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

(19) World Intellectual Property Organization
International Bureau

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

PCT

(10) International Publication Number
WO 2007/098548 A1

(51) International Patent Classification:
C07K 14/475 (2006.01) *C07K 16/22* (2006.01)
C07K 14/71 (2006.01) *C12N 15/19* (2006.01)
C07K 19/00 (2006.01) *A61P 35/00* (2006.01)
A61K 38/18 (2006.01) *C07K 16/28* (2006.01)

(21) International Application Number:
PCT/AU2007/000245

(22) International Filing Date: 1 March 2007 (01.03.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2006901037 1 March 2006 (01.03.2006) AU
2006903918 20 July 2006 (20.07.2006) AU

(71) Applicant *(for all designated States except US)*: **APOLLO LIFE SCIENCES LIMITED** [AU/AU]; Level 1, 147 Queen Street, Beaconsfield, New South Wales 2015 (AU).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **PRIEST, John, D.** [AU/AU]; 12b Grafton Street, Balmain, New South Wales 2041 (AU). **WATTS, Alan, D.** [AU/AU]; 31 Undercliffe Street, Dee Why, New South Wales 2099 (AU). **WHITTAKER, Jason, S.** [AU/AU]; 24 Maddison Street, Redfern East, New South Wales 2016 (AU). **DOMAGALA, Teresa, A.** [AU/AU]; 14/30 Garden Street, Alexandria, New South Wales 2015 (AU). **SIMPSON, Raina, J.** [AU/AU]; 10 Snowgum Street, Acacia Gardens, New South Wales 2763 (AU). **LIDDELL, Catherine, A.** [AU/AU]; 6 Caroline Street, Kingsgrove, New South Wales 2208 (AU). **PILKINGTON, Glenn, R.** [AU/AU]; 6 Allara Street, Rye, Victoria 3941 (AU).

(74) Agents: **HUGHES, E., John, L.** et al.; DAVIES COLLISON CAVE, 1 Nicholson Street, Melbourne, Victoria 3000 (AU).

(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, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, 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.

(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, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, 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).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** A MOLECULE AND CHIMERIC MOLECULES THEREOF

(57) **Abstract:** The present invention relates generally to the fields of proteins, diagnostics, therapeutics and nutrition. More particularly, the present invention provides an isolated protein molecule that comprises a mast/stem cell growth factor receptor (SCFR) or chimeric molecules thereof comprising at least a portion of the SCFR molecule, such as SCFR-Fc, or a ligand thereof, such as stem cell factor (SCF); wherein the protein or chimeric molecule thereof has a profile of measurable physiochemical parameters, wherein the profile is indicative of, associated with or forms the basis of one or more pharmacological traits. The present invention further contemplates the use of the isolated protein or chimeric molecule thereof in a range of diagnostic, prophylactic, therapeutic, nutritional and/or research applications.

WO 2007/098548 A1

A MOLECULE AND CHIMERIC MOLECULES THEREOF

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

The present invention relates generally to the fields of proteins, diagnostics, therapeutics and nutrition. More particularly, the present invention provides an isolated protein molecule that comprises a mast/stem cell growth factor receptor (SCFR) or chimeric molecules thereof comprising at least a portion of the SCFR molecule, such as SCFR-Fc, or a ligand thereof, such as stem cell factor (SCF); wherein the protein or chimeric molecule thereof has a profile of measurable physiochemical parameters, wherein the profile is indicative of, associated with or forms the basis of one or more pharmacological traits. The present invention further contemplates the use of the isolated protein or chimeric molecule thereof in a range of diagnostic, prophylactic, therapeutic, nutritional and/or research applications.

DESCRIPTION OF THE PRIOR ART

Reference to any prior art in this specification is not, and should not be taken as an acknowledgment or any form of suggestion that this prior art forms a part of the common general knowledge.

Tyrosine kinase receptors and their ligands can exert a variety of effects on cells including survival, proliferation, migration and differentiation, thereby regulating various processes such as hematopoiesis, angiogenesis, wound healing the development of the central nervous system, oncogenesis and inflammation.

The tyrosine kinase class of receptors include mast/stem cell binding factor receptor, also known as stem cell factor receptor, c-KIT or CD117, and is a proto-oncogene that encodes

a type III receptor tyrosine kinase. Stem Cell Factor Receptor (SCFR) is related to the platelet-derived growth factor (PDGF)/colony-stimulating factor 1 (CSF-1) (c-fms) receptor subfamily. Structurally, SCFR is divided into an extracellular and intracellular region. The intra-cellular region is comprised of a juxtamembrane domain (JMD), a kinase domain, which is split by a kinase insert (KI) in the ATP-binding region and the phosphotransferase domain, and a carboxyl tail. The extracellular domain of SCFR contains 5 Ig-like loops and the mature protein has a molecular weight of 145 kDa.

SCFR is thought to play an important role in hematopoiesis, spermatogenesis, melanogenesis and, more recently, in carcinogenesis. SCFR is expressed by hematopoietic cells in the embryonic liver throughout development, and by more committed progenitors, such as myeloid, erythroid, megakaryocytic, natural killer, and dendritic progenitor cells, as well as mature mast cells and pro-B and -T cells. Binding of SCFR to its ligand (stem cell factor; SCF) induces receptor dimerization, autophosphorylation and subsequent signal transduction. In addition, receptor binding generates conformational changes in SCFR that increase receptor-receptor interactions and lead to further stabilization of the dimeric complex. After initiation of the kinase activity, tyrosine residues located primarily outside the kinase domain are phosphorylated and serve as docking sites for signal transduction molecules with phosphotyrosine binding or Src homology 2 (SH2) domains.

The expression of SCFR is predominately limited to normal embryonic development, however, recent studies have implicated SCF-SCFR signaling in carcinogenesis. Inappropriate expression or activation of SCFR can be classified as loss-of-function mutations or gain-of-function mutations and are associated with a variety of diseases in humans. Autosomal-dominant piebaldism is associated with loss-of-function mutations in SCFR. This syndrome results in deafness, megacolon, and abnormalities in pigmentation of skin and hair. Activating mutations in SCFR have been identified primarily in the JMD and in the second part of the kinase domain, and are associated with gastrointestinal stromal cell tumors and mastocytosis, respectively. There are also reports of activating mutations in some forms of germ cell tumors and core binding factor leukemias. Overexpression of SCFR has previously been documented in myeloid leukemia,

neuroblastoma, breast tumor, colon tumors, gynecological tumors, testicular germ cell tumors and small cell lung carcinoma (SCLC) as well as in other conditions such as acne.

Stem Cell Factor (SCF) is also known as C-kit ligand, mast cell growth factor (MGF) and steel factor. SCF is a glycoprotein exhibiting 5 potential N-linked glycosylation sites and is expressed as a type I membrane protein. A soluble form of SCF is produced by proteolytic cleavage of the extracellular domain.

SCF is expressed primarily by fibroblasts, however, expression has also been detected in keratinocytes, mature granulocytes, Sertoli cells and bone marrow stromal cells. Functionally, SCF is involved in steady state maintenance of hematopoiesis and mediates the proliferation of myeloid, erythroid and lymphoid progenitors in bone marrow cultures and has been shown to act synergistically with colony stimulating factors such as G-CSF and erythropoietin as well as interleukins. Treatment of non-human primates with SCF resulted in an expansion of mast cells as well as an increase in megakaryocytes and circulating platelet counts. SCF factor is also involved in gametogenesis, melanogenesis and the membrane bound form of SCF mediates cell-cell adhesion.

SCF contributes to the generation and survival of mast cells from CD34+ progenitor cells and plays a role in mast cell degranulation, resulting in the release of histamine, proinflammatory cytokines and chemokines. Furthermore, SCF also induces eosinophil activation. This suggests SCF is involved in allergic disease states and elevated levels of SCF are detected in inflammatory conditions in human and animal models such as asthma.

The biological effector functions exerted by SCFR or chimeric SCFR molecules, or a ligands thereof, such as SCF, may have significant potential as therapeutic agents to modulate physiological processes. However, minor changes to the molecule such as primary, secondary, tertiary or quaternary structure and co- or post-translational modification patterns can have a significant impact on the activity, secretion, antigenicty and clearance of the protein. It is possible, therefore, that the proteins can be generated with specific primary, secondary, tertiary or quaternary structure, or co- or post-translational structure or make-up that confer unique or particularly useful properties.

There is consequently a need to evaluate the physiochemical properties of proteins under different conditions of production to determine whether they have useful physiochemical characteristics or other pharmacological traits.

The problem to date is that production of commercially available proteins are carried out in cells derived from species that are evolutionary distant to humans, cells such as bacteria, yeast, fungi, and insect. These cells express proteins that either lack glycosylation or exhibit glycosylation repertoires that are distinct to human cells and this impacts substantially on their clinical utility. For example, proteins expressed in yeast or fungi systems such as *Aspergillus* possess a high density of mannose which makes the protein therapeutically useless (Herscovics *et al. FASEB J* 7:540-550, 1993).

Even in non-human mammalian expression systems such as Chinese hamster ovary (CHO) cells, significant differences in the glycosylation patterns are documented compared with that of human cells. For example, most mammals, including rodents, express the enzyme (α 1,3) galactotransferase, which generates Gal (α 1,3)-Gal (β 1,4)-GlcNAc oligosaccharides on glycoproteins. However, in humans, apes and Old World monkeys, the expression of this enzyme has become inactivated through a frameshift mutation in the gene (Larsen *et al. J Biol Chem 265*:7055-7061, 1990). Although most of the CHO cell lines used for recombinant protein synthesis, such as Dux-B11, have inactivated the gene expressing (α 1,3) Galactotransferase, they still lack a functional (α 2, 6) sialyltransferase enzyme for synthesis of (α 2, 6)-linked terminal sialic acids which are present in human cells. Furthermore, the sialic acid motifs present on CHO cell expressed glycoproteins proteins are prone to degradation by a CHO cell endogenous sialidase (Gramer *et al. Biotechnology (N.Y.) 13(7)*:692-8, 1995).

As a result, proteins produced from these non-human expression systems will exhibit physiochemical and pharmacological characteristics such as half-life, antigenicity, stability and functional potency that are distinct from human cell-derived proteins.

The recent advancement of stem cell technology has substantially increased the potential for utilizing stem cells in applications such as transplantation therapy, drug screening,

toxicology studies and functional genomics. However, stem cells are routinely maintained in culture medium that contains non-human proteins and are therefore not suitable for clinical applications due to the possibility of contamination with non-human infectious material. Furthermore, culturing of stem cells in non-human derived media may result in the incorporation of non-human carbohydrate moieties thus compromising transplant application (Martin *et al. Nature Med 11(2)*:228-232, 2005). Hence, the use of specific human-derived proteins in the maintenance and/or differenttiation of stem cells will ameliorate the incorporation of xenogeneic proteins and enhance stem cell clinical utility.

Accordingly, there is a need to develop proteins and their receptors which have particularly desired physiochemical and pharmacological properties for use in diagnostic, prophylactic, therapeutic and/or nutritional research applications and the present invention provides proteins belonging to the SCFR family, or a ligand thereof, such as SCF, and related proteins for clinical, commercial and research applications.

SUMMARY OF THE INVENTION

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group of elements or integers but not the exclusion of any other element or integer or group of elements or integers.

Nucleotide and amino acid sequences are referred to by a sequence identifier number (SEQ ID NO:). The SEQ ID NOs: correspond numerically to the sequence identifiers <400>1 (SEQ ID NO:1), <400>2 (SEQ ID NO:2), etc. A summary of the sequence identifiers is provided in Table 1. A sequence listing is provided after the claims.

The present invention relates generally to an isolated protein or chimeric molecule thereof in or related to the SCFR family, or a ligand thereof, such as SCF, comprising a profile of physiochemical parameters, wherein the profile is indicative of, associated with, or forms the basis of one or more distinctive pharmacological traits. More particularly, the present invention provides an isolated protein or chimeric molecule thereof selected from the list of SCFR, SCFR-Fc, SCF, comprising a physiochemical profile comprising a number of measurable physiochemical parameters, $\{[P_x]_1, [P_x]_2, \ldots [P_x]_n,\}$, wherein P_x represents a measurable physiochemical parameter and "n" is an integer ≥ 1, wherein each parameter between and including $[P_x]_1$ to $[P_x]_n$ is a different measurable physiochemical parameter, wherein the value of any one or more of the measurable physiochemical characteristics is indicative of, associated with, or forms the basis of, a distinctive pharmacological trait, T_y, or series of distinctive pharmacological traits $\{[T_y]_1, [T_y]_2, \ldots . [T_y]_m\}$ wherein T_y represents a distinctive pharmacological trait and m is an integer ≥ 1 and each of $[T_y]_1$ to $[T_y]_m$ is a different pharmacological trait.

As used herein the term "distinctive" with regard to a pharmacological trait of a protein or chimeric molecule thereof of the present invention refers to one or more pharmacological traits of a protein or chimeric molecule thereof which are distinctive for the particular physiochemical profile. In a particular embodiment, one or more of the pharmacological traits of an isolated protein or chimeric molecule thereof is different from, or distinctive

relative to a form of the same protein or chimeric molecule thereof produced in a prokaryotic or lower eukaryotic cell or even a higher eukaryotic cell of a non-human species. In another embodiment, the pharmacological traits of a subject isolated protein or chimeric molecule thereof contribute to a desired functional outcome. As used herein, the term "measurable physiochemical parameters" or Px refers to one or more measurable characteristics of the isolated protein or chimeric molecule thereof. In a particular embodiment of the present invention, the measurable physiochemical parameters of a subject isolated protein or chimeric molecule thereof contribute to or are otherwise responsible for the derived pharmacological trait, Ty.

An isolated protein or chimeric molecule of the present invention comprises physiochemical parameters (P_x) which taken as a whole define protein molecule or chimeric molecule. The physiochemical parameters may be selected from the group consisting of apparent molecular weight (P_1), isoelectric point (pI) (P_2), number of isoforms (P_3), relative intensities of the different number of isoforms (P_4), percentage by weight carbohydrate (P_5), observed molecular weight following N-linked oligosaccharide deglycosylation (P_6), observed molecular weight following N-linked and O-linked oligosaccharide deglycosylation (P_7), percentage acidic monosaccharide content (P_8), monosaccharide content (P_9), sialic acid content (P_{10}), sulfate and phosphate content (P_{11}), Ser/Thr : GalNAc ratio (P_{12}), neutral percentage of N-linked oligosaccharide content (P_{13}), acidic percentage of N-linked oligosaccharide content (P_{14}), neutral percentage of O-linked oligosaccharide content (P_{15}), acidic percentage of O-linked oligosaccharide content (P_{16}), ratio of N-linked oligosaccharides (P_{17}), ratio of O-linked oligosaccharides (P_{18}), structure of N-linked oligosaccharide fraction (P_{19}), structure of O-linked oligosaccharide fraction (P_{20}), position and make up of N-linked oligosaccharides (P_{21}), position and make up of O-linked oligosaccharides (P_{22}), co-translational modification (P_{23}), post-translational modification (P_{24}), acylation (P_{25}), acetylation (P_{26}), amidation (P_{27}), deamidation (P_{28}), biotinylation (P_{29}), carbamylation or carbamoylation (P_{30}), carboxylation (P_{31}), decarboxylation (P_{32}), disulfide bond formation (P_{33}), fatty acid acylation (P_{34}), myristoylation (P_{35}), palmitoylation (P_{36}), stearoylation (P_{37}), formylation (P_{38}), glycation (P_{39}), glycosylation (P_{40}), glycophosphatidylinositol anchor (P_{41}), hydroxylation (P_{42}), incorporation of selenocysteine (P_{43}), lipidation (P_{44}), lipoic acid addition (P_{45}),

methylation (P_{46}), N- or C-terminal blocking (P_{47}), N- or C-terminal removal (P_{48}), nitration (P_{49}), oxidation of methionine (P_{50}), phosphorylation (P_{51}), proteolytic cleavage (P_{52}), prenylation (P_{53}), farnesylation (P_{54}), geranyl geranylation (P_{55}), pyridoxal phosphate addition (P_{56}), sialylation (P_{57}), desialylation (P_{58}), sulfation (P_{59}), ubiquitinylation or ubiquitination (P_{60}), addition of ubiquitin-like molecules (P_{61}), primary structure (P_{62}), secondary structure (P_{63}), tertiary structure (P_{64}), quaternary structure (P_{65}), chemical stability (P_{66}), thermal stability (P_{67}). A list of these parameters is summarized in Table 2.

In an embodiment, an SCF of the present invention is characterized by a profile of one or more of the following physiochemical parameters (P_x) and pharmacological traits (T_y), comprising:

- an apparent molecular weight (P_1) of about 1 to 250, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250 kDa and in one embodiment, 18 to 55 kDa;
- a pI (P_2) range of about 2 to about 14 such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and in one embodiment, 3 to 7.5;
- about 2 to 100 isoforms (P_3), such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 isoforms and in one embodiment 15 to 86 isoforms;
- a percentage by weight carbohydrate (P_5) of about 0 to 99%, such as 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% and in one embodiment, 0 to 62%;

- an observed molecular weight after the N-linked oligosaccharides are removed (P_6) of about 18 to 40 kDa;
- an observed molecular weight after the N-linked and O-linked oligosaccharides are removed (P_7) of about 18 to 35 kDa;
- a site of N-glycosylation (P_{21}) which includes N-145 (numbering from the start of the signal sequence) identified by PMF after PNGase treatment;
- an immunoreactivity profile (T_{13}) that is distinct from that of a human SCF molecule expressed in a non-human cell system, and in one embodiment, the protein concentration of the SCF of the present invention is underestimated when assayed using a quantitative immunoassay which includes a protein standard of a human SCF molecule expressed in *E. coli* cells;
- a biological activity that is distinct from that of a human SCF expressed in a non-human cell system, and in one embodiment, the ability of SCF of the present invention to induce proliferation (T_{32}) in M-07e cells is 1.5-3.0 fold more potent than a human SCF expressed in *E. coli* cells;
- a biological activity that is distinct from that of a human SCF expressed in a non-human cell system, and in one embodiment, the ability of SCF of the present invention to produce a greater yield of viable cells (T_{73}) from $CD34^+$ human haematopoietic cells than that produced by treatment with rhSCF expressed in *E. coli* cells;
- a biological activity that is distinct from that of a human SCF expressed in a non-human cell system, and in one embodiment, the ability of SCF of the present invention to induce a 2-fold greater activation of STAT5 (T_{45}) in human $CD34^+$ haematopoietic stem cells than that induced by a rhSCF expressed in *E. coli* cells;
- a biological activity that is distinct from that of a human SCF expressed in a non-human cell system, and in one embodiment, the ability of SCF of the present invention to induce a 4.5-fold greater activation of Akt (T_{47}) in human $CD34^+$ haematopoietic stem cells than that induced by a rhSCF expressed in *E. coli* cells.

In another embodiment, SCFR-Fc of the present invention is characterized by a profile of one or more of the following physiochemical parameters (P_x) and pharmacological traits (T_y), comprising:

- an apparent molecular weight (P_1) of about 1 to 250, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250 kDa and in one embodiment, 55 to 180 kDa;
- a pI (P_2) range of about 2 to about 14 such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and in one embodiment, 3.5 to 8.0;
- about 2 to 150 isoforms (P_3), such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150 isoforms and in one embodiment 10 to 109 isoforms;
- a percentage by weight carbohydrate (P_5) of about 0 to 99%, such as 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% and in one embodiment, 0 to 54%;
- an observed molecular weight of the molecule after the N-linked oligosaccharides are removed (P_6) of about 65 to 135 kDa;
- an observed molecular weight of the molecule after the N-linked and O-linked oligosaccharides are removed (P_7) of about 65 to 120 kDa;
- sites of N-glycosylation (P_{21}) which include N-130 and N-613 (numbering from the start of the signal sequence);
- an ability to inhibit rh SCF- rh SCF (expressed from human cells) induced activation of STAT3 (T_{45}) in M-07e cells;

- a biological activity that is distinct from that of a rh SCFR molecule expressed in a non-human cell system, and in one embodiment, a greater ability of SCFR-Fc of the present invention to inhibit SCF induced activation of p42 and p44 MAPKs (T_{46}) in M-07e cells than a rh SCFR molecule expressed from *Sf*21 insect cells;
- a biological activity that is distinct from that of a rh SCFR molecule expressed in a non-human cell system, and in one embodiment, a greater ability of SCFR-Fc of the present invention to inhibit SCF induced activation of p42 and p44 MAPKs (T_{46}) in human $CD34^+$ haematopoietic stem cells than a rh SCFR molecule expressed from *Sf*21 insect cells;

In another embodiment, the present invention contemplates an isolated form of SCFR such as a soluble extra-cellular domain of SCFR. Such a molecule is also referred to herein as either soluble SCFR or SCFR-ECD of the present invention.

In an embodiment, a SCFR-ECD of the present invention is characterized by a profile of one or more of the following physiochemical parameters (P_x) and pharmacological traits (T_y), comprising:

- an apparent molecular weight (P_1) of about 1 to 250, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250 kDa and in one embodiment, 55 to 150 kDa;
- a pI (P_2) range of about 2 to about 14 such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and in one embodiment, 3.5 to 7.5;
- about 2 to 100 isoforms (P_3), such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 isoforms and in one embodiment 15 to 63 isoforms;

- a percentage by weight carbohydrate (P_5) of about 0 to 99%, such as 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% and in one embodiment, 0 to 63%;
- an observed molecular weight of the molecule after the N-linked oligosaccharides are removed (P_6) of about 42 to 85 kDa;
- an observed molecular weight of the molecule after the N-linked and O-linked oligosaccharides are removed (P_7) of about 42 to 75 kDa;
- four sites of N-glycosylation (P_{21}) which include N-130 N-367, N-463 and N-486 (numbering from the start of the signal sequence);
- a biological activity that is distinct from that of a human SCFR-ECD expressed in a non-human cell system, and in one embodiment, a greater ability of SCFR-ECD of the present invention to inhibit SCF induced activation of p42 and p44 MAPKs (T_{46}) M-07e cells than a rh SCFR-ECD expressed from *Sf*21 insect cells;
- a biological activity that is distinct from that of a human SCFR-ECD expressed in a non-human cell system, and in one embodiment, a greater ability of SCFR-ECD of the present invention to inhibit SCF induced activation of p42 and p44 MAPKs (T_{46}) in human $CD34^+$ haematopoietic stem cells than a rh SCFR-ECD expressed from *Sf*21 insect cells.

In a particular embodiment, the present invention contemplates an isolated form of protein or chimeric molecule thereof in or related to the SCFR family, including SCFR-Fc, or a ligand thereof, such as SCF. An isolated protein or chimeric molecule of the present invention comprises distinctive pharmacological traits selected from the group comprising or consisting of therapeutic efficiency (T_1), effective therapeutic dose ($TCID_{50}$) (T_2), bioavailability (T_3), time between dosages to maintain therapeutic levels (T_4), rate of absorption (T_5), rate of excretion (T_6), specific activity (T_7), thermal stability (T_8), lyophilization stability (T_9), serum/plasma stability (T_{10}), serum half-life (T_{11}), solubility in blood stream (T_{12}), immunoreactivity profile (T_{13}), immunogenicity (T_{14}), inhibition by neutralizing antibodies (T_{14A}), side effects (T_{15}), receptor/ligand binding affinity (T_{16}),

receptor/ligand activation (T_{17}), tissue or cell type specificity (T_{18}), ability to cross biological membranes or barriers (i.e. gut, lung, blood brain barriers, skin etc) (T_{19}), angiogenic ability (T_{19A}), tissue uptake (T_{20}), stability to degradation (T_{21}), stability to freeze-thaw (T_{22}), stability to proteases (T_{23}), stability to ubiquitination (T_{24}), ease of administration (T_{25}), mode of administration (T_{26}), compatibility with other pharmaceutical excipients or carriers (T_{27}), persistence in organism or environment (T_{28}), stability in storage (T_{29}), toxicity in an organism or environment and the like (T_{30}).

In addition, the protein or chimeric molecule of the present invention may have altered biological effects on different cells types (T_{31}), including without being limited to human primary cells, such as lymphocytes, erythrocytes, retinal cells, hepatocytes, neurons, keratinocytes, endothelial cells, endodermal cells, ectodermal cells, mesodermal cells, epithelial cells, kidney cells, liver cells, bone cells, bone marrow cells, lymph node cells, dermal cells, fibroblasts, T-cells, B-cells, plasma cells, natural killer cells, macrophages, granulocytes, neutrophils, Langerhans cells, dendritic cells, eosinophils, basophils, mammary cells, lobule cells, prostate cells, lung cells, oesophageal cells, pancreatic cells, Beta cells (insulin secreting cells), hemangioblasts, muscle cells, oval cells (hepatocytes), mesenchymal cells, brain microvessel endothelial cells, astrocytes, glial cells, various stem cells including adult and embryonic stem cells, various progenitor cells; and other human immortal, transformed or cancer cell lines.

The biological effects on the cells include effects on proliferation (T_{32}), differentiation (T_{33}), apoptosis (T_{34}), growth in cell size (T_{35}), cytokine adhesion (T_{36}), cell adhesion (T_{37}), cell spreading (T_{38}), cell motility (T_{39}), migration and invasion (T_{40}), chemotaxis (T_{41}), cell engulfment (T_{42}), signal transduction (T_{43}), recruitment of proteins to receptors/ligands (T_{44}), activation of the JAK/STAT pathway (T_{45}), activation of the Ras-erk pathway (T_{46}), activation of the AKT pathway (T_{47}), activation of the PKC pathway (T_{48}), activation of the PKA pathway (T_{49}), activation of src (T_{50}), activation of fas (T_{51}), activation of TNFR (T_{52}), activation of NFkB (T_{53}), activation of p38MAPK (T_{54}), activation of c-fos (T_{55}), secretion (T_{56}), receptor internalization (T_{57}), receptor cross-talk (T_{58}), up or down regulation of surface markers (T_{59}), alteration of FACS front/side scatter profiles (T_{60}), alteration of subgroup ratios (T_{61}), differential gene expression (T_{62}), cell

necrosis (T_{63}), cell clumping (T_{64}), cell repulsion (T_{65}), binding to heparin sulfates (T_{66}), binding to glycosylated structures (T_{67}), binding to chondroitin sulfates (T_{68}), binding to extracellular matrix (such as collagen, fibronectin) (T_{69}), binding to artificial materials (such as scaffolds) (T_{70}), binding to carriers (T_{71}), binding to co-factors (T_{72}) the effect alone or in combination with other proteins on stem cell proliferation, differentiation and/or self-renewal (T_{73}) and the like. These are summarized in Table 3.

The present invention further provides a chimeric molecule comprising an isolated protein or a fragment thereof, such as an extra-cellular domain of a membrane bound protein, linked to the constant (Fc) or framework region of a human immunoglobulin via one or more protein linker. Such a chimeric molecule is also referred to herein as protein-Fc. Examples of such protein-Fc contemplated by the present invention include SCFR-Fc. Such protein-Fc has a profile of measurable physiochemical parameters indicative of or associated with one or more distinctive pharmacological traits of the isolated protein-Fc. Other chimeric molecules contemplated by the present invention include the protein or protein-Fc or a fragment thereof, linked to a lipid moiety such as a polyunsaturated fatty acid molecule. Such lipid moieties may be linked to an amino acid residue in the backbone of the molecule or to a side chain of such an amino acid residue.

The present invention further provides a chimeric molecule comprising an isolated protein or a fragment thereof, such as an extra-cellular domain of a membrane bound protein, linked to the constant (Fc) or framework region of a mammalian immunoglobulin *via* one or more protein linker. In another aspect, the mammal Fc or framework region of the immunoglobulin is derived from a mammal selected from the group consisting of primates, including humans, marmosets, orangutans and gorillas, livestock animals (e.g. cows, sheep, pigs, horses, donkeys), laboratory test animals (e.g. mice, rats, guinea pigs, hamsters, rabbits, companion animals (e.g. cats, dogs) and captured wild animals (e.g. rodents, foxes, deer, kangaroos). In another embodiment the Fc or framework region is a human immunoglobulin. In a particular embodiment the mammal is a human. Such a chimeric molecule is also referred to herein as protein-Fc. Other chimeric molecules contemplated by the present invention include the protein or protein-Fc or a fragment thereof linked to a lipid moiety such as a polyunsaturated fatty acid molecule. Such lipid

moieties may be linked to an amino acid residue in the background of the molecule or to a side chain of such an amino acid residue. The chimeric molecules of the present invention, including SCFR-Fc, have a profile of measurable physiochemical parameters indicative of or associated with one or more distinctive pharmacological traits of the isolated protein-Fc.

Accordingly, the present invention provides an isolated polypeptide encoded by a nucleotide sequence selected from the list consisting of SEQ ID NOs: 23, 25, 27, 31, 33, 35, 37, 39, 41, 43, or a nucleotide sequence having at least about 90% identity to any one of the above-listed sequence or a nucleotide sequence capable of hybridizing to any one of the above sequences or their complementary forms under high stringency conditions.

Yet another aspect of the present invention provides an isolated polypeptide comprising an amino acid sequence selected from the list consisting of SEQ ID NOs: 24, 26, 28, 32, 34, 36, 38, 40, 42, 44, or an amino acid sequence having at least about 90% similarity to one or more of the above sequences.

The present invention further contemplates a pharmaceutical composition comprising at least part of the protein or chimeric molecule thereof, together with a pharmaceutically acceptable carrier, co-factor and/or diluent.

With respect to the primary structure, the present invention provides an isolated protein or chimeric molecule thereof, or a fragment thereof, encoded by a nucleotide sequence selected from the list consisting of SEQ ID NOs: 23, 25, 27, 31, 33, 35, 37, 39, 41, 43, or a nucleotide sequence having at least about 90% identity to any one of the above-listed sequence or a nucleotide sequence capable of hybridizing to any one of the above sequences or their complementary forms under high stringency conditions.

Still, another aspect of the present invention provides an isolated nucleic acid molecule encoding protein or chimeric molecule thereof or a functional part thereof comprising a sequence of nucleotides having at least 90% similarity selected from the list consisting of SEQ ID NOs: 23, 25, 27, 31, 33, 35, 37, 39, 41, 43 or after optimal alignment and/or being

capable of hybridizing to one or more of SEQ ID NOs: 23, 25, 27, 31, 33, 35, 37, 39, 41, 43 or their complementary forms under high stringency conditions.

In a particular embodiment, the present invention is directed to an isolated nucleic acid molecule comprising a sequence of nucleotides encoding a protein or chimeric molecule in or related to the SCFR family, including SCFR-Fc, or a ligand thereof, such as SCF, or a fragment thereof, having an amino acid sequence substantially as set forth in one or more of SEQ ID NOs: 24, 26, 28, 32, 34, 36, 38, 40, 42, 44 or an amino acid sequence having at least about 90% similarity to one or more of SEQ ID NOs: 24, 26, 28, 32, 34, 36, 38, 40, 42, 44 after alignment.

In another aspect, the present invention provides an isolated nucleic acid molecule encoding a protein or chimeric molecule in or related to the SCFR family, or a fragment thereof, comprising a sequence of nucleotides selected from the group consisting of SEQ ID NOs: 23, 25, 27, 31, 33, 35, linked directly or *via* one or more nucleotide sequences encoding protein linkers known in the art to nucleotide sequences encoding the constant (Fc) or framework region of a human immunoglobulin, substantially as set forth in one or more of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17 or 19. In a particular embodiment, the nucleotide sequences encoding protein linker comprises nucleotide sequences selected from IP, GSSNT, TRA or VDGIQWIP.

In another aspect, the present invention provides an isolated protein in or related to the the SCFR family, or a fragment thereof, comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 24, 26, 28, 32, 34, 36 linked directly or *via* one or more protein linkers known in the art, to the constant (Fc) or framework region of a human immunoglobulin, substantially as set forth in one or more of SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, 18 or 20.

The present invention further extends to uses of an isolated protein or chimeric molecule thereof or nucleic acid molecules encoding same in diagnostic, prophylactic, therapeutic, nutritional and/or research applications. More particularly, the present invention extends to a method of treating or preventing a condition or ameliorating the symptoms of a condition

in an animal subject, said method comprising administering to said animal subject an effective amount of an isolated protein or chimeric molecule thereof.

In addition, the present invention extends to uses of a protein or chimeric molecule thereof for screening small molecules, which may have a variety of diagnostic, prophylactic, therapeutic, nutritional and/or research applications.

The present invention further contemplates using an isolated protein or chimeric molecule thereof as immunogens to generate antibodies for therapeutic or diagnostic applications.

The present invention further contemplates using an isolated protein or chimeric molecule thereof in culture mediums for stem cells used in stem cell or related therapy.

The subject invention also provides the use of a protein or chimeric molecule thereof in the manufacture of a formulation for diagnostic, prophylactic, therapeutic, nutritional and/or research applications.

The subject invention also provides a human derived protein or chimeric molecule thereof for use as a standard protein in an immunoassay and kits thereof. The subject invention also extends to a method for determining the level of human cell-expressed human protein or chimeric molecule thereof in a biological preparation.

TABLE 1

Sequence Identifier

Sequence Identifier	**Sequence**
SEQ ID NO:1	Human IgG1 Fc nucleotide sequence
SEQ ID NO:2	Human IgG1 Fc amino acid sequence
SEQ ID NO:3	Human IgG1 Fc nucleotide sequence (variant)
SEQ ID NO:4	Human IgG1 Fc amino acid sequence (variant)
SEQ ID NO:5	Human IgG2 Fc nucleotide sequence
SEQ ID NO:6	Human IgG2 Fc amino acid sequence
SEQ ID NO:7	Human IgG3 Fc nucleotide sequence
SEQ ID NO:8	Human IgG3 Fc amino acid sequence
SEQ ID NO:9	Human IgG4 Fc nucleotide sequence
SEQ ID NO:10	Human IgG4 Fc amino acid sequence
SEQ ID NO:11	Human IgA1 Fc nucleotide sequence
SEQ ID NO:12	Human IgA1 Fc amino acid sequence
SEQ ID NO:13	Human IgA2 Fc nucleotide sequence
SEQ ID NO:14	Human IgA2 Fc amino acid sequence
SEQ ID NO:15	Human IgM Fc nucleotide sequence
SEQ ID NO:16	Human IgM Fc amino acid sequence
SEQ ID NO:17	Human IgE Fc nucleotide sequence
SEQ ID NO:18	Human IgE Fc amino acid sequence
SEQ ID NO:19	Human IgD Fc nucleotide sequence
SEQ ID NO:20	Human IgD Fc amino acid sequence
SEQ ID NO: 21	SCF forward primer (nucleotide sequence)
SEQ ID NO: 22	SCF reverse primer (nucleotide sequence)
SEQ ID NO: 23	SCF nucleotide sequence for signal peptide
SEQ ID NO: 24	SCF amino acid sequence for signal peptide
SEQ ID NO: 25	SCF nucleotide sequence for mature peptide
SEQ ID NO: 26	SCF amino acid sequence for mature peptide
SEQ ID NO: 27	SCF nucleotide sequence for signal peptide + mature peptide
SEQ ID NO: 28	SCF amino acid sequence for signal peptide + mature peptide

Sequence Identifier	Sequence
SEQ ID NO: 29	SCFR forward primer (nucleotide sequence)
SEQ ID NO: 30	SCFR reverse primer (nucleotide sequence)
SEQ ID NO: 31	SCFR nucleotide sequence for signal peptide
SEQ ID NO: 32	SCFR amino acid sequence for signal peptide
SEQ ID NO: 33	SCFR nucleotide sequence for mature peptide
SEQ ID NO: 34	SCFR amino acid sequence for mature peptide
SEQ ID NO: 35	SCFR nucleotide sequence for signal peptide + mature peptide
SEQ ID NO: 36	SCFR amino acid sequence for signal peptide + mature peptide
SEQ ID NO: 37	SCFR-Fc nucleotide sequence (mature peptide + GSSNT linker + IgG1 Fc)
SEQ ID NO: 38	SCFR-Fc amino acid sequence (mature peptide + GSSNT linker + IgG1 Fc)
SEQ ID NO: 39	SCFR-Fc nucleotide sequence (mature peptide + GSSNT linker + IgG1 Fc (variant))
SEQ ID NO: 40	SCFR-Fc amino acid sequence (mature peptide + GSSNT linker + IgG1 Fc (variant))
SEQ ID NO: 41	SCFR-Fc nucleotide sequence for whole construct (signal peptide + mature peptide + GSSNT linker + IgG1 Fc)
SEQ ID NO: 42	SCFR-Fc amino acid sequence for whole construct (signal peptide + mature peptide + GSSNT linker + IgG1 Fc)
SEQ ID NO: 43	SCFR-Fc nucleotide sequence for whole construct (signal peptide + mature peptide + GSSNT linker + IgG1 Fc (variant))
SEQ ID NO: 44	SCFR-Fc amino acid sequence for whole construct (signal peptide + mature peptide + GSSNT linker + IgG1 Fc (variant))
SEQ ID NO: 45	Human IgG1 Fc forward primer (for pIRESbleo GSSNT cloning) (nucleotide sequence)
SEQ ID NO: 46	Human IgG1 Fc reverse primer (for pIRESbleo GSSNT cloning) (nucleotide sequence)
SEQ ID NO: 47	SCFR-Fc reverse primer (nucleotide sequence)

TABLE 2

List of physiochemical parameters

P_x	**Physiochemical Parameter**	**SCF**	**SCFR-ECD**	**SCFR-Fc**
P_1	Apparent molecular weight	18 to 55 kDa	55 to 150 kDa	55 to 180 kDa
P_2	Isoelectric point (pI)	3 to 7.5	3.5 to 7.5	3.5 to 8.0
P_3	Number of isoforms	15 to 86	15 to 63	10 to 109
P_4	Relative intensities of the different number of isoforms			
P_5	Percentage by weight carbohydrate	0 to 62%	0 to 63%	0 to 54%
P_6	Observed molecular weight following N-linked oligosaccharide deglycosylation	18 to 40 kDa	42 to 85 kDa	65 to 135 kDa
P_7	Observed molecular weight following N-linked oligosaccharide deglycosylation and O-linked oligosaccharide deglycosylation	18 to 35 kDa	42 to 75 kDa	65 to 120 kDa
P_8	Percentage acidic			

P_x	Physiochemical Parameter	SCF	SCFR-ECD	SCFR-Fc
	monosaccharide content			
P_9	Monosaccharide content			
P_{10}	Sialic acid content			
P_{11}	Sulfate and phosphate content			
P_{12}	Ser/Thr : GalNAc ratio			
P_{13}	Neutral percentage of N-linked oligosaccharide content			
P_{14}	Acidic percentage of N-linked oligosaccharide content			
P_{15}	Neutral percentage of O-linked oligosaccharide content			
P_{16}	Acidic percentage of O-linked oligosaccharide content			
P_{17}	Ratio of N-linked oligosaccharides			
P_{18}	Ratio of O-linked oligosaccharides			
P_{19}	Structure of N-			

P_x	Physiochemical Parameter	SCF	SCFR-ECD	SCFR-Fc
	linked fraction			
P_{20}	Structure of O-linked fraction			
P_{21}	Position and make up of N-linked oligosaccharides	Includes N-145 (numbering from the start of the signal sequence).	Includes N-130, N-367, N-463 and N-486 (numbering from the start of the signal sequence).	Includes N-130 and N-613 (numbered from the start of the signal sequence)
P_{22}	Position and make up of O-linked oligosaccharides			
P_{23}	Co-translational modification			
P_{24}	Post-translational modification			
P_{25}	Acylation			
P_{26}	Acetylation			
P_{27}	Amidation			
P_{28}	Deamidation			
P_{29}	Biotinylation			
P_{30}	Carbamylation or carbamoylation			
P_{31}	Carboxylation			
P_{32}	Decarboxylation			
P_{33}	Disulfide bond formation			
P_{34}	Fatty acid acylation			
P_{35}	Myristoylation			
P_{36}	Palmitoylation			
P_{37}	Stearoylation			

P_x	**Physiochemical Parameter**	**SCF**	**SCFR-ECD**	**SCFR-Fc**
P_{38}	Formylation			
P_{39}	Glycation			
P_{40}	Glycosylation			
P_{41}	Glycophosphatidy linositol anchor			
P_{42}	Hydroxylation			
P_{43}	Incorporation of selenocysteine			
P_{44}	Lipidation			
P_{45}	Lipoic acid addition			
P_{46}	Methylation			
P_{47}	N or C terminal blocking			
P_{48}	N or C terminal removal			
P_{49}	Nitration			
P_{50}	Oxidation of methionine			
P_{51}	Phosphorylation			
P_{52}	Proteolytic cleavage			
P_{53}	Prenylation			
P_{54}	Farnesylation			
P_{55}	Geranyl geranylation			
P_{56}	Pyridoxal phosphate addition			
P_{57}	Sialylation			
P_{58}	Desialylation			
P_{59}	Sulfation			

P_x	**Physiochemical Parameter**	**SCF**	**SCFR-ECD**	**SCFR-Fc**
P_{60}	Ubiquitinylation or ubiquitination			
P_{61}	Addition of ubiquitin-like molecules			
P_{62}	Primary structure			
P_{63}	Secondary structure			
P_{64}	Tertiary structure			
P_{65}	Quaternary structure			
P_{66}	Chemical stability			
P_{67}	Thermal stability			

TABLE 3

List of Pharmacological traits

T_y	**Pharmacological trait**	**SCF**	**SCFR-ECD**	**SCFR-Fc**
T_1	Therapeutic efficiency			
T_2	Effective therapeutic dose ($TCID_{50}$)			
T_3	Bioavailability			
T_4	Time between dosages to maintain therapeutic levels			
T_5	Rate of absorption			
T_6	Rate of excretion			
T_7	Specific activity			
T_8	Thermal stability			
T_9	Lyophilization stability			
T_{10}	Serum/plasma stability			
T_{11}	Serum half-life			
T_{12}	Solubility in blood stream			

T_y	**Pharmacological trait**	**SCF**	**SCFR-ECD**	**SCFR-Fc**
T_{13}	Immunoreactivity Profile	Distinct from that of a rh SCF molecule expressed in a non-human system.		
T_{14}	Immunogenicity			
T_{14A}	Inhibitable by neutralizing antibodies			
T_{15}	Side effects			
T_{16}	Receptor/ligand binding affinity			
T_{17}	Receptor/ligand activation			
T_{18}	Tissue or cell type specificity			
T_{19}	Ability to cross biological membranes or barriers (i.e. gut, lung, blood brain barriers, skin etc)			
T_{19A}	Angiogenic ability			
T_{20}	Tissue uptake			
T_{21}	Stability to degradation			
T_{22}	Stability to freeze-thaw			
T_{23}	Stability to proteases			
T_{24}	Stability to ubiquitination			
T_{25}	Ease of administration			
T_{26}	Mode of administration			
T_{27}	Compatibility with other pharmaceutical excipients or carriers			
T_{28}	Persistence in organism or environment			
T_{29}	Stability in storage			
T_{30}	Toxicity in an organism			

T_y	Pharmacological trait	SCF	SCFR-ECD	SCFR-Fc
	or environment and the like			
T_{31}	Altered biological effects on different cells types			
T_{32}	Proliferation	1.5 to 3-fold more potent than a human rhSCF expressed in *E. coli* cells in the proliferation of M-07e cells		
T_{33}	Differentiation			
T_{34}	Apoptosis			
T_{35}	Growth in cell size			
T_{36}	Cytokine adhesion			
T_{37}	Cell adhesion			
T_{38}	Cell spreading			
T_{39}	Cell motility			
T_{40}	Migration and invasion			
T_{41}	Chemotaxis			
T_{42}	Cell engulfment			
T_{43}	Signal transduction			
T_{44}	Recruitment of proteins to receptors/ligands			
T_{45}	Activation of the JAK/STAT pathway	2-fold greater activation of STAT5 in human CD34+ haematopoietic stem cells than that induced by a rhSCF expressed in E. coli cells		An ability to inhibit rh SCF- rh SCF (expressed from human cells) induced activation of STAT3 in M-07e cells
T_{46}	Activation of the Ras-erk pathway		A greater ability to inhibit SCF induced activation of p42 and p44 MAPKs in	A greater ability to inhibit SCF induced activation of p42 and p44 MAPKs in

T_y	Pharmacological trait	SCF	SCFR-ECD	SCFR-Fc
			CD34+ haematopoietic stem cells than a rh SCFR-ECD expressed from Sf21 insect cells; a greater ability to inhibit SCF induced activation of p42 and p44 MAPKs in M-07e cells than a rh SCFR-ECD expressed from Sf21 insect cells	CD34+ haematopoietic stem cells than a rh SCFR molecule expressed from Sf21 insect cells; a greater ability to inhibit SCF induced activation of p42 and p44 MAPKs in M-07e cells than a rh SCFR-ECD expressed from Sf21 insect cells
T_{47}	Activation of the AKT pathway	4.5-fold greater activation of Akt in human CD34+ haematopoietic stem cells than that induced by a rhSCF expressed in E. coli cells		
T_{48}	Activation of the PKC pathway and PKA pathway			
T_{49}	Activation of the PKA pathway			
T_{50}	Activation of src			
T_{51}	Activation of fas			
T_{52}	Activation of TNFR			
T_{53}	Activation of NFkB			
T_{54}	Activation of p38MAPK			
T_{55}	Activation of c-fos			
T_{56}	Secretion			
T_{57}	Receptor internalization			
T_{58}	Receptor cross-talk			
T_{59}	Up or down regulation of surface markers			
T_{60}	Alteration of FACS			

T_y	Pharmacological trait	SCF	SCFR-ECD	SCFR-Fc
	front/side scatter profiles			
T_{61}	Alteration of subgroup ratios			
T_{62}	Differential gene expression			
T_{63}	Cell necrosis			
T_{64}	Cell clumping			
T_{65}	Cell repulsion			
T_{66}	Binding to heparin sulfates			
T_{67}	Binding to glycosylated structures			
T_{68}	Binding to chondroitin sulfates			
T_{69}	Binding to extracellular matrix (such as collagen, fibronectin)			
T_{70}	Binding to artificial materials (such as scaffolds)			
T_{71}	Binding to carriers			
T_{72}	Binding to co-factors			
T_{73}	The effect alone or in combination with other proteins on stem cell proliferation, differentiation and/or self-renewal.	Greater yield of viable cells from CD34+ human haematopoietic stem cells than that produced by treatment with rhSCF expressed in *E. coli* cells		

A list of abbreviations commonly used herein is provided in Tables 4 and 5.

