SignalP V1.1 program (Nielsen et al, (1997) *Int. J. Neural Syst.* 8, 581-599). SEQ ID NO: 8 is the SCFA1 polypeptide of SEQ ID NO: 2 that lacks the signal peptide i.e. SEQ ID NO: 8 is the dominant mature form of SCFA1. SCFA1 contains a predicted furin protease cleavage site between amino acids 32 and 33 of SEQ ID NO: 2. Therefore, it is possible that two forms of SCFA1 exist: a dominant mature form (SEQ ID NO: 8), which lacks the signal peptide, and a mature form (SEQ ID NO: 10), which lacks the signal peptide (amino acids 1 to 21) and amino acids residues between amino acid 22 and amino acid 32 of SEQ ID NO: 2.

Using the Pfam software program (Sonnhammer et al., *Nucleic Acids Res.*, Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) the SCFA1 polypeptide (SEQ ID NO: 2) was examined for domains with homology to known peptide domains. SCFA1 contains 2 furin-like domains. The furin-like domains of SCFA1 (SEQ ID NO: 12 and 14) are encoded by the polynucleotides of SEQ ID NO: 11 and 13, respectively, and the polypeptide domains span from amino acid 40 to 83, and from amino acid 94-133 of SEQ ID NO: 2. These furin-like repeat domains have been found in a variety of eukaryotic proteins involved in the mechanism of signal transduction by receptor tyrosine kinases (Raz et al. *Genetics* 129:191-201 (1991)). Thus, the furin-like cysteine-rich domains of SCFA1 may effect the proliferation of the intestinal epithelium.

SCFA1 polypeptide of SEQ ID NO: 2 is expected to have a thrombospondin type 1 domain (TSP1) (SEQ ID NO: 16) encoded by the nucleotide sequence of SEQ ID NO: 15). The thrombospondin type 1 domain contained within SEQ ID NO: 2 is predicted to be from amino acid residue 150-206.

5 Thrombospondins are a family of extracellular matrix proteins that are involved in cell-cell and cell-matrix communication (Lawler et al., *Curr. Opin. Cell Bio.* 12:634-640 (2000)). More than five different thrombospondins are known with distinct patterns of tissue distribution. Some tissues like heart, cartilage, and brain express most of the thrombospondin gene products. Thrombospondin-1 is a major
10 constituent of blood platelets. Thrombospondin-1 appears to function at the cell surface to bring together membrane proteins and cytokines and other soluble factors. Membrane proteins that bind thrombospondin-1 include integrins, integrin-associated protein (CD47), CD36, proteoglycans. Transforming growth factor β (TGF β) and platelet-derived growth factor also bind thrombospondin-1.

15 Thrombospondin-1 is a large protein with many distinct domains. It contains a globular domain at both amino and carboxy terminus, a region of homology with procollagen, and three types of repeated sequence motifs termed thrombospondin (TSP) type 1, type 2, and type 3 repeats. TSP1 repeat has been found in various different proteins including, complement components (C6, C7, C8A etc.)

20 extracellular matrix proteins like ADAMTS, mindin, axonal guidance molecule like F-spondin semaphorins, and also SCO-spondin, and TRAP proteins of Plasmodium.

 Thrombospondin type 1 repeat can activate TGF β epithelial tissues which are involved in regulation of cell growth, differentiation, adhesion, migration, and death. TSP1 is further involved in protein binding, heparin binding, cell attachment, neurite
25 outgrowth, inhibition of proliferation, inhibition of angiogenesis, and activation of apoptosis. TSP1 domains of *Plasmodium circumsporozoite* (CS) protein and TRAP proteins are implicated in salivary gland invasion by the sporozoite.

 TSP1 sequences are characterized by conserved cysteines, closely spaced tryptophans, and a cluster of basic residues. Spatial configuration of TSP1
30 sequences shows two β -sheet domains which are shown to bind heparin (Kilpelainen et al (2000) *J. Biol Chem.* 275, 13564-13570, incorporated herein by reference). A similar spatial fold has been described for heparin-binding growth associated molecule (HB-GAM). HB-GAM is identical to mitogenic and neurite

outgrowth-promoting protein pleiotrophin; osteoblast specific factor-1; heparin-binding neurotrophic factor; and heparin affin regulatory peptide. Expression of HB-GAM was shown to be associated with extracellular matrix of axonal tracts and synapses, and also with basement membranes outside of brain and in the cartilage matrix. Recently, N-syndecan has been shown to be receptor for HB-GAM in brain and has been suggested to play roles in regulation of hippocampal long-term potentiation, a form of brain plasticity implicated in memory and learning. Therefore, TSP1 containing proteins may act as growth promoters and may exhibit SCFA1 activities.

In addition, thrombospondin, synthesized in bone marrow and deposited within the extracellular matrix, functions as a cytoadhesion molecule for primary pluripotent progenitor cells, as well as for hematopoietic progenitor cells committed to erythroid, granulocytic, and megakaryocytic lineages. Thus thrombospondins may be important in blood cell development (Long and Dixit (1990) *Blood* 75, 2311-2318, incorporated herein by reference).

SCFA1 polypeptides and polynucleotides may be used to induce proliferation or differentiation of gastrointestinal crypt cells. They may also be used in the treatment of conditions where epithelialization is required, such as for the treatment of gastrointestinal disorders including chemotherapy and radiation therapy-induced mucositis, mucositis of the oropharynx, lips and esophagus, inflammatory bowel disease, and other conditions including wounds, burns, ophthalmic disorders, and any disorder where stimulation of epithelial cell proliferation or regeneration is desired. The polynucleotides and polypeptides may further be utilized to generate new tissues and organs that may aid patients in need of transplanted tissues.

4.1 DEFINITIONS

In describing the present invention the following terms will be employed and are intended to be defined as indicated below.

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

In accordance with the present invention, the term "SCFA1 protein(s) or "SCFA1 polypeptide(s) refers to the full-length protein defined by amino acids Met¹ to His²⁷² (SEQ ID NO: 2), fragments and analogs thereof.

The term "full-length SCFA1" or "native SCFA1" as used herein all refer to the polypeptide that contains 272 amino acid residues (SEQ ID NO: 2).

The term "fragment" refers to a polypeptide derived from the native SCFA1 that does not include the entire sequence of SCFA1. Such a fragment may be a truncated version of the full-length molecule, for example SEQ ID NO: 4 as well as internally deleted polypeptides. An SCFA1 fragment may have SCFA1 bioactivity as determined by the effect of SCFA1 on the proliferation of epithelial cells *in vitro* or *in vivo*, as described herein.

The term "analog" refers to derivatives of the reference molecule. The analog may retain biological activity, as described above. In general, the term "analog" refers to compounds having a native polypeptide sequence and structure with one or more amino acid additions, substitutions (generally conservative in nature) and/or deletions, relative to the native molecule, so long as the modifications do not destroy activity. Preferably, the analog has at least the same biological activity as the parent molecule, and may even display enhanced activity over the parent molecule.

Methods for making polypeptide analogs are known in the art. Particularly preferred analogs include substitutions that are conservative in nature, i.e., those substitutions that take place within a family of amino acids that are related in their side chains. Specifically, amino acids are generally divided into four families: (1) acidic: aspartate and glutamate; (2) basic: lysine, arginine, histidine; (3) non-polar: alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar: glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine.

Phenylalanine, tryptophan, and tyrosine are sometimes classified as aromatic amino acids. For example, it is reasonably predictable that an isolated replacement of leucine with isoleucine or valine; aspartate with glutamate; threonine with serine; or a similar conservative replacement of an amino acid with a structurally related amino acid will preserve the biological activity of SCFA1.

Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and

minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant analogs encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be

selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

5 The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as PEGylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

10 The terms "polypeptide" and "protein" refer to a polymer of amino acid residues and are not limited to a minimum length of the product. The terms also include, unless otherwise indicated, modifications of the polypeptide that do not change the sequence of amino acids, for example, glycosylated, acetylated and phosphorylated forms. A polypeptide or protein, for purposes of the present invention, may be synthetically or
15 recombinantly produced, as well as isolated from natural sources.

 The terms "purified" and "isolated" mean, when referring to a polypeptide or polynucleotide, that the indicated molecule is present in the substantial absence of other biological macromolecules of the same type. The term "purified" as used herein preferably means at least 75% by weight, more preferably at least 85% by
20 weight, more preferably still at least 95% by weight, and most preferably at least 98% by weight, of biological macromolecules of the same type are present in the sample. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight of the indicated biological macromolecules present but water, buffers, and other small molecules, especially molecules having
25 a molecular weight of less than 1000 daltons, can be present.

 An "isolated polynucleotide which encodes a particular polypeptide" refers to a nucleic acid molecule which is substantially free of other nucleic acid molecules that do not encode the subject polypeptide; however, the molecule may include some additional bases or moieties which do not deleteriously affect the basic characteristics
30 of the composition.

 The term "naturally occurring polypeptide" means polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide

including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or a processing sequence.

The term "dominant mature protein coding sequence" means a polynucleotide sequence which encodes a peptide or protein without any leader/signal sequence. The "dominant mature protein portion" refers to the portion of the protein without the leader/signal sequence. The "mature" form refers to a SCFA1 polypeptide that lacks the leader/signal sequence and the sequence to the furin cleavage site. The peptide may have the leader sequence and/or the furin cleavage site removed during processing in the cell or the protein may have been produced synthetically or using a polynucleotide only encoding for the mature protein coding sequence. It is contemplated that the mature or dominant mature protein portion may or may not include an initial methionine residue. The initial methionine is often removed during processing of the peptide.

The term "isolated" means a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other components normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., *E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant polypeptide" means a polypeptide, which has been prepared by recombinant DNA techniques as described herein. In general, the gene coding for the desired polypeptide is cloned and then expressed in transformed organisms, as described below. The host organism expresses the foreign gene to produce the polypeptide under expression conditions. Alternatively, the promoter controlling expression of an endogenous polypeptide can be altered to render a recombinant polypeptide.

The term "active" means those forms of the polypeptide that retain the biologic and/or immunologic activities of any naturally occurring polypeptide.

According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "biologically active" or "biological activity" refers to the capability of the natural, recombinant or synthetic SCFA1, or any peptide thereof, to induce a specific biological response in appropriate animals or cells.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) *Cytokine* 4(2):134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) *Annu. Rev. Immunol.* 16:27-55)

The terms "polynucleotide" or "nucleic acid molecule" mean a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule and thus includes double- and single-stranded DNA and RNA. It also includes known types of modifications, for example, labels which are known in the art, methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., methyl

phosphonates, phosphotriesters, phosphoamidates, carbamates, etc.) and with charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, for example proteins (including for e.g., nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), those with
5 intercalators (e.g., acridine, psoralen, etc.), those containing chelates (e.g., metals, radioactive metals, boron, oxidative metals, etc.), those containing alkylators, those with modified linkages (e.g., alpha anomeric nucleic acids, etc.), as well as unmodified forms of the polynucleotide. Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short
10 oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment",
15 "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably
20 less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides.
25 Preferably the fragments can be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to a portion of SEQ ID NO:
30 SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15.

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, *PCR*

Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from any of the nucleic acid sequences of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15. The sequence information can be a segment of SEQ ID NO SEQ ID NO1, 3, 5, 7, 9, 11, 13 or 15 that uniquely identifies or represents the sequence information of SEQ ID NO: SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4^{20} possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match ($1 \div 4^{25}$) times the increased probability for mismatch at each nucleotide position (3×25). The probability that an eighteen-mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" mean functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The terms "recombinant DNA molecule," or "recombinant polynucleotide" mean a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation: (1) is not associated with all or a portion of a polynucleotide with which it is associated in nature, (2) is linked to a polynucleotide other than that to which it is linked in nature, or (3) does not occur in nature. Thus, the term encompasses "synthetically derived" nucleic acid molecules.

The terms "complementary" or "complementarity" mean the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "stringent" means conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base

oligonucleotides), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

The phrase "substantially equivalent," for both nucleotide and amino acid sequences, means a sequence that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (*i.e.*, the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, *e.g.*, mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more than 5% (95% sequence identity). Substantially equivalent, *e.g.*, mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 90% sequence identity. Substantially equivalent nucleotide sequence of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, and most preferably at least about 95% identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (*e.g.*, via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, *e.g.*, using the Jotun Hein method (Hein, J. (1990) *Methods Enzymol.* 183:626-645). Identity between sequences can also be

determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "vector" means a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. The term "expression vector" includes plasmids, cosmids or phages capable of synthesizing the SCFA1 protein encoded by the respective recombinant gene carried by the vector. Preferred vectors are those capable of autonomous replication and expression of nucleic acids to which they are linked.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration.

The term "transfection" means taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

The term "transcriptional regulatory elements" and transcriptional regulatory sequences" are used interchangeably to refer to DNA sequences necessary for the expression of an operably linked coding sequence in a particular organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, enhancers, splicing signals and polyadenylation signals. These terms are intended to encompass all elements that promote or regulate transcription, including promoters, core elements required for basic interaction of RNA polymerase and transcription factors, upstream elements, enhancers, and response elements (Lewin, "Genes V" (Oxford University Press, Oxford) pages 847-873).

A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then optionally trans-RNA spliced and translated into the protein encoded by the coding sequence.

The term "expression modulating fragment," EMF, means a series of nucleotides that modulates the expression of an operably linked ORF or another EMF.

A sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs is nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The term "recombinant expression vehicle or vector" means a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A

pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "treat" or "treatment" mean both therapeutic and prophylactic or preventative measures, wherein the object is to prevent or lessen an undesired physiological change or condition, such as chemotherapy or radiation therapy-induced mucositis. For the purposes of this invention, beneficial or desired clinical results include, but are not limited to alleviation of symptoms, diminishment of extent of the disease, stabilized state of the disease, whether detectable or undetectable.

The term "disorder" means any condition that would benefit from treatment with a molecule identified using the transgenic animal model of the invention. This includes chronic and acute disorders or diseases including those pathological conditions which predispose the mammal to the disorder in question. Non-limiting examples of disorders to be treated herein include mucositis, inflammatory bowel disease and skin lesions. A preferred disorder to be treated in accordance with the present invention is mucositis.

The term "inflammatory bowel disease (IBD)" means idiopathic or chronic inflammatory disease of either or both the small intestine and large bowel, and includes Crohn's disease, ulcerative colitis, IBD caused by infectious agents, and antibiotic associated IBD.

The term "mucositis" means inflammation of the mucous membranes of the alimentary tract including the oropharynx and lips, esophagus, and large and small intestine.

The term "Short Bowel Syndrome" or "SBS" means a condition of nutritional malabsorption resulting from anatomical or functional loss of a significant length of the small intestine.

The terms "effective amount" or "pharmaceutically effective amount" mean a nontoxic but sufficient amount of the agent to provide the desired biological result. That result can be reduction and/or alleviation of the signs, symptoms, or causes of a disease, or any other desired alteration of a biological system. For example, an effective amount of a SCFA1 fragment for use with the present methods is an amount sufficient to stimulate epithelial cell stimulation or proliferation, and preferably an amount sufficient to cause increased regeneration of the gastrointestinal epithelium in a subject suffering from chemotherapy or radiation therapy-induced mucositis,

inflammatory bowel disease, or other disorders where epithelial cell proliferation is desired. Such amounts are described below. An appropriate "effective" amount in any individual case may be determined by one of ordinary skill in the art using routine experimentation.

5 The term "pharmaceutically acceptable" or "pharmacologically acceptable" means a material which is not biologically or otherwise undesirable, i.e., the material may be administered to an individual without causing any undesirable biological effects or interacting in a deleterious manner with any of the components of the composition in which it is contained.

10 The term "physiological pH" or a "pH in the physiological range" means a pH in the range of approximately 7.0 to 8.0 inclusive. Preferred physiological pH is in the range of approximately 7.2 to 7.6 inclusive.

 The term "non-human mammal" means all members of the class Mammalia except humans. "Mammal" refers to any animal classified as a mammal, including
15 humans, non-human primates such as chimpanzees, and other apes and monkey species; farm animals such as cattle, horses, sheep, goats, swine; domestic animals such as rabbits, dogs, and cats; laboratory animals including rodents, such as rats, mice and guinea pigs, and the like. Examples of non-mammals include, but are not limited to, birds, fish and the like.

20 The term "subject" encompasses mammals and non-mammals. The term does not denote a particular age or gender.

4.2 COMPOSITIONS OF THE INVENTION

4.2.1 NUCLEIC ACID COMPOSITIONS

25 The invention is based on the discovery that compositions comprising the epithelial cell growth factor polypeptide, SCFA1, and the polynucleotides encoding the SCFA1 polypeptide stimulate the growth and proliferation of intestinal epithelial cells including crypt cells. Therefore, the use of these compositions for the diagnosis and treatment of conditions wherein stimulation of epithelial cell
30 proliferation or regeneration is desired, is contemplated.

 The isolated polynucleotides of the invention include, but are not limited to a polynucleotide comprising any of the nucleotide sequences of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15; a polynucleotide comprising the full length protein coding

sequence of SEQ ID NO: 1; (for example coding for SEQ ID NO: 2); and a polynucleotide comprising the nucleotide sequence encoding the protein coding sequence of the polypeptides of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14 or 16. The polynucleotides of the present invention also include, but are not limited to, a
5 polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15 (b) a polynucleotide encoding any one of the polypeptides of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14 or 16 (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog of any of the
10 proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14 or 16. Domains of interest include catalytic and substrate binding domains.

The polynucleotides of the invention include naturally occurring or wholly or
15 partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include the entire coding region of the cDNA or may represent a portion of the coding region of the cDNA.

The present invention also provides compositions comprising genes corresponding to the cDNA sequences disclosed herein. The corresponding genes
20 can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequences can be obtained using methods known in the art. For example, full
25 length cDNA or genomic DNA that corresponds to any of the polynucleotide of SEQ ID NO: 2 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15 may be used as the basis
30 for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The polynucleotides for use herein also include polynucleotides with nucleotide sequences that are substantially equivalent to the polynucleotides recited

above. Polynucleotides according to the invention can have, e.g., at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, or 89%, more typically at least about 90%, 91%, 92%, 93%, or 94% and even more typically at least about 95%, 96%, 97%, 98% or 99% sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that are selective for (i.e. specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor result for the nucleic acids of the present invention can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altschul, S.F. *J Mol. Evol.* 36 290-300 (1993) and Altschul S.F. et al. *J. Mol. Biol.* 21:403-410 (1990)).

