REPUBLIC OF SOUTH AFRICA
PATENTS ACT, 1978

FORM P.8
(to be lodged in duplicate)

A&~~A Ref: 155004~~

PUBLICATION PARTICULARS AND ABSTRACT

(Section 32(3)(a) - Regulations 22(1)(g) and 31)

06-02-2008

21	01	PATENT APPLICATION NO	22	LODGING DATE	43	ACCEPTANCE DATE 25-07-2007
		2006/04555		2 June 2006		24-1-07

51	INTERNATIONAL CLASSIFICATION	NOT FOR PUBLICATION
	A01N	CLASSIFIED BY: **ISA**

71	FULL NAME(S) OF APPLICANT(S)

Neose Technologies, Inc.

72	FULL NAME(S) OF INVENTOR(S)

DEFREES, Shawn
ZOPF, David A.
BOWE, Caryn
WU, Bingyuan
CLAUSEN, Henrik
WANG, Zhi-Guang
SCHWARTZ, Marc

EARLIEST PRIORITY CLAIMED	COUNTRY		NUMBER		DATE	
	33	**US**	31	**60/526,796**	32	**3 December 2003**

NOTE: The country must be indicated by its International Abbreviation - see schedule 4 of the Regulations

54	TITLE OF INVENTION

Glycopegylated granulocyte colony stimulating factor

57	ABSTRACT (NOT MORE THAN 150 WORDS)	NUMBER OF SHEETS 144

The sheet(s) containing the abstract is/are attached.

If no classification is furnished, Form P.9 should accompany this form.
~~The figure of the drawing to which the abstract refers is attached.~~

A&A P208

- 2006/04555

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 June 2005 (23.06.2005)

PCT

(10) International Publication Number
WO 2005/055946 A3

(51) **International Patent Classification:**
A01N 43/04 (2006.01)

(21) **International Application Number:**
PCT/US2004/041004

(22) **International Filing Date:**
3 December 2004 (03.12.2004)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**

60/526,796	3 December 2003 (03.12.2003)	US
60/539,387	26 January 2004 (26.01.2004)	US
60/555,813	23 March 2004 (23.03.2004)	US
60/570,282	11 May 2004 (11.05.2004)	US
60/592,744	29 July 2004 (29.07.2004)	US
60/614,518	29 September 2004 (29.09.2004)	US
60/623,387	29 October 2004 (29.10.2004)	US

(71) **Applicant** *(for all designated States except US)*: **NEOSE TECHNOLOGIES, INC.** [US/US]; 102 Witmer Road, Horsham, PA 19044 (US).

(72) **Inventors; and**

(75) **Inventors/Applicants** *(for US only)*: **DEFREES, Shawn** [US/US]; 126 Filly Drive, North Wales, PA 19454 (US). **CLAUSEN, Henrik** [DK/DK]; Soebrinken 6, DK-2840 Holte (DK). **ZOPF, David, A.** [US/US]; 560 Beechtree Lane, Wayne, PA 19087 (US). **WANG, Zhi-Guang** [US/US]; 1473 Golden Drive, Dresher, PA 19015 (US). **BOWE, Caryn** [US/US]; 276 Cherry Lane, Doylestown, PA 19044 (US).

(74) **Agents: MANN, Jeffry, S.** et al.; Morgan Lewis & Bockius LLP, 2 Palo Alto Square, Suite 700, 3000 El Camino Real, Palo Alto, CA 94304 (US).

(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, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, 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, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) **Date of publication of the international search report:**
30 March 2006

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: GLYCOPEGYLATED GRANULOCYTE COLONY STIMULATING FACTOR

(57) **Abstract:** The present invention provides conjugates between Granulocyte Colony Stimulating Factor and PEG moieties. The conjugates are linked via an intact glycosyl linking group that is interposed between and covalently attached to the peptide and the modifying group. The conjugates are formed from both glycosylated and unglycosylated peptides by the action of a glycosyltransferase. The glycosyltransferase ligates a modified sugar moiety onto either an amino acid or glycosyl residue on the peptide. Also provided are pharmaceutical formulations including the conjugates. Methods for preparing the conjugates are also within the scope of the invention.

WO 2005/055946 A3

GLYCOPEGYLATED GRANULOCYTE COLONY STIMULATING FACTOR

CROSS-REFERENCES TO RELATED APPLICATIONS

[0001] The present application claims priority to U.S. Provisional Patent Application No. 60/526,796, filed on December 3, 2003; U.S. Provisional Patent Application No. 60/555,813, filed March 23, 2004; U.S. Provisional Patent Application No. 60/570,282, filed May 11, 2004; U.S. Provisional Patent Application No. 60/539,387, filed January 26, 2004; U.S. Provisional Patent Application No. 60/592,744, filed July 29, 2004; U.S. Provisional Patent Application No. 60/614,518, filed September 29, 2004; and U.S. Provisional Patent Application No. 60/623,387, filed October 29, 2004 each of which is incorporated herein by reference in their entirety for all purposes.

BACKGROUND OF THE INVENTION

[0002] Granulocyte colony stimulating factor (G-CSF) is a glycoprotein which stimulates the survival, proliferation, differentiation and function of neutrophil granulocyte progenitor cells and mature neutrophils. The two forms of recombinant human G-CSF in clinical use are potent stimulants of neutrophil granulopoiesis and have demonstrated efficacy in preventing infectious complications of some neutropenic states. They can be used to accelerate neutrophil recovery from myelosuppressive treatments.

[0003] G-CSF decreases the morbidity of cancer chemotherapy by reducing the incidence of febrile neutropenia, the morbidity of high-dose chemotherapy supported by marrow transplantation, and the incidence and duration of infection in patients with severe chronic neutropenia. Further, G-CSF has recently been shown to have therapeutic when administered after the onset of myocardial infarction.

[0004] The human form of G-CSF was cloned by groups from Japan and the U.S.A. in 1986 (*see e.g.*, Nagata et al. *Nature* **319**: 415-418, 1986). The natural human glycoprotein exists in two forms, one of 175 and the other of 178 amino acids. The more abundant and more active 175 amino acid form has been used in the development of pharmaceutical products by recombinant DNA technology.

[0005] The recombinant human G-CSF synthesised in an *E. coli* expression system is called filgrastim. The structure of filgrastim differs slightly from the natural glycoprotein. The other form of recombinant human G-CSF is called *lenograstim* and is synthesised in Chinese hamster ovary (CHO) cells.

[0006] hG-CSF is a monomeric protein that dimerizes the G-CSF receptor by formation of a 2:2 complex of 2 G-CSF molecules and 2 receptors (Horan et al. *Biochemistry*, 35(15): 4886-96 (1996)). The following hG-CSF residues have been identified by X-ray crystalographic studies as being part of the receptor binding interfaces: G4, P5, A6, S7, S8, L9, P10, Q11, S12, L15, K16, E19, Q20, L108, D109, D112, T115, T116, Q119, E122, E123, and L124 (*see e.g.*, Aritomi et al., (1999) *Nature* **401**: 713).

[0007] The commercially available forms of rhG-CSF have a short-term pharmacological effect and must often be administered more once a day for the duration of the leukopenic state. A molecule with a longer circulation half-life would decrease the number of administrations necessary to alleviate the leukopenia and prevent consequent infections. Another problem with currently available rG-CSF products is the occurrence of dose-dependent bone pain. Since bone pain is experienced by patients as a significant side effect of treatment with rG-CSF, it would be desirable to provide a rG-CSF product that does not cause bone pain, either by means of a product that inherently does not have this effect or that is effective in a sufficiently small dose that no bone pain is caused. Thus, there is clearly a need for improved recombinant G-CSF molecules.

[0008] Protein-engineered variants of hG-CSF have been reported (U.S. Pat. No. 5,581,476, U.S. 5,214,132, U.S. 5,362,853, U.S. 4,904,584 and Riedhaar-Olson et al. Biochemistry 35: 9034-9041, 1996). Modification of hG-CSF and other polypeptides so as to introduce at least one additional carbohydrate chain as compared to the native polypeptide has also been reported (U.S. Pat. No. 5,218,092). In addition, polymer modifications of native hG-CSF, including attachment of PEG groups, have been reported and studied (*see e.g.*, Satake-Ishikawa et al., (1992) *Cell Structure and Function* **17**: 157; Bowen et al. (1999) *Experimental Hematology* **27**: 425; U.S. Pat. No. 5,824,778, U.S. 5,824,784, WO 96/11953, WO 95/21629, and WO 94/20069).

[0009] The attachment of synthetic polymers to the peptide backbone in an attempt to improve the pharmacokinetic properties of glycoprotein therapeutics is known in the art. An exemplary polymer that has been conjugated to peptides is poly(ethylene glycol) ("PEG"). The use of PEG to derivatize peptide therapeutics has been demonstrated to reduce the immunogenicity of the peptides. For example, U.S. Pat. No. 4,179,337 (Davis *et al.*) discloses non-immunogenic polypeptides such as enzymes and peptide hormones coupled to polyethylene glycol (PEG) or polypropylene glycol. In addition to reduced immunogenicity, the clearance time in circulation is prolonged due to the increased size of the PEG-conjugate of the polypeptides in question.

[0010] The principal mode of attachment of PEG, and its derivatives, to peptides is a non-specific bonding through a peptide amino acid residue (*see e.g.,* U.S. Patent No. 4,088,538 U.S. Patent No. 4,496,689, U.S. Patent No. 4,414,147, U.S. Patent No. 4,055,635, and PCT WO 87/00056). Another mode of attaching PEG to peptides is through the non-specific oxidation of glycosyl residues on a glycopeptide (*see e.g.*, WO 94/05332).

[0011] In these non-specific methods, poly(ethyleneglycol) is added in a random, non-specific manner to reactive residues on a peptide backbone. Of course, random addition of PEG molecules has its drawbacks, including a lack of homogeneity of the final product, and the possibility for reduction in the biological or enzymatic activity of the peptide. Therefore, for the production of therapeutic peptides, a derivitization strategy that results in the formation of a specifically labeled, readily characterizable, essentially homogeneous product is superior. Such methods have been developed.

[0012] Specifically labeled, homogeneous peptide therapeutics can be produced *in vitro* through the action of enzymes. Unlike the typical non-specific methods for attaching a synthetic polymer or other label to a peptide, enzyme-based syntheses have the advantages of regioselectivity and stereoselectivity. Two principal classes of enzymes for use in the synthesis of labeled peptides are glycosyltransferases (*e.g.*, sialyltransferases, oligosaccharyltransferases, N-acetylglucosaminyltransferases), and glycosidases. These enzymes can be used for the specific attachment of sugars which can be subsequently modified to comprise a therapeutic moiety. Alternatively, glycosyltransferases and modified glycosidases can be used to directly transfer modified sugars to a peptide backbone (*see e.g.*, U.S. Patent 6,399,336, and U.S. Patent Application Publications 20030040037, 20040132640, 20040137557, 20040126838, and 20040142856, each of which are incorporated by reference herein). Methods combining both chemical and enzymatic synthetic elements are also known (*see e.g.*, Yamamoto et al. *Carbohydr. Res.* **305**: 415-422 (1998) and U.S.

Patent Application Publication 20040137557 which is incorporated herein by reference).

[0013] In response to the need for improved therapeutic G-CSF, the present invention provides a glycopegylated G-CSF that is therapeutically active and which has pharmacokinetic parameters and properties that are improved relative to an identical, or closely analogous, G-CSF peptide that is not glycopegylated. Furthermore, the invention provides method for producing cost effectively and on an industrial scale the improved G-CSF peptides of the invention.

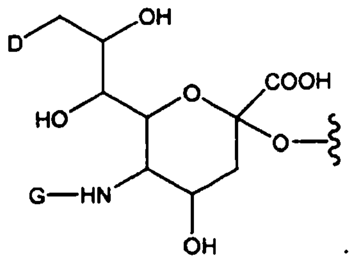
SUMMARY OF THE INVENTION

[0014] It has now been discovered that the controlled modification of Granulocyte colony stimulating factor (G-CSF) with one or more poly(ethylene glycol) moieties affords a novel G-CSF derivative with pharmacokinetic properties that are improved relative to the corresponding native (un-pegylated) G-CSF (**FIG. 3**). Moreover, the pharmacological activity of the glycopegylated G-CSF is approximately the same as the commercially available mono-pegylated filgrastim (**FIG. 4**).

[0015] In an exemplary embodiment, "glycopeglyated" G-CSF molecules of the invention are produced by the enzyme mediated formation of a conjugate between a glycosylated or non-glycosylated G-CSF peptide and an enzymatically transferable saccharyl moiety that includes a poly(ethylene glycol) moiety within its structure The PEG moiety is attached to the saccharyl moiety directly (i.e., through a single group formed by the reaction of two reactive groups) or through a linker moiety, e.g., substituted or unsubstituted alkyl, substituted or unsubstituted heteroalkyl, etc. An exemplary transferable PEG-saccharyl structure is set forth in **FIG. 5**.

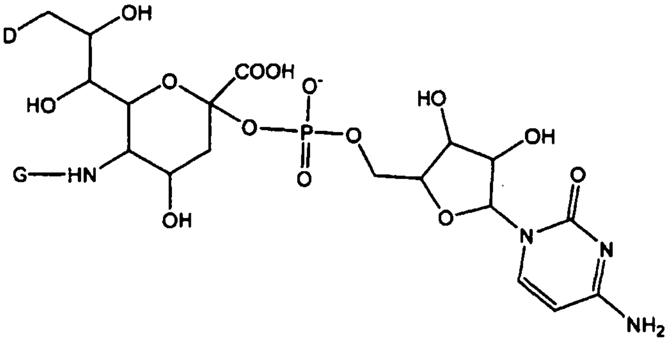
[0016] Thus, in one aspect, the present invention provides a conjugate between a PEG moiety, *e.g.*, PEG and a peptide that has an *in vivo* activity similar or otherwise analogous to art-recognized G-CSF. In the conjugate of the invention, the PEG moiety is covalently attached to the peptide via an intact glycosyl linking group. Exemplary intact glycosyl linking groups include sialic acid moieties that are derivatized with PEG.

[0017] In one exemplary aspect, the present invention provides a G-CSF peptide that includes the moiety:



[0018] In the formula above, D is -OH or R^1-L-HN-. The symbol G represents R^1-L- or -C(O)(C_1-C_6)alkyl. R^1 is a moiety comprising a straight-chain or branched poly(ethylene glycol) residue; and L is a linker which is a member selected from a bond, substituted or unsubstituted alkyl and substituted or unsubstituted heteroalkyl. Generally, when D is OH, G is R^1-L-, and when G is –C(O)(C_1-C_6)alkyl, D is R^1-L-NH-. In the modified sialic acid structures set forth herein, COOH also represents COO^- and/or a salt thereof.

[0019] In another aspect, the invention provides a method of making a PEG-ylated G-CSF comprising the moiety above. The method of the invention includes (a) contacting a substrate G-CSF peptide with a PEG-sialic acid donor and an enzyme that transfers the PEG-sialic acid onto an amino acid or glycosyl residue of the G-CSF, under conditions appropriate for the transfer. An exemplary PEG-sialic acid donor moiety has the formula:



[0020] In one embodiment the host is mammalian cell. In other embodiments the host cell is an insect cell, plant cell, a bacteria or a fungi.

[0021] The pharmacokinetic properties of the compounds of the invention are readily varied by altering the structure, number or position of the glycosylation site(s) of the peptide. Thus, it is within the purview of the present application to add one or more mutation that inserts an O- or N-linked glycosylation site into the G-CSF peptide that is not present in the wild type. Antibodies to these mutants and their glycosylated final products and intermediates are also within the scope of the present invention.

[0022] In another aspect, the invention provides a G-CSF conjugate having a population of PEG moiety moieties, *e.g.*, PEG, covalently bound thereto through an intact glycosyl linking group. In the conjugate of the invention, essentially each member of the population is bound via the glycosyl linking group to a glycosyl residue of the peptide, and each glycosyl residue has the same structure.

[0023] In exemplary embodiment, the present invention provides a G-CSF conjugate having a population of PEG moiety moieties, *e.g.*, PEG, covalently bound thereto through an intact glycosyl linking group. In the conjugate of the invention, essentially each member of the population is bound to an amino acid residue of the peptide, and each of the amino acid residues to which the polymer is bound has the same structure. For example, if one peptide includes an Thr linked glycosyl residue, at least about 70%, 80%, 90%, 95%, 97%, 99%, 99.2%, 99.4%, 99.6%, or more preferably 99.8% of the peptides in the population will have the same glycosyl residue covalently bound to the same Thr residue. The discussion above is equally relevant for both O-glycosylation and N-glycosylation sites.

[0024] Also provided is a pharmaceutical composition. The composition includes a pharmaceutically acceptable carrier and a covalent conjugate between a non-naturally-occurring, PEG moiety and a glycosylated or non-glycosylated G-CSF peptide.

[0025] Other objects and advantages of the invention will be apparent to those of skill in the art from the detailed description that follows.

DESCRIPTION OF THE DRAWINGS

[0026] **FIG. 1** is the structure of G-CSF, showing the presence and location of a potential glycosylation at Thr 133 (Thr 134 if a methionine is present).

[0027] **FIG. 2** is a scheme showing an exemplary embodiment of the invention in which a carbohydrate residue on a G-CSF peptide is remodeled by enzymatically adding a GalNAc moiety to the glycosyl residue at Thr 133 (Thr 134 is methionine is present) prior to adding a saccharyl moiety derivatized with PEG.

[0028] **FIG. 3** is a plot comparing the in vivo residence lifetimes of unglycosylated G-CSF, Neulasta™ and enzymatically glycopegylated G-CSF.

[0029] **FIG. 4** is a plot comparing the activities of the species shown in **FIG. 3**.

[0030] **FIG. 5** is a synthetic scheme for producing an exemplary PEG-glycosyl linking group precursor (modified sugar) of us in preparing the conjugates of the invention.

[0031] **FIG. 6** shows exemplary G-CSF amino acid sequences. SEQ ID NO:1 is the 175 amin cid variant, wherein the first amino acid is methionine and there is a threonine residue at Thr 134. SEQ ID NO:2 is a 174 amino acid variant which has the same sequence as the 175 amino acid variant execpt thet the leading methionine is missing, thus the sequence begins with T and there is a Threonine residue at position 133.

[0032] **FIG. 7** illustrates some exemplary modified sugar nucleotides useful in the practice of the invention.

[0033] **FIG. 8** illustrates further exemplary modified sugar nucleotides useful in the practice of the invention.

[0034] **FIG. 9** demonstrates production of recombinant GCSF in bacteria grown in various media and induced with IPTG.

[0035] **FIG. 10** provides Western blot analysis of refolded GCSF after SP-sepharose chromatography.

[0036] **FIG. 11** is a table of sialyl transferases that are of use for transferring to an acceptor the modified sialic acid species set forth herein and unmodified sialic acid.

DETAILED DESCRIPTION OF THE INVENTION AND THE PREFERRED EMBODIMENTS

Abbreviations

[0037] PEG, poly(ethyleneglycol); PPG, poly(propyleneglycol); Ara, arabinosyl; Fru, fructosyl; Fuc, fucosyl; Gal, galactosyl; GalNAc, N-acetylgalactosaminyl; Glc, glucosyl; GlcNAc, N-acetylglucosaminyl; Man, mannosyl; ManAc, mannosaminyl acetate; Xyl, xylosyl; and NeuAc, sialyl (N-acetylneuraminyl); M6P, mannose-6-phosphate; Sia, sialic acid, N-acetylneuraminyl, and derivatives and analogues thereof.

Definitions

[0038] Unless defined otherwise, all technical and scientific terms used herein generally have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Generally, the nomenclature used herein and the laboratory procedures in cell culture, molecular genetics, organic chemistry and nucleic acid chemistry and hybridization are those well known and commonly employed in the art. Standard techniques are used for nucleic acid and peptide synthesis. The techniques and procedures are generally performed according to conventional methods in the art and various general references (*see generally*, Sambrook *et al.* MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed. (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., which is incorporated herein by reference), which are provided throughout this document. The nomenclature used herein and the laboratory procedures in analytical chemistry, and organic synthetic described below are those well known and commonly employed in the art. Standard techniques, or modifications thereof, are used for chemical syntheses and chemical analyses.

[0039] All oligosaccharides described herein are described with the name or abbreviation for the non-reducing saccharide (*i.e.*, Gal), followed by the configuration of the glycosidic bond (α or β), the ring bond (1 or 2), the ring position of the reducing saccharide involved in the bond (2, 3, 4, 6 or 8), and then the name or abbreviation of the reducing saccharide (*i.e.*, GlcNAc). Each saccharide is preferably a pyranose. For a review of standard glycobiology nomenclature *see, Essentials of Glycobiology* Varki *et al.* eds. CSHL Press (1999).

[0040] Oligosaccharides are considered to have a reducing end and a non-reducing end, whether or not the saccharide at the reducing end is in fact a reducing sugar. In accordance with accepted nomenclature, oligosaccharides are depicted herein with the non-reducing end on the left and the reducing end on the right.

[0041] The term "sialic acid" refers to any member of a family of nine-carbon carboxylated sugars. The most common member of the sialic acid family is N-acetyl-neuraminic acid (2-keto-5-acetamido-3,5-dideoxy-D-glycero-D-galactononulopyranos-1-onic acid (often abbreviated as Neu5Ac, NeuAc, or NANA). A second member of the family is N-glycolyl-neuraminic acid (Neu5Gc or NeuGc), in which the N-acetyl group of NeuAc is hydroxylated. A third sialic acid family member is 2-keto-3-deoxy-nonulosonic acid (KDN) (Nadano *et al.* (1986) *J. Biol. Chem.* **261**: 11550-11557; Kanamori *et al., J. Biol. Chem.* **265**: 21811-21819 (1990)). Also included are 9-substituted sialic acids such as a 9-O-C_1-C_6 acyl-Neu5Ac like 9-O-lactyl-Neu5Ac or 9-O-acetyl-Neu5Ac, 9-deoxy-9-fluoro-Neu5Ac and 9-azido-9-deoxy-Neu5Ac. For review of the sialic acid family, *see, e.g.*, Varki, *Glycobiology* **2**: 25-40 (1992); *Sialic Acids: Chemistry, Metabolism and Function*, R. Schauer, Ed. (Springer-Verlag, New York (1992)). The synthesis and use of sialic acid compounds in a sialylation procedure is disclosed in international application WO 92/16640, published October 1, 1992.

[0042] The term "Granuloctye Colony Stimulating Factor" or "Granuloctye Colony Stimulating Factor peptide", or "G-CSF" or "G-CSF peptide" refers to any wild type or mutated peptide, recombinant, or native, or any fragment of G-CSF that has an activity that is or that mimics that of native GCSF. The term also generally encompasses non-peptide G-CSF mimetics. In an exemplary embodiment a G-CSF peptide has the amino acid sequence shown in SEQ ID NO:1. In other exemplary embodiments a G-CSF peptide has a sequence selected from SEQ ID NOs:3-11.

[0043] The term "Granuloctye Colony Stimulating Factor activity" refers to any activity including but not limited to, receptor binding and activation, inhibition of receptor binding, or any biochemical or physiological reaction that is normally affected by the action of wild-type Granuloctye Colony Stimulating Factor. Granuloctye Colony Stimulating Factor activity can arise from the action of any Granuloctye Colony Stimulating Factor peptide, as defined above.

[0044] "Peptide" refers to a polymer in which the monomers are amino acids and are joined together through amide bonds, alternatively referred to as a polypeptide. Additionally, unnatural amino acids, for example, β-alanine, phenylglycine and homoarginine are also included. Amino acids that are not gene-encoded may also be used in the present invention. Furthermore, amino acids that have been modified to include reactive groups, glycosylation sites, polymers, therapeutic moieties, biomolecules and the like may also be used in the invention. All of the amino acids used in the present invention may be either the D - or L -isomer. The L -isomer is generally preferred. In addition, other peptidomimetics are also useful in the present invention. As used herein, "peptide" refers to both glycosylated and unglycosylated peptides. Also included are petides that are incompletely glycosylated by a system that expresses the peptide. For a general review, *see,* Spatola, A. F., in CHEMISTRY AND BIOCHEMISTRY OF AMINO ACIDS, PEPTIDES AND PROTEINS, B. Weinstein, eds., Marcel Dekker, New York, p. 267 (1983).

[0045] The term “peptide conjugate,” refers to species of the invention in which a peptide is conjugated with a modified sugar as set forth herein.

[0046] The term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, *e.g.*, hydroxyproline, γ-carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, *i.e.*, an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, *e.g.*, homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (*e.g.*, norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that function in a manner similar to a naturally occurring amino acid. As used herein, “amino acid,” whether it is in a linker or a component of a peptide sequence refers to both the D- and L-isomer of the amino acid as well as mixtures of these two isomers.

[0047] As used herein, the term “modified sugar,” refers to a naturally- or non-naturally-occurring carbohydrate that is enzymatically added onto an amino acid or a glycosyl residue of a peptide in a process of the invention. The modified sugar is selected from a number of enzyme substrates including, but not limited to sugar nucleotides (mono-, di-, and tri-phosphates), activated sugars (*e.g.*, glycosyl halides, glycosyl mesylates) and sugars that are neither activated nor nucleotides. The “modified sugar” is covalently functionalized with a “modifying group.” Useful modifying groups include, but are not limited to, PEG moieties, therapeutic moieties, diagnostic moieties, biomolecules and the like. The modifying group is preferably not a naturally occurring, or an unmodified carbohydrate. The locus of functionalization with the modifying group is selected such that it does not prevent the “modified sugar” from being added enzymatically to a peptide.

[0048] The term "water-soluble" refers to moieties that have some detectable degree of solubility in water. Methods to detect and/or quantify water solubility are well known in the art. Exemplary PEG moieties include peptides, saccharides, poly(ethers), poly(amines), poly(carboxylic acids) and the like. Peptides can have mixed sequences of be composed of a single amino acid, *e.g.* poly(lysine). Similarly, saccharides can be of mixed sequence or composed of a single saccharide subunit, e.g, dextran, amylose, chitosan, and poly(sialic acid). An exemplary poly(ether) is poly(ethylene glycol). Poly(ethylene imine) is an exemplary polyamine, and poly(acrylic) acid is a representative poly(carboxylic acid)

[0049] The term, "glycosyl linking group," as used herein refers to a glycosyl residue to which an agent (*e.g.*, PEG moiety, therapeutic moiety, biomolecule) is covalently attached. In the methods of the invention, the "glycosyl linking group" becomes covalently attached to a glycosylated or unglycosylated peptide, thereby linking the agent to an amino acid and/or glycosyl residue on the peptide. A "glycosyl linking group" is generally derived from a "modified sugar" by the enzymatic attachment of the "modified sugar" to an amino acid and/or glycosyl residue of the peptide. An "intact glycosyl linking group" refers to a linking group that is derived from a glycosyl moiety in which the individual saccharide monomer that links the conjugate is not degraded, *e.g.*, oxidized, *e.g.*, by sodium metaperiodate. "Intact glycosyl linking groups" of the invention may be derived from a naturally occurring oligosaccharide by addition of glycosyl unit(s) or removal of one or more glycosyl unit from a parent saccharide structure.

[0050] The term "targeting moiety," as used herein, refers to species that will selectively localize in a particular tissue or region of the body. The localization is mediated by specific recognition of molecular determinants, molecular size of the targeting agent or conjugate, ionic interactions, hydrophobic interactions and the like. Other mechanisms of targeting an agent to a particular tissue or region are known to those of skill in the art. Exemplary targeting moieties include antibodies, antibody fragments, transferrin, HS-glycoprotein, coagulation factors, serum proteins, β-glycoprotein, G-CSF, GM-CSF, M-CSF, EPO and the like.

[0051] As used herein, "pharmaceutically acceptable carrier" includes any material, which when combined with the conjugate retains the conjugates' activity and is non-reactive with the subject's immune systems. Examples include, but are not limited to, any of the standard pharmaceutical carriers such as a phosphate buffered saline solution, water, emulsions such as oil/water emulsion, and various types of wetting agents. Other carriers may also include sterile solutions, tablets including coated tablets and capsules. Typically such carriers contain excipients such as starch, milk, sugar, certain types of clay, gelatin, stearic acid or salts thereof, magnesium or calcium stearate, talc, vegetable fats or oils, gums, glycols, or other known excipients. Such carriers may also include flavor and color additives or other ingredients. Compositions comprising such carriers are formulated by well known conventional methods.