TABLE 4

Abbreviations and alternate names

Abbreviation	Description
AAA	Amino Acid Analysis
AFC	Affinity Chromatography
bFGF	Basic Fibroblast Growth Factor, FGF2
BSA	Bovine Serum Albumin
cDLC	Combinatorial Dye Ligand Chromatography
CRD	Carbohydrate Recognition Domain
CSF	Colony Stimulating Factor
DCS	Donor Calf Serum
DeoxGlc	2-deoxyglucose
DLC	Dye Ligand pseudoaffinity Chromatography
DSC	Differential Scanning Calorimetry
ECD	Extracellular domain
EGF	Epidermal Growth Factor
ELISA	Enzyme-Linked Immunosorbent Assays
EPO	Erythropoietin
EST	Expressed Sequence Tags
Fc	Fragment Crystallizable or Immunoglobulin constant region
FCS	Fetal Calf Serum
FGF2	Basic Fibroblast Growth Factor, bFGF
FTIS	Fourier Transform Infrared Spectroscopy
Fuc	Fucose
G-CSF	Granulocyte Colony Stimulating Factor
Gal	Galactose
GalNAc, galactosamine	2-deoxy, 2 amino galactose
GFC	Gel Filtration Chromatography
GlcA	Glucuronic acid
GlcNAc, glucosamine	2-deoxy, 2 amino glucose

Abbreviation	Description
Glc	Glucose
GM-CSF	Granulocyte-Macrophage Colony Stimulating Factor
HBS	Hepes Buffered Saline
hES	Human Embryonic Stem Cells
HIC	Hydrophobic Interaction Chromatography
HPAEC-PAD	High-pH anion-exchange chromatography with pulsed amperometric detection
HPLC	High Pressure Liquid Chromatography or High Performance Liquid Chromatography
HSA	Human Serum Albumin
HTS	High Throughput Screening
IdoA	Iduronic acid
IEC	Ion Exchange Chromatography
IEF	Isoelectric focussing
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
lacNAc	N-acetyl lactosamine
lacdiNAc	N,N'-diacetyllactosediamine
LC	Liquid Chromatography
MALDI-TOF	Matrix-Assisted Laser Desorption Ionization - Time of Flight
Man	Mannose
MCC	Metal Chelating Chromatography
MS	Mass Spectroscopy
NacSial, NeuAc or NeuNAc	N-acetyl neuraminic acid
NGlySial, NeuGc or NeuGly	N-glycolyl neuraminic acid
PBS	Phosphate Buffered Saline
PCS	Photon Correlation Spectroscopy
PDGF-AA	Platelet Derived Growth Factor A homodimer

Abbreviation	Description
PNGase	Peptide-N4-(N-acetyl-β-D-glucosaminyl) Asparagine Amidase
RMLP	Receptor Mediated Ligand Chromatography
RPC	Reversed Phase Chromatography
SDS PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
Sia	Sialic acid
TCA	Trichloroacetic acid
TFF	Tangential flow filtration
TGF	Transforming Growth Factor
TNF	Tumor Necrosis Factor
TNFR	Tumor Necrosis Factor Receptor
Xyl	Xylose

TABLE 5

Abbreviations for amino acids

Amino Acid	3 Letter Code	1 Letter Code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

TABLE 6

Stem cell list

Cell type
General Stem Cell Types
Embryonic stem cells
Somatic stem cells
Germ stem cells
Human embryonic stem cells
Human epidermal stem cells
Adipose derived stem cells
Brain
Adult neural stem cells
Human neurons
Human astrocytes
Epidermis
Human keratinocyte stem cells
Human keratinocyte transient amplifying cells
Human melanocyte stem cells
Human melanocytes
Skin
Human foreskin fibroblasts
Pancreas
Human duct cells
Human pancreatic islets
Human pancreatic beta-cells
Kidney
Human adult renal stem cells
Human embryonic renal epithelial stem cells
Human kidney epithelial cells
Liver
Human hepatic oval cells

Cell type
Human hepatocytes
Human bile duct epithelial cells
Human embryonic endodermal stem cells
Human adult hepatocyte stem cells (existence controversial)
Breast
Human mammary epithelial stem cells
Lung
Bone marrow-derived stem cells
Human lung fibroblasts
Human bronchial epithelial cells
Human alveolar type II pneumocytes
Muscle
Human skeletal muscle stem cells (satellite cells)
Heart
Human cardiomyocytes
Bone marrow mesenchymal stem cells
Simple Squamous Epithelial cells
Descending Aortic Endothelial cells
Aortic Arch Endothelial cells
Aortic Smooth Muscle cells
Eye
Limbal stem cells
Corneal epithelial cells
CD34+ hematopoietic stem cells
Mesenchymal stem cells
Osteoblasts (precursor is mesenchymal stem cell)
Peripheral blood mononuclear progenitor cells (hematopoietic stem cells)
Osteoclasts (precursor is above cell type)
Stromal cells
Spleen
Human splenic precursor stem cells

Cell type
Human splenocytes
Immune cells
Human CD4+ T-cells
Human CD8+ T-cells
Human NK cells
Human monocytes
Human macrophages
Human dendritic cells
Human B-cells
Nose
Goblet cells (mucus secreting cells of the nose)
Pseudostriated ciliated columnar cells (located below olfactory region in the nose)
Pseudostratified ciliated epithelium (cells that line the nasopharangeal tubes)
Trachea
Stratified Epithelial cells (cells that line and structure the trachea)
Ciliated Columnar cells (cells that line and structure the trachea)
Goblet cells (cells that line and structure the trachea)
Basal cells (cells that line and structure the trachea)
Oesophagus
Cricopharyngeus muscle cells
Reproduction
Female primary follicles
Male spermatogonium

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a diagrammatic representation of the cloning process for inserting cDNA encoding a protein of the present invention into the pIRESbleo3 or pIRESbleo3-Fc vector.

Figure 2 is a graphical representation comparing the proliferation of M-07e cells by SCF of the present invention (filled circles) and rh SCF expressed using *E. coli* cells (filled squares).

Figure 3 is a graphical representation comparing numbers of viable nucleated cells following treatment of $CD34^+$ human haematopoietic cells with SCF of the present invention (black bar) and recombinant human SCF expressed in *E. coli* cells (open bar).

Figure 4 is a graphical representation comparing viable nucleated cell numbers following treatment of $CD34^+$ human haematopoietic cells with SCF of the present invention + rhG-CSF expressed in human cells (black bar) and rh SCF expressed in *E. coli* cells + rhG-CSF expressed in *E.coli* (open bar).

Figure 5 is a graphical representation comparing the induction of cell surface marker expression (CD11b, CD15, CD184, CD15) associated with granulocytic differentiation following treatment of $CD34^+$ human haematopoietic cells with SCF of the present invention + rhG-CSF expressed in human cells (black bar) and rh SCF expressed in *E. coli* cells + rhG-CSF expressed in *E.coli* (open bar).

Figure 6 is a graphical representation comparing the induction of colony formation and colony size following treatment of $CD34^+$ human haematopoietic cells with SCF of the present invention + rhG-CSF expressed in human cells (black columns) and rh SCF expressed in *E. coli* cells + rhG-CSF expressed in *E.coli* (cross-hatched columns). hcx = human cell expressed protein.

Figure 7 contains two graphical representations comparing the induction of glycophorin A and CD71 expression of $CD34^+$ human haematopoietic cells following treatment with SCF of the present invention + human cell expressed EPO (black bars) and rhSCF expressed in *E. coli* + rhEPO expressed in *E. coli* (white bars). Error bars represent SD.

Figure 8 shows 4 graphical representations comparing the STAT5 and Akt activation in $CD34^+$ human haematopoietic cells following treatment with SCF of the present invention +/- human cell expressed EPO (black bars) and rhSCF expressed in *E. coli* +/- rhEPO expressed in *E. coli* (white bars). Columns in each graph represent 0min, 10min and 30min, as read from left to right.

Figure 9 shows two graphical representations of STAT3 and MAPK activation following treatment of M-07e cells with SCF +/- SCFR molecules. (a) SCFR-Fc of the present invention (diagonal hatching) inhibited SCF-mediated activation of STAT3 phosphorylation (black); (b) Both SCFR-Fc (diagonal hatching) and SCF-ECD (cross hatching) of the present invention inhibited SCF-mediated activation (black) of MAPK, whereas SCF-ECD expressed in S*f*21 insect cells did not (horizontal hatching). Open bar = unstimulated control.

Figure 10 shows two graphical representations of STAT3 and MAPK activation following treatment of human $CD34^+$ haematopoietic stem cells with SCF +/- SCFR molecules. (a) SCFR-Fc of the present invention (diagonal hatching) inhibited SCF-mediated activation of STAT3 phosphorylation (black); (b) Both SCFR-Fc (diagonal hatching) and SCF-ECD (cross hatching) of the present invention inhibited SCF-mediated activation (black) of MAPK, whereas SCF-ECD expressed in S*f*21 insect cells did not (horizontal hatching). Open bar = unstimulated control.

Figure 11 is a graphical representation showing the *in vitro* comparison of immunoreactivity profiles between SCF of the present invention and a human SCF molecule expressed using non-human systems. OD-antibody dilution plots for the SCF of the present invention (filled circles) and a recombinant human SCF molecule expressed in *E. coli* (open circles).

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

It is to be understood that unless otherwise indicated, the subject invention is not limited to specific formulations, manufacturing methods, diagnostic methods, assay protocols, nutritional protocols, or research protocols or the like as such may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting.

It must be noted that, as used in the subject specification, the singular forms "a", "an" and "the" include plural aspects unless the context already dictates otherwise. Thus, for example, reference to "a protein", "a cytokine" or "a chimeric molecule" or "a receptor" includes a single protein, cytokine or receptor or chimeric molecule as well as two or more proteins, cytokines or receptors or chimeric molecules; a "physiochemical parameter" includes a single parameter as well as two or more parameters and so forth.

The terms "compound", "active agent", "chemical agent", "pharmacologically active agent", "medicament", "active" and "drug" are used interchangeably herein to refer to a chemical compound and in particular a protein or chimeric molecule thereof that induces a desired pharmacological and/or physiological effect. The terms also encompass pharmaceutically acceptable and pharmacologically active ingredients of those active agents specifically mentioned herein including but not limited to salts, esters, amides, prodrugs, active metabolites, analogs and the like. When the terms "compound", "active agent", "chemical agent" "pharmacologically active agent", "medicament", "active" and "drug" are used, then it is to be understood that this includes the active agent *per se* as well as pharmaceutically acceptable, pharmacologically active salts, esters, amides, prodrugs, metabolites, analogs, etc.

Reference to a "compound", "active agent", "chemical agent" "pharmacologically active agent", "medicament", "active" and "drug" includes combinations of two or more actives such as two or more cytokines. A "combination" also includes multi-part such as a two-part composition where the agents are provided separately and given or dispensed separately or admixed together prior to dispensation.

For example, a multi-part pharmaceutical pack may have two or more proteins or chimeric molecules in or related to the SCFR receptor family, or a ligand thereof, selected from the group comprising SCF, SCFR-ECD, SCFR-Fc, separately maintained.

The terms "effective amount" and "therapeutically effective amount" of an agent as used herein mean a sufficient amount of the protein or chimeric molecule thereof, alone or in combination with other agents to provide the desired therapeutic or physiological effect or outcome. Undesirable effects, e.g. side effects, are sometimes manifested along with the desired therapeutic effect; hence, a practitioner balances the potential benefits against the potential risks in determining what is an appropriate "effective amount". The exact amount required will vary from subject to subject, depending on the species, age and general condition of the subject, mode of administration and the like. Thus, it may not be possible to specify an exact "effective amount". However, an appropriate "effective amount" in any individual case may be determined by one of ordinary skill in the art using only routine experimentation.

By "pharmaceutically acceptable" carrier, excipient or diluent is meant a pharmaceutical vehicle comprised of a material that is not biologically or otherwise undesirable, i.e. the material may be administered to a subject along with the selected active agent without causing any or a substantial adverse reaction. Carriers may include excipients and other additives such as diluents, detergents, coloring agents, wetting or emulsifying agents, pH buffering agents, preservatives, and the like.

Similarly, a "pharmacologically acceptable" salt, ester, amide, prodrug or derivative of a compound as provided herein is a salt, ester, amide, prodrug or derivative that this not biologically or otherwise undesirable.

The terms "treating" and "treatment" as used herein refer to reduction in severity and/or frequency of symptoms of the condition being treated, elimination of symptoms and/or underlying cause, prevention of the occurrence of symptoms of the condition and/or their underlying cause and improvement or remediation or amelioration of damage following a

condition.

"Treating" a subject may involve prevention of a condition or other adverse physiological event in a susceptible individual as well as treatment of a clinically symptomatic individual by ameliorating the symptoms of the condition.

A "subject" as used herein refers to an animal, in a particular embodiment, a mammal and in a further embodiment human who can benefit from the pharmaceutical formulations and methods of the present invention. There is no limitation on the type of animal that could benefit from the presently described pharmaceutical formulations and methods. A subject regardless of whether a human or non-human animal may be referred to as an individual, patient, animal, host or recipient. The compounds and methods of the present invention have applications in human medicine, veterinary medicine as well as in general, domestic or wild animal husbandry.

As indicated above, in a particular embodiment, the animals are humans or other primates such as orangutans, gorillas, marmosets, livestock animals, laboratory test animals, companion animals or captive wild animals, as well as avian species.

Examples of laboratory test animals include mice, rats, rabbits, guinea pigs and hamsters. Rabbits and rodent animals, such as rats and mice, provide a convenient test system or animal model. Livestock animals include sheep, cows, pigs, goats, horses and donkeys. Non-mammalian animals such as avian species, fish, and amphibians including *Xenopus* spp prokaryotes and non-mammalian eukaryotes.

The term "cytokine" is used in its most general sense and includes any of various proteins secreted by cells to regulate the immune system, modulate the functional activities of individual cells and/or tissues, and/or induce a range of physiological responses. As used herein the term "cytokine" should be understood to refer to a "complete" cytokine as well as fragments, derivatives or homologs or chimeras thereof comprising one or more amino acid additions, deletions or substitutions, but which substantially retain the biological activity of the complete cytokine.

A "cytokine receptor" is a cell membrane associated or soluble portion of the cytokine receptor involved in cytokine signalling or regulation. As used herein the term "cytokine receptor" should be understood to refer to a "complete" cytokine receptor as well as fragments, derivatives or homologs or chimeras thereof comprising one or more amino acid additions, deletions or substitutions, but which substantially retain the biological activity of the complete cytokine receptor.

The term "protein" is used in its most general sense and includes cytokines and cytokine receptors. As used herein, the term "protein" should be understood to refer to a "complete" protein as well as fragments, derivatives or homologs or chimeras thereof comprising one or more amino acid additions, deletions or substitutions, but which substantially retain the biological activity of the complete protein.

The present invention contemplates an isolated protein or chimeric molecule thereof having a profile of measurable physiochemical parameters (P_x), wherein the profile is indicative of, associated with or forms the basis of one or more distinctive pharmacological traits (T_y). The isolated protein or chimeric molecule is a protein in or related to the SCFR receptor family, or a ligand thereof, selected from the group comprising SCF, SCFR-ECD, SCFR-Fc. As used herein, the terms SCF, SCFR-ECD, SCFR-Fc includes reference to the whole polypeptide as well as fragments thereof.

More particularly, the present invention provides an isolated protein or chimeric molecule thereof having a physiochemical profile comprising an array of measurable physiochemical parameters, $\{[P_x]_1, [P_x]_2,...[P_x]_n,\}$, wherein P_x represents a measurable physiochemical parameter and "n" is an integer ≥ 1, wherein each of $[P_x]_1$ to $[P_x]_n$ is a different measurable physiochemical parameter, wherein the value of any one or more of the measurable physiochemical characteristics is indicative of, associated with, or forms the basis of, a distinctive pharmacological trait, T_y, or a number of distinctive pharmacological traits $\{[T_y]_1, [T_y]_2, ...[T_y]_m\}$ wherein T_y represents a distinctive pharmacological trait and m is an integer ≥ 1 and each of $[T_y]_1$ to $[T_y]_m$ is a different pharmacological trait.

As used herein, the term "measurable physiochemical parameters" (P_x) refers to one or more measurable characteristics of an isolated protein or chimeric molecule thereof. Exemplary "distinctive measurable physiochemical parameters" include, but are not limited to apparent molecular weight (P_1), isoelectric point (pI) (P_2), number of isoforms (P_3), relative intensities of the different number of isoforms (P_4), percentage by weight carbohydrate (P_5), observed molecular weight following N-linked oligosaccharide deglycosylation (P_6), observed molecular weight following N-linked and O-linked oligosaccharide deglycosylation (P_7), percentage acidic monosaccharide content (P_8), monosaccharide content (P_9), sialic acid content (P_{10}), sulfate and phosphate content (P_{11}), Ser/Thr:GalNAc ratio (P_{12}), neutral percentage of N-linked oligosaccharide content (P_{13}), acidic percentage of N-linked oligosaccharide content (P_{14}), neutral percentage of O-linked oligosaccharide content (P_{15}), acidic percentage of O-linked oligosaccharide content (P_{16}), ratio of N-linked oligosaccharides (P_{17}), ratio of O-linked oligosaccharides (P_{18}), structure of N-linked oligosaccharide fraction (P_{19}), structure of O-linked oligosaccharide fraction (P_{20}), position and make up of N-linked oligosaccharides (P_{21}), position and makeup of O-linked oligosaccharides (P_{22}), co-translational modification (P_{23}), post-translational modification (P_{24}), acylation (P_{25}), acetylation (P_{26}), amidation (P_{27}), deamidation (P_{28}), biotinylation (P_{29}), carbamylation or carbamoylation (P_{30}), carboxylation (P_{31}), decarboxylation (P_{32}), disulfide bond formation (P_{33}), fatty acid acylation (P_{34}), myristoylation (P_{35}), palmitoylation (P_{36}), stearoylation (P_{37}), formylation (P_{38}), glycation (P_{39}), glycosylation (P_{40}), glycophosphatidylinositol anchor (P_{41}), hydroxylation (P_{42}), incorporation of selenocysteine (P_{43}), lipidation (P_{44}), lipoic acid addition (P_{45}), methylation (P_{46}), N or C terminal blocking (P_{47}), N or C terminal removal (P_{48}), nitration (P_{49}), oxidation of methionine (P_{50}), phosphorylation (P_{51}), proteolytic cleavage (P_{52}), prenylation (P_{53}), farnesylation (P_{54}), geranyl geranylation (P_{55}), pyridoxal phosphate addition (P_{56}), sialylation (P_{57}), desialylation (P_{58}), sulfation (P_{59}), ubiquitinylation or ubiquitination (P_{60}), addition of ubiquitin-like molecules (P_{61}), primary structure (P_{62}), secondary structure (P_{63}), tertiary structure (P_{64}), quaternary structure (P_{65}), chemical stability (P_{66}), thermal stability (P_{67}). A summary of these parameters is provided is Table 2.

The term "distinctive pharmacological traits" would be readily understood by one of skill in the art to include any pharmacological or clinically relevant property of the protein or chimeric molecule of the present invention. Exemplary "pharmacological traits" which in no way limit the invention include: therapeutic efficiency (T_1), effective therapeutic dose ($TCID_{50}$) (T_2), bioavailability (T_3), time between dosages to maintain therapeutic levels (T_4), rate of absorption (T_5), rate of excretion (T_6), specific activity (T_7), thermal stability (T_8), lyophilization stability (T_9), serum/plasma stability (T_{10}), serum half-life (T_{11}), solubility in blood stream (T_{12}), immunoreactivity profile (T_{13}), immunogenicity (T_{14}), inhibition by neutralizing antibodies (T_{14A}), side effects (T_{15}), receptor/ligand binding affinity (T_{16}), receptor/ligand activation (T_{17}), tissue or cell type specificity (T_{18}), ability to cross biological membranes or barriers (i.e. gut, lung, blood brain barriers, skin etc) (T_{19}), angiogenic ability (T_{19A}), tissue uptake (T_{20}), stability to degradation (T_{21}), stability to freeze-thaw (T_{22}), stability to proteases (T_{23}), stability to ubiquitination (T_{24}), ease of administration (T_{25}), mode of administration (T_{26}), compatibility with other pharmaceutical excipients or carriers (T_{27}), persistence in organism or environment (T_{28}), stability in storage (T_{29}), toxicity in an organism or environment and the like (T_{30}).

In addition, the protein or chimeric molecule of the present invention may have altered biological effects on different cells types (T_{31}), including but not limited to human primary cells, such as lymphocytes, erythrocytes, retinal cells, hepatocytes, neurons, keratinocytes, endothelial cells, endodermal cells, ectodermal cells, mesodermal cells, epithelial cells, kidney cells, liver cells, bone cells, bone marrow cells, lymph node cells, dermal cells, fibroblasts, T-cells, B-cells, plasma cells, natural killer cells, macrophages, neutrophils, granulocytes Langerhans cells, dendritic cells, eosinophils, basophils, mammary cells, lobule cells, prostate cells, lung cells, oesophageal cells, pancreatic cells, Beta cells (insulin secreting cells), hemangioblasts, muscle cells, oval cells (hepatocytes), mesenchymal cells, brain microvessel endothelial cells, astrocytes, glial cells, various stem cells including adult and embryonic stem cells, various progenitor cells; and other human immortal, transformed or cancer cell lines. The biological effects on the cells include effects on proliferation (T_{32}), differentiation (T_{33}), apoptosis (T_{34}), growth in cell size (T_{35}), cytokine adhesion (T_{36}), cell adhesion (T_{37}), cell spreading (T_{38}), cell motility (T_{39}), migration and invasion (T_{40}), chemotaxis (T_{41}), cell engulfment (T_{42}), signal transduction

(T_{43}), recruitment of proteins to receptors/ligands (T_{44}), activation of the JAK/STAT pathway (T_{45}), activation of the Ras-erk pathway (T_{46}), activation of the AKT pathway (T_{47}), activation of the PKC pathway (T_{48}), activation of the PKA pathway (T_{49}), activation of src (T_{50}), activation of fas (T_{51}), activation of TNFR (T_{52}), activation of NFkB (T_{53}), activation of p38MAPK (T_{54}), activation of c-fos (T_{55}), secretion (T_{56}), receptor internalization (T_{57}), receptor cross-talk (T_{58}), up or down regulation of surface markers (T_{59}), alteration of FACS front/side scatter profiles (T_{60}), alteration of subgroup ratios (T_{61}), differential gene expression (T_{62}), cell necrosis (T_{63}), cell clumping (T_{64}), cell repulsion (T_{65}), binding to heparin sulfates (T_{66}), binding to glycosylated structures (T_{67}), binding to chondroitin sulfates (T_{68}), binding to extracellular matrix (such as collagen, fibronectin) (T_{69}), binding to artificial materials (such as scaffolds) (T_{70}), binding to carriers (T_{71}), binding to co-factors (T_{72}), the effect alone or in combination with other proteins on stem cell proliferation, differentiation and/or self-renewal (T_{73}) and the like. A summary of these traits is provided in Table 3.

As used herein the term "distinctive" with regard to a pharmacological trait of a protein or a chimeric molecule of the present invention refers to one or more pharmacological traits of the protein or chimeric molecule thereof, which are distinctive for the particular physiochemical profile. In a particular embodiment, one or more of the pharmacological traits of the isolated protein or chimeric molecule thereof is different from, or distinctive relative to a form of the same protein or chimeric molecule produced in a prokaryotic or lower eukaryotic cell or even a higher non-human eukaryotic cell. In a particular embodiment, the pharmacological traits of the subject isolated protein or chimeric molecule thereof are substantially similar to or functionally equivalent to a naturally occurring protein.

As used herein the term "prokaryote" refers to any prokaryotic cell, which includes any bacterial cell (including actinobacterial cells) or archaeal cell. The meaning of the term "non-mammalian eukaryote", as used herein is self-evident. However, for clarity, this term specifically includes any non-mammalian eukaryote including: yeasts such as *Saccharomyces* spp. or *Pichea* spp.; other fungi; insects, including *Drosophila* spp. and

insect cell cultures; fish, including *Danio* spp.; amphibians, including *Xenopus* spp.; plants and plant cell cultures.

Reference to a "stem cell" includes embryonic or adult stem cells and includes those stem cells listed in Table 6. A protein or chimeric molecule of the present invention may be used alone or in a cocktail of proteins to induce one or more of stem cell proliferation, differentiation or self-renewal.

Primary structure of a protein or chimeric molecule thereof may be measured as an amino acid sequence. Secondary structure may be measured as the number and/or relative position of one or more protein secondary structures such as α-helices, parallel β-sheets, antiparallel β-sheets or turns. Tertiary structure describes the folding of the polypeptide chain to assemble the different secondary structure elements in a particular arrangement. As helices and sheets are units of secondary structure, so the domain is the unit of tertiary structure. In multi-domain proteins, tertiary structure includes the arrangement of domains relative to each other. Accordingly, tertiary structure may be measured as the presence, absence, number and/or relative position of one or more protein "domains". Exemplary domains which in no way limit the present invention include: lone helices, helix-turn-helix domains, four helix bundles, DNA binding domains, three helix bundles, Greek key helix bundles, helix-helix packing domains, β-sandwiches, aligned β-sandwiches, orthogonal β-sandwiches, β-barrels, up and down antiparallel β-sheets, Greek key topology domains, jellyroll topology domains, β-propellers, β-trefoils, β-Helices, Rossman folds, α/β horseshoes, α/β barrels, α+β topologies, disulphide rich folds, serine proteinase inhibitor domains, sea anemone toxin domains, EGF-like domains, complement C-module domain, wheat plant toxin domains, Naja (Cobra) neurotoxin domains, green mamba anticholinesterase domains, Kringle domains, mucin like region, globular domains, spacer regions. Quaternary structure is described as the arrangement of different polypeptide chains within the protein structure, with each chain possessing individual primary, secondary and tertiary structure elements. Examples include either homo- or hetro-oligomeric multimerization (e.g. dimerization or trimerization). In one embodiment, the molecule of the present invention selected from the group comprising SCF, SCFR-ECD, SCFR-Fc, exists as a homo- or hetro- dimer, trimer or oligomer.

With respect to the primary structure, the present invention provides an isolated protein or chimeric molecule thereof, or a fragment thereof, encoded by a nucleotide sequence selected from the list consisting of SEQ ID NOs: 23, 25, 27, 31, 33, 35, 37, 39, 41, 43, or a nucleotide sequence having at least about 90% identity to any one of the above-listed sequence or a nucleotide sequence capable of hybridizing to any one of the above sequences or their complementary forms under high stringency conditions.

Still, another aspect of the present invention provides an isolated nucleic acid molecule encoding protein or chimeric molecule thereof or a functional part thereof comprising a sequence of nucleotides having at least 90% similarity selected from the list consisting of SEQ ID NOs: 23, 25, 27, 31, 33, 35, 37, 39, 41, 43 or after optimal alignment and/or being capable of hybridizing to one or more of SEQ ID NOs: 23, 25, 27, 31, 33, 35, 37, 39, 41, 43 or their complementary forms under high stringency conditions.

In a particular embodiment, the present invention is directed to an isolated nucleic acid molecule comprising a sequence of nucleotides encoding a protein or chimeric molecule thereof, or a fragment thereof, having an amino acid sequence substantially as set forth in one or more of SEQ ID NOs: 24, 26, 28, 32, 34, 36, 38, 40, 42, 44 or an amino acid sequence having at least about 90% similarity to one or more of SEQ ID NOs: 24, 26, 28, 32, 34, 36, 38, 40, 42, 44 after optimal alignment.

In another aspect, the present invention provides an isolated nucleic acid molecule encoding a protein molecule, or a fragment thereof, comprising a sequence of nucleotides selected from the group consisting of SEQ ID NOs: 23, 25, 27, 31, 33, 35, linked directly or *via* one or more nucleotide sequences encoding protein linkers known in the art to nucleotide sequences encoding the constant (Fc) or framework region of a human immunoglobulin, substantially as set forth in one or more of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17 or 19. In a particular embodiment, the nucleotide sequences encoding protein linker comprises nucleotide sequences selected from IP, GSSNT, TRA or VDGIQWIP.

In another aspect, the present invention provides an isolated protein molecule, or a fragment thereof, comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 24, 26, 28, 32, 34, 36 linked directly or *via* one or more protein linkers known in the art, to the constant (Fc) or framework region of a human immunoglobulin, substantially as set forth in one or more of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20.

Another aspect of the present invention provides an isolated protein or chimeric molecule thereof, or a fragment thereof, comprising an amino acid sequence selected from the list consisting of SEQ ID NOs: 24, 26, 28, 32, 34, 36, 38, 40, 42, 44, or an amino acid sequence having at least about 90% similarity to one or more of the above sequences.

In a particular embodiment, percentage amino acid similarity or nucleotide identity levels include at least about 61% or at least about 62% or at least about 63% or at least about 64% or at least about 65% or at least about 66% or at least about 67% or at least about 68% or at least about 69% or at least about 70% or at least about 71% or at least about 72% or at least about 73% or at least about 74% or at least about 75% or at least about 76% or at least about 77% or at least about 78% or at least about 79% or at least about 80% or at least about 81% or at least about 82% or at least about 83% or at least about 84% or at least about 85% or at least about 86% or at least about 87% or at least about 88% or at least about 89% or at least about 90% or at least about 91% or at least about 92% or at least about 93% or at least about 94% or at least about 95% or at least about 96% or at least about 97% or at least about 98% or at least about 99% similarity or identity.

A "derivative" of a polypeptide of the present invention also encompasses a portion or a part of a full-length parent polypeptide, which retains partial transcriptional activity of the parent polypeptide and includes a variant. Such "biologically-active fragments" include deletion mutants and small peptides, for example, for at least 10, in a particular embodiment, at least 20 and in a further embodiment at least 30 contiguous amino acids, which exhibit the requisite activity. Peptides of this type may be obtained through the application of standard recombinant nucleic acid techniques or synthesized using conventional liquid or solid phase synthesis techniques. For example, reference may be

made to solution synthesis or solid phase synthesis as described, for example, in Chapter 9 entitled "*Peptide Synthesis*" by Atherton and Shephard which is included in a publication entitled "*Synthetic Vaccines*" edited by Nicholson and published by Blackwell Scientific Publications. Alternatively, peptides can be produced by digestion of an amino acid sequence of the invention with proteinases such as endoLys-C, endoArg-C, endoGlu-C and staphylococcus V8-protease. The digested fragments can be purified by, for example, high performance liquid chromatographic (HPLC) techniques. Any such fragment, irrespective of its means of generation, is to be understood as being encompassed by the term "derivative" as used herein.

The term "variant" refers, therefore, to nucleotide sequences displaying substantial sequence identity with reference nucleotide sequences or polynucleotides that hybridize with a reference sequence under stringency conditions that are defined hereinafter. The terms "nucleotide sequence", "polynucleotide" and "nucleic acid molecule" may be used herein interchangeably and encompass polynucleotides in which one or more nucleotides have been added or deleted, or replaced with different nucleotides. In this regard, it is well understood in the art that certain alterations inclusive of mutations, additions, deletions and substitutions can be made to a reference nucleotide sequence whereby the altered polynucleotide retains the biological function or activity of the reference polynucleotide or the encoded polypeptide. The term "variant" also includes naturally occurring allelic variants.

The nucleic acid molecules of the present invention may be in the form of a vector or other nucleic acid construct.

In one embodiment, the vector is DNA and may optionally comprise a selectable marker.

Examples of selectable markers include genes conferring resistance to compounds such as antibiotics, genes conferring the ability to grow on selected substrates, genes encoding proteins that produce detectable signals such as luminescence. A wide variety of such markers are known and available, including, for example, antibiotic resistance genes such as the neomycin resistance gene (*neo*) and the hygromycin resistance gene (*hyg*).

Selectable markers also include genes conferring the ability to grown on certain media substrates such as the *tk* gene (thymidine kinase) or the *hprt* gene (hypoxanthine phosphoribosyltransferase) which confer the ability to grow on HAT medium (hypoxanthine, aminopterin and thymidine); and the bacterial *gpt* gene (guanine/xanthine phosphoribosyltransferase) which allows growth on MAX medium (mycophenolic acid, adenine and xanthine). Other selectable markers for use in mammalian cells and plasmids carrying a variety of selectable markers are described in Sambrook *et al. Molecular Cloning - A Laboratory Manual*, Cold Spring Harbour, New York, USA, 1990.

The selectable marker may depend on its own promoter for expression and the marker gene may be derived from a very different organism than the organism being targeted (e.g. prokaryotic marker genes used in targeting mammalian cells). However, it is favorable to replace the original promoter with transcriptional machinery known to function in the recipient cells. A large number of transcriptional initiation regions are available for such purposes including, for example, metallothionein promoters, thymidine kinase promoters, β-actin promoters, immunoglobulin promoters, SV40 promoters and human cytomegalovirus promoters. A widely used example is the pSV2-*neo* plasmid which has the bacterial neomycin phosphotransferase gene under control of the SV40 early promoter and confers in mammalian cells resistance to G418 (an antibiotic related to neomycin). A number of other variations may be employed to enhance expression of the selectable markers in animal cells, such as the addition of a poly(A) sequence and the addition of synthetic translation initiation sequences. Both constitutive and inducible promoters may be used.

The genetic construct of the present invention may also comprise a 3' non-translated sequence. A 3' non-translated sequence refers to that portion of a gene comprising a DNA segment that contains a polyadenylation signal and any other regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. Polyadenylation signals are commonly recognized by the presence of homology to the canonical form 5' AATAAA-3' although variations are not uncommon.

Accordingly, a genetic construct comprising a nucleic acid molecule of the present invention, operably linked to a promoter, may be cloned into a suitable vector for delivery to a cell or tissue in which regulation is faulty, malfunctioning or non-existent, in order to rectify and/or provide the appropriate regulation. Vectors comprising appropriate genetic constructs may be delivered into target eukaryotic cells by a number of different means well known to those skilled in the art of molecular biology.

The term "similarity" as used herein includes exact identity between compared sequences at the nucleotide or amino acid level. Where there is non-identity at the nucleotide level, "similarity" includes differences between sequences which result in different amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. Where there is non-identity at the amino acid level, "similarity" includes amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. In a particular embodiment, nucleotide and sequence comparisons are made at the level of identity rather than similarity.

Terms used to describe sequence relationships between two or more polynucleotides or polypeptides include "reference sequence", "comparison window", "sequence similarity", "sequence identity", "percentage of sequence similarity", "percentage of sequence identity", "substantially similar" and "substantial identity". A "reference sequence" is at least 12 but frequently 15 to 18 and often at least 25 or above, such as 30 monomer units, inclusive of nucleotides and amino acid residues, in length. Because two polynucleotides may each comprise (1) a sequence (i.e. only a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a "comparison window" to identify and compare local regions of sequence similarity. A "comparison window" refers to a conceptual segment of typically 12 contiguous residues that is compared to a reference sequence. The comparison window may comprise additions or deletions (i.e. gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window

may be conducted by computerized implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA) or by inspection and the best alignment (i.e. resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul *et al.* (*Nucl Acids Res 25:*389, 1997). A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel *et al.* (*In: Current Protocols in Molecular Biology*, John Wiley & Sons Inc. 1994-1998).

The terms "sequence similarity" and "sequence identity" as used herein refers to the extent that sequences are identical or functionally or structurally similar on a nucleotide-by-nucleotide basis or an amino acid-by-amino acid basis over a window of comparison. Thus, a "percentage of sequence identity", for example, is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (e.g. A, T, C, G, I) or the identical amino acid residue (e.g. Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys and Met) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity. For the purposes of the present invention, "sequence identity" will be understood to mean the "match percentage" calculated by the DNASIS computer program (Version 2.5 for windows; available from Hitachi Software Engineering Co., Ltd., South San Francisco, California, USA) using standard defaults as used in the reference manual accompanying the software. Similar comments apply in relation to sequence similarity.

Reference herein to a low stringency includes and encompasses from at least about 0 to at least about 15% v/v formamide and from at least about 1 M to at least about 2 M salt for hybridization, and at least about 1 M to at least about 2 M salt for washing conditions. Generally, low stringency is at from about 25-30°C to about 42°C, such as 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 and 42°C. The temperature may be altered

and higher temperatures used to replace formamide and/or to give alternative stringency conditions. Alternative stringency conditions may be applied where necessary, such as medium stringency, which includes and encompasses from at least about 16% v/v to at least about 30% v/v formamide, such as 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 and 30% and from at least about 0.5 M to at least about 0.9 M salt, such as 0.5, 0.6, 0.7, 0.8 or 0.9 M for hybridization, and at least about 0.5 M to at least about 0.9 M salt, such as 0.5, 0.6, 0.7, 0.8 or 0.9 M for washing conditions, or high stringency, which includes and encompasses from at least about 31% v/v to at least about 50% v/v formamide, such as 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 and 50% and from at least about 0.01 M to at least about 0.15 M salt, such as 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10, 0.11, 0.12, 0.13, 0.14 and 0.15 M for hybridization, and at least about 0.01 M to at least about 0.15 M salt, such as 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10, 0.11, 0.12, 0.13, 0.14 and 0.15 M for washing conditions. In general, washing is carried out $T_m = 69.3 + 0.41$ (G+C)% (Marmur and Doty, *J Mol Biol 5:*109, 1962). However, the T_m of a duplex DNA decreases by 1°C with every increase of 1% in the number of mismatch base pairs (Bonner and Laskey, *Eur J Biochem 46:*83, 1974. Formamide is optional in these hybridization conditions. Accordingly, in a particular embodiment levels of stringency are defined as follows: low stringency is 6 x SSC buffer, 0.1% w/v SDS at 25-42°C; a moderate stringency is 2 x SSC buffer, 0.1% w/v SDS at a temperature in the range 20°C to 65°C; high stringency is 0.1 x SSC buffer, 0.1% w/v SDS at a temperature of at least 65°C.

As used herein, the terms "co- or post-translational modifications" refer to covalent modifications occurred during or after translation of the peptide chain. Exemplary co- or post-translational modifications include but are not limited to acylation (including acetylation), amidation or deamidation, biotinylation, carbamylation (or carbamoylation), carboxylation or decarboxylation, disulfide bond formation, fatty acid acylation (including myristoylation, palmitoylation and stearoylation), formylation, glycation, glycosylation, hydroxylation, incorporation of selenocysteine, lipidation, lipoic acid addition, methylation, N- or C-terminal blocking, N- or C-terminal removal, nitration, oxidation of methionine, phosphorylation, proteolytic cleavage, prenylation (including farnesylation,

geranyl geranylation), pyridoxal phosphate addition, sialylation or desialylation, sulfation, ubiquitinylation (or ubiquitination) or addition of ubiquitin-like proteins.

Acylation involves the hydrolysis of the N-terminus initiator methionine and the addition of an acetyl group to the new N-termino amino acid. Acetyl Co-A is the acetyl donor for acylation.

Amidation is the covalent linkage of an amide group to the carboxy terminus of a peptide and is frequently required for biological activity and stability of a protein. Deamidation is the hydrolytic removal of an amide group. Deamidation of amide containing amino acid residues is a rare modification that is performed by the organism to re-arrange the 3D structure and alter the charge ratio/pI.

Biotinylation is a technique whereby biotinyl groups are incorporated into molecules, either that catalyzed by holocarboxylase synthetase during enzyme biosynthesis or that undertaken *in vitro* to visualise specific substrates by incubating them with biotin-labeled probes and avidin or streptavidin that has been linked to any of a variety of substances amenable to biochemical assay.

Carbamylation (or carbamoylation) is the transfer of the carbamoyl from a carbamoyl-containing molecule (e.g., carbamoyl phosphate) to an acceptor moiety such as an amino group.

Carboxylation of glutamic acid residues is a vitamin K dependent reaction that results in the formation of a gamma carboxyglutamic acid (Gla residue). Gla residues within several proteins of the blood-clotting cascade are necessary for biological function of the proteins. Carboxylation can also occur to aspartic acid residues.

Disulfide bonds are covalent linkages that form when the thiol groups of two cysteine residues are oxidized to a disulfide. Many mammalian proteins contain disulfide bonds, and these are crucial for the creation and maintenance of tertiary structure of the protein, and thus biological activity.

Protein synthesis in bacteria involves formylation and deformylation of N-terminal methionines. This formylation/deformylation cycle does not occur in cytoplasm of eukaryotic cells and is a unique feature of bacterial cells. In addition to the hydroxylation that occurs on glycine residues as part of the amidation process, hydroxylation can also occur in proline and lysine residues catalysed by prolyl and lysyl hydroxylase (Kivirikko *et al. FASEB Journal 3:*1609-1617, 1989).

Glycation is the uncontrolled, non-enzymatic addition of glucose or other sugars to the amino acid backbone of protein.

Glycosylation is the addition of sugar units to the polypeptide backbone and is further described hereinafter.

Hydroxylation is a reaction which is dependent on vitamin C as a co-factor. Adding to the importance of hydroxylation as a post- translation modification is that hydroxy-lysine serves as an attachment site for glycosylation.

Selenoproteins are proteins which contain selenium as a trace element by the incorporation of a unique amino acid, selenocysteine, during translation. The tRNA for selenocysteine is charged with serine and then enzymatically selenylated to produce the selenocysteinyl-tRNA. The anticodon of selenocysteinyl-tRNA interacts with a stop codon in mRNA (UGA) instead of a serine codon. An element in the 3' non-translated region (UTR) of selenoprotein mRNAs determines whether UGA is read as a stop codon or as a selenocysteine codon.

Lipidation is a generic term that encompasses the covalent attachment of lipids to proteins, this includes fatty acid acylation and prenylation.

Fatty acid acylation involves the covalent attachment of fatty acids such as the 14 carbon Myristic acid (Myristoylation), the 16 carbon Palmitic acid (Palmitoylation) and the 18 carbon Stearic acid (Stearoylation). Fatty acids are linked to proteins in the pre-Golgi

compartment and may regulate the targeting of proteins to membranes (Blenis and Resh *Curr Opin Cell Biol 5(6):*984-9, 1993). Fatty acid acylation is, therefore, important in the functional activity of a protein (Bernstein *Methods Mol Biol 237*:195-204, 2004).

Prenylation involves the addition of prenyl groups, namely the 15 carbon farnesyl or the 20 carbon geranyl-geranyl group to acceptor proteins. The isoprenoid compounds, including farnesyl diphosphate or geranylgeranyl diphosphate, are derived from the cholesterol biosynthetic pathway. The isoprenoid groups are attached by a thioether link to cysteine residues within the consensus sequence CAAX, (where A is any aliphatic amino acid, except alanine) located at the carboxy terminus of proteins. Prenylation enhances proteins ability to associate with lipid membranes and all known GTP-binding and hydrolyzing proteins (G proteins) are modified in this way, making prenylation crucial for signal transduction (Rando *Biochim Biophys Acta 1300(1):*5-16, 1996; Gelb *et al. Curr Opin Chem Biol 2(1):*40-8, 1998).

Lipoic acid is a vitamin-like antioxidant that acts as a free radical scavenger. Lipoyl-lysine is formed by attaching lipoic acid through an amide bond to lysine by lipoate protein ligase.

Protein methylation is a common modification that can regulate the activity of proteins or create new types of amino acids. Protein methyltransferases transfer a methyl group from S-adenosyl-L-methionine to nucleophilic oxygen, nitrogen, or sulfur atoms on the protein. The effects of methylation fall into two general categories. In the first, the relative levels of methyltransferases and methylesterases can control the extent of methylation at a particular carboxyl group, which in turn regulates the activity of the protein. This type of methylation is reversible. The second group of protein methylation reactions involves the irreversible modification of sulfur or nitrogen atoms in the protein. These reactions generate new amino acids with altered biochemical properties that alter the activity of the protein (Clarke *Curr Opin Cell Biol 5*:977 983, 1993).

Protein nitration is a significant post-translational modification, which operates as a transducer of nitric oxide signalling. Nitration of proteins modulates catalytic activity, cell signalling and cytoskeletal organization.

Phosphorylation refers to the addition of a phosphate group by protein kinases. Serine, threonine and tyrosine residues are the amino acids subject to phosphorylation. Phosphorylation is a critical mechanism, which regulates biological activity of a protein.

A majority of proteins are also modified by proteolytic cleavage. This may simply involve the removal of the initiation methionine. Other proteins are synthesized as inactive precursors (proproteins) that are activated by limited or specific proteolysis. Proteins destined for secretion or association with membranes (preproteins) are synthesized with a signal sequence of 12-36 predominantly hydrophobic amino acids, which is cleaved following passage through the ER membrane.

Pyridoxal phosphate is a co-enzyme derivative of vitamin B6 and participates in transaminations, decarboxylations, racemizations, and numerous modifications of amino acid side chains. All pyridoxal phosphate-requiring enzymes act via the formation of a Schiff base between the amino acid and coenzyme. Most enzymes responsible for attaching the pyridoxal-phosphate group to the lysine residue are self activating.

Sialylation refers to the attachment of sialic acid to the terminating positions of a glycoprotein via various sialyltransferase enzymes; and desialylation refers the removal of sialic acids. Sialic acids include but are not limited to, N-acetyl neuraminic acid (NeuAc) and N-glycolyl neuraminic acid (NeuGc). Sialyl structures that result from the sialylation of glycoproteins include sialyl Lewis structures, for example, sialyl Lewis a and sialyl Lewis x, and sialyl T structures, for example, Sialyl-TF and Sialyl Tn.

Sulfation occurs at tyrosine residues and is catalyzed by the enzyme tyrosylprotein sulfotransferase which occurs in the *trans*-Golgi network. It has been determined that 1 in 20 of the proteins secreted by HepG2 cells and 1 in 3 of those secreted by fibroblasts contain at least one tyrosine sulfate residue. Sulfation has been shown to influence

biological activity of proteins. Of particular interest is that the CCR5, a major HIV co-receptor, was shown to be tyrosine-sulfated and that sulfation of one or more tyrosine residues in the N-terminal extracellular domain of CCR5 are required for optimal binding of MIP-1 alpha/CCL3, MIP-1 beta/CCL4, and RANTES/CCL5 and for optimal HIV co-receptor function (Moore *J Biol Chem 278(27):*24243-24246, 2003). Sulfation can also occur on sugars. In addition, sulfation of a carbohydrate moiety of a glycoprotein can occur by the action of glycosulfotransferases such as GalNAc(β1-4)GlcNAc(β1-2)Manα4 sulfotransferase.

Post-translational modifications can encompass protein-protein linkages. Ubiquitin is a 76 amino acid protein which both self associates and covalently attaches to other proteins in mammalian cells. The attachment is via a peptide bond between the C-terminus of ubiquitin and the amino group of lysine residues in other proteins. Attachment of a chain of ubiquitin molecules to a protein targets it for proteolysis by the proteasome and is an important mechanism for regulating the steady state levels of regulatory proteins e.g. with respect to the cell cycle (Wilkinson *Annu Rev Nutr 15*:161-89, 1995). In contrast, mono-ubiquitination can play a role in direct regulation of protein function. Ubiquitin-like proteins that can also be attached covalently to proteins to influence their function and turnover include NEDD-8, SUMO-1 and Apg12.

Glycosylation is the addition of sugar residues in the polypeptide backbone. Sugar residues, such as monosaccharides, disaccharides and oligosaccharides include but are not limited to: fucose (Fuc), galactose (Gal), glucose (Glc), galactosamine (GalNAc), glucosamine (GlcNAc), mannose (Man), N-acetyl-lactosamine (lacNAc) N,N'-diacetyllactosediamine (lacdiNAc). These sugar units can attach to the polypeptide back bones in at least seven ways, namely:

(1) via an N-glycosidic bond to the R-group of an asparagine residue in the consensus sequence Asn-X-Ser; Asn-X-Thr; or Asn-X-Cys (N-glycosylation; N-linked glycans);

(2) via an O-glycosidic bond to the R-group of serine, threonine, hydroxyproline, tyrosine or hydroxylysine (O-glycosylation; O-linked glycans);

(3) via the R-group of tyrosine in C-linked mannose;

(4) as a glycophosphatidylinositol anchor used to secure some proteins to cell membranes;

(5) as a single monosaccharide attachment of GlcNAc to the R-group of serine or threonine. This linkage is often reversibly associated with attachment of inorganic phosphate (Yin-o-Yang);

(6) attachment of a linear polysaccharide to serine, threonine or asparagine (proteoglycans); and

(7) *via* a S-glycosidic bond to the R-group of cysteine.