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encodes proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode analogs of the described nucleic acids. These amino acid sequence analogs may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence analogs are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, *e.g.*, by substituting first with conservative choices (*e.g.*, hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (*e.g.*, hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the amino acid sequences for use with the invention are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., *supra*, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising SEQ ID NO: 3 or the dominant mature or mature protein coding sequences, coding for any one of SEQ ID NO: 4, 6, 8, 10, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15 or a fragment thereof or any other SCFA1 polynucleotide. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 or 15 or a fragment thereof is inserted. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for

example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, and pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate

phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

In addition to the use of expression vectors in the practice of the present invention, the present invention further includes novel expression vectors comprising promoter elements operatively linked to polynucleotide sequences encoding a protein of interest.

4.2.2 HOSTS

The present invention further provides host cells genetically engineered with the vectors of this invention, which may be, for example, a cloning vector or an expression vector that contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The vector may be, for example, in the form of a plasmid, a viral particle, a phage etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying SCFA1 genes. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the SCFA1 polypeptides. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs

derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by
5 reference.

Various mammalian cell culture systems can be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell* 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example,
10 the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin,
15 early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as
20 necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

25 A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937,
30 HaK or Jurkat cells. Preferably, SCFA1 proteins are expressed in Chinese Hamster Ovary (CHO) cells, and human embryonic kidney 293 cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains

include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, *Pichia pastoris* or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial strain capable of
5 expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

4.2.3 CHIMERIC AND FUSION PROTEINS

The invention also provides SCFA1 chimeric or fusion proteins. As used herein, a SCFA1 "chimeric protein" or "fusion protein" comprises a SCFA1 polypeptide operatively-linked to a non-SCFA1 polypeptide. A "SCFA1
15 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a SCFA1 protein, whereas a "non-SCFA1 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the SCFA1 protein, e.g., a protein that is different from the SCFA1 protein and that is derived from the same or a different organism. Within a SCFA1
20 fusion protein the SCFA1 polypeptide can correspond to all or a portion of a SCFA1 protein. In one embodiment, a SCFA1 fusion protein comprises at least one biologically active portion of a SCFA1 protein. In another embodiment, a SCFA1 fusion protein comprises at least two biologically active portions of a SCFA1 protein. In yet another embodiment, a SCFA1 fusion protein comprises at least three
25 biologically active portions of a SCFA1 protein. Within the fusion protein, the term "operatively-linked" is intended to indicate that the SCFA1 polypeptide and the non-SCFA1 polypeptide are fused in-frame with one another. The non-SCFA1 polypeptide can be fused to the N-terminus or C-terminus of the SCFA1 polypeptide.

30 In one embodiment, the fusion protein is a GST SCFA1 fusion protein in which the SCFA1 sequences are fused to the C-terminus of the GST (glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant SCFA1 polypeptides. Preferably, the SCFA1 polypeptide is fused with

a V5-His tag for easy detection with an anti-V5 antibody and for rapid purification as described in the examples. SEQ ID NO: 6, which is encoded by the polynucleotide of SEQ ID NO: 5, represents SCFA1 fusion protein that has a V5-His6 tag.

5 In another embodiment, the fusion protein is a SCFA1 protein containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of SCFA1 can be increased through use of a heterologous signal sequence. For example the signal sequence of SCF1 can be replaced with the signal sequence from the VJ2-C region of the mouse IgKappa-chain. For example SEQ ID NO: 6, which is encoded by the
10 polynucleotide of SEQ ID NO: 5, contains an IgK signal sequence. The polynucleotide of SEQ ID NO: 5 was used as described in the examples to express SCFA1 (SEQ ID NO: 6) *in vivo* and determine the biological activity of the polypeptides. A SCFA1 chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for
15 the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene
20 can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, e.g., Ausubel, et al. (eds.) *Current*
25 *Protocols In Molecular Biology*, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide).

4.2.4 POLYPEPTIDE COMPOSITIONS

30 The pharmaceutical compositions of the invention comprise isolated SCFA1 polypeptides that include, but are not limited to, a polypeptide comprising: the amino acid sequence set forth as any one of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, or 16, or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:

1, 3, 5, 7, 9, 11, 13, or 15. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, or 15; or (b) polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, or 15 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, or 16; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, or 89%, more typically at least about 90%, 91%, 92%, 93%, or 94% and even more typically at least about 95%, 96%, 97%, 98% or 99%, most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, or 16.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., *Bio/Technology* 10, 773-778 (1992) and in R. S. McDowell, et al., *J. Amer. Chem. Soc.* 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides full-length, dominant mature forms (for example, without a signal sequence or precursor sequence) or mature forms (for example, lacking the signal sequence and the furin cleavage site) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. The term "degenerate variant" means nucleotide fragments which differ from a nucleic acid fragment of the present invention (*e.g.*, an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetic protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide

in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells, which are transformed with SCFA1-encoding DNA to produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, et al., in *Molecular Cloning: A Laboratory Manual*; Ausubel et al., *Current Protocols in Molecular Biology*. Polypeptide fragments that retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for e.g., small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the

peptide or DNA sequence can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues
5 may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are
10 important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the
15 protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological methodologies may also be easily made by those skilled in the
20 art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and
25 methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBat™ kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of
30 the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture

medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-
5 toyopearl™ or Cibacrom blue 3GA Sepharose™; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion
10 protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific
15 antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP- HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the
20 protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include SCFA1 analogs. This embraces
25 fragments of SCFA1 polypeptide, as well as SCFA1 polypeptides which comprise one or more amino acids deleted, inserted, or substituted. Also, analogs of the SCFA1 polypeptide of the invention embrace fusions of the SCFA1 polypeptides or modifications of the SCFA1 polypeptides, wherein the SCFA1 polypeptide or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic
30 agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the SCFA1 polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to the small intestine, e.g., antibodies to the small intestine, or

antibodies to receptor and ligands expressed on gastrointestinal cells. Other moieties which may be fused to SCFA1 polypeptide include therapeutic agents which are used for treatment, for example cytokines or other medications, of gastrointestinal disorders, and other conditions as recited herein.

5

4.2.5 GENE THERAPY

The invention provides gene therapy to treat the diseases cited herein. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, *Nature*, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, *Science*, 244: 1275-1281 (1989); Verma, *Scientific American*: 68-84 (1990); and Miller, *Nature*, 357: 455-460 (1992).

As discussed above, a "vector" is any means for the transfer of a nucleic acid according to the invention into a host cell. Preferred vectors are viral vectors, such as retroviruses, herpes viruses, adenoviruses and adeno-associated viruses. Thus, a gene or nucleic acid sequence encoding a SCFA1 protein or polypeptide domain fragment thereof is introduced *in vivo*, *ex vivo*, or *in vitro* using a viral vector or through direct introduction of DNA. Expression in targeted tissues can be effected by targeting the transgenic vector to specific cells, such as with a viral vector or a receptor ligand, or by using a tissue-specific promoter, or both.

Viral vectors commonly used for *in vivo* or *ex vivo* targeting and therapy procedures are DNA-based vectors and retroviral vectors. Methods for constructing and using viral vectors are known in the art [see, e.g., Miller and Rosman, *BioTechniques* 7:980-990 (1992)]. Preferably, the viral vectors are replication defective, that is, they are unable to replicate autonomously in the target cell. In general, the genome of the replication defective viral vectors which are used within the scope of the present invention lack at least one region which is necessary for the replication of the virus in the infected cell. These regions can either be eliminated (in whole or in part), be rendered non-functional by any technique known to a person skilled in the art. These techniques include the total removal,

substitution (by other sequences, in particular by the inserted nucleic acid), partial deletion or addition of one or more bases to an essential (for replication) region. Such techniques may be performed *in vitro* (on the isolated DNA) or *in situ*, using the techniques of genetic manipulation or by treatment with mutagenic agents.

- 5 Preferably, the replication defective virus retains the sequences of its genome which are necessary for encapsulating the viral particles.

DNA viral vectors include an attenuated or defective DNA virus, such as but not limited to herpes simplex virus (HSV), papillomavirus, Epstein-Barr virus (EBV), adenovirus, adeno-associated virus (AAV), and the like. Defective viruses, which
10 entirely or almost entirely lack viral genes, are preferred. Defective virus is not infective after introduction into a cell. Use of defective viral vectors allows for administration to cells in a specific, localized area, without concern that the vector can infect other cells. Thus, a specific tissue can be specifically targeted. Examples of particular vectors include, but are not limited to, a defective herpes virus 1 (HSV1) vector [Kaplitt et al., *Molec. Cell. Neurosci.* 2:320-330 (1991)], defective herpes virus
15 vector lacking a glyco-protein L gene [Patent Publication RD 371005 A], or other defective herpes virus vectors [International Patent Publication No. WO 94/21807, published Sep. 29, 1994; International Patent Publication No. WO 92/05263, published Apr. 2, 1994]; an attenuated adenovirus vector, such as the vector
20 described by Stratford-Perricaudet et al. [*J. Clin. Invest.* 90:626-630 (1992); see also La Salle et al., *Science* 259:988-990 (1993)]; and a defective adeno-associated virus vector [Samulski et al., *J. Virol.* 61:3096-3101 (1987); Samulski et al., *J. Virol.* 63:3822-3828 (1989); Lebkowski et al., *Mol. Cell. Biol.* 8:3988-3996 (1988)].

Preferably, for *in vivo* administration, an appropriate immunosuppressive
25 treatment is employed in conjunction with the viral vector, e.g., adenovirus vector, to avoid immuno-deactivation of the viral vector and transfected cells. For example, immunosuppressive cytokines, such as interleukin-12 (IL-12), interferon- γ (IFN- γ), or anti-CD4 antibody, can be administered to block humoral or cellular immune responses to the viral vectors [see, e.g., Wilson, *Nature Medicine* (1995)]. In
30 addition, it is advantageous to employ a viral vector that is engineered to express a minimal number of antigens.

In a preferred embodiment, the vector is an adenovirus vector. As shown in the Examples, the adenovirus vector has shown itself to be particularly effective for

delivery of the SCFA1 polypeptide, as shown by the unexpectedly efficient effects of stimulating intestinal epithelial cell proliferation resulting in marked, diffuse thickening of the mucosa by crypt epithelial hyperplasia and a marked increase in crypt length and complex branching. Adenoviruses are eukaryotic DNA viruses that can be modified to efficiently deliver a nucleic acid of the invention to a variety of cell types. Various serotypes of adenovirus exist. Of these serotypes, preference is given, within the scope of the present invention, to using type 2 or type 5 human adenoviruses (Ad 2 or Ad 5) or adenoviruses of animal origin (see WO94/26914). Those adenoviruses of animal origin which can be used within the scope of the present invention include adenoviruses of canine, bovine, murine (example: Mav1, Beard et al., *Virology* 75 (1990) 81), ovine, porcine, avian, and simian (example: SAV) origin.

Preferably, the replication defective adenoviral vectors of the invention comprise the ITRs, an encapsidation sequence and the nucleic acid of interest. Still more preferably, at least the E1 region of the adenoviral vector is non-functional. Other regions may also be modified, in particular the E3 region (WO95/02697), the E2 region (WO94/28938), the E4 region (WO94/28152, WO94/12649 and WO95/02697), or in any of the late genes L1-L5.

In a preferred embodiment, the adenoviral vector has a deletion in the E1 and E3 region. Examples of E1-deleted adenoviruses are disclosed in EP 185,573, the contents of which are incorporated herein by reference.

The replication defective recombinant adenoviruses according to the invention can be prepared by any technique known to the person skilled in the art (Levrero et al., *Gene* 101 (1991) 195, EP 185 573; Graham, *EMBO J.* 3 (1984) 2917). In particular, they can be prepared by homologous recombination between an adenovirus and a plasmid, which carries, *inter alia*, the DNA sequence of interest. The homologous recombination is effected following cotransfection of the said adenovirus and plasmid into an appropriate cell line. The cell line which is employed should preferably (i) be transformable by the said elements, and (ii) contain the sequences which are able to complement the part of the genome of the replication defective adenovirus, preferably in integrated form in order to avoid the risks of recombination. Examples of cell lines which may be used are the human embryonic kidney cell line 293 (Graham et al., *J. Gen. Virol.* 36 (1977) 59) which

contains the left-hand portion of the genome of an Ad5 adenovirus (12%) integrated into its genome, and cell lines which are able to complement the E1 and E4 functions, as described in applications WO94/26914 and WO95/02697.

Recombinant adenoviruses are recovered and purified using standard molecular biological techniques, which are well known to one of ordinary skill in the art.

Promoters that may be used in the present invention include both constitutive promoters and regulated (inducible) promoters. The promoter may be naturally responsible for the expression of the nucleic acid. It may also be from a heterologous source. In particular, it may be promoter sequences of eukaryotic or viral genes. For example, it may be promoter sequences derived from the genome of the cell which it is desired to infect. Likewise, it may be promoter sequences derived from the genome of a virus, including the adenovirus used. In this regard, there may be mentioned, for example, the promoters of the E1A, MLP, CMV and RSV genes and the like.

In addition, the promoter may be modified by addition of activating or regulatory sequences or sequences allowing a tissue-specific or predominant expression (enolase and GFAP promoters and the like). Moreover, when the nucleic acid does not contain promoter sequences, it may be inserted, such as into the virus genome downstream of such a sequence.

Some promoters useful for practice of this invention are ubiquitous promoters (e.g., HPRT, vimentin, actin, tubulin), intermediate filament promoters (e.g., desmin, neurofilaments, keratin, GFAP), therapeutic gene promoters (e.g., MDR type, CFTR, factor VIII), tissue-specific promoters (e.g., actin promoter in smooth muscle cells), promoters which are preferentially activated in dividing cells, promoters which respond to a stimulus (e.g., steroid hormone receptor, retinoic acid receptor), tetracycline-regulated transcriptional modulators, cytomegalovirus immediate-early, retroviral LTR, metallothionein, SV-40, E1a, and MLP promoters. Tetracycline-regulated transcriptional modulators and CMV promoters are described in WO 96/01313, U.S. Pat. Nos. 5,168,062 and 5,385,839, the contents of which are incorporated herein by reference.

Thus, the promoters which may be used to control gene expression include, but are not limited to, the cytomegalovirus (CMV) promoter, the SV40 early promoter region (Benoist and Chambon, 1981, *Nature* 290:304-310), the promoter

contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto, et al., 1980, *Cell* 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, *Proc. Natl. Acad. Sci. U.S.A.* 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster et al., 1982, *Nature* 296:39-42); prokaryotic

5 expression vectors such as the b-lactamase promoter (Villa-Kamaroff, et al., 1978, *Proc. Natl. Acad. Sci. U.S.A.* 75:3727-3731), or the tac promoter (DeBoer, et al., 1983, *Proc. Natl. Acad. Sci. U.S.A.* 80:21-25); see also "Useful proteins from recombinant bacteria" in *Scientific American*, 1980, 242:74-94; promoter elements from yeast or other fungi such as the Gal 4 promoter, the ADC (alcohol

10 dehydrogenase) promoter, PGK (phosphoglycerol kinase) promoter, alkaline phosphatase promoter; and the animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: elastase I gene control region which is active in pancreatic acinar cells (Swift et al., 1984, *Cell* 38:639-646; Ornitz et al., 1986, *Cold Spring Harbor Symp. Quant. Biol.* 50:399-409;

15 MacDonald, 1987, *Hepatology* 7:425-515); insulin gene control region which is active in pancreatic beta cells (Hanahan, 1985, *Nature* 315:115-122), immunoglobulin gene control region which is active in lymphoid cells (Grosschedl et al., 1984, *Cell* 38:647-658; Adames et al., 1985, *Nature* 318:533-538; Alexander et al., 1987, *Mol. Cell. Biol.* 7:1436-1444), mouse mammary tumor virus control region

20 which is active in testicular, breast, lymphoid and mast cells (Leder et al., 1986, *Cell* 45:485-495), albumin gene control region which is active in liver (Pinkert et al., 1987, *Genes and Devel.* 1:268-276), alpha-fetoprotein gene control region which is active in liver (Krumlauf et al., 1985, *Mol. Cell. Biol.* 5:1639-1648; Hammer et al., 1987, *Science* 235:53-58), alpha 1-antitrypsin gene control region which is active in

25 the liver (Kelsey et al., 1987, *Genes and Devel.* 1:161-171), beta-globin gene control region which is active in myeloid cells (Mogam et al., 1985, *Nature* 315:338-340; Kollias et al., 1986, *Cell* 46:89-94), myelin basic protein gene control region which is active in oligodendrocyte cells in the brain (Readhead et al., 1987, *Cell* 48:703-712), myosin light chain-2 gene control region which is active in skeletal muscle (Sani,

30 1985, *Nature* 314:283-286), and gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason et al., 1986, *Science* 234:1372-1378).

Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes. In addition to the use of viral vectors in the practice of the present invention, the present invention further provides a vector, the pAdenoVator-CMV5-Intron vector, comprising operator and promoter elements operatively linked to polynucleotide sequences encoding a protein of interest and is described in detail in Examples.

4.2.6 CRYPT CELL AND TISSUE GROWTH ACTIVITY

The SCFA1 polypeptide of the invention exhibits growth factor activity and is involved in the proliferation and differentiation of intestinal crypt cells. SCFA1 may also exhibit growth factor activity on other epithelial cells of the gastrointestinal tract. Administration of the polypeptide of the invention to crypt cells *in vivo* or *ex vivo* may maintain and expand cell populations in a totipotent state which may be useful for re-engineering damaged or diseased tissues, transplantation, and manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat tissues for grafting such gastrointestinal cells.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors, basic fibroblast growth factor (bFGF), keratinocyte growth factor-2 (KGF2), and glucagons-like peptide 2 (GLP-2).

Intestinal epithelial cells including crypt cells can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated cell lines that are useful as is or that can then be differentiated into the desired mature cell types.

5 These stable cell lines can also serve as a source of undifferentiated mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in crypt cell populations that regulate crypt proliferation and/or maintenance.

10 Expansion and maintenance of epithelial stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate crypt cells in culture to give rise to gastrointestinal epithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders
15 inflammation caused by ionizing radiation, chemotherapy, infection and inflammation.

Expression of the polypeptide of the invention and its effect on crypt cells can also be manipulated to achieve controlled differentiation of the crypt cells into more differentiated cell types. A broadly applicable method of obtaining pure populations
20 of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker.

In vitro cultures of intestinal epithelial cells including crypt cells can be used to determine if the polypeptide of the invention exhibits growth factor activity. Crypt cells are isolated from disaggregated colonic crypts from human and murine colonic
25 mucosa, and the clonogenic activity of SCFA1 can be assessed using the method described by Whitehead et al., *Gastroenterology* 117:858-865 (1999), which is herein incorporated by reference in its entirety. Growth factor activity may be assayed in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines.

30 The compositions of the present invention may also be useful for proliferation of intestinal epithelial cells including crypt cells and for regeneration of oral and gastrointestinal tissue, i.e. for the treatment of injuries sustained by the epithelial layer which involve degeneration, death or trauma to epithelial crypt cells. More

specifically, a composition may be used in the treatment of diseases of the gastrointestinal tract as recited herein.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like. Assays for wound healing activity include, without limitation, those described in: Winter, *Epidermal Wound Healing*, pp. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, *J. Invest. Dermatol* 71:382-84 (1978).