[0052] As used herein, "administering," means oral administration, administration as a suppository, topical contact, intravenous, intraperitoneal, intramuscular, intralesional, intranasal or subcutaneous administration, or the implantation of a slow-release device *e.g.*, a mini-osmotic pump, to the subject. Adminsitration is by any route including parenteral, and transmucosal (*e.g.*, oral, nasal, vaginal, rectal, or transdermal). Parenteral administration includes, *e.g.*, intravenous, intramuscular, intra-arteriole, intradermal, subcutaneous, intraperitoneal, intraventricular, and intracranial. Moreover, where injection is to treat a tumor, *e.g.*, induce apoptosis, administration may be directly to the tumor and/or into tissues surrounding the tumor. Other modes of delivery include, but are not limited to, the use of liposomal formulations, intravenous infusion, transdermal patches, etc.

The term "ameliorating" or "ameliorate" refers to any indicia of success in the treatment of a pathology or condition, including any objective or subjective parameter such as abatement, remission or diminishing of symptoms or an improvement in a patient's physical or mental well-being. Amelioration of symptoms can be based on objective or subjective parameters; including the results of a physical examination and/or a psychiatric evaluation.

[0053] The term “therapy” refers to"treating" or "treatment" of a disease or condition including preventing the disease or condition from occurring in an animal that may be predisposed to the disease but does not yet experience or exhibit symptoms of the disease (prophylactic treatment), inhibiting the disease (slowing or arresting its development), providing relief from the symptoms or side-effects of the disease (including palliative treatment), and relieving the disease (causing regression of the disease).

[0054] The term “effective amount” or “an amount effective to”or a "therapeutically effective amount" or any gramatically equivalent term means the amount that, when administered to an animal for treating a disease, is sufficient to effect treatment for that disease.

[0055] The term “isolated” refers to a material that is substantially or essentially free from components, which are used to produce the material. For peptide conjugates of the invention, the term “isolated” refers to material that is substantially or essentially free from components, which normally accompany the material in the mixture used to prepare the peptide conjugate. “Isolated” and “pure” are used interchangeably. Typically, isolated peptide conjugates of the invention have a level of purity preferably expressed as a range. The lower end of the range of purity for the peptide conjugates is about 60%, about 70% or about 80% and the upper end of the range of purity is about 70%, about 80%, about 90% or more than about 90%.

[0056] When the peptide conjugates are more than about 90% pure, their purities are also preferably expressed as a range. The lower end of the range of purity is about 90%, about 92%, about 94%, about 96% or about 98%. The upper end of the range of purity is about 92%, about 94%, about 96%, about 98% or about 100% purity.

[0057] Purity is determined by any art-recognized method of analysis (*e.g.*, band intensity on a silver stained gel, polyacrylamide gel electrophoresis, HPLC, or a similar means).

[0058] “Essentially each member of the population,” as used herein, describes a characteristic of a population of peptide conjugates of the invention in which a selected percentage of the modified sugars added to a peptide are added to multiple, identical acceptor sites on the peptide. “Essentially each member of the population” speaks to the “homogeneity” of the sites on the peptide conjugated to a modified sugar and refers to conjugates of the invention, which are at least about 80%, preferably at least about 90% and more preferably at least about 95% homogenous.

[0059] “Homogeneity,” refers to the structural consistency across a population of acceptor moieties to which the modified sugars are conjugated. Thus, in a peptide conjugate of the invention in which each modified sugar moiety is conjugated to an acceptor site having the same structure as the acceptor site to which every other modified sugar is conjugated, the peptide conjugate is said to be about 100% homogeneous. Homogeneity is typically expressed as a range. The lower end of the range of homogeneity for the peptide conjugates is about 60%, about 70% or about 80% and the upper end of the range of purity is about 70%, about 80%, about 90% or more than about 90%.

[0060] When the peptide conjugates are more than or equal to about 90% homogeneous, their homogeneity is also preferably expressed as a range. The lower end of the range of homogeneity is about 90%, about 92%, about 94%, about 96% or about 98%. The upper end of the range of purity is about 92%, about 94%, about 96%, about 98% or about 100% homogeneity. The purity of the peptide conjugates is typically determined by one or more methods known to those of skill in the art, *e.g.*, liquid chromatography-mass spectrometry (LC-MS), matrix assisted laser desorption mass time of flight spectrometry (MALDITOF), capillary electrophoresis, and the like.

[0061] "Substantially uniform glycoform" or a "substantially uniform glycosylation pattern," when referring to a glycopeptide species, refers to the percentage of acceptor moieties that are glycosylated by the glycosyltransferase of interest (*e.g.*, fucosyltransferase). For example, in the case of a α1,2 fucosyltransferase, a substantially uniform fucosylation pattern exists if substantially all (as defined below) of the Galβ1,4-GlcNAc-R and sialylated analogues thereof are fucosylated in a peptide conjugate of the invention. It will be understood by one of skill in the art, that the starting material may contain glycosylated acceptor moieties (*e.g.*, fucosylated Galβ1,4-GlcNAc-R moieties). Thus, the calculated percent glycosylation will include acceptor moieties that are glycosylated by the methods of the invention, as well as those acceptor moieties already glycosylated in the starting material.

[0062] The term "substantially" in the above definitions of "substantially uniform" generally means at least about 40%, at least about 70%, at least about 80%, or more preferably at least about 90%, and still more preferably at least about 95% of the acceptor moieties for a particular glycosyltransferase are glycosylated.

[0063] Where substituent groups are specified by their conventional chemical formulae, written from left to right, they equally encompass the chemically identical substituents, which would result from writing the structure from right to left, *e.g.*, $-CH_2O-$ is intended to also recite $-OCH_2-$.

[0064] The term "alkyl," by itself or as part of another substituent means, unless otherwise stated, a straight or branched chain, or cyclic hydrocarbon radical, or combination thereof, which may be fully saturated, mono- or polyunsaturated and can include di- and multivalent radicals, having the number of carbon atoms designated (*i.e.* C_1-C_{10} means one to ten carbons). Examples of saturated hydrocarbon radicals include, but are not limited to, groups such as methyl, ethyl, n-propyl, isopropyl, n-butyl, t-butyl, isobutyl, sec-butyl, cyclohexyl, (cyclohexyl)methyl, cyclopropylmethyl, homologs and isomers of, for example, n-pentyl, n-hexyl, n-heptyl, n-octyl, and the like. An unsaturated alkyl group is one having one or more double bonds or triple bonds. Examples of unsaturated alkyl groups include, but are not limited to, vinyl, 2-propenyl, crotyl, 2-isopentenyl, 2-(butadienyl), 2,4-pentadienyl, 3-(1,4-pentadienyl), ethynyl, 1- and 3-propynyl, 3-butynyl, and the higher homologs and isomers. The term "alkyl," unless otherwise noted, is also meant to include those derivatives of alkyl defined in more detail below, such as "heteroalkyl." Alkyl groups that are limited to hydrocarbon groups are termed "homoalkyl".

[0065] The term "alkylene" by itself or as part of another substituent means a divalent radical derived from an alkane, as exemplified, but not limited, by $-CH_2CH_2CH_2CH_2-$, and further includes those groups described below as "heteroalkylene." Typically, an alkyl (or alkylene) group will have from 1 to 24 carbon atoms, with those groups having 10 or fewer carbon atoms being preferred in the present invention. A "lower alkyl" or "lower alkylene" is a shorter chain alkyl or alkylene group, generally having eight or fewer carbon atoms.

[0066] The terms "alkoxy," "alkylamino" and "alkylthio" (or thioalkoxy) are used in their conventional sense, and refer to those alkyl groups attached to the remainder of the molecule via an oxygen atom, an amino group, or a sulfur atom, respectively.

[0067] The term "heteroalkyl," by itself or in combination with another term, means, unless otherwise stated, a stable straight or branched chain, or cyclic hydrocarbon radical, or combinations thereof, consisting of the stated number of carbon atoms and at least one heteroatom selected from the group consisting of O, N, Si and S, and wherein the nitrogen and sulfur atoms may optionally be oxidized and the nitrogen heteroatom may optionally be quaternized. The heteroatom(s) O, N and S and Si may be placed at any interior position of the heteroalkyl group or at the position at which the alkyl group is attached to the remainder of the molecule. Examples include, but are not limited to, $-CH_2-CH_2-O-CH_3$, $-CH_2-CH_2-NH-CH_3$, $-CH_2-CH_2-N(CH_3)-CH_3$, $-CH_2-S-CH_2-CH_3$, $-CH_2-CH_2,-S(O)-CH_3$, $-CH_2-CH_2-S(O)_2-CH_3$, $-CH=CH-O-CH_3$, -

$Si(CH_3)_3$, $-CH_2-CH=N-OCH_3$, and $-CH=CH-N(CH_3)-CH_3$. Up to two heteroatoms may be consecutive, such as, for example, $-CH_2-NH-OCH_3$ and $-CH_2-O-Si(CH_3)_3$. Similarly, the term "heteroalkylene" by itself or as part of another substituent means a divalent radical derived from heteroalkyl, as exemplified, but not limited by, $-CH_2-CH_2-S-CH_2-CH_2-$ and $-CH_2-S-CH_2-CH_2-NH-CH_2-$. For heteroalkylene groups, heteroatoms can also occupy either or both of the chain termini (*e.g.*, alkyleneoxy, alkylenedioxy, alkyleneamino, alkylenediamino, and the like). Still further, for alkylene and heteroalkylene linking groups, no orientation of the linking group is implied by the direction in which the formula of the linking group is written. For example, the formula $-C(O)_2R'-$ represents both $-C(O)_2R'-$ and $-R'C(O)_2-$.

[0068] The terms "cycloalkyl" and "heterocycloalkyl", by themselves or in combination with other terms, represent, unless otherwise stated, cyclic versions of "alkyl" and "heteroalkyl", respectively. Additionally, for heterocycloalkyl, a heteroatom can occupy the position at which the heterocycle is attached to the remainder of the molecule. Examples of cycloalkyl include, but are not limited to, cyclopentyl, cyclohexyl, 1-cyclohexenyl, 3-cyclohexenyl, cycloheptyl, and the like. Examples of heterocycloalkyl include, but are not limited to, 1 –(1,2,5,6-tetrahydropyridyl), 1-piperidinyl, 2-piperidinyl, 3-piperidinyl, 4-morpholinyl, 3-morpholinyl, tetrahydrofuran-2-yl, tetrahydrofuran-3-yl, tetrahydrothien-2-yl, tetrahydrothien-3-yl, 1 –piperazinyl, 2-piperazinyl, and the like.

[0069] The terms "halo" or "halogen," by themselves or as part of another substituent, mean, unless otherwise stated, a fluorine, chlorine, bromine, or iodine atom. Additionally, terms such as "haloalkyl," are meant to include monohaloalkyl and polyhaloalkyl. For example, the term "halo(C_1-C_4)alkyl" is mean to include, but not be limited to, trifluoromethyl, 2,2,2-trifluoroethyl, 4-chlorobutyl, 3-bromopropyl, and the like.

[0070] The term "aryl" means, unless otherwise stated, a polyunsaturated, aromatic, substituent that can be a single ring or multiple rings (preferably from 1 to 3 rings), which are fused together or linked covalently. The term "heteroaryl" refers to aryl groups (or rings) that contain from one to four heteroatoms selected from N, O, and S, wherein the nitrogen and sulfur atoms are optionally oxidized, and the nitrogen atom(s) are optionally quaternized. A heteroaryl group can be attached to the remainder of the molecule through a heteroatom. Non-limiting examples of aryl and heteroaryl groups include phenyl, 1-naphthyl, 2-naphthyl, 4-biphenyl, 1-pyrrolyl, 2-

pyrrolyl, 3-pyrrolyl, 3-pyrazolyl, 2-imidazolyl, 4-imidazolyl, pyrazinyl, 2-oxazolyl, 4-oxazolyl, 2-phenyl-4-oxazolyl, 5-oxazolyl, 3-isoxazolyl, 4-isoxazolyl, 5-isoxazolyl, 2-thiazolyl, 4-thiazolyl, 5-thiazolyl, 2-furyl, 3-furyl, 2-thienyl, 3-thienyl, 2-pyridyl, 3-pyridyl, 4-pyridyl, 2-pyrimidyl, 4-pyrimidyl, 5-benzothiazolyl, purinyl, 2-benzimidazolyl, 5-indolyl, 1-isoquinolyl, 5-isoquinolyl, 2-quinoxalinyl, 5-quinoxalinyl, 3-quinolyl, tetrazolyl, benzo[b]furanyl, benzo[b]thienyl, 2,3-dihydrobenzo[1,4]dioxin-6-yl, benzo[1,3]dioxol-5-yl and 6-quinolyl. Substituents for each of the above noted aryl and heteroaryl ring systems are selected from the group of acceptable substituents described below.

[0071] For brevity, the term "aryl" when used in combination with other terms (*e.g.*, aryloxy, arylthioxy, arylalkyl) includes both aryl and heteroaryl rings as defined above. Thus, the term "arylalkyl" is meant to include those radicals in which an aryl group is attached to an alkyl group (*e.g.*, benzyl, phenethyl, pyridylmethyl and the like) including those alkyl groups in which a carbon atom (*e.g.*, a methylene group) has been replaced by, for example, an oxygen atom (*e.g.*, phenoxymethyl, 2-pyridyloxymethyl, 3-(1-naphthyloxy)propyl, and the like).

[0072] Each of the above terms (*e.g.*, "alkyl," "heteroalkyl," "aryl" and "heteroaryl") is meant to include both substituted and unsubstituted forms of the indicated radical. Preferred substituents for each type of radical are provided below.

[0073] Substituents for the alkyl and heteroalkyl radicals (including those groups often referred to as alkylene, alkenyl, heteroalkylene, heteroalkenyl, alkynyl, cycloalkyl, heterocycloalkyl, cycloalkenyl, and heterocycloalkenyl) are generically referred to as "alkyl group substituents," and they can be one or more of a variety of groups selected from, but not limited to: -OR', =O, =NR', =N-OR', -NR'R", -SR', -halogen, -SiR'R"R"', -OC(O)R', -C(O)R', -CO_2R', -CONR'R", -OC(O)NR'R", -NR"C(O)R', -NR'-C(O)NR"R"', -NR"$C(O)_2R'$, -NR-C(NR'R"R"")=NR"", -NR-C(NR'R")=NR"', -S(O)R', -$S(O)_2R'$, -$S(O)_2NR'R"$, -$NRSO_2R'$, -CN and –NO_2 in a number ranging from zero to (2m'+1), where m' is the total number of carbon atoms in such radical. R', R", R"' and R"" each preferably independently refer to hydrogen, substituted or unsubstituted heteroalkyl, substituted or unsubstituted aryl, *e.g.*, aryl substituted with 1-3 halogens, substituted or unsubstituted alkyl, alkoxy or thioalkoxy groups, or arylalkyl groups. When a compound of the invention includes more than one R group, for example, each of the R groups is independently selected as are each R', R", R"' and R"" groups when more than one of these groups is present. When R' and R" are attached to the same nitrogen atom, they can be combined with the nitrogen atom to form a 5-, 6-, or 7-membered ring. For example, -NR'R" is meant to include, but not be limited to, 1-pyrrolidinyl and 4-morpholinyl. From the above discussion of substituents, one of skill in the art will understand that the term "alkyl" is meant to include groups including carbon atoms bound to groups other than hydrogen groups, such as haloalkyl (*e.g.*, $-CF_3$ and $-CH_2CF_3$) and acyl (*e.g.*, $-C(O)CH_3$, $-C(O)CF_3$, $-C(O)CH_2OCH_3$, and the like).

[0074] Similar to the substituents described for the alkyl radical, substituents for the aryl and heteroaryl groups are generically referred to as "aryl group substituents." The substituents are selected from, for example: halogen, -OR', =O, =NR', =N-OR', -NR'R", -SR', -halogen, -SiR'R"R"', -OC(O)R', -C(O)R', $-CO_2R'$, -CONR'R", -OC(O)NR'R", -NR"C(O)R', -NR'-C(O)NR"R"', $-NR"C(O)_2R'$, -NR-C(NR'R"R"')=NR"", -NR-C(NR'R")=NR"', -S(O)R', $-S(O)_2R'$, $-S(O)_2NR'R"$, $-NRSO_2R'$, -CN and $-NO_2$, -R', $-N_3$, $-CH(Ph)_2$, fluoro(C_1-C_4)alkoxy, and fluoro(C_1-C_4)alkyl, in a number ranging from zero to the total number of open valences on the aromatic ring system; and where R', R", R"' and R"" are preferably independently selected from hydrogen, substituted or unsubstituted alkyl, substituted or unsubstituted heteroalkyl, substituted or unsubstituted aryl and substituted or unsubstituted heteroaryl. When a compound of the invention includes more than one R group, for example, each of the R groups is independently selected as are each R', R", R"' and R"" groups when more than one of these groups is present. In the schemes that follow, the symbol X represents "R" as described above.

[0075] Two of the substituents on adjacent atoms of the aryl or heteroaryl ring may optionally be replaced with a substituent of the formula $-T-C(O)-(CRR')_q-U-$, wherein T and U are independently -NR-, -O-, -CRR'- or a single bond, and q is an integer of from 0 to 3. Alternatively, two of the substituents on adjacent atoms of the aryl or heteroaryl ring may optionally be replaced with a substituent of the formula $-A-(CH_2)_r-B-$, wherein A and B are independently -CRR'-, -O-, -NR-, -S-, -S(O)-, $-S(O)_2-$, $-S(O)_2NR'-$ or a single bond, and r is an integer of from 1 to 4. One of the single bonds of the new ring so formed may optionally be replaced with a double bond. Alternatively, two of the substituents on adjacent atoms of the aryl or heteroaryl ring may optionally be replaced with a substituent of the formula $-(CRR')_s-X-(CR"R"')_d-$, where s and d are independently integers of from 0 to 3, and X is -O-, -NR'-, -S-, -S(O)-, $-S(O)_2-$, or $-S(O)_2NR'-$. The substituents R, R', R" and

R''' are preferably independently selected from hydrogen or substituted or unsubstituted (C_1-C_6)alkyl.

[0076] As used herein, the term "heteroatom" is meant to include oxygen (O), nitrogen (N), sulfur (S) and silicon (Si).

Introduction

[0077] The present invention encompasses a method for the modification of the glycan structure on G-CSF. G-CSF is well known in the art as a cytokine produced by activated T-cells, macrophages, endothelial cells, and stromal fibroblasts. G-CSF primarily acts on the bone marrow to increase the production of inflammatory leukocytes, and further functions as an endocrine hormone to initiate the replenishment of neutrophils consumed during inflammatory functions. G-CSF also has clinical applications in bone marrow replacement following chemotherapy.

[0078] The present invention provides a conjugate of granulocyte colony stimulating factor (G-CSF). The invention provides conjugates of glycosylated and unglycosylated peptides having granulocyte colony stimulating activity. The conjugates may be additionally modified by further conjugation with diverse species such as therapeutic moieties, diagnostic moieties, targeting moieties and the like.

[0079] The present invention further includes a method for remodeling and/or modifying G-CSF. G-CSF is a valuable tool in treatment of numerous diseases, but as stated above, its clinical efficacy has been hampered by its relatively poor pharmacokinetics

[0080] In exemplary embodiments, a G-CSF peptide of the invention may be administered to patients for the purposed of preventing infection in cancer patients undergoing certain types of radiation therapy, chemotherapy, and bone marrow transplantations, to mobilize progenitor cells for collection in peripheral blood progenitor cell transplantations, for treatment of severe chronic or relative leukopenia, irrespective of cause, and to support treatment of patients with acute myeloid leukaemia. Additionally, the polypeptide conjugate or composition of the invention may be used for treatment of AIDS or other immunodeficiency diseases as well as bacterial infections.

[0081] G-CSF has been cloned and sequenced. In an exemplary embodiment, G-CSF has an the amino acid sequence according to SEQ ID NO:1. The skilled artisan will readily appreciate that the present invention is not limited to the sequences depicted herein, as variants of G-CSF, as discussed hereinabove.

[0082] Thus, the present invention further encompasses G-CSF variants, as well known in the art. As an example, but in no way meant to be limiting to the present invention, a G-CSF variant has been described in U.S. Patent No. 6,166,183, in which a G-CSF comprising the natural complement of lysine residues and further linked to one or two polyethylene glycol molecules is described. Additionally, U.S. Patent Nos. 6,004,548, 5,580,755, 5,582,823, and 5,676,941 describe a G-CSF variant in which one or more of the cysteine residues at position 17, 36, 42, 64, and 74 are replaced by alanine or alternatively serine. U.S. Patent No. 5,416,195 describes a G-CSF molecule in which the cysteine at position 17, the aspartic acid at position 27, and the serines at positions 65 and 66 are substituted with serine, serine, proline, and proline, respectively. Other variants are well known in the art, and are described in, for example, U.S. Patent No. 5,399,345. Still further variants have an amino acid selected from SEQ ID Nos:3-11.

[0083] The expression and activity of a modified G-CSF molecule of the present invention can be assayed using methods well known in the art, and as described in, for example, U.S. Patent No. 4,810,643. As an example, activity can be measured using radio-labeled thymidine uptake assays. Briefly, human bone marrow from healthy donors is subjected to a density cut with Ficoll-Hypaque (1.077 g/ml, Pharmacia, Piscataway, NJ) and low density cells are suspended in Iscove's medium (GIBCO, La Jolla, CA) containing 10% fetal bovine serum, glutamine and antibiotics. About 2 x 10^4 human bone marrow cells are incubated with either control medium or the G-CSF or the present invention in 96-well flat bottom plates at about 37° C in 5% CO_2 in air for about 2 days. Cultures are then pulsed for about 4 hours with 0.5 µCi/well of ^3H-thymidine (New England Nuclear, Boston, Mass.) and uptake is measured as described in, for example, Ventua, et al.(1983, Blood 61:781). An increase in ^3H-thymidine incorporation into human bone marrow cells as compared to bone marrow cells treated with a control compound is an indication of a active and viable G-CSF compound.

[0084] As discussed above, the conjugates of the invention are formed by the enzymatic attachment of a modified sugar to the glycosylated or unglycosylated G-CSF peptide. The modified sugar, when interposed between the G-CSF peptide and the modifying group on the sugar becomes what may be referred to herein *e.g.*, as an "intact glycosyl linking group." Using the exquisite selectivity of enzymes, such as glycosyltransferases, the present method provides peptides that bear a desired group at one or more specific locations. Thus, according to the present invention, a modified sugar is attached directly to a selected locus on the G-CSF peptide chain or, alternatively, the modified sugar is appended onto a carbohydrate moiety of a glycopeptide. Peptides in which modified sugars are bound to both a glycopeptide carbohydrate and directly to an amino acid residue of the G-CSF peptide backbone are also within the scope of the present invention.

[0085] In contrast to known chemical and enzymatic peptide elaboration strategies, the methods of the invention, make it possible to assemble peptides and glycopeptides that have a substantially homogeneous derivatization pattern; the enzymes used in the invention are generally selective for a particular amino acid residue or combination of amino acid residues of the G-CSF peptide. The methods are also practical for large-scale production of modified peptides and glycopeptides. Thus, the methods of the invention provide a practical means for large-scale preparation of glycopeptides having preselected uniform derivatization patterns. The methods are particularly well suited for modification of therapeutic peptides, including but not limited to, glycopeptides that are incompletely glycosylated during production in cell culture cells (*e.g.*, mammalian cells, insect cells, plant cells, fungal cells, yeast cells, or prokaryotic cells) or transgenic plants or animals.

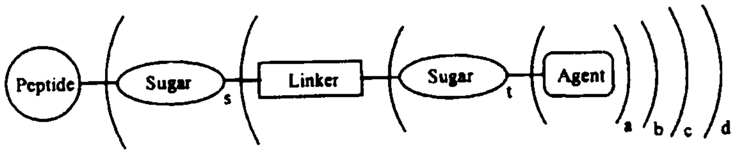
[0086] The present invention also provides conjugates of glycosylated and unglycosylated G-CSF peptides with increased therapeutic half-life due to, for example, reduced clearance rate, or reduced rate of uptake by the immune or reticuloendothelial system (RES). Moreover, the methods of the invention provide a means for masking antigenic determinants on peptides, thus reducing or eliminating a host immune response against the peptide. Selective attachment of targeting agents can also be used to target a peptide to a particular tissue or cell surface receptor that is specific for the particular targeting agent.

The Conjugates

[0087] In a first aspect, the present invention provides a conjugate between a selected modifying group and a G-CSF peptide.

[0088] The link between the G-CSF peptide and the selected moiety includes an intact glycosyl linking group interposed between the peptide and the selected moiety. As discussed herein, the selected moiety is essentially any species that can be attached to a saccharide unit, resulting in a "modified sugar" that is recognized by an appropriate transferase enzyme, which appends the modified sugar onto the G-CSF peptide. The saccharide component of the modified sugar, when interposed between the G-CSF peptide and a selected moiety, becomes an "intact glycosyl linking group." The glycosyl linking group is formed from any mono- or oligo-saccharide that, after modification with a selected moiety, is a substrate for an appropriate transferase.

[0089] The conjugates of the invention will typically correspond to the general structure:



in which the symbols a, b, c, d and s represent a positive, non-zero integer; and t is either 0 or a positive integer. The "agent" is typically a water-soluable moiety, e.g., a PEG moiety. The linker can be any of a wide array of linking groups, *infra*. Alternatively, the linker may be a single bond or a "zero order linker."

[0090] In an exemplary embodiment, the selected modifying group is a water-soluble polymer, *e.g.*, m-PEG. The water-soluble polymer is covalently attached to the G-CSF peptide via a glycosyl linking group, which is covalently attached to an amino acid residue or a glycosyl residue of the G-CSF peptide. The invention also provides conjugates in which an amino acid residue and a glycosyl residue are modified with a glycosyl linking group.

[0091] An exemplary water-soluble polymer is poly(ethylene glycol), *e.g.*, methoxy-poly(ethylene glycol). The poly(ethylene glycol) used in the present invention is not restricted to any particular form or molecular weight range. For unbranched poly(ethylene glycol) molecules the molecular weight is preferably between 500 and 100,000. A molecular weight of 2,000-60,000 is preferably used and more preferably of from about 5,000 to about 30,000.

[0092] In another embodiment the poly(ethylene glycol) is a branched PEG having more than one PEG moiety attached. Examples of branched PEGs are described in U.S. Pat. No. 5,932,462; U.S. Pat. No. 5,342,940; U.S. Pat. No. 5,643,575; U.S. Pat.

No. 5,919,455; U.S. Pat. No. 6,113,906; U.S. Pat. No. 5,183,660; WO 02/09766; Kodera Y., *Bioconjugate Chemistry* 5: 283-288 (1994); and Yamasaki et al., *Agric. Biol. Chem.*, **52**: 2125-2127, 1998. Other useful branched PEG structures are disclosed herein.

[0093] In an exemplary embodiment the molecular weight of each poly(ethylene glycol) of the branched PEG is equal to or greater than about 2,000, 5,000, 10,000, 15,000, 20,000, 40,000 or 60,000 daltons.

[0094] The peptides of the present invention include at least on N- or O-linked glycosylation site. In addition to providing conjugates that are formed through an enzymatically added glycosyl linking group, the present invention provides conjugates that are highly homogenous in their substitution patterns. Using the methods of the invention, it is possible to form peptide conjugates in which essentially all of the modified sugar moieties across a population of conjugates of the invention are attached to multiple copies of a structurally identical amino acid or glycosyl residue. Thus, in a second aspect, the invention provides a peptide conjugate having a population of water-soluble polymer moieties, which are covalently bound to the G-CSF peptide through an intact glycosyl linking group. In a preferred conjugate of the invention, essentially each member of the population is bound via the glycosyl linking group to a glycosyl residue of the G-CSF peptide, and each glycosyl residue of the G-CSF peptide to which the glycosyl linking group is attached has the same structure.

[0095] Also provided is a peptide conjugate having a population of water-soluble polymer moieties covalently bound thereto through a glycosyl linking group. In a preferred embodiment, essentially every member of the population of water soluble polymer moieties is bound to an amino acid residue of the G-CSF peptide via a glycosyl linking group, and each amino acid residue having a glycosyl linking group attached thereto has the same structure.

[0096] The present invention also provides conjugates analogous to those described above in which the G-CSF peptide is conjugated to a therapeutic moiety, diagnostic moiety, targeting moiety, toxin moiety or the like via an intact glycosyl linking group. Each of the above-recited moieties can be a small molecule, natural polymer (*e.g.*, polypeptide) or synthetic polymer.