N-linked glycans are beta linked to the amide nitrogen of an asparagine residue in the consensus sequence – Asn-Xaa-Ser/Thr/Cys where Xaa ≠ Pro and nearly always have an N-acetylglucosamine residue at the reducing terminus. They are biologically synthesised by the transfer of Glc3Man9GlcNAc2 from dolichol-pyrophosphate-Glc3Man9GlcNAc2 to asparagine. This is followed by enzymatic processing to form the final N-linked glycan structures. This method of synthesis means that the core pentasaccharide Man(α1-6)[Man(α1-3)]Man(β1-4)GlcNAc(β1-4)GlcNAc is always present in N-linked glycans.

N-linked glycans are divided into three distinct classes: high-mannose type, complex type and hybrid type; all are formed by elongation of the core. High-mannose type N-glycans result when the core structure is predominantly substituted by mannose. Complex type N-glycans contain the core structure substituted by one or more of the sugars N-acetylglycosamine, galactose, fucose or sialic acid. Hybrid type N-glycans have structural features of both the high-mannose and complex type chains. Within each of the three classes of N-glycans there exists a wide diversity of structures which comes from the variation in the number and position of individual sugar residues, the degree of branching, and the level of phosphorylation and sulfation.

O-linked glycans are usually composed of N-acetylgalactosamine linked to serine or threonine. This type of O-linked glycan is also called mucin-type as they are often found on mucins. Unlike N-linked glycans, mucin-type O-linked glycans are synthesised one residue at a time and at least 8 core classes have been described.

Three domains can be distinguished within the larger mucin-type O-glycan chains: a core, backbone and peripheral region. The core region consists of the N-acetylgalactosamine residue attached to serine or threonine plus the sugar residue/s linked directly. The backbone region is defined as the series of galactose and N-acetylglucosamine residues linked to the core. There are two common backbone repeats: Gal(β1-3)GlcNAc (type 1) and Gal(β1-4)GlcNAc (type 2). Both backbone structures may be combined to form a long linear chain or a branched structure e.g., Gal(β1-4)GlcNAc(β1-6)[Gal(β1-4)GlcNAc(β1-3)Gal-{rest of chain}.

The peripheral region can include terminal residues such as fucose, galactose, N-acetlygalactosamine, N-acetylglucosamine, N-acetylneuraminic acid and sulfate. These terminal residues determine most of the characteristics of the O-linked glycans. For example, sialic acid residues and sulfate groups add a negative charge to the protein, increase its solubility and modify its conformation. Additionally, peripheral residues of glycosylation structures can comprise one or more of the following carbohydrate antigenic determinants in Table 7.

TABLE 7

List of carbohydrate antigenic determinants

Antigenic Name	**Antigenic Glycan Structure**
Blood group H(O), type 1	Fuc(α1-2)Gal(β1-3)GlcNAc-R
Blood group H(O), type 2	Fuc(α1-2)Gal(β1-4)GlcNAc-R
Blood group A, type 1	GalNAc(α1-3)[Fuc(α1-2)]Gal(β1-3)GlcNAc-R
Blood group A, type 2	GalNAc(α1-3)[Fuc(α1-2)]Gal(β1-4)GlcNAc-R
Blood group B, type 1	Gal(α1-3)[Fuc(α1-2)]Gal(β1-3)GlcNAc-R
Blood group B, type 2	Gal(α1-3)[Fuc(α1-2)]Gal(β1-4)GlcNAc-R
Blood group i	$[Gal(\beta 1\text{-}4)GlcNAc(\beta 1\text{-}3)]_n Gal(\beta 1\text{-}R$

Antigenic Name	Antigenic Glycan Structure
Blood group I	Gal(β1-4)GlcNAc(β1-3)[Gal(β1-4)GlcNAc(β1-6)]Gal(β1-4)GlcNAc(β1-3)Gal(β1-R
Lewis a (Le^a)	Gal(β1-3)[Fuc(α1-4)]GlcNAc-R
Sialyl Lewis a (sLe^a)	NeuAc(α2-3)Gal(β1-3)[Fuc(α1-4)]GlcNAc-R
Lewis b (Le^b)	Fuc(α1-2)Gal(β1-3)[Fuc(α1-4)]GlcNAc-R
Lewis x (Le^X)	Gal(β1-4)[Fuc(α1-3)]GlcNAc-R
Sialyl Lewis x (sLe^X)	NeuAc(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc-R
Lewis y (Le^y)	Fuc(α1-2)Gal(β1-4)[Fuc(α1-3)]GlcNAc-R
Forssman	GalNAc(α1-3)GalNAc(β1-3)Gal-R
Thomsen-Friedenreich (TF or T)	Gal(β1-3)GalNAc(α1-O)-Ser/Thr
Sialyl-TF (sTF) or Sialyl-T (sT)	Gal(β1-3)[NeuAc(α2-6)]GalNAc(α1-O)-Ser/Thr
Tn	GalNAc(α1-O)-Ser/Thr
Sialyl Tn (sTn)	NeuAc(α2-6)GalNAc(α1-O)-Ser/Thr

The carbohydrates will also contain several antennary structures, including mono, bi, tri and tetra outer structures.

Glycosylation may be measured by the presence, absence or pattern of N-linked glycosylation, O-linked glycosylation, C-linked mannose structure, and glycophosphatidylinositol anchor; the percentage of carbohydrate by mass; Ser/Thr – GalNAc ratio; the proportion of mono, bi, tri and tetra sugar structures or by lectin or antibody binding.

Sialylation of a protein may be measured by the immunoreactivity of the protein with an antibody directed against a particluar sialyl structure. For example, Lewis x specific antibodies react with CEACAM1 expressed from granulocytes but not with recombinant human CEACAM1 expressed in 293 cells (Lucka *et al. Glycobiology 15(1)*:87-100, 2005).

Alternatively, the presence of sialylated structures on a protein may be detected by a combination of glycosidase treatment followed by a suitable measurement procedure such as mass spectroscopy (MS), high performance liquid chromatography (HPLC) or glyco mass fingerprinting (GMF).

The apparent molecular weight of a protein includes all elements of a protein complex (cofactors and non-covalently bonded domains) and all co- or post-translational modifications (addition or removal of covalently bonded groups to and from peptide). Apparent molecular weight is often affected by co- or post-translational modifications. A protein's apparent molecular weight may be determined by SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis), which is also the second dimension on its two-dimensional counterpart, 2D-PAGE (two-dimensional polyacrylamide gel electrophoresis). It may be determined more accurately, however, by mass spectrometry (MS)- either by Matrix-Assisted Laser Desorption Ionization - Time of Flight (MALDI-TOF) MS, which produces charged molecular ions or the more sensitive Electrospray Ionization (ESI) MS, which produces multiple-charged peaks. The apparent molecular weights of the protein or chimeric molecule thereof may be within the range of 1 to 1000 kDa. Accordingly, the isolated protein or chimeric molecule of the present invention has a apparent molecular weight of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277,

278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853,

854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000 kDa. The molecular weight or molecular mass of a protein may be determined by any convenient means such as electrophoresis, mass spectrometry, gradient ultracentrifugation.

The isoelectric point (or pI) of a protein is the pH at which the protein carries no net charge. This attribute may be determined by isoelectric focusing (IEF), which is also the first dimension of 2D-PAGE. Experimentally determined pI values are affected by a range of co- or post-translational modifications and therefore the difference between an experimental pI and theoretical pI may be as high as 5 units. Accordingly, an isolated protein or chimeric molecule of the present invention may have a pI of 0, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, 7.0, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 8.0, 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5, 9.6, 9.7, 9.8, 9.9, 10.0, 10.1, 10.2, 10.3, 10.4, 10.5, 10.6, 10.7, 10.8, 10.9, 11.0, 11.1, 11.2, 11.3, 11.4, 11.5, 11.6, 11.7, 11.8, 11.9, 12.0, 12.1, 12.2, 12.3, 12.4, 12.5, 12.6, 12.7, 12.8, 12.9, 13.0, 13.1, 13.2, 13.3, 13.4, 13.5, 13.6, 13.7, 13.8, 13.9, or 14.0.

As used herein, the term "isoform" means multiple molecular forms of a given protein, and includes proteins differing at the level of (1) primary structure (such as due to alternate RNA splicing, or polymorphisms); (2) secondary structure (such as due to different co- or post translational modifications); and/or (3) tertiary or quaternary structure (such as due to different sub-unit interactions, homo- or hetero- oligomeric multimerization). In particular, the term "isoform" includes glycoform, which encompasses a protein or

chimeric molecule thereof having a constant primary structure but differing at the level of secondary or tertiary structure or co-or post-translational modification such as different glycosylation forms.

Chemical stability of a protein may be measured as the "half-life" of the protein in a particular solvent or environment. Typically, proteins with a molecular weight of less than 50 kDa have a half-life of approximately 5 to 20 minutes. The proteins or chimeric molecules of the present invention are contemplated to have a half-life of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100 hours. Another particularly convenient measure of chemical stability is the resistance of a protein or chimeric molecule thereof to protease digestion, such as trypsin or chymotrypsin digestion.

The binding affinity of a protein or chimeric molecule thereof to its ligand or receptor may be measured as the equilibrium dissociation constant (Kd) or functionally equivalent measure.

The solubility of a protein may be measured as the amount of protein that is soluble in a given solvent and/or the rate at which the protein dissolves. Furthermore, the rate and or level of solubility of a protein or chimeric molecule thereof in solvents of differing properties such as polarity, pH, temperature and the like may also provide measurable physiochemical characteristics of the protein or chimeric molecule thereof.

Any "measurable physiochemical parameters" may be determined, measured, quantified or qualified using any methods known to one of skill in the art. Described below is a range of methodologies which may be useful in determining, measuring, quantifying or qualifying one or more measurable physiochemical parameters of an isolated protein or chimeric molecule thereof. However, it should be understood that the present invention is in no way

limited to the particular methods described, or to the measurable physiochemical parameters that are measurable using these methods.

Glycoproteins can be said to have two basic components that interact with each other to create the molecule as a whole- the amino acid sequence and the carbohydrate or sugar side chains. The carbohydrate component of the molecule exists in the form of monosaccharide or oligosaccharide side chains attached to the amine side chain of Asn or the hydroxyl side chain of Ser/Thr residues of the amino acid backbone by N- or O-linkages, respectively. A monosaccharide is the term given to the smallest unit of a carbohydrate that is regarded as a sugar, having the basic formula of $(CH_2O)_n$ and most often forming a ring structure of 5 or 6 atoms (pentoses and hexoses respectively). Oligosaccharides are combinations of monosaccharides forming structures of varying complexities that may be either linear or branched but which generally do not have long chains of tandem repeating units (such as is the case for polysaccharides). The level of branching that the oligosaccharide contains as well as the terminal and branching substitutions dramatically affect the properties of the glycoprotein as a whole, and play an important role in the biological function of the molecule. Oligosaccharides are manufactured and attached to the amino acid backbone in the endoplasmic reticulum (ER) and Golgi apparatus of the cell. Different organisms and cell types have different ratios of glycotransferases and endoglycosidases and exoglycosidases and therefore produce different oligosaccharide structures. One of the fundamental defence mechanisms of the body is the detection and destruction of aberrant isoforms and as such it is important to have correct glycosylation of a biological therapeutic not only to increase effectiveness but also to decrease detection by neutralizing antibodies.

Glycan chains are often expressed in a branched fashion, and even when they are linear, such chains are often subject to various modifications. Thus, the complete sequencing of oligosaccharides is difficult to accomplish by a single method and therefore requires iterative combinations of physical and chemical approaches that eventually yield the details of the structure under study.

Determination of the glycosylation pattern of a protein can be performed using a number of different systems, for example using SDS-PAGE. This technique relies on the fact that glycosylated proteins often migrate as diffuse bands by SDS-PAGE. Differentiation between different isoforms are performed by treating a protein with a series of agents. For example, a marked decrease in band width and change in migration position after digestion with peptide-N4-(N-acetyl-β-D-glucosaminyl) asparagine amidase (PNGase) is considered diagnostic of N-linked glycosylation.

To determine the composition of N-linked glycosylation, N-linked oligosaccharides are removed from the protein with PNGase cloned from *Flavobacterium meningosepticum* and expressed in *E. coli*. The removed N-linked oligosaccharides may be recovered from Alltech Carbograph SPE Carbon columns (Deerfield, Illinois, USA) as described by Packer *et al. Glycoconj J 5(8)*:737-47, 1998. The sample can then be taken for monosaccharide analysis, sialic acid analysis or sulfate analysis on a Dionex system with a GP50 pump ED50 pulsed Amperometric or conductivity detector and a variety of pH anion exchange columns.

The extent of O-linked glycosylation may be determined by first removing O-linked oligosaccharides from the target protein by β-elimination. The removed O-linked oligosaccharides may be recovered from Alltech Carbograph SPE Carbon columns (Deerfield, Illinois, USA) as described by Packer *et al.* (1998, *supra*). The sample can then be taken for monosaccharide analysis, sialic acid analysis or sulfate analysis on a Dionex system with a GP50 pump ED50 pulsed Amperometric or conductivity detector and a variety of pH anion exchange columns.

Monosaccharide subunits of an oligosaccharide have variable sensitivities to acid and thus can be released from the target protein by mild trifluoro-acetic acid (TFA) conditions, moderate TFA conditions, and strong hydrochloric acid (HCl) conditions. The monosaccharide mixtures are then separated by high pH anion exchange chromatography (HPAEC) using a variety of column media, and detected using pulsed amperometric electrochemical detection (PAD).

High-pH anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) has been extensively used to determine monosaccharide composition. Fluorophore-based labeling methods have been introduced and many are available in kit form. A distinct advantage of fluorescent methods is an increase in sensitivity (about 50-fold). One potential disadvantage is that different monosaccharides may demonstrate different selectivity for the fluorophore during the coupling reaction, either in the hydrolyzate or in the external standard mixture. However, the increase in sensitivity and the ability to identify which monosaccharides are present from a small portion of the total amount of available glycoprotein, as well as the potential for greater sensitivity using laser-induced fluorescence, makes this approach attractive. In addition a conductivity detector may be used to determine the sulfate and phosphate composition. By using standards, the peak areas can be calculated to total amounts of each monosaccharide present. These data can indicate the level of N- and O-linked glycosylation, the extent of sialylation, and in combination with amino acid composition, percent by weight glycosylation, percent by weight acidic glycoproteins.

Monosaccharide composition analysis of small amounts of protein is best performed with PVDF (PSQ) membranes, after electroblotting, or, if smaller aliquots are to be analyzed, on dot blots. PVDF is an ideal matrix for carbohydrate analysis because neither monosaccharides nor oligosaccharides bind to the membrane, once released by acid or enzymatic hydrolysis.

Determination of the oligosaccharide content of the target molecule is performed by a number of techniques. The sugars are first removed from the amino acid backbone by enzymatic (such as digestion with PNGase)) or chemical (such as beta elimination with hydroxide) means. The sugars may be stabilised by reduction or labeled with a fluorophore for ease of detection. The resultant free oligosaccharides are then separated either by high pH anion exchange chromatography with pulsed amperometric electrochemical detection (HPAEC-PAD), which can be used with known standards to determine the ratios of the various structures and levels of sialylation, or by fluorophore assisted carbohydrate electrophoresis (FACE) a process similar to SDS-PAGE separation of proteins. In this process the oligosaccharides are labeled with a fluorophore that imparts electrophoretic

mobility. They are separated on high percentage polyacrylamide gels and the resultant band pattern provides a profile of the oligosaccharide content of the target molecule. By using standards it is possible to gain some information on the actual structures present or the bands can be cut and analysed using mass spectrometry to determine each of their structures.

Fluorophore assisted carbohydrate electrophoresis (FACE) is a polyacrylamide gel electrophoresis system designed to separate individual oligosaccharides that have been released from a glycoconjugate. Oligosaccharides are removed from the sample protein by either chemical or enzymatic means in such a way as to retain the reducing terminus. Oligosaccharides are then either digested into monosaccharides or left intact and labeled with a fluorophore (either charged or non charged). High percentage polyacrylamide gels and various buffer systems are used to migrate the oligosaccharides/monosaccharides which migrate relative to their size/composition in much the same way as proteins. Sugars are visualised by densitometry and relative amounts of sugars can be determined by fluorophore detection. This process is compatible with MALDI-TOF MS, hence the method can be used to elucidate actual structures.

Quartz crystal microbalance and surface plasmon resonance (QCM and SPR, respectively) are two methods of obtaining biological information through the physiochemical properties of a molecule. Both measure protein-protein interactions indirectly through the change that the interaction causes in the physical characteristics of a prefabricated chip. In QCM a single crystal quartz wafer is treated with a receptor/antibody etc which interacts with the ligand of interest. This chip is oscillated by the microbalance and the frequency of the chip recorded. The protein of interest is allowed to pass over the chip and the interaction with the bound molecule causes the frequency of the wafer to change. By changing the conditions by which the ligand interacts with the chip, it is possible to determine the binding characteristics of the target molecule.

Apparent molecular weight is also a physiochemical property which can be used to determine the similarities between the protein or chimeric molecule of the present invention and those produced using alternative means.

As used herein, the term "molecular weight" is defined as the sum of atomic weights of the constituent atoms in a molecule, sometimes also referred to as "molecular mass" (Mr). Molecular weight can be determined theoretically by summing the atomic masses of the constituent atoms in a molecule. The term "apparent molecular weight" is defined as the molecular weight determined by one or more analytical techniques such as SDS page or ultra centrifugation and depends on the relationship between the molecule and the detection system. The apparent molecular weight of a protein or chimeric molecule thereof can be determined using any one of a range of experimental methods. Analytical methods for determining the molecular weight of a protein include, without being limited to, size-exclusion chromatography (SEC), gel electrophoresis, Rayleigh light scattering, analytical ultracentrifugation, and, to some extent, time-of-flight mass spectrometry.

Gel electrophoresis is a process of determining some of the physiochemical properties (specifically apparent molecular weight and pI) of a protein and in the case of 2 dimensional electrophoresis to separate the molecule into isoforms, thereby providing information on the post-translational modifications of the protein product. Specifically, electrophoresis is the process of forcing a charged molecule (such as protein or DNA) to migrate through a gel matrix (most commonly polyacrylamide or agarose) by applying an electric potential through its body. The most common forms of electrophoresis used on proteins are isoelectric focussing, native, and SDS polyacrylamide gel electrophoresis. In isoelectric focussing a protein is placed into a polyacrylamide gel that has a pH gradient across its length. The protein will migrate to the point in the gel where it has a net charge of zero thereby giving its isoelectric point.

Glyco mass fingerprinting (GMF) is the process by which the oligosaccharide profile of a protein or one of its isoforms is identified by electrophoresis followed by specific mass spectrometric techniques. Sample protein is purified either by 1D SDS-PAGE for total profile determination or 2D gel electrophoresis for specific isoform characterization. The protein band/spot is excised from the gel and de-stained to remove contaminants. The sugars are released by chemical or enzymatic means and desalted/separated using a nanoflow LC system and a graphitised carbon column. The LC flow can be directly

injected into an electrospray mass spectrometer that is used to determine the mass and subsequently the identity of the oligosaccharides present on the sample. This provides a profile or fingerprint of each isoform which can be combined with quantitative techniques such as Dionex analysis to determine the total composition of the molecule being tested.

Primary structure can be evaluated in determining the physiochemical properties of the protein or chimeric molecule of the present invention.

The primary structure of a protein or chimeric molecule thereof can be assayed using one or more of the following systems.

Information on the primary structure of a protein or chimeric molecule thereof can be determined using a combination of mass spectrometry (MS), DNA sequencing, amino acid composition, protein sequencing and peptide mass fingerprinting.

To determine the sequence of the amino acid backbone either N-terminal chemical sequencing, tandem mass spectrometry sequencing, or a combination of both is used. N-terminal chemical sequencing utilises Edman chemistry (Edman P "Sequence determination" *Mol Biol Biochem Biophys 8:*211-55, 1970), which states that the peptide bond between the N-terminal amino acid and the amino acid in position 2 of the protein is weaker than all other peptide bonds in the sequence. By using moderate acidic conditions the N-terminal amino acid is removed derivatised with a fluorophore (FTIC) and the retention time on a reversed-phase HPLC column determined, and compared to a standard to identify what the amino acid is. This method will determine the actual primary structure of the molecule but is not quantitative. Alternatively tandem mass spectrometry in conjunction with nanoflow liquid chromatography may be used (LC-MS/MS). In this process the protein is digested into peptides using specific endoproteases and the molecular weight of the peptides determined. High energy collision gases such as nitrogen or argon are then used to break the peptide bonds and the masses of the resultant peptides measured. By calculating the change in mass of the peptides it is possible to determine the sequence of each of the peptides (each amino acid has a unique mass). By using different proteases

the peptides may then be overlapped to determine their order and thus the entire sequence of the protein.

Clearly, the combination of enzymatic digestion, chemical derivatization, liquid chromatography (LC)/MS and tandem MS provides an extremely powerful tool for AA sequence analysis. For example, the detailed structure of recombinant soluble CD4 receptor was characterized by a combination of methods, which confirmed over 95% of the primary sequence of this 369 AA glycoprotein and showed the whole nature of both N-and C-termini, the positions of attachment of the glycans, the structures of the glycans and the correct assignment of the disulfide bridges (Carr *et al. J Biol Chem 264(35)*:21286-21295, 1989).

Mass spectrometry (MS) is the process of measuring the mass of a molecule through extrapolation of its behavior in a charged environment under a vacuum. MS is very useful in stability studies and quality control. The method first requires digestion of samples by proteolytic enzymes (trypsin, V8 protease, chymotrypsin, subtilisin, and clostripain) (Franks *et al. Characterization of proteins*, Humana Press, Clifton, NJ, 1988; Hearn *et al. Methods in Enzymol 104*:190-212, 1984) and then separation of digested samples by reverse phase chromatography (RPC). With tryptic digestion in conjunction with LC-MS, the peptide map can be used to monitor the genetic stability, the homogeneity of production lots, and protein stability during fermentation, purification, dosage form manufacture and storage.

Before a mass analysis, several ways are used to interface a HPLC to a mass spectrometer: 1) direct liquid injection; 2) supercritical fluid; 3) moving belt system; 4) thermospray. The HPLC/MS interface used in Caprioli's work used a fused silica capillary column to transport the eluate from the column to the tip of the sample probe in the ionization chamber of the mass spectrometer. The probe tip is continuously bombarded with energetic Xe atoms, causing sputtering of the sample solution as it emerges from the tip of the capillary. The mass is then analyzed by the instrument (Caprioli *et al. Biochem Biophys Res Commun 146*:291-299, 1987).

MS/MS and LC/MS interfaces expand the potential applications of MS. MS/MS allows direct identification of partial to full sequence for peptides up to 25 AAs, sites of deamidation and isomerization (Carr *et al. Anal Chem 63*:2802-2824, 1991). Coupled with RPC or capillary electrophoresis (CE), MS can perform highly sensitive analysis of proteins (Figeys and Aebersold, *Electrophoresis 19*:885-892, 1998; Nguyen *et al. J Chromatogr A 705*:21-45, 1995). LC/MS allows LC methodology to separate peptides before entering the MS, such as the continuous flow FAB interfaced with microbore HPLC (Caprioli *et al.* 1987, *supra*). The latter "interface" allows the sequencing of individual peptides from complex mixtures: Fragmentation of the peptides selected by the first MS is followed by passing through a cloud of ions in a collision cell: CID (collision induced dissociation). The collision generates characteristic set of fragments, from which the sequence may be deduced, without knowing other information, such as the cDNA sequence. In a single MS experiment, an unfractionated mixture of peptides (e.g. from an enzyme digest) is injected and the masses of the major ions are compared with those predicted from the cDNA sequence. The sequence of the recombinant human interleukin-2 was verified by fast atom bombardment (FAB)-MS analysis of CNBr and proteolytic digests (Fukuhara *et al. J Biol Chem 260*:10487-10494, 1985).

Electrospray ionization MS (ESI-MS) uses an aerosol of solution protein to introduce into a needle under a high voltage, generating a series of charged peaks of the same molecules with various charges. Because each peak generated from the differently charged species produces an estimation of the molecular weights, these estimations can be combined to increase the overall precision of the molecular weight estimation. Matrix Assisted Laser Desorption Ionization MS (MALDI-MS) uses a high concentration of a chromophore. A higher intensity laser pulse will be absorbed by the matrix and the energy absorbed evaporates part of the matrix and carries the protein sample with it into the vapor phase almost entirely. The resulting ions are then analyzed in a time of flight MS. The mild ionization may enhance the capacity of the method to provide quaternary structure information. MALDI-MS can be run rapidly in less than 15 minutes. It does not need to fragment the molecules and the result is easy to interpret as a densitometric scan of an SDS-PAGE gel, with a mass range up to over 100kDa.

Amino acid sequence can be predicted by sequencing DNA that encodes a protein or chimeric molecule thereof. However, occasionally the actual protein sequence may be different. Traditionally, DNA sequencing reactions are just like the PCR reactions for replicating DNA (DNA denaturation, replication). By DNA cloning technology, the gene is cloned, and the nucleotide sequence determined.

The amino acid sequence of a protein or chimeric molecule thereof can be assayed using one or more of the following systems.

Full sequence description of the protein or chimeric molecule thereof is usually required to describe the product. Amino acid sequencing includes: in gel tryptic digestion, fractionation of the digested peptides by RPC-HPLC, screening the peptide peaks that have the most symmetrical absorbance profile by MALDI-TOF MS, and the first peptide (N-terminal) by Edman degradation. Edman chemically derived primary sequence data is the classical method to identify proteins at the molecular level. MALDI-TOF MS can be used for N-terminal sequence analysis. However, all enzymatic digests for HPLC and peptide sequencing are recommended to first be subjected to MS or MS/MS protein identification to decrease the time and cost. The internal amino acid sequences from SDS-PAGE-separated proteins are obtained by elution of the peptides with HPLC separation after an in situ tryptic or lysyl endopeptidase digestion in the gel matrix.

Internal sequencing of the standard peptide is recommended to be run with the analyzed samples to maintain the instruments at the peak performance. More than 80% of higher eukaryotic proteins are reported to have blocked amino-termini that prevent direct amino acid sequencing. When a blocked eukaryotic protein is encountered, the presence of the sequence of the internal standard assures that the instrument is operating properly.

Edman degradation can be used for direct N-terminal sequencing with a chemical procedure, which derivatizes the N-terminal amino acids to release the amino acids and expose the amino terminal of the next AAs. The Edman sequencing includes: 1). By microbore HPLC, N-terminal sequence analysis is repeated by Edman chemistry cycles. Every cycle of the Edman chemistry can identify one amino acid. 2). After in-gel or

PVDF bound protein digestions followed by HPLC separation of the resulting peptides, internal protein sequence analysis is conducted by Edman degradation chemistry.

Microbore HPLC and capillary HPLC are used for analysis and purification of peptide mixtures using RPC-HPLC. In-gel samples and PVDF samples are purified using different columns. MALDI-TOF MS analysis can be used for N-terminal analysis after HPLC fractionation. The selection criteria are: 1) The apparent purity of the HPLC fraction. 2) The mass and thus the estimated length of the peptide. The peptide mass information is useful for confirming the Edman sequencing amino acid assignments, and also in the possible detection of co- or post-translational modifications.

In-gel digests are suitable for purification on the higher sensitivity HPLC system. The internal protein sequence analysis is first enzymatically digested by SDS-PAGE. Proteins in an SDS-PAGE mini-gel can be reliably digested in-gel only with trypsin. The peptide fragments are purified by RPC-HPLC and then analyzed by MALDI-TOF MS, screening for peptides suitable for Edman sequence analysis. Proteins in a gel can only be analyzed by internal sequencing analysis, but very accurate peptides masses can be obtained, which provides additional information useful in both amino acid assignment and database searching.

PVDF-bound proteins are suitable for both N-terminal and internal Edman sequencing analysis. PVDF-bound proteins are digested with the proper enzyme (trypsin, endoproteinase Lys-C, endoproteinase Glu-C, clostripain, endoproteinase Asp-N, thermolysin) and a non-ionic detergent such as hydrogenated Triton X-100. In PVDF bound proteins, the detergents used for releasing digested peptides from the membrane can interfere with MALDI-TOF MS analysis. Before the enzyme is added, Cys is reduced with DTT and alkylated with iodoacetamide to generate carboxyamidomethyl Cys, which can be identified during N-terminal sequence analysis.

To determine the amino acid composition of a protein or chimeric molecule thereof, the sample is hydrolyzed using phenol catalyzed strong hydrochloric acid (HCl) acidic conditions in the gaseous phase. Once the hydrolysis is performed the liberated amino

acids are derivatised with a fluorophore compound that imparts a specific reversed phase characteristic on the combined molecule. The derivatized amino acids are separated using reversed phase high performance liquid chromatography (RP-HPLC) and detected with a fluorescence detector. By using external and internal standards it is possible to calculate the amount of each amino acid present in the sample from the observed peak area. This information may be used to determine sample identity and to quantify the amount of protein present in the sample. For instance, discrepancies between theoretical and actual results can be used to initially identify the possibility of a de-amidation site. In combination with monosaccharide analysis it may determine the composition % by weight glycosylation and percent by weight acidic glycoproteins. This method is limited in the information that it can provide on the actual sequence of the backbone however as there is inherent variability due to environmental contaminants and occasional destruction of amino acids. For example, it is not possible for this method to detect point mutations in the sequence.

Peptide mass fingerprinting (PMF) is another method by which the identity of a protein or chimeric molecule thereof may be determined. The procedure involves an initial separation of the sample by electophoretic means (either 1 or 2 dimensional), excision of the spot/band from the gel and digestion with a specific endoprotease (typically porcine trypsin). Peptides are eluted from the gel fragment and analysed by mass spectrometry to determine the peptide masses present. The resultant peptide masses are then compared to a database of theoretical mass fragments for all reported proteins (or in the case of constructs for the theoretical peptide masses of the designed sequence). The technique relies on the fact that the "fingerprint" of a protein (i.e. its combination of peptide masses) is unique. Identity can be confidently determined (greater than 90% accuracy) with as little as 4 peptides and 30% sequence coverage. Modifications such as lipid moieties and de-amidation can be identified during the PMF stage of analysis. Peaks that do not correspond to those of the identified protein are further analysed by tandem mass spectrometry (MS-MS), a technique that uses the energy created by the impact of a collision gas to break the weaker bond of the PTM. The newly freed molecule and the original peptide are then re-analysed for mass to identify the post-translational modification and the peptide fragment to which it was attached.

HPLC is classified into different modes depending on the size, charge, hydrophobicity, function or specific content of the target biomolecules. Generally, two or more chromatographic methods are used to purify a protein. It is of paramount importance to consider both the characteristics of the protein and the sample solvent when the chromatographic modes are selected.

Secondary structures of a protein or chimeric molecule of the present invention can also be evaluated in characterising their properties.

The secondary structure of a protein or chimeric molecule thereof can be assayed using one or more of the following systems.

To study the secondary structures of proteins, most commonly several spectroscopic methods should be applied and compared. Electromagnetic energy can be defined as a continuous waveform of radiation, depending on the size and shape of the wave. Different spectroscopic methods use different electromagnetic energy.

The wavelength, is the extent of a single wave of radiation (the distance between two successive maxima of the waves). When the radiant energy increases, the wavelength becomes shorter. The relationship between frequency and wavenumber is:

$$\text{Wavenumber } (cm^{-1}) = \text{Frequency } (s^{-1}) / \text{ The speed of light (cm/s).}$$

The absorption of electromagnetic radiation by molecules includes vibrational and rotational transitions, and electronic transitions. Infrared (IR) and Raman spectroscopy are most commonly used to measure the vibrational energies of molecules in order to determine secondary structure. However, they are different in their approach to determine molecular absorbance.

The energy of the scattered radiation is less than the incident radiation for the Stokes line. The energy of the scattered radiation is more than the incident radiation for the anti-Stokes

line. The energy increase or decrease from the excitation is related to the vibrational energy spacing in the ground electronic state of the molecule. Therefore, the wavenumber of the Stokes and anti-Stokes lines are a direct measure of the vibrational energies of the molecule.

Only the Stokes shift is observed in a Raman spectrum. The Stokes lines are at smaller wavenumbers (or higher wavelengths) than the exciting light. A high power excitation source, such as a laser, should be used to enhance the efficiency of Raman scattering. The excitation source should be monochromatic because we are interested in the energy (wavenumber) difference between the excitation and the Stokes lines.

For a vibrational motion to be IR active, the dipole moment of the molecule must change. Therefore, the symmetric stretch in carbon dioxide is not IR active because there is no change in the dipole moment. The asymmetric stretch is IR active due to a change in dipole moment. For a vibration to be Raman active, the polarizability of the molecule must change with the vibrational motion. The symmetric stretch in carbon dioxide is Raman active because the polarizability of the molecule change. Thus, Raman spectroscopy complements IR spectroscopy (Herzberg *et al. Infrared and Raman Spectra of Polyatomic Molecules,* Van Nostrand Reinhold, New York, NY, 1945). For example, IR is not able to detect a homonuclear diatomic molecule due to the lack of dipole moments, but Raman spectroscopy can detect it because the molecular polarizability is changed by stretching and contraction of the bond, further, the interactions between electrons and nuclei are changed.

For highly symmetric polyatomic molecules with a center of inversion (such as benzene), it is more likely that bands active in the IR spectrum are not active in the Raman spectrum or vice-versa. In molecules with little or no symmetry, modes are likely to be active in both infrared and Raman spectroscopy.

IR spectroscopy measures the wavelength and intensity of the absorption of infrared light by a sample. Infrared light is so energetic that it can excite the molecular vibrations to

higher energy levels. Both infrared and RAMAN spectroscopy measure the vibrations of bond lengths and angles.

IR characterizes vibrations in molecules by measuring the absorption of light of certain energies corresponding to the vibrational excitation of the molecule from v = 0 to v = 1 (or higher) states. There are selection rules that govern the ability of a molecule to be detected by infrared spectroscopy - Not all of the normal modes of vibration can be excited by infrared radiation (Herzberg *et al.* 1945, *supra*).

IR spectra can provide qualitative and quantitative information of the secondary structures of proteins, such as α-helix, β-sheet, β-turn and disordered structure. The most informative IR bands for protein analysis are amide I (1620-1700 cm^{-1}), amide II (1520-1580 cm^{-1}) and amide III (1220-1350 cm^{-1}). Amide I is the most intense absorption band in proteins. It consists of stretching vibration of the C=O (70-85% and C-N groups (10-20%). The exact band position is dictated by the backbone conformation and the hydrogen bonding pattern. Amide II is more complex than Amide I. Amide II is governed by in-plane N-H bending (40-60%), C-N (18-40%) and C-C (10%) stretching vibrations. Amide III bands are not very useful (Krimm and Bandekar, *Adv Protein Chem 38:*181-364, 1986). Most of the β-sheet structures of FTIR amide I band usually are located at about 1629 cm^{-1} with a minimum of 1615 cm^{-1} and a maximum of 1637 cm^{-1}; the minor component may show peaks around 1696 cm^{-1} (lowest value 1685 cm^{-1}). α-helix is mainly found at 1652 cm^{-1}. An absorption near 1680 cm^{-1} is now assigned to β-turns.

The principle of Raman scattering is different from that of infrared absorption. Raman spectroscopy measures the wavelength and intensity of inelastically scattered light from molecules. The Raman scattered light occurs at wavelengths that are shifted from the incident light by the energies of molecular vibrations.

To be Raman active, for the vibration to be inelastically scattered, a change in polarizability during the vibration is essential. In the symmetric stretch, the strength of electron binding is different between the minimum and maximum internuclear distances. Therefore the polarizability changes during the vibration, and this vibrational mode

scatters Raman light, the vibration is Raman active. In the asymmetric stretch the electrons are more easily polarized in the bond that expands but are less easily polarized in the bond that compresses. There is no overall change in polarizability and the asymmetric stretch is Raman inactive (Herzberg *et al.* 1945, *supra*).

Circular dichroism can be used to detect any asymmetrical structures, such as proteins. Optically active chromophores absorb different amount of right and left polarized light, this absorbance difference results in either a positive or negative absorption spectrum (Usually, the right polarized spectrum is subtracted from the left polarized spectrum). Commonly, the far UV or amide region (190-250nm) is mainly contributed from peptide bonds, providing information on the environment of the carbonyl group of the amide bond and consequently the secondary structure of the protein. α-helix usually displays two negative peaks at 208, 222 nm (Holzwarth *et al. J Am Chem Soc 178*:350, 1965), β-sheet displays one negative peak at 218 nm, random coils has a negative peak at 196 nm. Near UV region peaks are (250-350 nm) contributed from the environment of the aromatic chromophores (Phe, Tyr, Trp). Disulfide bonds give rise to minor CD bands around 250 nm.

Intense dichroism is commonly associated with the side-chain structures being held tightly in a highly folded, three-dimensional structure. Denaturation of the protein mostly releases the steric hindrance; a weaker CD spectrum is obtained along with an increasing degree of denaturation. For example, the side chain CD spectrum of hGH is quite sensitive to the partial denaturation by adding denaturants. Some reversible chemical alterations of the molecules, such as reduction of the disulfide bonds, or alkaline titrations will change the side-chain CD spectrum. For hGH, these spectral difference can be caused by entirely the removal of a chromophores, or by affecting changes in the particular chromophore's CD response, but not by the gross denaturation or conformational changes (Aloj *et al. J Biol Chem 247*:1146-1151, 1971).

UV absorption spectroscopy is one of the most significant methods to determine protein properties. It can provide information about protein concentrations and the immediate environments of chromophoric groups. Proteins functional groups, such as amino,

alcoholic (or phenolic) hydroxyl, carbonyl, carboxyl, or thiol can be transformed into strong chromophores. Visible and near UV spectroscopy are used to monitor two types of chromophores: metalloproteins (more than 400 nm) and proteins that contains Phe, Trp, Tyr residues (260-280nm). The change in UV or fluorescence signal can be negative or positive, depending on the protein sequence and solution properties.

Fluorescence measures the emission energy after the molecule has been irradiated into an excited state. Many proteins emitted fluorescence in the range of 300 to 400 nm when excited at 250 to 300 nm from their aromatic amino acids. Only proteins with Phe, Trp, Tyr residues can be measured with the order of intensity Trp>> Tyr>> Phe. Fluorescence spectra can reflect the microenvironments information that are affected by the folding of the proteins. For example, a buried Trp is usually in a hydrophobic environment and will fluoresce at maximum 325 to 330 nm range, but an exposed residue or free amino acids fluoresces at around 350 to 355 nm. An often used agent to probe protein unfolding is Bis-ANS. The fluorescence of Bis-ANS is pH-independent. Even though its signal is weak in water, it can be increased significantly by binding to unfolding-exposed hydrophobic sites in proteins (James and Bottomley *Arch Biochem Biophy 356*:296-300, 1998).

Effective quenching of Tyr and Trp in the folded proteins causes significant signal increase upon unfolding. A simple solute may cause the change also. To maximize detection sensitivity, a signal ratio can be used. For example, In the study of rFXIII unfolding, a ratio of fluorescence intensity at 350nm to that at 330nm was used (Kurochkin *et al. J Mol Biol 248*:414-430, 1995). Conformational changes may be studied by means of excitation-energy transfer between a fluorescent donor and an absorbing acceptor, because the efficiency of transfer depends on the distance between the two chromophores (Honroe *et al. Biochem J 258*:199-204, 1989). Fluorescence was used to probe a-Antitrypsin conformation (Kwon and Yu, *Biophim Biophys Acta 1335*:265-272, 1997), to determine Tm of HSA (Farruggia *et al. Int J Biol Macromol 20:*43-51, 1997), and to detect MerP unfolding interactions (Aronsson *et al. FEBS Lett 411*:359-364, 1997).

At neutral pH, the intensity of the fluorescence emission spectrum is in the order of Trp> Tyr. At acidic pH, due to the conformational changes which disrupts the energy transfer,

the fluorescence from Tyr dominates over Trp. Fluorescence studies also confirm the presence of intermediates in the guanidine-induced unfolding transition of the proteins.

Tertiary and quaternary structures of the physiochemical forms of a protein or chimeric molecule of the present invention are also important in ascertaining their function.

The tertiary and quaternary structures of a protein or chimeric molecule thereof can be assayed using one or more of the following systems.

NMR and X-ray crystallography are the most often used techniques to study the 3D structure of proteins. Other less detailed methods to probe protein tertiary structure include CD in near UV region, second-derivative of UV spectroscopy (Ackland *et al. J Chromatogr 540*:187-198, 1991) and fluorescence.

NMR is one of the main experimental methods for molecular structure and intermolecular interactions in structural biology. In addition to studying protein structures, NMR can also be utilised to study the carbohydrate structures of a protein or chimeric molecule of the present invention. NMR spectroscopy is routinely used by chemists to study chemical structure using simple one-dimensional techniques. The structure of more complicated molecules can also be determined by two-dimensional techniques. Time domain NMR are used to probe molecular dynamics in solutions. Solid state NMR is used to determine the molecular structure of solids. NMR can be used to study structural and dynamic properties of proteins, nucleic acids, a variety of low molecular weight compounds of biological, pharmacological and medical interests. However, not all nuclei possess the correct property in order to be read by NMR, i.e., not all nuclei posses spin, which is required for NMR. The spin causes the nucleus to produce an NMR signal, functioning as a small magnetic field.

The crystal structure of a protein or chimeric molecule thereof can be assayed using one or more of the following systems.

X-ray crystallography is an experimental technique that applies the fact that X-rays are diffracted by crystals. X-rays have the appropriate wavelength (in the Ångström range, ~10-8 cm) to be scattered by the electron cloud of an atom of comparable size. The electron density can be reconstructed based on the diffraction pattern obtained from X-ray scattering off the periodic assembly of molecules or atoms in the crystal. Additional phase information either from the diffraction data or from supplementing diffraction experiments should be obtained to complete the reconstruction. A model is then progressively built into the experimental electron density, refined against the data and the result is a very accurate molecular structure.

X ray diffraction has been developed to study the structure of all states of matter with any beam, e.g., ions, electrons, neutrons, and protons, with a wavelength similar to the distance between the atomic or molecular structures of interest.

Light scattering spectroscopy is based on the simple principle that larger particles scatter light more than the smaller particles. A slope base line in the 310-400nm region originates from light scattering when large particles, such as aggregates, present in the solution (Schmid *et al. Protein structure, a practical approach*, Creighton Ed., IRI Press, Oxford, England, 1989)

Light scattering spectroscopy can be used to estimate the molecular weight of a protein and is a simple tool to monitor protein quaternary structure or protein aggregation. The degree of protein aggregation can be indicated by simple turbidity measurement. Final product pharmaceutical solutions are subjected to inspection of clarity because most aggregated proteins are present as haze and opalescence. Quasielastic light scattering spectroscopy (QELSS), sometimes called photon correlation spectroscopy (PCS), or dynamic light scattering (DLS), is a noninvasive probe of diffusion in complex fluids for macromolecules (proteins, polysaccharides, synthetic polymers, micelles, colloidal particles and aggregations). In most cases, light scattering spectroscopy yields directly the mutual diffusion coefficient of the scattering species. When applied to dilute monodisperse solutions, the diffusion coefficient obtained by QELSS can estimate the size. With polydisperse system, it estimates the width of molecular weight distribution. For accurate

measurement, 200-500 mW laser power is mandatory, conventional Ar+/Kr+ gas lasers are widely used (Phillies *Anal Chem 62:*1049A-1057A, 1990). Protein aggregation was detected by human relaxin (Li *et al. Biochemistry 34:*5762-5772, 1995).

Stability of a protein or chimeric molecule thereof is also an important determinant of function. Methods for analysing such characteristics include DSC, TGA and freeze-dry cryostage microscopy, analysis of freeze-thaw resistance, and protease resistance.

A protein or chimeric molecule of the present invention may be more stable for lyophilization (freeze drying). Lyophilization is used to enhance the stability and/or shelf life of the product as it is stored in powder rather than liquid form. The process involves an initial freezing of the sample, then removal of the liquid by evaporation under vacuum. The end result is a dessicated "cake" of protein and excipients (other substances used in the formulation). The consistency of the resulting cake is critical for successful reconstitution. The lyophilization process can result in changes to the protein, especially aggregate formation though crosslinking, but also deamidation and other modifications. These can reduce efficacy by either losses, reduced activity or by inducing immune reactions against aggregates. In order to test lyophilization stability, the protein can be formulated for lyophilization using standard stabilizers (e.g. mannitol, trehalose, Tween 80, human serum albumin and the like). After lyophilization, the amount of protein recovered can be assayed by ELISA, while its activity can be assayed by a suitable bioassay. Aggregates of the protein can be detected by HPLC or Western blot analysis.

Prior to lyophilization, the Tg or Te (define Tg or Te) of the formulation should be determined to set the maximum allowable temperature of the product during primary drying. Also, information about the crystallinity or amorphousity of the formulation helps to design the lyophilization cycle in a more rationale manner. Product information on these thermal parameters can be obtained by using differential scanning calorimetry (DSC), thermogravimetric analysis (TGA) or freeze-dry cryostage microscope.

Differential Scanning Calorimetry (DSC) is a physical thermo-analytical method to measure, characterize and analyze thermal properties of materials and determine the heat

capacities, melting enthalpies and transition points accordingly. DSC scans through a temperature range at a linear rate. Individual heaters within the instrument provide heat to sample and reference pans separately, based on the "power compensated null balance" principle. During a physical transition, the absorption or evolution of the energy causes an imbalance in the amount of energy supplied to that of the sample holder. Depending on the varying thermal behavior of the sample, the energy will be taken or diffused from the sample, and the temperature difference will be sensed as an electrical signal to the computer. As a result, an automatic adjustment of the heaters makes the temperature of the sample holder identical to the reference holder. The electrical power needed for the compensation is equivalent to the calorimetric effect.

The purity of an organic substance can be estimated by DSC based on the shape and temperature of the DSC melting endotherm. The power-compensated DSC provides very high resolution compared to a heat flux DSC under the identical conditions. More well-defined and more accurate partial areas of melting can be generated from power-compensated DSC because the partial areas of melting are not "smeared" over a narrow temperature interval, as for the lesser-resolved heat flux DSC. The power-compensated DSC produces inherently better partial melting areas and therefore better purity analysis. By the help of StepScan DSC, the power-compensated DSC can provide a direct heat capacity measurement using the traditional and time-proven means without the need for deconvolution or the extraction of sine wave amplitudes.

Thermogravimetric Analysis (TGA) measures sample mass loss and the rate of weight loss as a function of temperature or time.

As DSC, freeze-dry cryostage can reach a wide temperature range rapidly. Currently, as an preformulation and formulation study tool, simulating the lyophilization cycle in a freeze dry cryostage provides the best platform to study thermal parameters of the protein formulations on a miniature scale. Freeze dry microscope can predict the influence of formulations and process factors on freezing and drying. Only a 2-3mL sample is required for a cryostage study, which makes this technique a valuable tool to study scarce, difficult-to-obtain drugs. It is a good tool to study the effect of freezing, rate, drying rate, thawing

rate on the lyophilization cycle. Annealing research may be advanced by the studies from freeze-dry cryostage microscope. Because of extensive applications of lyophilization technology, and larger demand to stabilize the extremely expensive drugs (such as proteins and gene therapy drugs), it is expected that an in-process microscopic monitor should be realized in the pharmaceutical industries soon.