4.2.7 IMMUNOMODULATORY ACTIVITY

A polypeptide of the present invention may exhibit activity relating to regulation of immune system components including, but not limited to cytokine production and/or activity, and/or cells of the immune system. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Regulation of cytokines and/or cells of the immune system may include increasing and/or decreasing levels of cytokines or numbers of particular cells of the immune system.

With such immunomodulatory activity, polypeptides of the invention may be used to treat various immune disorders. These disorders include, but are not limited to inflammatory bowel disease (IBD), which includes ulcerative colitis and/or Crohn's disease, and mucositis as a consequence of anti-cancer therapies including radiation treatment and/or chemotherapy. The cause of these immune disorders may be, for example, idiopathic (i.e. of unknown cause), genetic, by infectious agents (e.g. viruses, bacteria, fungi), and/or by damage induced by anti-cancer therapies (e.g. radiation therapy and/or chemotherapy).

Modulation of immune responses and/or components of the immune system may be accomplished in a number of ways. Down-regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process that requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or

anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of tolerizing agent.

5 Inflammatory bowel disease is almost always mediated by one of two pathways: excessive T helper 1 (Th1)-cell response associated with high levels of IL-12, IFN-gamma, and/or TNF or excessive T helper 2 (Th2)-cell response associated with high levels of IL-4, IL-5, and/or IL-13 (Bouma *et al.*, herein incorporated by reference in its entirety). Therefore a mechanism through which
10 polypeptides of the invention could mediate immunomodulatory activity in disease treatment would be to down-regulate the numbers of Th1 and/or Th2 cell populations. Alternatively, another activity could be to decrease the levels of cytokines (e.g. IL-12, IFN-gamma, TNF, IL-4, IL-5, and/or IL-13) that are associated with and/or mediate the inflammatory response.

15 The activity of the polypeptide of the present invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing
20 Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai *et al.*, *J. Immunol.* 137:3494-3500, 1986; Bertagnolli *et al.*, *J. Immunol.* 145:1706-1712, 1990; Bertagnolli *et al.*, *Cellular Immunology* 133:327-341, 1991; Bertagnolli, *et al.*, *J. Immunol.* 149:3778-3783, 1992; Bowman *et al.*, *J. Immunol.* 152:1756-
25 1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In *Current Protocols in Immunology*. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons,
30 Toronto. 1994; and Measurement of mouse and human interferon- γ , Schreiber, R. D. In *Current Protocols in Immunology*. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M.

- 5 Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., *Proc. Natl. Acad. Sci. USA* 77:6091-6095, 1980; Weinberger et al., *Eur. J. Immun.* 11:405-411, 1981; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Takai et al., *J. Immunol.* 140:508-512, 1988.)
- 10

4.2.8 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against a tumor or an infecting agent.

15

20

25 A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

30

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce

the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margules, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. *J. Clin. Invest.* 95:1370-1376, 1995; Lind et al. *APMIS* 103:140-146, 1995; Muller et al. *Eur. J. Immunol.* 25:1744-1748; Gruber et al. *J. Immunol.* 152:5860-5867, 1994; Johnston et al. *J. Immunol.* 153:1762-1768, 1994.

4.2.9 DRUG SCREENING

Screening for a useful compound involves administering the candidate compound over a range of doses to a non-human animal, and assaying at various time points for the effect(s) of the compound on the activity of the SCFA1 protein. The compound may be administered prior to or at the onset of abdominal distension. Administration may be oral, or by suitable injection, depending on the chemical nature of the compound being evaluated. The cellular response to the compound is evaluated over time using appropriate biochemical and /or histological assays.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other
5 libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol.*, 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol*,
10 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem*, 4(5):709-15 (1996) (alkylated dipeptides).

4.3 DISEASES AMENABLE TO SCFA1 THERAPY

In one aspect, the present invention provides pharmaceutical reagents and
15 methods useful for treating diseases and conditions wherein epithelialization is desired. SCFA1 polypeptides are useful to increase cytoprotection, proliferation and/or differentiation of epithelial cells of the oral and gastrointestinal tract. Specifically, SCFA1 polypeptides are useful to treat or prevent diseases or conditions that include without limitation gastrointestinal diseases, mucositis of the
20 gastrointestinal tract, mucositis of the oropharynx, lips and esophagus (oral mucositis), inflammatory bowel disease, short bowel syndrome, gastric and duodenal ulcers, erosions of the gastrointestinal tract including erosive gastritis, esophagitis, esophageal reflux and other conditions including wounds, burns, ophthalmic disorders, and any disorder where stimulation of epithelial cell proliferation or regeneration is desired.
25 Treatment of diseases that result in insufficient production of mucus throughout the oral and gastrointestinal tract is also contemplated.

Mucositis, which includes oral and gastrointestinal mucositis, is a complication of some cancer therapies in which the lining of the digestive system becomes inflamed. SCFA1 is useful for preventing and/or ameliorating the degeneration of the mucosa of the
30 alimentary tract that is caused by chemotherapy and /or radiation therapy given to a patient for the treatment of cancer, or is given as an adjuvant therapy following the removal of a tumor. Exemplary chemotherapeutic agents include, without limitation, BCNU, busulfan, carboplatin, cyclophosphamide, tannorubicin, doxorubicin, etoposide, 5-

fluorouracil, gemcytabine, ifophamide, irinotecan, melphalan, methotrexate, navelbine, totpotecan, and taxol, and exemplary treatment regimens include without limitation, BEAM (busulfan, etoposide, cytosine, arabinoside, methotrexate); cyclophosphamide and total body irradiation; cyclophosphamine, total body irradiation and etoposide;

5 cyclophosphamide and busulfan; and 5-fluorouracil with leucovorin or levamisole.

Treatment, pretreatment or post-treatment with SCFA1 is useful to generate a cytoprotective effect or regeneration or both, for example, of the mucosa of the small intestine and colon, allowing increased dosages of therapies while reducing their potential side effects.

10 Inflammatory bowel disease that can be treated with SCFA1 includes general inflammatory bowel disease that is characterized by chronic, relapsing, inflammatory disorders of unknown origin, Crohn's disease, dysplasia associated with inflammatory bowel disease, intermediate colitis, ulcerative colitis; non-infectious colitis including active colitis, antibiotic-associated colitis, collagenous colitis, 15 diversion colitis, eosinophilic colitis, graft versus host disease, granulomatous colitis, ischaemic colitis, hemorrhagic colitis, malakoplakia, necrotizing enterocolitis, radiation enterocolitis, typhlitis; infectious colitis including adenovirus and amebic colitis, balantidiasis, HSV/AIDS associated colitis, and colitis caused by trypanosomes, *E. coli*, *Mycobacterium avium intracellulare*, Sotavirus, Salmonella, 20 Shigella, *Campylobacter jejuni*, Clostridium, Botulinum, and colitis associated with schistosomiasis, spirochetosis, syphilis, trichuriasis, tuberculosis typhoid fever, *Vibrio cholera*, and Yersinia.

Short bowel syndrome is a group of problems affecting people who have had half or more of their small intestine removed. The most common reason for 25 removing part of the small intestine is to treat Crohn's disease. In addition, surgical resection of part of the intestine may be required to remove cancerous growths. Diarrhea is the main symptom of short bowel syndrome. Other symptoms include cramping, bloating, and heartburn. Many people with short bowel syndrome are malnourished because their remaining small intestine is unable to absorb enough 30 water, vitamins, and other nutrients from food. They may also become dehydrated, which can be life threatening. Problems associated with dehydration and malnutrition include weakness, fatigue, depression, weight loss, bacterial infections, and food sensitivities. Short bowel syndrome is treated through changes in diet,

intravenous feeding, vitamin and mineral supplements, and medicine to relieve symptoms. SCFA1 polypeptides may be useful to increase the proliferation of the unresected intestinal tissue, thereby increasing the absorptive surface area of the intestine, and ameliorate the symptoms associated with short bowel syndrome.

5 The cytoprotective and/or regenerative activity of SCFA1 polypeptides can be tested in *in vivo* models of radiation induced mucositis (Withers and Elkind, *Int J Radiat* 17:261-267 (1970), herein incorporated by reference; in *in vivo* chemotherapy-induced mucositis (Sorix et al., *Oral Surg Oral Med Oral Pathol* 69:437-443 (1990); Moore, *Cancer Chemother Pharmacol* 15:11-15 (1985); Farell et al., *Cell Prolif* 35:78-85 (2002), all of which are incorporated by reference in their
10 entirety); in a dextran sulfate sodium (DSS) model of colitis and small intestinal ulceration or inflammation (Jeffers et al., *Gastroenterology* 123:1151-1162 (2002), Han et al., *Am J Physiol Gastrointest Liver Physiol* 279:G1011-G1022 (2000); and in a surgical model of short bowel syndrome (SBS) (Scott et al. *Am J Physiol* G911-
15 G921 (1998); Helmrath et al., *J Am Coll Surg* 183:441-449 (1996)), herein incorporated by reference in their entirety).

 Comparisons of SCFA1 mRNA and protein expression levels between diseased cells, tissue and corresponding normal samples are made to determine if the subject is responsive to SCFA1 therapy. Methods for detecting and quantifying
20 the expression of SCFA1 polypeptide mRNA or protein use standard nucleic acid and protein detection and quantitation techniques that are well known in the art and are described in Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, NY (1989) or Ausubel, et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY (1989), both of which are
25 incorporated herein by reference in their entirety. Standard methods for the detection and quantification of SCFA1 mRNA include *in situ* hybridization using labeled SCFA1 riboprobes (Gemou-Engesaeth, et al., *Pediatrics* 109: E24-E32 (2002), herein incorporated by reference in its entirety), Northern blot and related techniques using SCFA1 polynucleotide probes (Kunzli, et al., *Cancer* 94: 228
30 (2002), herein incorporated by reference in its entirety , herein incorporated by reference in its entirety), RT-PCR analysis using SCFA1-specific primers (Angchaiskisir, et al., *Blood* 99:130 (2002)), and other amplification detection methods, such as branched chain DNA solution hybridization assay (Jardi, et al., *J.*

Viral Hepat. 8:465-471 (2001), herein incorporated by reference in its entirety), transcription-mediated amplification (Kimura, et al., *J. Clin. Microbiol.* 40:439-445 (2002)), microarray products, such as oligos, cDNAs, and monoclonal antibodies, and real-time PCR (Simpson, et al., *Molec. Vision*, 6:178-183 (2000), herein
5 incorporated by reference in its entirety). Standard methods for the detection and quantification of SCFA1 protein include western blot analysis (Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, NY (1989), Ausubel, et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY (1989)), immunocytochemistry (Racila, et al., *Proc. Natl. Acad. Sci.*
10 *USA* 95:4589-4594 (1998) *supra*), and a variety of immunoassays, including enzyme-linked immunosorbant assay (ELISA), radioimmuno assay (RIA), and specific enzyme immunoassay (EIA) (Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, NY (1989), Ausubel, et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY (1989)).

15 The diseases and conditions treatable by methods of the present invention preferably occur in mammals. Mammals include, for example, humans and other primates, as well as pet or companion animals such as dogs and cats, laboratory animals such as rats, mice and rabbits, and farm animals such as horses, pigs, sheep, and cattle.

20 4.3.1 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of
25 therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

One embodiment of the invention is the administration of an effective amount of SCFA1 polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be treated the peptides of the invention. While the
30 mode of administration is not particularly important, parenteral administration is preferred. Exemplary modes of administration are to deliver a subcutaneous or intravenous bolus. The dosage of SCFA1 polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be

expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01 μ g/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1 μ g/kg to 10 mg/kg of patient body weight. For parenteral administration, SCFA1 polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.3.2 PHARMACEUTICAL FORMULATIONS

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors and various growth factors such as any of the FGFs, epidermal growth factor (EGF), platelet-derived growth factor (PDGF), transforming growth factors (TGF- α and TGF- β), insulin-like growth factor (IGF), keratinocyte growth factor (KGF), and the like, as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or

other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredients of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.3.3 ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous (iv), intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal (IP), parenteral or intravenous injection. Subcutaneous or intravenous administration to the patient is preferred.

Alternatively, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into the tissue, often in a depot or sustained release formulation.

In another embodiment, the implantation of cells producing SCFA1 (cell therapy) into a subject in need of proliferation and/or stimulation of epithelial cells is contemplated. Cells that do not normally express SCFA1 or that express low levels of SCFA1 may be modified to produce therapeutic levels of SCFA1 by transformation with a polynucleotide that encodes SCFA1. The cells may be of the same species as the subject, or may be derived from a different species. Preferably, the cells are derived from the subject in need of SCFA1 therapy. Human or nonhuman cells may be implanted in a subject using a biocompatible, semi-permeable polymeric enclosure to allow release of SCFA1 protein, or may be implanted directly without encapsulation.

In another embodiment, *in vivo* gene therapy is contemplated. A nucleotide sequence encoding SCFA1 is introduced directly into a subject for secretion of the protein to prevent or treat the diseases as recited herein. The nucleotide encoding SCFA1 may be injected directly into the tissue to be treated, or it may be delivered
5 into the cells of the affected tissue by a viral vector e.g. adenovirus vector or retrovirus vector. Physical transfer of appropriate vectors containing a SCFA1-encoding nucleic acid may also be achieved by methods including liposome-mediated transfer, direct injection of naked DNA, receptor-mediated transfer, or microparticle bombardment.

10 The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the
15 polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

4.3.4 COMPOSITIONS/FORMULATIONS

20 Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a
25 manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other
30 active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95%

protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries,

suspensions and the like, for oral ingestion by a patient to be treated.

Pharmaceutical preparations for oral use can be obtained solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations, which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to

deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes.

Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents, which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of

the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the

patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μg to about 100 mg (preferably about 0.1 μg to about 10 mg, more preferably about 0.1 μg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention, which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as

polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredient of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- α and TGF- β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins or other active ingredient of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the

condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

4.3.5 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate *in vitro* assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC₅₀ as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-

maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient.

5 Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the
10 ratio between LD₅₀ and ED₅₀. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage may vary within this range
15 depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety
20 which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

25 Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

30 An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 µg/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 µg/kg to 25 mg/kg of patient body weight

daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction,
5 the manner of administration and the judgment of the prescribing physician.

4.3.6 DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test
10 sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the
15 complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the
20 invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the
25 sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

30 Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the

commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers,

5 Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention

10 include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed.

Methods for preparing protein extracts or membrane extracts of cells are well known
15 in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention.

Specifically, the invention provides a compartment kit to receive, in close

20 confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained
25 in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allow one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to
30 another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe.

Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats that are well known in the art.

4.3.7 SCREENING ASSAYS

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying modulatory agents which bind to a polypeptide encoded by an ORF corresponding to the nucleotide sequence set forth in SEQ ID NO: 1, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
- (b) determining whether the agent binds to said protein or said nucleic acid.

The modulatory agents may increase or decrease the proliferative activity of SCFA1 on epithelial cells.

In general, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the

complex drives expression of a target gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

5 Compounds identified via such methods can include compounds that modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds that modulate the expression of a polynucleotide of the invention (that is, increase or
10 decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

 The agents screened in the above assay can be, but are not limited to,
15 peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

 For random screening, agents such as peptides, carbohydrates,
20 pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention.

 Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical
25 agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., "Application of Synthetic Peptides: Antisense Peptides," In *Synthetic Peptides, A User's Guide*, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., *Biochemistry* 28:9230-8 (1989), or pharmaceutical agents, or the like.

30 In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF

allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by
5 binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods usually contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription
10 (triple helix - see Lee et al., *Nucl. Acids Res.* 6:3073 (1979); Cooney et al., *Science* 241:456 (1988); and Dervan et al., *Science* 251:1360 (1991)) or to the mRNA itself (antisense - Okano, *J. Neurochem.* 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA,
15 while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

20 Agents that bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents that bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

25 5. EXAMPLES

EXAMPLE 1

CLONING AND EXPRESSION OF SCFA1

In order to express SCFA1 polypeptide, the full-length SCFA1 DNA (SEQ ID NO: 1) was PCR amplified from a cDNA library that was constructed using human
30 mRNA from cerebellum (Ambion). The sequence was also isolated from pools of cDNAs that were derived from human mRNAs from various tissues.

The first round of PCR was performed using the forward primer (SEQ ID NO: 17) and reverse primer (SEQ ID NO: 18). A second round of PCR was then

performed using the primary PCR as template using the forward primer SEQ ID NO: 19 and reverse primer of SEQ ID NO: 20. The restriction sites NheI and XbaI, which are embedded in the forward and reverse primers (SEQ ID NO: 21 and 22) were used for subcloning the NheI-XbaI fragment into the mammalian expression vector plntron/Igk. The NheI-XbaI fragment contains polynucleotide sequence (SEQ ID NO: 3), which encodes amino acids 31 to 272 sequence (SEQ ID NO: 4) of the full-length SCFA1 (SEQ ID NO: 2). The mammalian expression vector plntron/IgK was constructed by introducing an engineered chimeric intron from the mammalian pCI expression vector (Promega, Madison, WI), and the IgK leader sequence derived from the pSectag vector into pcDNA3.1V5His vector (Invitrogen, Inc., Carlsbad, CA).

EXAMPLE 2

THE ADENOVIRAL VECTOR

The polynucleotide of SEQ ID NO: 5 was amplified from the plntron/Igk vector together with the Igk leader sequence and the V5His6 tag of the plntron/IgK vector, and cloned into the pAdenovator-CMVIntron adenoviral vector (SEQ ID NO:27) as follows. The polynucleotide sequence (SEQ ID NO: 5), which encodes plntron-SCFA1-V5His6 (SEQ ID NO: 6) was amplified from the plntron/IgK using the forward primer of SEQ ID NO: 23 and the reverse primer of SEQ ID NO: 24. The restriction enzyme sites XbaI and NotI that are contained in the primers were used to clone SEQ ID NO: 5 into the NheI and NotI sites of the pAdenovator-CMVIntron vector (SEQ ID NO: 27).

The adenoviral vector pAdenovator-CMVIntron was obtained by modifying the pAdenoVatorCMV5-IRES-GFP (Qbiogene, Carlsbad, CA, U.S. as follows. pAdenoVator-CMV5-IRES-GFP was digested with SpeI to remove its MCS, IRES and GFP and ligated with PCR amplified Intron-MCS-V5His-BGH polyA from pcDNA/Intron vector using forward primer (SEQ ID NO: 25) and reverse primer (SEQ ID NO:26).