[0097] Essentially any Granulocyte Colony Stimulating Factor peptide or agent, having any sequence, is of use as the peptide component of the conjugates of the present invention. Granulocyte Colony Stimulating Factor has been cloned and sequenced. In an exemplary embodiment, the G-CSF peptide has the sequence presented in SEQ ID NO:1:

MTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKL
CATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLA
GCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLD
VADFATTIWQQMEELGMAPALQPTQGAMPAFASAF
QRRAGGVLVASHLQSFLEVSYRVLRHLAQP **(SEQ ID NO: 1)**.

[0098] In another exemplary embodiment, the G-CSF peptide has the sequence presented in SEQ ID NO:2:

TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCA
TYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLAGC
LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVA
DFATTIWQQMEELGMAPALQPTQGAMPAFASAFQR
RAGGVLVASHLQSFLEVSYRVLRHLAQP **(SEQ ID NO: 2)**.

[0099] In other exemplary embodiments, the G-CSF peptide has a sequence presented in SEQ ID Nos:3-11, below.

MTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKL
VSECATYKLCHPEELVLLGHSLGIPWAPLSSCPSQAL
QLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTL
QLDVADFATTIWQQMEELGMAPALQPTQGAMPAFA
SAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP
(SEQ ID NO:3)

MAGPATQSPMKLMALQLLLWHSALWTVQEATPLGP
ASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKL
CHPEELVLLGHSLGIPWAPLSSCPSQALQLAGCLSQL
HSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFAT

TIWQQMEELGMAPALQPTQGAMPAFASAFQRRAGG
VLVASHLQSFLEVSYRVLRHLAQP **(SEQ ID NO:4)**

MAGPATQSPMKLMALQLLLWHSALWTVQEATPLGP
ASSLPQSFLLKCLEQVRKIQGDGAALQEKLVSECAT
YKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLAGCL
SQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVAD
FATTIWQQMEELGMAPALQPTQGAMPAFASAFQRR
AGGVLVASHLQSFLEVSYRVLRHLAQP **(SEQ ID NO:5)**

MVTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEK
LCATYKLCHPEELVLLGHTLGIPWAPLSSCPSQALQL
AGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQL
DVADFATTIWQQMEELGMAPALQPTQGAMPAFASA
FQRRAGGVLVASHLQSFLEVSYRVLRHLAQP**(SEQ ID NO:6)**;

MTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKL
CATYKLCHPEELVLLGHTLGIPWAPLSSCPSQALQLA
GCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLD
VADFATTIWQQMEELGMAPALQPTQGAMPAFASAF
QRRAGGVLVASHLQSFLEVSYRVLRHLAQP**(SEQ ID NO:7)**;

MVTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEK
LCATYKLCHPEELVLLGSSLGIPWAPLSSCPSQALQL
AGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQL
DVADFATTIWQQMEELGMAPALQPTQGAMPAFASA
FQRRAGGVLVASHLQSFLEVSYRVLRHLAQP**(SEQ ID NO:8)**;

MQTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEK
LCATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQL
AGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQL
DVADFATTIWQQMEELGMAPALQPTQGAMPAFASA
FQRRAGGVLVASHLQSFLEVSYRVLRHLAQP(**SEQ
ID NO:9**);

MTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKL
CATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLA
GCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLD
VADFATTIWQQMEELGMAPALQPTQGAMPAFASAF
QRRAGGVLVASHLQSFLEVSYRVLRHLAQPTQGAM
P; (**SEQ ID NO:10**) and

MTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKL
CATYKLCHPEELVLLGSSLGIPWAPLSSCPSQALQLA
GCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLD
VADFATTIWQQMEELGMAPTTTPTQTAMPAFASAF
QRRAGGVLVASHLQSFLEVSYRVLRHLAQP(**SEQ ID
NO:11**)

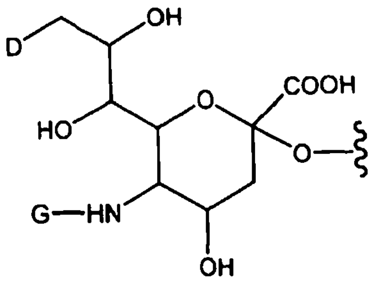
[0100] The present invention is in no way limited to the sequence set forth herein.

[0101] In an exemplary embodiment, the G-CSF peptides of the invention include at least one O-linked glycosylation site, which is glycosylated with a glycosyl residue that includes a PEG moiety. The PEG is covalently attached to the G-CSF peptide via an intact glycosyl linking group. The glycosyl linking group is covalently attached to either an amino acid residue or a glycosyl residue of the G-CSF peptide. Alternatively, the glycosyl linking group is attached to one or more glycosyl units of a glycopeptide. The invention also provides conjugates in which the glycosyl linking group is attached to both an amino acid residue and a glycosyl residue.

[0102] The PEG moiety is attached to an intact glycosyl linker directly, or via a non-glycosyl linker, e.g., substituted or unsubstituted alkyl, substituted or unsubstituted heteroalkyl.

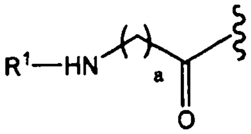
[0103] In a preferred embodiment, the G-CSF peptide comprises a moiety having the formula of Formula I.

Formula I



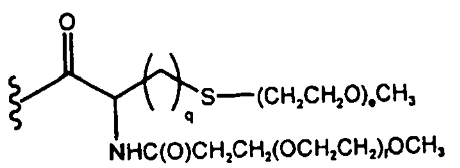
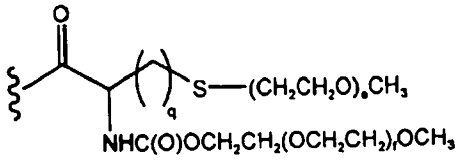
in which D is a member selected from -OH and R^1-L-HN-; G is a member selected from R^1-L- and -C(O)(C_1-C_6)alkyl; R^1 is a moiety comprising a member selected a moiety comprising a straight-chain or branched poly(ethylene glycol) residue; and L is a linker which is a member selected from a bond, substituted or unsubstituted alkyl and substituted or unsubstituted heteroalkyl,such that when D is OH, G is R^1-L-, and when G is –C(O)(C_1-C_6)alkyl, D is R^1-L-NH-. In the modified sialic acid structures set forth herein, COOH also represents COO^- and/or a salt thereof.

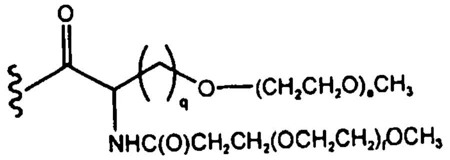
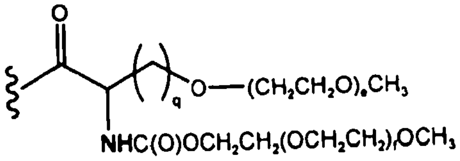
[0104] In one embodiment, a R^1-L has the formula:



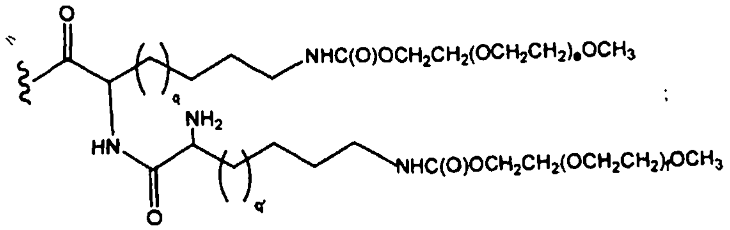
wherein a is an integer from 0 to 20.

[0105] In an exemplary embodiment, R^1 has a structure that is a member selected from:

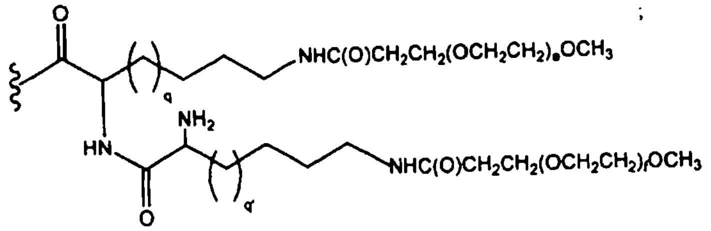

;



; and


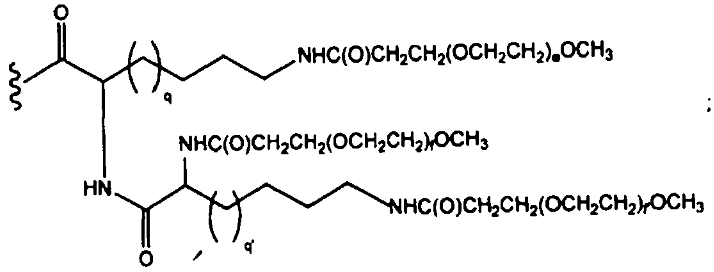
wherein e and f are integers independently selected from 1 to 2500; and q is an integer from 1 to 20. In other embodiments R^1 has a structure that is a member selected from:



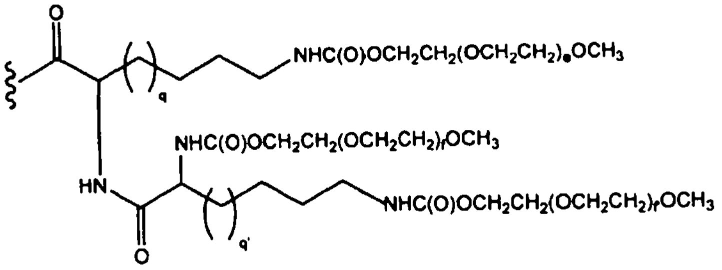
;



;

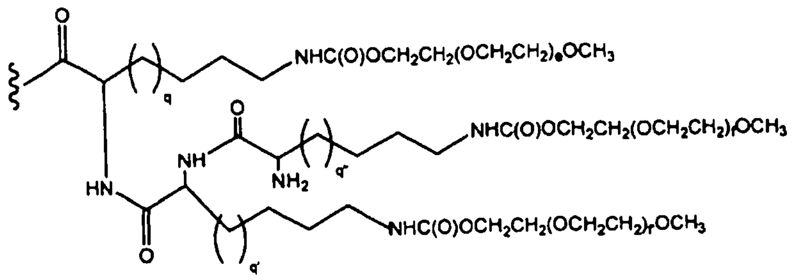


; and

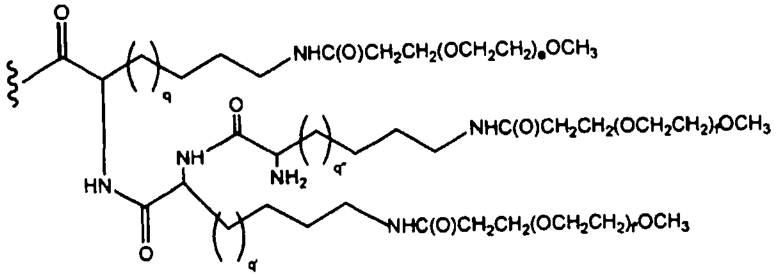


wherein e, f and f' are integers independently selected from 1 to 2500; and q and q' are integers independently selected from 1 to 20.

[0106] In still another embodiment, the invention provides a Facto IX peptide conjugate wherein R^1 has a structure that is a member selected from:



; and



wherein e, f and f' are integers independently selected from 1 to 2500; and q, q' and q''are integers independently selected from 1 to 20.

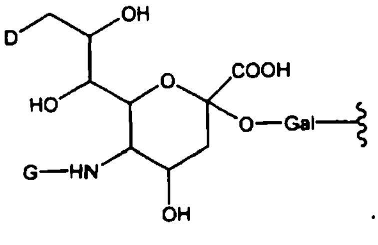
[0107] In other embodiments, R^1 has a structure that is a member selected from:

$\wr\!-C(O)CH_2CH_2(OCH_2CH_2)_eOCH_3$; and

$\wr\!-C(O)OCH_2CH_2(OCH_2CH_2)_fOCH_3$

wherein e and f are integers independently selected from 1 to 2500.

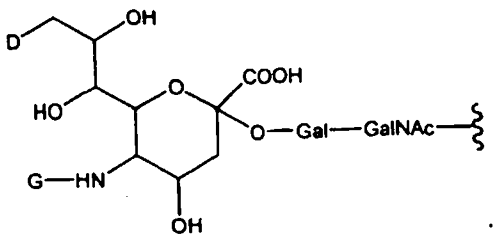
[0108] In another exemplary embodiment, the invention provides a peptide comprising a moiety having the formula:



.

The Gal can be attached to an amino acid or to a glycosyl residue that is directly or indirectly (e.g., through a glycosyl residue) attached to an amino acid.

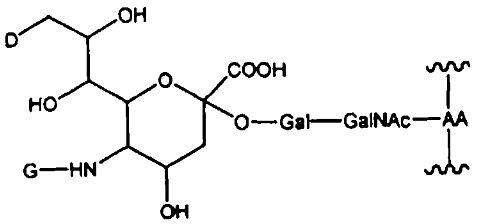
[0109] In other embodiments, the moiety has the formula:



.

The GalNAc can be attached to an amino acid or to a glycosyl residue that is directly or indirectly (e.g., through a glycosyl residue) attached to an amino acid.

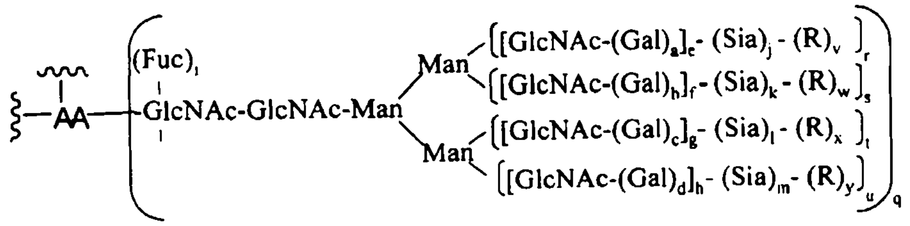
[0110] In a still further exemplary embodiment the peptide comprises a moiety according to the formula



wherein AA is an amino acid residue of said peptide and, in each of the above structures, D and G are as described herein.

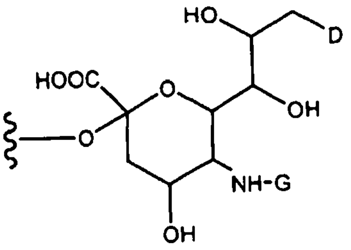
[0111] An exemplary amino acid residue of the G-CSF peptide at which one or more of the above species can be conjugated include serine and threonine, e.g., threonine 133 of SEQ. ID. NO.: 1.

[0112] In another exemplary embodiment, the invention provides a G-CSF conjugate that includes a glycosyl residue having the formula:



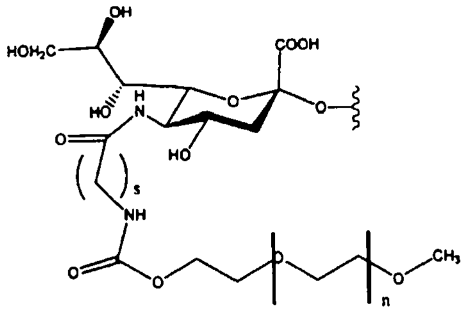
wherein a, b, c, d, i, r, s, t, and u are integers independently selected from 0 and 1. The index q is 1. The indices e, f, g, and h are independently selected from the integers from 0 to 6. The indices j, k, l, and m are independently selected from the integers from 0 and 100. The indices v, w, x, and y are independently selected from 0 and 1, and at least one of v, w, x and y is 1. The symbol AA represents an amino acid residue of the G-CSF peptide.

[0113] The symbol Sia-(R) represents a group that has the formula:



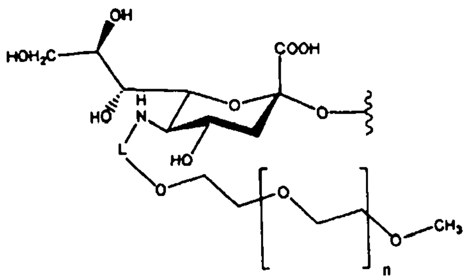
wherein D is selected from -OH and R^1-L-HN-. The symbol G is represents R^1-L- or -C(O)(C_1-C_6)alkyl. R^1 represents a moiety that includes a straight-chain or branched poly(ethylene glycol) residue. L is a linker which is a member selected from a bond, substituted or unsubstituted alkyl and substituted or unsubstituted heteroalkyl. In general, when D is OH, G is R^1-L-, and when G is –C(O)(C_1-C_6)alkyl, D is R^1-L-NH-.

[0114] In another exemplary embodiment, the PEG-modified sialic acid moiety in the conjugate of the invention has the formula:



in which the index "s" represents an integer from 0 to 20, and n is an integer from 1 to 2500. In a preferred embodiment, s is equal to 1; and the m-PEG moiety has a molecular weight of about 20 kD.

[0115] In a still further exemplary embodiment, the PEG-modified sialic acid in has the formula:



in which L is a substituted or unsubstituted alkyl or substituted or unsubstituted heteroalkyl linker moiety joining the sialic acid moiety and the PEG moiety.

[0116] In a preferred embodiment, at least two, more preferably three, more preferably four of the above-named asparagine residues is functionalized with the N-linked glycan chain shown above.

[0117] The conjugates of the invention include intact glycosyl linking groups that are mono- or multi-valent (*e.g.*, antennary structures). Thus, conjugates of the invention include both species in which a selected moiety is attached to a peptide via a monovalent glycosyl linking group and a multivalent linking group. Also included within the invention are conjugates in which more than one selected moiety is attached to a peptide via a multivalent linking group.

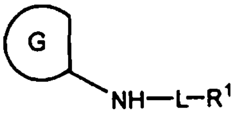
Modified Sugars

[0118] The present invention provides modified sugars, modified sugar nucleotides and conjugates of the modified sugars. In modified sugar compounds of the invention, the sugar moiety is preferably a saccharide, a deoxy-saccharide, an amino-saccharide, or an N-acyl saccharide. The term "saccharide" and its equivalents, "saccharyl," "sugar," and "glycosyl" refer to monomers, dimers, oligomers and polymers. The sugar moiety is also functionalized with a modifying group. The modifying group is conjugated to the sugar moiety, typically, through conjugation with an amine, sulfhydryl or hydroxyl, *e.g.*, primary hydroxyl, moiety on the sugar. In an exemplary embodiment, the modifying group is attached through an amine moiety on the sugar, *e.g.*, through an amide, a urethane or a urea that is formed through the reaction of the amine with a reactive derivative of the modifying group.

[0119] Any sugar can be utilized as the sugar core of the conjugates of the invention. Exemplary sugar cores that are useful in forming the compositions of the invention include, but are not limited to, glucose, galactose, mannose, fucose, and sialic acid. Other useful sugars include amino sugars such as glucosamine, galactosamine, mannosamine, the 5-amine analogue of sialic acid and the like. The sugar core can be a structure found in nature or it can be modified to provide a site for conjugating the modifying group. For example, in one embodiment, the invention provides a peptide conjugate comprising a sialic acid derivative in which the 9-hydroxy moiety is replaced with an amine. The amine is readily derivatized with an activated analogue of a selected modifying group.

[0120] In the discussion that follows the invention is illustrated by reference to the use of selected derivatives of sialic acid. Those of skill in the art will recognize that the focus of the discussion is for clarity of illustration and that the structures and compositions set forth are generally applicable across the genus of saccharide groups, modified saccharide groups, activated modified saccharide groups and conjugates of modified saccharide groups.

[0121] In an exemplary embodiment, the invention provides a peptide conjugate comprising a modified sugar amine that has the formula:



in which G is a glycosyl moiety, L is a bond or a linker and R^1 is the modifying group. Exemplary bonds are those that are formed between an NH_2 on the glycosyl moiety and a group of complementary reactivity on the modifying group. Thus, exemplary bonds include, but are not limited to NHR^1, OR^1, SR^1 and the like. For example, when R^1 includes a carboxylic acid moiety, this moiety may be activated and coupled with an NH_2 moiety on the glycosyl residue affording a bond having the structure $NHC(O)R^1$. Similarly, the OH and SH groups can be converted to the corresponding ether or thioether derivatives, respectively.

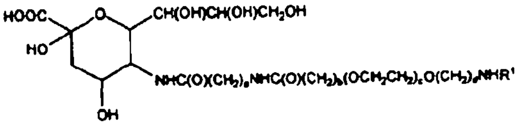
[0122] Exemplary linkers include alkyl and heteroalkyl moieties. The linkers include linking groups, for example acyl-based linking groups, *e.g.*, -C(O)NH-, -OC(O)NH-, and the like. The linking groups are bonds formed between components of the species of the invention, *e.g.*, between the glycosyl moiety and the linker (L), or between the linker and the modifying group (R^1). Other linking groups are ethers, thioethers and amines. For example, in one embodiment, the linker is an amino acid residue, such as a glycine residue. The carboxylic acid moiety of the glycine is converted to the corresponding amide by reaction with an amine on the glycosyl residue, and the amine of the glycine is converted to the corresponding amide or urethane by reaction with an activated carboxylic acid or carbonate of the modifying group.

[0123] Another exemplary linker is a PEG moiety or a PEG moiety that is functionalized with an amino acid residue. The PEG is to the glycosyl group through the amino acid residue at one PEG terminus and bound to R^1 through the other PEG terminus. Alternatively, the amino acid residue is bound to R^1 and the PEG terminus not bound to the amino acid is bound to the glycosyl group.

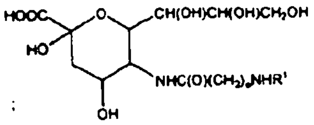
[0124] An exemplary species for NH-L-R^1 has the formula: $-NH\{C(O)(CH_2)_aNH\}_s\{C(O)(CH_2)_b(OCH_2CH_2)_cO(CH_2)_dNH\}_tR^1$, in which the indices s and t are independently 0 or 1. The indices a, b and d are independently integers from 0 to 20, and c is an integer from 1 to 2500. Other similar linkers are based on species in which the –NH moiety is replaced by, for example, -S, -O and –CH_2.

[0123] More particularly, the invention provides a peptide conjugate comprising compounds in which NH-L-R^1 is: $NHC(O)(CH_2)_aNHC(O)(CH_2)_b(OCH_2CH_2)_cO(CH_2)_dNHR^1$, $NHC(O)(CH_2)_b(OCH_2CH_2)_cO(CH_2)_dNHR^1$, $NHC(O)O(CH_2)_b(OCH_2CH_2)_cO(CH_2)_dNHR^1$, $NH(CH_2)_aNHC(O)(CH_2)_b(OCH_2CH_2)_cO(CH_2)_dNHR^1$, $NHC(O)(CH_2)_aNHR^1$, $NH(CH_2)_aNHR^1$, and NHR^1. In these formulae, the indices a, b and d are independently selected from the integers from 0 to 20, preferably from 1 to 5. The index c is an integer from 1 to 2500.

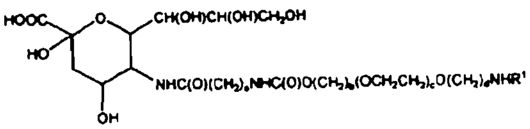
[0125] In an illustrative embodiment, G is sialic acid and selected compounds of the invention have the formulae:

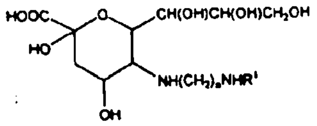


;

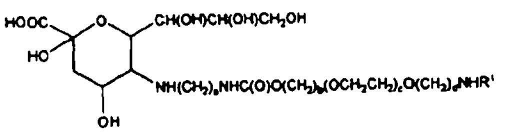


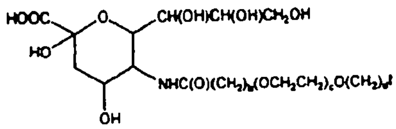
;

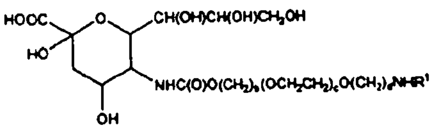




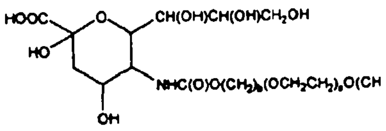
;



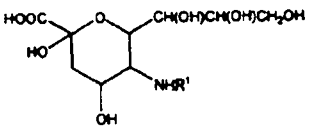




;



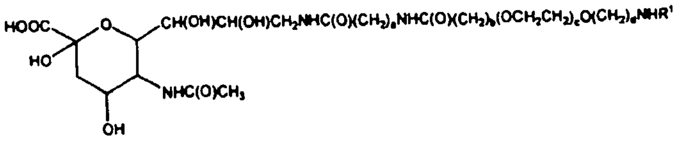
and



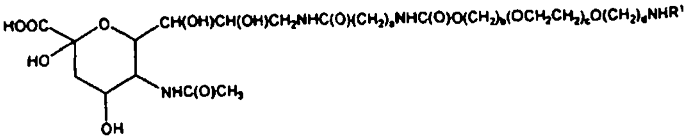
.

[0126] As those of skill in the art will appreciate, the sialic acid moiety in the exemplary compounds above can be replaced with any other amino-saccharide including, but not limited to, glucosamine, galactosamine, mannosamine, their N-acetyl derivatives, and the like.

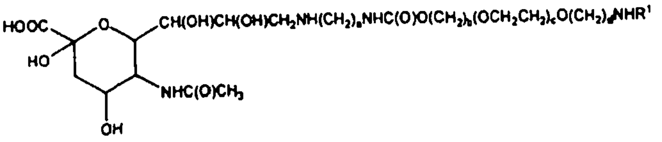
[0127] In another illustrative embodiment, a primary hydroxyl moiety of the sugar is functionalized with the modifying group. For example, the 9-hydroxyl of sialic acid can be converted to the corresponding amine and functionalized to provide a compound according to the invention. Formulae according to this embodiment include:



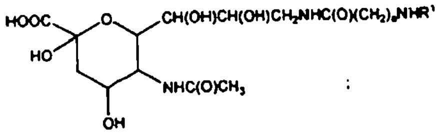
;



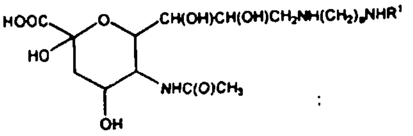
;



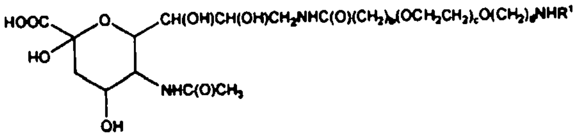
;



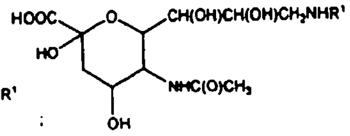
;

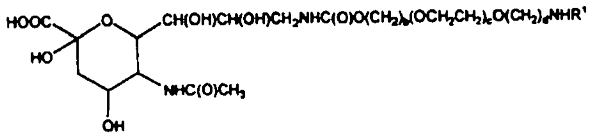


;



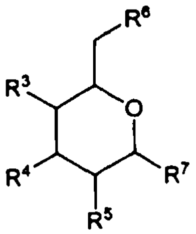
;





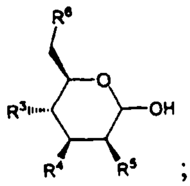
;

[0128] In a further exemplary embodiment, the invention provides a peptide conjugate comprising modified sugars in which the 6-hydroxyl position is converted to the corresponding amine moiety, which bears a linker-modifying group cassette such as those set forth above. Exemplary saccharyl groups that can be used as the core of these modified sugars include Gal, GalNAc, Glc, GlcNAc, Fuc, Xyl, Man, and the like. A representative modified sugar according to this embodiment has the formula:

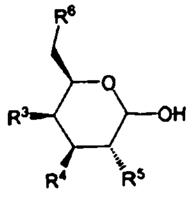


in which R^3-R^5 and R^7 are members independently selected from H, OH, C(O)CH$_3$, NH, and NH C(O)CH$_3$. R^6 is OR^1, NHR^1 or NH-L-R^1, which is as described above.

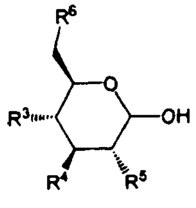
[0129] Selected conjugates of the invention are based on mannose, galactose or glucose, or on species having the stereochemistry of mannose, galactose or glucose. The general formulae of these conjugates are:



;



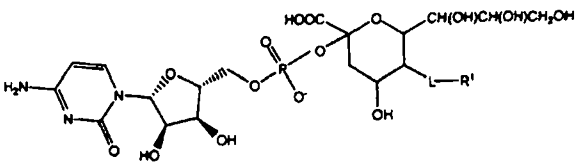
; and



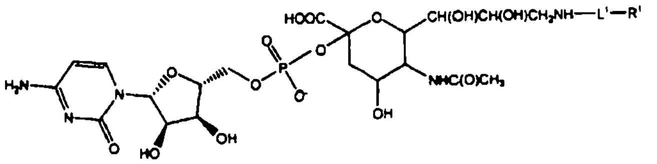
.