The freeze-thaw resistance of a protein or chimeric molecule thereof can be assayed using one or more of the following systems.

Co- or post translational modification such as glycosylation may protect proteins from repeated freeze/thaw cycles. To determine this, a protein or chimeric molecule of the present invention can be compared to carrier-free *E. coli*-produced counterparts. A protein or chimeric molecule thereof are diluted into suitable medium (e.g. cell growth medium, PBS or the like) then frozen by various methods, for instance, snap frozen in liquid nitrogen, slowly frozen by being placed at –70 degrees or rapidly frozen on dry ice. The samples are then thawed either rapidly at room temperature or slowly at 4 degrees. Some samples are then refrozen and the process are repeated for a number of cycles. The amount of protein present can be measured by ELISA, and the activity measured in a suitable bioassay chosen by a skilled artisan. The amount of activity/protein remaining is compared to the starting material to determine the resistance over many the freeze/thaw cycles.

A protein or chimeric molecule of the present invention may have altered thermal stability in solution. The thermal stability of the present invention may be determined *in vitro* as follows.

A protein or chimeric molecule of the present invention can be mixed into buffer e.g. phosphate buffered saline containing carrier protein e.g. human serum albumin and incubated at a particular temperature for a particular time (e.g. 37 degrees for 7 days). The amount of protein or chimeric molecule thereof remaining after this treatment can be determined by ELISA and compared to material stored at –70 degrees. The biological activity of the remaining protein or chimeric molecule thereof is determined by performing a suitable bioassay chosen by a person skilled in the relevant art.

The protease resistance of a protein or chimeric molecule thereof can be assayed using one or more of the following systems.

To compare protease resistance, solution containing a protein or chimeric molecule of the present invention and solution containing *E. coli* expressed counterparts can be incubated with a protease of choice (e.g. unpurified serum proteases, purified proteases, recombinant proteases) for different time periods. The amount of protein remaining is measured by an appropriate ELISA (e.g. one in which the epitopes recognized by the capture and detection antibodies are separated by the protease cleavage site), and the activity of the remaining protein or chimeric molecule thereof is determined by a suitable bioassay chosen by a skilled artisan.

The bioavailability of a protein or chimeric molecule thereof can be assayed using one or more of the following systems.

Bioavailability is the degree to which a drug or other substance becomes available to the target tissue after administration. Bioavailability may depend on half life of the drug or its ability to reach the target tissue.

Compositions comprising a protein or chimeric molecule of the present invention is injected subcutaneously or intramuscularly. The levels of the protein or its chimeric molecule can then be measured in the blood by ELISA or radioactive counts. Alternatively, the blood samples can be assayed for activity of the proteinby a suitable bioassay chosen by a skilled artisan, for instance, stimulation of proliferation of a particular target cell population. As the sample will be from plasma or serum, there may be a number of other molecules that could be responsible for the output activity. This can be controlled by using a neutralizing antibody to the protein being tested. Hence, any remaining bioactivity is due to the other serum components.

The stability or half-life of a protein or chimeric molecule thereof can be assayed using one or more of the following systems.

A protein or chimeric molecule of the present invention may have altered half-life in serum or plasma. The half-life of the present invention may be determined *in vitro* as follows. Composition containing the protein or chimeric molecule of the present invention can be mixed into human serum/plasma and incubated at a particular temperature for a particular time (e.g. 37 degrees for 4 hours, 12 hours etc). The amount of protein or chimeric molecule thereof remaining after this treatment can be determined by ELISA. The biological activity of the remaining protein or chimeric molecule thereof is determined by performing a suitable bioassay chosen by a person skilled in the relevant art. The serum chosen may be from a variety of human blood groups (e.g. A, B, AB, O etc.)

The half-life of a protein or chimeric molecule thereof can also be determined *in vivo*. Composition containing a protein or chimeric molecule thereof, which may be labeled by a radioactive tracer or other means, can be injected intravenously, subcutaneously, retro-orbitally, tail vein, intramuscularly or intraperitoneally) into the species of choice for the study, for instance, mouse, rat, pig, primate, human. Blood samples are taken at time points after injection and assayed for the presence of the protein or chimeric molecule thereof (either by ELISA or by TCA-precipitable radioactive counts). A comparison composition consisting of *E. coli* or CHO-produced protein or chimeric molecule thereof can be run as a control.

To determine the half-life of protein or chimeric molecule of the present invention, *in vivo*, male Wag/Rij rats, or other suitable animals can be injected intravenously with a protein or chimeric molecule thereof.

Just before the administration of the substrate, 200µl of EDTA blood is sampled as negative control. At various time points after the injection, 200µl EDTA blood can be taken from the animals using the same technique. After the last blood sampling, the animals are sacrificed. The specimen is centrifuged for 15 min at RT within 30 min of collection. The plasma samples are tested in a specific ELISA to determine the concentration of protein or chimeric molecule of the present invention in each sample.

A protein or chimeric molecule of the present invention may cross the blood brain barrier.

An *in vitro* assay to determine if protein or chimeric molecule of the present invention binds human brain endothelial cells can be tested using the following assays.

Radiolabeled protein or chimeric molecule of the present invention can be tested for its ability to bind to human brain capillary endothelial cells. An isolated protein or chimeric molecule of the present invention can be custom conjugated with radiolabel to a specific activity using a method known in the art, for instance, with ^{125}I by the chloramine T method, or with ^{3}H.

Primary cultures of human brain endothelial cells can be grown in flat-bottom 96-well plates until five days post-confluency then lightly fixed using acetone. Cells are lysed, transferred to glass fibre membranes. Radiolabeled protein or chimeric molecule of the present invention can be detected using a liquid scintillation counter.

In vivo assays for the determination of protein or chimeric molecule of the present invention binding to human brain endothelial cells can be tested using the following assays.

A human-specific protein or chimeric molecule of the present invention are tested for binding to human brain capillaries using sections of human brain tissue that are fresh frozen (without fixation), sectioned on a cryostat, placed on glass slides and fixed in acetone. Binding of ^{3}H-protein or chimeric molecule of the present invention is examined on brain sections using quantitative autoradiography.

In vivo assay can be used to measure tissue distribution and blood clearance of human-specific protein or chimeric molecule of the present invention in a primate system.

A protein or chimeric molecule of the present invention is used to determine the tissue distribution and blood clearance of ^{14}C -labeled protein or chimeric molecule of the present invention in 2 male cynomolgus monkeys or other suitable primates. protein or chimeric molecule of the present invention is administered concurrently with a ^{3}H -labeled control protein to the animals with an intravenous catheter. During the course of the study, blood samples are collected to determine the clearance of the proteins from the circulation. At 24

hours post-injection, the animals are euthanized and selected organs and representative tissues collected for the determination of isotope distribution and clearance by combustion. In addition, capillary depletion experiments are performed to samples from different regions of the brain in accordance with Triguero *et al. J of Neurochemistry 54*:1882-1888, 1990. This method removes greater than 90% of the vasculature from the brain homogenate (Triguero *et al.* cited supra).

The time-dependent redistribution of the radiolabeled protein or chimeric molecule of the present invention from the capillary fraction to the parenchyma fraction is consistent with the time dependent migration of a protein or chimeric molecule of the present invention across the blood-brain barrier.

A protein or chimeric molecule of the present invention may promote or inhibit angiogenesis.

The angiogenic potential of the protein or chimeric molecule of the present invention may be assessed methods known in the art. For example, the extent of angiogenesis may be measured by microvessel sprouting in a model of angiogenesis. In this assay, rat fat microvessel fragments (RFMFs) are isolated as described in Shepherd *et al. Arterioscler Thromb Vasc Biol 24:*898-904, 2004. Epididymal fat pads are harvested from euthanized animals, minced and digested in collagenase. RFMFs and single cells are separated from lipids and adipocytes by centrifugation and suspended in 0.1% BSA in PBS. The RFMF suspension is sequentially filtered to remove tissue debris, single cells, and red blood cells from the fragments. RFMFs are suspended in cold, pH-neutralized rat-tail type 1 collagen at 15,000 RFMF/ml and plated into wells (for example, 0.25 ml/well) of 48-well plate for culture. After polymerization of the collagen, an equal volume of DMEM containing 10% FBS is added to each gel. After formation of the gels, vascular extensions characteristic of angiogenic sprouts appear by day 4 of culture. These sprouts are readily distinguished from the parent vessel fragment by the absence of the rough, smooth-muscle associated appearance. The RFMF 3-D cultures can be treated with the protein or chimeric molecule of the present invention and vessel sprout lengths can be measured at day 5 and 6 of culture.

The angiogenic potential of the protein or chimeric molecule of the present invention may also be assessed by an *in vivo* angiogenesis assay described in Guedez *et al. Am J Pathol 162:*1431–1439, 2003. This assay consists of subcutaneous implantation of semiclosed silicone cylinders (angioreactors) into nude mice. Angioreactors are filled with extracellular matrix premixed with or without the protein or chimeric molecule of the present invention. Vascularization within angioreactors is quantified by the intravenous injection of fluorescein isothiocyanate (FITC)-dextran before their recovery, followed by spectrofluorimetry. Angioreactors examined by immunofluorescence is able to show cells and invading angiogenic vessels at different developmental stages.

A protein or chimeric molecule of the present invention may have a distinct immunoreactivity profile determined by immunoassay techniques, which involve the interaction of the molecule with one or more antibodies directed against the molecule. Examples of immunoassay techniques include enzyme-linked immunoabsorbant assays (ELISA), dot blots and immunochromatographic assays such as lateral flow tests or strip tests.

The level of the protein or chimeric molecule thereof may be measured using an immunoassay procedure, for example, a commerically purchased ELISA kit. The protein or chimeric molecule of the present invention may have a different immunoreactivity profile to non-human cell expressed protein or chimeric molecule thereof due to the specificity of the antibodies provided in an immunoassay kit. For instance, the capture and/or detection antibodies of the immunoassay may be antibodies specifically directed against non-human cell expressed human protein or chimeric molecule thereof.

In addition, incorrect folding of the non-human cell expressed human protein or chimeric molecule thereof may result in the exposure of antigenic epitopes which are not exposed on the correctly folded human cell expressed human protein or chimeric molecule thereof. Incorrect folding may arise through, for instance, overproduction of heterologous proteins in the cytoplasm of non-human cells, for example, *E. coli* (Baneyx *Current Opinion in Biotechnology, 10:*411–421, 1999). Further, non-human cell expressed human protein or

chimeric molecule thereof may have a different pattern of post-translational modifications to that of the protein or chimeric molecule of the present invention. For example, the non-human cell expressed human protein or chimeric molecule thereof may exhibit abnormal quantities and/or types of carbohydrate structures, phosphate, sulfate, lipid or other residues. This may result in the exposure of antigenic epitopes which are not exposed on the protein or chimeric molecule of the present invention. Conversely, an altered pattern of post-translational modifications may result in an absence of antigenic epitopes on the protein or chimeric molecule of the present invention which are exposed on the non-human cell expressed human protein or chimeric molecule thereof.

Any one of, or combination of, the above-mentioned factors may lead to inaccurate measurements of:

(a) naturally occurring human protein in laboratory samples or human tissues; or

(b) human cell expressed recombinant human protein or chimeric molecule thereof in laboratory samples, human tissues or in human embryonic stem cell (hES) culture media.

The immunoreactivity profile of a human cell expressed human protein or chimeric molecule thereof, as determined by the use of a suitable immunoassay, may provide an indication of the protein's immunogenicity in the human, as described hereinafter.

Most biologic products elicit a certain level of antibody response against them. The antibody response can, in some cases, lead to potentially serious side effects and/or loss of efficacy. For instance, some patients treated with recombinant protein or chimeric molecule thereof expressed from non-human cells may generate neutralizing antibodies particularly during long-term therapeutic use and thereby reducing the protein's efficacy and or contribute to side effects. The protein or chimeric protein molecule expressed from human cells is unlikely to generate neutralizing antibodies therefore increasing its therapeutic efficacy compared with non-human cell expressed protein or chimeric molecule thereof.

The immunogenicity of protein or chimeric molecule thereof can be assayed using one or more of the following systems.

Most biologic products elicit a certain level of antibody response against them. The antibody response can, in some cases, lead to potentially serious side effects and/or loss of efficacy. For instance, some patients treated with recombinant EPO will generate neutralizing antibodies that also cross-react with the patient's own EPO. In this case, they can develop pure red cell aplasia and be resistant to EPO treatment, resulting in a need for constant dialysis.

Immunogenicity is the property of being able to evoke an immune response within an organism. Immunogenicity depends partly upon the size of the substance in question and partly upon how unlike host molecules it is. A protein or chimeric molecule thereof may have altered immunogenicity due to its novel physiochemical characteristics. For instance, the glycosylation structure of a protein or chimeric molecule thereof may shield or obscure the epitope(s) recognized by the antibody and therefore preventing or reducing antibody binding to the protein or chimeric molecule thereof. Alternatively, some antibodies may recognize a glycopeptide epitope not present in the non-glycosylated version of the protein.

The ability of patient samples to recognize a protein or chimeric molecule thereof with a distinctive physiochemical form can be determined by various immunoassays, as described herein. A properly designed immunoassay involves considerations directing to appropriate detection, quantitation and characterization of antibody responses. A number of recommendations for the design and optimization of immunoassays are outlined in Mire-Sluis *et al. J Immunol Methods 289(1-2)*:1-16, 2004, which is incorporated by reference.

The use of protein or chimeric molecule thereof on therapeutic implants can be assayed using one or more of the following systems.

The present invention extends to the use of a protein or chimeric molecule thereof to manipulate stem cells. A major therapeutic use of stem cells is in regeneration of tissue, cartilage or bone. In one embodiment, the cells are likely to be introduced to the body in a

biocompatible three-dimensional matrix. The implant will consist of a mixture of cells, the scaffold, growth factors and accessory components such as biodegradable polymers, proteoglycans and the like. Incorporation of a protein or chimeric molecule thereof into these matrices during their construction is proposed to regulate the behavior of the cells. Such implants may be used for the formation of bone, the growth of neurons from progenitor cells, chondrocyte implantation for cartilage replacement and other applications. Human cell-derived proteins may reduce the quantity and/or variety of xenogeneic proteins from stem cell culture conditions and thereby reduce the risks of infection by non-human pathogens.

A protein or chimeric molecule of the present invention may interact differently with the matrix used for the formation of the implant, as well as regulating the cells incorporated within the implant. It is anticipated that the combination of a protein or chimeric molecule of the present invention with the implant components will result in one or more of the following pharmacological traits, such as higher proliferation, enhanced differentiation, maintenance in a desired state of differentiation, greater lineage specificity of differentiation, enhanced secretion of matrix components, better 3-dimensional structure formation, enhanced signaling, better structural performance, reduced toxicity, reduced side effects, reduced inflammation, reduced immune cell infiltrate, reduced rejection, longer duration of the implant, longer function of the implant, better stimulation of the cells surrounding the implant, better tissue regeneration, better organ function, or better tissue remodeling.

The effects of protein or chimeric molecule thereof on differential gene expression can be assayed using one or more of the following systems.

The differences in gene expression can be analyzed in cells exposed to a protein or chimeric molecule thereof.

Microarray technology enables the simultaneous determination of the mRNA expression of almost all genes in an organism's genome. This method uses gene "chips" in which oligonucleotides corresponding to the sequences of different genes are attached to a solid

support. Labeled cDNA derived from mRNA isolated from the cell or tissue of interest is incubated with the chips to allow hybridisation between cDNA and the attached complementary sequence. A control is also used, and following hybridisation and washing the signal from both is compared. Specialised software is used to determine which genes are up or down regulated or which have unchanged expression. Many thousands of genes can be analysed on each chip. For example using Affymetrix technology, the Human Genome U133 (HG-U133) Set, consisting of two GeneChip (registered trade mark) arrays, contains almost 45,000 probe sets representing more than 39,000 transcripts derived from approximately 33,000 well-substantiated human genes. The GeneChip (registered trade mark) Mouse Genome 430 2.0 contains over 39,000 transcripts on a single array.

This type of analysis reveals changes in the global mRNA expression pattern and therefore differences in the expression of genes not known to be controlled by a particular stimulus may be uncovered. This technology is hence suitable to analyze the induced gene expression associated with protein or chimeric molecule of the present invention.

The definition of known and novel genes regulated by the particular stimulus will assist in the identification of the biochemical pathways that are important in the biological activity of the particular protein or chimeric molecule of the present invention. This information will be useful in the identification of novel therapeutic targets.

The system could also be used to look at differences in gene expression induced by a protein or chimeric molecule of the present invention as compared to commercially available products.

The effects of protein or chimeric molecule thereof on binding ability can be assayed using one or more of the following systems.

The binding ability of a protein or chimeric molecule of the present invention to various substances, including extracellular matrix, artificial materials, heparin sulfates, carriers or co-factors can be investigated.

The effects of a protein or chimeric molecule thereof on the ability of a particular protein to bind an extracellular matrix can be determined using the following assays.

A surface is coated with extracellular matrix proteins, including but not limited to collagen, vitronectin, fibronectin, laminin, in an appropriate buffer. The unbound sites can be blocked by methods known in the art, for instance, by incubation with BSA solution. The surface is washed, for instance, with PBS solutions, then a solution containing the protein to be tested, for instance a protein or chimeric molecule of the present invention, is added to the surface. After coating, the surface is washed and incubated with an antibody that recognizes a protein or chimeric molecule thereof. Bound antibody is then detected, for instance, by an enzyme-linked secondary antibody that recognizes the primary antibody. The bound antibodies are visualized by incubating with the appropriate substrate and observing a colour change reaction. Glycosylated proteins may adhere more strongly to the extracellular matrix proteins than unglycosylated proteins.

Alternatively, an equivalent amount (specified by ELISA concentration or bioassay activity units) of a protein or chimeric molecule of the present invention, or a counterpart protein or chimeric molecule thereof expressed by non-human cells, are incubated with matrix coated wells, then following washing of the wells the amount bound is determined by ELISA. The amount bound can be indirectly measured by a drop in ELISA reactivity following incubation of the sample with the coated surface.

The ability of protein or chimeric molecule thereof to bind artificial materials can be assayed using one or more of the following systems.

In order to determine the binding ability of a protein or chimeric molecule thereof to artificial materials, a surface is coated with artificial material, including but not limited to metals, scaffolds, in an appropriate buffer. The surface is washed, for instance, with PBS solutions, then a solution containing the protein to be tested, for instance a protein or chimeric molecule of the present invention, is added to the surface. After coating, the surface is washed and incubated with an antibody that recognizes a protein or chimeric molecule thereof. Bound antibody is then detected, for instance, by a enzyme-linked secondary antibody that recognizes the primary antibody. The bound antibodies are

visualized by incubating with the appropriate substrate and observing a color change reaction.

Alternatively, an equivalent amount (specified by ELISA concentration or bioassay activity units) of a protein or chimeric molecule of the present invention, and a counterpart protein or chimeric molecule thereof expressed by non-human cells, are incubated with wells coated by artificial materials, the wells are then washed and the amount bound is determined by ELISA. The amount bound can be indirectly measured by a drop in ELISA reactivity following incubation of the sample with the coated surface.

Ability to bind to artificial surfaces may have biological consequences, for instance, in stent coating. Alternatively, a scaffold coated with a protein or chimeric molecule of the present invention is used to seed cells on. The cell growth and differentiation is then monitored and compared to uncoated or differentially coated scaffolds.

The ability of protein or chimeric molecule thereof to bind to heparin sulfates can be assayed using one or more of the following systems.

A protein or chimeric molecule of the present invention is expected to interact differentially with heparin sulfates due to their physiochemical form. These differences are expected to be evident in experimental models of cell proliferation, differentiation, migration and the like. The combination of a protein or chimeric molecule thereof with heparin sulfates is expected to have distinctive pharmacological traits for a given treatment. This may be an increase in serum half-life, bioavailability, reduced immune-related clearance, greater efficacy, reduced dosage fewer side effects and related advantages.

The ability of protein or chimeric molecule thereof to bind to carriers or co-factors can be assayed using one or more of the following systems.

Proteins or chimeric molecules thereof will be bound to other molecules when they are present in plasma. These molecules may be termed "carriers" or "co-factors" and will influence such factors as bioavailability or serum half life.

Incubating purified versions of the proteins in plasma and analyzing the resulting solution by size exclusion chromatography can determine the interaction of a protein or chimeric molecule of the present invention with their binding partners. If the protein or chimeric molecule thereof binds a co-factor, the resulting complex will have a larger molecular weight, resulting in an altered elution time. The complex can be compared for biological activity, *in vitro* or *in vivo* half-life and bioavailability.

The effects of protein or chimeric molecule thereof on bioassays can be assayed using one or more of the following systems.

Various bioassays can be performed to test the activity of a protein or chimeric molecule of the present invention, including assays on cell proliferation, cell differentiation, cell apoptosis, cell size, cytokine/cytokine receptor adhesion, cell adhesion, cell spreading, cell motility, migration and invasion, chemotaxis, ligand-receptor binding, receptor activation, signal transduction, and alteration of subgroup ratios.

The effects of protein or chimeric molecule thereof on cell proliferation can be assayed using one or more of the following systems.

Cells, in a particular embodiment, exponentially growing cells, are incubated in a growth medium in the presence of a protein or chimeric molecule of the present invention. This can be performed in flasks or 96 well plates. The cells are grown for a period of time and then the number of cells is determined by either a direct (e.g. cell counting) or an indirect (MTT, MTS, tritiated thymidine) method. The increase or decrease in proliferation is determined by comparison with a medium only control assay. Different concentrations of protein or chimeric molecule thereof can be used in parallel series of experiments to get a dose response profile. This can be used to determine the ED50 and ED100 (the dose required to generate the half maximal and maximal response effectively).

The effects of protein or chimeric molecule thereof on cell differentiation or maintenance of cells in an undifferentiated state can be assayed using one or more of the following systems.

Cells are incubated in a growth medium in the presence of a protein or chimeric molecule of the present invention. After a suitable period of time, the cells are assayed for indicators of differentiation. This may be the expression of particular markers on the cell surface, cytoplasmic markers, an alteration in the cell dimensions, shape or cytoplasmic characteristics. The markers may include proteins, sugar structures (e.g. glycosaminocglycans such as heparin sulfates, chondroitin sulfates etc.) lipids (glycosphingolipids or lipid bilayer components). These changes can be assayed by a number of techniques including microscopy, western blot, FACS staining or forward/side scatter profiles.

The effects of protein or chimeric molecule thereof on cell apoptosis can be assayed using one or more of the following systems.

Apoptosis is defined as programmed cell death, and is distinct from other methods of cell death such as necrosis. It is characterized by defined changes in the cells, such as activation of signaling pathways (e.g. Fas, TNFR) resulting in the activation of a subset of proteases know as caspases. Initiator caspase activation leads to the activation of the executioner caspases which cleave a variety of cellular proteins resulting in nuclear fragmentation, cleavage of nuclear lamins, blebbing of the cytoplasm and destruction of the cell. Apoptosis can be induced by protein ligands such as FasL, TNFa and lymphotoxin or by signals such as UV light and substances causing DNA damage.

Cells are incubated in a growth medium in the presence of protein or chimeric molecule thereof and or other agents as suitable for the assay. For instance, the presence of agents able to block transcription (actinomycin D) or translation (cycloheximide) may be required. Following incubation for an appropriate period, the number of cells is determined by a suitable method. A decrease in cell number may indicate apoptosis. Other indications of apoptosis may be obtained by staining of the cells, for instance, for annexins or

observing characteristic laddering patterns of DNA. Further evidence for the confirmation of apoptosis may be achieved by preventing the expression of apoptotic markers by incubating with cell permeable caspases inhibitors (e.g. z-VAD FMK), then assaying for apoptotic markers.

A protein or chimeric molecule of the present invention may prevent apoptosis by providing a survival signal through cellular survival pathways such as the Bcl2 or Akt pathways. Activation of these pathways can be confirmed by western blotting for an increase in cellular Bcl2 expression, or for an increase in the activated (phosphorylated) form of Akt using a phospho-specific antibody directed against Akt.

For this assay, cells are incubated in the presence or absence of the survival factor (e.g. IL-3 and certain immune cells). A proportion of cells incubated in the absence of the survival factor will die by apoptosis upon extended culture, whereas cells incubated in sufficient quantities of survival factor will survive or proliferate. Activation of the cellular pathways responsible for these effects can be determined by western blotting, immunocytochemistry and FACS analysis.

The effects of a protein or chimeric molecule thereof on the inhibition of apoptosis can be assayed using one or more of the following systems.

A protein or chimeric molecule of the present invention is tested for *in vitro* activity to protect rat-, mouse-and human cortical neural cells from cell death under hypoxic conditions and with glucose deprivation. For this, neural cell cultures are prepared from rat embryos. To evaluate the effects of the protein or chimeric molecule of the present invention, the cells are maintained in modular incubator chambers in a water-jacketed incubator for up to 48 hours at 37° C, in serum-free medium with 30 mM glucose and humidified 95% air/5%CO_2 (normoxia) or in serum-free medium without glucose and humidified 95% N_2/5% CO_2 (hypoxia and glucose deprivation), in the absence or presence of the protein or chimeric molecule of the present invention. The cell cultures are exposed to hypoxia and glucose deprivation for less than 24 hour and thereafter returned to normoxic conditions for the remainder of 24 hour. The cytotoxicity is analyzed by the

fluorescence of Alamar blue, which reports cell viability as a function of metabolic activity.

In another method, the neural cell cultures are exposed for 24 hours to 1 mM L-glutamate or a-amino-3-hydroxy-5-methylisoxazole-4-propionic acid (AMPA) under normoxic conditions, in the absence or presence of various concentrations of the protein or chimeric molecule of the present invention. The cytotoxicity is analyzed by the fluorescence of Alamar blue, which reports cell-viability as a function of metabolic activity.

A protein or its chimeric molecule may affect the growth, apoptosis, development, or differentiation of a variety of cells. These changes can be reflected by, among other measurable parameters, changes in the cell size and changes in cytoplasmic complexity, which are due to intracellular organelle development. For instance, keratinocytes induced to differentiate by suspension culture exhibit downregulation of surface markers such as β1 integrins, an increase in cell size and cytoplasmic complexity. The effects of a protein or chimeric molecule thereof on cell size, or cytoplasmic complexity can be assayed using one or more of the following systems.

FACS measures the amount of light scattered off by a cell when a beam of laser is incident on it. An argon laser providing light with a wavelength of 488nm is frequently used. The larger the size of the cell, the greater the disruption of the beam of light in the forward direction, hence the level of forward scatter corresponds to the size of the cell. In order to measure changes in cell size, cells treated with a protein or chimeric molecule of the present invention are diluted in sheath fluid and injected into the flow cytometer (FACSVantage SE, Becton Dickinson). Untreated cells act as a control. The cells pass through a beam of light and the amount of forward scattering of the light corresponds to the size of the cells.

Changes in intracellular organelle growth and development (cytoplasmic complexity) can also be measured by FACS. The intracellular organelles of the cell scatter light sideways. Hence, change in cytoplasmic complexity can be measured by the amount of side scattering of light by the cells by the above method, and the level of complexity of

intracellular organelles and the level of granularity of the cell can be estimated by measuring the level of side scatter of light given off by the cells.

The effect of a protein or chimeric molecule thereof on cell size or cytoplasmic complexity can be assessed by using FACS to compare the profiles given off by, for instance, 20,000 treated cells with the signals emitted by identical number of untreated cells. By comparing the signals from the different treated populations of cells, the relative changes in cell size and cytoplasmic complexity can be determined.

The effects of a protein or chimeric molecule thereof on cell growth, apoptosis, development, or differentiation can be assayed using one or more of the following systems.

Protein-induced apoptosis and changes in cell growth or cycles can be assessed by labeling the DNA of treated cells with dyes such as propidium iodine which has an excitation wavelength in the range of 488 nm and emission at 620 nm. Cells undergoing apoptosis has condensed DNA as well as different size and granularity. These factors give specific forward and size scatter profiles as well as fluorescence signal, and hence the population of cells undergoing apoptosis can be differentiated from normal cells. The amount of DNA in a cell also reflects which state of the cell cycle the cell is in. For instance, a cell in G_2 stage will have twice the amount of DNA as a cell in G_0 state. This will be reflected by a doubling of the fluorescence signal given off by a cell in G_2 phase. The effect of a protein or chimeric molecule thereof can be assessed by using FACS to compare the fluorescence signals given off by for instance, 20,000 treated cells with the signals emitted by identical number of untreated cells.

The protein or its chimeric molecule of the present invention may also alter the expression of various proteins. The effects of the protein or chimeric molecule thereof on protein expression by cells can be assayed using one or more of the following systems.

To assess the increase and decrease in expression of a protein in an entire cell, the cells can be fixed and permeabilised, then incubated with fluorescence conjugated antibody targeting the epitope of the protein of interest. A large variety of fluorescent labels can be

used with an Argon laser system. Fluorescent molecules such as FITC, Alexa Fluor 488, Cyanine 2, Cyanine 3 are commonly used for this experiment. This method can also be used to estimate the changes in expression of surface markers and proteins by labeling non-permeabilised cells where only the epitope exposed on the cell surface can be labeled with antibodies. The effect of a protein or chimeric molecule thereof can be assessed by using FACS to compare the fluorescence signals given off by, for instance, 20,000 treated cells with the signals emitted by identical number of untreated cells.

The effects of a protein or its chimeric molecule on ligand/receptor adhesion can be assayed using one or more of the following systems.

A protein or chimeric molecule of the present may be more or less adhesive to substrates compared to those of a previously known physiochemical form. The interaction may be with protein receptors for sugar structures (e.g. selectins, such as L-selectin and P-selectin), with extracellular matrix components such as fibronectin, collagens, vitronectins, and laminins, or with non-protein components such as sugar molecules (heparin sulfates, other glycosaminoglycans).

A protein or chimeric molecule thereof may also interact differently with non-biological origin materials such as tissue culture plastics, medical device components (e.g. stents or other implants) or dental materials. In the case of medical devices this may alter the engraftment rates, the interaction of the implant with particular classes of cell type or the type of linkage formed with the body.

Any suitable assays for protein adhesion can be employed. For instance, a solution containing a protein or chimeric molecule of the present invention is incubated with a binding partner, in a particular embodiment, on an immobilised surface. Following incubation, the amount of the protein or the chimeric molecule present in the solution is assayed by ELISA and the difference between the amount remaining and the starting material is what has bound to the binding partner. For instance, the interaction between the protein or the chimeric molecule and an extracellular matrix protein could be determined by first coating wells of a 96 well plate with the ECM protein (e.g. fibronectin). Non-specific binding is then blocked by incubation with a BSA solution. Following washing, a

known concentration of a protein or its chimeric molecule solution is added for a defined period. The solution is then removed and assayed for the amount of protein or its chimeric molecule remaining in solution. The amount bound to the ECM protein can be determined by incubating the wells with an antibody to a protein or its chimeric molecule, then detecting with an appropriate system (either a labeled secondary antibody or by biotin-avidin enzyme complexes such as those used for ELISA).

Methods for determining the amount bound to other surfaces may involve hydrolyzing a protein or its chimeric molecule from the inert implant surface, then measuring the amino acids present in the solution.

The effects of a protein or a chimeric molecule thereof on cell adhesion can be assayed using one or more of the following systems.

Cell adhesion to matrix (e.g. extracellular matrix components such as fibronectin, vitronectin, collagen, laminin etc.) is mediated at least in part by the integrin molecules. Integrin molecules consist of alpha and beta subunits, and the particular combinations of alpha and beta subunit give rise to the binding specificity to a particular ligand (e.g. a2b1 integrin binds collagen, a5b1 binds fibronectin etc). The integrins subunits have large extracellular domains responsible for binding ligand, and shorter cytoplasmic domains responsible for interaction with the cytoskeleton. In the presence of ligand, the cytoplasmic domains are responsible for the induction of signal transduction events ("outside in signaling"). The affinity of integrins for their ligands can be modulated by extracellular signaling events that in turn lead to changes in the cytoplasmic tails of the integrins ("inside out signaling").

Incubation with a protein or chimeric molecule of the present invention can potentially alter cell adhesion in a number of ways. First, it can alter qualitatively the expression of particular integrin subsets, leading to changes in binding ability. Secondly, the amount of a particular integrin expressed may alter, leading to altered cell binding to its target matrix. Thirdly, the affinity of a particular integrin may be altered without changing its surface expression (inside-out signaling). All these changes may alter the binding of cells to either a spectrum of ligands, or alter the binding to a particular ligand.

A protein or chimeric molecule of the present invention can be tested in Cell-ECM adhesion assays which are generally performed in 96 well plate. Wells are coated with matrix, then unbound sites within the wells are blocked with BSA. A defined number of cells are incubated with the coated wells, then unbound cells are washed away and the bound cells incubated in the presence or absence of the protein or the chimeric molecule thereof. The number of cells is determined by an indirect method such as MTT/MTS. Alternatively, the cells are labeled with a radioactive label (e.g. ^{51}Cr) and a known amount of radioactivity (i.e. cells) is added to each well. The amount of bound radioactivity is determined and calculated as a percentage of the amount loaded.

Cells also adhere to other cells, for instance, adhesion of one population of cells to a monolayer of another type of cells. To assay for this, the suspension cells added to the monolayer cells would be labeled with radioactivity. The cells are then incubated in the presence or absence of a protein or chimeric molecule thereof. The unbound cells would be washed away and the remaining mixed population of cells can be lysed and assayed for the amount of radioactivity present.

The effects of a protein or chimeric molecule thereof on cell spreading can be assayed using one or more of the following systems.

A protein or chimeric molecule of the present invention may have altered effects on cell spreading. Initiation of cell spreading is a key step in cell motility and invasive behavior. Cells spreading can be initiated *in vitro* in a number of ways. Plating a suspension of cells onto ECM components will result in attachment and ligand binding by integrin receptors. This initiates signal transduction events resulting in the activation of a family of the Cdc42, Rac and Rho small GTPases. Activation of these proteins results in actin polymerization and an extension of a lamellipodium, resulting in gradual flattening of the cells and contact of more integrins with their receptors. Eventually the cells have flattened totally and formed focal adhesions (large structures containing integrins and signaling proteins). Cell spreading can also be initiated by stimulation of adherent cells with growth factors, again resulting in activation of the Cdc42/Rac/Rho proteins and lamellipodium formation.

Cell spreading can be quantitated by examining a large number of cells at different time points following stimulation with a protein or chimeric molecule thereof. The area of each cell can be determined using image analysis programs and the percentage of cells spread as well as the degree of cell spreading can be compared with time. More rapid spreading may be initiated by a higher activation of the Cdc42/Rac/Rho pathways, alternatively, temporal, qualitative and quantitative differences in their activation may be observed with a protein or chimeric molecule of the present invention. This in turn may reflect differences in the signaling events induced by the protein or chimeric molecule of the present invention.

The effects of a protein or a chimeric molecule thereof on cell motility, migration and invasion can be assayed using one or more of the following systems.

Cells adherent to a tissue culture dish do not remain statically anchored to one spot, but rather constantly extend and retract portions of their cell body. When viewed under time-lapse photography, the cells can be observed to move around the dish, either as isolated single cells or as a cell colony. This motion may be either "random walk" (i.e. not directed in a particular direction), or directional. Both types of motion can be increased by the addition of growth factors. Time-lapse photography can be used to quantitate the overall distance covered by the cells in a given time period, as well as the overall directionality.

In the case of directional migration, cells will move towards a source of chemoattractant by sensing the chemical gradient and orienting their migration machinery towards it. In many instances, the chemoattractant is a growth factor. Directional migration can be quantitated by providing a source of chemoattractant (e.g. *via* a thin pipette) then imaging the cells migrating towards it with time-lapse photography.

An alternative system for determining directed migration is the Boyden chamber assay. In this assay, cells are placed in an upper chamber that is connected to a lower chamber *via* small holes in the partitioning membrane. Growth medium is put in both chambers, but chemoattractant is added only to the lower chamber, resulting in a diffusion gradient between the two chambers. The cells are attracted to the growth factor source and migrate through the holes in the separation membrane and on to the lower side of the membrane.

After a number of hours, the membrane is removed and the number of cells that has migrated onto the bottom of the membrane is determined.

The process of cellular invasion utilises many of the same components as migration. Cell invasion can be modeled using layers of extracellular matrix through which the cells invade. For instance, Matrigel is a mixture of basement membrane components (ECM components, growth factors etc.) that is liquid at 4 degrees but rapidly sets at 37 degrees to form a gel. This can be used to coat the upper surface of a Boyden chamber, and the chemoattractant added to the lower layer. For cells to pass onto the lower surface of the membrane, they must degrade the matrigel using enzymes such as collagenases and matrix metalloproteinases (MMPs) as well as migrating directionally towards the chemoattractant. This assay mimics the various processes required for cellular invasion.

The effects of a protein or a chimeric molecule thereof on chemotaxis can be assayed using one or more of the following systems.

The migration of cells toward the chemoattractant can be measured *in vitro* in a Boyden chamber. A protein or chimeric molecule of the present in invention is placed in the lower chamber and an appropriate target cell population is placed in the upper chamber. To mimic the *in vitro* process of immune cells migrating from the blood to sites of inflammation, migration through a layer of cells may be measured. Coating the upper surface of the well of the Boyden chamber with a confluent sheet of cells, for instance, epithelial, endothelial or fibroblastic cells, will prevent direct migration of immune cells through the holes in the well. Instead, the cells will need to adhere to the monolayer and migrate through it towards the protein to be tested. The presence of cells on the under surface of the Boyden chamber or in the medium in the lower well in only those wells treated with the protein or chimeric molecule thereof is indicative of the chemotactic ability of the protein or the chimeric molecule. To show that the effect is specific to a protein or chimeric molecule thereof, a neutralising antibody can be incubated with the protein in the lower chamber.

Alternatively, to test the ability of a substance (chemical, protein, sugar) to prevent chemotaxis, the substance is included in the lower chamber of the Boyden chamber along

with a solution containing known chemotactic ability (this may be a specific chemokine, conditioned medium from a cell source or cells secreting a range of chemokines). A susceptible target cell population is then added to the upper chamber and the assay performed as described above.

The effects of a protein or chimeric molecule thereof on ligand-receptor binding can be assayed using one or more of the following systems.

A protein or chimeric molecule of the present invention may have different ligand-receptor binding abilities. Ligand-receptor binding can be measured by various parameters, for instance, the dissociation constant (Kd), dissociation rate constant (off rate) (k^-), association rate constant (on rate) (k^+). Differences in ligand-receptor binding may correlate with different timing and activation of signaling, leading to different biological outcomes.

Ligand-receptor binding can be measured and analysed by either Scatchard plot or by other means such as Biacore.

For Scatchard analysis, a protein or its chimeric molecule, labeled with, for instance, radioactively labeled (eg, ^{125}I), is incubated in the presence of differing amounts of cold competitor of a protein or its chimeric molecule, with cells, or extracts thereof, expressing the corresponding ligand or receptor. The amount of specifically bound labeled protein or its chimeric molecule is determined and the binding parameters calculated.

For the Biacore, the corresponding recombinant ligand or receptor of the protein or its chimeric molecule is coupled to the detection unit. Solutions containing a protein or chimeric molecule thereof of choice are then passed over the detection cell and binding is determined by a change in the properties of the detection unit. On rates can be determined by passing solutions containing the protein or the chimeric molecule over the detection cell until a fixed reading is recorded (when the available sites are all occupied). A solution not containing the protein or the chimeric molecule is then passed over the cell and the protein dissociates from the corresponding ligand or receptor, giving the off rate.

The effects of a protein or chimeric molecule thereof on receptor activation can be assayed using one or more of the following systems.

Interaction with a protein or a chimeric molecule thereof and its corresponding ligand or receptor may be paralleled by differences in the signaling events induced from the cell's endogenous protein. The timing of interaction may be characteristic of a protein or chimeric molecule thereof as definitely on/off rates or dissociation constants.

Activated receptors are often internalized by the cells. The receptor/ligand complex can then be dissociated (e.g., be lowering the pH within cellular vesicles, resulting in detachment of the ligand) and the receptor recycled to the cell surface. Alternatively, the complex may be targeted for destruction. In this case the receptors are effectively down-regulated and unable to generate more signal, whereas when they are recycled they are able to repeat the signaling process. Differential receptor binding or activation may result in the receptor being switched from a destruction to a recycling pathway, resulting in a stronger biological response.

The effects of a protein or a chimeric molecule thereof on signal transduction can be assayed using one or more of the following systems.

Binding of ligands or receptors to the protein or its chimeric molecule thereof may initiate signaling, which may include reverse signaling, through a variety of cytoplasmic proteins. Reverse signaling occurs when a membrane-bound form of a ligand transduces a signal following binding by a soluble or membrane bound version of its receptor. Reverse signaling can also occur after binding of the membrane bound ligand by an antibody. These signaling events (including reverse signaling events) lead to changes in gene and protein expression. Hence, a protein or chimeric molecule of the present invention can induce or inhibit different signal transductions in various pathways or other signal transduction events, such as the activation of JAK/STAT pathway, Ras-erk pathway, AKT pathway, the activation of PKC, PKA, Src, Fas, TNFR, NFkB, p38MAPK, c-Fos, recruitment of proteins to receptors, receptor phosphorylation, receptor internalization, receptor cross-talk or secretion.

The ligands or receptors recruited to the protein or chimeric molecule thereof may be unique to the protein or chimeric molecule of the present invention, due to different conformations of the ligand or receptors being induced. One way of assaying for these differences is to immunoprecipitate the ligand or receptor using an antibody crosslinked to sepahrose beads. Following immunoprecipitation and washing, the proteins are loaded on a 2D gel and the comparative spot patterns are analysed. Different spots can be cut out and identified by mass spectrometry.

The effects of a protein or chimeric molecule thereof on up regulation and down regulation of surface markers can be assayed using one or more of the following systems.

Cells may have a variety of responses to the protein or chimeric molecule of the present invention. There are a range of proteins on cell surfaces responsible for communication between the cells and the extracellular environment. Through regulated processes of endocytosis and exocytosis, various proteins are transported to and from the cell surface. Typical proteins found on the cells surface includes receptors, binding proteins, regulatory proteins and signaling molecules. Changes in expression and degradation rate of the proteins also changes the level of the proteins on the cell surface. Some proteins are also stored in intracellular reservoirs where specific signals can induce trafficking of proteins between this storage and the cellular membrane.

Cells are incubated for an appropriate amount of time in medium containing a protein or chimeric molecule of the present invention and their responses can be compared with cells exposed to the same medium without the protein or chimeric molecule of the present invention. The proteins on the cell membrane can be solubilised and separated from the cells by centrifugation. The level of expression of a specific protein can be measured by Western blotting. Cells can also be labeled with fluorescence conjugated antibodies, and visualized under confocal microscopy system or counted by fluorescence activated cell sorting (FACS). This will detect any changes in expression and distribution of proteins on the cells. By using multiple antibodies, changes in protein interaction can also be studied by confocal microscopy and immuno-precipitation. Similarly, these experiments can be extended to *in vivo* animal models. Cells from specific part of animals treated with the

protein or chimeric molecule of the present invention may be extracted and examined with identical methodologies.

Cells induced to differentiate *in vitro* or *in vivo* by the addition of the protein or chimeric molecule of the present invention will express differentiation markers that distinguish them from the untreated cells. Some cells, for instance, progenitor or stem cells, can differentiate into many subpopulations, distinguishable by their surface markers. A protein or chimeric molecule of the present invention may stimulate the progenitor cells to differentiate into subgroups in a particular ratio.

The protein of the present invention and its chimeric molecule may have effects upon cell repulsion.

The effects of the protein or its chimeric molecule on the modulation of the growth and guidance of cells and neurons is a convenient assay for cell repulsion.

Disrupting the interactions between subunits and other components of a protein leads to a way to inhibit the biological effects of the protein or its chimeric molecule. Compounds inhibiting such biological effects are identified by a number of ways.

High throughput screening programs use a library of small chemical entities (chemicals or peptides) to generate lead compounds for clinical development. A number of assays can be used to screen a library compounds for their ability to affect a biologically relevant endpoint. Each potential compound in a library is tested with a particular assay in a single well, and the ability of the compound to affect the assay determined. Some examples of the assays are provided below:

For this assay, cells are plated into a microtitre plate (96 plate, 384 plate or the like). The cells will have a readout mechanism for activation of a protein or chimeric molecule thereof. This may involve assaying for cell growth, assaying for stimulation of a particular pathway (e.g., FRET based techniques), assaying for induction of a reporter gene (e.g., CAT, beta-galactosidase, fluorescent proteins), assaying for apoptosis and assaying for differentiation. Cells are then exposed to the protein or chimeric molecule of the present

invention in the presence or absence of a particular small molecule. The drug can be added before, after or during the addition of the protein or chimeric molecule thereof. After an appropriate period of time, the individual wells are read using an appropriate method (eg, Fluorescence for FRET or induction of fluorescent proteins, cell number by MTT, beta-galactosidase activity etc). Control wells without addition of any drug or cytokine serve as comparisons. Any molecule able to inhibit the receptor/cytokine complex will give a different readout to the control wells. Further experiments will be required to show specificity of the inhibition. Alternatively, the drug could affect the detection method by a non-cytokine, non-receptor mechanism (a false positive).

A ligand or receptor of the protein or chimeric molecule thereof is immobilised on a solid surface. A protein or its chimeric molecule and the compound to be tested are then added. This can be performed by adding a protein or its chimeric molecule first, then the compound; the compound first, then a protein or its chimeric molecule; or the compound and the protein or its chimeric molecule can be added together. Bound protein or the chimeric molecule is then detected by an appropriate detection antibody. The detection antibody can be labeled with an enzyme (e.g., alkaline phosphatase or Horse-radish peroxidase for colorimetric detection) or a fluorescent tag for fluorescence detection. Alternatively, a protein or its chimeric molecule can be labeled (e.g., Biotin, radioactive labeling) and be detected with an appropriate technique (e.g., for Biotin labeling, streptavidin linked to a colorimetric detection system, for radiolabeling the complex is solubilised and counted). Inhibition of protein binding is measured by a drop in the reading compared to the control wells.

Soluble ligands or receptors of the protein or chimeric molecules thereof are bound to beads. This binding reaction can be either an adsorption process or involve chemically linking them to the plate. The beads are incubated with the protein or the chimeric molecules and a candidate compound in an appropriate well. This can be performed as the protein or the chimeric molecules first, then compound; compound first then the protein or the chimeric molecules; or compound and the protein or the chimeric molecules together. A fluorescently labeled detection antibody that recognizes a protein or chimeric molecule thereof is then added. The unbound antibody is removed and the beads are passed through

a FACS. The amount of fluorescence detected will decrease if a compound inhibits the interaction of a protein or chimeric molecule thereof with its receptor.

To enable screening of multiple interactions between protein and its corresponding ligand/receptor against one inhibitory compound, the ability of the FACS machine to analyse scatter profiles is used. A bead with a larger diameter will have a different scatter profile to that of a smaller bead, and this can be separated out for analysis ("gating").