Transformation of linearized transfer vector into bacterial cells, BJ5183, (Qbiogene, Carlsbad, CA, U.S.A) which carries an AdEasy-1 plasmid that encodes Adenovirus-5 genome (E1/E3 deleted) was performed by electroporation according to the manufacturer's instructions. Recombinant adenovirus was generated and amplified in QBI-293A cells (Qbiogene, Carlsbad, CA, U.S.A) and purified by CsCl

banding as previously described (Garnier, A., J. Cote et al. 1994). Recombinant protein expression by 293A cells that had been infected with the recombinant adenovirus was measured by Western analysis using anti-V5 antibody (Invitrogen Inc., Carlsbad, CA). The titer of CsCl purified recombinant viruses was measured using the Adeno-X rapid titer kit (BD biosciences, Palo Alto, U.S.A.) according to the manufacturer's protocols. Briefly, a viral stock was tested by infecting 293A cells with serial dilutions of the recombinant adenovirus stock followed by fixation and staining of the transduced cells with mouse anti-hexon antibody 48 hours after infection. The signal was detected with a goat anti-mouse antibody conjugated to horseradish peroxidase and developed with metal-enhanced 3,3'-diaminobenzidine tetrahydro-chloride (DAB).

EXAMPLE 3

ADMINISTRATION OF RECOMBINANT ADENOVIRUS AS A MODEL TO EVALUATE THE BIOLOGICAL ACTIVITY OF SCFA1

The SCFA1 recombinant adenovirus was administered to normal mice to determine the effect SCFA1 on the intestinal and colonic epithelium. Prior to injection of adenovirus, BALB/c mice, 9-11 weeks of age, were anesthetized using isoflurane. 1×10^{10} viral particles per mouse were injected via the retro-orbital vein. The same titer of control virus (empty virus) was used in control animals. Three mice were used in each experimental group, and were sacrificed 3 days after receiving the virus injection.

4 hours before sacrifice, 1 mg of bromodeoxyuridine (BrdU) was injected intraperitoneally (IP) to determine *in vivo* proliferation of epithelial cells. Various tissues including small intestine, colon, spleen, liver and bone marrow were collected and fixed in formaline.

Paraffin embedded sections were stained with hematoxyline and eosin (H&E) for histological evaluation. Sections were also processed for BrdU immunohistochemistry according to the manufacturer's instruction (Oncogene Research product, Boston, U.S.A.) as previously described (McKinley, J. N. et al. 2000).

H&E staining of sections from the mid-jejunum of mice that had been sacrificed 3 days following the adenovirus injection showed that the small intestine of mice that had received the SCFA1 adenovirus (Figure 1b) was significantly

altered when compared to that of control mice (Figure 1a). The histological changes caused by SCFA1 included a marked elongation of the mucosal crypts (Figure 1b), when compared to the crypts from control animals (Figure 1a). Similarly, SCFA1 also caused a marked increase in the of Goblet cells in the colon from mice to which
5 SCFA1 had been administered (Figure 1d) when compared to control animals (Figure 1c).

EXAMPLE 4

EXPRESSION AND PURIFICATION OF RECOMBINANT SCFA1

10 SCFA1 (SEQ ID NO: 6) was expressed in HEK293 as follows.

A stable cell culture of HEK293 cells that had been transfected with the SCFA1 pcDNA/Intron construct comprising the DNA encoding the V5-His-tagged SCFA1 polypeptide (SEQ ID NO: 5) was grown in serum free 293 free-style media (GIBCO). A suspension culture was seeded at cell density of 1×10^6 cells/ml, and
15 harvested after 4-6 days. The level of the V5-His-tagged SCFA1 that had been secreted into the culture medium was assayed by ELISA.

The media containing the secreted SCFA1 protein was harvested and frozen at -80°C . The media was thawed at 4°C , and protease inhibitors, EDTA and Pefabloc (Roche, Basel, Switzerland) were added at a final concentration of 1mM
20 each to prevent degradation of SCFA1. The media were filtered through a $0.22 \mu\text{m}$ PES filter (Corning), and concentrated 10-fold using TFF system (Pall Filtron) with a 10 kDa molecular weight cut-off membrane. The buffers of the concentrated media were exchanged with 20 mM sodium phosphate, 0.5M NaCl, pH 7. The addition of 0.5 M NaCl in the phosphate buffer is crucial to keep full solubility of V5-His tagged
25 SCFA1 at pH 7 during purification. Following ultrafiltration and diafiltration, a mammalian protease inhibitor cocktail (Sigma) was added to a final dilution of 1:500 (v/v).

A HiTrap Ni^{2+} -chelating affinity column (Pharmacia) was equilibrated with 20 mM sodium phosphate, pH 7, 0.5 M NaCl. The buffer-exchanged media was filtered
30 with $0.22 \mu\text{m}$ PES filter and loaded onto Ni^{2+} -chelating affinity column. The Ni^{2+} column was washed with 10 column volumes (CV) of 20 mM imidazole for 10 column volumes and protein was eluted with a gradient of 20 mM to 300 mM imidazole over 35 CV. The fractions were analyzed by SDS-PAGE and Western

blot. Fractions containing V5-His tagged SCFA1 were analyzed and pooled to yield a SCFA1 protein solution that was between 75-80% pure.

The buffer containing the SCFA1 protein isolated using the Ni^{2+} column was exchanged with 20 mM sodium phosphate, 0.3 M Arginine, pH 7 to remove the NaCl. NaCl was replaced with 0.3 M Arginine in the phosphate buffer to maintain full solubility of V5-His tagged SCFA1 protein during the subsequent purification steps. The SCFA1 protein isolated using the Ni^{2+} column was loaded onto a SP Sepharose high performance cation exchange column (Pharmacia, Piscataway, NJ) that had been equilibrated with 20 mM sodium phosphate, 0.3 M Arginine, pH 7.

The column was washed with 0.1 M NaCl for 8 CV, and eluted with a gradient of 0.1 M to 1 M NaCl over 30 CV. Fractions containing V5-His tagged SCFA1 were pooled to yield a protein solution that was between 90-95 % pure.

The buffer of the pooled fractions was exchanged with 20mM sodium phosphate, pH 7, 0.15 M NaCl, the protein was concentrated to 1 or 2 mg/mL, and passed through a sterile 0.22 μm filter. The pure SCFA1 preparation was stored at -80°C .

In summary, the purification of V5-His-tagged SCFA1 from cultures of HEK293 included: 1) concentrating and diafiltering the SCFA1 protein present in the culture media, 2) performing Ni^{2+} -chelating affinity chromatography, and 3) SP cation exchange chromatography. The purification process yields a SCFA1 protein that is > 90 % pure. The overall recovery of the current purification process is approximately 50%. Addition of 0.5 M NaCl to the buffer during the purification process of media diafiltration and Ni column is crucial to keep SCFA1 fully soluble at pH 7. For binding SCFA1 onto the SP column, NaCl was removed, and 0.3 M Arginine was added to maintain high solubility and increase protein recovery. The addition of 0.5 M NaCl and 0.3 Arginine during the first and second purification steps showed to increase the overall recovery by at least from 25% to 50%.

EXAMPLE 5

IN VIVO BIOLOGICAL TESTING OF RECOMBINANT SCFA1 PROTEIN EXPRESSED IN HEK293

The *in vivo* biological effects of SCFA1 protein that was derived from HEK293 was evaluated in normal mice as follows.

To investigate whether purified recombinant SCFA1 protein could generate a phenotype similar to that observed in the mice that had been injected with recombinant adenovirus, 6-8 weeks old BALB/c mice were injected daily through tail vein with 100 μ g SCFA1 protein or formulation buffer as control for 3 days. On the fourth day, and two hours prior to being sacrificed, 4mg of bromodeoxyuridine (BrdU) was injected ip to determine the *in vivo* proliferative activity of SCFA1. Paraffin embedded sections of the small intestine and colon were stained with hematoxylin and eosin for histological evaluation. Sections were also processed for BrdU immunohistochemistry according to the manufacturer's instruction (Oncogene Research product, Boston, U.S.A.) and previously described (McKinley, J. N. et al. 2000). In all experiments, at least 3 animals were analyzed per group and experiments were repeated at least twice.

H&E staining of gastrointestinal sections showed a marked distention of mucosal crypts in the mid-jejunum of the small intestine of mice that had received recombinant human SCFA1 protein (Figure 2b), while normal intestinal morphology was observed in the saline control mice (Figure 2a). A significantly higher number of goblet cells was noted in the colon of the experimental mice (Figure 2d), when compared to the saline control animals (Figure 2c).

The proliferative effect of SCFA1 protein was confirmed by assaying BrdU incorporation in both small intestine and colon (Figure 3). BrdU immunohistochemistry revealed a marked proliferation of intestinal crypt epithelial cells in the small intestine (Figure 3b) and colon (Figure 3d) was seen in the mice that had received SCFA1, when compared to that of the saline control animals (Figure 3 a and c, respectively). These results are consistent with those obtained in the mice that received SCFA1 adenovirus (see Example above).

EXAMPLE 6

EFFECT OF SCFA1 ON INTRACELLULAR SIGNALING

The wnt/ β -catenin signaling pathway plays a pivotal role in development and homeostasis. In the small intestine, wnt signaling is known to play a critical role as a regulator of intestinal crypt proliferation by stabilizing β -catenin, which subsequently induces the transactivation of T-cell factor (TCF) target genes (Wetters et al., *Cell*

111:241-250 (2002); Batle et al., *Cell*, 111:251-263, (2002); Perreault et al., *J Biol Chem* 276:43328-43333 (2001); Booth et al., *Nat Med* 8:1360-1361 (2002)).

To evaluate the effect of SCFA1 on the wnt/ β -catenin signaling pathway, the stabilization of β -catenin was measured in various cultured cell lines. Cells were seeded at 1 million cells/well for a 6-well plate in Dulbecco's modified Eagle's medium supplemented with 10 % FBS. The following day, cells were grown in serum-free medium, and treated either with SCFA1 at indicated concentration or LiCl₂ (positive control) at 10 mM in low serum conditions (0.1% FBS) for 3 hours. Cytoplasmic fractions were prepared as described by Haertel-Weismann et al., (*J Biol Chem* 175:32046-32051 (2000)). The proteins were resolved by gradient (4-20%) SDS-PAGE, and the level of β -catenin was assessed using a β -catenin rabbit antibody (Abcam) that was visualized using a peroxidase conjugated secondary antibody (Cell Signaling). Among tested cell lines, SCFA1 induced the stabilization of β -catenin in HEK 293 cells in a dose-dependent manner (Figure 4).

The functional activation of β -catenin signaling by SCFA1, was also tested using the TCF reporter assay. The 16TCF-luciferase reporter construct was generated as described by DasGupta et al. (*Science* 308: 826-833 (2005)). A mutant form of 16TCF (m16TCF) was also generated and used as a control. HEK293 cells were seeded at 13,000 cells per well and transfected with either 17TCF-luciferase reporter construct or the corresponding mutant at 0.07 μ g per well. On following day, cells were serum starved for 8 hours and treated with SCFA1 or Wnt3A (R&D System) as a positive control at indicated concentrations. After 18 hours incubation at 37°C, CO₂ incubator, luciferase activity was measured by using Bright Glo substrate (Promega). As shown in Figure 5, SCFA1 activated TCF-mediated transcription to a level comparable to the activation obtained with Wnt3A protein. This result is consistent with the effect of SCFA1 on the stabilization of β -catenin described above, and suggests that SCFA1 functions to activate β -catenin signaling.

EXAMPLE 7

PROPHYLACTIC EFFECT OF SCFA1 ON RADIATION-INDUCED MUCOSITIS

The efficacy of SCFA1 as a prophylactic and therapeutic agent is tested in an animal model of radiation-induced mucositis.

Forty eight adult male BDF1 mice, aged 10-12 weeks, are used. On delivery from the supplier and prior to the experiment, the animals are housed for two weeks in individually ventilated cages on a 12 hour light:dark cycle to stabilize the circadian rhythm. Animals are allowed food and water *ad libitum*.

- 5 The animals are divided into 8 groups of 6 animals, and treated as follows:
1. Injected with 2 mg/kg SCFA1 iv at 72, 48, and 24 hours prior to being exposed to 13 Gy X-ray (whole body);
 2. Injected with 5 mg/kg SCFA1 iv at 72, 48, and 24 hours prior to being exposed to 13 Gy X-ray (whole body);
 - 10 3. Injected with 125 µg KGF iv at 72, 48, and 24 hours prior to being exposed to 13 Gy X-ray (whole body);
 4. Injected with saline vehicle iv at 72, 48, and 24 hours prior to being exposed to 13 Gy X-ray (whole body);
 5. Untreated, non-irradiated controls;
 - 15 6. Injected with 2 mg/kg SCFA1 iv 24, 48, and 72 hours post irradiation with 13 Gy X-rays (whole body);
 7. Injected with 5 mg/kg SCFA1 iv 24, 48, and 72 hours post irradiation with 13 Gy X-rays (whole body);
 8. Injected with saline iv at 24, 48, and 72 hours post irradiation with 13 Gy X-rays (whole body).
 - 20

All injections are given at the same time of day. Intestinal damage is induced using a single dose of 13 Gy X-irradiation (delivered at 0.7 Gy/min) at 15:00.

Four days after irradiation the animals are sacrificed. The small intestine is removed and fixed in Carnoy's fixative prior to processing for histological analysis.

- 25 Transverse sections 3 µm thick are cut and stained with haematoxylin and eosin. Immediately after sacrifice the duodenum, mid colon, liver, lung, tongue, spleen, stomach and pancreas are also removed and fixed in formal saline overnight prior to storage in 70% ethanol.

- 30 For each animal, ten intestinal circumferences are analyzed (60 per group) – a circumference is equivalent to a given length of intestine and therefore a convenient baseline unit of length. The number of surviving crypts per circumference is scored and the average per group determined.

The average crypt width (measured at its widest point) is also measured in order to correct for scoring errors due to crypt size difference. The correction is applied thus:

Corrected number of crypts / circumference=

$$5 \quad \frac{\text{Mean crypt width in untreated control}}{\text{Mean crypt width in treated animal}} \times \text{Mean number of surviving crypts in treatment group}$$

EXAMPLE 8

CHEMOTHERAPY-INDUCED MUCOSITIS

10 The efficacy of the recombinant human SCFA1 in treating chemotherapy-induced mucositis is evaluated in healthy and in tumor-bearing mice. The experimental protocol is based on that previously described by Boushey et al. (*Cancer Res* 61:687-693 (2001)).

15 One million CT26 murine colon carcinoma cells (ATCC, Manassas, VA, USA) are injected sc into syngeneic female BALB/c mice, and the tumors are allowed to develop for 5 days. Healthy and tumor-bearing animals are divided into experimental groups of 6 mice each and treated as follows:

1. Tumor bearing mice, vehicle (50% DMSO) injected ip from day 1 to day 5, saline injected iv from day 0 to day 7 (TVS)
- 20 2. Tumor bearing mice, vehicle (50% DMSO) injected ip from day 1 to day 5, 50 μ g SCFA1 in 100 μ l saline injected iv daily from day 0 to day 7 (TVG);
3. Tumor bearing mice, 50 mg/kg 5-FU injected ip from day 1 to day 5, saline injected iv from day 0 to day 7 (TDS);
- 25 4. Tumor bearing mice, 50 mg/kg 5-FU injected ip from day 1 to day 5, 50 μ g SCFA1 in 100 μ l saline injected iv from day 0 to day 7 (TDG);
5. Healthy mice, 50 mg/kg 5-FU injected ip from day 1 to day 5, saline injected iv from day 0 to day 7 (NDS); and
6. Healthy mice, 50 mg/kg 5-FU injected ip from day 1 to day 5, 50 μ g SCFA1 in 100 μ l saline injected iv from day 0 to day 7 (NDG).
- 30

On days 0, 2, 4, 6, and 8 measurements of animal body weight, severity of diarrhea, and size of the tumors are recorded. A diarrhea score of 0-3 reflects a corresponding worsening of the symptom from 0 being normal to 3 being severe.

The change in body weight is calculated as the percent body weight of that of the untreated group. The length, width and height of the tumor are measured with calipers, and the volume of the tumor is calculated as $(\text{length} \times \text{width} \times \text{height})/2$.

All animals are euthanized on day 8. The large and small intestine are removed and weighed, their length is measured, and the diameter of the mid-jejunum is recorded. A segment (1 cm) of the mid-jejunum is excised about 14-15 cm from the pylorus, and a segment (1 cm) of the transverse colon is excised at about 4 cm from the ileocaecal junction. The bowel segments are flushed and fixed using 10% neutral buffered formalin for histological analysis. Histological examination and morphometry of the mucosa are performed on tissue sections using the ImagePro Software (Imagepro, Ltd., Ashford, Middlesex, UK).

EXAMPLE 9

PROPHYLACTIC EFFECT OF SCFA1 ON CHEMOTHERAPY AND RADIATION-INDUCED ORAL MUCOSITIS

The effect of SCFA1 on the proliferation of the dorsal (buccal) and ventral epithelium of the tongue is studied in mice that are subjected to X-ray irradiation or dosed with 5-FU, as described in Examples 7 and 8, respectively.

Immunohistochemistry using monoclonal rat anti-mouse Ki67 antigen (Dako Ltd., High Wycombe, UK) is performed, according to manufacturer's instruction and the method previously described (Scholzen, T. et al. 2000), on paraffin embedded sections of tongue from non-irradiated and irradiated mice.

The epithelial proliferative index, which is calculated as the percent epithelial cells that stain positive for Ki67, is measured and is used to determine the amount of cellularity that is lost or gained caused by radiation to the ventral tongue epithelium with or without SCFA1 treatment.

Quantitative animal models of oral mucositis (e.g. Wardly et al., *Arch Oral Biol* 43:567-577 (1998); Potten et al., *Cell Prolif* 35:32-47 (2002)) can be used to study further the therapeutic properties of SCFA1, when administered in combination with other cytotoxic agent to further assess the potential role of SCFA1 in reducing the severity of the cellular depletion and to increase the rate of regeneration of the epithelial layers of the oral and intestinal epithelium.

EXAMPLE 10

THERAPEUTIC EFFECT OF SCFA1 ON DEXTRAN SULFATE SODIUM-INDUCED COLITIS

The efficacy of recombinant human SCFA1 (in treating colitis is tested in a mouse model of dextran sulfate sodium (DSS)-induced colitis, and compared to that the efficacy of GLP-2 (L'Heureux and Brubaker *J Pharmacol Exp Ther* 306:347-354 (2003); Kriegelstein et al., *J Clin Invest* 110:1773-1782 (2002); Siegmund et al., *J Pharmacol Exp Ther* 296:99-105 (2001)).