[0130] In another exemplary embodiment, the invention provides compounds as set forth above that are activated as the corresponding nucleotide sugars. Exemplary sugar nucleotides that are used in the present invention in their modified form include nucleotide mono-, di- or triphosphates or analogs thereof. In a preferred embodiment, the modified sugar nucleotide is selected from a UDP-glycoside, CMP-glycoside, or a GDP-glycoside. Even more preferably, the sugar nucleotide portion of the modified sugar nucleotide is selected from UDP-galactose, UDP-galactosamine, UDP-glucose, UDP-glucosamine, GDP-mannose, GDP-fucose, CMP-sialic acid, or CMP-NeuAc. In an exemplary embodiment, the nucleotide phosphate is attached to C-1.

[0131] Thus, in an illustrative embodiment in which the glycosyl moiety is sialic acid, the invention provides peptide conjugates that are formed using compounds having the formulae:

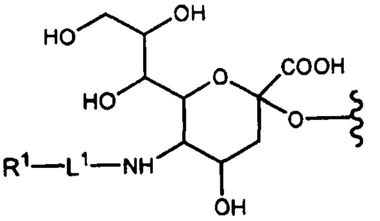


; and

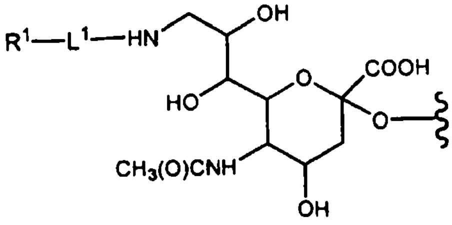


in which L-R^1 is as discussed above, and L^1-R^1 represents a linker bound to the modifying group. As with L, exemplary linker species according to L^1 include a bond, alkyl or heteroalkyl moieties. Exemplary modified sugar nucleotide compounds according to these embodiments are set forth in FIG. 1 and FIG. 2.

[0132] In another exemplary embodiment, the invention provides a conjugate formed between a modified sugar of the invention and a substrate, *e.g.*, a peptide, lipid, aglycone, etc., more particularly between a modified sugar and a glycosyl residue of a glycopeptide or a glycolipid. In this embodiment, the sugar moiety of the modified sugar becomes a glycosyl linking group interposed between the substrate and the modifying group. An exemplary glycosyl linking group is an intact glycosyl linking group, in which the glycosyl moiety or moieties forming the linking group are not degraded by chemical (*e.g.*, sodium metaperiodate) or enzymatic processes (*e.g.*, oxidase). Selected conjugates of the invention include a modifying group that is attached to the amine moiety of an amino-saccharide, *e.g.*, mannosamine, glucosamine, galactosamine, sialic acid etc. Exemplary modifying group-intact glycosyl linking group cassette according to this motif is based on a sialic acid structure, such as that having the formulae:

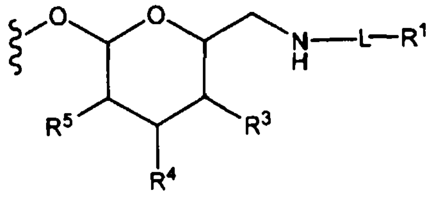


; and

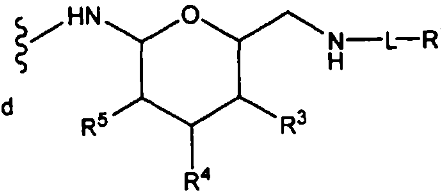


In the formulae above, R^1, L^1 and L^2 are as described above.

[0133] In still a further exemplary embodiment, the conjugate is formed between a substrate and the 1-position of a saccharyl moiety that in which the modifying group is attached through a linker at the 6-carbon position of the saccharyl moiety. Thus, illustrative conjugates according to this embodiment have the formulae:

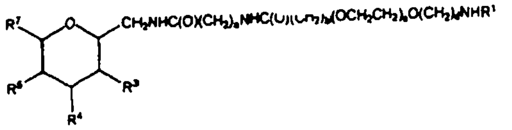


; and

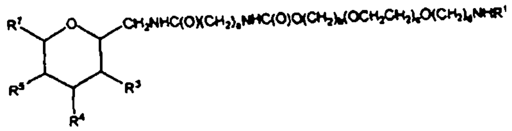


in which the radicals are as discussed above. Those of skill will appreciate that the modified saccharyl moieties set forth above can also be conjugated to a substrate at the 2, 3, 4, or 5 carbon atoms.

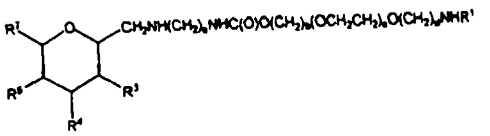
[0134] Illustrative compounds according to this embodiment include compounds having the formulae:



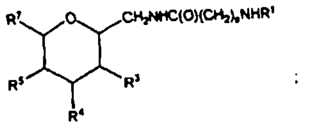
;



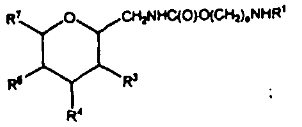
;



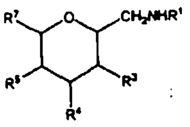
;



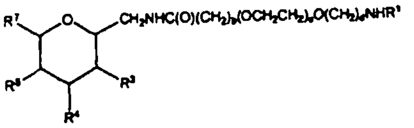
;

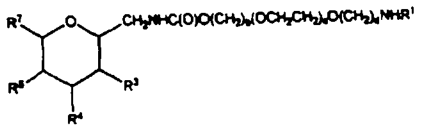


;



and

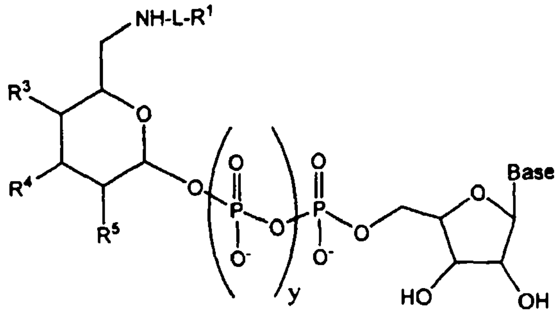




.

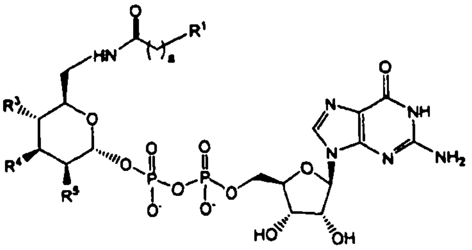
in which the R groups and the indices are as described above.

[0135] The invention also provides sugar nucleotides modified with L-R^1 at the 6-carbon position. Exemplary species according to this embodiment include:

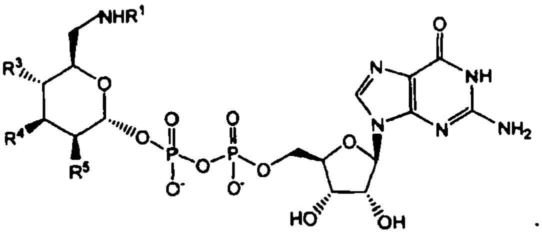


in which the R groups, and L, represent moieties as discussed above. The index "y" is 0, 1 or 2.

[0136] A further exemplary nucleotide sugar of the invention, based on a species having the stereochemistry of GDP mannose. An exemplary species according to this embodiment has the structure:

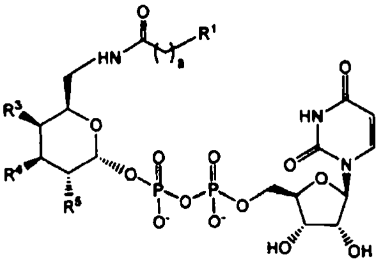


; and

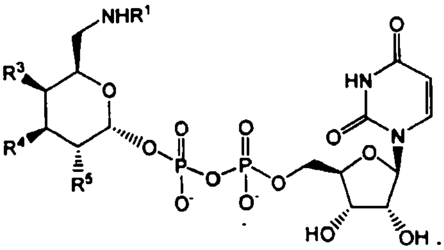


.

[0137] In a still further exemplary embodiment, the invention provides a conjugate, based on the stereochemistry of UDP galactose. An exemplary compound according to this embodiment has the structure:

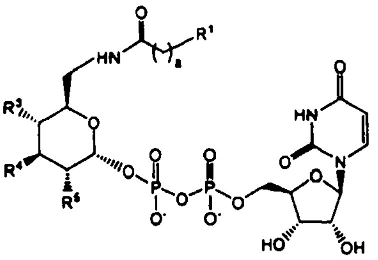


; and

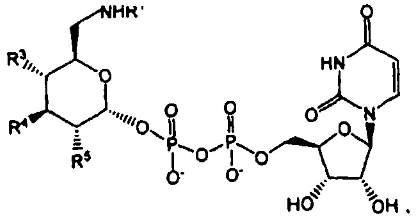


.

[0138] In another exemplary embodiment, the nucleotide sugar is based on the stereochemistry of glucose. Exemplary species according to this embodiment have the formulae:



; and



[0139] The modifying group, R^1, is any of a number of species including, but not limited to, water-soluble polymers, water-insoluble polymers, therapeutic agents, diagnostic agents and the like. The nature of exemplary modifying groups is discussed in greater detail hereinbelow.

Modifying Groups

Water-Soluble Polymers

[0140] Many water-soluble polymers are known to those of skill in the art and are useful in practicing the present invention. The term water-soluble polymer encompasses species such as saccharides (*e.g.*, dextran, amylose, hyalouronic acid, poly(sialic acid), heparans, heparins, etc.); poly (amino acids), *e.g.*, poly(aspartic acid) and poly(glutamic acid); nucleic acids; synthetic polymers (*e.g.*, poly(acrylic acid), poly(ethers), *e.g.*, poly(ethylene glycol); peptides, proteins, and the like. The present invention may be practiced with any water-soluble polymer with the sole limitation that the polymer must include a point at which the remainder of the conjugate can be attached.

[0141] Methods for activation of polymers can also be found in WO 94/17039, U.S. Pat. No. 5,324,844, WO 94/18247, WO 94/04193, U.S. Pat. No. 5,219,564, U.S. Pat. No. 5,122,614, WO 90/13540, U.S. Pat. No. 5,281,698, and more WO 93/15189, and for conjugation between activated polymers and peptides, *e.g.* Coagulation Factor VIII (WO 94/15625), hemoglobin (WO 94/09027), oxygen carrying molecule (U.S. Pat. No. 4,412,989), ribonuclease and superoxide dismutase (Veronese *at al.*, *App. Biochem. Biotech.* **11**: 141-45 (1985)).

[0142] Preferred water-soluble polymers are those in which a substantial proportion of the polymer molecules in a sample of the polymer are of approximately the same molecular weight; such polymers are "homodisperse."

[0143] The present invention is further illustrated by reference to a poly(ethylene glycol) conjugate. Several reviews and monographs on the functionalization and conjugation of PEG are available. *See,* for example, Harris, *Macronol. Chem. Phys.*

C25: 325-373 (1985); Scouten, *Methods in Enzymology* **135**: 30-65 (1987); Wong *et al.*, *Enzyme Microb. Technol.* **14**: 866-874 (1992); Delgado *et al.*, *Critical Reviews in Therapeutic Drug Carrier Systems* **9**: 249-304 (1992); Zalipsky, *Bioconjugate Chem.* **6**: 150-165 (1995); and Bhadra, *et al.*, *Pharmazie*, **57**:5-29 (2002). Routes for preparing reactive PEG molecules and forming conjugates using the reactive molecules are known in the art. For example, U.S. Patent No. 5,672,662 discloses a water soluble and isolatable conjugate of an active ester of a polymer acid selected from linear or branched poly(alkylene oxides), poly(oxyethylated polyols), poly(olefinic alcohols), and poly(acrylomorpholine).

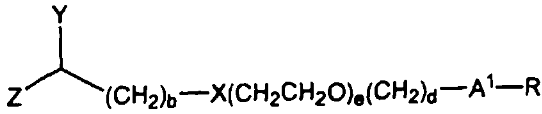
[0144] U.S. Patent No. 6,376,604 sets forth a method for preparing a water-soluble 1-benzotriazolylcarbonate ester of a water-soluble and non-peptidic polymer by reacting a terminal hydroxyl of the polymer with di(1-benzotriazoyl)carbonate in an organic solvent. The active ester is used to form conjugates with a biologically active agent such as a protein or peptide.

[0145] WO 99/45964 describes a conjugate comprising a biologically active agent and an activated water soluble polymer comprising a polymer backbone having at least one terminus linked to the polymer backbone through a stable linkage, wherein at least one terminus comprises a branching moiety having proximal reactive groups linked to the branching moiety, in which the biologically active agent is linked to at least one of the proximal reactive groups. Other branched poly(ethylene glycols) are described in WO 96/21469, U.S. Patent No. 5,932,462 describes a conjugate formed with a branched PEG molecule that includes a branched terminus that includes reactive functional groups. The free reactive groups are available to react with a biologically active species, such as a protein or peptide, forming conjugates between the poly(ethylene glycol) and the biologically active species. U.S. Patent No. 5,446,090 describes a bifunctional PEG linker and its use in forming conjugates having a peptide at each of the PEG linker termini.

[0146] Conjugates that include degradable PEG linkages are described in WO 99/34833; and WO 99/14259, as well as in U.S. Patent No. 6,348,558. Such degradable linkages are applicable in the present invention.

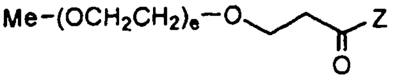
[0147] The art-recognized methods of polymer activation set forth above are of use in the context of the present invention in the formation of the branched polymers set forth herein and also for the conjugation of these branched polymers to other species, *e.g.*, sugars, sugar nucleotides and the like.

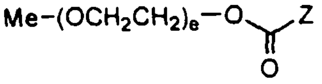
[0148] Exemplary poly(ethylene glycol) molecules of use in the invention include, but are not limited to, those having the formula:

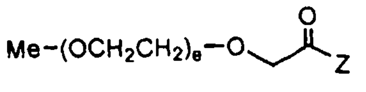


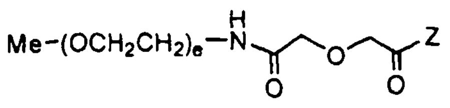
in which R^8 is H, OH, NH_2, substituted or unsubstituted alkyl, substituted or unsubstituted aryl, substituted or unsubstituted heteroaryl, substituted or unsubstituted heterocycloalkyl, substituted or unsubstituted heteroalkyl, *e.g.*, acetal, OHC-, $H_2N\text{-}(CH_2)_q\text{-}$, $HS\text{-}(CH_2)_q$, or $\text{-}(CH_2)_qC(Y)Z^1$. The index "e" represents an integer from 1 to 2500. The indices b, d, and q independently represent integers from 0 to 20. The symbols Z and Z^1 independently represent OH, NH_2, leaving groups, *e.g.*, imidazole, p-nitrophenyl, HOBT, tetrazole, halide, $S\text{-}R^9$, the alcohol portion of activated esters; $\text{-}(CH_2)_pC(Y^1)V$, or $\text{-}(CH_2)_pU(CH_2)_sC(Y^1)_v$. The symbol Y represents H(2), =O, =S, $=N\text{-}R^{10}$. The symbols X, Y, Y^1, A^1, and U independently represent the moieties O, S, $N\text{-}R^{11}$. The symbol V represents OH, NH_2, halogen, $S\text{-}R^{12}$, the alcohol component of activated esters, the amine component of activated amides, sugar-nucleotides, and proteins. The indices p, q, s and v are members independently selected from the integers from 0 to 20. The symbols R^9, R^{10}, R^{11} and R^{12} independently represent H, substituted or unsubstituted alkyl, substituted or unsubstituted heteroalkyl, substituted or unsubstituted aryl, substituted or unsubstituted heterocycloalkyl and substituted or unsubstituted heteroaryl.

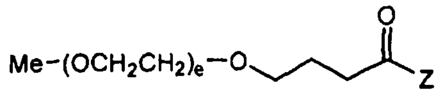
[0149] In other exemplary embodiments, the poly(ethylene glycol) molecule is selected from the following:

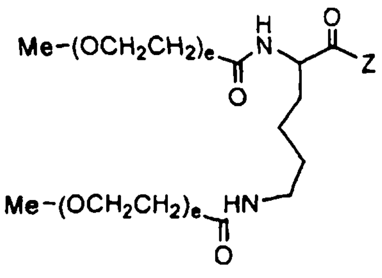




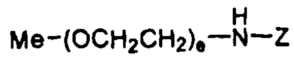






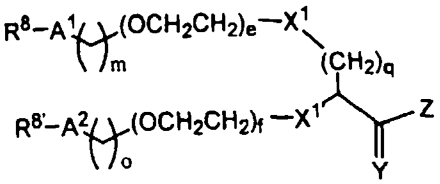


$Me\text{-}(OCH_2CH_2)_e\text{—}S\text{—}Z$



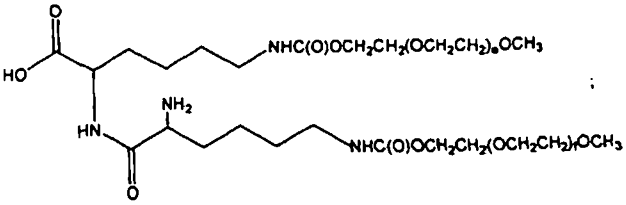
$Me\text{-}(OCH_2CH_2)_e\text{—}NH\text{—}Z$

[0150] The poly(ethylene glycol) useful in forming the conjugate of the invention is either linear or branched. Branched poly(ethylene glycol) molecules suitable for use in the invention include, but are not limited to, those described by the following formula:

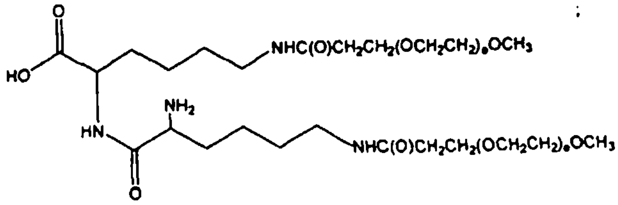


in which R^8 and $R^{8'}$ are members independently selected from the groups defined for R^8, above. A^1 and A^2 are members independently selected from the groups defined for A^1, above. The indices e, f, o, and q are as described above. Z and Y are as described above. X^1 and $X^{1'}$ are members independently selected from S, SC(O)NH, HNC(O)S, SC(O)O, O, NH, NHC(O), (O)CNH and NHC(O)O, OC(O)NH.

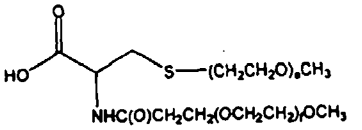
[0151] In other exemplary embodiments, the branched PEG is based upon a cysteine, serine or di-lysine core. Thus, further exemplary branched PEGs include:



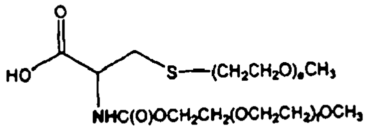
;

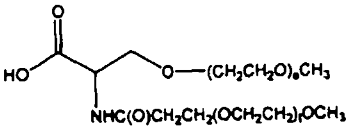


;

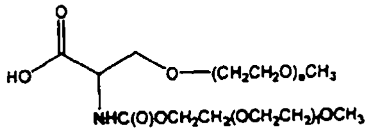


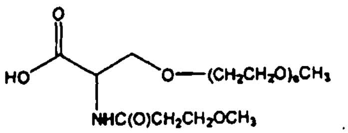
;



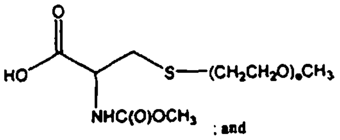


;

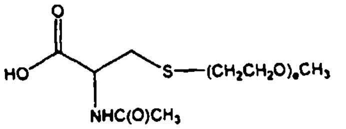




;

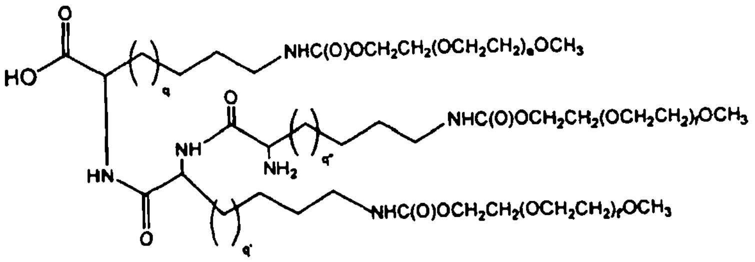


; and

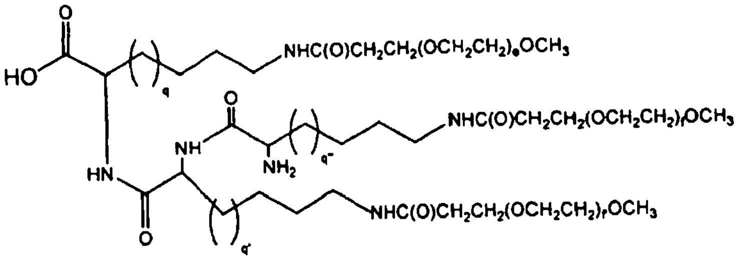


.

[0152] In yet another embodiment, the branched PEG moiety is based upon a tri-lysine peptide. The tri-lysine can be mono-, di-, tri-, or tetra-PEG-ylated. Exemplary species according to this embodiment have the formulae:



; and

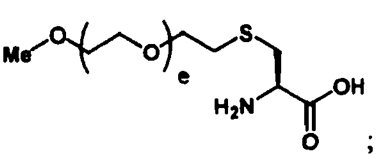


in which e, f and f' are independently selected integers from 1 to 2500; and q, q' and q" are independently selected integers from 1 to 20.

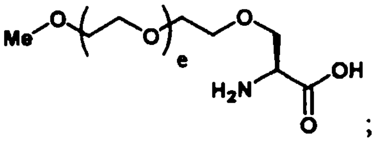
[0153] In exemplary embodiments of the invention, the PEG is m-PEG (5 kD, 10 kD, or 20kD). An exemplary branched PEG species is a serine- or cysteine-(m-PEG)$_2$ in which the m-PEG is a 20 kD m-PEG.

[0154] As will be apparent to those of skill, the branched polymers of use in the invention include variations on the themes set forth above. For example the di-lysine-PEG conjugate shown above can include three polymeric subunits, the third bonded to the α-amine shown as unmodified in the structure above. Similarly, the use of a tri-lysine functionalized with three or four polymeric subunits is within the scope of the invention.

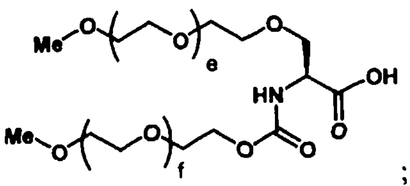
[0155] Specific embodiments according to the invention include:



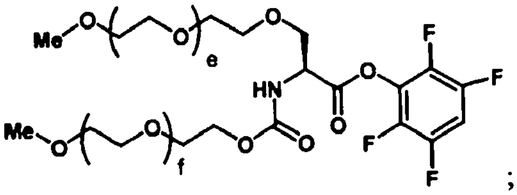
;

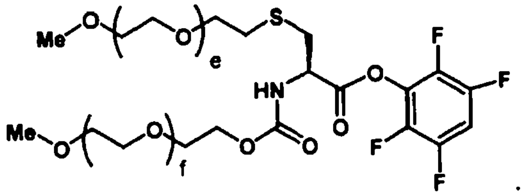


; and

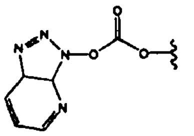

;

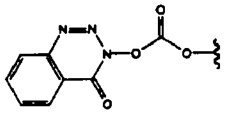
and carbonates and active esters of these species, such as:

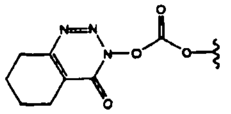

; and

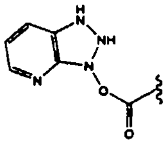

.

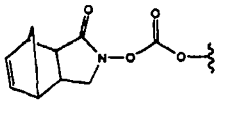
[0156] Other activating, or leaving groups, appropriate for activating linear PEGs of use in preparing the compounds set forth herein include, but are not limited to the species:

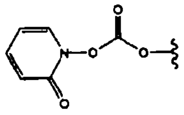

;

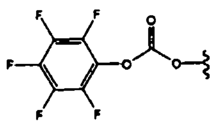

;

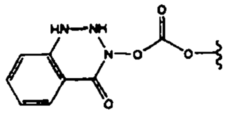

;


;


;


;

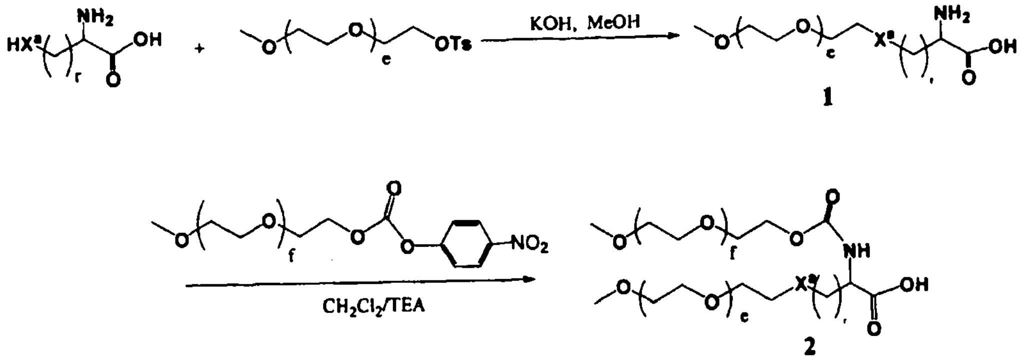

; and


.

[0157] PEG molecules that are activated with these and other species and methods of making the activated PEGs are set forth in WO 04/083259.

[0158] Those of skill in the art will appreciate that one or more of the m-PEG arms of the branched polymer can be replaced by a PEG moiety with a different terminus, *e.g.*, OH, COOH, NH_2, C_2-C_{10}-alkyl, etc. Moreover, the structures above are readily modified by inserting alkyl linkers (or removing carbon atoms) between the α-carbon atom and the functional group of the side chain. Thus, "homo" derivatives and higher homologues, as well as lower homologues are within the scope of cores for branched PEGs of use in the present invention.

[0159] The branched PEG species set forth herein are readily prepared by methods such as that set forth in the scheme below:



in which X^a is O or S and r is an integer from 1 to 5. The indices e and f are independently selected integers from 1 to 2500.

[0160] Thus, according to this scheme, a natural or unnatural amino acid is contacted with an activated m-PEG derivative, in this case the tosylate, forming **1** by alkylating the side-chain heteroatom X^a. The mono-functionalized m-PEG amino acid is submitted to N-acylation conditions with a reactive m-PEG derivative, thereby assembling branched m-PEG **2**. As one of skill will appreciate, the tosylate leaving group can be replaced with any suitable leaving group, *e.g.*, halogen, mesylate, triflate, etc. Similarly, the reactive carbonate utilized to acylate the amine can be replaced with an active ester, *e.g.*, N-hydroxysuccinimide, etc., or the acid can be activated *in situ* using a dehydrating agent such as dicyclohexylcarbodiimide, carbonyldiimidazole, etc.