A number of different proteins, one of which is the protein or chimeric molecule of the present invention, are each linked to beads of a particular diameter. A mixture of ligands/receptors to the above-mentioned proteins are then added to the bead mixture in the presence of one candidate compound. The bound ligands/receptors are then detected using a specific secondary antibodies that is fluorescently labeled. The antibodies can be all labeled with the same detection fluorophore. The ability of the compound to prevent binding of a protein to its ligand/receptor is then determined by running the sample though a FACS machine and gating for each known bead size. The individual binding results are then analysed separately. The major benefit of this method of analysis is that the screening each compound can be tested in parallel with a number of proteins to decrease the time taken for screening proportionally.

A protein or chimeric molecule thereof may also be characterised by its crystal structure. The physiochemical form of a protein or its chimeric molecule may provide a unique 3D crystal structure. In addition, the crystal structure of the protein-ligand/receptor complex may also be generated using a protein or chimeric molecule of the present invention. Since the present invention provides a protein or a chimeric molecule thereof which is substantially similar to a human naturally occurring form, the complex is likely to be a more reflective representation of the *in vivo* structure of the naturally occurring protein-ligand/receptor complex. Once a crystal structure has been obtained, interactions between a protein or its chimeric molecule and potential compounds inhibiting such interactions can be identified.

Once potential compounds are identified by high throughput screening or from the crystal structure of the protein-ligand/receptor complex, a process of rational drug design can begin.

There are several steps commonly taken in the design of a mimetic from a compound having a given desired property. First, the particular parts of the compound that are critical and/or important in determining the desired property are determined. In the case of a peptide, this can be done by systematically varying the amino acid residues in the peptide, e.g. by substituting each residue in turn. Alanine scans of peptides are commonly used to refine such peptide motifs. These parts or residues constituting the active region of the compound are known as its "pharmacophore".

Once the pharmacophore has been found, its structure is modeled according to its physical properties, e.g. stereochemistry, bonding, size and/or charge, using data from a range of sources, e.g. spectroscopic techniques, x-ray diffraction data and NMR. Computational analysis, similarity mapping (which models the charge and/or volume of a pharmacophore, rather than the bonding between atoms) and other techniques can be used in this modeling process.

In a variant of this approach, the three-dimensional structure of the ligand and its binding partner are modeled. This can be especially useful where the ligand and/or binding partner change conformation on binding, allowing the model to take account of this in the design of the mimetic. Modeling can be used to generate inhibitors which interact with the linear sequence or a three-dimensional configuration.

A template molecule is then selected onto which chemical groups which mimic the pharmacophore can be grafted. The template molecule and the chemical groups grafted onto it can conveniently be selected so that the mimetic is easy to synthesize, is likely to be pharmacologically acceptable, and does not degrade *in vivo*, while retaining the biological activity of the lead compound. Alternatively, where the mimetic is peptide-based, further stability can be achieved by cyclizing the peptide, increasing its rigidity. The mimetic or mimetics found by this approach can then be screened to see whether they have the target

property, or to what extent they exhibit it. Further optimization or modification can then be carried out to arrive at one or more final mimetics for *in vivo* or clinical testing.

The goal of rational drug design is to produce structural analogs of biologically active polypeptides of interest or of small molecules with which they interact (e.g. agonists, antagonists, inhibitors or enhancers) in order to fashion drugs which are, for example, more active or stable forms of the polypeptide, or which, e.g. enhance or interfere with the function of a polypeptide *in vivo*. See, e.g. Hodgson (*Bio/Technology 9:*19-21, 1991). In one approach, one first determines the three-dimensional structure of a protein of interest by x-ray crystallography, by computer modeling or most typically, by a combination of approaches. Useful information regarding the structure of a polypeptide may also be gained by modeling based on the structure of homologous proteins. An example of rational drug design is the development of HIV protease inhibitors (Erickson *et al. Science 249:*527-533, 1990). In addition, target molecules may be analyzed by an alanine scan (Wells, *Methods Enzymol 202:*2699-2705, 1991). In this technique, an amino acid residue is replaced by Ala and its effect on the peptide's activity is determined. Each of the amino acid residues of the peptide is analyzed in this manner to determine the important regions of the peptide.

It is also possible to isolate a target-specific antibody, selected by a functional assay and then to solve its crystal structure. In principle, this approach yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced banks of peptides. Selected peptides would then act as the pharmacore.

In one aspect, the protein or chimeric molecule of the present invention is used as an immunogen to generate antibodies. The physiochemical form of a protein or chimeric molecule of the present invention may raise antibodies to the protein or the chimeric molecule ; glycopeptides specific to the protein or chimeric molecule of the present

invention; or antibodies directed to another co- or post-translationally modified peptide within the protein or chimeric molecule thereof.

The protein of the present invention or its chimeric molecule may present epitopes not normally accessible (but possibly present) *in vivo*. For instance, there may be regions within a receptor domain that are normally in contact with another component of a heteromeric receptor. These epitopes may be used to generate monoclonal antibodies that cross react with the endogenous receptor. Such antibodies may block interaction of one receptor component with another and therefore prevent signal transduction. This may be therapeutically useful in the case of overexpression of a cytokine or receptor. The antibodies may also be therapeutically useful in diseases where the receptor is overexpressed and signals without needing the ligand.

The antibodies are also useful to detect the levels of the protein or chimeric molecule thereof during the treatment of the disease (e.g., serum levels for half-life determination).

In addition, the antibodies are useful as diagnostic for determining the presence of a protein or chimeric molecule of the present invention in a particular sample.

Reference to an "antibody" or "antibodies" includes reference to all the various forms of antibodies, including but not limited to: full antibodies (e.g. having an intact Fc region), including, for example, monoclonal antibodies; antigen-binding antibody fragments, including, for example, Fv, Fab, Fab' and $F(ab')_2$ fragments; humanized antibodies; human antibodies (e.g., produced in transgenic animals or through phage display); and immunoglobulin-derived polypeptides produced through genetic engineering techniques. Unless otherwise specified, the terms "antibody" or "antibodies" and as used herein encompasses both full antibodies and antigen-binding fragments thereof.

Unless stated otherwise, specificity in respect of an antibody of the present invention is intended to mean that the antibody binds substantially only to its target antigen with no appreciable binding to unrelated proteins. However, it is possible that an antibody will be designed or selected to bind to two or more related proteins. A related protein includes different splice variants or fragments of the same protein or homologous proteins from

different species. Such antibodies are still considered to have specificity for those proteins and are encompassed by the present invention. The term "substantially" means in this context that there is no detectable binding to a non-target antigen above basal, i.e. non-specific, levels.

The antibodies of the present invention may be prepared by well-known procedures. See, for example, Monoclonal Antibodies, *Hybridomas: A New Dimension in Biological Analyses*, Kennet *et al.* (eds.), Plenum Press, New York (1980); and *Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1988).

One method for producing an antibody comprises immunizing a non-human animal, such as a mouse or a transgenic mouse, with a protein or chimeric molecule of the present invention, or immunogenic parts thereof, such as, for example, a peptide containing the receptor binding domain, whereby antibodies directed against the polypeptide of a protein or its chimeric molecule, or immunogenic parts thereof, are generated in the animal. Various means of increasing the antigenicity of a particular protein or its chimeric molecule, such as administering adjuvants or conjugated antigens, comprising the antigen against which an antibody response is desired and another component, are well known to those in the art and may be utilized. Immunizations typically involve an initial immunization followed by a series of booster immunizations. Animals may be bled and the serum assayed for antibody titer. Animals may be boosted until the titer plateaus. Conjugates may be made in recombinant cell culture as protein fusions. Also, aggregating agents such as alum are suitably used to enhance the immune response.

Both polyclonal and monoclonal antibodies can be produced by this method. The methods for obtaining both types of antibodies are well known in the art. Polyclonal antibodies are less favored but are relatively easily prepared by injection of a suitable animal with an effective amount of a protein or chimeric molecule of the present invention, or immunogenic parts thereof, collecting serum from the animal and isolating specific antibodies to a protein or chimeric molecule thereof by any of the known

immunoadsorbent techniques. Antibodies produced by this technique are generally less favoured, because of the potential for heterogeneity of the product.

The use of monoclonal antibodies is particularly favored because of the ability to produce them in large quantities and the homogeneity of the product. Monoclonal antibodies may be produced by conventional procedures.

The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to polyclonal antibody preparations which typically include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. The modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies to be used in accordance with the present invention may be made by the hybridoma method first described by Kohler *et al. Nature 256*:495 (1975), or may be made by recombinant DNA methods (see, e.g., U.S. Pat. No. 4,816,567). The "monoclonal antibodies" may also be isolated from phage antibody libraries using for example, the techniques described in Clackson *et al. Nature 352*:624-628, 1991 and Marks *et al. J Mol Biol 222*:581-597, 1991.

The present invention contemplates a method for producing a hybridoma cell line which comprises immunizing a non-human animal, such as a mouse or a transgenic mouse, with a protein or chimeric molecule of the present invention; harvesting spleen cells from the immunized animal; fusing the harvested spleen cells to a myeloma cell line to generate hybridoma cells; and identifying a hybridoma cell line that produces a monoclonal antibody that binds a protein or chimeric molecule thereof.

Such hybridoma cell lines and the monoclonal antibodies produced by them are encompassed by the present invention. Monoclonal antibodies secreted by the hybridoma cell lines are purified by conventional techniques. Hybridomas or the monoclonal antibodies produced by them may be screened further to identify monoclonal antibodies with particularly desirable properties, such as the ability to inhibit cytokine-signaling through its receptor.

A protein or chimeric molecule thereof or immunogenic part thereof that may be used to immunize animals in the initial stages of the production of the antibodies of the present invention should be from a human-expressed source.

Antigen-binding fragments of antibodies of the present invention may be produced by conventional techniques. Examples of such fragments include, but are not limited to, Fab, Fab', $F(ab')_2$ and Fv fragments, including single chain Fv fragments (termed sFv or scFv). Antibody fragments and derivatives produced by genetic engineering techniques, such as disulfide stabilized Fv fragments (dsFv), single chain variable region domain (Abs) molecules, minibodies and diabodies are also contemplated for use in accordance with the present invention.

Such fragments and derivatives of monoclonal antibodies directed against a protein or chimeric molecule thereof may be prepared and screened for desired properties, by known techniques, including the assays herein described. The assays provide the means to identify fragments and derivatives of the antibodies of the present invention that bind to a protein or chimeric molecule thereof, as well as identify those fragments and derivatives that also retain the activity of inhibiting signaling by a protein or chimeric molecule thereof. Certain of the techniques involve isolating DNA encoding a polypeptide chain (or a portion thereof) of a mAb of interest, and manipulating the DNA through recombinant DNA technology. The DNA may be fused to another DNA of interest, or altered (e.g. by mutagenesis or other conventional techniques) to add, delete, or substitute one or more amino acid residues.

DNA encoding antibody polypeptides (e.g. heavy or light chain, variable region only or full length) may be isolated from B-cells of mice that have been immunized with a protein or chimeric molecule of the present invention. The DNA may be isolated using conventional procedures. Phage display is another example of a known technique whereby derivatives of antibodies may be prepared. In one approach, polypeptides that are components of an antibody of interest are expressed in any suitable recombinant expression system, and the expressed polypeptides are allowed to assemble to form antibody molecules.

Single chain antibodies may be formed by linking heavy and light chain variable region (Fv region) fragments *via* an amino acid bridge (short peptide linker), resulting in a single polypeptide chain. Such single-chain Fvs (scFvs) have been prepared by fusing DNA encoding a peptide linker between DNAs encoding the two variable region polypeptides (VL and VH). The resulting antibody fragments can form dimers or trimers, depending on the length of a flexible linker between the two variable domains (Kortt *et al. Protein Engineering 10:*423, 1997). Techniques developed for the production of single chain antibodies include those described in U.S. Patent No. 4,946,778; Bird (*Science 242:*423, 1988), Huston *et al.* (*Proc Natl Acad Sci USA 85:*5879, 1988) and Ward *et al.* (*Nature 334:*544, 1989). Single chain antibodies derived from antibodies provided herein are encompassed by the present invention.

In one embodiment, the present invention provides antibody fragments or chimeric, recombinant or synthetic forms of the antibodies that bind to the protein or chimeric molecule of the present invention and inhibit signaling by the protein or its chimeric molecule.

Techniques are known for deriving an antibody of a different subclass or isotype from an antibody of interest, i.e., subclass switching. Thus, IgG1 or IgG4 monoclonal antibodies may be derived from an IgM monoclonal antibody, for example, and vice versa. Such techniques allow the preparation of new antibodies that possess the antigen-binding properties of a given antibody (the parent antibody), but also exhibit biological properties associated with an antibody isotype or subclass different from that of the parent antibody.

Recombinant DNA techniques may be employed. Cloned DNA encoding particular antibody polypeptides may be employed in such procedures, e.g. DNA encoding the constant region of an antibody of the desired isotype.

The monoclonal production process described above may be used in animals, for example mice, to produce monoclonal antibodies. Conventional antibodies derived from such animals, for example murine antibodies, are known to be generally unsuitable for administration to humans as they may cause an immune response. Therefore, such antibodies may need to be modified in order to provide antibodies suitable for administration to humans. Processes for preparing chimeric and/or humanized antibodies are well known in the art and are described in further detail below.

The monoclonal antibodies herein specifically include "chimeric" antibodies in which the variable domain of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a non-human species (e.g., murine), while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from humans, as well as fragments of such antibodies, so long as they exhibit the desired biological activity (U.S. Pat. No. 4,816,567; and Morrison *et al. Proc Natl Acad Sci USA 81*:6851-6855, 1984).

"Humanized" forms of non-human (e.g., murine) antibodies are chimeric antibodies which contain minimal sequence derived from the non-human immunoglobulin. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which the complementarity determining regions (CDRs) of the recipient are replaced by the corresponding CDRs from a non-human species (donor antibody) such as mouse, rat, rabbit or nonhuman primate having the desired properties, for example specificity, and affinity. In some instances, framework region residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized antibodies may comprise residues which are not found in the recipient antibody or in the donor antibody. These modifications are made to further refine antibody performance. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the complementarity determining

regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework region residues are those of a human immunoglobulin sequence. The humanized antibody optionally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. For further details, see Jones *et al. Nature 321*:522-525, 1986; Reichmann *et al. Nature 332*:323-329, 1988; Presta, *Curr Op Struct Biol 2*:593-596, 1992; Liu *et al. Proc Natl Acad Sci USA 84*:3439, 1987; Larrick *et al. Bio/Technology 7*:934, 1989; and Winter and Harris, *TIPS 14*:139, 1993.

In a further embodiment, the present invention provides an immunoassay kit with the ability to assay the level of human protein expressed from human cells present in a biological preparation, including a biological preparation comprising the naturally occurring human protein.

A biological preparation which can be assayed using the immunoassay kit of the present invention includes but is not limited to laboratory samples, cells, tissues, blood, serum, plasma, urine, stool, saliva and sputum.

The immunoassay kit of the present invention comprises a solid phase support matrix, not limited to but including a membrane, dipstick, bead, gel, tube or a multi-well, flat-bottomed, round-bottomed or v-bottomed microplate, for example, a 96-well microplate; a preparation of antibody directed against the human protein of interest (the capture antibody); a preparation of blocking solution (for example, BSA or casein); a preparation of secondary antibody (the detection antibody), also directed against the human protein of interest and conjugated to a suitable detection molecule (for example, alkaline phosphatase); a solution of chromagenic substrate (for example, nitro blue tetrazolium); a solution of additional substrate (for example, 5-bromo-4-chloro-3-indolyl phosphate); a stock solution of substrate buffer (for example, 0.1M Tris-HCL (pH 7.5) and 0.1M NaCl, 50mM $MgCl_2$); a preparation of the protein or chimeric molecule of the present invention with known concentration (the standard); and instructions for use.

A suitable detection molecule may be chosen from the list consisting an enzyme, a dye, a fluorescent molecule, a chemiluminescent, an isotope or such agents as colloidal gold conjugated to molecules including, but not limited to, such molecules as staphylococcal protein A or streptococcal protein G.

In a particular embodiment, the capture and detection antibodies are monoclonal antibodies, the production of which comprises immunizing a non-human animal, such as a mouse or a transgenic mouse, with a protein or chimeric molecule of the present invention, followed by standard methods, as hereinbefore described. Monoclonal antibodies may alternatively be produced by recombinant methods, as hereinbefore described and may comprise human or chimeric antibody portions or domains.

In another embodiment, the capture and detection antibodies are polyclonal antibodies, the production of which comprises immunizing a non-human animal, such as a mouse, rabbit, goat or horse, with a protein or chimeric molecule of the present invention, followed by standard methods, as hereinbefore described.

The components of the immunoassay kit are provided in predetermined ratios, with the relative amounts of the various reagents suitably varied to provide for concentrations in solution of the reagents that substantially maximize the sensitivity of the assay. Particularly, the reagents may be provided as dry powders, usually lyophilized, including excipients, which on dissolution provide for each reagent solution having the appropriate concentration for combining with the biological preparation to be tested.

The instructions for use may detail the method for using the immunoassay kit of the present invention. For example, the instructions for use may describe the method for coating the solid phase support matrix with a prepared solution of capture antibody under suitable conditions, for example, overnight at 4°C. The instructions for use may further detail blocking non-specific protein binding sites with the prepared blocking solution; adding and incubating serially diluted sample containing the protein or chimeric protein of the present invention under suitable conditions, for example, 1 hour at 37°C or 2 hours at room temperature, followed by a series of washes using a suitable buffer known in the art,

for example, a solution of 0.05% Tween 20 in 0.1M PBS (pH 7.2). In addition, the instructions may provide that a preparation of detection antibody is applied followed by incubation under suitable conditions, for example, 1 hour at 37°C or 2 hours at room temperature, followed by a further series of washes. A working solution of detection buffer is prepared from the supplied detection substrate(s) and substrate buffer, then added to each well under a suitable conditions ranging from 5 minutes at room temperature to 1 hour at 37°C. The chromatogenic reaction may be halted with the addition of 1N NaOH or 2N H_2SO_4.

In an alternative embodiment, the instructions for use may provide the simultaneous addition of any combination of any or all of the above components to be added in predetermined ratios, with the relative amounts of the various reagents suitably varied to provide for concentrations in solution of the reagents that substantially maximize the formation of a measurable signal from formation of a complex.

The level of colored product, or fluorescent or chemiluminescent or radioactive or other signal generated by the bound, conjugated detection reagents can be measured using an ELISA-plate reader or spectrophotometer, at an appropriate optical density (OD), or as emitted light, using a spectrophotometer, fluorometer or flow cytometer, at an appropriate wavelength, or using a radioactivity counter, at an appropriate energy spectrum, or by a densitometer, or visually by comparison to a chart or guide. A serially diluted solution of the standard preparation is assayed in parallel with the above sample. A standard curve or chart is generated and the level of the protein or chimeric molecule thereof present within the sample can be interpolated from the standard curve or chart.

The subject invention also provides a human derived protein or chimeric molecule thereof for use as a standard protein in an immunoassay. The present invention further extends to a method for determining the level of human cell-expressed human protein or chimeric molecule thereof in a biological preparation comprising a suitable assay for measuring the human protein or the chimeric molecule wherein the assay comprises (a) combining the biological preparation with one or more antibodies directed against the human protein or chimeric molecule thereof; (b) determining the level of binding of the or each antibody to

the human protein or the chimeric molecule in the biological preparation; (c) combining a standard human protein or a chimeric molecule sample with one or more antibodies directed against the human protein or the chimeric molecule; (d) determining the level of binding of the or each antibody to the standard human protein or the chimeric molecule sample; (e) comparing the level of the or each antibody bound to the human protein or the chimeric molecule in the biological preparation to the level of the or each antibody bound to the standard human protein or chimeric molecule sample.

In particular, the standard human protein or chimeric molecule sample is a preparation comprising the protein or chimeric molecule of the present invention.

The biological preparation includes but is not limited to laboratory samples, cells, tissues, blood, serum, plasma, urine, stool, saliva and sputum. The biological preparation is bound to one or more capture antibody as described hereinbefore or by methods known in the art. For instance, the solid phase support matrix is first coated with a prepared solution of capture antibody under suitable conditions (for example, overnight at 4°C); followed by blocking non-specific protein binding sites with the prepared blocking solution; then adding and incubating serially diluted sample containing a protein or chimeric molecule of the present invention under suitable conditions (for example, 1 hour at 37°C or 2 hours at room temperature), followed by a series of washes using a suitable buffer known in the art (for example, a solution of 0.05% Tween 20 in 0.1M PBS (pH 7.2)).

The biological preparation is then combined with one or more detection antibodies conjugated to a suitable detection molecule as described herein. For instance, applying a preparation of detection antibody followed by incubation under suitable conditions (for example, 1 hour at 37°C or 2 hours at room temperature), followed by a further series of washes.

Determination of the level of binding may be carried out as described hereinbefore or by methods known in the art. For instance, a working solution of detection buffer is prepared from the detection substrate(s) and substrate buffer, then adding to each well under a

suitable conditions ranging from 5 minutes at room temperature to 1 hour at 37°C. The chromatogenic reaction may be halted with the addition of 1N NaOH or 2N H_2SO_4.

In a particular embodiment, the present invention contemplates an isolated protein or chimeric molecule as hereinbefore described.

In an embodiment, an SCF of the present invention is characterized by a profile of one or more of the following physiochemical parameters (P_x) and pharmacological traits (T_y), comprising:

- an apparent molecular weight (P_1) of about 1 to 250, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250 kDa and in one embodiment, 18 to 55 kDa;
- a pI (P_2) range of about 2 to about 14 such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and in one embodiment, 3 to 7.5;
- about 2 to 100 isoforms (P_3), such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 isoforms and in one embodiment 15 to 86 isoforms;
- a percentage by weight carbohydrate (P_5) of about 0 to 99%, such as 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% and in one embodiment, 0 to 62%;
- an observed molecular weight after the N-linked oligosaccharides are removed (P_6) of about 18 to 40 kDa;

- an observed molecular weight after the N-linked and O-linked oligosaccharides are removed (P_7) of about 18 to 35 kDa;
- a site of N-glycosylation (P_{21}) which includes N-145 (numbering from the start of the signal sequence) identified by PMF after PNGase treatment;
- an immunoreactivity profile (T_{13}) that is distinct from that of a human SCF molecule expressed in a non-human cell system, and in one embodiment, the protein concentration of the SCF of the present invention is underestimated when assayed using a quantitative immunoassay which includes a protein standard of a human SCF molecule expressed in *E. coli* cells;
- a biological activity that is distinct from that of a human SCF expressed in a non-human cell system, and in one embodiment, the ability of SCF of the present invention to induce proliferation (T_{32}) in M-07e cells is 1.5-3.0 fold more potent than a human SCF expressed in *E. coli* cells;
- a biological activity that is distinct from that of a human SCF expressed in a non-human cell system, and in one embodiment, the ability of SCF of the present invention to produce a greater yield of viable cells (T_{73}) from $CD34^+$ human haematopoietic cells than that produced by treatment with rhSCF expressed in *E. coli* cells;
- a biological activity that is distinct from that of a human SCF expressed in a non-human cell system, and in one embodiment, the ability of SCF of the present invention to induce a 2-fold greater activation of STAT5 (T_{45}) in human $CD34^+$ haematopoietic stem cells than that induced by a rhSCF expressed in *E. coli* cells;
- a biological activity that is distinct from that of a human SCF expressed in a non-human cell system, and in one embodiment, the ability of SCF of the present invention to induce a 4.5-fold greater activation of Akt (T_{47}) in human $CD34^+$ haematopoietic stem cells than that induced by a rhSCF expressed in *E. coli* cells.

In another embodiment, SCFR-Fc of the present invention is characterized by a profile of one or more of the following physiochemical parameters (P_x) and pharmacological traits (T_y), comprising:

- an apparent molecular weight (P_1) of about 1 to 250, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32,

33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250 kDa and in one embodiment, 55 to 180 kDa;

- a pI (P_2) range of about 2 to about 14 such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and in one embodiment, 3.5 to 8.0;
- about 2 to 150 isoforms (P_3), such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150 isoforms and in one embodiment 10 to 109 isoforms;
- a percentage by weight carbohydrate (P_5) of about 0 to 99%, such as 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% and in one embodiment, 0 to 54%;
- an observed molecular weight of the molecule after the N-linked oligosaccharides are removed (P_6) of about 65 to 135 kDa;
- an observed molecular weight of the molecule after the N-linked and O-linked oligosaccharides are removed (P_7) of about 65 to 120 kDa;
- sites of N-glycosylation (P_{21}) which include N-130 and N-613 (numbering from the start of the signal sequence);
- an ability to inhibit rh SCF- rh SCF (expressed from human cells) induced activation of STAT3 (T_{45}) in M-07e cells;
- a biological activity that is distinct from that of a rh SCFR molecule expressed in a non-human cell system, and in one embodiment, a greater ability of SCFR-Fc of

the present invention to inhibit SCF induced activation of p42 and p44 MAPKs (T_{46}) in M-07e cells than a rh SCFR molecule expressed from *Sf*21 insect cells;

- a biological activity that is distinct from that of a rh SCFR molecule expressed in a non-human cell system, and in one embodiment, a greater ability of SCFR-Fc of the present invention to inhibit SCF induced activation of p42 and p44 MAPKs (T_{46}) in human $CD34^+$ haematopoietic stem cells than a rh SCFR molecule expressed from *Sf*21 insect cells;

In another embodiment, the present invention contemplates an isolated form of SCFR such as a soluble extra-cellular domain of SCFR. Such a molecule is also referred to herein as either soluble SCFR or SCFR-ECD of the present invention.

In an embodiment, a SCFR-ECD of the present invention is characterized by a profile of one or more of the following physiochemical parameters (P_x) and pharmacological traits (T_y), comprising:

- an apparent molecular weight (P_1) of about 1 to 250, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250 kDa and in one embodiment, 55 to 150 kDa;
- a pI (P_2) range of about 2 to about 14 such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and in one embodiment, 3.5 to 7.5;
- about 2 to 100 isoforms (P_3), such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 isoforms and in one embodiment 15 to 63 isoforms;
- a percentage by weight carbohydrate (P_5) of about 0 to 99%, such as 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29,

30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% and in one embodiment, 0 to 63%;

- an observed molecular weight of the molecule after the N-linked oligosaccharides are removed (P_6) of about 42 to 85 kDa;
- an observed molecular weight of the molecule after the N-linked and O-linked oligosaccharides are removed (P_7) of about 42 to 75 kDa;
- four sites of N-glycosylation (P_{21}) which include N-130 N-367, N-463 and N-486 (numbering from the start of the signal sequence);
- a biological activity that is distinct from that of a human SCFR-ECD expressed in a non-human cell system, and in one embodiment, a greater ability of SCFR-ECD of the present invention to inhibit SCF induced activation of p42 and p44 MAPKs in M-07e cells than a rh SCFR-ECD expressed from *Sf*21 insect cells;
- a biological activity that is distinct from that of a human SCFR-ECD expressed in a non-human cell system, and in one embodiment, a greater ability of SCFR-ECD of the present invention to inhibit SCF induced activation of p42 and p44 MAPKs in human $CD34^+$ haematopoietic stem cells than a rh SCFR-ECD expressed from *Sf*21 insect cells.

The physiochemical form of the protein or chimeric molecule of the present invention may be achieved by modifying the host cell by a variety of ways known in the art, including but not limited to the introduction of one or more transgene into the host cell that encodes an enzyme or enzymes that will produce the desired physiochemical form. Such transgenes include various types of sialyltransferases, such as ST3Gal1, ST3Gal2, ST3Gal3, ST3Gal4, ST3Gal5, ST3Gal6, ST6Gal1, ST6Gal2, ST6GalNAc1, ST6GalNAc2, ST6GalNAc3, ST6GalNAc4, ST6GalNAc5, ST8Sia1, ST8Sia2, ST8Sia3, ST8Sia4, ST8Sia5, ST8Sia6; galactosyltransferases, such as GalT1, GalT2; fucosyltransferases such as FUT1, FUT2, FUT3, FUT4, FUT5, FUT6, FUT7, FUT8, FUT9, FUT10, FUT11; sulfotransferases; GlcNAc transferases such as GNT1, GNT2, GNT3, GNT4, GNT5; antenna-cleaving enzymes and endoglycosidases.

For instance, inefficient terminal sialylation of N-glycan structures that results in reduced serum half-life of an expressed protein such as recombinant human AchE can be ameliorated by the addition of a rat beta-galactoside alpha-2,6-sialyltransferase transgene to HEK 293 cells (Chitlaru *et al. J Biochem 336*:647-658, 1998; Chitlaru *et al. J Biochem 363*:619-631, 2002).

Similarly, inefficient formation of particular Lewis x groups such as sialyl Lewis x structures on N-glycan structures that results in reduced ligand binding of an expressed protein such as recombinant human PSGL-1 can be ameliorated by the addition of a fucosyltransferase transgene to HEK 293 cells (Fritz *et al. PNAS 95*:12283-12288, 1998).

In one embodiment, a protein or chimeric molecule thereof is produced using a human cell line transformed with either α-2,3 or α-2,6 sialytransferase, or both α-2,3 sialytransferase and α-2,6 sialytransferase ("sialylated-protein"). Examples of sialylated-protein include sialylated-SCF, sialylated-SCFR-ECD and sialylated-SCFR-Fc.

In particular, the sialylated-protein is characterized by a profile of physiochemical parameters (P_x) comprising one or more physiochemical parameters. Monosaccharide (P_9) and sialic acid contents (P_{10}) of the sialylated-protein are, when normalized to GalNAc, 1 to 0.1-100 NeuNAc; and when normalized to 3 times of mannose 3 to 0.1-100 NeuNAc. Neutral percentage of N-linked oligosaccharides (P_{13}) of the sialylated-protein is 0 to 99% such as 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98 or 99%. Acidic percentage of N-linked oligosaccharides (P_{14}) of the sialylated-protein is 1 to 100% such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100%. Neutral percentage of O-linked oligosaccharides (P_{15}) of the sialylated-protein is 0 to 99% such as 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13,

14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98 or 99%. Acidic percentage of O-linked oligosaccharides (P_{16}) of the sialylated-protein is 1 to 100% such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100%. The *in vivo* half-life (T_{11}) of the sialylated-protein is increased in comparison to the half-life of the protein or chimeric molecule of the invention expressed without the transgene.

In one embodiment, the sialylated-protein contains at least one of the structural formulae described herein or at least one of the structural formulae described herein where one or more NeuNAc linkage is a α 2,6 linkage in the N-linked fraction.

In one embodiment, the sialylated-protein contains at least one of the structural formulae described herein or at least one of the structural formulae described herein where one or more NeuNAc linkage is a α 2,6 linkage in the O-linked fraction.

In one embodiment, the protein or chimeric molecule thereof of the invention is produced using a human cell line transformed with FUT3 ("fucosylated-protein"). Examples of fucosylated-protein include fucosylated-SCF, fucosylated-SCFR-ECD and fucosylated-SCFR-Fc.

In particular, the fucosylated-protein is characterized by a profile of physiochemical parameters (P_x) comprising one or more of physiochemical parameters. Monosaccharide (P_9) and sialic acid contents (P_{10}) of the fucosylated-protein are, when normalized to GalNAc, 1 to 0.1-100 NeuNAc; and when normalized to 3 times of mannose 3 to 0.1-100 NeuNAc.

In one embodiment, the fucosylated-protein has a higher proportion of structure containing Lewis structures (such as Lewis a, Lewis b, Lewis x or Lewis y) or sialyl Lewis structures (such as sialyl Lewis a or sialyl Lewis x).

In one embodiment, the fucosylated-protein has altered binding affinity to ligands in comparison to the binding affinity of the protein or chimeric molecule of the invention expressed without the transgene.

Using respective forward primer and reverse primer for the protein molecule selected from SCF, SCFR-ECD and SCFR-Fc, the DNA encoding the relevant protein was amplified from an EST by Polymerase Chain Reaction (PCR) by methods known in the art, for example, according to the method of Invitrogen's PCR Super Mix High Fidelity (Cat. No.:10790-020). The amplicon is digested and ligated into the corresponding restriction enzyme sites of an appropriate vector, for instance, pIRESbleo3, pCMV-SPORT6, pUMCV3, pORF, pORF9, pcDNA3.1/GS, pCEP4, pIRESpuro3, pIRESpuro4, pcDNA3.1/Hygro(+), pcDNA3.1/Hygro(-), pEF6/V5-His. The ligated vector is transformed into an appropriate *E. coli* host cell, for instance, XLGold, ultracompetant cell (Strategene), XL-Blue, DH5α, DH10B or the like.

For the production of chimeric molecules, the DNA sequence for the Fc domain of an immunoglobulin, such as IgG1, IgG2, IgG3, IgG4, IgGA1, IgGA2, IgGM, IgGE, IgGD is amplified from the EST using the appropriate forward and reverse primers by PCR. The amplicon is cloned into the corresponding restriction enzyme sites of an appropriate vector, for instance, pIRESbleo3, pCMV-SPORT6, pUMCV3, pORF, pORF9, pcDNA3.1/GS, pCEP4, pIRESpuro3, pIRESpuro4, pcDNA3.1/Hygro(+), pcDNA3.1/Hygro(-), pEF6/V5-His. The DNA sequence of relevant protein is amplified and cloned into the corresponding restriction enzyme sites of the respective Fc-vector in frame with the Fc.

In a particular embodiment, the Fc receptor binding region or the complement activating region of the Fc region may be modified recombinantly, comprising one or more amino acid insertions, deletions or substitutions relative to the amino acid sequence of the Fc region. In addition, the receptor binding region or the complement activating region of the

Fc region may be modified chemically by changes to its glycosylation pattern, the addition or removal of carbohydrate moieties, the addition of polyunsaturated fatty acid moieties or other lipid based moieties to the amino acid backbone or to any associated co- or post-translational entities. The Fc region may also be in a truncated form, resulting from the cleavage by an enzyme including papain, pepsin or any other site-specific proteases. The Fc region may promote the spontaneous formation by the chimeric protein of a dimer, trimer or higher order multimer that is better capable of binding to its corresponding ligand or receptor.

Diagnostic digests using the appropriate restriction enzymes are performed to identify/isolate bacterial colonies containing the vector bearing the correct gene. Positive colonies are isolated and stored as Glycerol stocks at –70°C. The clone is then expanded to 750ml of sterile LB broth containing ampicillin (100µg/ml) at 37°C with shaking for 16 hours. The plasmid is prepared in accordance with methods known in the art, preferably, in accordance with a Qiagen Endofree Plasmid Mega Kit (Qiagen Mega Prep Kit #12381).

Human host cells suitable for the introduction of the cloned DNA sequence comprising a the protein or chimeric molecule of the present invention include but are not limited to HEK 293 and any derivatives thereof, HEK 293 c18, HEK 293-T, HEK 293 CEN4, HEK 293F, HEK 293FT, HEK 293E, AD- 293 (Stratagene), 293A (Invitrogen), Hela cells and any derivatives thereof, HepG2, PA-1 Jurkat, THP-1, HL-60, H9, HuT 78, Hep-2, Hep G2, MRC-5, PER.C6, SKO-007, U266, Y2 (Apollo), WI-38, WI-L2.

The physiochemical form of protein or chimeric molecule of the present invention may be achieved by modifying the host cell by a variety of ways known in the art, including but not limited to the introduction of a transgene into the host cell that encodes an enzyme or enzymes that will produce the desired physiochemical form. The introduction of specific DNA sequences can be used to optimize the integration of the cloned DNA sequence into the host cell genome, the various types of integration including but not limited to site-specific, targeted, direct or enzyme-mediated integration.

The DNA of protein or chimeric molecule thereof can be introduced into suitable host cells by various transfection methods known in the art, for instance, using chemical reagents such as DEAE-dextran, calcium phosphate, artificial liposomes, or by direct microinjection, electroporation, biolistic particle delivery or infection or transfection with viral constructs as described below.

DEAE-dextran is a cationic polymer that associates with negatively charged nucleic acids. An excess of positive charge, contributed by the polymer in the DNA/polymer complex allows the complex to come into closer association with the negatively charged cell membrane. Uptake of the complex is presumably by endocytosis. Other synthetic cationic polymers including polybrene, polyethyleneimine and dendrimers have also been used for transfection.

Calcium phosphate co-precipitation can be used for transient and stable transfection of a variety of cell types. The DNA is mixed with calcium chloride in a controlled manner and added to a buffered saline/phosphate solution and the mixture is incubated at room temperature. A precipitate is generated and is taken up by the cells via endocytosis or phagocytosis.

The most commonly used synthetic lipid component of liposomes for liposome-mediated gene delivery is one which has overall net positive charge at physiological pH. Often the cationic lipid is mixed with a neutral lipid such as L-dioleoyl phosphatidylethanolamine (DOPE). The cationic portion of the lipid molecule associates with the negatively charged nucleic acids, resulting in compaction of the nucleic acid in a liposome/nucleic acid complex. Uptake of the complex is by endocytosis.

Direct microinjection of DNA into cultured cells or nuclei is an effective, although laborious technique, which is not appropriate if a large number of transfected cells are required.

Electroporation utilizes an electric pulse, which generates pores that allow the passage of nucleic acids into the cells. This technique requires fine-tuning and optimization for

duration and strength of the pulse for each type of cell used. Commercially available electroporation device includes Amaxa Biosystems' Nucleofector Kits (Amaxa Biosystems, Germany).

This method relies upon high velocity delivery of nucleic acids on microprojectiles to recipient cells.

Infection or transfection with viral or retroviral constructs include the use of retrovirus, such as lentivirus, or DNA viruses, such as adenovirus. The process involves using a viral or retroviral vector to transfer a foreign gene to the host's cells.

In some embodiments, the protein or chimeric molecule thereof is produced by either transient methods or from stably transfected cell lines. Transient transfection is performed using either adherent or suspension cell lines. For adherent cell lines, the cells are grown in serum containing medium (between 2-10% serum) and in medium such as DMEM, DMEM/F12 (JRH). Serum used can be fetal calf serum (FCS), donor calf serum (DCS), new born calf serum (NBCS) or the like. Plasmid vectors are introduced into the cells by standard methods known in the art. In a particular embodiment, the DNA of the protein or chimeric molecule thereof is transfected using DEAE dextran or calcium phosphate precipitation. Following transfection, the cells are switched to an appropriate collection medium (e.g. serum free DMEM/F12) for collection of the expressed protein or chimeric molecule thereof.

Transient expression of the protein or chimeric molecule thereof from suspension cells can be performed by introducing the plasmid vector using the methods outlined above. The suspension cells can be grown in either serum containing medium, or in serum free medium (e.g. Freestyle medium (Invitrogen), CD293 medium (Invitrogen), Excell medium (JRH) or the like). The transfection can be performed in the absence of serum by transfecting in an appropriate media using a suitable transfection method, for instance, lipofectamine in OptiMEM medium.

Transient expression usually results in a peak of expression 2-3 days after transfection. Episomal vectors are replicated within the cell and give sustained expression. Therefore, to

obtain large amounts of product, episomal expression vectors are transfected into cells and the cells are expanded. A protein or chimeric molecule thereof is expressed into the medium, which is collected as the cells are expanded over a period of weeks. The expression medium can be serum containing or serum free and the cells can be either adherent or suspension adapted.

Stable clones are obtained by transfection of the expression vector into the cells, then selecting with an appropriate agent, for instance, phleomycin, hygromycin, puromycin, neomycin G418, methotrexate or the like. Stable clones will survive selection as the plasmid contains a resistance gene in addition to the gene encoding the protein or the chimeric molecule. One to two days after introduction of the gene, selection is begun on either the whole population of cells (stable pools) or on cells plated at clonal density. A non-transfected population of cells is also selected to determine the efficacy of cell killing by the selective agent. For adherent cells, the cells are allowed to grow on a tissue culture plate until visible separate clones are obtained. They are then removed from the plate by trypsinization, or physical removal and placed into tissue culture wells (eg, one clone per well of a 96 well plate). For suspension cells, limiting dilution cloning is performed subsequent to selection. The clones are then expanded, then either characterized and/or subjected to a further round of limiting dilution analysis.

Stable clones growing in serum containing medium can be adapted by gradual reduction of serum levels followed by detachment and growth under low serum in suspension. The serum levels are then reduced further until serum free status is achieved. Some growth media allow more rapid adaptation (e.g. a straight swap from serum containing adherent conditions to serum free suspension growth), an example of which is Invitrogen's CD293 media.

Following growth in serum free media, the clones can begin media optimization. The clones are tested for production characteristics, for example, integral viable cell number, in many different growth media until an optimum formulation or formulations are obtained. This may depend on the method of production of the product. For instance, the cells may

be expanded in one medium, then additives that enhance expression added prior to product collection.

The over-expressed protein or chimeric molecule may accumulate within host cells. Recovery of intracellular protein involves treatment of the host cells with lysis buffers including but not limited to buffers containing: NP40, Triton X-100, Triton X-114, sodium dodecyl sulfate (SDS), sodium cholate, sodium deoxycholate, CHAPS, CHAPSO, Brij-35, Brij-58, Tween-20, Tween-80, Octylglucoside and Octylthioglucoside. Alternative methods of host cell lysis may include sonication, homogenization, french press treatment and repeated cycles of freeze thawing and treatment of the cells with hypotonic solutions.

The final product can be produced in many different sorts of bioreactors, by way of non-limiting examples, including stirred tank, airlift, packed bed perfusion, microcarriers, hollow fibre, bag technologies, cell factories. The methods may be continuous culture, batch, fed batch or induction. Peptones may be added to low serum cultures to achieve increases in volumetric protein production.

The protein or chimeric molecule of the present invention is purified using a purification strategy specifically tailored for protein or chimeric molecule of the present invention. Purification methods include but are not limited to: tangential flow filtration (TFF); ammonium sulfate precipitation; size exclusion chromatography (SEC); gel filtration chromatography (GFC); affinity chromatography (AFC); Protein A Affinity Purification; Receptor mediated Ligand Chromatography (RMLC); dye ligand chromatography (DLC); ion exchange chromotogaphy (IEC), including anion or cation exchange chromatography (AEC or CEC); reversed-phase chromatography (RPC); hydrophobic interaction chromatography (HIC); metal chelating chromatography (MCC).

TFF is a rapid and efficient method for biomolecule separation and is used for concentrating, desalting, or fractionating samples. TFF can concentrate samples as large as hundreds of litres down to as little as 10 ml. In conjunction with a suitable molecular weight cut off membrane, TFF can separate and isolate biomolecules of differing size and molecular weight (nominal molecular weight cutoff (NMWC) 5 KDa, 10 KDa, 30 KDa,

100 KDa). The process of diafiltration involving dilution of the sample followed by re-concentration can be used to desalt or exchange the sample buffer.

Salting out or ammonium sulfate precipitation is useful for concentrating dilute solutions of proteins. It is also useful for fractionating a mixture of proteins. Increases in the ionic strength of a solution containing protein causes a reduction in the repulsive effect of like charges between protein molecules. It also reduces the forces holding the solvation shell around the protein molecules. When these forces are sufficiently reduced, the protein will precipitate; hydrophobic proteins precipitating at lower salt concentrations than hydrophilic proteins. Fractionation of protein mixtures by the stepwise increase in the ionic strength followed by centrifugation can be a very effective way of partly purifying proteins.

SEC separates proteins by size, based on the flow of the sample through a porous matrix. SEC has the same principle as GFC when it is used to separate molecules in aqueous systems. In SEC, molecules larger than pores of the packing elute with the solvent front first and are completely excluded. Intermediate sizes of molecules, between the completely excluded and the retained, pass through the pores of the matrix according to their sizes. Small molecules which freely pass in and out of the pores are retained. Therefore, different sizes of proteins have different elution volume and retention times. For structurally similar molecules, the larger the molecular sizes, the earlier they elute out. Before running any samples, a standard curve should be established to determine the working limits and reference retention time.

When the protein shapes are the same, molecular weight can be screened in the elutes from the column rapidly by UV absorption, fluorescence or light scattering, according to the packing materials of various pore sizes on the column. Photon correlation spectroscopy (PCS) has been usually performed on static samples and for liquid chromatographic detection. Low angle laser light scattering has also been coupled to chromatographic detection to detect the molecular weights directly, independent of the shapes of the proteins (Carr *et al. Anal Biochem 175:*492-499, 1988). SEC-HPLC was used to detect hGH degradation and aggregation (Pikal *et al. Pharm Res 8*:427-436, 1991). It was also

used for estimation of contamination in studying β-galactosidase (Yoshioka *et al. Pharm Res 10*:103-108, 1993).

AFC purifies biological molecules according to specific interactions between their chemical structures and the suitable affinity ligands. The target molecule is adsorbed by a complementary immobilized ligand specifically and reversibly. The ligand can be an inhibitor, substrate, analog or cofactor, or an antibody which can recognize the target molecules specifically. Subsequently, the adsorbed molecules are either eluted by competitive displacement, or by the conformation change through a pH or ionic strength shift.

Protein A Affinity Purification is an example of affinity purification utilising the affinity of certain bacterial proteins that bind generally to antibodies, regardless of the antibody's specificity to antigen. Protein A, Protein G and Protein L are three that have well characterised antibody-binding properties. These proteins have been produced recombinantly and used routinely for affinity purification of key antibody types from a variety of species. A genetically engineered recombinant form of Protein A and G, called Protein A/G, is also available. These antibody-binding proteins can be immobilized to support matrixes. This method has been modified to purify recombinant proteins that have had the Protein A binding region of an antibody (Fc region) linked to the target protein. Binding to the immobilised Protein A molecule is performed under physiological conditions and eluted by change in pH or ionic strength.

RMLC is a special kind of AFC utilising the inherent affinity of a receptor for its cognate target molecule. The receptor molecule is immobilised on a suitable chromatography support matrix via reactive amines, reactive hydrogens, carbonyl, carboxyl or sulfhydryl groups. In one example of RMLC, the receptor-Fc chimera molecule is immobilised on Protein A sepharose beads via affinity of the Fc portion of the receptor to the Protein A. This method has the advantage of immobilising the receptor in an orientation that exposes its ligand-binding site to its cognate cytokine. Adsorption of the target molecule to the receptor is performed under physiological conditions and elution is achieved by change in pH or ionic strength.

DLC is a kind of ALC utilizing the ability of reactive dyes to bind proteins in a selective and reversible manner. The dyes are generally monochlorotriazine compounds. The reactive chloro group allows easy immobilization of the triazine dye to a support matrix, such as Sepharose or agarose, and, more recently, to nylon membranes.

The initial discovery of the ability of these dyes to bind proteins came from the observation that blue dextran (a conjugate of cibacron blue FG-3A), used as a void volume marker on gel filtration columns, could retard the elution of certain proteins. A number of studies have been carried out on the specificity of the dyes for particular proteins, mostly using the prototype cibacron blue dye. The dyes appear to be most effective at binding proteins and enzymes that utilize nucleotide cofactors, such as kinases and dehydrogenases, although other proteins such as serum albumin also bind tightly. It has been proposed that the aromatic triazine dye structure resembles the nucleotide structure of nicotinamide adenine dinucleotide (NAD) and that the dye interacts with the dinucleotide fold in these proteins. In many cases, bound proteins can be eluted from the columns by a substrate or nucleotide cofactor in a competitive fashion, and dyes have been shown to compete for substrate-binding sites in free solution. It seems likely that these dyes can bind proteins by electrostatic and hydrophobic interactions and by more specific "pseudoaffinity" interactions with ligand-binding sites. Enhancing the specificity of dye ligands by modification to further resemble ligands (biomimetic dyes) has been successful in the purification of a number of dehydrogenases and proteases (McGettrick *et al. Methods Mol Biol 244*:151-7, 2004).