Six to eight-week old female BALB/c mice (Charles River Laboratories, Wilmington, MA, USA) are housed in ventilated cages and acclimated for one week to a 12 hour light:dark cycle. Twenty four mice having similar body weight (approximately 20 g; <5% variance) are housed in 4 cages and fed *ad libitum* a 4% DSS (v/w) drinking solution for 7 days.

On day 7, the body weight of each animal is recorded, and the scores for loss in body weight, the consistency of stools, and anal bleeding are determined as shown in the Table below.

TABLE 1

SCORE	Weight Loss (%)	Stool Consistency	Occult/Gross Rectal bleeding
0	None	Normal	Normal
1	0-5%		
2	5-10%	Loose	Hemoccult
3	10-20%		
4	>20%	Diarrhea	Gross

The scores are used to calculate the IBD activity index (IBDAI), which is used as an indicator of the severity of the colitis, and is calculated as the average of the scores given for the tabulated parameters. The scores for weight loss, stool consistency, and rectal bleeding are determined daily, and the IBDAI is recorded daily for the duration of the experiment.

On day 7, the 4% (v/w) DSS drinking solution is substituted with a 1% (v/w) DSS solution to maintain the disease activity without exacerbating the effect of the DSS. Sixteen of the DSS-fed animals are selected for consistent and comparable disease activity, and are divided into groups of 4 animals and are treated as follows:

1. Water, saline injected iv daily (10 am) for 7 days;
2. DSS (1%) for 7 days, saline injected iv daily (10 am) for 7 days;
3. DSS (1%) for 7 days, 100 μ g SCFA1 injected daily iv (10 am) for 7 days;
4. DSS (1%) for 7 days, 50 μ g SCFA1 injected daily iv (10 am) for 7 days; and
5. DSS (1%) for 7 days, 10 μ g GLP-2 injected sc twice daily (10 am and 6 pm) for 7 days.

On day 14, food is removed from the cages to allow for purging of the intestine, and the animals are sacrificed by cervical dislocation. All animals are injected with 4 mg/0.1ml BrdU two hours prior to sacrifice. The large and small intestine are removed and weighed, their lengths are measured, and the diameter of the mid-jejunum is recorded. A segment (1 cm) of the mid-jejunum is excised about 14-15 cm from the pylorus, and a segment (1 cm) of the transverse colon is excised at about 4 cm from the ileocaecal junction. The bowel segments are flushed and fixed using 10% neutral buffered formalin for histological analysis. Histological examination and morphometry of the mucosa is performed on tissue sections using the ImagePro Software (Imagepro, Ltd., Ashford, Middlesex, UK).

EXAMPLE 11

THERAPEUTIC EFFECT OF SCFA1 FOLLOWING MASSIVE INTESTINAL RESECTION

The effect of SCFA1 in augmenting the adaptive response to massive intestinal resection is tested in a rat animal model of short bowel syndrome. The animal model used in the study of the effects of enterotrophic agents has been described (Scott et al. *Am J Physiol* G911-G921 (1998); Helmrath et al., *J Am Coll Surg* 183:441-449 (1996)), and the experimental protocol is herein incorporated by reference).

The animals are divided into a resected group that will have a 75% surgical resection of the midjejunum, a sham-resected operated control group in which the intestine is sectioned and reanastomosed, and an unoperated control group. The animals are administered saline or SCFA1 at a dose of 2 mg/kg. The 75% intestinal resection is chosen to maximize any adaptive response and retention of equal portions of the proximal jejunum and distal ileum is based on the nutritional implications of removing the specialized absorptive capacity of the terminal ileum for

vitamin B12 and bile acids and the ileal brake. In the rat, the retention of 25% of the small intestine inclusive of a portion of distal ileum is sufficient to allow resected animals to achieve the same growth rate as control animals.

The morphological and functional response of the gut to resection and treatment with SCFA1 is assessed at 6, 14, and 21 days. Food intake and growth, gross and microscopic small intestinal morphology, and functional evaluation of mucosal absorptive characteristics are evaluated as described (Scott et al., *supra*).

EXAMPLE 12

EFFECT OF SCFA1 ON TNBS-INDUCED COLITIS

The hapten agent 2,4,6-trinitrobenzenesulfonic acid (TNBS) induces a chronic colitis that is characterized by severe, transmural inflammation associated with diarrhea, rectal prolapse, and weight loss. These clinical and histopathological features indicate that TNBS-induced colitis mimics important characteristics of human Crohn's disease (Neurath et al., *J Exp Med* 182:1281-1290 (1995)).

The therapeutic effect of SCFA1 is tested in mice with TNBS-induced colitis. Intestinal inflammation is induced in 6-8 week-old female BALBc mice by a single rectal administration of 1 mg TNBS, as described by Neurath et al, *supra*). A control animal group receives rectal administration of vehicle alone (45% ethanol). The mice are sacrificed after 7 days, and the induction of colitis by TNBS is assessed. Histologic changes are evaluated in H&E stained paraffin-embedded sections of the colon from the control and TNBS groups.

The therapeutic effect of SCFA1 is tested in the TNBS-treated animals by administering daily doses of up to 4 mg/Kg (100 µg/mouse; iv), beginning at day 3 following administration of TNBS. Animals from each group are sacrificed at day 7 or day 10. The tissues are removed, and histological evaluation, morphometric analysis, and proliferative and apoptotic indices are determined, as described in the examples above.

WE CLAIM:

1. A method of stimulating epithelial cell proliferation in a subject in need
5 thereof, said method comprising administering to said subject a therapeutically
effective amount of a composition comprising a stem cell factor-like protein A1
(SCFA1) polypeptide and a pharmaceutically acceptable carrier.
2. The method of claim 1, wherein the SCFA1 polypeptide comprises the
10 contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10, or a sequence
of amino acids with at least 90% sequence identity to the contiguous sequence of
amino acids of SEQ ID NO:2, 4, 6, 8 or 10.
3. The method of claim 3, wherein the SCFA1 polypeptide comprises the
15 contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10.
4. The method of any one of claims 1-3, wherein the method comprises
stimulating epithelial cell proliferation in the oral or gastrointestinal tract.
- 20 5. The method of claim 4, wherein the method comprises stimulating
epithelial cell proliferation in the esophagus.
6. The method of claim 4, wherein the method comprises stimulating
epithelial cell proliferation in the small intestine.
25
7. The method of claim 4, wherein the method comprises stimulating
epithelial cell proliferation in the large intestine.
8. The method of claim 4, wherein the method comprises stimulating
30 epithelial cell proliferation in the stomach.
9. The method of claim 4, wherein the method comprises stimulating
epithelial cell proliferation in the oral cavity.

10. A method of treating a gastrointestinal disorder or oral mucosal disorder in a mammalian subject in need thereof, said method comprising administering to said subject a therapeutically effective amount of a composition comprising a stem cell factor-like protein A1 (SCFA1) polypeptide and a pharmaceutically acceptable carrier.

11. The method of claim 10, wherein the SCFA1 polypeptide comprises the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10, or a sequence of amino acids with at least 90% sequence identity to the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8 or 10.

12. The method of claim 11, wherein the SCFA1 polypeptide comprises the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10.

13. The method of any one of claims 10-12, wherein the disorder is mucositis, inflammatory bowel disease, or short bowel syndrome.

14. A method for treating a mammalian subject at risk for damage to the epithelial cells lining at least a portion of the oral or gastrointestinal tract, the method comprising administering to said subject a therapeutically effective amount of a composition comprising a stem cell factor-like protein A1 (SCFA1) polypeptide and a pharmaceutically acceptable carrier.

15. The method of claim 14, wherein the SCFA1 polypeptide comprises the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10, or a sequence of amino acids with at least 90% sequence identity to the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8 or 10.

16. The method of claim 15, wherein the SCFA1 polypeptide comprises the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10.

17. The method of any one of claims 14-16 wherein the subject has undergone or will undergo radiation therapy.

18. The method of any one of claims 14-16 wherein the subject has undergone or will undergo chemotherapy.

19. An adenoviral vector comprising a polynucleotide encoding a stem cell factor-like protein A1 (SCFA1) polypeptide operably associated with an expression control sequence.

20. The adenoviral vector of claim 19, wherein the SCFA1 polypeptide comprises the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10, or a sequence of amino acids with at least 90% sequence identity to the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8 or 10.

21. The adenoviral vector of claim 20, wherein the SCFA1 polypeptide comprises the contiguous sequence of amino acids of SEQ ID NO:2, 4, 6, 8, or 10.

22. A pharmaceutical composition comprising the adenoviral vector of any one of claims 19-22 and a pharmaceutically acceptable carrier.

23. A method of stimulating epithelial cell proliferation in a subject in need thereof, said method comprising administering to said subject a therapeutically effective amount of the pharmaceutical composition of claim 22.

24. A method of stimulating epithelial cell proliferation in a subject in need thereof, said method comprising administering to said subject a therapeutically effective amount of a pharmaceutical composition comprising a vector encoding a stem cell factor-like protein A1 (SCFA1) polypeptide operably associated with an expression control sequence and a pharmaceutically acceptable carrier.

25. Use of a stem cell factor-like protein A1 (SCFA1) polypeptide in the manufacture of a medicament for stimulating epithelial cell proliferation in a subject in need thereof.

5 26. Use of a stem cell factor-like protein A1 (SCFA1) polypeptide in the manufacture of a medicament for treating a gastrointestinal disorder or oral mucosal disorder in a mammalian subject in need thereof.

10 27. Use of a stem cell factor-like protein A1 (SCFA1) polypeptide in the manufacture of a medicament for treating a mammalian subject at risk for damage to the epithelial cells lining at least a portion of the oral or gastrointestinal tract.

15 28. Use of an adenoviral vector comprising a polynucleotide encoding a stem cell factor-like protein A1 (SCFA1) polypeptide operably associated with an expression control sequence in the manufacture of a medicament for stimulating epithelial cell proliferation in a subject in need thereof.

20 29. Use of a vector comprising a polynucleotide encoding a stem cell factor-like protein A1 (SCFA1) polypeptide operably associated with an expression control sequence in the manufacture of a medicament for stimulating epithelial cell proliferation in a subject in need thereof.

25

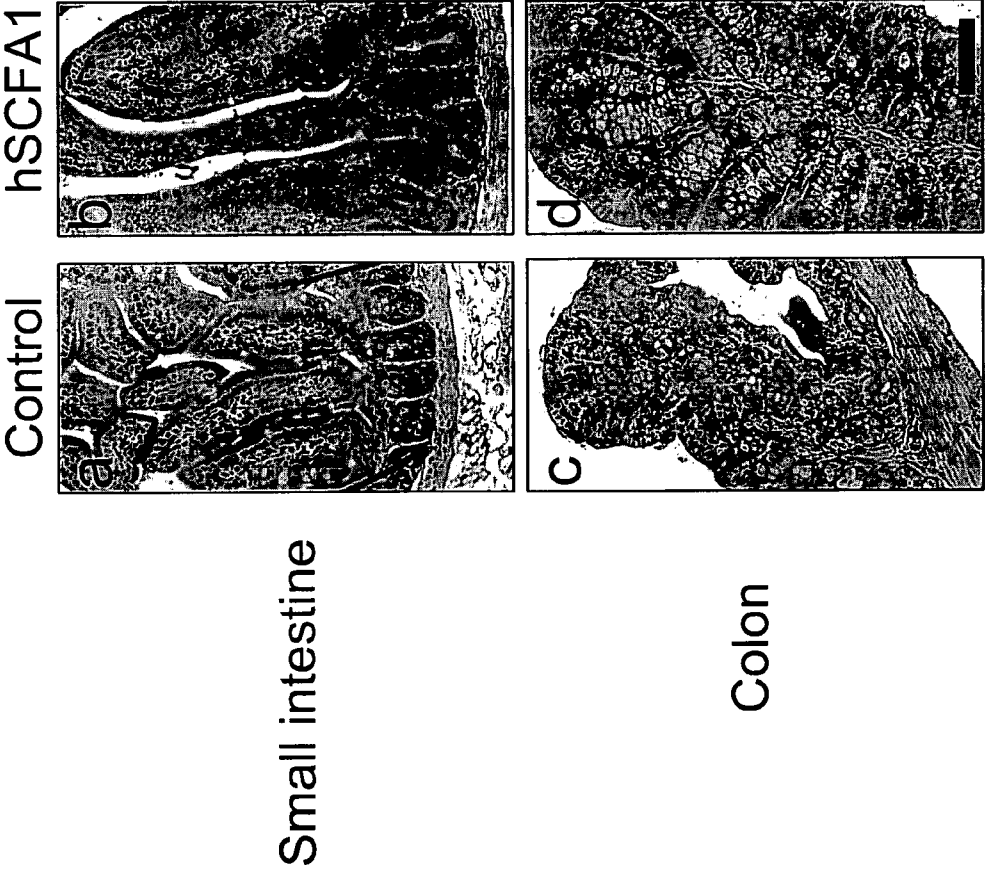


Figure 1

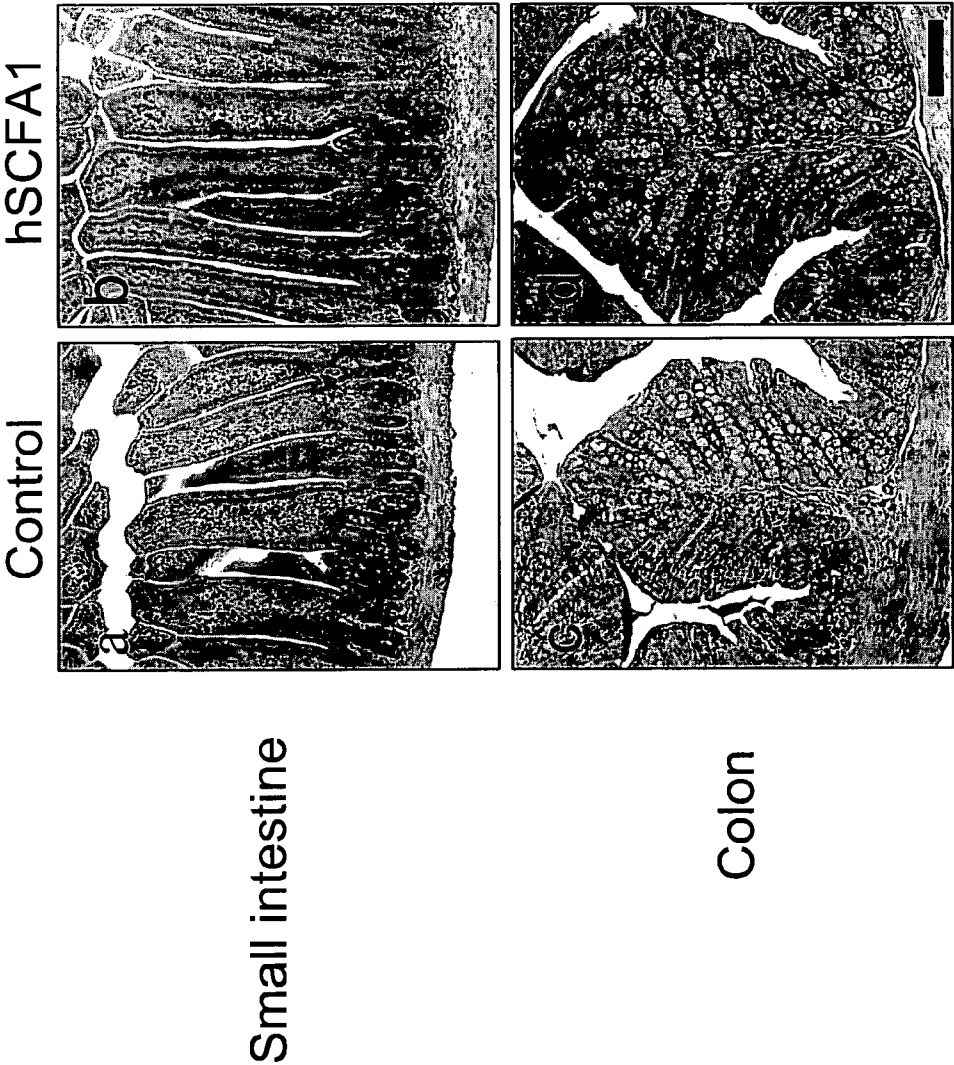


Figure 2

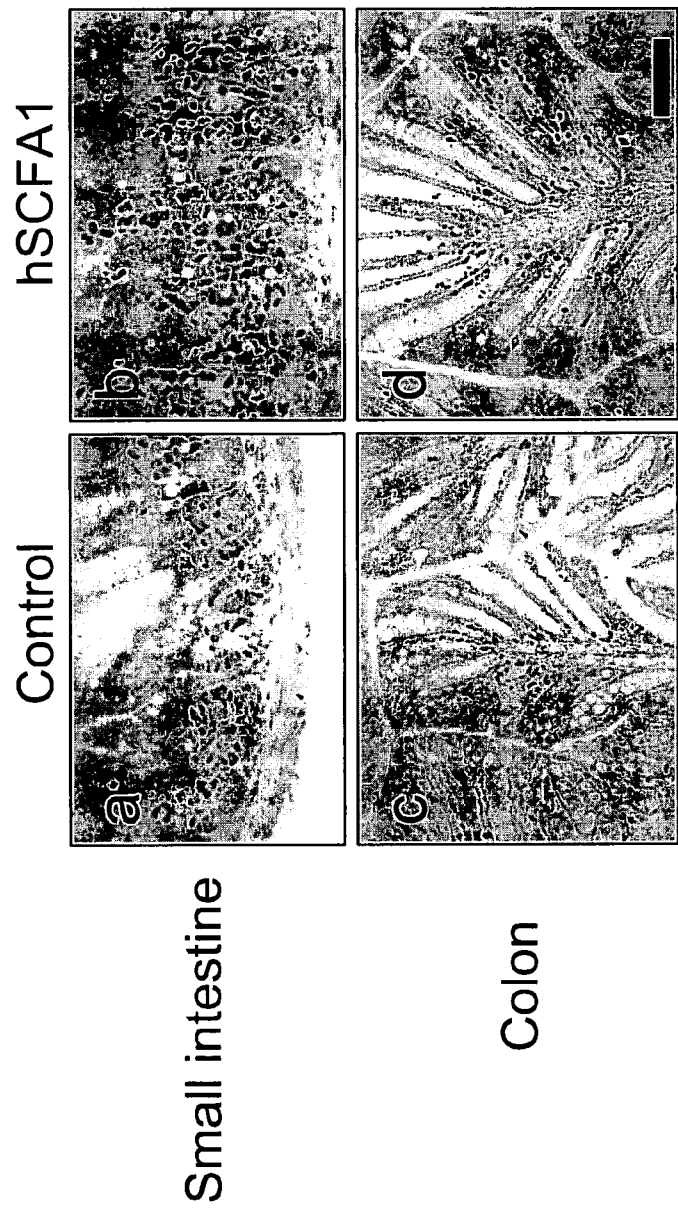


Figure 3

Effect of SCFA1 on β -catenin stabilization in HEK 293 cells

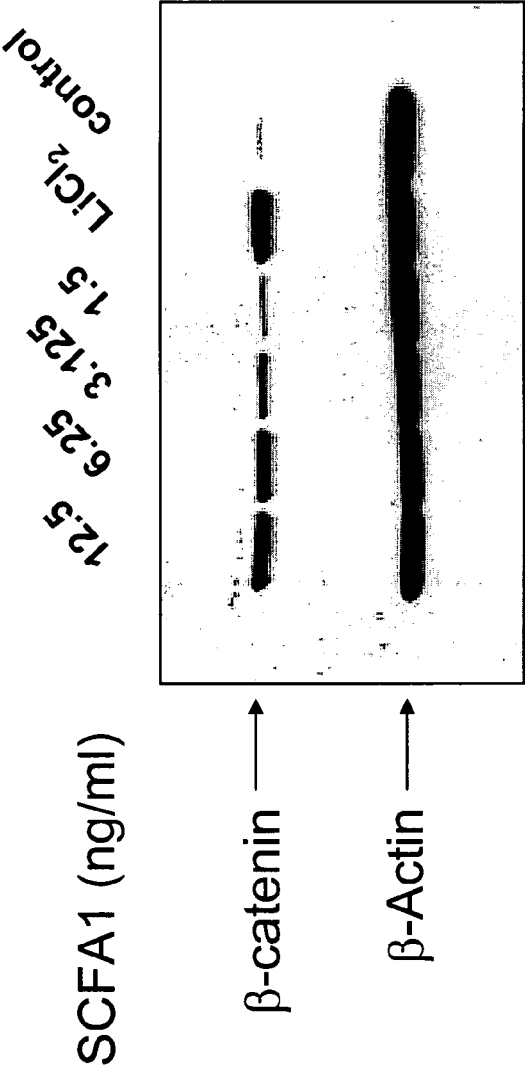


FIGURE 4

Effect of SCFA1 on β -catenin/TCF mediated trascriptional activation using 16TCF-luciferase reporter assay