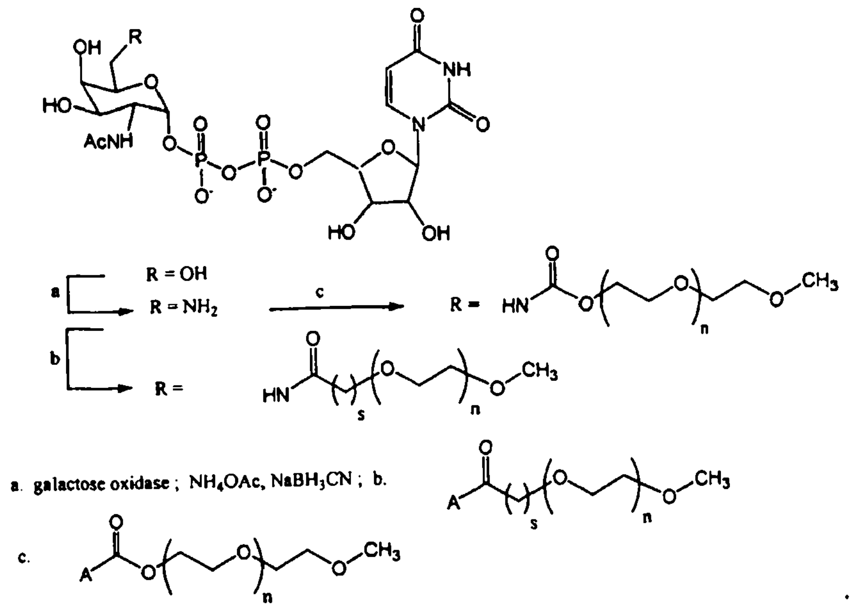
[0161] In an exemplary embodiment, the modifying group is a PEG moiety, however, any modifying group, *e.g.*, water-soluble polymer, water-insoluble polymer, therapeutic moiety, etc., can be incorporated in a glycosyl moiety through an appropriate linkage. The modified sugar is formed by enzymatic means, chemical means or a combination thereof, thereby producing a modified sugar. In an exemplary embodiment, the sugars are substituted with an active amine at any position that allows for the attachment of the modifying moiety, yet still allows the sugar to function as a substrate for an enzyme capable of coupling the modified sugar to the G-CSF peptide. In an exemplary embodiment, when galactosamine is the modified sugar, the amine moiety is attached to the carbon atom at the 6-position.

Water-soluble Polymer Modified Species

[0162] Water-soluble polymer modified nucleotide sugar species in which the sugar moiety is modified with a water-soluble polymer are of use in the present invention. An exemplary modified sugar nucleotide bears a sugar group that is modified through an amine moiety on the sugar. Modified sugar nucleotides, e.g., saccharyl-amine derivatives of a sugar nucleotide, are also of use in the methods of the invention. For example, a saccharyl amine (without the modifying group) can be enzymatically conjugated to a peptide (or other species) and the free saccharyl amine moiety subsequently conjugated to a desired modifying group. Alternatively, the modified sugar nucleotide can function as a substrate for an enzyme that transfers the modified sugar to a saccharyl acceptor on a substrate, e.g., a peptide, glycopeptide, lipid, aglycone, glycolipid, etc.

[0163] In one embodiment in which the saccharide core is galactose or glucose, R^5 is NHC(O)Y.

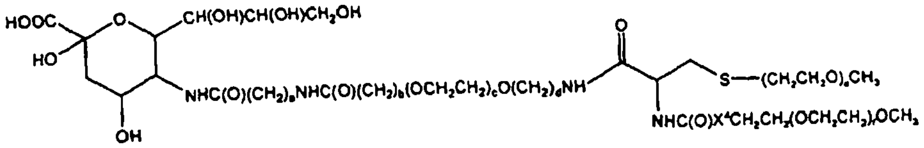
[0164] In an exemplary embodiment, the modified sugar is based upon a 6-amino-N-acetyl-glycosyl moiety. As shown below for N-acetylgalactosamine, the 6-amino-sugar moiety is readily prepared by standard methods.

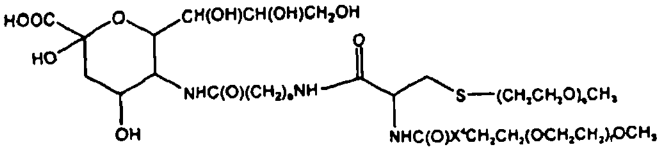


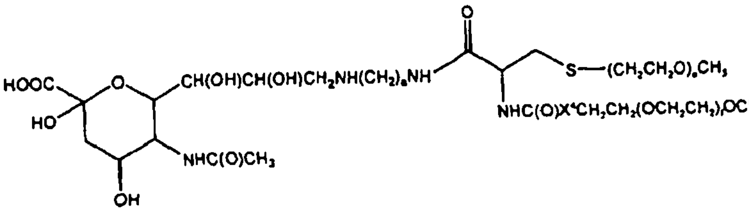
[0165] In the scheme above, the index n represents an integer from 1 to 2500, preferably from 10 to 1500, and more preferably from 10 to 1200. The symbol "A" represents an activating group, *e.g.*, a halo, a component of an activated ester (*e.g.*, a N-hydroxysuccinimide ester), a component of a carbonate (*e.g.*, p-nitrophenyl carbonate) and the like. Those of skill in the art will appreciate that other PEG-amide nucleotide sugars are readily prepared by this and analogous methods.

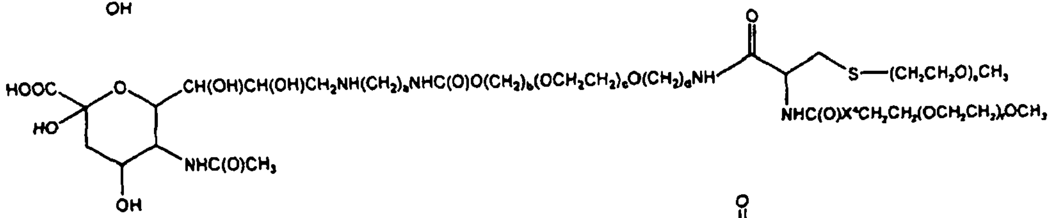
[0166] In other exemplary embodiments, the amide moiety is replaced by a group such as a urethane or a urea.

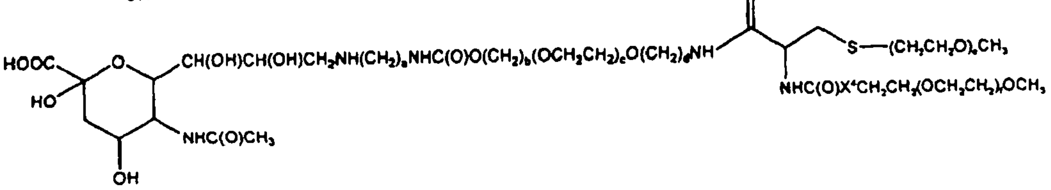
[0167] In still further embodiments, R^1 is a branched PEG, for example, one of those species set forth above. Illustrative compounds according to this embodiment include:

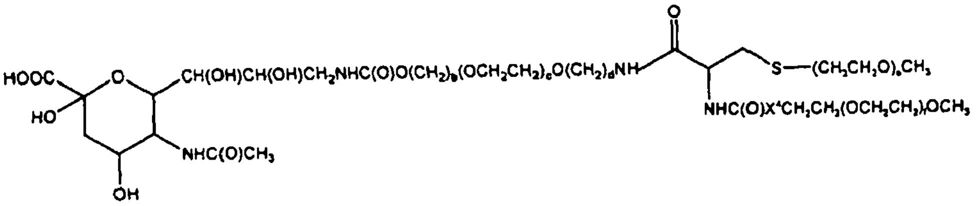

;


;


;

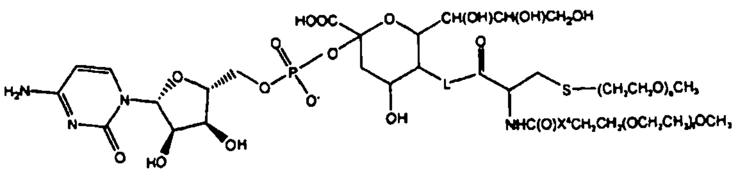

;



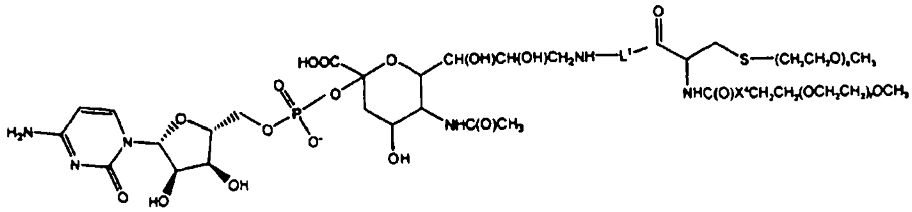


in which X^4 is a bond or O.

[0168] Moreover, as discussed above, the present invention provides peptide conjugates that are formed using nucleotide sugars that are modified with a water-soluble polymer, which is either straight-chain or branched. For example, compounds having the formula shown below are within the scope of the present invention:

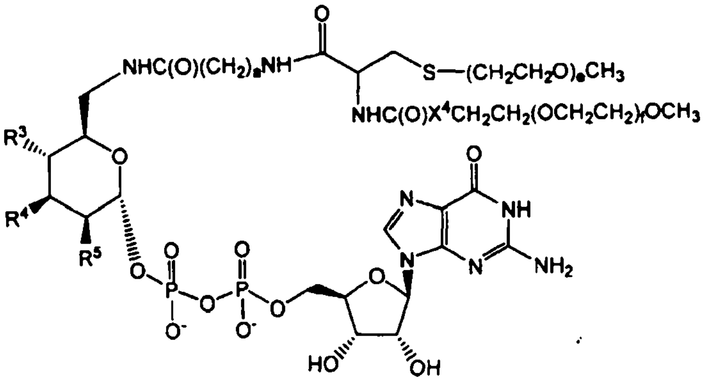


; and



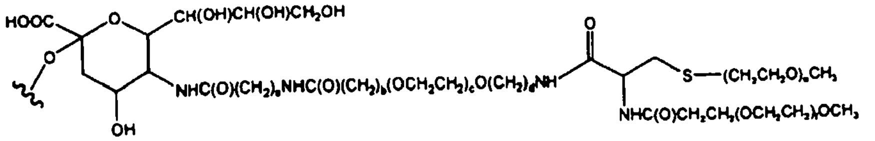
in which X^4 is O or a bond.

[0169] Similarly, the invention provides peptide conjugates that are formed using nucleotide sugars of those modified sugar species in which the carbon at the 6-position is modified:

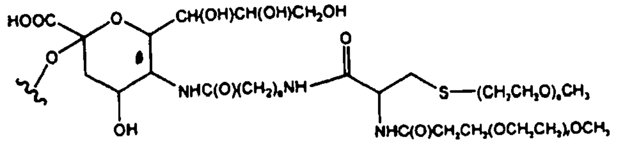


in which X^4 is a bond or O.

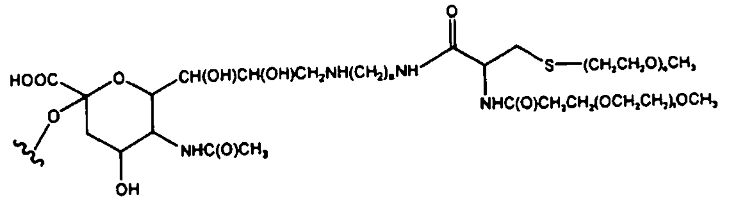
[0170] Also provided are conjugates of peptides and glycopeptides, lipids and glycolipids that include the compositions of the invention. For example, the invention provides conjugates having the following formulae:



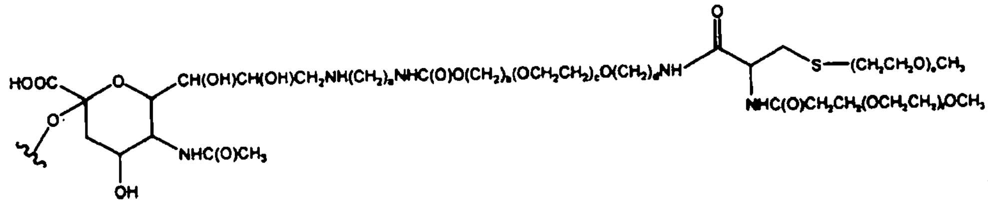
;



;



; and



.

Water-insoluble polymers

[0171] In another embodiment, analogous to those discussed above, the modified sugars include a water-insoluble polymer, rather than a water-soluble polymer. The conjugates of the invention may also include one or more water-insoluble polymers. This embodiment of the invention is illustrated by the use of the conjugate as a vehicle with which to deliver a therapeutic peptide in a controlled manner. Polymeric drug delivery systems are known in the art. *See*, for example, Dunn *et al.*, Eds. POLYMERIC DRUGS AND DRUG DELIVERY SYSTEMS, ACS Symposium Series Vol. 469, American Chemical Society, Washington, D.C. 1991. Those of skill in the art will appreciate that substantially any known drug delivery system is applicable to the conjugates of the present invention.

[0172] Representative water-insoluble polymers include, but are not limited to, polyphosphazines, poly(vinyl alcohols), polyamides, polycarbonates, polyalkylenes, polyacrylamides, polyalkylene glycols, polyalkylene oxides, polyalkylene terephthalates, polyvinyl ethers, polyvinyl esters, polyvinyl halides, polyvinylpyrrolidone, polyglycolides, polysiloxanes, polyurethanes, poly(methyl methacrylate), poly(ethyl methacrylate), poly(butyl methacrylate), poly(isobutyl methacrylate), poly(hexyl methacrylate), poly(isodecyl methacrylate), poly(lauryl methacrylate), poly(phenyl methacrylate), poly(methyl acrylate), poly(isopropyl acrylate), poly(isobutyl acrylate), poly(octadecyl acrylate) polyethylene, polypropylene, poly(ethylene glycol), poly(ethylene oxide), poly (ethylene terephthalate), poly(vinyl acetate), polyvinyl chloride, polystyrene, polyvinyl pyrrolidone, pluronics and polyvinylphenol and copolymers thereof.

[0173] Synthetically modified natural polymers of use in conjugates of the invention include, but are not limited to, alkyl celluloses, hydroxyalkyl celluloses, cellulose ethers, cellulose esters, and nitrocelluloses. Particularly preferred members of the broad classes of synthetically modified natural polymers include, but are not limited to, methyl cellulose, ethyl cellulose, hydroxypropyl cellulose, hydroxypropyl methyl cellulose, hydroxybutyl methyl cellulose, cellulose acetate, cellulose propionate, cellulose acetate butyrate, cellulose acetate phthalate, carboxymethyl cellulose, cellulose triacetate, cellulose sulfate sodium salt, and polymers of acrylic and methacrylic esters and alginic acid.

[0174] These and the other polymers discussed herein can be readily obtained from commercial sources such as Sigma Chemical Co. (St. Louis, MO.), Polysciences (Warrenton, PA.), Aldrich (Milwaukee, WI.), Fluka (Ronkonkoma, NY), and BioRad (Richmond, CA), or else synthesized from monomers obtained from these suppliers using standard techniques.

[0175] Representative biodegradable polymers of use in the conjugates of the invention include, but are not limited to, polylactides, polyglycolides and copolymers thereof, poly(ethylene terephthalate), poly(butyric acid), poly(valeric acid), poly(lactide-co-caprolactone), poly(lactide-co-glycolide), polyanhydrides, polyorthoesters, blends and copolymers thereof. Of particular use are compositions that form gels, such as those including collagen, pluronics and the like.

[0176] The polymers of use in the invention include "hybrid' polymers that include water-insoluble materials having within at least a portion of their structure, a bioresorbable molecule. An example of such a polymer is one that includes a water-insoluble copolymer, which has a bioresorbable region, a hydrophilic region and a plurality of crosslinkable functional groups per polymer chain.

[0177] For purposes of the present invention, "water-insoluble materials" includes materials that are substantially insoluble in water or water-containing environments. Thus, although certain regions or segments of the copolymer may be hydrophilic or even water-soluble, the polymer molecule, as a whole, does not to any substantial measure dissolve in water.

[0178] For purposes of the present invention, the term "bioresorbable molecule" includes a region that is capable of being metabolized or broken down and resorbed and/or eliminated through normal excretory routes by the body. Such metabolites or break down products are preferably substantially non-toxic to the body.

[0179] The bioresorbable region may be either hydrophobic or hydrophilic, so long as the copolymer composition as a whole is not rendered water-soluble. Thus, the bioresorbable region is selected based on the preference that the polymer, as a whole, remains water-insoluble. Accordingly, the relative properties, *i.e.*, the kinds of functional groups contained by, and the relative proportions of the bioresorbable region, and the hydrophilic region are selected to ensure that useful bioresorbable compositions remain water-insoluble.

[0180] Exemplary resorbable polymers include, for example, synthetically produced resorbable block copolymers of poly(α-hydroxy-carboxylic acid)/poly(oxyalkylene, (*see*, Cohn *et al.*, U.S. Patent No. 4,826,945). These copolymers are not crosslinked and are water-soluble so that the body can excrete the degraded block copolymer compositions. *See*, Younes *et al.*, *J Biomed. Mater. Res.* **21**: 1301-1316 (1987); and Cohn *et al.*, *J Biomed. Mater. Res.* **22**: 993-1009 (1988).

[0181] Presently preferred bioresorbable polymers include one or more components selected from poly(esters), poly(hydroxy acids), poly(lactones), poly(amides), poly(ester-amides), poly (amino acids), poly(anhydrides), poly(orthoesters), poly(carbonates), poly(phosphazines), poly(phosphoesters), poly(thioesters), polysaccharides and mixtures thereof. More preferably still, the bioresorbable polymer includes a poly(hydroxy) acid component. Of the poly(hydroxy) acids, polylactic acid, polyglycolic acid, polycaproic acid, polybutyric acid, polyvaleric acid and copolymers and mixtures thereof are preferred.

[0182] In addition to forming fragments that are absorbed *in vivo* ("bioresorbed"), preferred polymeric coatings for use in the methods of the invention can also form an excretable and/or metabolizable fragment.

[0183] Higher order copolymers can also be used in the present invention. For example, Casey *et al.*, U.S. Patent No. 4,438,253, which issued on March 20, 1984, discloses tri-block copolymers produced from the transesterification of poly(glycolic acid) and an hydroxyl-ended poly(alkylene glycol). Such compositions are disclosed for use as resorbable monofilament sutures. The flexibility of such compositions is controlled by the incorporation of an aromatic orthocarbonate, such as tetra-p-tolyl orthocarbonate into the copolymer structure.

[0184] Other polymers based on lactic and/or glycolic acids can also be utilized. For example, Spinu, U.S. Patent No. 5,202,413, which issued on April 13, 1993, discloses biodegradable multi-block copolymers having sequentially ordered blocks of polylactide and/or polyglycolide produced by ring-opening polymerization of lactide and/or glycolide onto either an oligomeric diol or a diamine residue followed by chain extension with a di-functional compound, such as, a diisocyanate, diacylchloride or dichlorosilane.

[0185] Bioresorbable regions of coatings useful in the present invention can be designed to be hydrolytically and/or enzymatically cleavable. For purposes of the present invention, "hydrolytically cleavable" refers to the susceptibility of the copolymer, especially the bioresorbable region, to hydrolysis in water or a water-containing environment. Similarly, "enzymatically cleavable" as used herein refers to the susceptibility of the copolymer, especially the bioresorbable region, to cleavage by endogenous or exogenous enzymes.

[0186] When placed within the body, the hydrophilic region can be processed into excretable and/or metabolizable fragments. Thus, the hydrophilic region can include, for example, polyethers, polyalkylene oxides, polyols, poly(vinyl pyrrolidine), poly(vinyl alcohol), poly(alkyl oxazolines), polysaccharides, carbohydrates, peptides, proteins and copolymers and mixtures thereof. Furthermore, the hydrophilic region can also be, for example, a poly(alkylene) oxide. Such poly(alkylene) oxides can include, for example, poly(ethylene) oxide, poly(propylene) oxide and mixtures and copolymers thereof.

[0187] Polymers that are components of hydrogels are also useful in the present invention. Hydrogels are polymeric materials that are capable of absorbing relatively large quantities of water. Examples of hydrogel forming compounds include, but are not limited to, polyacrylic acids, sodium carboxymethylcellulose, polyvinyl alcohol, polyvinyl pyrrolidine, gelatin, carrageenan and other polysaccharides, hydroxyethylenemethacrylic acid (HEMA), as well as derivatives thereof, and the like. Hydrogels can be produced that are stable, biodegradable and bioresorbable.

Moreover, hydrogel compositions can include subunits that exhibit one or more of these properties.

[0188] Bio-compatible hydrogel compositions whose integrity can be controlled through crosslinking are known and are presently preferred for use in the methods of the invention. For example, Hubbell *et al.*, U.S. Patent Nos. 5,410,016, which issued on April 25, 1995 and 5,529,914, which issued on June 25, 1996, disclose water-soluble systems, which are crosslinked block copolymers having a water-soluble central block segment sandwiched between two hydrolytically labile extensions. Such copolymers are further end-capped with photopolymerizable acrylate functionalities. When crosslinked, these systems become hydrogels. The water soluble central block of such copolymers can include poly(ethylene glycol); whereas, the hydrolytically labile extensions can be a poly(α-hydroxy acid), such as polyglycolic acid or polylactic acid. *See*, Sawhney *et al.*, *Macromolecules* **26**: 581-587 (1993).

[0189] In another preferred embodiment, the gel is a thermoreversible gel. Thermoreversible gels including components, such as pluronics, collagen, gelatin, hyalouronic acid, polysaccharides, polyurethane hydrogel, polyurethane-urea hydrogel and combinations thereof are presently preferred.

[0190] In yet another exemplary embodiment, the conjugate of the invention includes a component of a liposome. Liposomes can be prepared according to methods known to those skilled in the art, for example, as described in Eppstein *et al.*, U.S. Patent No. 4,522,811, which issued on June 11, 1985. For example, liposome formulations may be prepared by dissolving appropriate lipid(s) (such as stearoyl phosphatidyl ethanolamine, stearoyl phosphatidyl choline, arachadoyl phosphatidyl choline, and cholesterol) in an inorganic solvent that is then evaporated, leaving behind a thin film of dried lipid on the surface of the container. An aqueous solution of the active compound or its pharmaceutically acceptable salt is then introduced into the container. The container is then swirled by hand to free lipid material from the sides of the container and to disperse lipid aggregates, thereby forming the liposomal suspension.

[0191] The above-recited microparticles and methods of preparing the microparticles are offered by way of example and they are not intended to define the scope of microparticles of use in the present invention. It will be apparent to those of skill in the art that an array of microparticles, fabricated by different methods, are of use in the present invention.

[0192] The structural formats discussed above in the context of the water-soluble polymers, both straight-chain and branched are generally applicable with respect to the water-insoluble polymers as well. Thus, for example, the cysteine, serine, dilysine, and trilysine branching cores can be functionalized with two water-insoluble polymer moieties. The methods used to produce these species are generally closely analogous to those used to produce the water-soluble polymers.

[0193] The *in vivo* half-life of therapeutic glycopeptides can also be enhanced with PEG moieties such as polyethylene glycol (PEG). For example, chemical modification of proteins with PEG (PEGylation) increases their molecular size and decreases their surface- and functional group-accessibility, each of which are dependent on the size of the PEG attached to the protein. This results in an improvement of plasma half-lives and in proteolytic-stability, and a decrease in immunogenicity and hepatic uptake (Chaffee *et al. J. Clin. Invest.* **89**: 1643-1651 (1992); Pyatak *et al. Res. Commun. Chem. Pathol Pharmacol.* **29**: 113-127 (1980)). PEGylation of interleukin-2 has been reported to increase its antitumor potency *in vivo* (Katre *et al. Proc. Natl. Acad. Sci. USA.* **84**: 1487-1491 (1987)) and PEGylation of a F(ab')2 derived from the monoclonal antibody A7 has improved its tumor localization (Kitamura *et al. Biochem. Biophys. Res. Commun.* **28**: 1387-1394 (1990)). Thus, in another preferred embodiment, the *in vivo* half-life of a peptide derivatized with a PEG moiety by a method of the invention is increased relevant to the *in vivo* half-life of the non-derivatized peptide.

[0194] The increase in peptide *in vivo* half-life is best expressed as a range of percent increase in this quantity. The lower end of the range of percent increase is about 40%, about 60%, about 80%, about 100%, about 150% or about 200%. The upper end of the range is about 60%, about 80%, about 100%, about 150%, or more than about 250%.

[0195] In an exemplary embodiment, the present invention provides a PEGylated FSH (**FIG. 1**, **FIG. 2** and **FIG. 5**).

The Methods

[0196] In addition to the conjugates discussed above, the present invention provides methods for preparing these and other conjugates. Thus, in a further aspect, the invention provides a method of forming a covalent conjugate between a selected moiety and an G-CSF peptide. Additionally, the invention provides methods for targeting conjugates of the invention to a particular tissue or region of the body.

[0197] In exemplary embodiments, the conjugate is formed between a PEG moiety (or an enzymatically transferable glycosyl moiety comprising the PEG moiety), and a glycosylated or non-glycosylated peptide. The PEG is conjugated to the G-CSF peptide via an intact glycosyl linking group, which is interposed between, and covalently linked to both the G-CSF peptide and the PEG moiety, or to a PEG-non-glycosyl linker (*e.g.*, substituted or unsubstituted alkyl, substituted or unsubstituted heteroalkyl) construct. The method includes contacting the G-CSF peptide with a mixture containing a modified sugar and a glycosyltransferase for which the modified sugar is a substrate. The reaction is conducted under conditions sufficient to form a covalent bond between the modified sugar and the G-CSF peptide. The sugar moiety of the modified sugar is preferably selected from nucleotide sugars, activated sugars and sugars, which are neither nucleotides nor activated.

[0198] The acceptor peptide (glycosylated or non-glycosylated) is typically synthesized *de novo*, or recombinantly expressed in a prokaryotic cell (*e.g.*, bacterial cell, such as *E. coli*) or in a eukaryotic cell such as a mammalian, yeast, insect, fungal or plant cell. The G-CSF peptide can be either a full-length protein or a fragment. Moreover, the G-CSF peptide can be a wild type or mutated peptide. In an exemplary embodiment, the G-CSF peptide includes a mutation that adds one or more N- or O-linked glycosylation sites to the peptide sequence.

[0199] In an exemplary embodiment, Factor IX is O-glycosylated and functionalized with a water-soluble polymer in the following manner. The peptide is either produced with an available amino acid glycosylation site or, if glycosylated, the glycosyl moiety is trimmed off to exposed the amino acid. For example, a serine or threonine is α-1 N-acetyl amino galactosylated (GalNAc) and the NAc-galactosylated peptide is sialylated with a sialic acid-modifying group cassette using ST6GalNAcT1. Alternatively, the NAc-galactosylated peptide is galactosylated using Core-1-GalT-1 and the product is sialylated with a sialic acid-modifying group cassette using ST3GalT1. An exemplary conjugate according to this method has the following linkages: Thr-α-1-GalNAc-β-1,3-Gal-α2,3-Sia*, in which Sia* is the sialic acid-modifying group cassette.

[0200] In the methods of the invention, such as that set forth above, using multiple enzymes and saccharyl donors, the individual glycosylation steps may be performed separately, or combined in a "single pot" reaction. For example, in the three enzyme reaction set forth above the GalNAc tranferase, GalT and SiaT and their donors may be combined in a single vessel. Alternatively, the GalNAc reaction can be performed alone and both the GalT and SiaT and the appropriate saccharyl donors added as a single step. Another mode of running the reactions involves adding each enzyme and an appropriate donor sequentially and conducting the reaction in a "single pot" motif. Combinations of each of the methods set forth above are of use in preparing the compounds of the invention.

[0201] In the conjugates of the invention, particularly the glycopegylated N-linked glycans, the Sia-modifying group cassette can be linked to the Gal in an α-2,6, or α-2,3 linkage.

[0202] The method of the invention also provides for modification of incompletely glycosylated peptides that are produced recombinantly. Employing a modified sugar in a method of the invention, the G-CSF peptide can be simultaneously further glycosylated and derivatized with, *e.g.*, a PEG moiety, therapeutic agent, or the like. The sugar moiety of the modified sugar can be the residue that would properly be conjugated to the acceptor in a fully glycosylated peptide, or another sugar moiety with desirable properties.

[0203] G-CSF peptides modified by the methods of the invention can be synthetic or wild-type peptides or they can be mutated peptides, produced by methods known in the art, such as site-directed mutagenesis. Glycosylation of peptides is typically either N-linked or O-linked. An exemplary N-linkage is the attachment of the modified sugar to the side chain of an asparagine residue. The tripeptide sequences asparagine-X-serine and asparagine-X-threonine, where X is any amino acid except proline, are the recognition sequences for enzymatic attachment of a carbohydrate moiety to the asparagine side chain. Thus, the presence of either of these tripeptide sequences in a polypeptide creates a potential glycosylation site. O-linked glycosylation refers to the attachment of one sugar (*e.g.*, N-aceylgalactosamine, galactose, mannose, GlcNAc, glucose, fucose or xylose) to a the hydroxy side chain of a hydroxyamino acid, preferably serine or threonine, although 5-hydroxyproline or 5-hydroxylysine may also be used.

[0204] For example, in one embodiment, G-CSF is expressed in a mammalian system and modified by treatment of sialidase to trim back terminal sialic acid residues, followed by PEGylation using ST3Gal3 and a donor of PEG-sialic acid.