Ion Exchange Chromatography (IEC) purifies proteins using protein retention on columns resulting from the electrostatic interactions between the ion exchange column matrix and the proteins. When the pH of the mobile phase is above the pI of the target protein will be negatively charged and will interact with an anion exchange column (AEC). When the pH of the mobile phase is below the pI of the target protein the protein will be positively charged and a cation exchange column (CEC) should be used. The target proteins are eluted by increasing the concentrations of a counter ion with the same charge as the target molecule.

RPC separates biological molecules according to the hydrophobic interactions between the molecule and a chromatographic support matrix. Ionizable compounds are best analyzed in their neutral form by controlling the pH of the separation. Mobile phase additives, such as trifluoroacetic acid, increase protein hydrophobicity by forming ion pairs which strongly adsorb to the stationary phase. By changing the polarity of the mobile phase, the biological molecules are eluted from the chromatographic support.

HIC is similar to RPC, but with a larger nominal pore size. In HIC, the elution solvent uses an aqueous salt solution, instead of the aqueous or organic mobile phases used in RPC. Also, the order of sample elution is reversed from that obtained from RPC. The surfaces of proteins consist of hydrophilic residues and hydrophobic "patches", which are usually located in the interior of the folded proteins to stabilize the proteins. When the hydrophobic patches become exposed to the aqueous environment, they will disrupt the normal solvation properties of the protein, which is thermodynamically unfavorable. In the aqueous mobile phase, the higher the concentrations of inorganic salts (e.g. ammonium sulfate), the higher surface tension, thereby increasing the strength of hydrophobic interactions between the hydrophobic groups of the HIC resin and the proteins, which are adsorbed. However, while descending the salt concentration gradient, the surface tension of the aqueous mobile phase is decreased, thus reducing the hydrophobic interaction, resulting in the proteins desorbing from the hydrophobic groups of the column.

MCC is a technique in which proteins are separated on the basis of their affinity for chelated metal ions. Various metal ions including but not limited to Cu^{2+}, Co^{2+}, Zn^{2+}, Mn^{2+}, Mg^{2+} or Ni^{2+} are immobilized on the stationary phase of a chromatographic support via a covalently bound chelating ligand (e.g. iminodiacetic acid). Free coordination sites of the metal ions are used to bind different proteins and peptides. Elution can occur by displacement of the protein with a competitive molecule or by changing the pH. For instance, a lowering of the pH in the buffer results in a reduced binding affinity of the protein-metal ion complex and desorption of the protein. Alternatively, bound proteins can be eluted from the column using a descending pH gradient, in the form of a step gradient or as linear gradient.

The physiochemical form of the protein or chimeric molecule of the present invention may be achieved by chemical and/or enzymatic modification to the expressed molecule in a variety of ways known in the art.

The present invention contemplates chemical or enzymatic coupling of carbohydrates to the peptide chain of a protein or chimeric molecule at a time after the protein or chimeric molecule is expressed and purified. Chemical and/or enzymatic coupling procedures may be used to modify, increase or decrease the number or profile of carbohydrate substituents. Depending on the coupling mode used, the sugar(s) may be attached to (a) amide group of arginine, (b) free carboxyl groups, (c) sulfhydroxyl groups such as those of cysteine, (d) hydroxyl groups such as those of serine, threonine, hydroxylysine or hydroxyproline, (e) aromatic residues such as those of phenylalanine, tyrosine, or tryptophan, (f) the amide group of glutamine, or (g) the amino groups such as those of histidine, arginine or lysine. Additions can be carried out chemically or enzymatically. For example serial addition of sugar units to the protein or chimeric molecule thereof can be performed using appropriate recombinant glycosyltransferases. Glycosyltransferases can also be used to add sugars that have covalently attached substituents. For example, sialic acid with covalently attached polyethylene glycol (PEG) can be transferred by a sialyltransferase to a terminal galactosyl residue to increase molecular size and serum half-life.

The carbohydrate side chain of a protein or chimeric molecule can also be modified chemically or enzymatically to incorporate a variety of functionalities, including phosphate, sulfate, hydroxyl, carboxylate, O-sulfate and N-acetyl groups.

Carbohydrates present on a protein or chimeric molecule thereof may also be removed chemically or enzymatically. Trifluoromethanesulfonic acid or an equivalent compound can be used for chemical deglycosylation. This treatment can result in the cleavage of most or all sugars, except the linking sugar, while leaving the polypeptide intact. Individual sugars or the entire chain can also be removed from a protein or chimeric molecule thereof by a variety of endoglycosidases and exoglycosidases.

The glycan component of a protein or a chimeric moleculemay be modified synthetically by treatment with sialidases, or mild acid treatment to remove any residual sialic acids; treatment with exo- or endo- glycosidases to trim down the antennae of N-linked oligosaccharides or shorten O-linked oligosaccharides. It may also be treated with fucosidases or sulfatases to remove side groups such as fucose and sulfate. Pseudo glycan structures such as polyethylene glycol or dextrans may be chemically added to the amino acid backbone, or a glycotransferase cocktail can be used with sugar-dUDP precursors to synthetically add sugar subunits to the glycan.

The present invention contemplates a protein or chimeric molecule thereof chemically or enzymatically coupled to radionuclides. Such protein or chimeric molecule may be selected from the list comprising SCF, SCFR-ECD, SCFR-Fc.

Iodination procedures may be used to attach iodine isotopes (e.g. ^{123}I) to the peptide chain of the protein or chimeric molecule thereof. In particular, the isotope(s) may be attached to a (a) phenolic ring of a tyrosine, or (b) the imidazole ring of a histidine on the peptide chain of the protein or the chimeric molecule thereof. Iodination may be performed using the Chloramine-T, iodine monochloride, triiodide, electrolytic, enzymatic, conjugation, demetallation, iodogen or iodo-bead methods.

Technetium labeling procedures may be used to attach ^{99m}Tc to the protein or chimeric molecule of the present invention using a method known in the art, for instance, by the reduction of $^{99m}TcO_4^-$ with a reducing agent (e.g. stannous chloride) followed by ^{99m}Tc labelling of the protein or the chimeric molecule via a bifunctional chelating agent, for instance, diethylenetriamine pentaacetic acid (DTPA).

The present invention contemplates a protein or chimeric molecule thereof chemically or enzymatically coupled to chemotherapeutic agents. Suitable agents (e.g. zoledronic acid) may be conjugated to the the protein or the chimeric molecule thereof using methods known in the art, for instance, by a N-hydroxysulfosuccinimide enhanced carbodiimide-mediated coupling reaction.

The present invention contemplates a protein or chimeric molecule thereof chemically or enzymatically coupled to toxins. Suitable toxins, including melittin, vanous toxin, truncated pseudomonas exotoxin, ricin, gelonin and diptheria toxin may be conjugated to the protein or the chimeric molecule using a method known in the art, for instance, by maleimide or carbodiimide coupling chemistry.

An isolated protein or chimeric molecule thereof described herein may be delivered to the subject by any means that produces contact of the isolated protein or the chimeric molecule with the target receptor or ligand in the subject. In a particular embodiment, a protein or chimeric molecule thereof is delivered to the subject as a "pharmaceutical composition".

In another aspect, the present invention contemplates a pharmaceutical composition comprising one or more isolated proteins or chimeric protein molecules as hereinbefore described together with a pharmaceutically acceptable carrier or diluent.

Composition forms suitable for injectable use include sterile aqueous solutions (where water soluble) and sterile powders for the extemporaneous preparation of sterile injectable solutions. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dilution medium comprising, for example, water, ethanol, polyol (for example, glycerol, propylene glycol and liquid polyethylene glycol, and the like), suitable mixtures thereof and vegetable oils. The proper fluidity can be maintained, for example, by the use of surfactants. The preventions of the action of microorganisms can be brought about by various anti-bacterial and anti-fungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thirmerosal and the like. In many cases, it will be favorable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminium monostearate and gelatin.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with the active ingredient and optionally other

active ingredients as required, followed by filtered sterilization or other appropriate means of sterilization. In the case of sterile powders for the preparation of sterile injectable solutions, suitable methods of preparation include vacuum drying and the freeze-drying technique which yield a powder of active ingredient plus any additionally desired ingredient.

When the active agent is suitably protected, it may be orally administered, for example, with an inert diluent or with an assimilable edible carrier, or it may be enclosed in hard or soft shell gelatin capsule, or it may be compressed into tablets, or it may be incorporated directly with the food of the diet or administered *via* breast milk. For oral therapeutic administration, the active ingredient may be incorporated with excipients and used in the form of ingestible tablets, buccal tablets, troches, capsules, elixirs, suspensions, syrups, wafers and the like. Such compositions and preparations should contain at least 1% by weight of active agent. The percentage of the compositions and preparations may, of course, be varied and may conveniently be between about 5 to about 80% of the weight of the unit. The amount of active agent in such therapeutically useful compositions is such that a suitable dosage will be obtained. In a particular embodiment, compositions or preparations according to the present invention are prepared so that an oral dosage unit form contains between about 0.1 μg and 200 mg of modulator. Alternative dosage amounts include from about 1 μg to about 1000 mg and from about 10 μg to about 500 mg. These dosages may be per individual or per kg body weight. Administration may be per hour, day, week, month or year.

The tablets, troches, pills, capsules and the like may also contain the components as listed hereafter. A binder such as gum, acacia, corn starch or gelatin; excipients such as dicalcium phosphate; a disintegrating agent such as corn starch, potato starch, alginic acid and the like; a lubricant such as magnesium stearate; and a sweetening agent such as sucrose, lactose or saccharin may be added or a flavouring agent such as peppermint, oil of wintergreen or cherry flavouring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills or capsules may be coated with shellac, sugar or both. A syrup or

elixir may contain the active compound, sucrose as a sweetening agent, methyl and propylparabens as preservatives, a dye and flavouring such as cherry or orange flavour. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compound(s) may be incorporated into sustained-release preparations and formulations.

The present invention also contemplates topical formulations. In a topical composition, the active agent may be suspended within a cream or lotion or wax or other liquid solution such that topical application of the cream or lotion or wax or liquid solution results in the introduction of the active agent to a biological surface in the subject. The term "biological surface" as used herein, contemplates any surface on or within the organism. Examples of "biological surfaces" to which the topical compositions of the present invention may be applied include any epithelial surface such as the skin, respiratory tract, gastrointestinal tract and genitourinary tract.

In addition to traditional cream, emulsion, patch or spray formulations, the agents of the present invention may also be delivered topically and/or transdermally using a range of iontophoric or poration based methodologies.

"Iontophoresis" is predicated on the ability of an electric current to cause charged particles to move. A pair of adjacent electrodes placed on the skin set up an electrical potential between the skin and the capillaries below. At the positive electrode, positively charged drug molecules are driven away from the skin's surface toward the capillaries. Conversely, negatively charged drug molecules would be forced through the skin at the negative electrode. Because the current can be literally switched on and off and modified, iontophoretic delivery enables rapid onset and offset, and drug delivery is highly controllable and programmable.

Poration technologies, use high-frequency pulses of energy, in a variety of forms (such as radio frequency radiation, laser, heat or sound) to temporarily disrupt the stratum corneum, the layer of skin that stops many drug molecules crossing into the bloodstream. It is important to note that unlike iontophoresis, the energy used in poration technologies is not

used to transport the drug across the skin, but facilitates its movement. Poration provides a "window" through which drug substances can pass much more readily and rapidly than they would normally.

Pharmaceutically acceptable carriers and/or diluents include any and all solvents, dispersion media, coatings, anti-bacterial and anti-fungal agents, isotonic and absorption delaying agents and the like. The use of such media and agents for pharmaceutical active substances is well known in the art and except insofar as any conventional media or agent is incompatible with the modulator; their use in the pharmaceutical compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

In another embodiment, the pharmaceutical composition comprising an isolated SCF of the present invention can be used, alone or in conjunction with other biologics, drugs or therapies in the same manner as a rh SCF expressed by non-human cell line, such as, a rh SCF expressed by E. coli, yeast, or CHO, in the treatment of diseases including: myelodysplastic syndromes, anemias, Diamond-Blackfan anemia, Fanconi's anemia, aplastic anemia, dyskeratosis congenita, thrombocytopenia, amegakaryocytic thrombocytopenia, transient erythroblastopenia of childhood, acute myocardial infarction, multiple myeloma, treatment of adult and pediatric marrow failure states, treatment of disorders of skin pigmentation, non-Hodgkin's lymphoma, Hodgkin's disease and carcinoma of the breast.

In an additional embodiment, the pharmaceutical composition comprising an isolated SCFR of the present invention can be used alone or in conjunction with other drugs or therapies in the same manner as a rh SCFR molecule or chimeric molecule thereof expressed by non-human cell line, such as, a rh SCFR molecule or chimeric molecule expressed by E. coli, yeast, or CHO for treatment of gastrointestinal stromal cell tumors and mastocytosis, myeloid leukemia, neuroblastoma, breast tumor, colon tumors, gynecological tumors, testicular germ cell tumors and small cell lung carcinoma (SCLC) and acne.

However, the pharmaceutical composition of the present invention has higher pharmaceutical efficacy, increased thermal stability, increased serum half-life or higher solubility in the bloodstream when compared with the protein or chimeric molecule thereof expressed in non-human cell lines. The present invention also shows reduced risks for immune-related clearance or related side effects. Because of these improved properties, the composition of the present invention can be administered at a lower frequency than a protein or chimeric molecule expressed in non-human cell lines. Decreased frequency of administration is anticipated to enhance patient compliance resulting in improved treatment outcomes. The quality of life of the patient is also elevated.

Accordingly, in one embodiment, the pharmaceutical composition of the present invention can be administered in a therapeutically effective amount to patients in the same way a protein or chimeric molecule expressed in non-human cell lines is administered. The therapeutic amount is that amount of the composition necessary for the desired *in vivo* activity. The exact amount of composition administered is a matter of preference subject to such factors as the exact type of condition being treated, the condition of the patient being treated and the other ingredients in the composition. The pharmaceutical compositions containing the isoforms of the protein or chimeric molecule of the present invention may be formulated at a strength effective for administration by various means to a human patient experiencing one or more of the above disease conditions. Average therapeutically effective amounts of the composition may vary. Effective doses are anticipated to range from 0.1ng/kg body weight to 20μg/kg body weight; or based upon the recommendations and prescription of a qualified physician.

The present invention further extends to uses of the isolated protein or the chimeric molecule comprising at least part of the protein or chimeric molecule thereof and a composition comprising same in a variety of therapeutic and/or diagnostic applications.

More particularly, the present invention extends to a method of treating or preventing a condition in a mammalian subject, wherein the condition can be ameliorated by increasing the amount or activity of the protein or chimeric molecule of the present invention, the method comprising administering to said mammalian subject an effective amount of an

isolated protein, a chimeric molecule comprising the protein, a fragment or an extracellular domain thereof or a composition comprising the isolated protein or the chimeric molecule.

The subject invention also provides a pharmaceutical composition comprising one or more molecules of the present invention selected from the list of SCF, SCFR-ECD, SCFR-Fc, for use, alone or in combination with other proteins expressed from human cell lines, for example, bFGF, EPO, G-CSF, in the culturing hES and other human progenitor cells, thus reducing the potential risks of transferring animal-derived infectious agents to hES and other human progenitor cells, and allowing hES and other human progenitor cells to be applicable for therapeutic applications.

The present invention is further described by the following non-limiting examples.

EXAMPLE 1

(a) Production of a Vector-Fc Construct

The DNA sequence encoding the Fc domain of human IgG1 was amplified from EST cDNA library (Clone ID 6277773, Invitrogen) by Polymerase Chain Reaction (PCR), using forward primer and reverse primer pair (SEQ ID NOs: 45, 46, respectively) which results in the incorporation of restriction enzyme sites BamH1 and BstX1, respectively. The resulting amplicons were cloned into the corresponding enzyme sites of pIRESbleo3 (Cat. No. 6989-1, BD Biosciences) to produce pIRESbleo3-Fc constructs. Digestion of pIRESbleo3-Fc constructs with BamH1 and BstX1 released an expected size insert of approximately 780 bp as determined by gel electrophoresis.

(b) Production of a DNA construct expressing a Protein or a Protein-Fc

The DNA sequence encoding the protein or the extra cellular domain thereof was amplified from an EST cDNA library or cloned cDNA sequence by PCR, using forward primer and reverse primers that incorporated restriction enzyme sites according to Table 8. After amplification, the amplicon was digested with suitable restriction enzymes and cloned into an expression vector as per Table 8, to produce the vector-Protein or vector-Protein-Fc constructs. Where a construct encoding a Protein-Fc was produced, the DNA sequence encoding the protein was cloned upstream of the Fc nucleotide sequence, such that the two sequences were fused in-frame so that when the protein was expressed it was fused directly or by a linker to the Fc domain. Suitable restriction enzymes were used to digest the vector containing the DNA sequence encoding the Protein or the Protein-Fc to release the expected size fragments as shown in Table 8. Vector-Protein or vector-Protein-Fc constructs were sequenced to confirm the integrity of the cloning procedures as herein described.

(c) Preparation of Megaprep vector-Protein or vector-Protein-Fc

750ml of sterile LB broth containing ampicillin (100μg/ml) was inoculated with 750μl of overnight culture of *E. Coli* transformed with vector-Protein or vector-Protein-Fc. The culture was incubated at 37°C with shaking for 16 hours. Plasmid was prepared in accordance with a Qiagen Endofree Plasmid Mega Kit (Qiagen Mega Prep Kit #12381).

TABLE 8

Protein-Fc and relevant cloning information

Protein	cDNA Source	Forward Primer	Reverse Primer	Restriction Enzyme sites	Vector	Size (bp)
SCF	Clone ID RG001444 (Invitrogen) pcDNA3.1/GS	21	22	EcoRV, Bam HI	pIRESbleo3 (Cat. No. 6989-1, BD Biosciences))	838
SCFR	Clone ID 4375615, (Invitrogen)	29	30	NotI, BamHI	pIRESbleo3	1581
SCFR-Fc	Clone ID 4375615, (Invitrogen)	29	47	NotI, BamHI	pIRESbleo3-Fc	1581

Alternatively, the nucleotide sequence of the Protein that was cloned into the vector (such as pIRESbleo3 or pCEP4) can be amplified with primers that incorporate restriction sites allowing the cloning of the DNA sequence encoding the Protein upstream of the Fc nucleotide sequence in a vector-Fc, such that the Protein and the Fc nucleotide sequences are fused in-frame directly or by a linker.

EXAMPLE 2

Production and Purification of Target Molecules of the Present Invention

(i) Production of Target Molecules of the Present Invention

At day 0, five 500 cm^2 tissue culture dishes (Corning) were seeded with 3×10^7 cells of transformed embryonal human kidney cell line, for example HEK 293 or derivatives thereof, HEK 293 c18, HEK 293T, 293 CEN4, HEK 293F, HEK 293FT, HEK 293E, AD-

293 (Stratagene), or 293A (Invitrogen). Cells were seeded in 90 ml per plate of Dulbecco's Modified Eagle's Medium/Ham's Nutrient Mixture F12 (DMEM/F12) (JRH Biosciences), the medium being supplemented with either 10% (v/v) heat-inactivated fetal calf serum (FCS, JRH Biosciences) or 10% (v/v) donor calf serum (DCS, JRH Biosciences), 4 mM L-glutamine (Amresco), 1 %(v/v) Penicillin-Streptomycin (Penicillin G 5000 U/ml, Streptomycin Sulphate 5 mg/ml) (JRH Biosciences). The plates were incubated at 37 °C and 5% CO_2 overnight.

At day 1, transfection was performed using calcium phosphate. Before transfection, the medium in each plate was replaced with 120 ml of fresh DMEM/F12 supplemented with 10% (v/v) FCS or DCS, 4 mM L-glutamine, 1% (v/v) Penicillin-Streptomycin. Calcium phosphate / DNA precipitate was prepared by adding 1200 µg of plasmid DNA harbouring the gene for the target molecule (as described above in Example 1) and 3000 µl of 2.5 M $CaCl_2$ in sterile 1× TE (10mM Tris and 1 mM EDTA pH 7.3 (Sigma)) to a final volume of 30 ml (solution A). Solution A was added drop-wise to 30 ml of 2× HEPES Buffered Saline (273.8 mM NaCl, 10 mM KCl, 11.1 mM (+)-D-Glucose, 41.96 mM HEPES free acid and 1.5 mM $Na_2HPO_4.7H_2O$ pH 7.06-7.12 (Sigma)) (solution B) with a 10 ml pipette. During the course of addition, bubbles were gently blown through solution B. The mixture was incubated at 25 °C for 20 minutes and vortexed. 12 ml of the mixture was added drop-wise to each plate. After 4 hours the medium containing the transfection mixture was removed and 100 ml of DMEM/F12 supplemented with 10% (v/v) FCS or DCS, 4 mM L-glutamine, 50 U/ml Penicillin G and 50 µg/ml Streptomycin Sulphate, and a final concentration of 3.5 mM HCl, with the medium having a final pH of 7, was added to each plate. The plates were incubated at 37 °C and 5% CO_2 overnight.

At day 2, the cell culture supernatant was discarded. The contents in the plates were washed twice with 50 ml of DMEM/F12 medium per plate and 100 ml of fresh serum-free DMEM/F12 medium supplemented with 40 mM N-acetyl-D-mannosamine (New Zealand Pharmaceuticals), 10 mM L-Glutamine (Amresco), 0.5 g/L Mannose (Sigma) and 1% (v/v) Penicillin-Streptomycin was added to each plate. The plates were incubated at 37 °C and 5% CO_2 overnight.

At day 3, the cell culture supernatant was collected and 100 ml fresh serum-free DMEM/F12 medium supplemented with 40 mM N-acetyl-D-mannosamine, 10 mM L-Glutamine, 4.1 g/L Mannose, and 1% (v/v) Penicillin-Streptomycin was added to each plate. The plates were incubated at 37 °C and 5% CO_2 overnight. 100 mM PMSF and 500 mM EDTA were added to the collected cell culture supernatant and the mixture was stored at 4 °C.

At day 4, the cell culture supernatant was collected. 100 mM PMSF and 500 mM EDTA was added to the collected cell culture supernatant and combined with the day 3 collection before particulate removal using a 0.45 micron low-protein binding filter (Durapore, Millipore). The mixture was either stored at 4 °C or used immediately. For long-term storage, the supernatant was kept at –70 °C.

(ii) Purification of SCF and SCFR-ECD of the Present Invention

One litre of filtered cell culture supernatants containing SCF or SCFR-ECD was concentrated 10 fold using a tangential flow filtration (TFF) device (Pelicon XL, Ultracell, Millipore). Sample was pumped at 150 ml/min across 150 cm^2 of regenerated cellulose membrane, with a nominal molecular weight cut-off of 5 KDa until the sample had concentrated down to a volume of 100 ml. Each concentrated sample was diafiltered by the addition of an equal volume of buffer 20 mM Tris pH 8.5 followed by another concentration down to 100 ml. This diafiltration step was repeated twice with a final concentration to 100 ml. The concentrated diafiltered sample was then filtered through a 0.45 micron low-protein binding filter (Durapore, Millipore).

Purification of SCF or SCFR-ECD was achieved by passing the concentrated cell culture supernatant from the TFF over an anion exchange column (SCF: Uno Q12 column; SCFR-ECD: MacroPrep Q beads, MT 20 column; both from Bio-Rad Laboratories) pre-equilibrated with 20 mM Tris pH 8.5. The bound SCF or SCFR-ECD was then eluted from the respective column with an optimised gradient from 20 mM Tris pH 8.5 to 20 mM Tris pH 8.5 containing 1 M NaCl. The resulting fractions were analysed for apparent molecular weight and level of purity by silver-stained SDS-PAGE using 4 – 20 % Tris-Glycine gels

(Invitrogen) and by Western analysis with anti-human SCF or anti-human SCFR-ECD antibodies (R & D Systems), respectively. The SCFR-ECD containing fractions were then pooled and concentrated to 1–2 ml using a centrifugal filter device (Amicon Ultra, Millipore or Vivaspin).

Reverse phase chromatography (RPC) was performed on the anion exchange fractions containing SCF using a Jupiter 5μ C5 300A 250 × 4.6 mm (Phenomenex) column pre-equilibrated with 0.15 %(v/v) trifluoroacetic acid (TFA). The bound SCF was eluted with a linear 10-column volume gradient from 0.15 %(v/v) TFA to 0.125 %(v/v) TFA containing 60 %(v/v) isopropanol. The run was carried out at a flow rate of 0.5 ml/min. The eluted fractions were assayed by silver-stained SDS-PAGE using 4 – 20 % Tris-Glycine gels (Invitrogen) and by anti-human SCF Western blot.

Size exclusion chromatography was performed on the concentrated anion exchange chromatography fractions containing SCFR-ECD using a Superdex 200 preparative grade 16/70 column (Pharmacia, Uppsala, Sweden), as described below in Example 2(v) and Table 9. Eluted fractions were assayed by silver-stained SDS-PAGE using 4 – 20% Tris-Glycine gels and by Western analysis using anti-human SCFR ECD antibody (R & D Systems). Apparent molecular weight and level of purity are recorded in Table 9.

(iii) Purification of SCFR-Fc of the Present Invention

Collected medium was adjusted to pH 8 by the addition of 2 M Tris-HCl pH 8 (Sigma) and filtered (Durapore, 0.45 μm, Millipore). One litre of pH-adjusted medium containing a target molecule of the present invention was passed by gravity flow over a Protein A Sepharose column (Pharmacia) with a 1-ml bed volume that had been pre-equilibrated to pH 8 with Dulbecco's phosphate buffered saline (DPBS) (SAFC Biosciences). After washing with 20 column volumes of column buffer (DPBS pH 8), the target molecule of the present invention was eluted in 1-ml fractions with 0.1 M Citric Acid (Sigma) at pH 4, then at pH 2.2. The fractions were immediately neutralised by the addition of 100 μl and 400 μl respectively of 2 M Tris-HCl pH 9 (Sigma). Fractions were analysed by silver-stained SDS-PAGE using 4 – 20 % gradient Tris-Glycine gels (Invitrogen). Pure fractions

containing the target molecule were pooled and concentrated to less than 1 ml for size exclusion chromatography using a centrifugal filter device (Amicon Ultra, Millipore).

Size exclusion chromatography (SEC) was performed on the concentrated samples using a Superdex 200 preparative grade 16/70 column (Pharmacia, Uppsala, Sweden). An isocratic flow of 1 %(w/v) ammonium bicarbonate was used at a flow rate of 1 ml/min. Total run time was approximately 120 min. The eluted fractions were assayed by silver-stained SDS-PAGE using 4 – 20 % Tris-Glycine gels (Invitrogen). Peak elution ranges and target molecule elution times are shown in Table 9. Fractions containing target molecules were pooled and concentrated to less than 2 ml using a centrifugal filter device (Amicon Ultra, Millipore).

The apparent MW and purity of purified target molecules, as determined by silver stained SDS PAGE, are listed in Table 9. Table 9 also contains final concentrations of purified target molecules, as determined by absorption at 280 nm using the relevant molar extinction co-efficient.

TABLE 9

Protein and protein-Fc purification data

Protein	SEC peak elution range (min)	Target molecule elution time (min)	Apparent MW (kDa)	% Purity	Final concentration (micrograms/ ml)	Molar Extinction Coefficient (M^{-1} cm^{-1})
SCF	-	-	15 to 40	>95	1000	10220
SCFR-ECD	50 to 80	65	70 to 110	>95	1900	69050
SCF R-Fc	40 to 110	50	100 to 150	>95	2226	104835

EXAMPLE 3

(i) Characterization of Target Molecules of the Present Invention by Two-Dimensional Polyacrylamide Electrophoresis

Each sample collected from Example 2 was buffer exchanged by dialysis or desalting column (Pharmacia HR 10/10 Fast Desalting Column) into repurified (18 MOhm) water and dried using a SpeedVac concentrator. Alternatively, samples underwent precipitation, for example, TCA precipitation, using methods known in the art. Dried sample was then re-dissolved into 240 microlitres MSD buffer (5M urea, 2M thiourea, 65mM DTT, 2% (w/v) CHAPS, 2% (w/v) sulfobetaine 3-10, 0.2% (v/v) carrier ampholytes, 40mM Tris, 0.002% (w/v) bromophenol blue, water). Samples were then reduced and alkylated by incubation in 5mM tributylphosphine, 1M acrylamide at room temperature for 1 hour then centrifuged at 15000g for 8 minutes.

Isoelectric focusing (IEF) was performed using either precast 11 cm or precast 17 cm gel pH 3-10 immobolised pH gradient IEF strips (BioRad or Amersham). The IEF strips were re-hydrated in the sample in a rehydration tray (Amersham) at room temperature until all the solution was taken up by the strips. The IEF strips were placed into the focusing chamber and covered with paraffin oil. IEF was performed (100 V for 1 hour, 200V for 1 hour, 600V for 2 hours, 1000 V for 2 hours, 2000 V for 2 hours, 3500 V for 12 hours and 100 V for up to 12 hours until the strips had reached approxiamtely 35 kV hours in the case of 11cm strips; 85kV hours in the case of 17cm strips (using the same V ramp up procedure)).

Following isoelectric focusing the strips were reduced and alkylated before being applied to a second dimension gel. The strips were incubated in 1M Tris/HCl pH 8.8, 6M urea, 2% (w/v) SDS, 2% (v/v) glycerol, 5mM tributylphosphine (TBP), 2.5% (v/v) acrylamide solution for at least 20 minutes.

The 11cm strips were separated on the second dimension by Criterion pre poured (11 x 8cm 1mm thick) 10-20% Tris HCl gradient gels (BioRad). 17cm strips were separated on

17 x 17 cm, 1.5mm thick, self poured 10-20% Tris glycine gradient gels. Precision or Kaleidoscope molecular weight markers (BioRad) were also applied to the gel.

The SDS-PAGE was run using either a Criterion or Protean II electrophoresis system (BioRad) (40mA per gel for approximately 1 hour (until the buffer front was about to run off the end of the gel) for 11 cm gels and 15mA constant current per gel for 21 hours for 17cm gels). The buffer used was 192 mM glycine, 0.1% (w/v) SDS, 24.8 mM Tris base at pH 8.3.

The completed second dimension gels were stained using Deep Purple fluorescent stain. The gels were fixed in a 7.5% acetic acid (v/v) in 10% methanol for a minimum of 1 hour before being incubated in 300mM Na2CO3, 35mM NaHCO3 for 2 x 30 minutes, then incubated in 1:200 dilution Deep Purple stain for at least 1 hour in the dark with agitation. The gels were then destained by 2 x 15 minute incubations in 10% MeOH, 7% HAc. The gel was imaged using a Typhoon imager (Amersham) or a LAS3000 imager (Fuji).

The software ImageJ (http://rsb.info.nih.gov/ij/) was used to analyse the relative intensities of protein spots on each gel. Densitometry was performed on the spots within a selected area of gel and a background subtraction was conducted using the appropriate region of gel lacking protein spots. A volume integration was performed on each protein spot of interest from which the centre of mass for the spot was calculated. Relative percentage intensities were calculated for each protein spot and by normalising the combined value of the intensities of all spots to 100%, the intensity of each protein spot relative to the other spots in the gel was determined.

The charge of the isoforms (pKa values) were determined by measuring the respective distance of the spots from the left side of each gel using ImageJ. Since the relationship between the pI values of the strip and the physical distance of the gel was linear, the pI values corresponding to the different pKa values of the isoform spots were readily determined.

The major protein spots in the resulting gels correspond to isoforms of molecules of the present invention. The low intensity spots may represent a molecule of the present invention or low level contaminants, however, these cannot be confirmed by PMF due to the low intensity. Tables 10 to 13 show key properties of these isoforms for paricular molecules of the present invention: the pI values (± 1.0), the apparent molecular weights (± 20%), and the relative intensities (± 20% of the actual value or ± 2% of the total, whichever is larger). The values listed correspond to the intensity weighted center within the selected area of gel containing the spot and hence, are only reflective of the pI and molecular weight of the protein at one particular reading within the selected area of the gel. Taking into consideration the inherent variability of size and position of protein spots within 2D gels, the apparent molecular weights, the pI values and the number of isoforms for the molecules of the present invention were determined (Table 14).

TABLE 10

Molecular weights and pI values of isoforms of SCF

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
2	3.34	±1.00	46.77	±9.35	0.22	±2.00
3	3.31	±1.00	40.84	±8.17	0.78	±2.00
4	3.45	±1.00	40.60	±8.12	0.72	±2.00
5	3.66	±1.00	40.09	±8.02	0.84	±2.00
6	4.02	±1.00	39.74	±7.95	0.91	±2.00
7	4.24	±1.00	39.62	±7.92	2.66	±2.00
8	4.40	±1.00	38.78	±7.76	1.44	±2.00
9	4.57	±1.00	38.44	±7.69	2.65	±2.00
10	4.71	±1.00	38.32	±7.66	1.77	±2.00
11	4.83	±1.00	38.21	±7.64	1.19	±2.00
12	4.97	±1.00	38.00	±7.60	1.25	±2.00
13	3.38	±1.00	35.78	±7.16	0.46	±2.00
14	3.67	±1.00	35.32	±7.06	0.54	±2.00

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
15	4.02	±1.00	34.68	±6.94	1.12	±2.00
16	4.23	±1.00	33.98	±6.80	3.98	±2.00
17	4.39	±1.00	33.57	±6.71	3.49	±2.00
18	4.57	±1.00	32.69	±6.54	9.51	±2.00
19	4.71	±1.00	32.43	±6.49	6.75	±2.00
20	4.83	±1.00	32.34	±6.47	5.67	±2.00
21	4.95	±1.00	32.17	±6.43	5.48	±2.00
22	5.06	±1.00	31.99	±6.40	4.01	±2.00
23	5.19	±1.00	31.65	±6.33	5.00	±2.00
24	5.32	±1.00	31.24	±6.25	3.86	±2.00
25	5.43	±1.00	31.11	±6.22	3.73	±2.00
26	5.54	±1.00	30.84	±6.17	2.40	±2.00
27	4.57	±1.00	28.35	±5.67	0.48	±2.00
28	4.72	±1.00	27.89	±5.58	1.07	±2.00
29	4.85	±1.00	27.40	±5.48	1.05	±2.00
30	4.96	±1.00	27.24	±5.45	1.51	±2.00
31	5.08	±1.00	27.02	±5.40	2.34	±2.00
32	5.20	±1.00	26.77	±5.35	2.78	±2.00
33	5.32	±1.00	26.39	±5.28	2.92	±2.00
34	5.43	±1.00	26.29	±5.26	2.43	±2.00
35	5.52	±1.00	26.33	±5.27	1.88	±2.00
36	5.61	±1.00	25.96	±5.19	0.98	±2.00
37	5.75	±1.00	25.85	±5.17	1.36	±2.00
38	5.87	±1.00	25.97	±5.19	0.39	±2.00
39	6.08	±1.00	25.89	±5.18	0.14	±2.00
40	5.58	±1.00	23.37	±4.67	0.77	±2.00
41	5.75	±1.00	23.38	±4.68	1.04	±2.00
42	5.87	±1.00	23.33	±4.67	0.22	±2.00

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
43	6.08	±1.00	23.25	±4.65	0.09	±2.00
44	5.58	±1.00	21.09	±4.22	0.20	±2.00
45	5.73	±1.00	21.11	±4.22	0.66	±2.00
46	5.85	±1.00	21.23	±4.25	0.13	±2.00
47	6.08	±1.00	21.06	±4.21	0.06	±2.00
48	5.59	±1.00	18.83	±3.77	0.04	±2.00
49	5.74	±1.00	18.84	±3.77	0.16	±2.00
50	6.09	±1.00	18.57	±3.71	0.06	±2.00
51	4.30	±1.00	71.19	±14.24	0.05	±2.00
52	4.62	±1.00	68.76	±13.75	0.30	±2.00
53	4.74	±1.00	67.65	±13.53	0.55	±2.00
54	4.86	±1.00	65.95	±13.19	0.39	±2.00
55	4.98	±1.00	63.81	±12.76	0.67	±2.00
56	5.26	±1.00	60.08	±12.02	1.00	±2.00
57	5.39	±1.00	58.07	±11.61	1.47	±2.00
58	5.49	±1.00	56.35	±11.27	1.02	±2.00
59	5.59	±1.00	55.30	±11.06	0.88	±2.00
60	6.11	±1.00	48.75	±9.75	0.52	±2.00

TABLE 11

Molecular weights and pI values of isoforms of SCF

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
2	3.72	±1.00	48.74	±9.75	0.12	±2.00
3	4.03	±1.00	48.48	±9.70	0.27	±2.00
4	4.26	±1.00	48.26	±9.65	0.14	±2.00

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
5	4.36	±1.00	48.22	±9.64	0.27	±2.00
6	4.49	±1.00	48.20	±9.64	0.51	±2.00
7	4.63	±1.00	47.99	±9.60	0.18	±2.00
8	4.71	±1.00	48.20	±9.64	0.17	±2.00
9	3.71	±1.00	41.78	±8.36	0.38	±2.00
10	4.03	±1.00	41.65	±8.33	0.89	±2.00
11	4.27	±1.00	41.39	±8.28	0.91	±2.00
12	4.38	±1.00	41.31	±8.26	1.38	±2.00
13	4.51	±1.00	41.29	±8.26	2.22	±2.00
14	4.62	±1.00	41.17	±8.23	1.19	±2.00
15	4.70	±1.00	41.13	±8.23	1.36	±2.00
16	4.80	±1.00	41.06	±8.21	1.48	±2.00
17	4.89	±1.00	41.02	±8.20	1.15	±2.00
18	4.97	±1.00	40.87	±8.17	1.39	±2.00
19	5.06	±1.00	40.86	±8.17	1.35	±2.00
20	5.17	±1.00	40.68	±8.14	1.51	±2.00
21	5.30	±1.00	40.32	±8.06	1.52	±2.00
22	5.41	±1.00	40.15	±8.03	1.06	±2.00
23	5.55	±1.00	39.87	±7.97	0.96	±2.00
24	5.72	±1.00	39.28	±7.86	0.26	±2.00
25	3.71	±1.00	35.92	±7.18	0.04	±2.00
26	4.04	±1.00	35.31	±7.06	0.33	±2.00
27	4.28	±1.00	35.33	±7.07	0.48	±2.00
28	4.39	±1.00	35.06	±7.01	0.79	±2.00
29	4.51	±1.00	34.63	±6.93	1.93	±2.00
30	4.63	±1.00	34.37	±6.87	1.64	±2.00
31	4.71	±1.00	34.09	±6.82	1.91	±2.00
32	4.80	±1.00	33.83	±6.77	2.47	±2.00

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
33	4.88	±1.00	33.44	±6.69	2.36	±2.00
34	4.97	±1.00	33.14	±6.63	3.25	±2.00
35	5.06	±1.00	33.01	±6.60	2.85	±2.00
36	5.17	±1.00	32.68	±6.54	3.98	±2.00
37	5.30	±1.00	32.38	±6.48	3.92	±2.00
38	5.42	±1.00	32.20	±6.44	3.66	±2.00
39	5.57	±1.00	32.04	±6.41	4.05	±2.00
40	5.71	±1.00	31.89	±6.38	2.20	±2.00
41	5.84	±1.00	32.01	±6.40	2.32	±2.00
42	5.99	±1.00	32.45	±6.49	2.01	±2.00
43	6.18	±1.00	32.33	±6.47	1.71	±2.00
44	6.31	±1.00	32.00	±6.40	0.54	±2.00
45	6.44	±1.00	32.64	±6.53	0.42	±2.00
46	6.66	±1.00	32.20	±6.44	1.22	±2.00
47	3.72	±1.00	30.03	±6.01	0.09	±2.00
48	4.03	±1.00	29.52	±5.90	0.37	±2.00
49	4.26	±1.00	29.15	±5.83	0.26	±2.00
50	4.37	±1.00	28.74	±5.75	0.34	±2.00
51	4.50	±1.00	28.19	±5.64	0.62	±2.00
52	4.62	±1.00	27.81	±5.56	0.54	±2.00
53	4.71	±1.00	27.60	±5.52	0.52	±2.00
54	4.79	±1.00	27.29	±5.46	0.78	±2.00
55	4.88	±1.00	26.39	±5.28	0.74	±2.00
56	4.97	±1.00	26.30	±5.26	1.20	±2.00
57	5.06	±1.00	26.00	±5.20	1.32	±2.00
58	5.16	±1.00	25.63	±5.13	1.87	±2.00
59	5.29	±1.00	25.42	±5.08	2.62	±2.00
60	5.41	±1.00	25.19	±5.04	2.23	±2.00

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
61	5.55	±1.00	24.91	±4.98	3.24	±2.00
62	5.70	±1.00	24.67	±4.93	2.14	±2.00
63	5.83	±1.00	24.42	±4.88	2.23	±2.00
64	5.99	±1.00	24.16	±4.83	2.40	±2.00
65	6.20	±1.00	24.18	±4.84	1.50	±2.00
66	6.33	±1.00	23.70	±4.74	1.00	±2.00
67	6.48	±1.00	24.75	±4.95	0.26	±2.00
68	6.71	±1.00	23.98	±4.80	1.93	±2.00
69	4.95	±1.00	20.99	±4.20	0.16	±2.00
70	5.05	±1.00	20.73	±4.15	0.21	±2.00
71	5.16	±1.00	20.57	±4.11	0.30	±2.00
72	5.27	±1.00	20.36	±4.07	0.42	±2.00
73	5.40	±1.00	20.42	±4.08	0.45	±2.00
74	5.55	±1.00	20.06	±4.01	0.64	±2.00
75	5.70	±1.00	19.79	±3.96	0.35	±2.00
76	5.84	±1.00	19.37	±3.87	0.57	±2.00
77	6.00	±1.00	19.07	±3.81	0.62	±2.00
78	6.25	±1.00	19.00	±3.80	0.18	±2.00
79	6.35	±1.00	18.91	±3.78	0.32	±2.00
80	6.73	±1.00	18.80	±3.76	0.39	±2.00
81	4.91	±1.00	63.46	±12.69	0.13	±2.00
82	4.98	±1.00	62.69	±12.54	0.31	±2.00
83	5.07	±1.00	61.32	±12.26	0.43	±2.00
84	5.19	±1.00	59.81	±11.96	0.53	±2.00
85	5.32	±1.00	58.14	±11.63	0.64	±2.00
86	5.43	±1.00	57.19	±11.44	0.28	±2.00
87	5.55	±1.00	55.81	±11.16	0.15	±2.00

TABLE 12

Molecular weights and pI values of isoforms of SCFR-ECD

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
2	4.44	±1.00	127.39	±25.48	0.04	±2.00
3	4.48	±1.00	126.31	±25.26	0.07	±2.00
4	4.53	±1.00	124.53	±24.91	0.13	±2.00
5	4.59	±1.00	122.72	±24.54	0.21	±2.00
6	4.64	±1.00	120.64	±24.13	0.26	±2.00
7	4.69	±1.00	118.69	±23.74	0.35	±2.00
8	4.75	±1.00	116.61	±23.32	0.42	±2.00
9	4.81	±1.00	114.84	±22.97	0.61	±2.00
10	4.86	±1.00	112.64	±22.53	0.64	±2.00
11	4.92	±1.00	112.48	±22.50	0.66	±2.00
12	4.97	±1.00	110.71	±22.14	0.88	±2.00
13	5.04	±1.00	110.25	±22.05	0.94	±2.00
14	5.10	±1.00	108.85	±21.77	0.85	±2.00
15	5.17	±1.00	108.17	±21.63	1.13	±2.00
16	5.24	±1.00	106.31	±21.26	1.06	±2.00
17	5.31	±1.00	104.97	±20.99	1.01	±2.00
18	5.38	±1.00	102.96	±20.59	1.26	±2.00
19	5.45	±1.00	101.89	±20.38	1.09	±2.00
20	5.52	±1.00	100.66	±20.13	1.28	±2.00
21	5.61	±1.00	99.83	±19.97	1.25	±2.00
22	5.69	±1.00	97.83	±19.57	1.08	±2.00
23	5.77	±1.00	95.92	±19.18	1.24	±2.00
24	5.86	±1.00	95.04	±19.01	0.97	±2.00
25	5.94	±1.00	93.18	±18.64	1.12	±2.00
26	6.04	±1.00	89.50	±17.90	1.19	±2.00

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
27	6.13	±1.00	87.86	±17.57	0.77	±2.00
28	6.23	±1.00	85.21	±17.04	0.48	±2.00
29	6.33	±1.00	83.41	±16.68	0.28	±2.00
30	6.50	±1.00	86.52	±17.30	4.02	±2.00
31	6.62	±1.00	82.99	±16.60	0.75	±2.00
32	4.34	±1.00	107.61	±21.52	0.08	±2.00
33	4.39	±1.00	107.22	±21.44	0.17	±2.00
34	4.44	±1.00	105.99	±21.20	0.30	±2.00
35	4.49	±1.00	104.24	±20.85	0.38	±2.00
36	4.53	±1.00	102.84	±20.57	0.52	±2.00
37	4.59	±1.00	101.00	±20.20	0.83	±2.00
38	4.64	±1.00	98.84	±19.77	0.97	±2.00
39	4.69	±1.00	97.06	±19.41	1.21	±2.00
40	4.75	±1.00	95.44	±19.09	1.41	±2.00
41	4.81	±1.00	93.61	±18.72	2.03	±2.00
42	4.86	±1.00	91.80	±18.36	1.91	±2.00
43	4.92	±1.00	90.88	±18.18	2.19	±2.00
44	4.97	±1.00	89.50	±17.90	2.86	±2.00
45	5.04	±1.00	88.24	±17.65	3.13	±2.00
46	5.10	±1.00	87.07	±17.41	3.05	±2.00
47	5.17	±1.00	86.17	±17.23	4.09	±2.00
48	5.24	±1.00	84.98	±17.00	3.88	±2.00
49	5.31	±1.00	83.84	±16.77	3.62	±2.00
50	5.38	±1.00	82.75	±16.55	4.18	±2.00
51	5.45	±1.00	82.19	±16.44	3.85	±2.00
52	5.52	±1.00	81.57	±16.31	4.33	±2.00
53	5.61	±1.00	81.12	±16.22	4.45	±2.00
54	5.69	±1.00	80.23	±16.05	3.64	±2.00

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
55	5.78	±1.00	79.37	±15.87	3.75	±2.00
56	5.86	±1.00	78.46	±15.69	3.26	±2.00
57	5.95	±1.00	77.31	±15.46	2.87	±2.00
58	6.04	±1.00	75.80	±15.16	2.76	±2.00
59	6.13	±1.00	75.03	±15.01	2.19	±2.00
60	6.23	±1.00	74.26	±14.85	2.11	±2.00
61	6.34	±1.00	73.45	±14.69	1.87	±2.00
62	6.43	±1.00	73.11	±14.62	0.98	±2.00
63	6.57	±1.00	72.42	±14.48	0.68	±2.00
64	6.70	±1.00	81.43	±16.29	0.44	±2.00

TABLE 13

Molecular weights and pI values of isoforms of SCFR-Fc

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
2	5.58	±1.00	156.24	±31.25	0.14	±2.00
3	5.62	±1.00	153.79	±30.76	0.28	±2.00
4	5.66	±1.00	152.51	±30.50	0.36	±2.00
5	5.71	±1.00	150.11	±30.02	0.63	±2.00
6	5.77	±1.00	147.97	±29.59	0.68	±2.00
7	5.82	±1.00	147.77	±29.55	0.83	±2.00
8	5.87	±1.00	143.94	±28.79	0.80	±2.00
9	5.93	±1.00	142.98	±28.60	1.08	±2.00
10	5.99	±1.00	141.12	±28.22	1.08	±2.00
11	6.04	±1.00	140.90	±28.18	1.10	±2.00

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
12	6.10	±1.00	139.21	±27.84	1.24	±2.00
13	6.17	±1.00	139.12	±27.82	1.23	±2.00
14	6.23	±1.00	138.19	±27.64	1.15	±2.00
15	6.29	±1.00	138.14	±27.63	1.39	±2.00
16	6.36	±1.00	135.84	±27.17	1.13	±2.00
17	6.43	±1.00	134.34	±26.87	1.10	±2.00
18	6.54	±1.00	129.88	±25.98	1.62	±2.00
19	6.64	±1.00	126.34	±25.27	1.50	±2.00
20	6.72	±1.00	123.35	±24.67	1.37	±2.00
21	6.81	±1.00	121.09	±24.22	0.84	±2.00
22	5.37	±1.00	131.88	±26.38	0.34	±2.00
23	5.43	±1.00	130.34	±26.07	0.52	±2.00
24	5.47	±1.00	128.93	±25.79	0.64	±2.00
25	5.52	±1.00	125.82	±25.16	0.91	±2.00
26	5.57	±1.00	125.03	±25.01	1.46	±2.00
27	5.62	±1.00	124.24	±24.85	1.38	±2.00
28	5.67	±1.00	122.16	±24.43	1.87	±2.00
29	5.72	±1.00	121.84	±24.37	1.93	±2.00
30	5.77	±1.00	119.93	±23.99	2.73	±2.00
31	5.83	±1.00	119.23	±23.85	2.76	±2.00
32	5.88	±1.00	117.42	±23.48	3.19	±2.00
33	5.94	±1.00	116.03	±23.21	3.39	±2.00
34	5.99	±1.00	115.29	±23.06	3.68	±2.00
35	6.05	±1.00	113.96	±22.79	3.86	±2.00
36	6.11	±1.00	113.68	±22.74	4.36	±2.00
37	6.17	±1.00	112.91	±22.58	4.62	±2.00
38	6.24	±1.00	112.31	±22.46	4.57	±2.00
39	6.30	±1.00	112.13	±22.43	5.10	±2.00

Spot No	Isoelectric Point (pI)		Molecular Weight (kDa)		Relative Intensity (%) (Normalized Value)	
		Range		Range		Range
40	6.37	±1.00	110.78	±22.16	4.69	±2.00
41	6.43	±1.00	110.37	±22.07	4.87	±2.00
42	6.50	±1.00	108.45	±21.69	4.70	±2.00
43	6.57	±1.00	107.02	±21.40	5.43	±2.00
44	6.65	±1.00	104.95	±20.99	5.32	±2.00
45	6.73	±1.00	104.88	±20.98	4.60	±2.00
46	6.81	±1.00	104.14	±20.83	2.41	±2.00
47	6.90	±1.00	106.53	±21.31	1.11	±2.00

TABLE 14

Molecular weights, pI and number of isoforms of target molecules of the present invention

Molecule	MW (kDa)	pI	Number of isoforms
SCF	18 to 55	3 to 7.5	15 to 86
SCFR-ECD	55 to 150	3.5 to 7.5	15 to 63
SCFR-Fc	80 to 180	4.0 to 8.0	10 to 46

(ii) One-Dimensional Polyacrylamide Electrophoresis

Each collected sample from Example 2 was dried and then re-solubilised into 60ml of 1D sample buffer (10% glycerol, 0.1% SDS, 10mM DTT, 63mM Tris-HCl; i.e. under reduced conditions) and heated at 100°C for 5 minutes. For PNGaseF treatment, a 30 µL aliquot of the sample was taken and NP40 added to a final concentration of 0.5 %. 5 µL of PNGaseF was added and the sample was incubated at 37 °C for 3 hours. For glycosidase cocktail treatment of the sample, an aliquot was taken and NP40 was added to a final concentration of 0.5%. 1µL of PNGase F, and 1 µL each of Sialidase A (neuramidase), O-Glycanase, beta (1-4)-Galactosidase and beta N-Acetylglucosaminidase was added. Treated and untreated samples were incubated at 37 °C for 3 hours. Treated and untreated samples were

run on a pre-cast Tris gel, for example, a Tris 4-20% gradient gel (BioRad) or Tris HCl gradient gel (Invitrogen). Precision molecular weight markers (BioRad catalogue number 161-0363) were also applied to the gel. Criterion 4-20% or 18% gels were used for 1D SDS-PAGE (BioRad catalogue numbers: 345-0033 or 345-0024). The SDS-PAGE was run using either a Mini Protean II or a Criterion electrophoresis system (BioRad) at 200 V for approximately 1 h or until the buffer front was about to run off the end of the gel. The buffer used was 192 mM glycine, 0.1% (w/v) SDS, 24.8 mM Tris base at pH 8.3. The completed gels were fixed for at least 30 minutes in 10% MeOH and 7% HAc. The gel was then stained using Sypro Ruby gel stain (BioRad) for at least 3 h and destained with 10% MeOH and 7% HAc for at least 30 minutes. Alternatively the gels were stained using Deep Purple (Amersham) as per the manufacturers instructions. The gel was imaged using a FX laser densitometer (BioRad) and the appropriate filter.