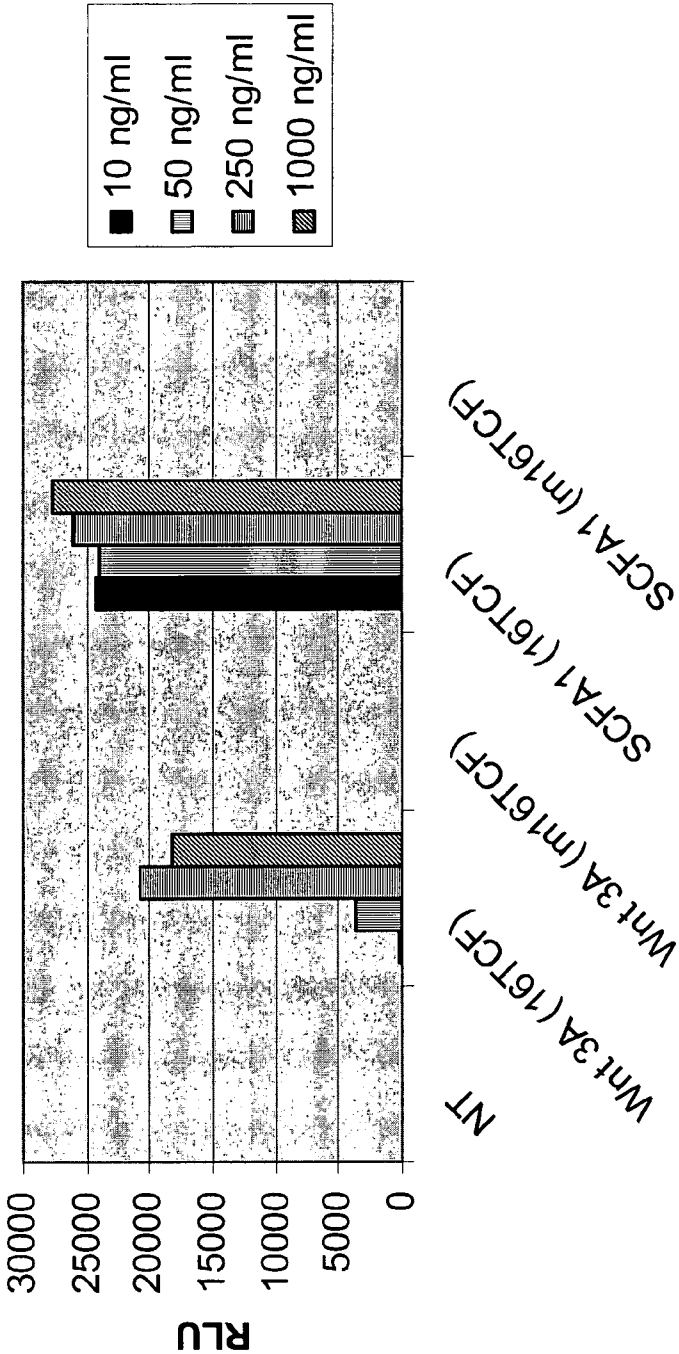


FIGURE 5

SEQUENCE LISTING

<110> Emtage, Peter C.R.
Funk, Walter

<120> Stem cell factor-like protein SCFA1 and uses thereof

<130> NUVO-27/PCT

<150> 60/724,908

<151> 2005-10-07

<160> 27

<170> PatentIn version 3.3

<210> 1

<211> 1343

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (291)..(1106)

<400> 1

```

gcggccgccc cggcggctcc tggaaccccc gttcgcggcg atgccagcca ccccagcgaa      60
gccgccgcag ttcagtgcct ggataatttg aaagtacaat agttgggttc cctgtccacc      120
cgccccactt cgcttgccat cacagcacgc ctatcggatg tgagaggaga agtcccgcgtg      180
ctcgggcact gtctatatac gcctaacacc tacatatatt ttaaaaacat taaatataat      240
taacaatcaa aagaaagagg agaaaggaag ggaagcatta ctgggttact atg cac      296
                                     Met His
                                     1

ttg cga ctg att tct tgg ctt ttt atc att ttg aac ttt atg gaa tac      344
Leu Arg Leu Ile Ser Trp Leu Phe Ile Ile Leu Asn Phe Met Glu Tyr
                                     5
                                     10
                                     15

atc ggc agc caa aac gcc tcc cgg gga agg cgc cag cga aga atg cat      392
Ile Gly Ser Gln Asn Ala Ser Arg Gly Arg Arg Gln Arg Arg Met His
    20
    25
    30

cct aac gtt agt caa ggc tgc caa gga ggc tgt gca aca tgc tca gat      440
Pro Asn Val Ser Gln Gly Cys Gln Gly Gly Cys Ala Thr Cys Ser Asp
    35
    40
    45

tac aat gga tgt ttg tca tgt aag ccc aga cta ttt ttt gct ctg gaa      488
Tyr Asn Gly Cys Leu Ser Cys Lys Pro Arg Leu Phe Phe Ala Leu Glu
    55
    60
    65

aga att ggc atg aag cag att gga gta tgt ctc tct tca tgt cca agt      536
Arg Ile Gly Met Lys Gln Ile Gly Val Cys Leu Ser Ser Cys Pro Ser
    70
    75
    80

gga tat tat gga act cga tat cca gat ata aat aag tgt aca aaa tgc      584
Gly Tyr Tyr Gly Thr Arg Tyr Pro Asp Ile Asn Lys Cys Thr Lys Cys
    85
    90
    95

aaa gct gac tgt gat acc tgt ttc aac aaa aat ttc tgc aca aaa tgt      632
Lys Ala Asp Cys Asp Thr Cys Phe Asn Lys Asn Phe Cys Thr Lys Cys
    100
    105
    110

```

aaa agt gga ttt tac tta cac ctt gga aag tgc ctt gac aat tgc cca 680
 Lys Ser Gly Phe Tyr Leu His Leu Gly Lys Cys Leu Asp Asn Cys Pro
 115 120 125 130
 gaa ggg ttg gaa gcc aac aac cat act atg gag tgt gtc agt att gtg 728
 Glu Gly Leu Glu Ala Asn Asn His Thr Met Glu Cys Val Ser Ile Val
 135 140 145
 cac tgt gag gtc agt gaa tgg aat cct tgg agt cca tgc acg aag aag 776
 His Cys Glu Val Ser Glu Trp Asn Pro Trp Ser Pro Cys Thr Lys Lys
 150 155 160
 gga aaa aca tgt ggc ttc aaa aga ggg act gaa aca cgg gtc cga gaa 824
 Gly Lys Thr Cys Gly Phe Lys Arg Gly Thr Glu Thr Arg Val Arg Glu
 165 170 175
 ata ata cag cat cct tca gca aag ggt aac cta tgt ccc cca aca aat 872
 Ile Ile Gln His Pro Ser Ala Lys Gly Asn Leu Cys Pro Pro Thr Asn
 180 185 190
 gag aca aga aag tgt aca gtg caa agg aag aag tgt cag aag gga gaa 920
 Glu Thr Arg Lys Cys Thr Val Gln Arg Lys Lys Cys Gln Lys Gly Glu
 195 200 205 210
 cga gga aaa aaa gga agg gag agg aaa aga aaa aaa cct aat aaa gga 968
 Arg Gly Lys Lys Gly Arg Glu Arg Lys Arg Lys Lys Pro Asn Lys Gly
 215 220 225
 gaa agt aaa gaa gca ata cct gac agc aaa agt ctg gaa tcc agc aaa 1016
 Glu Ser Lys Glu Ala Ile Pro Asp Ser Lys Ser Leu Glu Ser Ser Lys
 230 235 240
 gaa atc cca gag caa cga gaa aac aaa cag cag cag aag aag cga aaa 1064
 Glu Ile Pro Glu Gln Arg Glu Asn Lys Gln Gln Gln Lys Lys Arg Lys
 245 250 255
 gtc caa gat aaa cag aaa tcg gta tca gtc agc act gta cac 1106
 Val Gln Asp Lys Gln Lys Ser Val Ser Val Ser Thr Val His
 260 265 270
 tagagggttc catgagatta ttgtagactc atgatgctgc tatctcaacc agatgcccag 1166
 gacaggtgct ctagccatta ggaccacaaa tggacatgct agttattgct ctgtctaaac 1226
 aacattccca gtagttgcta tattcttcat acaagcatag ttaacaacaa agagccaaaa 1286
 gatcaaagaa gggatacttt cagatggttg tcttgtgtgc ttctctgcat ttttaaa 1343

<210> 2
 <211> 272
 <212> PRT
 <213> Homo sapiens

<400> 2

Met His Leu Arg Leu Ile Ser Trp Leu Phe Ile Ile Leu Asn Phe Met
1 5 10 15

Glu Tyr Ile Gly Ser Gln Asn Ala Ser Arg Gly Arg Arg Gln Arg Arg
20 25 30

Met His Pro Asn Val Ser Gln Gly Cys Gln Gly Gly Cys Ala Thr Cys
35 40 45

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

<210> 3
 <211> 723
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(723)

<400> 3	
aga atg cat cct aac gtt agt caa ggc tgc caa gga ggc tgt gca aca	48
Arg Met His Pro Asn Val Ser Gln Gly Cys Gln Gly Gly Cys Ala Thr	
1 5 10 15	
tgc tca gat tac aat gga tgt ttg tca tgt aag ccc aga cta ttt ttt	96
Cys Ser Asp Tyr Asn Gly Cys Leu Ser Cys Lys Pro Arg Leu Phe Phe	
20 25 30	
gct ctg gaa aga att ggc atg aag cag att gga gta tgt ctc tct tca	144
Ala Leu Glu Arg Ile Gly Met Lys Gln Ile Gly Val Cys Leu Ser Ser	
35 40 45	
tgt cca agt gga tat tat gga act cga tat cca gat ata aat aag tgt	192
Cys Pro Ser Gly Tyr Tyr Gly Thr Arg Tyr Pro Asp Ile Asn Lys Cys	
50 55 60	
aca aaa tgc aaa gct gac tgt gat acc tgt ttc aac aaa aat ttc tgc	240
Thr Lys Cys Lys Ala Asp Cys Asp Thr Cys Phe Asn Lys Asn Phe Cys	
65 70 75 80	
aca aaa tgt aaa agt gga ttt tac tta cac ctt gga aag tgc ctt gac	288
Thr Lys Cys Lys Ser Gly Phe Tyr Leu His Leu Gly Lys Cys Leu Asp	
85 90 95	
aat tgc cca gaa ggg ttg gaa gcc aac aac cat act atg gag tgt gtc	336
Asn Cys Pro Glu Gly Leu Glu Ala Asn Asn His Thr Met Glu Cys Val	
100 105 110	
agt att gtg cac tgt gag gtc agt gaa tgg aat cct tgg agt cca tgc	384
Ser Ile Val His Cys Glu Val Ser Glu Trp Asn Pro Trp Ser Pro Cys	
115 120 125	
acg aag aag gga aaa aca tgt ggc ttc aaa aga ggg act gaa aca cgg	432
Thr Lys Lys Gly Lys Thr Cys Gly Phe Lys Arg Gly Thr Glu Thr Arg	
130 135 140	
gtc cga gaa ata ata cag cat cct tca gca aag ggt aac cta tgt ccc	480
Val Arg Glu Ile Ile Gln His Pro Ser Ala Lys Gly Asn Leu Cys Pro	
145 150 155 160	
cca aca aat gag aca aga aag tgt aca gtg caa agg aag aag tgt cag	528
Pro Thr Asn Glu Thr Arg Lys Cys Thr Val Gln Arg Lys Lys Cys Gln	
165 170 175	
aag gga gaa cga gga aaa aaa gga agg gag agg aaa aga aaa aaa cct	576
Lys Gly Glu Arg Gly Lys Lys Gly Arg Glu Arg Lys Arg Lys Lys Pro	
180 185 190	
aat aaa gga gaa agt aaa gaa gca ata cct gac agc aaa agt ctg gaa	624
Asn Lys Gly Glu Ser Lys Glu Ala Ile Pro Asp Ser Lys Ser Leu Glu	
195 200 205	
tcc agc aaa gaa atc cca gag caa cga gaa aac aaa cag cag cag aag	672
Ser Ser Lys Glu Ile Pro Glu Gln Arg Glu Asn Lys Gln Gln Gln Lys	
210 215 220	
aag cga aaa gtc caa gat aaa cag aaa tcg gta tca gtc agc act gta	720
Lys Arg Lys Val Gln Asp Lys Gln Lys Ser Val Ser Val Ser Thr Val	
225 230 235 240	
cac	723
His	

<210> 4

<211> 241
 <212> PRT
 <213> Homo sapiens

<400> 4

Arg Met His Pro Asn Val Ser Gln Gly Cys Gln Gly Gly Cys Ala Thr
 1 5 10 15

Cys Ser Asp Tyr Asn Gly Cys Leu Ser Cys Lys Pro Arg Leu Phe Phe
 20 25 30

Ala Leu Glu Arg Ile Gly Met Lys Gln Ile Gly Val Cys Leu Ser Ser
 35 40 45

Cys Pro Ser Gly Tyr Tyr Gly Thr Arg Tyr Pro Asp Ile Asn Lys Cys
 50 55 60

Thr Lys Cys Lys Ala Asp Cys Asp Thr Cys Phe Asn Lys Asn Phe Cys
 65 70 75 80

Thr Lys Cys Lys Ser Gly Phe Tyr Leu His Leu Gly Lys Cys Leu Asp
 85 90 95

Asn Cys Pro Glu Gly Leu Glu Ala Asn Asn His Thr Met Glu Cys Val
 100 105 110

Ser Ile Val His Cys Glu Val Ser Glu Trp Asn Pro Trp Ser Pro Cys
 115 120 125

Thr Lys Lys Gly Lys Thr Cys Gly Phe Lys Arg Gly Thr Glu Thr Arg
 130 135 140

Val Arg Glu Ile Ile Gln His Pro Ser Ala Lys Gly Asn Leu Cys Pro
 145 150 155 160

Pro Thr Asn Glu Thr Arg Lys Cys Thr Val Gln Arg Lys Lys Cys Gln
 165 170 175

Lys Gly Glu Arg Gly Lys Lys Gly Arg Glu Arg Lys Arg Lys Lys Pro
 180 185 190

Asn Lys Gly Glu Ser Lys Glu Ala Ile Pro Asp Ser Lys Ser Leu Glu
 195 200 205

Ser Ser Lys Glu Ile Pro Glu Gln Arg Glu Asn Lys Gln Gln Gln Lys
 210 215 220

Lys Arg Lys Val Gln Asp Lys Gln Lys Ser Val Ser Val Ser Thr Val
 225 230 235 240

His

<210> 5
 <211> 888
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Chimeric Igk-SCFA1-V5His tag

<220>
 <221> CDS
 <222> (1)..(888)

<400> 5
 atg gag aca gac aca ctc ctg cta tgg gta ctg ctg ctc tgg gtt cca 48
 Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
 1 5 10 15
 ggt tcc act ggt gac gct agc aga atg cat cct aac gtt agt caa ggc 96
 Gly Ser Thr Gly Asp Ala Ser Arg Met His Pro Asn Val Ser Gln Gly
 20 25 30
 tgc caa gga ggc tgt gca aca tgc tca gat tac aat gga tgt ttg tca 144
 Cys Gln Gly Gly Cys Ala Thr Cys Ser Asp Tyr Asn Gly Cys Leu Ser
 35 40 45
 tgt aag ccc aga cta ttt ttt gct ctg gaa aga att ggc atg aag cag 192
 Cys Lys Pro Arg Leu Phe Phe Ala Leu Glu Arg Ile Gly Met Lys Gln
 50 55 60
 att gga gta tgt ctc tct tca tgt cca agt gga tat tat gga act cga 240
 Ile Gly Val Cys Leu Ser Ser Cys Pro Ser Gly Tyr Tyr Gly Thr Arg
 65 70 75 80
 tat cca gat ata aat aag tgt aca aaa tgc aaa gct gac tgt gat acc 288
 Tyr Pro Asp Ile Asn Lys Cys Thr Lys Cys Lys Ala Asp Cys Asp Thr
 85 90 95
 tgt ttc aac aaa aat ttc tgc aca aaa tgt aaa agt gga ttt tac tta 336
 Cys Phe Asn Lys Asn Phe Cys Thr Lys Cys Lys Ser Gly Phe Tyr Leu
 100 105 110
 cac ctt gga aag tgc ctt gac aat tgc cca gaa ggg ttg gaa gcc aac 384
 His Leu Gly Lys Cys Leu Asp Asn Cys Pro Glu Gly Leu Glu Ala Asn
 115 120 125
 aac cat act atg gag tgt gtc agt att gtg cac tgt gag gtc agt gaa 432
 Asn His Thr Met Glu Cys Val Ser Ile Val His Cys Glu Val Ser Glu
 130 135 140
 tgg aat cct tgg agt cca tgc acg aag aag gga aaa aca tgt ggc ttc 480
 Trp Asn Pro Trp Ser Pro Cys Thr Lys Lys Gly Lys Thr Cys Gly Phe
 145 150 155 160
 aaa aga ggg act gaa aca cgg gtc cga gaa ata ata cag cat cct tca 528
 Lys Arg Gly Thr Glu Thr Arg Val Arg Glu Ile Ile Gln His Pro Ser
 165 170 175
 gca aag ggt aac cta tgt ccc cca aca aat gag aca aga aag tgt aca 576
 Ala Lys Gly Asn Leu Cys Pro Pro Thr Asn Glu Thr Arg Lys Cys Thr
 180 185 190
 gtg caa agg aag aag tgt cag aag gga gaa cga gga aaa aaa gga agg 624
 Val Gln Arg Lys Lys Cys Gln Lys Gly Glu Arg Gly Lys Lys Lys Gly Arg