[0205] In another exemplary embodiment, G-CSF expressed in mammalian cells is first treated with sialidase to trim back terminal sialic acid residues, then PEGylated using ST3Gal3 and a donor of PEG-sialic acid, and then sialylated using ST3Gal3 and a sialic acid donor.

[0206] G-CSF expressed in a mammalian system can also be treated with sialidase and galactosidase to trim back its sialic acid and galactose residues, then galactosylated using a galactose donor and a galactosyltransferase, and then PEGylated using ST3Gal3 and a donor of PEG-sialic acid.

[0207] In yet another examplary embodiment, the G-CSF is not first treated with sialidase, but is glycopegylated using a sialic acid transfer reaction with the modifying group-sialic acid cassette, and an enzyme such as ST3Gal3.

[0208] In a further exemplary embodiment, G-CSF is expressed in insect cells and modified in the following procedure: N-acetylglucosamine is first added to G-CSF using an appropriate N-acetylglucosamine donor and one or more of GnT-I, II, IV, and V; G-CSF is then PEGylated using a donor of PEG-galactose and a galactosyltransferase.

[0209] G-CSF produced in yeast can also be glycopegylated. For example, G-CSF is first treated with endoglycanase to trim back the glycosyl groups, galactosylated using a galactose donor and a galactosyltransferase, and is then PEGylated with ST3Gal3 and a donor of PEG-sialic acid.

[0210] Addition of glycosylation sites to a peptide or other structure is conveniently accomplished by altering the amino acid sequence such that it contains one or more glycosylation sites. The addition may also be made by the incorporation of one or more species presenting an –OH group, preferably serine or threonine residues, within the sequence of the G-CSF peptide (for O-linked glycosylation sites). The addition may be made by mutation or by full chemical synthesis of the G-CSF peptide. The G-CSF peptide amino acid sequence is preferably altered through changes at the DNA level, particularly by mutating the DNA encoding the peptide at preselected bases such that codons are generated that will translate into the desired amino acids. The DNA mutation(s) are preferably made using methods known in the art.

[0211] In an exemplary embodiment, the glycosylation site is added by shuffling polynucleotides. Polynucleotides encoding a candidate peptide can be modulated with DNA shuffling protocols. DNA shuffling is a process of recursive recombination and mutation, performed by random fragmentation of a pool of related genes, followed by reassembly of the fragments by a polymerase chain reaction-like process. *See, e.g.*, Stemmer, *Proc. Natl. Acad. Sci. USA* 91:10747-10751 (1994); Stemmer, *Nature* 370:389-391 (1994); and U.S. Patent Nos. 5,605,793, 5,837,458, 5,830,721 and 5,811,238.

[0212] The present invention also provides means of adding (or removing) one or more selected glycosyl residues to a peptide, after which a modified sugar is conjugated to at least one of the selected glycosyl residues of the peptide. The present embodiment is useful, for example, when it is desired to conjugate the modified sugar to a selected glycosyl residue that is either not present on a peptide or is not present in a desired amount. Thus, prior to coupling a modified sugar to a peptide, the selected glycosyl residue is conjugated to the G-CSF peptide by enzymatic or chemical coupling. In another embodiment, the glycosylation pattern of a glycopeptide is altered prior to the conjugation of the modified sugar by the removal of a carbohydrate residue from the glycopeptide. *See*, for example WO 98/31826.

[0213] Addition or removal of any carbohydrate moieties present on the glycopeptide is accomplished either chemically or enzymatically. Chemical deglycosylation is preferably brought about by exposure of the polypeptide variant to the compound trifluoromethanesulfonic acid, or an equivalent compound. This treatment results in the cleavage of most or all sugars except the linking sugar (N-acetylglucosamine or N-acetylgalactosamine), while leaving the peptide intact. Chemical deglycosylation is described by Hakimuddin *et al., Arch. Biochem. Biophys.* **259**: 52 (1987) and by Edge *et al., Anal. Biochem.* **118**: 131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptide variants can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura *et al., Meth. Enzymol.* **138**: 350 (1987).

[0214] Chemical addition of glycosyl moieties is carried out by any art-recognized method. Enzymatic addition of sugar moieties is preferably achieved using a modification of the methods set forth herein, substituting native glycosyl units for the modified sugars used in the invention. Other methods of adding sugar moieties are disclosed in U.S. Patent No. 5,876,980, 6,030,815, 5,728,554, and 5,922,577.

[0215] Exemplary attachment points for selected glycosyl residue include, but are not limited to: (a) consensus sites for N- and O-glycosylation; (b) terminal glycosyl moieties that are acceptors for a glycosyltransferase; (c) arginine, asparagine and histidine; (d) free carboxyl groups; (e) free sulfhydryl groups such as those of cysteine; (f) free hydroxyl groups such as those of serine, threonine, or hydroxyproline; (g) aromatic residues such as those of phenylalanine, tyrosine, or tryptophan; or (h) the amide group of glutamine. Exemplary methods of use in the present invention are described in WO 87/05330 published Sep. 11, 1987, and in Aplin and Wriston, CRC CRIT. REV. BIOCHEM., pp. 259-306 (1981).

The Methods

[0216] In addition to the conjugates discussed above, the present invention provides methods for preparing these and other conjugates. Moreover, the invention provides methods of preventing, curing or ameliorating a disease state by administering a conjugate of the invention to a subject at risk of developing the disease or a subject that has the disease.

[0217] Thus, the invention provides a method of forming a covalent conjugate between a selected moiety and a G-CSF peptide.

[0218] In exemplary embodiments, the conjugate is formed between a water-soluble polymer, a therapeutic moiety, targeting moiety or a biomolecule, and a glycosylated or non-glycosylated G-CSF peptide. The polymer, therapeutic moiety or biomolecule is conjugated to the G-CSF peptide via a glycosyl linking group, which is interposed between, and covalently linked to both the peptide and the modifying group (e.g., water-soluble polymer). The method includes contacting the G-CSF peptide with a mixture containing a modified sugar and an enzyme, e.g., a glycosyltransferase, that conjugates the modified sugar to the substrate (e.g., peptide, aglycone, glycolipid). The reaction is conducted under conditions appropriate to form a covalent bond between the modified sugar and the G-CSF peptide.

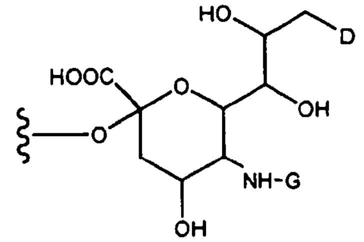
[0219] The acceptor G-CSF peptide is typically synthesized *de novo*, or recombinantly expressed in a prokaryotic cell (*e.g.*, bacterial cell, such as *E. coli*) or in a eukaryotic cell such as a mammalian, yeast, insect, fungal or plant cell. The G-CSF peptide can be either a full-length protein or a fragment. Moreover, the G-CSF peptide can be a wild type or mutated peptide. In an exemplary embodiment, the G-

CSF peptide includes a mutation that adds one or more N- or O-linked glycosylation sites to the peptide sequence.

[0220] The method of the invention also provides for modification of incompletely glycosylated G-CSF peptides that are produced recombinantly. Many recombinantly produced glycoproteins are incompletely glycosylated, exposing carbohydrate residues that may have undesirable properties, e.g., immunogenicity, recognition by the RES. Employing a modified sugar in a method of the invention, the peptide can be simultaneously further glycosylated and derivatized with, e.g., a water-soluble polymer, therapeutic agent, or the like. The sugar moiety of the modified sugar can be the residue that would properly be conjugated to the acceptor in a fully glycosylated peptide, or another sugar moiety with desirable properties.

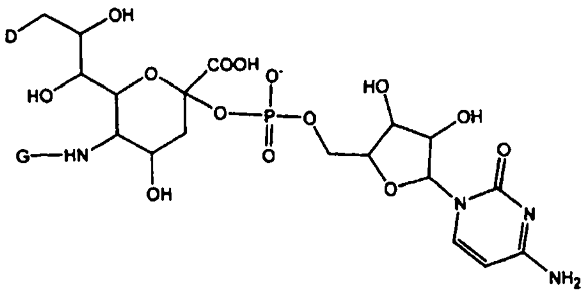
[0221] Exemplary methods of modifying peptides of use in the present invention are set forth in WO04/099231, WO 03/031464, and the references set forth therein.

[0222] In an exemplary embodiment, the invention provides a method of making a PEG-ylated G-CSF comprising the moiety:



wherein D is -OH or R^1-L-HN-. The symbol G represents R^1-L- or -C(O)(C_1-C_6)alkyl. R^1 is a moiety comprising a a straight-chain or branched poly(ethylene glycol) residue. The symbol L represents a linker selected from a bond, substituted or unsubstituted alkyl and substituted or unsubstituted heteroalkyl. In general, when D is OH, G is R^1-L-, and when G is –C(O)(C_1-C_6)alkyl, D is R^1-L-NH-. The method of the invention includes, (a) contacting a substrate G-CSF peptide with a PEG-sialic acid donor and an enzyme that is capable of transferring the PEG-sialic acid moiety from the donor to the substrate G-CSF peptide.

[0223] An exemplary PEG-sialic acid donor is a nucleotide sugar such as that having the formula:



and an enzyme that transfers the PEG-sialic acid onto an amino acid or glycosyl residue of the G-CSF peptide, under conditions appropriate for the transfer.

[0224] In one embodiment the substrate G-CSF peptide is expressed in a host cell prior to the formation of the conjugate of the invention. An exemplary host cell is a mammalian cell. In other embodiments the host cell is an insect cell, plant cell, a bacteria or a fungi.

[0225] The method presented herein is applicable to each of the G-CSF conjugates set forth in the sections above.

[0226] G-CSF peptides modified by the methods of the invention can be synthetic or wild-type peptides or they can be mutated peptides, produced by methods known in the art, such as site-directed mutagenesis. Glycosylation of peptides is typically either N-linked or O-linked. An exemplary N-linkage is the attachment of the modified sugar to the side chain of an asparagine residue. The tripeptide sequences asparagine-X-serine and asparagine-X-threonine, where X is any amino acid except proline, are the recognition sequences for enzymatic attachment of a carbohydrate moiety to the asparagine side chain. Thus, the presence of either of these tripeptide sequences in a polypeptide creates a potential glycosylation site. O-linked glycosylation refers to the attachment of one sugar (*e.g.*, N-acetylgalactosamine, galactose, mannose, GlcNAc, glucose, fucose or xylose) to the hydroxy side chain of a hydroxyamino acid, preferably serine or threonine, although unusual or non-natural amino acids, e.g., 5-hydroxyproline or 5-hydroxylysine may also be used.

[0227] Addition of glycosylation sites to a peptide or other structure is conveniently accomplished by altering the amino acid sequence such that it contains one or more glycosylation sites. The addition may also be made by the incorporation of one or more species presenting an –OH group, preferably serine or threonine residues, within the sequence of the peptide (for O-linked glycosylation sites). The addition may be made by mutation or by full chemical synthesis of the peptide. The peptide amino acid sequence is preferably altered through changes at the DNA level, particularly by mutating the DNA encoding the peptide at preselected bases such that codons are generated that will translate into the desired amino acids. The DNA mutation(s) are preferably made using methods known in the art.

[0228] In an exemplary embodiment, the glycosylation site is added by shuffling polynucleotides. Polynucleotides encoding a candidate peptide can be modulated with DNA shuffling protocols. DNA shuffling is a process of recursive recombination and mutation, performed by random fragmentation of a pool of related genes, followed by reassembly of the fragments by a polymerase chain reaction-like process. *See, e.g.*, Stemmer, *Proc. Natl. Acad. Sci. USA* 91:10747-10751 (1994); Stemmer, *Nature* 370:389-391 (1994); and U.S. Patent Nos. 5,605,793, 5,837,458, 5,830,721 and 5,811,238.

[0229] Exemplary methods of adding or removing glycosylation sites, and adding or removing glycosyl structures or substructures are described in detail in WO04/099231, WO03/031464 and related U.S. and PCT applications.

[0230] The present invention also utilizes means of adding (or removing) one or more selected glycosyl residues to a G-CSF peptide, after which a modified sugar is conjugated to at least one of the selected glycosyl residues of the peptide. Such techniques are useful, for example, when it is desired to conjugate the modified sugar to a selected glycosyl residue that is either not present on a G-CSF peptide or is not present in a desired amount. Thus, prior to coupling a modified sugar to a peptide, the selected glycosyl residue is conjugated to the G-CSF peptide by enzymatic or chemical coupling. In another embodiment, the glycosylation pattern of a glycopeptide is altered prior to the conjugation of the modified sugar by the removal of a carbohydrate residue from the glycopeptide. *See*, for example WO 98/31826.

[0231] Exemplary attachment points for selected glycosyl residue include, but are not limited to: (a) consensus sites for N-linked glycosylation, and sites for O-linked glycosylation; (b) terminal glycosyl moieties that are acceptors for a glycosyltransferase; (c) arginine, asparagine and histidine; (d) free carboxyl groups; (e) free sulfhydryl groups such as those of cysteine; (f) free hydroxyl groups such as those of serine, threonine, or hydroxyproline; (g) aromatic residues such as those of phenylalanine, tyrosine, or tryptophan; or (h) the amide group of glutamine. Exemplary methods of use in the present invention are described in WO 87/05330 published Sep. 11, 1987, and in Aplin and Wriston, CRC CRIT. REV. BIOCHEM., pp. 259-306 (1981).

[0232] The PEG modified sugars are conjugated to a glycosylated or non-glycosylated peptide using an appropriate enzyme to mediate the conjugation. Preferably, the concentrations of the modified donor sugar(s), enzyme(s) and acceptor peptide(s) are selected such that glycosylation proceeds until the desired degree of modification of the acceptor is achieved. The considerations discussed below, while set forth in the context of a sialyltransferase, are generally applicable to other glycosyltransferase reactions.

[0233] A number of methods of using glycosyltransferases to synthesize desired oligosaccharide structures are known and are generally applicable to the instant invention. Exemplary methods are described, for instance, WO 96/32491, Ito *et al.*, *Pure Appl. Chem.* **65**: 753 (1993), U.S. Pat. Nos. 5,352,670, 5,374,541, 5,545,553, and commonly owned U.S. Pat. Nos. 6,399,336, and 6,440,703 which are incorporated herein by reference.

[0234] The present invention is practiced using a single glycosyltransferase or a combination of glycosyltransferases. For example, one can use a combination of a sialyltransferase and a galactosyltransferase. In those embodiments using more than one enzyme, the enzymes and substrates are preferably combined in an initial reaction mixture, or the enzymes and reagents for a second enzymatic reaction are added to the reaction medium once the first enzymatic reaction is complete or nearly complete. By conducting two enzymatic reactions in sequence in a single vessel, overall yields are improved over procedures in which an intermediate species is isolated. Moreover, cleanup and disposal of extra solvents and by-products is reduced.

[0235] In a preferred embodiment, each of the first and second enzyme is a glycosyltransferase. In another preferred embodiment, one enzyme is an endoglycosidase. In an additional preferred embodiment, more than two enzymes are used to assemble the modified glycoprotein of the invention. The enzymes are used to alter a saccharide structure on the G-CSF peptide at any point either before or after the addition of the modified sugar to the peptide.

[0236] In another embodiment, the method makes use of one or more exo- or endoglycosidase. The glycosidase is typically a mutant, which is engineered to form glycosyl bonds rather than rupture them. The mutant glycanase typically includes a substitution of an amino acid residue for an active site acidic amino acid residue. For example, when the endoglycanase is endo-H, the substituted active site residues will typically be Asp at position 130, Glu at position 132 or a combination thereof. The amino acids are generally replaced with serine, alanine, asparagine, or glutamine.

[0237] The mutant enzyme catalyzes the reaction, usually by a synthesis step that is analogous to the reverse reaction of the endoglycanase hydrolysis step. In these embodiments, the glycosyl donor molecule (*e.g.*, a desired oligo- or mono-saccharide structure) contains a leaving group and the reaction proceeds with the addition of the donor molecule to a GlcNAc residue on the protein. For example, the leaving group can be a halogen, such as fluoride. In other embodiments, the leaving group is a Asn, or a Asn-peptide moiety. In yet further embodiments, the GlcNAc residue on the glycosyl donor molecule is modified. For example, the GlcNAc residue may comprise a 1,2 oxazoline moiety.

[0238] In a preferred embodiment, each of the enzymes utilized to produce a conjugate of the invention are present in a catalytic amount. The catalytic amount of a particular enzyme varies according to the concentration of that enzyme's substrate as well as to reaction conditions such as temperature, time and pH value. Means for determining the catalytic amount for a given enzyme under preselected substrate concentrations and reaction conditions are well known to those of skill in the art.

[0239] The temperature at which an above process is carried out can range from just above freezing to the temperature at which the most sensitive enzyme denatures. Preferred temperature ranges are about 0 °C to about 55 °C, and more preferably about 20 ° C to about 37 °C. In another exemplary embodiment, one or more components of the present method are conducted at an elevated temperature using a thermophilic enzyme.

[0240] The reaction mixture is maintained for a period of time sufficient for the acceptor to be glycosylated, thereby forming the desired conjugate. Some of the conjugate can often be detected after a few hours, with recoverable amounts usually being obtained within 24 hours or less. Those of skill in the art understand that the rate of reaction is dependent on a number of variable factors (*e.g*, enzyme concentration, donor concentration, acceptor concentration, temperature, solvent volume), which are optimized for a selected system.

[0241] The present invention also provides for the industrial-scale production of modified peptides. As used herein, an industrial scale generally produces at least one gram of finsihed, purified conjugate.

[0242] In the discussion that follows, the invention is exemplified by the conjugation of modified sialic acid moieties to a glycosylated peptide. The exemplary modified sialic acid is labeled with PEG. The focus of the following discussion on the use of PEG-modified sialic acid and glycosylated peptides is for clarity of illustration and is not intended to imply that the invention is limited to the conjugation of these two partners. One of skill understands that the discussion is generally applicable to the additions of modified glycosyl moieties other than sialic acid. Moreover, the discussion is equally applicable to the modification of a glycosyl unit with agents other than PEG including other PEG moieties, therapeutic moieties, and biomolecules.

[0243] An enzymatic approach can be used for the selective introduction of PEGylated or PPGylated carbohydrates onto a peptide or glycopeptide. The method utilizes modified sugars containing PEG, PPG, or a masked reactive functional group, and is combined with the appropriate glycosyltransferase or glycosynthase. By selecting the glycosyltransferase that will make the desired carbohydrate linkage and utilizing the modified sugar as the donor substrate, the PEG or PPG can be introduced directly onto the G-CSF peptide backbone, onto existing sugar residues of a glycopeptide or onto sugar residues that have been added to a peptide.

[0244] An acceptor for the sialyltransferase is present on the G-CSF peptide to be modified by the methods of the present invention either as a naturally occurring structure or one placed there recombinantly, enzymatically or chemically. Suitable acceptors, include, for example, galactosyl acceptors such as Galβ1,4GlcNAc, Galβ1,4GalNAc, Galβ1,3GalNAc, lacto-N-tetraose, Galβ1,3GlcNAc, Galβ1,3Ara, Galβ1,6GlcNAc, Galβ1,4Glc (lactose), and other acceptors known to those of skill in the art (*see, e.g.*, Paulson *et al.*, *J. Biol. Chem.* **253**: 5617-5624 (1978)).

[0245] In one embodiment, an acceptor for the sialyltransferase is present on the glycopeptide to be modified upon *in vivo* synthesis of the glycopeptide. Such glycopeptides can be sialylated using the claimed methods without prior modification of the glycosylation pattern of the glycopeptide. Alternatively, the methods of the invention can be used to sialylate a peptide that does not include a suitable acceptor; one first modifies the G-CSF peptide to include an acceptor by methods known to those of skill in the art. In an exemplary embodiment, a GalNAc residue is added by the action of a GalNAc transferase.

[0246] In an exemplary embodiment, the galactosyl acceptor is assembled by attaching a galactose residue to an appropriate acceptor linked to the G-CSF peptide, e.g., a GlcNAc. The method includes incubating the G-CSF peptide to be modified with a reaction mixture that contains a suitable amount of a galactosyltransferase (*e.g.*, galβ1,3 or galβ1,4), and a suitable galactosyl donor (*e.g.*, UDP-galactose). The reaction is allowed to proceed substantially to completion or, alternatively, the reaction is terminated when a preselected amount of the galactose residue is added. Other methods of assembling a selected saccharide acceptor will be apparent to those of skill in the art.

[0247] In yet another embodiment, glycopeptide-linked oligosaccharides are first "trimmed," either in whole or in part, to expose either an acceptor for the sialyltransferase or a moiety to which one or more appropriate residues can be added to obtain a suitable acceptor. Enzymes such as glycosyltransferases and endoglycosidases (*see*, for example U.S. Patent No. 5,716,812) are useful for the attaching and trimming reactions.

[0248] In the discussion that follows, the method of the invention is exemplified by the use of modified sugars having a PEG moiety attached thereto. The focus of the discussion is for clarity of illustration. Those of skill will appreciate that the discussion is equally relevant to those embodiments in which the modified sugar bears a therapeutic moiety, biomolecule or the like.

[0249] In an exemplary embodiment of the invention in which a carbohydrate residue is "trimmed" prior to the addition of the modified sugar high mannose is trimmed back to the first generation biantennary structure. A modified sugar bearing a PEG moiety is conjugated to one or more of the sugar residues exposed by the "trimming back." In one example, a PEG moiety is added via a GlcNAc moiety conjugated to the PEG moiety. The modified GlcNAc is attached to one or both of the terminal mannose residues of the biantennary structure. Alternatively, an unmodified GlcNAc can be added to one or both of the termini of the branched species.

[0250] In another exemplary embodiment, a PEG moiety is added to one or both of the terminal mannose residues of the biantennary structure via a modified sugar having a galactose residue, which is conjugated to a GlcNAc residue added onto the terminal mannose residues. Alternatively, an unmodified Gal can be added to one or both terminal GlcNAc residues.

[0251] In yet a further example, a PEG moiety is added onto a Gal residue using a modified sialic acid.

[0252] In another exemplary embodiment, a high mannose structure is "trimmed back" to the mannose from which the biantennary structure branches. In one example, a PEG moiety is added via a GlcNAc modified with the polymer. Alternatively, an unmodified GlcNAc is added to the mannose, followed by a Gal with an attached PEG moiety. In yet another embodiment, unmodified GlcNAc and Gal residues are sequentially added to the mannose, followed by a sialic acid moiety modified with a PEG moiety.

[0253] In a further exemplary embodiment, high mannose is "trimmed back" to the GlcNAc to which the first mannose is attached. The GlcNAc is conjugated to a Gal residue bearing a PEG moiety. Alternatively, an unmodified Gal is added to the GlcNAc, followed by the addition of a sialic acid modified with a water-soluble sugar. In yet a further example, the terminal GlcNAc is conjugated with Gal and the GlcNAc is subsequently fucosylated with a modified fucose bearing a PEG moiety.

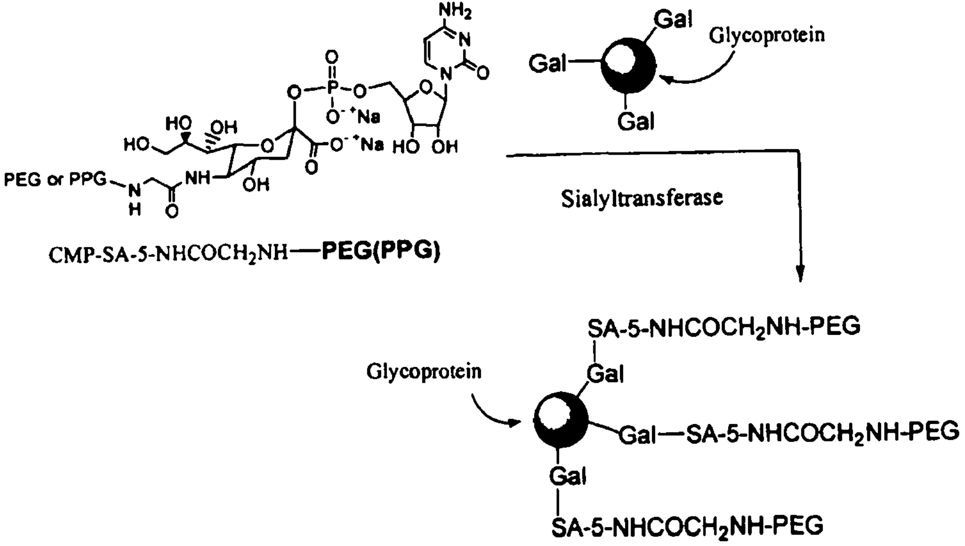
[0254] High mannose may also be trimmed back to the first GlcNAc attached to the Asn of the peptide. In one example, the GlcNAc of the GlcNAc-$(Fuc)_a$ residue is conjugated wit ha GlcNAc bearing a water soluble polymer. In another example, the GlcNAc of the GlcNAc-$(Fuc)_a$ residue is modified with Gal, which bears a water soluble polymer. In a still further embodiment, the GlcNAc is modified with Gal, followed by conjugation to the Gal of a sialic acid modified with a PEG moiety.

[0255] Other exemplary embodiments are set forth in commonly owned U.S. Patent application Publications: 20040132640; 20040063911; 20040137557; U.S. Patent application Nos: 10/369,979; 10/410,913; 10/360,770; 10/410,945 and PCT/US02/32263 each of which is incorporated herein by reference.

[0256] The examples set forth above provide an illustration of the power of the methods set forth herein. Using the methods described herein, it is possible to "trim back" and build up a carbohydrate residue of substantially any desired structure. The modified sugar can be added to the termini of the carbohydrate moiety as set forth above, or it can be intermediate between the peptide core and the terminus of the carbohydrate.

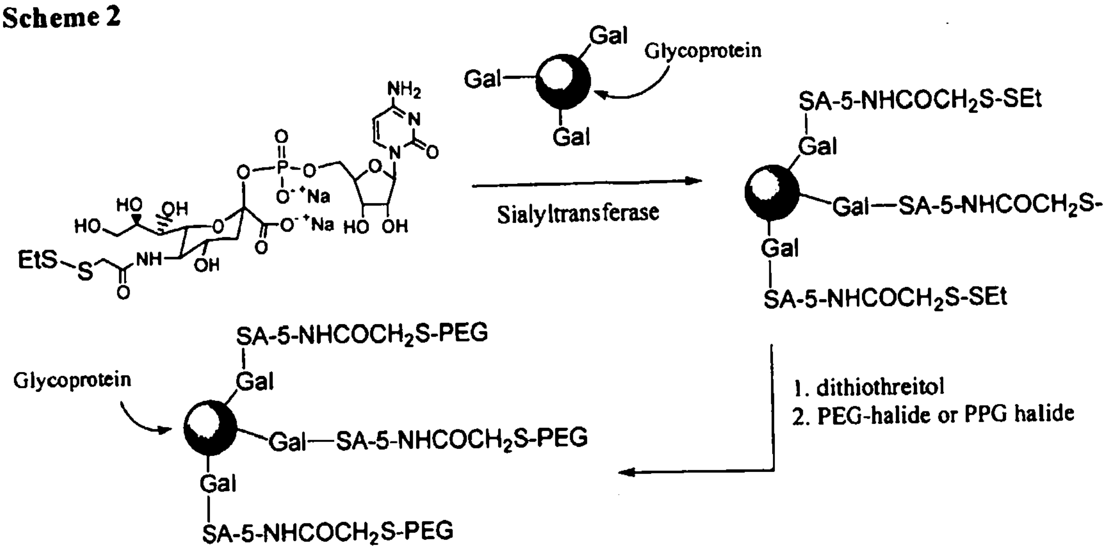
[0257] In an exemplary embodiment, an existing sialic acid is removed from a G-CSF glycopeptide using a sialidase, thereby unmasking all or most of the underlying galactosyl residues. Alternatively, a peptide or glycopeptide is labeled with galactose residues, or an oligosaccharide residue that terminates in a galactose unit. Following the exposure of or addition of the galactose residues, an appropriate sialyltransferase is used to add a modified sialic acid. The approach is summarized in Scheme 1.