The apparent molecular weights of particular molecules of the present invention (as observed by SDS-PAGE) in the untreated state, following the release of N-linked oligosaccharides (by PNGase treatment) and following the release of N-linked oligosaccharides (by PNGase treatment) and O-linked oligosaccharides (by glycosidase cocktail) are listed in Table 15.

TABLE 15

Molecular weights of isoforms of target molecules of the present invention following 1D gel analysis

Molecule	MW (kDa) following the release of N-linked oligosaccharides	MW (kDa) following the release of N-linked and O-linked oligosaccharides
SCF	18 to 40	18 to 35
SCFR-ECD	42 to 85	42 to 75
SCFR-Fc	65 to 135	65 to 120

(iii) N-Terminal Sequencing

Protein bands are cut from gels prepared above (either two-dimensional gels or one-dimensional gels), placed into a 0.5ml tube and 100ml extraction buffer added (100mM Sodium acetate, 0.1%SDS, 50mM DTT pH 5.5). The gel slices are incubated at 37°C for 16 hours with shaking. The supernatant is applied to a ProSorb membrane (ABI) as per the manufacturers instruction and sequenced using an automated 494 Protein Sequencer (Applied Biosystems) as per the manufacturers instructions. The sequence generated is used to confirm the identity of a human target molecule of the present invention.

(iv) Peptide Mass Fingerprinting

Protein bands were cut from the gel prepared above (either from a two-dimensional gel or a one-dimensional gel) and washed with 25ml of wash buffer (50% acetonitrile in 50mM NH4HCO3). The gel pieces were left at room temperature for at least 1 hour and dried by vacuum centrifugation for 30 minutes. The gel pieces and 12ml of trypsin solution (20mg trypsin, 1200ml NH4HCO3) was placed in each sample well and incubated at 4°C for 1 hour. The remaining trypsin solution was removed and 20ml 50mM NH4HCO3 was added. The mixture was incubated overnight at 37°C with gentle shaking. The peptide samples were concentrated and desalted using C18 Zip-Tips (Millipore, Bedford, MA) or pre-fabricated micro-columns containing Poros R2 (Perseptive Biosystems, Framingham, MA) chromatography resin. Bound peptides were eluted in 0.8 µl of matrix solution (cyano-4-hydroxy cinnamic acid (Sigma), 8 mg/ml in 70% acetonitrile / 1% formic acid) directly onto a target plate. Peptide mass fingerprints of tryptic peptides were generated by matrix-assisted laser desorption / ionisation time-of-flight mass spectrometry (MALDI-TOF MS) using a Perseptive Biosystems Voyager DE-STR. Spectra were obtained in reflectron mode using an accelerating voltage of 20 kV. Mass calibration was performed using trypsin autolysis peaks, 2211.11 Da and 842.51 Da as internal standards. Data generated from peptide mass fingerprinting (PMF) was used to confirm the identity of the protein. Searches (primarily of Homo sapien (Human) and mammalian entries) were performed in databases such the SWISS-PROT and TrEMBL, via the program PeptIdent (www.expasy.ch/tools/peptident.html). Identification parametres included peptide mass

tolerance of 0.1Da, a maximum of one missed tryptic cleavage per peptide, and the methionine sulfoxide and cysteine-acrylamide modifications. Identifications were based on the number of matching peptide masses and the total percentage of the amino acid sequence that those peptides covered, in comparison to other database entries. Generally, a peptide match with at least 30% total sequence coverage was required for confidence in identification, but very low and high mass proteins, and those resulting from protein fragmentation, may not always meet this criterion, therefore requiring further identification.

Where inconclusive or no protein identification could be obtained from MALDI-TOF PMF analysis, the remaining peptide mixture or the identical spot cut from a replicate gel was subjected to tryptic digest and analysed by electrospray ionization tandem MS (ESI-MS/MS). For ESI-MS/MS, peptides were eluted from Poros R2 micro-columns in 1-2 microlitres of 70% acetonitrile, 1% formic acid directly into borosilicate nanoelectrospray needles (Micromass, Manchester, UK). Tandem MS was performed using a Q-Tof hybrid quadrupole / orthogonal-acceleration TOF mass spectrometer (Micromass). Nanoelectrospray needles containing the sample were mounted in the source and stable flow obtained using capillary voltages of 900-1200V. Precursor ion scans were performed to detect mass to charge ratio (m/z) values for peptides within the mixture. The m/z of each individual precursor ion was selected for fragmentation and collided with argon gas using collision energies of 18-30eV. Fragment ions (corresponding to the loss of amino acids from the precursor peptide) were recorded and processed using MassLynx Version 3.4 (Micromass). Amino acid sequences were deduced by the mass differences between y- or b-ion 'ladder' series using the program MassSeq (Micromass) and confirmed by manual interpretation. Peptide sequences were then used to search the NCBI and TrEMBL databases using the program BLASTP "short nearly exact matches". A minimum of two matching peptides were required to provide confidence in a given identification.

The identity of the respective gel spots were confirmed to be human SCF, SCFR-ECD and SCFR-Fc.

Further, an observed 1Da shift in the masses of tryptic peptides indicated the asparagine residues (N) of NX(S/T/C) motifs found in the theoretical amino acid sequence of target molecules of the present invention were modified to aspartic acid (D), consistent with the known ability of PNGase F to induce an N to D residue modification upon removal of associated N-linked oligosaccharides. This observation allowed for the identification of confirmed sites of N-glycosylation of the following target molecules of the present invention (Table 16).

TABLE 16

Confirmed sites of N-glycosylation of target molecules of the present invention

Molecule	Sites of N-glycosylation (numbered from the start of the signal sequence).
SCF	N-145
SCFR-ECD	N-130, N-367, N-463, N-486
SCFR-Fc	N-130, N-617

EXAMPLE 4

Analysis of Amino Acid, Monosaccharide, Oligosaccharide, Phosphate, Sulfate and Isoform Composition of Target Molecules of the Present Invention

(a) Preparation of Samples for Amino Acid, Monosaccharide, Oligosaccharide, Phosphate, Sulfate and Isoform Analysis

For characterisation of monosaccharide and oligosaccharide glycosylation and phosphate and sulfate post-translational modifications, the saccharides of the target molecule are first removed from the polypeptide backbone by hydrolytic or enzymatic means. The sample buffer components are also removed and exchanged with water to avoid inhibition of the hydrolysis and enzymatic reactions before analysis began. A solution of purified target molecule in PBS is dialysed extensively against 4 litres of deionised ultrafiltered water (18 MOhm) for four days with two changes per day using a regenerated cellulose dialysis

membrane (Spectrapore) with a nominal molecular weight cut-off (NMWC) of 5 KDa. After dialysis the solution is dried using a Savant Speed Vac (New York, USA). The dried down sample is then resuspended in 2 ml of deionised ultrafiltered water (18 MOhm) and divided into aliquots for the various analyses.

(b) Analysis of Amino Acid Composition by the Gas Phase Hydrolysis Method

Amino acids in the samples are analysed using precolumn derivatisation with 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate (AQC). The stable fluorescent amino acid derivatives are separated and quantified by reversed phise (C18) HPLC. The procedure employed is based on the Waters AccQTag amino acid analysis methodology.

Three samples of the target molecule preparation are taken and dried in a Speed Vac. The dried samples are then hydrolysed for 24 hours at 110°C. After hydrolysis the samples are dried again before derivatisation as follows. The dried samples are re-dissolved in 10 μL of an internal amino acid standard solution (α-aminobutyric acid, AABA), 35 μL of borate buffer is added followed by 15 μL of AQC derivatising reagent. The reaction mix is heated at 50°C for 12 minutes in a heating block. The derivatised amino acid sample is transferred to the autosampler of a HPLC system consisting of a Waters Alliance 2695 Separation Module, a Waters 474 Fluorescence Detector and a Waters 2487 Dual λ Absorbance Detector in series. The control and analysis software is Waters Empower Pro Module (Waters Corporation, Milford. MA, USA). The samples are passed over a Waters AccQTag column (15cm x 3.9mm ID) using chromatographic parameters (i.e. suitable eluents and gradient flows) known in the art.

(c) Analysis of Neutral and Amino Monosaccharide composition

Two samples of the target molecule preparation are taken and treated in two different ways to liberate monosaccharides. Each treatment, as described below, is performed in triplicate.

1. Hydrolysed with 2 M trifluroacetic acid (TFA) heated to 100 degrees C for four hours to release neutral sugars (galactose, glucose, fucose and mannose).

2. Hydrolysed with 4 M HCl heated to 100 degrees C for four hours to release amino sugars (N-acetyl-galactosamine, N-acetyl-glucosamine).

All of the hydrolysates are lyophilised using a Speed Vac system, redissolved in 200 µl water containing 0.8 nmols of internal standard. For neutral and amino sugars the internal standard is 2-deoxy-glucose. The samples are then centrifuged at 10,000 g for 30 minutes to remove protein debris. The supernatant is transferred to a fresh tube and analysed by high pH anion exchange chromatography using a Dionex LC 50 system with a GP50 pump and an ED50 pulsed amperometric detector (Dionex Ltd). Analysis of neutral and amino sugars is performed using a Dionex CarboPac PA-20 column. Elution is performed with an isocratic hydroxide concentration of 10 mM over 20 minutes. This is achieved with the Dionex EG50 eluent generation system.

(d) Analysis of Acidic Monosaccharide Composition

A sample of the target molecule preparation is taken and treated in the following way to liberate sialic acid monosaccharides. The treatment is performed in triplicate. The sample is hydrolysed with 0.1 M TFA at 80 degrees C for 40 minutes to release N-Acetyl and N-Glycolyl neuraminic acid. The hydrolysates are lyophilised using a Speed Vac, redissolved in 200 microlitres water containing 0.8 nmols of internal standard. For sialic acid analysis the internal standard is lactobionic acid. Samples are then centrifuged at 10,000 g for 30 minutes. The supernatant is transferred to a fresh tube and analysed by high pH anion exchange chromatography using a Dionex LC 50 system with a GP50 pump and an ED50 pulsed amperometric detector. Analysis of sialic acids is performed using a Dionex CarboPac PA1 using using chromatographic parameters (i.e. suitable eluents and gradient flows) known in the art.

(e) Analysis of Oligosaccharide Composition

For analysis of oligosaccharide composition two samples of the target molecule preparation are taken in triplicate and treated in one of the following ways:

1. Release of N-linked oligosaccharides is achieved with the enzyme Peptide-N4-(N-acetyl-beta-D-glucosaminyl) Asparagine Amidise (PNGase). First, a 1/5th volume of denaturation solution (2 % SDS (Sigma)/1 M β-mercaptoethanol (Sigma)) is added to the sample. The sample is heated to 100 °C for 5 minutes. A 1/10th volume of 15 % Triton-X100 (Sigma) is added to the sample. The sample is mixed gently and allowed to cool to room temperature. 25 Units of PNGase (Sigma) is added and incubated overnight at 37°C.

2. Release of O-linked oligosaccharides is achieved by the process of β-elimination. First, a 1/2 volume of 4M sodium borohydride (freshly made) (Sigma) solution is added to the sample. A half volume of 0.4 M NaOH (BDH, HPLC grade) is added to the sample. The sample is incubated at 50° C for 16 hours. The sample is cooled on ice and a half volume of 0.4 M acetic acid (Sigma) is added to the sample.

Both the N-linked and O-linked samples are further processed to remove buffer components using a Carbo Pac graphitised carbon SPE column. The column equilibration and elution conditions are as follows:

Firstly, the column is pre-equilibrated with 1 column volume of 80 % acetonitrile (Sigma) followed by two column volumes of H_2O. The sample is loaded under gravity flow and the column washed with two column volumes of H_2O. To elute neutral oligosaccharides 2 ml of 50 % acetonitrile is applied to the column. To elute acidic oligosaccharides 2 ml of 50 % acetonitrile/0.1% formic acid is applied to the column. Any remaining oligosaccharides are eluted by the addition of 2 ml of 80 % acetonitrile/0.1 % formic acid. Individual fractions from the SPE columns containing the neutral or acidic N-linked oligosaccharides and the neutral or acidic O-linked oligosaccharides are dried down to completion using a Speed Vac. The samples are redissolved in 200 microlitres water and analysed by high pH anion exchange chromatography using a Dionex LC 20 system with a GP50 pump and an ED50 pulsed amperometric detector. Analysis of neutral and acidic oligosaccharides is performed using a CarboPac PA100 column and chromatographic parameters (i.e. suitable eluents and gradient flows) known in the art.

(f) Analysis of Sulfate and Phosphate Composition.

Sulfate/phosphate analysis is performed essentially by the method described by Harrison and Packer (Harrison and Packer *Methods Mol Biol. 125*:211-216, 2000). A sample of the target molecule preparation is taken for sulfate/phosphate analysis and hydrolysed in 4 M HCl at 100 degrees C for four hours. The HCl is removed by drying the samples in a Speed Vac system. Samples are then redissolved into 200 µl H_2O. A suitable volume of sample is injected onto a Dionex LC 50 system with a GP50 pump and a ED50 conductivity detector. Separation is performed by a Dionex IonPac IS11 Anion exchange column using chromatographic parameters (i.e. suitable eluents and gradient flows) known in the art.

(g) Further Separation of Protein Isoforms

Further separation of target molecule isoforms is performed using a pellicular anion exchange column. A suitable volume of sample, for example, 24 microliters, is separated through a ProPac SAX-10 column (Dionex Ltd) using a Dionex SUMMIT system with UV-Vis detector (Dionex Ltd). Separation is performed using suitable eluents and gradients known in the art. Target molecule isoforms are found to elute in a pattern of distinct peaks.

EXAMPLE 5

Glyco Mass Fingerprinting

A target molecule of the present invention is separated using 2D gel electrophoretic techniques as in Example 3 and blotted onto polyvinyl difluorethane (PVDF) membrane. The spots are stained using one of a standard array of protein stains (Colloidal Coomassie Blue, Sypro Ruby or Deep Purple), and the isoform relative amounts quantified using densitometry algorithms. The individual spots are excised and treated with an array of deglycosylating enzymes and/or chemical means, as appropriate, to remove the oligosaccharides present according to methods described in this document. Once the oligosaccharides are removed, they are separated and analysed on a liquid chromatography-electrospray mass spectrometry system (LC-MS) using a graphitised

carbon column and organic solvent (MeCN) gradient elution system. The generated peak profile that is generated is a "fingerprint" of the oligosaccharides present on the isoform. Furthermore, the mass spectrometry system simultaneously generates information on the mass of each of the sugars present in the sample which is used to identify their structure through pattern matching with the GlycoSuite database. In addition, individual mass peaks can be fragmented multiple times to give MS^n spectra. These fragments allow structural prediction using methods known in the art, for example, by the use of the GlycosidIQ software package.

The above separation, deglycosylation and analysis procedures are repeated using a corresponding target molecule expressed in a non-human cell system, e.g. *E. coli,* yeast or CHO cells and the respective glyco mass fingerprints are found to be significantly different. At a structural level, such a result indicates different patterns of glycan structures present on the target molecule of the present invention and the corresponding non-human cell expressed target molecule.

EXAMPLE 6

Fluorophore Assisted Carbohydrate Electrophoresis

Oligosaccharide profiles of the target molecule are derived using the fluorophore assisted carbohydrate electrophoresis protocols (FACE protocols). The oligosaccharides from the target cytokine are hydrolysed from the amino acid backbone using ammonium hydroxide and subsequently labelled using the fluorophore 8-aminonaphthalene-1,3,6-trisulfonic acid (ANTS). Polyacrylamide gel electrophoresis is used to separate the species and standards used to identify an oligosaccharide profile that is typical of the target molecule. Further, the oligosaccharides are identified using matrix assisted laser desorption and ionisation – time of flight mass spectrometry (MALDI-TOF) relying on the fluorophore and a specific matrix to ionise each sugar. The mass of each sugar is determined and potential structures identified using the GlycoSuite database. The potential sugar structures are further characterised by tandem mass spectrometric techniques, which allows partial or complete characterisation of the oligosaccharides present and their relative amounts. Further, the

process is repeated using the isoforms identified by 2D gel electrophoresis to generate a profile of the oligosaccharides present on each of the isoforms isolated.

EXAMPLE 7

QCM and SPR

The binding characteristics and activity of the target molecule is determined using either quartz crystal microbalance (QCM) or surface plasmon resonance (SPR). In both cases a suitable receptor for the molecule is bound to a wafer using the chemistry described by the manufacturer. The target molecule is dissolved into a suitable biological buffer and allowed to interact with the receptor on the chip by passing the buffer over it. Changes in the total protein mass on the surface of the wafer are measured either by change of oscillation frequency (in the case of QCM) or changes in the light scattering qualities of the chip (in the case of SPR). The chip is then treated with the biological buffer alone to observe the release of the target molecule back into solution. The rate at which the receptors reach saturation and complete disassociation is then used to calculate the binding curve of the target molecule.

EXAMPLE 8

Generation of a Transgenic Host Cell Line

(a) Transgenic Host Cell Line with alpha-2,6-sialyltransferase

The cDNA coding for alpha-2,6-sialyltransferase (alpha 2,6ST) is amplified by PCR from poly(A)-primed cDNA. The PCR product is ligated into a suitable vector, for instance pIRESpuro4 or pCEP4, to generate an alpha 2,6ST plasmid. The cloned cDNA is sequenced and its identity verified by comparison with the published alpha-2,6ST cDNA sequence. DNA sequencing is performed using known methods.

Mammalian host cells, including cell clones of the same lineage that express high levels of target molecule (cell line-target molecule) are transfected with the alpha 2,6ST plasmid, which also carries an antibiotic resistance marker. Selection of stably transfected cells is

performed by incubaton of the cells in the presence of the antibiotic; colonies of antibiotic-resistant cells that appear subsequent to transfection are pooled and examined for intracellular alpha 2,6ST activity. To isolate individual cell clones expressing alpha 2,6ST, cell pools are cloned by a limiting dilution process as described by Kronman (*Gene 121*: 295–304, 1992). Individual cell clones are chosen at random, cells expanded and clones tested for alpha 2,6ST activity.

Cell pellets are washed, resuspended in lysis buffer and left on ice prior to sonication. The cell lysate is centrifuged and the clear supernatant is assayed for protein concentration (via known methods) and sialyltransferase activity. Sialyltransferase activity is assayed by known methods, for example the method detailed by Datta *et al.* (*J Biol Chem 270:*1497–1500, 1995).

Expressed target molecule is purified from high-expressing alpha 2,6ST cell line-target molecule cells and subjected to *in vitro* and/or *in vivo* half-life bioassays (see Example 10). Target molecule from high-expressing alpha 2,6ST cell displays an increased *in vitro* and/or *vivo* half-life in comparison to target molecule derived from the same parent cell line without any subsequent transgene manuipulation or target molecule derived from other cell lines.

(b) Transgenic Host Cell Line with fucosyltransferase

The cDNA coding for a fucosyltransferase (FT) such as FUT1, FUT2, FUT3, FUT4, FUT5, FUT6, FUT7, FUT8, FUT9, FUT10, FUT11 is amplified by PCR from poly(A)-primed cDNA. The PCR product is ligated into a suitable vector, for instance pIRESpuro4 or pCEP4, to generate an alpha 2,6ST plasmid. The cloned cDNA is sequenced and its identity verified by comparison with the published FT cDNA sequence. DNA sequencing is performed using known methods.

Human host cells, including cell clones of the same lineage that express high levels of target molecule molecule (cell line-target molecule) are transfected with the FT plasmid, which also carries an antibiotic resistance marker. Selection of stably transfected cells is

performed by incubation of the cells in the presence of the antibiotic; colonies of antibiotic-resistant cells that appear subsequent to transfection are pooled and examined for intracellular FT activity. To isolate individual cell clones expressing FT, cell pools are cloned by a limiting dilution process as described by Kronman (*Gene 121*: 295–304, 1992); Individual cell clones are chosen at random, cells expanded and clones tested for FT activity.

Cell pellets are washed, resuspended in lysis buffer and left on ice prior to sonication. The cell lysate is centrifuged and the clear supernatant is assayed for protein concentration (via known methods) and FT activity. FT activity is assayed by known methods, for example the method detailed by Mas *et al.* (*Glycobiology 8(6)*: 605-13, 1998).

Expressed target molecule is purified from high-expressing FT cell line-target molecule cells. A Lewis x-specific antibody, such as L5 and a sialyl Lewis x-specific antibody such as KM93, HECA493, 2H5 or CSLEX are used to test the presence of Lewis x or sialyl Lewis x structures according to methods known in the art, for example, as detailed in Lucka *et al.* (*Glycobiology 15(1)*:87, 2005). Alternatively, the presence of Lewis x or sialyl Lewis x structures may be detected by treating the sample with appropriate glycosidases and detecting the effect of the glycosidases on parameters such as mass using MS or retention time using HPLC. Glyco mass fingerprinting, as described in Example 5, may also be employed to predict the presence of Lewis x or sialyl Lewis x structures.

EXAMPLE 9

Differential Gene Expression

Differences in gene expression can be analyzed using a target cell line of the target molecule. The target cells are grown to the appropriate density and treated with a range of concentration of target molecule or buffer control for a number of hours, for instance, 72 hours.

At various time points RNA is harvested, purified, and reverse transcribed according to Affymetrix protocols. Labelled cRNA (e.g. biotin labelled) is then prepared and hybridised

to expression arrays e.g. U133 GeneChips. Following washing and signal amplification, the GeneChips are scanned using a GeneChip scanner (Affymetrix) and the hybridisation intensities and fold change information at various time points is obtained using GeneChip software (Affymetrix).

The target molecule induces unique gene expression and results in different mRNA profiles upon comparison with profiles induced by cytokines or receptors produced from different sources e.g. *E. coli*, yeast or CHO cells.

EXAMPLE 10

Determining the Half-Life of the Target Molecule of the Present Invention

The half-life of the target molecule is determined in an *in vitro* system. Composition containing target molecule is mixed into human serum/plasma and incubated at a particular temperature for a particular time (e.g. 37 degrees for 4h, 12h etc). The amount of target molecule remaining after this treatment is determined by ELISA methods or dot blot methods known in the art. The biological activity of the remaining target molecule is determined by performing a suitable bioassay chosen by a person skilled in the relevant art. The serum chosen may be from a variety of human blood groups (eg A, B, AB, O etc.).

The half-life of target molecule is also determined in an *in vivo* system. Composition containing target molecule is labelled by a radioactive tracer (or other means) and injected intravenously, subcutaneously, retro-orbitally, intramuscularly or intraperitonally into the species of choice for the study, for instance, mouse, rat, pig, primate or human. Blood samples are taken at time points after injection and assayed for the presence of target molecule (either by ELISA methods, dot blot methods or by trichloroacetic acid (TCA)-precipitable label e.g. radioactive counts). A comparison composition consisting of target molecule produced from other sources eg *E. coli*, yeast, or CHO cells can be run as a control.

EXAMPLE 11

In Vivo Studies using the Target Molecule of the Present Invention

The individual subjects of the *in vivo* studies described herein are warm-blooded vertebrate animals, which includes humans.

The clinical trial is subjected to rigorous controls to ensure that individuals are not unnecessarily put at risk and that they are fully informed about their role in the study.

Preferably to account for the psychological effects of receiving treatments, the trial is conducted in a double-blinded fashion. Volunteers are randomly assigned to placebo or target molecule treatment groups. Furthermore, the relevant clinicans are blinded as to the treatment regime administered to a given subject to prevent from being biased in their post-treatment observations. Using this randomization approach, each volunteer has the same chance of being given either the new treatment or the placebo.

Volunteers receive either the target molecule or placebo for an appropriate period with biological parameters associated with the indicated disease state or condition being measured at the beginning (baseline measurements before any treatment), end (after the final treatment), and at regular intervals during the study period. Such measurements include the levels of target molecule in body fluids, tissues or organs compared to pre-treatment levels. Other measurements include, but are not limited to, indices of the disease state or condition being treated, body weight, blood pressure, serum titers of pharmacologic indicators of disease such as specific disease indicators or toxicity as well as ADME (absorption, distribution, metabolism and excretion) measurements.

Information recorded for each patient includes age (years), gender, height (cm), family history of disease state or condition (yes/no), motivation rating (some/moderate/great) and number and type of previous treatment regimens for the indicated disease or condition.

Volunteers taking part in this study are adults aged 18 to 65 years and roughly an equal number of males and females participate in the study. Volunteers with certain

characteristics are equally distributed for placebo and target molecule treatment. In general, the volunteers treated with placebo have little or no response to treatment, whereas the volunteers treated with the target molecule show positive trends in their disease state or condition index at the conclusion of the study.

EXAMPLE 12

(a) Comparing the *in vitro* bioactivities of SCF of the present invention and rh SCF expressed using non-human systems

M-07e cells were used to compare the bioactivities of SCF of the present invention and human SCF expressed using non-human systems.

Cells were plated in wells of a 96-well tissue culture plates at a concentration of 3 x 10^5 cells/ml cells in 200 µl medium. Cells were treated with various concentrations of SCF of the present invention and incubated for 4 days at 37 °C. Cell numbers were then measured using an MTS assay.

For the MTS assay, a CellTiter 96 Aqueous One Solution Cell Proliferation Assay (Promega) was employed. In this assay a tetrazolium compound MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) in the presence of an electron coupling reagent (phenazine methosulfate) was bioreduced by the cells into a formazan product. The concentration of the formazan was determined by reading the absorbance of the resultant solution at 490nm by a spectrophotometer (E Max precision microplate reader, Molecular Devices).

The above assay was repeated using a human SCF expressed in a non-human cell system (*E. coli*; R&D Systems Catalogue No. 255-SC).

The respective ED_{50}s were calculated after curve fitting the absorbance and the SCF concentration values using a 4-parameter fit equation.

The ED_{50} of SCF of the present invention was 10-20 ng/ml, 1.5-3.0 fold lower than the ED_{50} of the human SCF expressed in *E. coli* (20-30 ng/ml). Thus, the SCF of the present invention displays 1.5-3 fold greater proliferative activity than human SCF expressed in *E. coli* (Figure 2).

(b) Comparing the *in vitro* bioactivities of SCF of the present invention and human SCF expressed using non-human systems to stimulate proliferation of $CD34^+$ cells from human umbilical cord blood.

Human umbilical cord blood, collected at the time of delivery, was processed using standard techniques to isolate $CD34^+$ human haematopoietic cells. $CD34^+$ cells were plated into multi-well plates at 1 x 10^4 cells/ml in the presence of SCF of the present invention (100 ng/ml). Plates were incubated at 37 °C in a 5% CO_2 incubator for 1-2 weeks. Cells were harvested and viable nucleated cells were counted using a haemocytomter and Trypan Blue via standard methods.

The above assay was repeated using a recombinant human SCF expressed in *E. coli* cells (Peprotech Catalogue # 300-07).

After 1-2 weeks culture, SCF of the present invention yields more viable cells than that produced by treatment with rhSCF expressed in *E. coli* (Figure 3).

(c) Comparing the SCF of the present invention, in combination with G-CSF, with rhSCF expressed in non-human systems to stimulate differentiation of human $CD34^+$ cells isolated from cord blood

Human umbilical cord blood, collected at the time of delivery, was processed using standard techniques to isolate $CD34^+$ human haematopoietic cells. $CD34^+$ cells were plated into multi-well plates at 1 x10^4 cells/ml in the presence of SCF of the present invention (10-200 ng/ml) in combination with recombinant human cell expressed G-CSF (20-100 ng/ml) (Apollo Cytokine Research Catalogue # 1001C). Plates were incubated at 37 °C in a 5% CO_2 incubator for 1-2 weeks. Cells were harvested, counted and subjected

to surface antigen analysis, including CD11b, CD14, CD15 and CD184 markers, by flow cytometry using standard methods.

The above assay was repeated using a recombinant human SCF (rhSCF) expressed in *E. coli* cells (Peprotech Catalogue # 300-07).

After 1-2 weeks culture, SCF of the present invention in combination with G-CSF expressed in human cells yielded on average up to 52 % ($p<0.00021$) more viable cells than that produced by treatment with rhSCF expressed in *E. coli* in combination with rhG-CSF expressed in *E. coli* (Peprotech Catalogue # 300-23) (Figure 4).

Treatment SCF of the present invention in combination with human cell expressed rhG-CSF yielded higher percentages of CD11b, CD14, CD15, and CD184 expression than that observed following treatment with rhSCF and rhG-CSF expressed in *E. coli*. (Figure 5).

In a colony-forming assay, 100 $CD34^+$ human cord blood cells were plated with 100 ng/ml SCF of the present invention in combination with 100 ng/ml human cell expressed G-CSF. After 1-2 weeks growth colonies were counted and assessed for size.

The above assay was repeated using a rhSCF (Peprotech Catalogue # 300-07) together with rhG-CSF purified from *E. coli* cells (Peprotech Catalogue # 300-23).

SCF of the present invention yielded more colonies than rhSCF expressed in *E. coli*, and 30% more of these were large colonies (greater than 100 cells) $p<0.005$, indicating that SCF of the present invention had a greater capacity to induce proliferation than rhSCF expressed in *E. coli* (Figure 6).

Such results indicate that the treatment of human cord blood-derived cells with SCF of the present invention in combination with human cell expressed G-CSF has a greater capacity to induce granulocytic proliferation and maturation than treatment with rhSCF and rhG-CSF expressed in *E. coli*.

(d) Comparing the ability of SCF of the present invention in combination with human cell expressed erythropoietin and rhSCF expressed using non-human systems to stimulate differentiation of $CD34^+$ cells from human umbilical cord blood.

$CD34^+$ cells were plated into multi-well plates at 1 x 10^4 cells/ml/well in the presence of SCF of the present invention (10-200 ng/ml) in combination with 20-100 ng/ml recombinant human cell expressed erythropoietin (EPO: Apollo Cytokine Research Cat # 3005B). Plates were incubated at 37 °C in a 5% CO_2 incubator for 1-2 weeks. Cells were harvested and subjected to surface antigen analysis, including the erythroid markers CD71 and glycophorin A, by flow cytometry using standard methods.

The above assay was repeated using a recombinant human SCF expressed in *E. coli* cells (Peprotech Catalogue # 300-07) in combination with rhEPO expressed in CHO cells (Peprotech, Cat # 600-05).

SCF of the present invention in combination with human cell expressed EPO (Apollo Cytokine Research Cat # 3005B) resulted in an average of 2.7-fold more cells expressing glycophorin A ($p<0.05$) and 2.1 fold more cells expressing CD71 ($p<0.05$) than cells treated with rhSCF purified from *E. coli* in combination with rhEPO, indicating a more mature erythroid phenotype if SCF of the present invention was included in the protocol (Figure 7).

(e) Comparing the ability of SCF of the present invention in combination with human cell expressed erythropoietin and rhSCF expressed using non-human systems to stimulate cell signalling in erythroid progenitors derived from human umbilical cord blood cells.

SCF alone and in combination with EPO is known to activate signalling pathways in erythroid progenitor cells. These include STAT proteins (signal transducers and activators of transcription), which regulate cell growth and differentiation, and also Akt1 (also known as protein kinase B), which promotes cell survival by inhibiting apoptosis. Both of these

proteins are activated by phosphorylation of specific amino acid residues, which occurs within the cell.

Human $CD34^+$ haematopoietic stem cells from umbilical cord blood were expanded for 7-10 days in the presence of 50ng/ml EPO (Apollo Cytokine Research Catalogue# 3005B), 50 ng/ml SCF (Apollo Cytokine Research Catalogue# 3020B) and 20% foetal calf serum to generate a suitable number of erythroid progenitors. Cells were harvested and incubated for 0-30 min with either 100 ng/ml SCF of the present invention, or 100 ng/ml SCF of the present invention in combination with 50ng/ml human cell expressed EPO (Apollo Cytokine Research Cat # 3005B). Cells were then lysed and total protein extracts were separated by SDS-PAGE followed by western blotting according to standard chromatography techniques. The western blots were probed with an anti-phospho STAT5 antibody to measure STAT5 activation (Y694; Becton Dickinson Cat #9351), and a duplicate blot was probed with a STAT5 antibody (Becton Dickinson Cat #9310) to allow for normalisation against total STAT5 loading onto the SDS gel.

The same samples were separated by SDS-PAGE followed by western blotting according to standard chromatography techniques. The western blots were probed with an anti-phospho Akt (Cell Signaling Technologies #9271) to measure Akt activation, and a duplicate western blot was probed with an Akt antibody (Cell Signaling Technologies #9272) to allow for normalization against total Akt loading onto the SDS gel.

The above assay was repeated using 100 ng/ml recombinant human SCF expressed in *E. coli* cells (Peprotech Catalogue # 300-07) with or without 50 ng/ml rhEPO expressed in CHO cells (Peprotech, Cat # 600-05).

Treatment with SCF of the present invention induced a 3.5-fold activation of STAT5, whereas rhSCF expressed in *E. coli* only induced a 1.8-fold activation of STAT5 (Figure 8). When SCF of the present invention was used in combination with human cell expressed EPO, a 20-fold activation of STAT5 was observed, whereas rhSCF expressed in *E. coli* in combination with rhEPO expressed in *E. coli* only induced a 10-13 fold activation of STAT5 (Figure 8).

Similarly, SCF of the present invention induced a 42-fold activation of Akt whereas rhSCF expressed in *E. coli* only induced a 9-fold activation of Akt (Figure 8). When SCF of the present invention was used in combination with human cell expressed EPO, up to 37-fold activation of Akt was observed, whereas rhSCF expressed in *E. coli* in combination with rhEPO expressed in *E. coli* only induced a 11-fold activation of STAT5 (Figure 8).

(f) SCFR-Fc of the present invention is able to inhibit rhSCF induced STAT3 activation in human leukemia cells.

SCF is known to rapidly activate STAT3 in M-07e cells, as part of signalling mechanisms leading to cell proliferation. 0-500 ng/ml SCFR-Fc of the present invention was incubated with 50 ng/ml rh SCF (Apollo Cytokine Research; Cat No. 3020) for 1-2 hours at 37 °C under appropriate conditions to facilitate complete binding. This was then added to the M-07e cells for 0-10 minutes, after which cells were washed, lysed and total protein extracts were separated by SDS-PAGE followed by western blotting according to standard chromatography techniques. The western blots were probed with an anti-phospho STAT3 antibody to measure STAT3 activation (s727; Becton Dickinson Cat#612542), and a duplicate blot was probed with a STAT3 antibody (Becton Dickinson Cat#610189) to allow for normalisation against total STAT3 loading onto the SDS gel.

Addition of SCFR-Fc of the present invention was able to inhibit rh SCF (expressed from human cells) induced activation of STAT3 in M-07e cells (Figure 9).

(g) Comparing the ability of SCFR molecules of the present invention and rhSCF molecules expressed in non-human systems to inhibit MAPK activation in human leukemia cells.

SCF is also known to activate MAPK in M-07e cells, as part of signalling mechanisms leading to cell proliferation. 0-500 ng/ml SCFR-Fc or SCFR-ECD of the present invention was incubated with 50 ng/ml SCF for 1-2 hours at 37 °C under appropriate conditions to facilitate complete binding. This was then added to the M-07e cells for 0-10 minutes, after

which cells were washed, lysed and total protein extracts were separated by SDS-PAGE followed by western blotting according to standard chromatography techniques.

Similarly a western blot was probed with an anti-phospho MAPK antibody (Cell Signaling Technologies Cat # 4377) to measure MAPK (p44 and p42 ERK) activation, and a duplicate blot was probed with a MAPK antibody (Cell Signaling Technologies Cat # 9102), to allow for normalisation against total ERK loading onto the SDS gel.

The above assay was repeated using rhSCFR-ECD expressed in a non-human cell system (*Sf*21 insect cells) sourced from R&D Systems (Cat #332-SR).

SCFR-Fc and SCFR-ECD of the present invention were both able to inhibit SCF induced activation (3.7-fold 2-fold respectively) and of the p42 and p44 MAPKs in M-07e cells, whereas the rh SCFR-ECD purified from non-human cells did not inhibit MAPK activation (Figure 9).

(h) Inhibition of STAT3 activation by SCF R-Fc of the present invention in human erythroid progenitor cells.

SCF is known to rapidly activate STAT3 in erythroid progenitor cells, as part of signalling mechanisms leading to cell proliferation. Human $CD34^+$ haematopoietic stem cells were expanded for 7-10 days in the presence of 50ng/ml EPO (Apollo Cytokine Research Catalogue# 3005B), 50 ng/ml SCF (Apollo Cytokine Research Catalogue# 3020B) and 20% foetal calf serum to generate a suitable number of erythroid progenitors. 0 or 500 ng/ml SCFR-Fc of the present invention was incubated with 50 ng/ml SCF for 1-2 hours at 37 °C under appropriate conditions to facilitate complete binding. This was then added to the expanded erythroid progenitors for 0- 10 minutes, after which cells were washed, lysed and total protein extracts were separated by SDS-PAGE followed by western blotting according to standard chromatography techniques. The western blots were probed with an anti-phospho STAT3 antibody to measure STAT3 activation (s727; Becton Dickinson Cat#612542), and a duplicate blot was probed with a STAT3 antibody Becton Dickinson Cat#610189 to allow for normalisation against total STAT3 loading onto the SDS gel.

Addition of SCFR-Fc of the present invention was able to inhibit SCF-induced activation of STAT3 in erythroid progenitor cells (Figure 10).

(i) Comparing the abilities of SCFR molecules of the present invention and SCFR molecules expressed in a non-human system to inhibit MAPK activation in human erythroid progenitor cells.

SCF is known to rapidly activate MAPK in erythroid progenitor cells, as part of signalling mechanisms leading to cell proliferation. Human $CD34^+$ haematopoietic stem cells were expanded for 7-10 days in the presence of 50ng/ml EPO (Apollo Cytokine Research Catalogue# 3005B), 50 ng/ml SCF (Apollo Cytokine Research Catalogue# 3020B) and 20 % foetal calf serum to generate a suitable number of erythroid progenitors. 0 or 500 ng/ml SCFR-Fc or SCFR-ECD of the present invention was incubated with 50 ng/ml SCF for 1-2 hours at 37 °C under appropriate conditions to facilitate complete binding. This was then added to the expanded erythroid progenitors for 0- 10 minutes, after which cells were washed, lysed and total protein extracts were separated by SDS-PAGE followed by western blotting according to standard chromatography techniques. Similarly a western blot was probed with an anti-phospho MAPK antibody (Cell Signaling Technologies Cat # 4377) to measure MAPK (p44 and p42 ERK) activation, and a duplicate blot was probed with a MAPK antibody (Cell Signaling Technologies Cat # 9102) to allow for normalisation against total MAPK loading onto the SDS gel.

The above assay was repeated using SCFR-ECD expressed in a non-human cell system (*Sf*21 insect cells) sourced from R&D Systems (Catalogue #332-SR).

SCFR-Fc and SCFR-ECD of the present invention were both able to inhibit SCF induced activation of the p42 and p44 MAPKs (7.7-fold and 1.9-fold respectively), whereas rhSCFR-ECD purified from non-human cells did not (Figure 10).

EXAMPLE 13

In vitro comparison of Immunoreactivity Profiles between SCF of the Present Invention and a human SCF molecule expressed using non-human systems

Protein estimation of SCF of the present invention was determined by the A280 absorbance method using the calculated extinction coefficient (e), based on the peptide sequence, and the measured molecular mass based on SDS-PAGE analysis.

SCF of the present invention, standardised using the A280 assay results, was diluted and tested in a Peprotech human SCF ELISA kit (Cat # 900-K34) in accordance with the manufacturer's instructions. The above-mentioned ELISA kit employs a human SCF expressed in *E.coli* as a standard.

The ELISA kit results indicate a clear underestimate of the SCF of the present invention concentration (Figure 11). This result indicates different immunoreactivity profiles of SCF of the present invention and a non-human cell expressed human SCF molecule.

At a structural level, such a result indicates different immunoreactivity profiles of SCF of the present invention and a non-human cell expressed human SCF molecule.

EXAMPLE 14

Further Purification of Target Molecule of the Present Invention and Peptide Mass Fingerprinting by ESI-MS/MS

In addition to the purification protocol as described in Example 2, purification of the target molecule of the present invention is further performed by RP-HPLC, using a commercially available column. Eluting proteins are monitored by the absorbance at 215 or 280nm and collected with correction being made for the delay due to tubing volume between the flow cell and the collection port.

A gel piece containing the protein sample from a 1D or 2D gel is digested in trypsin solution as described in Example 3. Alternatively, a solution containing the protein sample is digested with trypsin in an ammonium bicarbonate buffer (10-25 mM, pH 7.5-9). The solution is incubated at 37° C overnight. The reaction is then stopped by adding acetic acid until the pH is in the range 4-5. The peptide samples are concentrated and desalted using C18 Zip-Tips (Millipore, Bedford, MA) or pre-fabricated micro-columns containing Poros R2 chromatography resin (Perspetive Biosystems, Framingham, MA) as described in Example 3.

The protein sample (2-5 μl) is injected onto a micro C18 precolumn and washed with 0.1% formic acid at 30 μl/min to concentrate and desalt. After a 3 min wash the pre-column is switched into line with the analytical column containing C18 RP silica (Atlantis, 75μm x 100mm, Waters Corporation). Peptides are eluted from the column using a linear solvent gradient, with steps, from H_2O:CH_3CN (95:5; + 0.1% formic acid) to H_2O:CH_3CN (20:80, + 0.1% formic acid) at 200 nl/min over a 40 min period. The LC eluent is subject to positive ion nanoflow electrospray analysis on a Micromass QTOF Ultima mass spectrometer (Micromass, Manchester, UK).