195						200						205						
gag	agg	aaa	aga	aaa	aaa	cct	aat	aaa	gga	gaa	agt	aaa	gaa	gca	ata	672		
Glu	Arg	Lys	Arg	Lys	Lys	Pro	Asn	Lys	Gly	Glu	Ser	Lys	Glu	Ala	Ile			
	210					215					220							
cct	gac	agc	aaa	agt	ctg	gaa	tcc	agc	aaa	gaa	atc	cca	gag	caa	cga	720		
Pro	Asp	Ser	Lys	Ser	Leu	Glu	Ser	Ser	Lys	Glu	Ile	Pro	Glu	Gln	Arg			
225					230					235					240			
gaa	aac	aaa	cag	cag	cag	aag	aag	cga	aaa	gtc	caa	gat	aaa	cag	aaa	768		
Glu	Asn	Lys	Gln	Gln	Gln	Lys	Lys	Arg	Lys	Val	Gln	Asp	Lys	Gln	Lys			
			245						250					255				
tcg	gta	tca	gtc	agc	act	gta	cac	agt	cta	gag	ggc	ccg	cgg	ttc	gaa	816		
Ser	Val	Ser	Val	Ser	Thr	Val	His	Ser	Leu	Glu	Gly	Pro	Arg	Phe	Glu			
			260					265					270					
ggt	aag	cct	atc	cct	aac	cct	ctc	ctc	ggt	ctc	gat	tct	acg	cgt	acc	864		
Gly	Lys	Pro	Ile	Pro	Asn	Pro	Leu	Leu	Gly	Leu	Asp	Ser	Thr	Arg	Thr			
		275					280					285						
ggt	cat	cat	cac	cat	cac	cat	tga									888		
Gly	His	His	His	His	His	His												
	290					295												

<210> 6
 <211> 295
 <212> PRT
 <213> Artificial sequence
 <220>
 <223> Chimeric Igk-SCFA1-V5His tag
 <400> 6

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

Asn His Thr Met Glu Cys Val Ser Ile Val His Cys Glu Val Ser Glu
 130 135 140
 Trp Asn Pro Trp Ser Pro Cys Thr Lys Lys Gly Lys Thr Cys Gly Phe
 145 150 155 160
 Lys Arg Gly Thr Glu Thr Arg Val Arg Glu Ile Ile Gln His Pro Ser
 165 170 175
 Ala Lys Gly Asn Leu Cys Pro Pro Thr Asn Glu Thr Arg Lys Cys Thr
 180 185 190
 Val Gln Arg Lys Lys Cys Gln Lys Gly Glu Arg Gly Lys Lys Gly Arg
 195 200 205
 Glu Arg Lys Arg Lys Lys Pro Asn Lys Gly Glu Ser Lys Glu Ala Ile
 210 215 220
 Pro Asp Ser Lys Ser Leu Glu Ser Ser Lys Glu Ile Pro Glu Gln Arg
 225 230 235 240
 Glu Asn Lys Gln Gln Gln Lys Lys Arg Lys Val Gln Asp Lys Gln Lys
 245 250 255
 Ser Val Ser Val Ser Thr Val His Ser Leu Glu Gly Pro Arg Phe Glu
 260 265 270
 Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr Arg Thr
 275 280 285
 Gly His His His His His His
 290 295

<210> 7
 <211> 753
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(753)

<400> 7		
caa aac gcc tcc cgg gga agg cgc cag cga aga atg cat cct aac gtt		48
Gln Asn Ala Ser Arg Gly Arg Arg Gln Arg Arg Met His Pro Asn Val		
1 5 10 15		
agt caa ggc tgc caa gga ggc tgt gca aca tgc tca gat tac aat gga		96
Ser Gln Gly Cys Gln Gly Gly Cys Ala Thr Cys Ser Asp Tyr Asn Gly		
20 25 30		
tgt ttg tca tgt aag ccc aga cta ttt ttt gct ctg gaa aga att ggc		144
Cys Leu Ser Cys Lys Pro Arg Leu Phe Phe Ala Leu Glu Arg Ile Gly		

35					40					45						
atg	aag	cag	att	gga	gta	tgt	ctc	tct	tca	tgt	cca	agt	gga	tat	tat	192
Met	Lys	Gln	Ile	Gly	Val	Cys	Leu	Ser	Ser	Cys	Pro	Ser	Gly	Tyr	Tyr	
	50					55					60					
gga	act	cga	tat	cca	gat	ata	aat	aag	tgt	aca	aaa	tgc	aaa	gct	gac	240
Gly	Thr	Arg	Tyr	Pro	Asp	Ile	Asn	Lys	Cys	Thr	Lys	Cys	Lys	Ala	Asp	
65					70					75					80	
tgt	gat	acc	tgt	ttc	aac	aaa	aat	ttc	tgc	aca	aaa	tgt	aaa	agt	gga	288
Cys	Asp	Thr	Cys	Phe	Asn	Lys	Asn	Phe	Cys	Thr	Lys	Cys	Lys	Ser	Gly	
				85					90					95		
ttt	tac	tta	cac	ctt	gga	aag	tgc	ctt	gac	aat	tgc	cca	gaa	ggg	ttg	336
Phe	Tyr	Leu	His	Leu	Gly	Lys	Cys	Leu	Asp	Asn	Cys	Pro	Glu	Gly	Leu	
			100				105						110			
gaa	gcc	aac	aac	cat	act	atg	gag	tgt	gtc	agt	att	gtg	cac	tgt	gag	384
Glu	Ala	Asn	Asn	His	Thr	Met	Glu	Cys	Val	Ser	Ile	Val	His	Cys	Glu	
		115					120					125				
gtc	agt	gaa	tgg	aat	cct	tgg	agt	cca	tgc	acg	aag	aag	gga	aaa	aca	432
Val	Ser	Glu	Trp	Asn	Pro	Trp	Ser	Pro	Cys	Thr	Lys	Lys	Gly	Lys	Thr	
	130					135					140					
tgt	ggc	ttc	aaa	aga	ggg	act	gaa	aca	cgg	gtc	cga	gaa	ata	ata	cag	480
Cys	Gly	Phe	Lys	Arg	Gly	Thr	Glu	Thr	Arg	Val	Arg	Glu	Ile	Ile	Gln	
145					150					155					160	
cat	cct	tca	gca	aag	ggg	aac	cta	tgt	ccc	cca	aca	aat	gag	aca	aga	528
His	Pro	Ser	Ala	Lys	Gly	Asn	Leu	Cys	Pro	Pro	Thr	Asn	Glu	Thr	Arg	
				165					170					175		
aag	tgt	aca	gtg	caa	agg	aag	aag	tgt	cag	aag	gga	gaa	cga	gga	aaa	576
Lys	Cys	Thr	Val	Gln	Arg	Lys	Lys	Cys	Gln	Lys	Gly	Glu	Arg	Gly	Lys	
			180					185					190			
aaa	gga	agg	gag	agg	aaa	aga	aaa	aaa	cct	aat	aaa	gga	gaa	agt	aaa	624
Lys	Gly	Arg	Glu	Arg	Lys	Arg	Lys	Lys	Pro	Asn	Lys	Gly	Glu	Ser	Lys	
		195					200					205				
gaa	gca	ata	cct	gac	agc	aaa	agt	ctg	gaa	tcc	agc	aaa	gaa	atc	cca	672
Glu	Ala	Ile	Pro	Asp	Ser	Lys	Ser	Leu	Glu	Ser	Ser	Lys	Glu	Ile	Pro	
	210					215					220					
gag	caa	cga	gaa	aac	aaa	cag	cag	cag	aag	aag	cga	aaa	gtc	caa	gat	720
Glu	Gln	Arg	Glu	Asn	Lys	Gln	Gln	Gln	Lys	Lys	Arg	Lys	Val	Gln	Asp	
225					230					235					240	
aaa	cag	aaa	tcg	gta	tca	gtc	agc	act	gta	cac						753
Lys	Gln	Lys	Ser	Val	Ser	Val	Ser	Thr	Val	His						
				245					250							

<210> 8
 <211> 251
 <212> PRT
 <213> Homo sapiens

<400> 8

Gln Asn Ala Ser Arg Gly Arg Arg Gln Arg Arg Met His Pro Asn Val
 1 5 10 15

Ser Gln Gly Cys Gln Gly Gly Cys Ala Thr Cys Ser Asp Tyr Asn Gly

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

<210> 9
 <211> 720
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS

<222> (1)..(720)

<400> 9

atg cat cct aac gtt agt caa ggc tgc caa gga ggc tgt gca aca tgc	48
Met His Pro Asn Val Ser Gln Gly Cys Gln Gly Gly Cys Ala Thr Cys	
1 5 10 15	
tca gat tac aat gga tgt ttg tca tgt aag ccc aga cta ttt ttt gct	96
Ser Asp Tyr Asn Gly Cys Leu Ser Cys Lys Pro Arg Leu Phe Phe Ala	
20 25 30	
ctg gaa aga att ggc atg aag cag att gga gta tgt ctc tct tca tgt	144
Leu Glu Arg Ile Gly Met Lys Gln Ile Gly Val Cys Leu Ser Ser Cys	
35 40 45	
cca agt gga tat tat gga act cga tat cca gat ata aat aag tgt aca	192
Pro Ser Gly Tyr Tyr Gly Thr Arg Tyr Pro Asp Ile Asn Lys Cys Thr	
50 55 60	
aaa tgc aaa gct gac tgt gat acc tgt ttc aac aaa aat ttc tgc aca	240
Lys Cys Lys Ala Asp Cys Asp Thr Cys Phe Asn Lys Asn Phe Cys Thr	
65 70 75 80	
aaa tgt aaa agt gga ttt tac tta cac ctt gga aag tgc ctt gac aat	288
Lys Cys Lys Ser Gly Phe Tyr Leu His Leu Gly Lys Cys Leu Asp Asn	
85 90 95	
tgc cca gaa ggg ttg gaa gcc aac aac cat act atg gag tgt gtc agt	336
Cys Pro Glu Gly Leu Glu Ala Asn Asn His Thr Met Glu Cys Val Ser	
100 105 110	
att gtg cac tgt gag gtc agt gaa tgg aat cct tgg agt cca tgc acg	384
Ile Val His Cys Glu Val Ser Glu Trp Asn Pro Trp Ser Pro Cys Thr	
115 120 125	
aag aag gga aaa aca tgt ggc ttc aaa aga ggg act gaa aca cgg gtc	432
Lys Lys Gly Lys Thr Cys Gly Phe Lys Arg Gly Thr Glu Thr Arg Val	
130 135 140	
cga gaa ata ata cag cat cct tca gca aag ggt aac cta tgt ccc cca	480
Arg Glu Ile Ile Gln His Pro Ser Ala Lys Gly Asn Leu Cys Pro Pro	
145 150 155 160	
aca aat gag aca aga aag tgt aca gtg caa agg aag aag tgt cag aag	528
Thr Asn Glu Thr Arg Lys Cys Thr Val Gln Arg Lys Lys Cys Gln Lys	
165 170 175	
gga gaa cga gga aaa aaa gga agg gag agg aaa aga aaa aaa cct aat	576
Gly Glu Arg Gly Lys Lys Gly Arg Glu Arg Lys Arg Lys Lys Lys Pro Asn	
180 185 190	
aaa gga gaa agt aaa gaa gca ata cct gac agc aaa agt ctg gaa tcc	624
Lys Gly Glu Ser Lys Glu Ala Ile Pro Asp Ser Lys Ser Leu Glu Ser	
195 200 205	
agc aaa gaa atc cca gag caa cga gaa aac aaa cag cag cag aag aag	672
Ser Lys Glu Ile Pro Glu Gln Arg Glu Asn Lys Gln Gln Gln Lys Lys	
210 215 220	
cga aaa gtc caa gat aaa cag aaa tcg gta tca gtc agc act gta cac	720
Arg Lys Val Gln Asp Lys Gln Lys Ser Val Ser Val Ser Thr Val His	
225 230 235 240	

<210> 10

<211> 240

<212> PRT

12/21

<213> Homo sapiens

<400> 10

Met His Pro Asn Val Ser Gln Gly Cys Gln Gly Gly Cys Ala Thr Cys
 1 5 10 15

Ser Asp Tyr Asn Gly Cys Leu Ser Cys Lys Pro Arg Leu Phe Phe Ala
 20 25 30

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

Pro Ser Gly Tyr Tyr Gly Thr Arg Tyr Pro Asp Ile Asn Lys Cys Thr
 50 55 60

Lys Cys Lys Ala Asp Cys Asp Thr Cys Phe Asn Lys Asn Phe Cys Thr
 65 70 75 80

Lys Cys Lys Ser Gly Phe Tyr Leu His Leu Gly Lys Cys Leu Asp Asn
 85 90 95

Cys Pro Glu Gly Leu Glu Ala Asn Asn His Thr Met Glu Cys Val Ser
 100 105 110

Ile Val His Cys Glu Val Ser Glu Trp Asn Pro Trp Ser Pro Cys Thr
 115 120 125

Lys Lys Gly Lys Thr Cys Gly Phe Lys Arg Gly Thr Glu Thr Arg Val
 130 135 140

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

Thr Asn Glu Thr Arg Lys Cys Thr Val Gln Arg Lys Lys Cys Gln Lys
 165 170 175

Gly Glu Arg Gly Lys Lys Gly Arg Glu Arg Lys Arg Lys Lys Pro Asn
 180 185 190

Lys Gly Glu Ser Lys Glu Ala Ile Pro Asp Ser Lys Ser Leu Glu Ser
 195 200 205

Ser Lys Glu Ile Pro Glu Gln Arg Glu Asn Lys Gln Gln Gln Lys Lys
 210 215 220

Arg Lys Val Gln Asp Lys Gln Lys Ser Val Ser Val Ser Thr Val His
 225 230 235 240

<210> 11

<211> 132

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(132)

<400> 11

ggc	tgc	caa	gga	ggc	tgt	gca	aca	tgc	tca	gat	tac	aat	gga	tgt	ttg	48
Gly	Cys	Gln	Gly	Gly	Cys	Ala	Thr	Cys	Ser	Asp	Tyr	Asn	Gly	Cys	Leu	
1				5					10					15		

tca	tgt	aag	ccc	aga	cta	ttt	ttt	gct	ctg	gaa	aga	att	ggc	atg	aag	96
Ser	Cys	Lys	Pro	Arg	Leu	Phe	Phe	Ala	Leu	Glu	Arg	Ile	Gly	Met	Lys	
			20					25					30			

cag	att	gga	gta	tgt	ctc	tct	tca	tgt	cca	agt	gga					132
Gln	Ile	Gly	Val	Cys	Leu	Ser	Ser	Cys	Pro	Ser	Gly					
		35					40									

<210> 12

<211> 44

<212> PRT

<213> Homo sapiens

<400> 12

Gly	Cys	Gln	Gly	Gly	Cys	Ala	Thr	Cys	Ser	Asp	Tyr	Asn	Gly	Cys	Leu
1				5					10					15	

Ser	Cys	Lys	Pro	Arg	Leu	Phe	Phe	Ala	Leu	Glu	Arg	Ile	Gly	Met	Lys
			20					25					30		

Gln	Ile	Gly	Val	Cys	Leu	Ser	Ser	Cys	Pro	Ser	Gly
		35					40				

<210> 13

<211> 120

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(120)

<400> 13

aag	tgt	aca	aaa	tgc	aaa	gct	gac	tgt	gat	acc	tgt	ttc	aac	aaa	aat	48
Lys	Cys	Thr	Lys	Cys	Lys	Ala	Asp	Cys	Asp	Thr	Cys	Phe	Asn	Lys	Asn	
1				5					10					15		

ttc	tgc	aca	aaa	tgt	aaa	agt	gga	ttt	tac	tta	cac	ctt	gga	aag	tgc	96
Phe	Cys	Thr	Lys	Cys	Lys	Ser	Gly	Phe	Tyr	Leu	His	Leu	Gly	Lys	Cys	
			20				25						30			

ctt	gac	aat	tgc	cca	gaa	ggg	ttg									120
Leu	Asp	Asn	Cys	Pro	Glu	Gly	Leu									
		35					40									

<210> 14

<211> 40

<212> PRT

<213> Homo sapiens

<400> 14

Lys Cys Thr Lys Cys Lys Ala Asp Cys Asp Thr Cys Phe Asn Lys Asn
 1 5 10 15

Phe Cys Thr Lys Cys Lys Ser Gly Phe Tyr Leu His Leu Gly Lys Cys
 20 25 30

Leu Asp Asn Cys Pro Glu Gly Leu
 35 40

<210> 15

<211> 261

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(261)

<400> 15

gtc agt gaa tgg aat cct tgg agt cca tgc acg aag aag gga aaa aca 48
 Val Ser Glu Trp Asn Pro Trp Ser Pro Cys Thr Lys Lys Gly Lys Thr
 1 5 10 15

tgt ggc ttc aaa aga ggg act gaa aca cgg gtc cga gaa ata ata cag 96
 Cys Gly Phe Lys Arg Gly Thr Glu Thr Arg Val Arg Glu Ile Ile Gln
 20 25 30

cat cct tca gca aag ggt aac cta tgt ccc cca aca aat gag aca aga 144
 His Pro Ser Ala Lys Gly Asn Leu Cys Pro Pro Thr Asn Glu Thr Arg
 35 40 45

aag tgt aca gtg caa agg aag aag tgt cag aag gga gaa cga gga aaa 192
 Lys Cys Thr Val Gln Arg Lys Lys Cys Gln Lys Gly Glu Arg Gly Lys
 50 55 60

aaa gga agg gag agg aaa aga aaa aaa cct aat aaa gga gaa agt aaa 240
 Lys Gly Arg Glu Arg Lys Arg Lys Lys Pro Asn Lys Gly Glu Ser Lys
 65 70 75 80

gaa gca ata cct gac agc aaa 261
 Glu Ala Ile Pro Asp Ser Lys
 85

<210> 16

<211> 87

<212> PRT

<213> Homo sapiens

<400> 16

Val Ser Glu Trp Asn Pro Trp Ser Pro Cys Thr Lys Lys Gly Lys Thr
 1 5 10 15

Cys Gly Phe Lys Arg Gly Thr Glu Thr Arg Val Arg Glu Ile Ile Gln
 20 25 30

His Pro Ser Ala Lys Gly Asn Leu Cys Pro Pro Thr Asn Glu Thr Arg
 35 40 45

Lys Cys Thr Val Gln Arg Lys Lys Cys Gln Lys Gly Glu Arg Gly Lys
 50 55 60

Lys Gly Arg Glu Arg Lys Arg Lys Lys Pro Asn Lys Gly Glu Ser Lys
 65 70 75 80

Glu Ala Ile Pro Asp Ser Lys
 85

<210> 17
 <211> 25
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Forward primer