Scheme 1



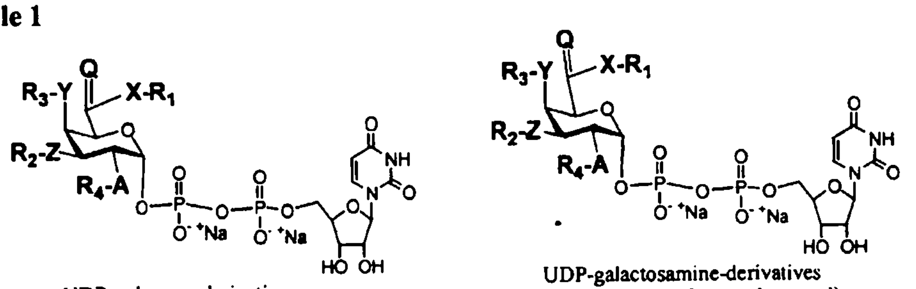
[0258] In yet a further approach, summarized in Scheme 2, a masked reactive functionality is present on the sialic acid. The masked reactive group is preferably unaffected by the conditions used to attach the modified sialic acid to the G-CSF. After the covalent attachment of the modified sialic acid to the G-CSF peptide, the mask is removed and the G-CSF peptide is conjugated with an agent such as PEG. The agent is conjugated to the peptide in a specific manner by its reaction with the unmasked reactive group on the modified sugar residue.

Scheme 2



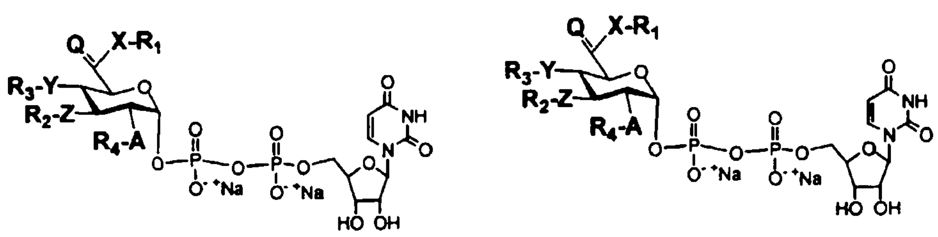
[0259] Any modified sugar set forth herein can be used with its appropriate glycosyltransferase, depending on the terminal sugars of the oligosaccharide side chains of the glycopeptide (Table 1). As discussed above, the terminal sugar of the glycopeptide required for introduction of the PEGylated structure can be introduced naturally during expression or it can be produced post expression using the appropriate glycosidase(s), glycosyltransferase(s) or mix of glycosidase(s) and glycosyltransferase(s).

Table 1



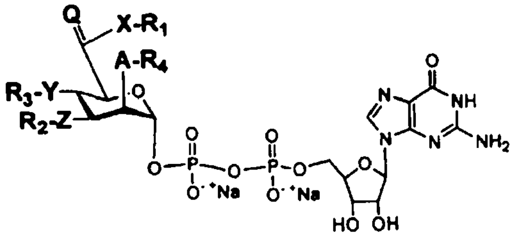
UDP-galactose-derivatives

UDP-galactosamine-derivatives (when A = NH, R_4 may be acetyl)

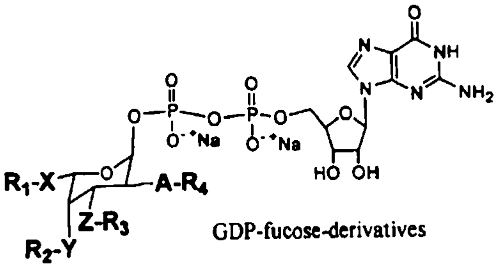


UDP-Glucose-derivatives

UDP-Glucosamine-derivatives (when A = NH, R_4 may be acetyl)



GDP-Mannose-derivatives



GDP-fucose-derivatives

$X = O, NH, S, CH_2, N\text{-}(R_{1\text{-}5})_2.$
$Y = X;\ Z = X;\ A = X;\ B = X.$

$Q = H_2, O, S, NH, N\text{-}R.$

$R, R_{1\text{-}4}$ = H, Linker-M, M.

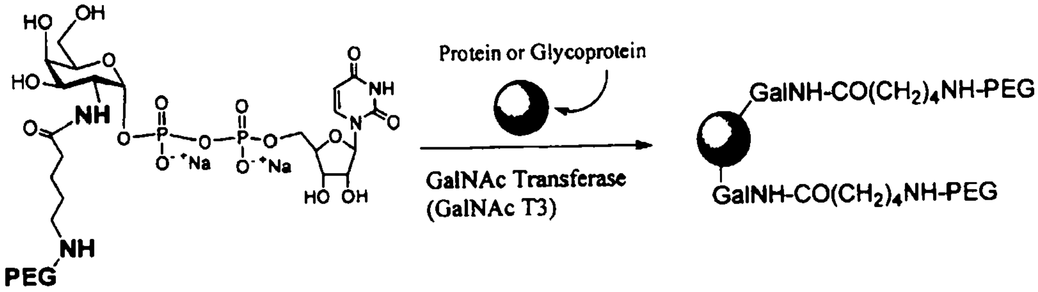
M =PEG, e.g., m-PEG

[0260] In a further exemplary embodiment, UDP-galactose-PEG is reacted with bovine milk β1,4-galactosyltransferase, thereby transferring the modified galactose to the appropriate terminal N-acetylglucosamine structure. The terminal GlcNAc residues on the glycopeptide may be produced during expression, as may occur in such expression systems as mammalian, insect, plant or fungus, but also can be produced by treating the glycopeptide with a sialidase and/or glycosidase and/or glycosyltransferase, as required.

[0261] In another exemplary embodiment, a GlcNAc transferase, such as GNT1-5, is utilized to transfer PEGylated-GlcN to a terminal mannose residue on a glycopeptide. In a still further exemplary embodiment, an the N- and/or O-linked glycan structures are enzymatically removed from a glycopeptide to expose an amino acid or a terminal glycosyl residue that is subsequently conjugated with the modified sugar. For example, an endoglycanase is used to remove the N-linked structures of a glycopeptide to expose a terminal GlcNAc as a GlcNAc-linked-Asn on the glycopeptide. UDP-Gal-PEG and the appropriate galactosyltransferase is used to introduce the PEG-galactose functionality onto the exposed GlcNAc.

[0262] In an alternative embodiment, the modified sugar is added directly to the G-CSF peptide backbone using a glycosyltransferase known to transfer sugar residues to the peptide backbone. This exemplary embodiment is set forth in Scheme 3. Exemplary glycosyltransferases useful in practicing the present invention include, but are not limited to, GalNAc transferases (GalNAc T1-14), GlcNAc transferases, fucosyltransferases, glucosyltransferases, xylosyltransferases, mannosyltransferases and the like. Use of this approach allows the direct addition of modified sugars onto peptides that lack any carbohydrates or, alternatively, onto existing glycopeptides. In both cases, the addition of the modified sugar occurs at specific positions on the peptide backbone as defined by the substrate specificity of the glycosyltransferase and not in a random manner as occurs during modification of a protein's peptide backbone using chemical methods. An array of agents can be introduced into proteins or glycopeptides that lack the glycosyltransferase substrate peptide sequence by engineering the appropriate amino acid sequence into the polypeptide chain.

Scheme 3



[0263] In each of the exemplary embodiments set forth above, one or more additional chemical or enzymatic modification steps can be utilized following the conjugation of the modified sugar to the peptide. In an exemplary embodiment, an enzyme (*e.g.*, fucosyltransferase) is used to append a glycosyl unit (*e.g.*, fucose) onto the terminal modified sugar attached to the G-CSF peptide. In another example, an enzymatic reaction is utilized to "cap" sites to which the modified sugar failed to conjugate. Alternatively, a chemical reaction is utilized to alter the structure of the conjugated modified sugar. For example, the conjugated modified sugar is reacted with agents that stabilize or destabilize its linkage with the peptide component to which the modified sugar is attached. In another example, a component of the modified sugar is deprotected following its conjugation to the peptide. One of skill will appreciate that there is an array of enzymatic and chemical procedures that are useful in the methods of the invention at a stage after the modified sugar is conjugated to the G-CSF peptide. Further elaboration of the modified sugar-peptide conjugate is within the scope of the invention.

Enzymes

[0264] In addition to the enzymes discussed above in the context of forming the acyl-linked conjugate, the glycosylation pattern of the conjugate and the starting substrates (*e.g.*, peptides, lipids) can be elaborated, trimmed back or otherwise modified by methods utilizing other enzymes. The methods of remodeling peptides and lipids using enzymes that transfer a sugar donor to an acceptor are discussed in great detail in DeFrees, WO 03/031464 A2, published April 17, 2003. A brief summary of selected enzymes of use in the present method is set forth below.

Glycosyltransferases

[0265] Glycosyltransferases catalyze the addition of activated sugars (donor NDP- or NMP-sugars), in a step-wise fashion, to a protein, glycopeptide, lipid or glycolipid or to the non-reducing end of a growing oligosaccharide. N-linked glycopeptides are synthesized via a transferase and a lipid-linked oligosaccharide donor Dol-PP-$NAG_2Glc_3Man_9$ in an en block transfer followed by trimming of the core. In this case the nature of the "core" saccharide is somewhat different from subsequent attachments. A very large number of glycosyltransferases are known in the art.

[0266] The glycosyltransferase to be used in the present invention may be any as long as it can utilize the modified sugar as a sugar donor. Examples of such enzymes include Leloir pathway glycosyltransferase, such as galactosyltransferase, N-acetylglucosaminyltransferase, N-acetylgalactosaminyltransferase, fucosyltransferase, sialyltransferase, mannosyltransferase, xylosyltransferase, glucurononyltransferase and the like.

[0267] For enzymatic saccharide syntheses that involve glycosyltransferase reactions, glycosyltransferase can be cloned, or isolated from any source. Many cloned glycosyltransferases are known, as are their polynucleotide sequences. *See, e.g.*, "The WWW Guide To Cloned Glycosyltransferases," (http://www.vei.co.uk/TGN/gt_guide.htm). Glycosyltransferase amino acid sequences and nucleotide sequences encoding glycosyltransferases from which the amino acid sequences can be deduced are also found in various publicly available databases, including GenBank, Swiss-Prot, EMBL, and others.

[0268] Glycosyltransferases that can be employed in the methods of the invention include, but are not limited to, galactosyltransferases, fucosyltransferases, glucosyltransferases, N-acetylgalactosaminyltransferases, N-acetylglucosaminyltransferases, glucuronyltransferases, sialyltransferases, mannosyltransferases, glucuronic acid transferases, galacturonic acid transferases, and oligosaccharyltransferases. Suitable glycosyltransferases include those obtained from eukaryotes, as well as from prokaryotes.

[0269] DNA encoding glycosyltransferases may be obtained by chemical synthesis, by screening reverse transcripts of mRNA from appropriate cells or cell line cultures, by screening genomic libraries from appropriate cells, or by combinations of these procedures. Screening of mRNA or genomic DNA may be carried out with oligonucleotide probes generated from the glycosyltransferases gene sequence. Probes may be labeled with a detectable group such as a fluorescent group, a radioactive atom or a chemiluminescent group in accordance with known procedures and used in conventional hybridization assays. In the alternative, glycosyltransferases gene sequences may be obtained by use of the polymerase chain reaction (PCR) procedure, with the PCR oligonucleotide primers being produced from the glycosyltransferases gene sequence. *See*, U.S. Pat. No. 4,683,195 to Mullis *et al.* and U.S. Pat. No. 4,683,202 to Mullis.

[0270] The glycosyltransferase may be synthesized in host cells transformed with vectors containing DNA encoding the glycosyltransferases enzyme. Vectors are used either to amplify DNA encoding the glycosyltransferases enzyme and/or to express DNA which encodes the glycosyltransferases enzyme. An expression vector is a replicable DNA construct in which a DNA sequence encoding the glycosyltransferases enzyme is operably linked to suitable control sequences capable of effecting the expression of the glycosyltransferases enzyme in a suitable host. The need for such control sequences will vary depending upon the host selected and the transformation method chosen. Generally, control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding sites, and sequences which control the termination of transcription and translation. Amplification vectors do not require expression control domains. All that is needed is the ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants.

[0271] In an exemplary embodiment, the invention utilizes a prokaryotic enzyme. Such glycosyltransferases include enzymes involved in synthesis of lipooligosaccharides (LOS), which are produced by many gram negative bacteria (Preston *et al., Critical Reviews in Microbiology* **23(3)**: 139-180 (1996)). Such enzymes include, but are not limited to, the proteins of the *rfa* operons of species such as *E. coli* and *Salmonella typhimurium*, which include a β1,6 galactosyltransferase and a β1,3 galactosyltransferase (*see, e.g.*, EMBL Accession Nos. M80599 and M86935 (*E. coli*); EMBL Accession No. S56361 (*S. typhimurium*)), a glucosyltransferase (Swiss-Prot Accession No. P25740 (*E. coli*), an β1,2-glucosyltransferase (*rfa*J)(Swiss-Prot Accession No. P27129 (*E. coli*) and Swiss-Prot Accession No. P19817 (*S. typhimurium*)), and an β1,2-N-acetylglucosaminyltransferase (*rfa*K)(EMBL Accession No. U00039 (*E. coli*). Other glycosyltransferases for which amino acid sequences are known include those that are encoded by operons such as *rfa*B, which have been characterized in organisms such as *Klebsiella pneumoniae, E. coli, Salmonella typhimurium, Salmonella enterica, Yersinia enterocolitica, Mycobacterium leprosum*, and the *rh1* operon of *Pseudomonas aeruginosa.*

[0272] Also suitable for use in the present invention are glycosyltransferases that are involved in producing structures containing lacto-N-neotetraose, D-galactosyl-β-1,4-N-acetyl-D-glucosaminyl-β-1,3-D-galactosyl-β-1,4-D-glucose, and the P^k blood group trisaccharide sequence, D-galactosyl-α-1,4-D-galactosyl-β-1,4-D-glucose, which have been identified in the LOS of the mucosal pathogens *Neisseria gonnorhoeae* and *N. meningitidis* (Scholten *et al., J. Med. Microbiol.* **41**: 236-243 (1994)). The genes from *N. meningitidis* and *N. gonorrhoeae* that encode the glycosyltransferases involved in the biosynthesis of these structures have been identified from *N. meningitidis* immunotypes L3 and L1 (Jennings *et al., Mol. Microbiol.* **18**: 729-740 (1995)) and the *N. gonorrhoeae* mutant F62 (Gotshlich, *J. Exp. Med.* **180**: 2181-2190 (1994)). In *N. meningitidis*, a locus consisting of three genes, *lgtA*, *lgtB* and *lg E*, encodes the glycosyltransferase enzymes required for addition of the last three of the sugars in the lacto-*N*-neotetraose chain (Wakarchuk *et al., J. Biol. Chem.* **271**: 19166-73 (1996)). Recently the enzymatic activity of the *lgtB* and *lgtA* gene product was demonstrated, providing the first direct evidence for their proposed glycosyltransferase function (Wakarchuk *et al., J. Biol. Chem.* **271(45)**: 28271-276 (1996)). In *N. gonorrhoeae*, there are two additional genes, *lgtD* which adds β-D-GalNAc to the 3 position of the terminal galactose of the lacto-*N*-

neotetraose structure and *lgtC* which adds a terminal α-D-Gal to the lactose element of a truncated LOS, thus creating the P^k blood group antigen structure (Gotshlich (1994), *supra.*). In *N. meningitidis*, a separate immunotype L1 also expresses the P^k blood group antigen and has been shown to carry an *lgtC* gene (Jennings *et al.*, (1995), *supra.*). *Neisseria* glycosyltransferases and associated genes are also described in USPN 5,545,553 (Gotschlich). Genes for α1,2-fucosyltransferase and α1,3-fucosyltransferase from *Helicobacter pylori* has also been characterized (Martin *et al.*, *J. Biol. Chem.* **272**: 21349-21356 (1997)). Also of use in the present invention are the glycosyltransferases of *Campylobacter jejuni* (*see*, for example, http://afmb.cnrs-mrs.fr/~pedro/CAZY/gtf_42.html).

Fucosyltransferases

[0273] In some embodiments, a glycosyltransferase used in the method of the invention is a fucosyltransferase. Fucosyltransferases are known to those of skill in the art. Exemplary fucosyltransferases include enzymes, which transfer L-fucose from GDP-fucose to a hydroxy position of an acceptor sugar. Fucosyltransferases that transfer non-nucleotide sugars to an acceptor are also of use in the present invention.

[0274] In some embodiments, the acceptor sugar is, for example, the GlcNAc in a Galβ(1→3,4)GlcNAcβ- group in an oligosaccharide glycoside. Suitable fucosyltransferases for this reaction include the Galβ(1→3,4)GlcNAcβ1-α(1→3,4)fucosyltransferase (FTIII E.C. No. 2.4.1.65), which was first characterized from human milk (*see*, Palcic, *et al.*, *Carbohydrate Res.* **190**: 1-11 (1989); Prieels, *et al.*, *J. Biol. Chem.* **256**: 10456-10463 (1981); and Nunez, *et al.*, *Can. J. Chem.* **59**: 2086-2095 (1981)) and the Galβ(1→4)GlcNAcβ- αfucosyltransferases (FTIV, FTV, FTVI) which are found in human serum. FTVII (E.C. No. 2.4.1.65), a sialyl α(2→3)Galβ((1→3)GlcNAcβ fucosyltransferase, has also been characterized. A recombinant form of the Galβ(1→3,4) GlcNAcβ- α(1→3,4)fucosyltransferase has also been characterized (*see*, Dumas, *et al.*, *Bioorg. Med. Letters* **1**: 425-428 (1991) and Kukowska-Latallo, *et al.*, *Genes and Development* **4**: 1288-1303 (1990)). Other exemplary fucosyltransferases include, for example, α1,2 fucosyltransferase (E.C. No. 2.4.1.69). Enzymatic fucosylation can be carried out by the methods described in Mollicone, *et al.*, *Eur. J. Biochem.* **191**: 169-176 (1990) or U.S. Patent No. 5,374,655.

Cells that are used to produce a fucosyltransferase will also include an enzymatic system for synthesizing GDP-fucose.

Galactosyltransferases

[0275] In another group of embodiments, the glycosyltransferase is a galactosyltransferase. Exemplary galactosyltransferases include α(1,3) galactosyltransferases (E.C. No. 2.4.1.151, *see, e.g.*, Dabkowski *et al., Transplant Proc.* **25:**2921 (1993) and Yamamoto *et al. Nature* **345**: 229-233 (1990), bovine (GenBank j04989, Joziasse *et al., J. Biol. Chem.* **264**: 14290-14297 (1989)), murine (GenBank m26925; Larsen *et al., Proc. Nat'l. Acad. Sci. USA* **86**: 8227-8231 (1989)), porcine (GenBank L36152; Strahan *et al., Immunogenetics* **41**: 101-105 (1995)). Another suitable α1,3 galactosyltransferase is that which is involved in synthesis of the blood group B antigen (EC 2.4.1.37, Yamamoto *et al., J. Biol. Chem.* **265**: 1146-1151 (1990) (human)). Yet a further exemplary galactosyltransferase is core Gal-T1.

[0276] Also suitable for use in the methods of the invention are β(1,4) galactosyltransferases, which include, for example, EC 2.4.1.90 (LacNAc synthetase) and EC 2.4.1.22 (lactose synthetase) (bovine (D'Agostaro *et al., Eur. J. Biochem.* **183**: 211-217 (1989)), human (Masri *et al., Biochem. Biophys. Res. Commun.* **157**: 657-663 (1988)), murine (Nakazawa *et al.*, J. Biochem. **104**: 165-168 (1988)), as well as E.C. 2.4.1.38 and the ceramide galactosyltransferase (EC 2.4.1.45, Stahl *et al., J. Neurosci. Res.* **38**: 234-242 (1994)). Other suitable galactosyltransferases include, for example, α1,2 galactosyltransferases (from *e.g., Schizosaccharomyces pombe*, Chapell *et al., Mol. Biol. Cell* **5**: 519-528 (1994)).

Sialyltransferases

[0277] Sialyltransferases are another type of glycosyltransferase that is useful in the recombinant cells and reaction mixtures of the invention. Cells that produce recombinant sialyltransferases will also produce CMP-sialic acid, which is a sialic acid donor for sialyltransferases. Examples of sialyltransferases that are suitable for use in the present invention include ST3Gal III (*e.g.*, a rat or human ST3Gal III), ST3Gal IV, ST3Gal I, ST3GalII, ST6Gal I, ST3Gal V, ST6Gal II, ST6GalNAc I, ST6GalNAc II, and ST6GalNAc III (the sialyltransferase nomenclature used herein is as described in Tsuji *et al., Glycobiology* 6: v-xiv (1996)). An exemplary α(2,3)sialyltransferase referred to as α(2,3)sialyltransferase (EC 2.4.99.6) transfers sialic acid to the non-reducing terminal Gal of a Galβ1→3Glc disaccharide or glycoside. *See*, Van den Eijnden *et al., J. Biol. Chem.* **256:** 3159 (1981), Weinstein *et al., J. Biol. Chem.* **257:** 13845 (1982) and Wen *et al., J. Biol. Chem.* **267:** 21011 (1992). Another exemplary α2,3-sialyltransferase (EC 2.4.99.4) transfers sialic acid to the non-reducing terminal Gal of the disaccharide or glycoside. *see*, Rearick *et al., J. Biol. Chem.* **254:** 4444 (1979) and Gillespie *et al., J. Biol. Chem.* **267:** 21004 (1992). Further exemplary enzymes include Gal-β-1,4-GlcNAc α-2,6 sialyltransferase (*See*, Kurosawa *et al. Eur. J. Biochem.* **219:** 375-381 (1994)).

[0278] Preferably, for glycosylation of carbohydrates of glycopeptides the sialyltransferase will be able to transfer sialic acid to the sequence Galβ1,4GlcNAc-, the most common penultimate sequence underlying the terminal sialic acid on fully sialylated carbohydrate structures (*see*, Table 2).

Table 2: Sialyltransferases which use the Galβ1,4GlcNAc sequence as an acceptor substrate

Sialyltransferase	**Source**	**Sequence(s) formed**	**Ref.**
ST6Gal I	Mammalian	NeuAcα2,6Galβ1,4GlCNAc-	1
ST3Gal III	Mammalian	NeuAcα2,3Galβ1,4GlCNAc- NeuAcα2,3Galβ1,3GlCNAc-	1
ST3Gal IV	Mammalian	NeuAcα2,3Galβ1,4GlCNAc- NeuAcα2,3Galβ1,3GlCNAc-	1
ST6Gal II	Mammalian	NeuAcα2,6Galβ1,4GlCNA	
ST6Gal II	photobacterium	NeuAcα2,6Galβ1,4GlCNAc-	2
ST3Gal V	*N. meningitides* *N. gonorrhoeae*	NeuAcα2,3Galβ1,4GlCNAc-	3

1) Goochee *et al., Bio/Technology* 9: 1347-1355 (1991)
2) Yamamoto *et al., J. Biochem.* **120**: 104-110 (1996)
3) Gilbert *et al., J. Biol. Chem.* **271**: 28271-28276 (1996)

[0279] An example of a sialyltransferase that is useful in the claimed methods is ST3Gal III, which is also referred to as α(2,3)sialyltransferase (EC 2.4.99.6). This enzyme catalyzes the transfer of sialic acid to the Gal of a Galβ1,3GlcNAc or Galβ1,4GlcNAc glycoside (*see, e.g.*, Wen *et al., J. Biol. Chem.* **267**: 21011 (1992);

Van den Eijnden *et al., J. Biol. Chem.* **256**: 3159 (1991)) and is responsible for sialylation of asparagine-linked oligosaccharides in glycopeptides. The sialic acid is linked to a Gal with the formation of an α-linkage between the two saccharides. Bonding (linkage) between the saccharides is between the 2-position of NeuAc and the 3-position of Gal. This particular enzyme can be isolated from rat liver (Weinstein *et al., J. Biol. Chem.* **257**: 13845 (1982)); the human cDNA (Sasaki *et al.* (1993) *J. Biol. Chem.* **268**: 22782-22787; Kitagawa & Paulson (1994) *J. Biol. Chem.* **269**: 1394-1401) and genomic (Kitagawa *et al.* (1996) *J. Biol. Chem.* **271**: 931-938) DNA sequences are known, facilitating production of this enzyme by recombinant expression. In a preferred embodiment, the claimed sialylation methods use a rat ST3Gal III.

[0280] Other exemplary sialyltransferases of use in the present invention include those isolated from Campylobacter jejuni, including CST-I and CST-II and those forming α (2,3) linkages. *See*, e.g, WO99/49051.

[0281] Sialyltransferases other those listed in Table 2, are also useful in an economic and efficient large-scale process for sialylation of commercially important glycopeptides. As a simple test to find out the utility of these other enzymes, various amounts of each enzyme (1-100 mU/mg protein) are reacted with asialo-α_1 AGP (at 1-10 mg/ml) to compare the ability of the sialyltransferase of interest to sialylate glycopeptides relative to either bovine ST6Gal I, ST3Gal III or both sialyltransferases. Alternatively, other glycopeptides or glycopeptides, or N-linked oligosaccharides enzymatically released from the peptide backbone can be used in place of asialo-α_1 AGP for this evaluation. Sialyltransferases with the ability to sialylate N-linked oligosaccharides of glycopeptides more efficiently than ST6Gal I are useful in a practical large-scale process for peptide sialylation.

[0282] These and additional sialyltransferases are set forth in **FIG. 11**, is a table of sialyl transferases that are of use for transferring to an acceptor the modified sialic acid species set forth herein and unmodified sialic acid.

GalNAc transferases

[0283] N-acetylgalactosaminyltransferases are of use in practicing the present invention, particularly for binding a GalNAc moiety to an amino acid of the O-linked glycosylation site of the peptide. Suitable N-acetylgalactosaminyltransferases include, but are not limited to, α(1,3) N-acetylgalactosaminyltransferase, β(1,4) N-acetylgalactosaminyltransferases (Nagata *et al.*, *J. Biol. Chem.* **267**: 12082-12089 (1992) and Smith *et al.*, *J. Biol Chem.* **269**: 15162 (1994)) and polypeptide N-acetylgalactosaminyltransferase (Homa *et al.*, *J. Biol. Chem.* **268**: 12609 (1993)).

[0284] Production of proteins such as the enzyme GalNAc T_{I-XX} from cloned genes by genetic engineering is well known. See, eg., U.S. Pat. No. 4,761,371. One method involves collection of sufficient samples, then the amino acid sequence of the enzyme is determined by N-terminal sequencing. This information is then used to isolate a cDNA clone encoding a full-length (membrane bound) transferase which upon expression in the insect cell line Sf9 resulted in the synthesis of a fully active enzyme. The acceptor specificity of the enzyme is then determined using a semiquantitative analysis of the amino acids surrounding known glycosylation sites in 16 different proteins followed by in vitro glycosylation studies of synthetic peptides. This work has demonstrated that certain amino acid residues are overrepresented in glycosylated peptide segments and that residues in specific positions surrounding glycosylated serine and threonine residues may have a more marked influence on acceptor efficiency than other amino acid moieties.

Cell-Bound Glycosyltransferases

[0285] In another embodiment, the enzymes utilized in the method of the invention are cell-bound glycosyltransferases. Although many soluble glycosyltransferases are known (see, for example, U.S. Pat. No. 5,032,519), glycosyltransferases are generally in membrane-bound form when associated with cells. Many of the membrane-bound enzymes studied thus far are considered to be intrinsic proteins; that is, they are not released from the membranes by sonication and require detergents for solubilization. Surface glycosyltransferases have been identified on the surfaces of vertebrate and invertebrate cells, and it has also been recognized that these surface transferases maintain catalytic activity under physiological conditions. However, the more recognized function of cell surface glycosyltransferases is for intercellular recognition (Roth, MOLECULAR APPROACHES to SUPRACELLULAR PHENOMENA, 1990).

[0286] Methods have been developed to alter the glycosyltransferases expressed by cells. For example, Larsen *et al.*, *Proc. Natl. Acad. Sci. USA* **86**: 8227-8231 (1989), report a genetic approach to isolate cloned cDNA sequences that determine expression of cell surface oligosaccharide structures and their cognate glycosyltransferases. A cDNA library generated from mRNA isolated from a murine cell line known to express UDP-galactose:.β.-D-galactosyl-1,4-N-acetyl-D-glucosaminide α-1,3-galactosyltransferase was transfected into COS-1 cells. The transfected cells were then cultured and assayed for α 1-3 galactosyltransferase activity.

[0287] Francisco *et al., Proc. Natl. Acad. Sci. USA* **89**: 2713-2717 (1992), disclose a method of anchoring β-lactamase to the external surface of *Escherichia coli*. A tripartite fusion consisting of (i) a signal sequence of an outer membrane protein, (ii) a membrane-spanning section of an outer membrane protein, and (iii) a complete mature β-lactamase sequence is produced resulting in an active surface bound β-lactamase molecule. However, the Francisco method is limited only to procaryotic cell systems and as recognized by the authors, requires the complete tripartite fusion for proper functioning.