Tandem MS is performed using a Q-Tof hybrid quadrupole / orthogonal-acceleration TOF mass spectrometer (Micromass). The QTOF is operated in a data dependent acquisition mode (DDA). A TOFMS survey scan was acquired (m/z 400-2000, 1.0s), with the three largest multiply charged ions (counts >15) in the survey scan sequentially subjected to MS/MS analysis. MS/MS spectra were accumulated for 8 s (m/z 50-2000).

The LC/MS/MS data are searched using Mascot (Matrix Science, London, UK) and Protein Lynx Global Server ("PLGS") (Micromass). The protein sample is anticipated to be the target molecule.

EXAMPLE 15

(a) Immunogenicity in non-human animals

(i) Animal immunization with target protein

Separate groups of non-human animals, for example, mice are immunized either sub-cutaneously, intramuscularly or intraperitoneally (IP) with 1-100ug of protein of the present invention and the protein expressed in non-human cells, respectively. Animals receive a secondary immunization one month following immunization. Prior to immunization, protein is emulsified in an adjuvant, for example, complete Freud's adjuvant for the primary immunization and incomplete Freud's adjuvant for the secondary immunization.

(ii) Detection of antibodies directed to target protein

For the detection of antibody response, animals from each group are bled from the tail and sera pooled. Protein-specific antibodies are detected by a solid phase ELISA using 50ng/well of protein of the present invention. Different immunoglobulin isotypes are detected by using labelled detection antibodies raised against IgG1, IgG2, IgG2b, IgG3, IgM, IgA, IgD. Alternatively, antibody response is measured against protein of the present invention blotted onto a membrane either as a dot blot or Western blot. Detection of different immunoglobulin isotypes are detected as described above. It is anticipated that the protein of the present invention will elicit an antibody response that is distinct to that of protein expressed in non-human cells.

(iii) T cell proliferation assay

Immunised animals are euthanised and spleen cells prepared. A suitable number of spleen cells, for example, 5 x 10^5 cells, from animals immunized with protein of the present invention are cultured with various concentrations of protein of the present invention while and equivalent number of spleen cells from animals immunized with protein expressed in

non-human cells are cultured with various concentrations of protein expressed in non-human cells. For T cell proliferation assays, spleen cells are cultured for 96 hours and treated with 1μCi [^{3}H] thymidine (6-7 μCi/umol) during the final 16 hours. The cells are harvested onto filter strips and [^{3}H] thymidine incorporation determined using standard methods. It is anticipated that the protein of the present invention will elicit a different proliferation response compared to the protein expressed in non-human cells.

(iv) IFN gamma assay

For the IFN gamma assay, culture supernatant from spleen cells incubated with either the protein of the present invention or protein expressed in non-human cells are harvested at 96 hours and IFN gamma production is detected by a sandwich ELISA, for example, a R&D Systems anti-IFN gamma Quantikine® ELISA kit (Cat # DIF50) in accordance with the manufacturer's instructions. It is anticipated that IFN gamma production will be different in culture supernatant derived from cells incubated with protein of the present invention compared with culture supernatant derived from cells incubated with protein expressed in non-human cells.

(b) *In vitro* Human Immunogenicity assays

(i) Human T-Cell response assay

Human dendritic cells and $CD4^+$ T cells are prepared from human blood as described in Stickler *et al. Toxicological Sciences 77*:280-289, 2004. Co-cultures of dendritic cells and $CD4^+$ T cells are plated out in 96 well plates containing 2 x 10^4 dendritic cells and 2 x 10^5 $CD4^+$ T cells. The protein of the present invention and[protein expressed in non-human cells undergo enzymatic digestion into peptide fragments using a suitable enzyme determined by cleavage site prediction software, for example, Peptide Cutter (http://au.expasy.org/tools/peptidecutter). The resulting peptide fragments are purified by a suitable technique, for example, liquid chromatography and added to the co-cultures to a final concentration of 5ug/ml. Cultures are incubated for 5 days and 0.5uCi ^{3}H thymidine

is then added to each culture. The cells are harvested onto filter strips and cell proliferation is determined by [^{3}H] thymidine incorporation.

It is anticipated that the peptides derived from protein of the present invention will elicit a weaker proliferation response compared to peptides derived from the protein expressed in non-human cells.

(ii) Human antibody response assay

Human donors undergoing treatment with protein expressed in non-human cells are bled and sera prepared. Protein-specific antibodies are detected by a solid phase ELISA against both 50ng/well of protein of the present invention and protein expressed in non-human cells. Different immunoglobulin isotypes are detected by using labelled detection antibodies raised against human IgG1, IgG2, IgG3, IgG4, IgM, IgA, IgD.

Alternatively, antibody response is measured against protein of the present invention and protein expressed in non-human cells blotted onto a membrane either as a dot blot or Western blot. Detection of different immunoglobulin isotypes are detected as described above.

It is anticipated that the immunoglobulin present in the sera of people treated with protein expressed in non-human cells will bind to protein expressed in non-human cells while either binding weakly or not binding with protein of the present invention.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to, or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

BIBLIOGRAPHY

Ackland *et al. Chromatogr 540*:187-198, 1991

Aloj *et al. J BiolChem* 247:1146-1151, 1971

Altschul *et al. Nucl Acids Res 25:*389, 1997

Antibodies: A Laboratory Manual, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1988)

Aronsson *et al. FEBS Lett 411*:359-364, 1997

Atherton and Shephard *Synthetic Vaccines 9:* Blackwell Scientific Publications

Ausubel *et al. In: Current Protocols in Molecular Biology* John Wiley & Sons Inc. 1994-1998)

Baneyx *Current Opinion in Biotechnology, 10*:411–421, 1999

Bernstein *Methods Mol Biol 237*:195-204, 2004

Bird *Science 242:*423, 1988

Blenis and Resh *Curr Opin Cell Biol 5(6):*984-9, 1993

Bonner and Laskey *Eur J Biochem 46:*83, 1974

Caprioli *et al. Biochem Biophys Res Commun 146*:291-299, 1987

Carr *et al. Anal Biochem 175:* 492-499, 1988

Carr *et al. Anal Chem 63*:2802-2824, 1991

Carr *et al. J Biol Chem 264(35)*:21286-21295, 1989

Chitlaru *et al. J Biochem 336*:647-658, 1998

Chitlaru *et al. J Biochem 363*:619-631, 2002

Clackson *et al. Nature 352*:624-628, 1991

Clarke *Curr Opin Cell Biol 5*:977 983, 1993

Datta *et al. J Biol Chem 270:*1497–1500, 1995

Edman *Mol Biol Biochem Biophys 8:*211-55, 1970

Erickson *et al. Science 249:*527-533, 1990

Farruggia *et al. Int J Biol Macromol 20:*43-51, 1997

Figeys and Aebersold, *Electrophoresis 19*:885-892, 1998

Franks *et al. Characterization of proteins*, Humana Press, Clifton, NJ, 1988

Fritz *et al. PNAS 95*:12283-12288, 1998

Fukuhara *et al. J Biol Chem 260*:10487-10494, 1985

Gelb *et al. Curr Opin Chem Biol 2(1)*:40-8, 1998

Gramer *et al. Biotechnology (N.Y.) 13(7)*:692-8, 1995

Guedez *et al. Am J Pathol 162:*1431–1439, 2003

Harrison and Packer *Methods Mol Biol 125:*211-216, 2000

Hearn *et al. Methods in Enzymol 104*:190-212, 1984

Herscovics *et al. FASEB J 7*:540-550, 1993

Herzberg *et al. Infrared and Raman Spectra of Polyatomic Molecules*, Van Nostrand Reinhold, New York, NY, 1945

Hodgson *Bio/Technology 9:*19-21, 1991

Holzwarth *et al. J Am Chem.Soc 178*:350, 1965

Honroe *et al. Biochem J 258*:99-204, 1989

Huston *et al. Proc Natl Acad Sci USA 85:*5879, 1988

James and Bottomley *Arch Biochem Biophy 356*:296-300, 1998

Jones *et al. Nature 321*:522-525, 1986

Kennet *et al. Monoclonal Antibodies, Hybridomas: A New Dimension in Biological Analyses*, Plenum Press, New York, 1980

Kivirikko *et al. FASEB Journal 3:*1609-1617, 1989

Kohler *et al. Nature 256*:495, 1975

Kortt *et al. Protein Engineering 10:*423, 1997

Krimm and Bandekar *Adv Protein Chem 38:*181-364, 1986

Kronman *Gene 121:*295–304, 1992

Kurochkin *et al. J Mol Biol 248*:414-430, 1995

Kwon and Yu *Biophim Biophys Acta 1335*:265-272, 1997

Larrick *et al. Bio/Technology 7*:934, 1989

Larsen *et al. J Biol Chem 265*:7055-7061, 1990

Li *et al. Biochemistry 34*:5762-5772, 1995

Liu *et al. J Immunol 158:*604-613, 1994

Lucka *et al. Glycobiology 15(1)*:87-100, 2005

Marks *et al. J Mol Biol 222*:581-597, 1991

Marmur and Doty *J Mol Biol 5:*109, 1962

Martin *et al. Nature Med 11(2)*:228-232, 2005

Mas *et al. Glycobiology 8(6)*: 605-13, 1998

McGettrick *et al. Methods Mol Biol 244*:151-7 2004

Mire-Sluis *et al. J Immunol Methods 289(1-2)*:1-16, 2004

Moore *J Biol Chem 278(27):*24243-24246, 2003

Morrison *et al. Proc Natl Acad Sci USA 81*:6851-6855, 1984

Nguyen *et al. J Chromatogr A. 705*:21-45, 1995

Packer *et al. Glycoconj J 5(8)*:737-47, 1998

Phillies *Anal Chem 62:*1049A-1057A, 1990

Pikal *et al. Pharm Res 8*:427-436, 1991

Presta *Curr Op Struct Biol 2*:593-596, 1992

Rando *Biochim Biophys Acta 1300(1)*:5-16, 1996

Reichmann *et al. Nature 332*:323-329, 1988

Sambrook *et al. Molecular Cloning - A Laboratory Manual*, Cold Spring Harbour, New York, USA, 1990

Schmid *et al. Protein structure, a practical approach*, Creighton Ed., IRI Press, Oxford, England, 1989

Shepherd *et al. Arterioscler Thromb Vasc Biol 24:*898-904, 2004

Stickler *et al. Toxicological Sciences 77:*280-289, 2004

Triguero *et al. J Neurochemistry 54*:1882-1888 1990

United States Patent No. 4,816,567

Ward *et al. Nature 334:*544, 1989

Wells *Methods Enzymol 202:*2699-2705, 1991

Wilkinson *Annu Rev Nutr 15*:161-89, 1995

Winter and Harris *TIPS 14:*139, 1993

Yoshioka *et al. Pharm Res 10*:103-108, 1993

CLAIMS

1. An isolated protein comprising a profile of measurable physiochemical parameters, wherein said profile is indicative of, associated with or forms the basis of one or more distinctive pharmacological traits, wherein said isolated protein comprises a physiochemical profile comprising a number of measurable physiochemical parameters, $\{[P_x]_1, [P_x]_2,...[P_x]_n,\}$, wherein P_x represents a measurable physiochemical parameter and "n" is an integer ≥ 1, wherein each of $[P_x]_1$ to $[P_x]_n$ is a different measurable physiochemical parameter, wherein the value of any one of the measurable physiochemical characteristics or an array of values of more than one measurable physiochemical characteristics is indicative of, associated with, or forms the basis of, a distinctive pharmacological trait, T_y, or an array of distinctive physiochemical traits $\{[T_y]_1, [T_y]_2, ...[T_y]_m\}$ wherein T_y represents a distinctive pharmacological trait and m is an integer ≥ 1 and each of $[T_y]_1$ to $[T_y]_m$ is a different pharmacological trait, wherein the isolated protein is selected from the group comprising SCF and SCFR.

2. The isolated protein of Claim 1, wherein said protein comprises one or more of the measurable physiochemical parameters set forth in Table 2.

3. The isolated protein of Claim 1 wherein said protein comprises one or more of the distinctive pharmacological traits set forth in Table 3.

4. A chimeric molecule comprising the isolated protein of Claim 1, or fragment thereof, fused to one or more peptide, polypeptide or protein moieties.

5. The chimeric molecule of Claim 4 wherein the peptide, polypeptide or protein moiety comprises the constant (Fc) or framework region of a human immunoglobulin.

6. The chimeric molecule of Claim 4 wherein the chimeric molecule is SCFR-Fc.

7. A pharmaceutical composition comprising the isolated protein or chimeric molecule of any one of Claims 1 to 6.

8. A method of treating or preventing a condition in a mammalian subject, wherein said condition can be ameliorated by increasing the amount or activity of a protein, said method comprising administering to said mammalian subject an effective amount of an isolated protein according to any one of Claims 1 to 3, a chimeric molecule according to any one of Claims 4 to 6 or the pharmaceutical composition of Claim 7.

9. A nucleotide sequence selected from the list consisting of SEQ ID NOs: 23, 25, 27, 31, 33, 35, 37, 39, 41, 43, or a nucleotide sequence having at least about 90% identity to any one of the above-listed sequences or a nucleotide sequence capable of hybridizing to any one of the above sequences or their complementary forms under high stringency conditions.

10. An isolated protein or chimeric molecule encoded by a nucleotide sequence selected from the list consisting of SEQ ID NOs: 23, 25, 27, 31, 33, 35, 37, 39, 41, 43, or a nucleotide sequence having at least about 90% identity to any one of the above-listed sequence or a nucleotide sequence capable of hybridizing to any one of the above sequences or their complementary forms under high stringency conditions.

11. An isolated nucleic acid molecule encoding a protein or chimeric molecule or a functional part thereof comprising a sequence of nucleotides having at least 90% similarity SEQ ID NOs: 23, 25, 27, 31, 33, 35, 37, 39, 41, 43 or after optimal alignment and/or being capable of hybridizing to one or more of SEQ ID NOs: 23, 25, 27, 31, 33, 35, 37, 39, 41, 43 or their complementary forms under high stringency conditions.

12. An isolated nucleic acid molecule comprising a sequence of nucleotides encoding a protein or chimeric molecule having an amino acid sequence substantially as set forth in one or more of SEQ ID NOs: 24, 26, 28, 32, 34, 36, 38, 40, 42, 44 or an amino acid sequence having at least about 90% similarity to one or more of SEQ ID NOs: 24, 26, 28, 32, 34, 36, 38, 40, 42, 44 after optimal alignment.

13. A kit for determining the level of human cell expressed human protein or chimeric molecule present in a biological preparation comprising (a) a solid phase support matrix; (b) one or more antibodies directed against a human protein according to any one of Claims 1 to 3 or chimeric molecule according to any one of Claims 4 to 6; (c) a blocking solution; (d) one or more stock solutions of substrate; (e) a solution of substrate buffer; (f) a standard human protein or chimeric molecule sample; and (g) instructions for use.

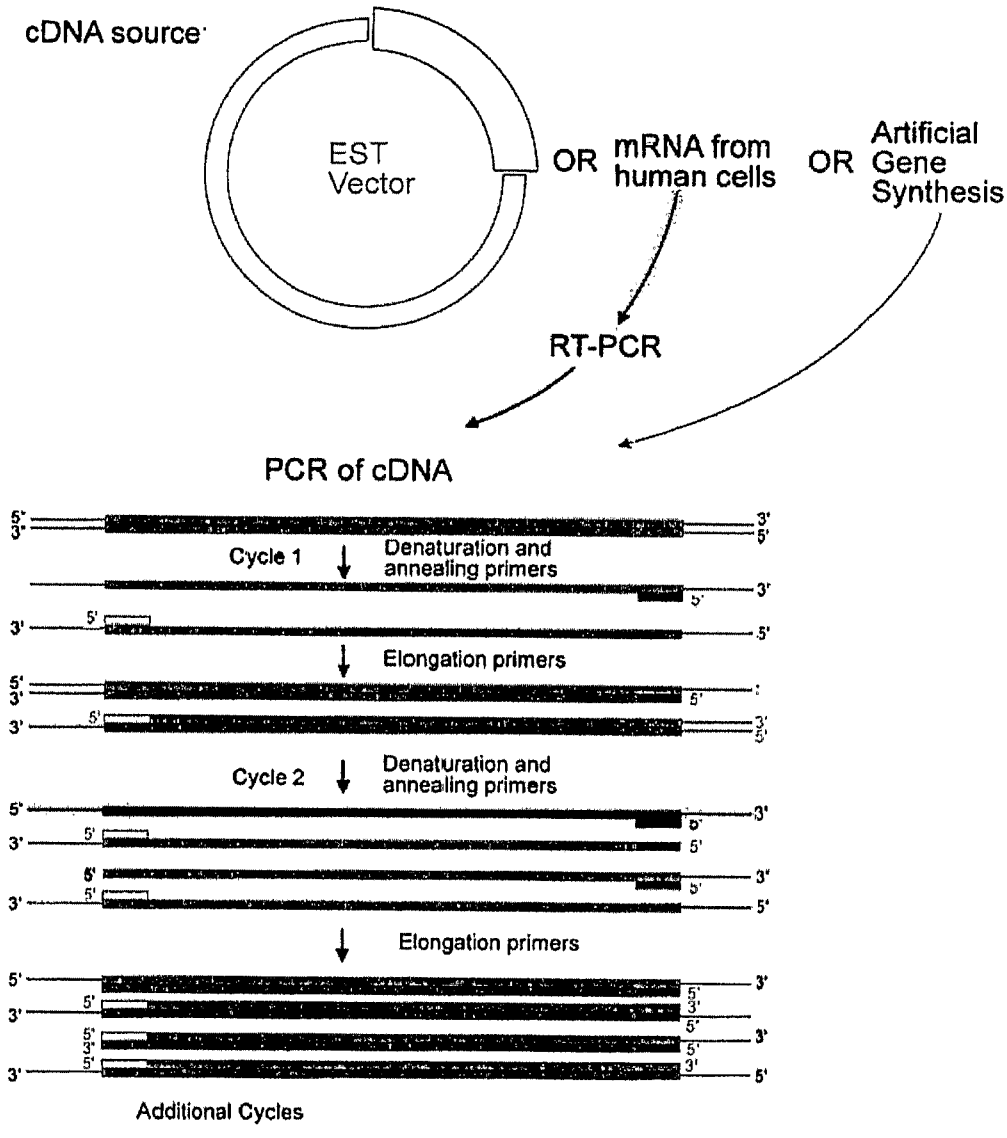
14. The kit of Claim 13, wherein the standard human protein or chimeric molecule sample is a preparation of the isolated protein of any one of Claim 2 or 3 or the chimeric molecule of Claim 4.

15. The kit of Claim 13 or 14, wherein the or each antibody is derived from an immunization of a mammal with a preparation comprising the isolated protein of any one of Claims 2 or 3 or the chimeric molecule of Claim 4.

16. The kit of any of Claims 13 to 15, wherein the human cell expressed human protein is naturally occurring human SCF or SCFR.

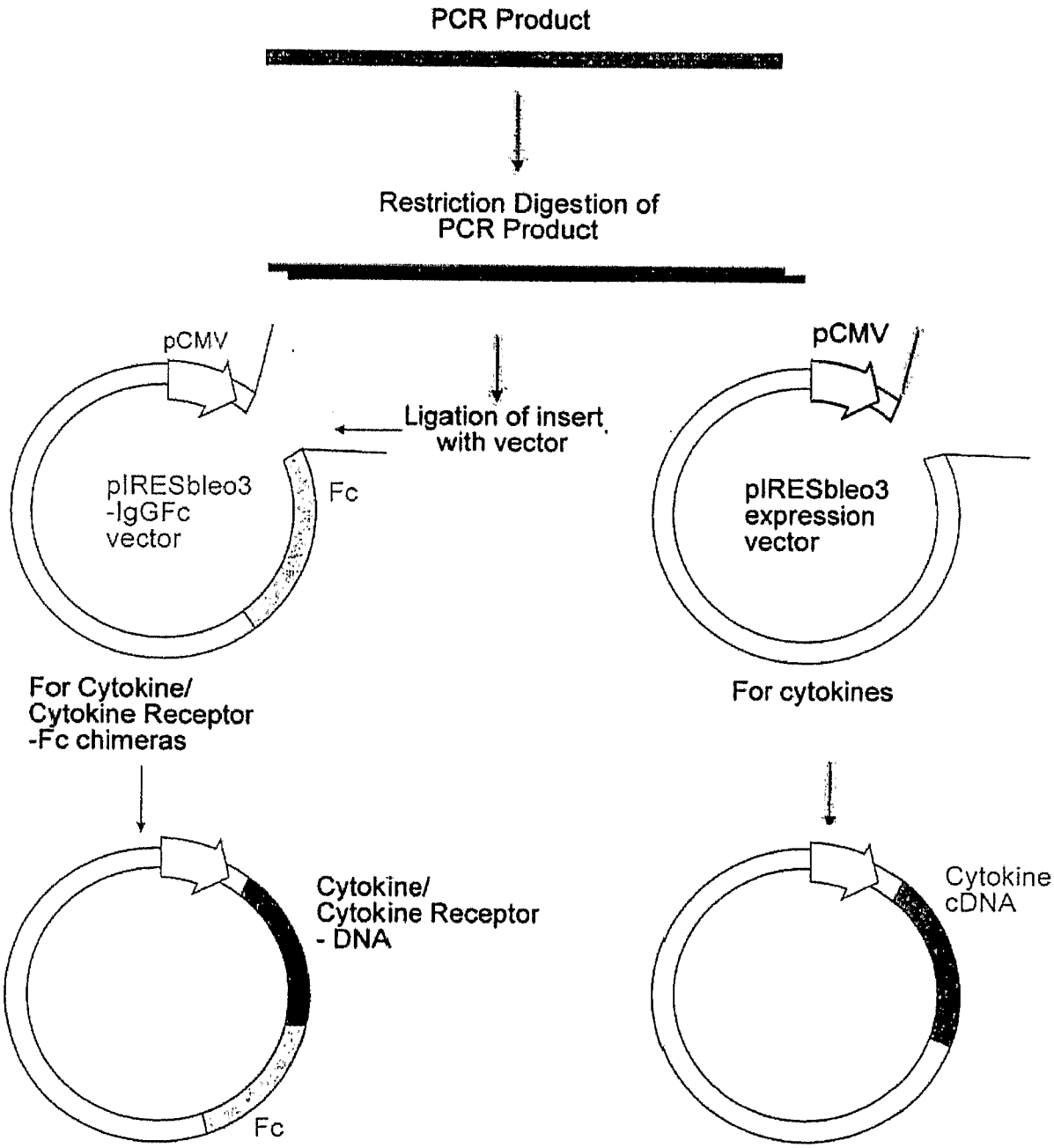
Figure 1 (i)

Flow Diagram of Cloning Process



Large quantity of cDNA insert (PCR product)

Figure 1 (ii)


PCR Product
Restriction Digestion of PCR Product
pCMV
pIRESbleo3 -IgGFc vector
Fc
Ligation of insert with vector
pCMV
pIRESbleo3 expression vector
For Cytokine/ Cytokine Receptor -Fc chimeras
For cytokines
Cytokine/ Cytokine Receptor - DNA
Fc
Cytokine cDNA

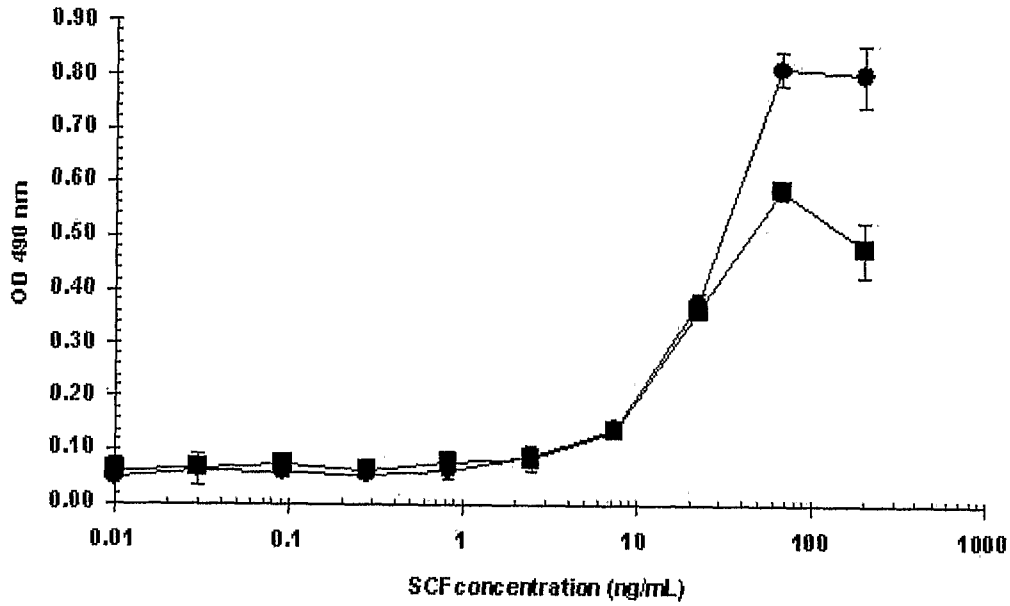

0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
OD 490 nm
0.01
0.1
1
10
100
1000
SCF concentration (ng/mL)

Figure 2

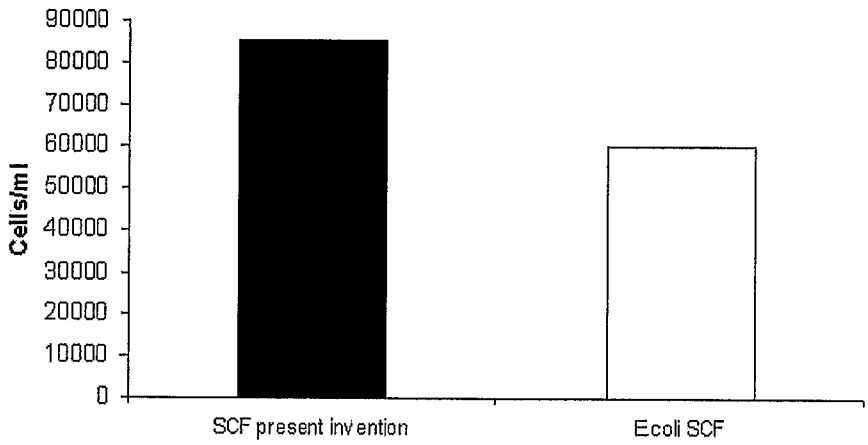

90000
80000
70000
60000
50000
40000
30000
20000
10000
0
Cells/ml
SCF present invention
Ecoli SCF

Figure 3

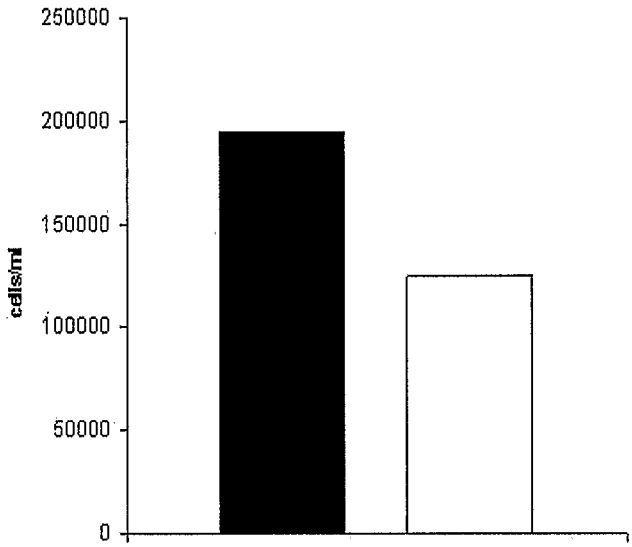


Figure 4

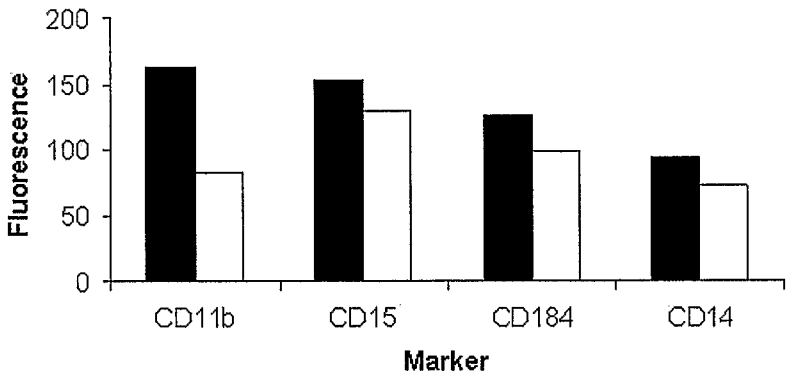


Figure 5

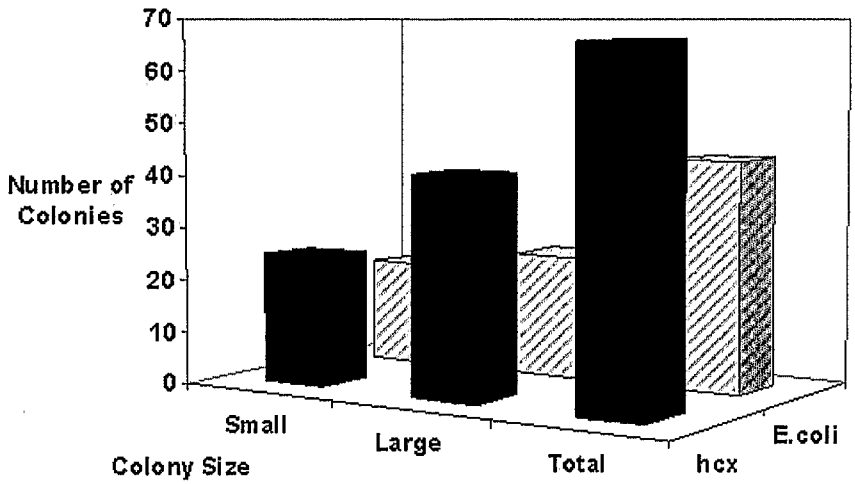

70
60
50
40
30
20
10
0
Number of Colonies
Small
Large
Total
E.coli
hcx
Colony Size

Figure 6

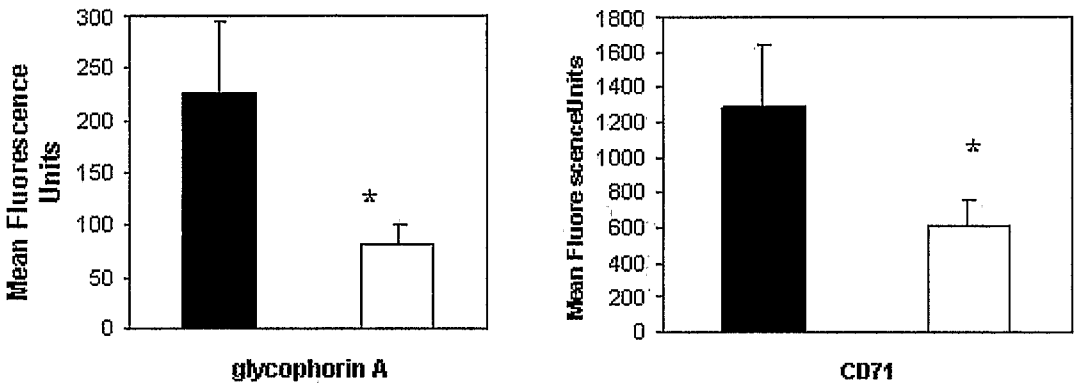

Mean Fluorescence Units
300
250
200
150
100
50
0
*
glycophorin A
Mean Fluorescence Units
1800
1600
1400
1200
1000
800
600
400
200
0
*
CD71

Figure 7

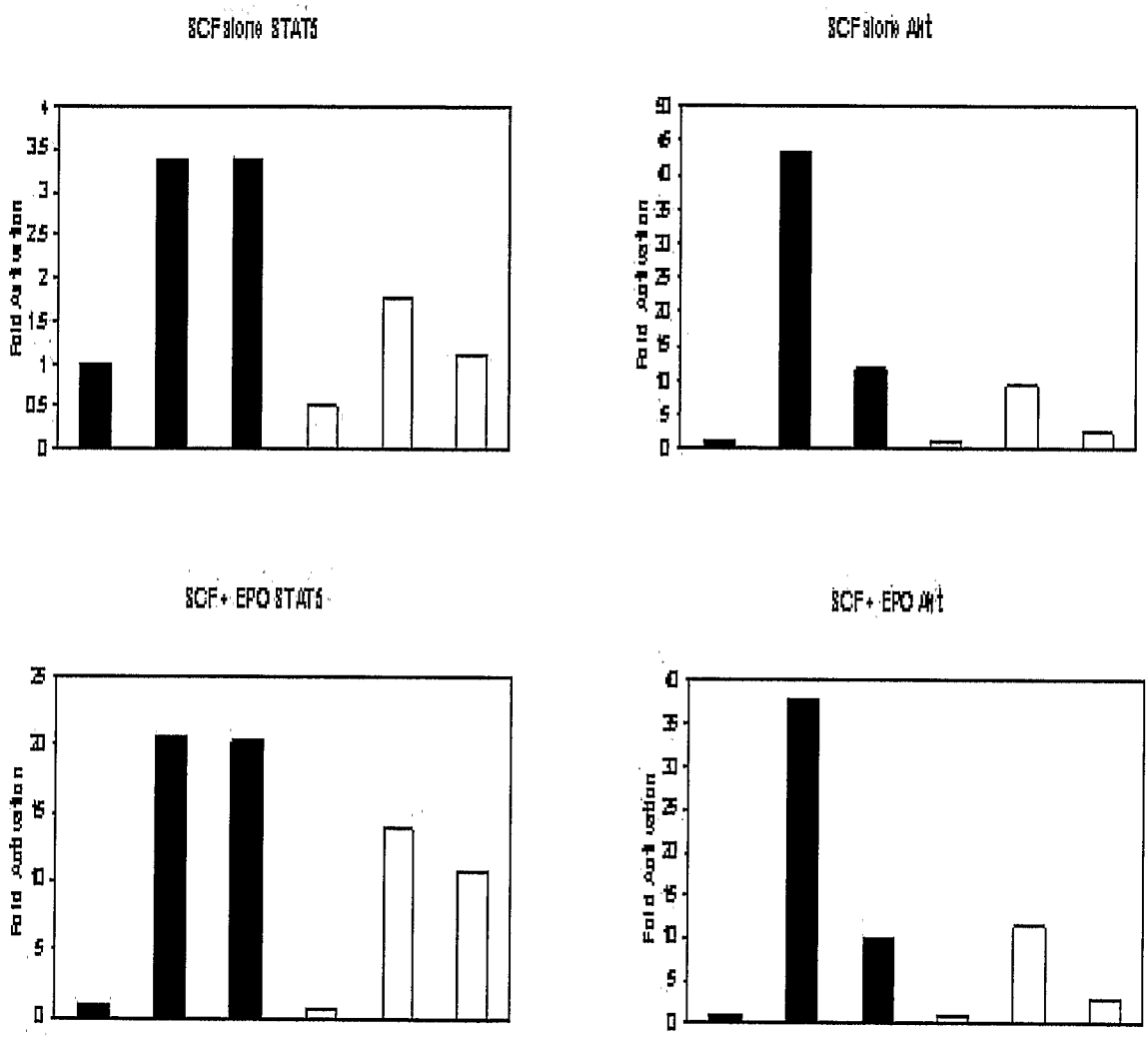


Figure 8

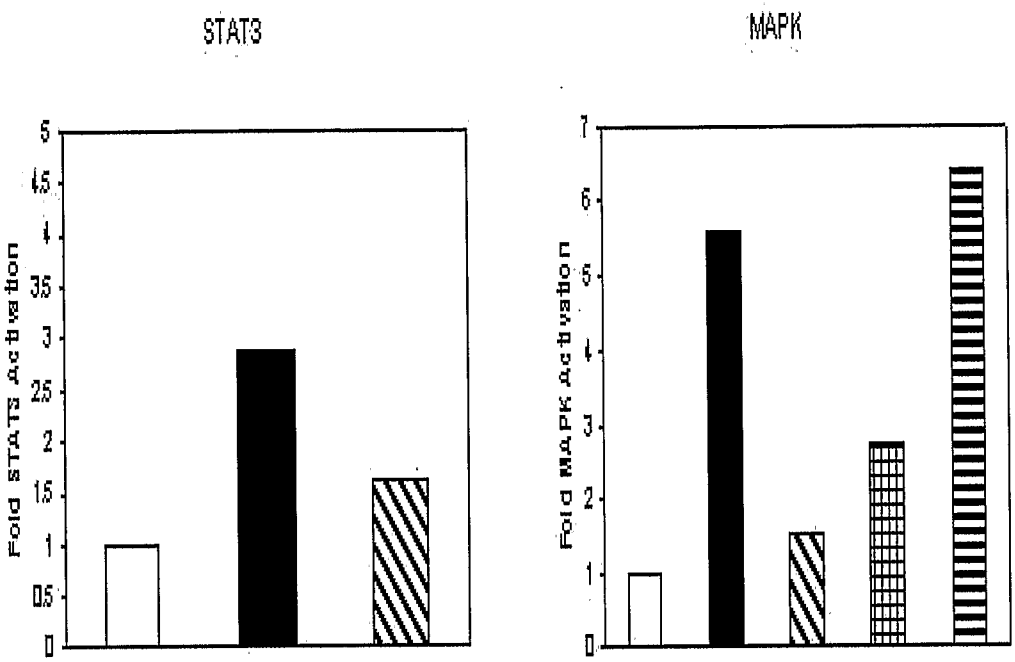

STAT3
MAPK
Fold STAT3 activation
Fold MAPK activation
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
0
1
2
3
4
5
6
7

Figure 9

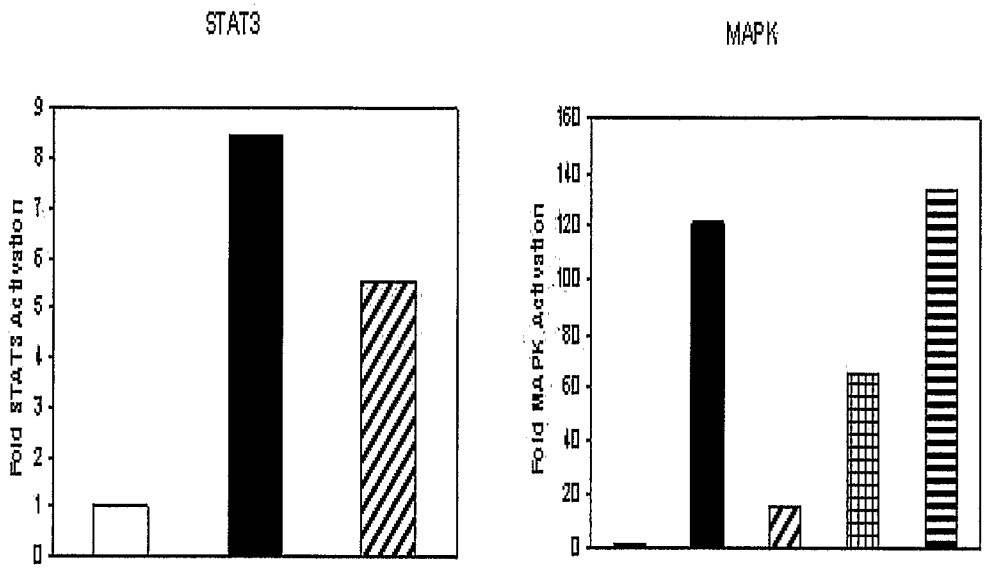

STAT3
Fold STAT3 activation
0
1
2
3
4
5
6
7
8
9
MAPK
Fold MAPK activation
0
20
40
60
80
100
120
140
160

Figure 10

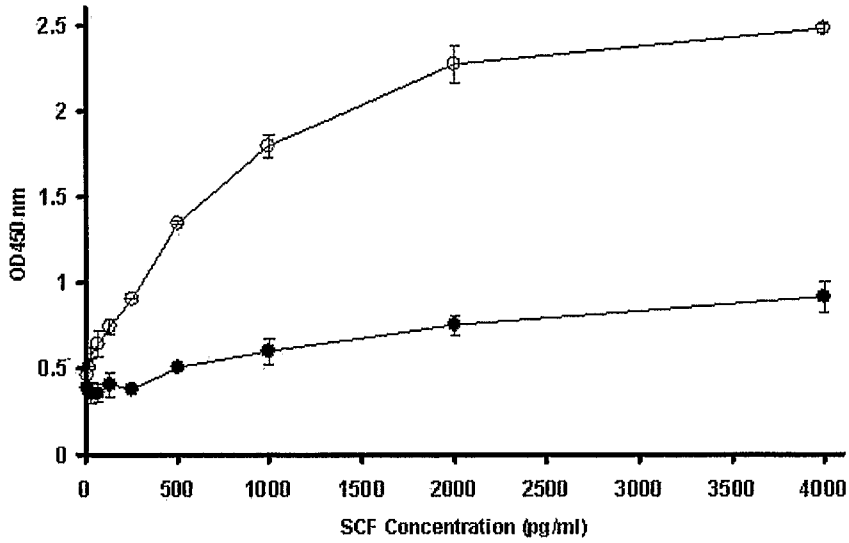

2.5
2
1.5
1
0.5
0
OD450 nm
0
500
1000
1500
2000
2500
3000
3500
4000
SCF Concentration (pg/ml)

Figure 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2007/000245

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

C07K 14/475 (2006.01) *C07K 14/71* (2006.01) *C07K 19/00* (2006.01) *C07K 16/28* (2006.01)
A61K 38/18 (2006.01) *C07K 16/22* (2006.01) *C12N 15/19* (2006.01) *A61P 35/00* (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
STN (CA, WPIDS, Medline, Biosis): Stem Cell Factor, Stem Cell Factor Receptor, HEK293, glycosylation

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 1998/017810 A2 G. D. SEARLE & CO. 30 April 1998 SEQ ID 465, Claim 1	1-16
X	WO 2003/038100 A1 CRUCELL HOLLAND B.V. 8 May 2003 Page 24 (last quarter of the page) and Claims 52, 53	1-3, 7, 8, 13-16
X	WO 2004/022593 A2 NAUTILIS BIOTECH 18 March 2004 Page 6 lines 4-24, page 98 lines 3-10, Claims 103-105	1-3, 7-16

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search
28 May 2007

Date of mailing of the international search report
13 JUL 2007

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
E-mail address: pct@ipaustralia.gov.au
Facsimile No. (02) 6285 3929

Authorized officer
Katherine Moerman
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No : (02) 6283 2714

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 1991/005795 A1 AMGEN INC. 2 May 1991 Example 11, 15	1-3, 7-16
X	Summerfield A. et al., 'C-kit positive porcine bone marrow progenitor cells identified and enriched using recombinant stem cell factor', Journal of Immunological Methods, 2003, 280(1-2), 113-123 Materials and Methods (2.1)	1-3
X	Langley K. E. et al., 'Soluble Stem Cell Factor in Human Serum', Blood, 1993, 81, 656-660 Abstract	1-3
X	Philo J. S. et al., 'Human Stem Cell Factor Dimer Forms a Complex with Two Molecules of the Extracellular Domain of its Receptor, Kit', Journal of Biological Chemistry, 1996, 271(12), 6895-902 Abstract, Page 6895-6896 and 6901	1-3
X	Blume-Jensen P. et al, 'The Kit receptor promotes cell survival via activation of PI 3-kinase and subsequent Akt- mediated phosphorylation of Bad on Ser136', Current Biology, 1998, 8(13), 779-782 Supplementary Materials – Material and Methods	1-3
X	Tsujimura T et al., 'Substitution of an Aspartic Acid Results in Constitutive Activation of c-kit receptor tyrosine kinase in a rat tumor mast cell line RBL-2H3', International Archives of Allergy and Immunology, 1995, 106(4), 377-385 Abstract, Discussion	1-3
X	Majumder S. et al., '*c-kit* Protein, a transmembrane Kinase: Identification in Tissues and Characterisation' Mol Cell Biol., 1988, 8(11), 4896-4903 Abstract, Discussion	1-3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2007/000245

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: **1-8, 13-16 (all in part)**
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
See Separate Sheet.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Multiple inventions are covered by Claims 1-16. In response to an Invitation to Pay Additional Fees, the Applicants elected to pay one (1) additional search fee and specifically requested that SCF and SCFR be searched.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☒ No protest accompanied the payment of additional search fees.

Supplemental Box
(To be used when the space in any of Boxes I to VIII is not sufficient)

Continuation of Box No: II

Claims 1-6 relate to an extremely large number of possible proteins as the physiochemical parameters place no limitation on the claims. Support within the meaning of Article 6 of the PCT and/or disclosure with the meaning of Article 5 of the PCT is to be found for only a very small proportion of the proteins claimed.

For this reason a meaningful search over the breadth of the claims is not possible.

In response to the Invitation to Pay Additional Fees, the Applicant elected to pay one (1) additional search fee and specifically requested that Stem Cell Factor (SCF) and Stem Cell Factor Receptor (SCFR) be searched. The search has been restricted accordingly.

Even with this restriction, a complete search was not possible. The prior art may not use the same parameters to describe the same proteins.

Please note that claims, or parts of claims, for which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT).

INTERNATIONAL SEARCH REPORT Information on patent family members	International application No. PCT/AU2007/000245

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member					
WO	1998/017810	AU	51652/98	BR	9713668	CA	2268742
		CZ	9901323	EP	0935663	KR	2000005281
		NO	991948	NZ	335382	PL	333023
		RU	2005105360	US	6967092	US	2007081979
		ZA	9709607				
WO	2003/038100	AU	2002224199	AU	2002225509	AU	2002307635
		AU	2002351444	BR	0213402	CA	2465007
		CA	2468957	CN	1578838	CN	1599794
		EP	1440157	EP	1465987	KR	2005004082
		KR	2005004467	KR	2006005764	MX	PA04003940
		NO	20042209	NZ	532438	NZ	533124
		US	6855544	US	7132280	US	2003092160
		US	2005106722	US	2005123564	US	2005164386
		US	2005164917	US	2005170463	US	2005181359
		US	2006099685	US	2007054394	US	2007117742
		WO	2003/048197	WO	2003/048348	WO	2003/050286
		WO	2003/089468	WO	2006/070011	ZA	200403209
WO	2004/022593	AU	2003263552	AU	2003267700	CA	2498284
		CA	2498319	EP	1539950	EP	1539960
		US	2004132977	US	2005202438	US	2006020116
		US	2006020396	WO	2004/022747		
WO	1991/005795	AU	17712/97	AU	60603/94	AU	65410/90
		CA	2026915	CA	2267626	CA	2267643
		CA	2267651	CA	2267658	CA	2267668
		CA	2267670	CA	2267671	CN	1051937
		CN	1306007	CZ	9700488	CZ	9700490
		CZ	9700491	EP	0423980	EP	0676470
		EP	0992579	EP	1241258	FI	20011804
		HK	1010397	HU	62011	IE	903562
		IE	20010893	IL	95905	LV	10462
		NO	912321	NO	964445	NO	982350
		NZ	235571	PT	95524	SG	43009

SG	59931	SK	98099	SK	98199
SK	500890	US	6204363	US	6207417
US	6207454	US	6207802	US	6218148
US	6248319	US	6759215	US	6841147
US	6852313	US	6967029	US	7144731
US	2002018763	US	2002031491	US	2004181044
US	2005080250	US	2005261175	US	2007071714
ZA	9007921				

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

END OF ANNEX