<400> 17
 gcactgtctatatacgctta acacc 25

<210> 18
 <211> 28
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Reverse primer

<400> 18
 gcaataactg acatgtccat ttgtggtc 28

<210> 19
 <211> 36
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Forward primer

<400> 19
 ccggctagca gaatgcatcc taacgttagt caaggc 36

<210> 20
 <211> 34
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Reverse primer

<400> 20
 tgctctagac tgtgtacagt gctgactgat accg 34

<210> 21
 <211> 35
 <212> DNA

<213> Artificial sequence

<220>

<223> Forward primer

<400> 21

tgctctagac accatggaga cagacacact cctgc

35

<210> 22

<211> 30

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer

<400> 22

ggcgctagcg tcaccagtgg aacctggaac

30

<210> 23

<211> 35

<212> DNA

<213> Artificial sequence

<220>

<223> Forward primer

<400> 23

tgctctagac accatggaga cagacacact cctgc

35

<210> 24

<211> 38

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer

<400> 24

ccatgcggcc gccttggtta gctctgtctg tagctagg

38

<210> 25

<211> 28

<212> DNA

<213> Artificial sequence

<220>

<223> Forward primer

<400> 25

caccctagg tcaatattgg ccattagc

28

<210> 26

<211> 28

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer

<400> 26

caccctagg taggcatccc cagcatgc

28

<210> 27
 <211> 8914
 <212> DNA
 <213> Artificial sequence

<220>
 <223> pAdenovator-CMVIntron(Transfer Vector)

<400> 27
 taacatcatc aataatatac cttatTTTTgg attgaagcca atatgataat gaggggggtgg 60
 agtttTgtgac gtggcgcggg gcgtgggaac ggggcgggtg acgtagtagt gtggcggaag 120
 tgtgatgttg caagtgtggc ggaacacatg taagcgacgg atgtggcaaa agtgacgttt 180
 ttggtgtgCG cCGgtgtaca caggaagtga caattttcgc gcggttttag gcggatgttg 240
 tagtaaattt gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga 300
 ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg taatactgcg atctatacat 360
 tgaatcaata ttggcaatta gccatattag tcattgggtta tatagcataa atcaatattg 420
 gctattggcc attgcatacg ttgtatctat atcataatat gtacatttat attggctcat 480
 gtccaatatg accgccatgt tgacattgat tattgactag gtcaatattg gccattagcc 540
 atattattca ttggttatat agcataaatc aatattggct attggccatt gcatacgttg 600
 tatctatatc ataatatgta ctttatattt ggctcatgtc caatatgacc gccatgtttg 660
 cattgattat tgactagtta ttaatagtaa tcaattacgg ggtcattagt tcatagccca 720
 tatatggagt tccgcgttac ataacttacg gtaaattggc cgcttggtg accgccaac 780
 gacccccgcc cattgacgtc aataatgacg tatgttccca tagtaacgcc aatagggact 840
 ttccattgac gtcaatgggt ggagtattta cggtaaactg cccacttggc agtacatcaa 900
 gtgtatcata tgccaagtcc gccccctatt gacgtcaatg acggtaaattg gcccgcctgg 960
 cattatgccc agtacatgac cttacgggac tttcctactt ggcagtacat ctacgtatta 1020
 gtcacgcta ttaccatggt gatgcggttt tggcagtaca ccaatgggcg tggatagcgg 1080
 tttgactcac ggggatttcc aagtctccac cccattgacg tcaatgggag tttgttttgg 1140
 cacaaaatc aacgggactt tccaaaatgt cgtaataacc ccgccccgtt gacgcaaattg 1200
 ggcggtaggc gtgtacggtg ggaggtctat ataagcagag ctcgtttagt gaaccgtcag 1260
 atcactagaa gctttattgc ggtagtttat cacagttaaa ttgctaacgc agtcagtgct 1320
 tctgacacaa cagtctcgaa cttaagctgc agaagttggt cgtgaggcac tgggcaggta 1380
 agtatcaagg ttacaagaca ggtttaagga gaccaataga aactgggctt gtcgagacag 1440
 agaagactct tgcgtttctg ataggcacct attgggtctta ctgacatcca ctttgccttt 1500
 ctctccacag gtgtccactc ccagttcaat tacagctctt aaggctagag tacttaatac 1560
 gactcactat aggctagcct cgagaattca cgcgtggtac cgagctcgga tccactagtc 1620
 cagtgtggtg gaattgccct taagggcaat tctgcagata tccagcacag tggcggccgc 1680

tcgagtctag	agggcccgcg	gttcgaaggt	aagcctatcc	ctaaccctct	cctcgggtctc	1740
gattctacgc	gtaccgggtca	tcatcaccat	caccattgag	ttcaaaccgc	ctgatcagcc	1800
tcgactgtgc	cttctagttg	ccagccatct	gttgtttgcc	cctcccccg	gccttccttg	1860
accctggaag	gtgccactcc	cactgtcctt	tcctaataaa	atgaggaaat	tgcatacgcat	1920
tgtctgagta	gggtgtcattc	tattctgggg	gggtgggggtg	ggcaggacag	caagggggag	1980
gattgggaag	acaatagcag	gcatgctggg	gatgcctacc	tagtcactac	tctgtgctat	2040
gggtgttcaat	gcttttcaag	atacccggt	catatgaaac	ggcatgactt	tttcaagagt	2100
gccatgccc	aagggttatgt	acaggaaagg	accatcttct	tcaaagatga	cggcaactac	2160
aagacacgtg	ctgaagtcaa	gtttgaaggt	gatacccttg	ttaatagaat	cgagttaaaa	2220
ggatttgact	tcaaggaaga	tggcaacatt	ctgggacaca	aattggaata	caactataac	2280
tcacacaatg	tatacatcat	ggcagacaaa	caaaagaatg	gaatcaaagt	gaacttcaag	2340
acccgccaca	acattgaaga	tggaaagcgt	caactagcag	accattatca	acaaaatact	2400
ccaattggcg	atggccctgt	ccttttacca	gacaaccatt	acctgtccac	acaatctgcc	2460
ctttcgaaag	atcccaacga	aaagagagac	cacatgggtcc	ttcttgagtt	tgtaacagct	2520
gctgggatta	cacatggcat	ggatgaactg	tacaactgag	gatccccga	cctcgacctc	2580
tggctaataa	aggaaattta	ttttcattgc	aatagtgtgt	tggaaatttt	tgtgtctctc	2640
actcggaagg	acatatggga	gggcaaatca	tttgggtcgag	atccctcgga	gatcggatct	2700
gggcgtgggt	aagggtggga	aagaatatat	aagggtgggg	tcttatgtag	ttttgtatct	2760
gttttgcagc	agccgccgcc	gccatgagca	ccaactcgtt	tgatggaagc	attgtgagct	2820
catatttgac	aacgcgcatg	cccccatggg	ccgggggtgcg	tcagaatgtg	atgggctcca	2880
gcattgatgg	tcgccccgtc	ctgcccgcga	actctactac	cttgacctac	gagaccgtgt	2940
ctggaacgcc	gttggagact	gcagcctccg	ccgccgcttc	agccgctgca	gccaccgccc	3000
gcgggattgt	gactgacttt	gctttcctga	gcccgccttc	aagcagtgca	gcttcccgtt	3060
catccgcccg	cgatgacaag	ttgacggctc	ttttggcaca	attggattct	ttgacccggg	3120
aacttaatgt	cgtttctcag	cagctgttgg	atctgcgcca	gcaggtttct	gccctgaagg	3180
cttcctcccc	tcccaatgcg	gtttaaaaca	taaataaaaa	accagactct	gtttggattt	3240
ggatcaagca	agtgtcttgc	tgtctttatt	taggggtttt	gcgcgcgcgg	taggcccggg	3300
accagcggtc	tcggtcgttg	agggtcctgt	gtattttttc	caggacgtgg	taaagggtgac	3360
tctggatggt	cagatacatg	ggcataagcc	cgtctctggg	gtggaggtag	caccactgca	3420
gagcttcatg	ctgcgggggtg	gtgttgtaga	tgatccagtc	gtagcaggag	cgctgggcgt	3480
gggtgcctaaa	aatgtctttc	agtagcaagc	tgattgccag	gggcaggccc	ttggtgtaag	3540
tgttttacaaa	gcggttaagc	tgggatgggt	gcatacgtgg	ggatatgaga	tgcatacttg	3600
actgtatttt	taggttggct	atgttcccag	ccatatccct	ccggggattc	atgttgtgca	3660
gaaccaccag	cacagtgtat	ccggtgcact	tgggaaattt	gtcatgtagc	ttagaaggaa	3720

atgcgtggaa	gaacttggag	acgcccttgt	gacctccaag	attttccatg	cattcgtcca	3780
taatgatggc	aatgggcca	cgggcggcgg	cctgggcgaa	gatatttctg	ggatcactaa	3840
cgtcatagtt	gtgttccagg	atgagatcgt	cataggccat	ttttacaaag	cgcgggcgga	3900
gggtgccaga	ctgcggtata	atggttccat	ccggcccagg	ggcgtagtta	ccctcacaga	3960
tttgcatttc	ccacgctttg	agttcagatg	gggggatcat	gtctacctgc	ggggcgatga	4020
agaaaacggt	ttccggggta	ggggagatca	gctgggaaga	aagcaggttc	ctgagcagct	4080
gcgacttacc	gcagccggtg	ggcccgtaaa	tcacacctat	taccgggtgc	aactggtagt	4140
taagagagct	gcagctgccg	tcattccctga	gcaggggggc	cacttcgtta	agcatgtccc	4200
tgactcgc	gttttccctg	accaaattccg	ccagaaggcg	ctcgccgccc	agcgatagca	4260
gttcttgcaa	ggaagcaaag	tttttcaacg	gtttgagacc	gtccgccgta	ggcatgcttt	4320
tgagcgtttg	accaagcagt	tccaggcggt	cccacagctc	ggtcacctgc	tctacggcat	4380
ctcgatccag	catatctcct	cgtttcgcgg	gttggggcgg	ctttcgctgt	acggcagtag	4440
tcggtgctcg	tccagacggg	ccagggtcat	gtctttccac	gggcgcaggg	tcctcgtcag	4500
cgtagtctgg	gtcacggtga	aggggtgcgc	tccgggctgc	gcgctggcca	gggtgcgctt	4560
gaggctggtc	ctgctgggtg	tgaagcgctg	ccggtcttcg	ccctgcgcgt	cgccaggta	4620
gcatttgacc	atggtgtcat	agtccagccc	ctccgcggcg	tggcccttgg	cgcgagctt	4680
gcccttgagg	gaggcgccgc	acgaggggca	gtgcagactt	ttgagggcgt	agagcttggg	4740
cgcgagaaat	accgattccg	gggagtaggc	atccgcgccg	caggccccgc	agacggtctc	4800
gcattccacg	agccagggtga	gctctggccg	tccggggtca	aaaaccagg	ttcccccatg	4860
ctttttgatg	cgtttcttac	ctctggtttc	catgagccgg	tgtccacgct	cggtgacgaa	4920
aaggctgtcc	gtgtccccgt	atacagactt	gagagggagt	ttaaacgaat	tcaatagctt	4980
gttgcatggg	cggcgatata	aaatgcaagg	tgctgtctaa	aaaatcaggc	aaagcctcgc	5040
gcaaaaaaga	aagcacatcg	tagtcatgct	catgcagata	aaggcaggta	agctccggaa	5100
ccaccacaga	aaaagacacc	atttttctct	caaacatgtc	tgcgggtttc	tgcataaaca	5160
caaaataaaa	taacaaaaaa	acattttaa	attagaagcc	tgtcttaca	caggaaaaac	5220
aacccttata	agcataagac	ggactacggc	catgccggcg	tgaccgtaaa	aaaactggtc	5280
accgtgatta	aaaagcacca	ccgacagctc	ctcggtcatg	tccggagtca	taatgtaaga	5340
ctcggtaaac	acatcagggt	gattcatcgg	tcagtgtctaa	aaagcgaccg	aaatagcccg	5400
ggggaataca	taccgcagg	cgtagagaca	acattacagc	ccccatagga	ggtataacaa	5460
aattaatagg	agagaaaaac	acataaacac	ctgaaaaacc	ctcctgccta	ggcaaaatag	5520
caccctccc	ctccagaaca	acatacagcg	cttcacagcg	gcagcctaac	agtcagcctt	5580
accagtaaaa	aagaaaacct	attaaaaaaa	caccactcga	cacggcacca	gctcaatcag	5640
tcacagtgt	aaaaagggcc	aagtgcagag	cgagtatata	taggactaaa	aaatgacgta	5700

acggttaaag	tccacaaaaa	acaccagaa	aaccgcacgc	gaacctacgc	ccagaaacga	5760
aagccaaaaa	accacaaact	tcctcaaate	gtcacttccg	ttttcccacg	ttacgtaact	5820
tcccatttta	agaaaactac	aattcccaac	acatacaagt	tactccgccc	taaaacctac	5880
gtcacccgcc	ccgttcccac	gccccgcgcc	acgtcacaaa	ctccaccccc	tcattatcat	5940
attggcttca	atccaaaata	aggtatatta	ttgatgatgt	taattaacat	gcatggatcc	6000
atatgcggtg	tgaaataccg	cacagatgcg	taaggagaaa	ataccgcac	aggcgctctt	6060
ccgcttcctc	gtcactgac	tcgctgcgct	cggtcgttcg	gctgcggcga	gcggtatcag	6120
ctcactcaaa	ggcggtata	cggttatcca	cagaatcagg	ggataacgca	ggaaagaaca	6180
tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcggtg	ctggcgtttt	6240
tccataggct	ccgccccct	gacgagcatc	acaaaaatcg	acgctcaagt	cagagggtggc	6300
gaaacccgac	aggactataa	agataccagg	cgtttcccc	tggaagctcc	ctcgtgcgct	6360
ctcctgttcc	gaccctgccg	cttaccggat	acctgtccgc	ctttctccct	tcgggaagcg	6420
tggcgctttc	tcatagctca	cgctgtagg	atctcagttc	gggtgtaggtc	gttcgctcca	6480
agctgggctg	tgtgcacgaa	cccccgctt	agcccgaccg	ctgcgcctta	tccggttaact	6540
atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	actggcagca	gccactggta	6600
acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	tggtggccta	6660
actacggcta	cactagaagg	acagtatttg	gtatctgcgc	tctgctgaag	ccagttacct	6720
tcggaaaaag	agttggtagc	tcttgatccg	gcaaacaac	caccgctggg	agcgggtggt	6780
tttttgtttg	caagcagcag	attacgcgca	gaaaaaaagg	atctcaagaa	gaccccttga	6840
tcttttctac	ggggtctgac	gtcagtgga	acgaaaactc	acgttaaggg	attttggtca	6900
tgagattatc	aaaaaggatc	ttcacctaga	tcctttttaa	ttaaaaatga	agttttaaat	6960
caatctaaag	tatatatgag	taaacttgg	ctgacagtta	ccaatgctta	atcagtgagg	7020
cacctatctc	agcgatctgt	ctatttcggt	catccatagt	tgctgactc	cccgtcgtgt	7080
agataactac	gatacgggag	ggcttaccat	ctggccccag	tgctgcaatg	ataccgcgag	7140
accacgctc	accggctcca	gatttatcag	caataaacca	gccagccgga	agggccgagc	7200
gcagaagtgg	tcctgcaact	ttatccgcct	ccatccagtc	tattaattgt	tgccgggaag	7260
ctagagtaag	tagttcgcca	gttaatagtt	tgcgcaacgt	tggtgccatt	gctgcagcca	7320
tgagattatc	aaaaaggatc	ttcacctaga	tccttttcac	gtagaaagcc	agtccgcaga	7380
aacggtgctg	accccggatg	aatgtcagct	actgggctat	ctggacaagg	gaaaacgcaa	7440
gcgcaaagag	aaagcaggta	gcttgcaagt	ggcttacatg	gcgatagcta	gactgggcgg	7500
ttttatggac	agcaagcgaa	ccggaattgc	cagctggggc	gccctctggt	aagggtggga	7560
agccctgcaa	agtaaaactg	atggctttct	cgccgccaag	gatctgatgg	cgaggggat	7620
caagctctga	tcaagagaca	ggatgaggat	cgtttcgcat	gattgaacaa	gatggattgc	7680
acgcaggttc	tccggccgct	tgggtggaga	ggctattcgg	ctatgactgg	gcacaacaga	7740

caatcggtctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggagc ccggttcttt	7800
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca agacgaggca gcgcggctat	7860
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgct actgaagcgg	7920
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg	7980
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc	8040
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga	8100
tggagccgg tcttgctgat caggatgatc tggacgaaga gcatcagggg ctgcgccag	8160
ccgaactgtt cgccaggctc aaggcgagca tgcccgacgg cgaggatctc gtcgtgacct	8220
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg	8280
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata	8340
ttgctgaaga gcttggcggc gaatgggctg accgcttcct cgtgctttac ggtatcgccg	8400
ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagtcttc tgaattttgt	8460
taaaattttt gttaaatcag ctcatTTTTT aaccaatagg ccgaaatcgg caacatccct	8520
tataaatcaa aagaatagac cgcgataggg ttgagtgttg ttccagtttg gaacaagagt	8580
ccactattaa agaacgtgga ctccaacgtc aaagggcgaa aaaccgtcta tcagggcgat	8640
ggcccactac gtgaaccatc acccaaatca agttttttgc ggtcgagggt ccgtaaagct	8700
ctaaatcgga accctaaagg gagccccga tttagagctt gacggggaaa gccggcgaac	8760
gtggcgagaa aggaagggaa gaaagcgaaa ggagcgggag ctagggcgct ggcaagtgt	8820
gcggtcacgc tgcgcgtaac caccacacc gcgcgcttaa tgcgccgcta cagggcgct	8880
ccattcgcca ttcaggatcg aattaattct taat	8914