Sulfotransferases

[0288] The invention also provides methods for producing peptides that include sulfated molecules, including, for example sulfated polysaccharides such as heparin, heparan sulfate, carragenen, and related compounds. Suitable sulfotransferases include, for example, chondroitin-6-sulphotransferase (chicken cDNA described by Fukuta *et al., J. Biol. Chem.* **270**: 18575-18580 (1995); GenBank Accession No. D49915), glycosaminoglycan N-acetylglucosamine N-deacetylase/N-sulphotransferase 1 (Dixon *et al., Genomics* **26**: 239-241 (1995); UL18918), and glycosaminoglycan N-acetylglucosamine N-deacetylase/N-sulphotransferase 2 (murine cDNA described in Orellana *et al., J. Biol. Chem.* **269**: 2270-2276 (1994) and Eriksson *et al., J. Biol. Chem.* **269**: 10438-10443 (1994); human cDNA described in GenBank Accession No. U2304).

Glycosidases

[0289] This invention also encompasses the use of wild-type and mutant glycosidases. Mutant β-galactosidase enzymes have been demonstrated to catalyze the formation of disaccharides through the coupling of an α-glycosyl fluoride to a galactosyl acceptor molecule. (Withers, U.S. Pat. No. 6,284,494; issued Sept. 4, 2001). Other glycosidases of use in this invention include, for example, β-glucosidases, β-galactosidases, β-mannosidases, β-acetyl glucosaminidases, β-N-acetyl galactosaminidases, β-xylosidases, β-fucosidases, cellulases, xylanases, galactanases, mannanases, hemicellulases, amylases, glucoamylases, α-glucosidases, α-galactosidases, α-mannosidases, α-N-acetyl glucosaminidases, α-N-acetyl galactose-aminidases, α-xylosidases, α-fucosidases, and neuraminidases/sialidases.

Immobilized Enzymes

[0290] The present invention also provides for the use of enzymes that are immobilized on a solid and/or soluble support. In an exemplary embodiment, there is provided a glycosyltransferase that is conjugated to a PEG via an intact glycosyl linker according to the methods of the invention. The PEG-linker-enzyme conjugate is optionally attached to solid support. The use of solid supported enzymes in the methods of the invention simplifies the work up of the reaction mixture and purification of the reaction product, and also enables the facile recovery of the enzyme. The glycosyltransferase conjugate is utilized in the methods of the invention. Other combinations of enzymes and supports will be apparent to those of skill in the art.

Fusion Proteins

[0291] In other exemplary embodiments, the methods of the invention utilize fusion proteins that have more than one enzymatic activity that is involved in synthesis of a desired glycopeptide conjugate. The fusion polypeptides can be composed of, for example, a catalytically active domain of a glycosyltransferase that is joined to a catalytically active domain of an accessory enzyme. The accessory enzyme catalytic domain can, for example, catalyze a step in the formation of a nucleotide sugar that is a donor for the glycosyltransferase, or catalyze a reaction involved in a glycosyltransferase cycle. For example, a polynucleotide that encodes a glycosyltransferase can be joined, in-frame, to a polynucleotide that encodes an enzyme involved in nucleotide sugar synthesis. The resulting fusion protein can then catalyze not only the synthesis of the nucleotide sugar, but also the transfer of the sugar moiety to the acceptor molecule. The fusion protein can be two or more cycle enzymes linked into one expressible nucleotide sequence. In other embodiments the fusion protein includes the catalytically active domains of two or more glycosyltransferases. See, for example, 5,641,668. The modified glycopeptides of the present invention can be readily designed and manufactured utilizing various suitable fusion proteins (*see*, for example, PCT Patent Application PCT/CA98/01180, which was published as WO 99/31224 on June 24, 1999.)

Preparation of Modified Sugars

[0292] In general, the sugar moiety or sugar moiety-linker cassette and the PEG or PEG-linker cassette groups are linked together through the use of reactive groups, which are typically transformed by the linking process into a new organic functional group or unreactive species. The sugar reactive functional group(s), is located at any position on the sugar moiety. Reactive groups and classes of reactions useful in practicing the present invention are generally those that are well known in the art of bioconjugate chemistry. Currently favored classes of reactions available with reactive sugar moieties are those, which proceed under relatively mild conditions. These include, but are not limited to nucleophilic substitutions (*e.g.*, reactions of amines and alcohols with acyl halides, active esters), electrophilic substitutions (*e.g.*, enamine reactions) and additions to carbon-carbon and carbon-heteroatom multiple bonds (*e.g.*, Michael reaction, Diels-Alder addition). These and other useful reactions are discussed in, for example, March, ADVANCED ORGANIC CHEMISTRY, 3rd Ed., John Wiley & Sons, New York, 1985; Hermanson, BIOCONJUGATE TECHNIQUES, Academic Press, San Diego, 1996; and Feeney *et al.*, MODIFICATION OF PROTEINS; Advances in Chemistry Series, Vol. 198, American Chemical Society, Washington, D.C., 1982.

[0293] Useful reactive functional groups pendent from a sugar nucleus or modifying group include, but are not limited to:

(a) carboxyl groups and various derivatives thereof including, but not limited to, N-hydroxysuccinimide esters, N-hydroxybenztriazole esters, acid halides, acyl imidazoles, thioesters, p-nitrophenyl esters, alkyl, alkenyl, alkynyl and aromatic esters;

(b) hydroxyl groups, which can be converted to, *e.g.*, esters, ethers, aldehydes, *etc.*

(c) haloalkyl groups, wherein the halide can be later displaced with a nucleophilic group such as, for example, an amine, a carboxylate anion, thiol anion, carbanion, or an alkoxide ion, thereby resulting in the covalent attachment of a new group at the functional group of the halogen atom;

(d) dienophile groups, which are capable of participating in Diels-Alder reactions such as, for example, maleimido groups;

(e) aldehyde or ketone groups, such that subsequent derivatization is possible via formation of carbonyl derivatives such as, for example, imines, hydrazones, semicarbazones or oximes, or via such mechanisms as Grignard addition or alkyllithium addition;

(f) sulfonyl halide groups for subsequent reaction with amines, for example, to form sulfonamides;

(g) thiol groups, which can be, for example, converted to disulfides or reacted with acyl halides;

(h) amine or sulfhydryl groups, which can be, for example, acylated, alkylated or oxidized;

(i) alkenes, which can undergo, for example, cycloadditions, acylation, Michael addition, *etc*; and

(j) epoxides, which can react with, for example, amines and hydroxyl compounds.

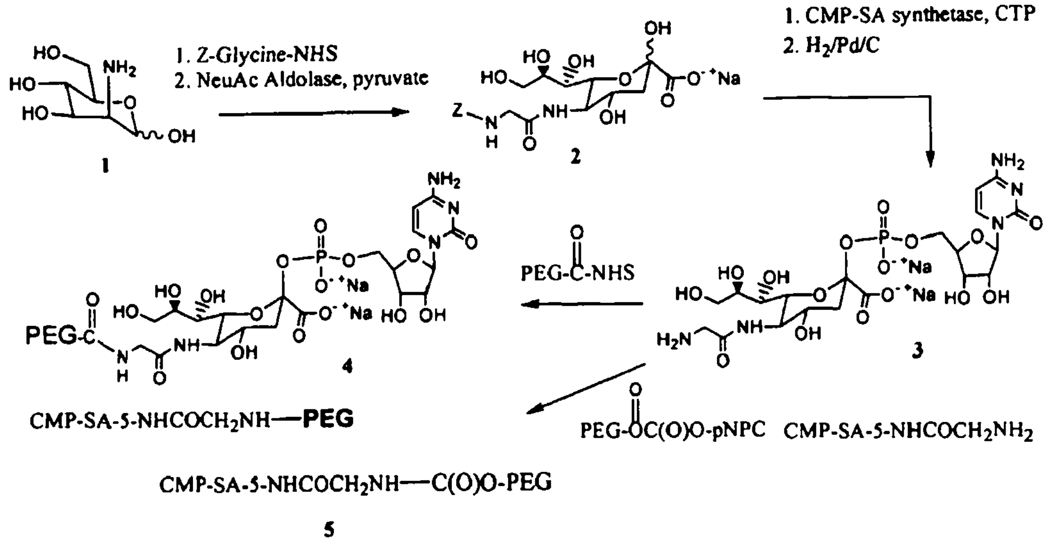
[0294] The reactive functional groups can be chosen such that they do not participate in, or interfere with, the reactions necessary to assemble the reactive sugar nucleus or modifying group. Alternatively, a reactive functional group can be protected from participating in the reaction by the presence of a protecting group. Those of skill in the art understand how to protect a particular functional group such that it does not interfere with a chosen set of reaction conditions. For examples of useful protecting groups, *see*, for example, Greene *et al.*, PROTECTIVE GROUPS IN ORGANIC SYNTHESIS, John Wiley & Sons, New York, 1991.

[0295] In the discussion that follows, a number of specific examples of modified sugars that are useful in practicing the present invention are set forth. In the exemplary embodiments, a sialic acid derivative is utilized as the sugar nucleus to which the modifying group is attached. The focus of the discussion on sialic acid derivatives is for clarity of illustration only and should not be construed to limit the scope of the invention. Those of skill in the art will appreciate that a variety of other sugar moieties can be activated and derivatized in a manner analogous to that set forth using sialic acid as an example. For example, numerous methods are available for modifying galactose, glucose, N-acetylgalactosamine and fucose to name a few sugar substrates, which are readily modified by art recognized methods. *See*, for example, Elhalabi *et al., Curr. Med. Chem.* **6**: 93 (1999); and Schafer *et al., J. Org. Chem.* **65**: 24 (2000)).

[0296] In an exemplary embodiment, the G-CSF peptide that is modified by a method of the invention is a glycopeptide that is produced in mammalian cells (e.g., CHO cells) or in a transgenic animal and thus, contains N- and/or O-linked oligosaccharide chains, which are incompletely sialylated. The oligosaccharide chains of the glycopeptide lacking a sialic acid and containing a terminal galactose residue can be PEGylated, PPGylated or otherwise modified with a modified sialic acid.

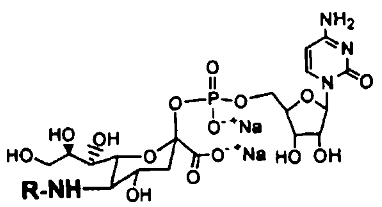
[0297] In Scheme 4, the amino glycoside **1**, is treated with the active ester of a protected amino acid (*e.g.*, glycine) derivative, converting the sugar amine residue into the corresponding protected amino acid amide adduct. The adduct is treated with an aldolase to form α-hydroxy carboxylate **2**. Compound **2** is converted to the corresponding CMP derivative by the action of CMP-SA synthetase, followed by catalytic hydrogenation of the CMP derivative to produce compound **3**. The amine introduced via formation of the glycine adduct is utilized as a locus of PEG attachment by reacting compound **3** with an activated PEG or PPG derivative (*e.g.*, PEG-C(O)NHS, PEG-OC(O)O-p-nitrophenyl), producing species such as **4** or **5**, respectively.

Scheme 4

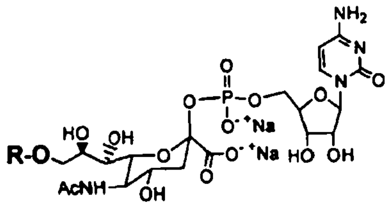


[0298] Table 3 sets forth representative examples of sugar monophosphates that are derivatized with a PEG moiety. Certain of the compounds of Table 3 are prepared by the method of Scheme 4. Other derivatives are prepared by art-recognized methods. *See*, for example, Keppler *et al.*, *Glycobiology* **11**: 11R (2001); and Charter *et al.*, *Glycobiology* **10**: 1049 (2000)). Other amine reactive PEG and PPG analogues are commercially available, or they can be prepared by methods readily accessible to those of skill in the art.

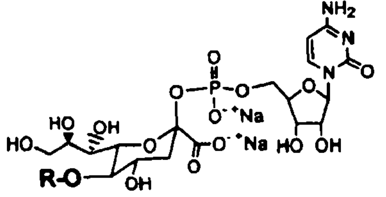
Table 3



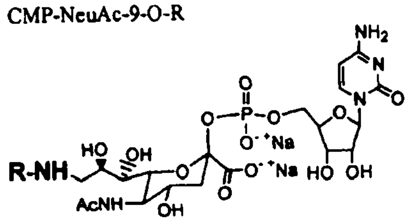
CMP-SA-5-NH-R



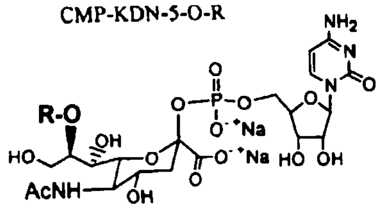
CMP-NeuAc-9-O-R



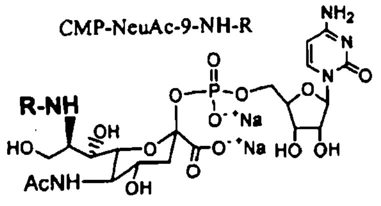
CMP-KDN-5-O-R



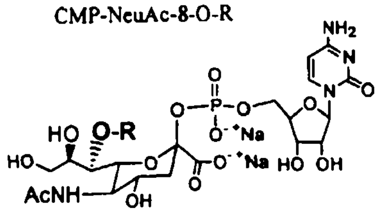
CMP-NeuAc-9-NH-R



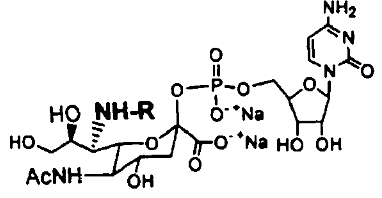
CMP-NeuAc-8-O-R



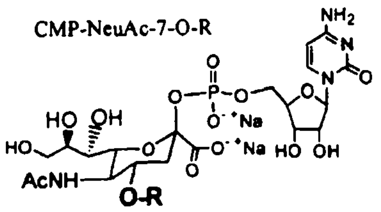
CMP-NeuAc-8-NH-R



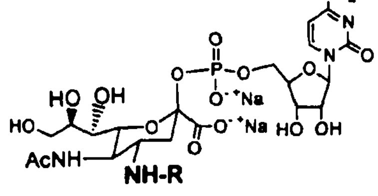
CMP-NeuAc-7-O-R



CMP-NeuAc-7-NH-R

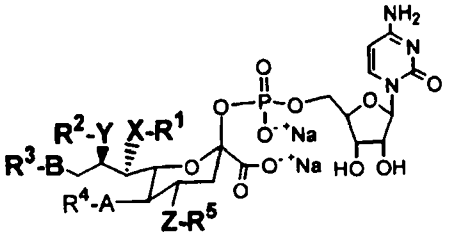


CMP-NeuAc-4-O-R



CMP-NeuAc-4-NH-R

[0299] The modified sugar phosphates of use in practicing the present invention can be substituted in other positions as well as those set forth above. Presently preferred substitutions of sialic acid are set forth in the formula below:



in which X is a linking group, which is preferably selected from –O-, -N(H)-, -S, -CH$_2$-, and N(R)$_2$, in which each R is a member independently selected from R^1-R^5. The symbols Y, Z, A and B each represent a group that is selected from the group set forth above for the identity of X. X, Y, Z, A and B are each independently selected and, therefore, they can be the same or different. The symbols R^1, R^2, R^3, R^4 and R^5 represent H, a PEG moiety, therapeutic moiety, biomolecule or other moiety. Alternatively, these symbols represent a linker that is bound to a PEG moiety, therapeutic moiety, biomolecule or other moiety.

[0300] Exemplary moieties attached to the conjugates disclosed herein include, but are not limited to, PEG derivatives (*e.g.*, acyl-PEG, acyl-alkyl-PEG, alkyl-acyl-PEG carbamoyl-PEG, aryl-PEG), PPG derivatives (*e.g.*, acyl-PPG, acyl-alkyl-PPG, alkyl-acyl-PPG carbamoyl-PPG, aryl-PPG), therapeutic moieties, diagnostic moieties, mannose-6-phosphate, heparin, heparan, SLe$_x$, mannose, mannose-6-phosphate, Sialyl Lewis X, FGF, VFGF, proteins, chondroitin, keratan, dermatan, albumin, integrins, antennary oligosaccharides, peptides and the like. Methods of conjugating the various modifying groups to a saccharide moiety are readily accessible to those of skill in the art (POLY (ETHYLENE GLYCOL CHEMISTRY : BIOTECHNICAL AND BIOMEDICAL APPLICATIONS, J. Milton Harris, Ed., Plenum Pub. Corp., 1992; POLY (ETHYLENE GLYCOL) CHEMICAL AND BIOLOGICAL APPLICATIONS, J. Milton Harris, Ed., ACS Symposium Series No. 680, American Chemical Society, 1997; Hermanson, BIOCONJUGATE TECHNIQUES, Academic Press, San Diego, 1996; and Dunn *et al.*, Eds. POLYMERIC DRUGS AND DRUG DELIVERY SYSTEMS, ACS Symposium Series Vol. 469, American Chemical Society, Washington, D.C. 1991).

Linker Groups (Cross-linking Groups)

[0301] Preparation of the modified sugar for use in the methods of the present invention includes attachment of a PEG moiety to a sugar residue and preferably, forming a stable adduct, which is a substrate for a glycosyltransferase. Thus, it is often preferred to use a linker, e.g., one formed by reaction of the PEG and sugar moiety with a cross-linking agent to conjugate the PEG and the sugar. Exemplary bifunctional compounds which can be used for attaching modifying groups to carbohydrate moieties include, but are not limited to, bifunctional poly(ethyleneglycols), polyamides, polyethers, polyesters and the like. General approaches for linking carbohydrates to other molecules are known in the literature. *See,* for example, Lee *et al., Biochemistry* **28**: 1856 (1989); Bhatia *et al., Anal. Biochem.* **178**: 408 (1989); Janda *et al., J. Am. Chem. Soc.* **112**: 8886 (1990) and Bednarski *et al.*, WO 92/18135. In the discussion that follows, the reactive groups are treated as benign on the sugar moiety of the nascent modified sugar. The focus of the discussion is for clarity of illustration. Those of skill in the art will appreciate that the discussion is relevant to reactive groups on the modifying group as well.

[0302] A variety of reagents are used to modify the components of the modified sugar with intramolecular chemical crosslinks (for reviews of crosslinking reagents and crosslinking procedures see: Wold, F., *Meth. Enzymol.* **25**: 623-651, 1972; Weetall, H. H., and Cooney, D. A., In: ENZYMES AS DRUGS. (Holcenberg, and Roberts, eds.) pp. 395-442, Wiley, New York, 1981; Ji, T. H., *Meth. Enzymol.* **91**: 580-609, 1983; Mattson *et al., Mol. Biol. Rep.* **17**: 167-183, 1993, all of which are incorporated herein by reference). Preferred crosslinking reagents are derived from various zero-length, homo-bifunctional, and hetero-bifunctional crosslinking reagents. Zero-length crosslinking reagents include direct conjugation of two intrinsic chemical groups with no introduction of extrinsic material. Agents that catalyze formation of a disulfide bond belong to this category. Another example is reagents that induce condensation of a carboxyl and a primary amino group to form an amide bond such as carbodiimides, ethylchloroformate, Woodward's reagent K (2-ethyl-5-phenylisoxazolium-3'-sulfonate), and carbonyldiimidazole. In addition to these chemical reagents, the enzyme transglutaminase (glutamyl-peptide γ-glutamyltransferase; EC 2.3.2.13) may be used as zero-length crosslinking reagent. This enzyme catalyzes acyl transfer reactions at carboxamide groups of protein-bound glutaminyl residues, usually with a primary amino group as substrate. Preferred homo- and hetero-bifunctional reagents contain two identical or two dissimilar sites, respectively, which may be reactive for amino, sulfhydryl, guanidino, indole, or nonspecific groups.

Purification of G-CSF Conjugates

Refolding insoluble G-CSF

[0303] Many recombinant proteins expressed in bacteria are expressed as insoluble aggregates in bacterial inclusion bodies. Inclusion bodies are protein deposits found in both the cytoplasmic and periplasmic space of bacteria. (See, *e.g.*, Clark, *Cur. Op. Biotech.* 12:202-207 (2001)). Recombinant G-CSF proteins are expressed in bacterial inclusion bodies, and methods for refolding these proteins to produce active G-CSF proteins are provided herein.

A. Conditions for refolding active G-CSF

[0304] To produce active G-CSF proteins from bacterial cells, G-CSF proteins are expressed in bacterial inclusion bodies, the bacteria are harvested, disrupted and the inclusion bodies are isolated and washed. In one embodiment, three washes are performed: a first wash in a buffer at a pH between 6.0 and 9.0; a monovalent salt, *e.g.*, sodium chloride; a nonionic detergent, *e.g.*, Triton X-100; an ionic detergent, *e.g.*, sodium deoxycholate; and EDTA; a second w ash in a detergent free buffer, and a third wash in H_20. The proteins within the inclusion bodies are then solubilized. Solubilization can be performed using denaturants, guanidiniunl chloride or urea; extremes of pH; or detergents or any combination of these. In one embodiment of 5-6M guanidine HCl or urea are used to solubilize GCSF. In ..mother embodiment, DTT is added.

[0305] After solubilization, denaturants are removed from the GCSF protein mixture. Denaturant removal can be done by a variety of methods, including dilution into a refolding buffet- or buffer exchange methods. Buffer exchange methods include dialysis, diafiltration, g.el filtration, and immobilization of the protein onto a solid support. (*See, e.g.*, Clark, *Cur. Op. Biotech.* 12:202-207 (2001)). Any of the above methods can be combined to remove denaturants.

[0306] Disulfide bond formation in the GCSF proteins is promoted by addition of a refolding buffer comprising a redox couple. Redox couples include reduced and oxidized glutathionc ((-JSF-I/GSSG), cysteine/cystine, cysteamine/cystamine, DTT/GSSG, and DTE/GSSG. (*See, e.g.*, Clark, *Cur. Op. Biotech.* 12:202-207 (2001)). In one embodiment the redox couple is GSH/GSSG at a ratio of 10:1.

[0307] Refolding can be performed in buffers at pH's ranging from, for example, 6.0 to 10.0. Refolding buffers can include other additives to enhance refolding, *e.g.*, L-arginine (0.4-1 M); PEG; low concentrations of denaturants, such as urea (1-2M) and guanidinium chloride (0.5-1.5 M); and detergents (*e.g.*, Chaps, SDS, CTAB, lauryl maltoside, Tween 80, and Triton X-100).

[0308] Alter refolding, the GCSF protein can be dialyzed to remove the redox couple or other unwanted buffer components. In one embodiment, dialysis is performed using a buffer including sodium acetae, glycerol, and a non-ionic detergent, *e.g.*, Tween-80. After dialysis the GCSF protein can be further purified, and/or concentrated by ion exchange chromatography. In one embodiment, an SP-sepharose cation exchange resin is used.

[0309] Those of skill will recognize that a protein has been refolded correctly when the refolded protein has detectable biological activity. For a GCSF protein, biological activity can be measured using a variety of methods. For example, biologically active GCSF proteins are substrates for the O-linked glycosylation described in U.S. Patent Applications 60/535 284, filed January 8, 2004; 60/544411, filed February 12, 2004; and Attorney Docket Number 019957-018820US, filed February 20, 2004; each of which is herein incorporated by reference for all purposes. GCSF protein activity can also be measured using cell proliferation assays or white blood cell (WBC) assays in rats. (Also described in U.S. Patent Applications 60/535284, filed January 8, 2004; 60/544411, filed February 12, 2004; and Attorney Docket Number 019957-018820US, filed February 20, 2004; each of which is herein incorporated by reference for all purposes.) The proliferation assays and the WBC assays can be done before or after O-linked glycosylation of the refolded GCSF proteins.

Other Methods for Isolating Conjugates of the Invention

[0310] Alternatively, the products produced by the above processes can be used without purification. However, it is usually preferred to recover the product. Standard, well-known techniques for recovery of glycosylated saccharides such as thin or thick layer chromatography, column chromatography, ion exchange chromatography, or membrane filtration can be used. It is preferred to use membrane filtration, more preferably utilizing a reverse osmotic membrane, or one or more column chromatographic techniques for the recovery as is discussed hereinafter and in the literature cited herein. For instance, membrane filtration wherein the membranes have molecular weight cutoff of about 3000 to about 10,000 can be used to remove proteins such as glycosyl transferases. Nanofiltration or reverse osmosis can then be used to remove salts and/or purify the product saccharides (*see, e.g.*, WO 98/15581). Nanofilter membranes are a class of reverse osmosis membranes that pass monovalent salts but retain polyvalent salts and uncharged solutes larger than about 100 to about 2,000 Daltons, depending upon the membrane used. Thus, in a typical application, saccharides prepared by the methods of the present invention will be retained in the membrane and contaminating salts will pass through.

[0311] If the modified glycoprotein is produced intracellularly, as a first step, the particulate debris, either host cells or lysed fragments, is removed, for example, by centrifugation or ultrafiltration; optionally, the protein may be concentrated with a commercially available protein concentration filter, followed by separating the polypeptide variant from other impurities by one or more steps selected from immunoaffinity chromatography, ion-exchange column fractionation (*e.g.*, on diethylaminoethyl (DEAE) or matrices containing carboxymethyl or sulfopropyl groups), chromatography on Blue-Sepharose, CM Blue-Sepharose, MONO-Q, MONO-S, lentil lectin-Sepharose, WGA-Sepharose, Con A-Sepharose, Ether Toyopearl, Butyl Toyopearl, Phenyl Toyopearl, or protein A Sepharose, SDS-PAGE chromatography, silica chromatography, chromatofocusing, reverse phase HPLC (*e.g.*, silica gel with appended aliphatic groups), gel filtration using, *e.g.*, Sephadex molecular sieve or size-exclusion chromatography, chromatography on columns that selectively bind the polypeptide, and ethanol or ammonium sulfate precipitation.

[0312] Modified glycopeptides produced in culture are usually isolated by initial extraction from cells, enzymes, etc., followed by one or more concentration, salting-out, aqueous ion-exchange, or size-exclusion chromatography steps. Additionally, the modified glycoprotein may be purified by affinity chromatography. Finally, HPLC may be employed for final purification steps.

[0313] A protease inhibitor, *e.g.*, methylsulfonylfluoride (PMSF) may be included in any of the foregoing steps to inhibit proteolysis and antibiotics may be included to prevent the growth of adventitious contaminants.

[0314] Within another embodiment, supernatants from systems which produce the modified glycopeptide of the invention are first concentrated using a commercially available protein concentration filter, for example, an Amicon or Millipore Pellicon ultrafiltration unit. Following the concentration step, the concentrate may be applied to a suitable purification matrix. For example, a suitable affinity matrix may comprise a ligand for the peptide, a lectin or antibody molecule bound to a suitable support. Alternatively, an anion-exchange resin may be employed, for example, a matrix or substrate having pendant DEAE groups. Suitable matrices include acrylamide, agarose, dextran, cellulose, or other types commonly employed in protein purification. Alternatively, a cation-exchange step may be employed. Suitable cation exchangers include various insoluble matrices comprising sulfopropyl or carboxymethyl groups. Sulfopropyl groups are particularly preferred.

[0315] Finally, one or more RP-HPLC steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups, may be employed to further purify a polypeptide variant composition. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a homogeneous modified glycoprotein.

[0316] The modified glycopeptide of the invention resulting from a large-scale fermentation may be purified by methods analogous to those disclosed by Urdal *et al.*, *J. Chromatog.* **296**: 171 (1984). This reference describes two sequential, RP-HPLC steps for purification of recombinant human IL-2 on a preparative HPLC column. Alternatively, techniques such as affinity chromatography may be utilized to purify the modified glycoprotein.

Pharmaceutical Compositions

[0317] In another aspect, the invention provides a pharmaceutical composition. The pharmaceutical composition includes a pharmaceutically acceptable diluent and a covalent conjugate between a non-naturally-occurring, PEG moiety, therapeutic moiety or biomolecule and a glycosylated or non-glycosylated peptide. The polymer, therapeutic moiety or biomolecule is conjugated to the G-CSF peptide via an intact glycosyl linking group interposed between and covalently linked to both the G-CSF peptide and the polymer, therapeutic moiety or biomolecule.

[0318] Pharmaceutical compositions of the invention are suitable for use in a variety of drug delivery systems. Suitable formulations for use in the present invention are found in *Remington's Pharmaceutical Sciences*, Mace Publishing Company, Philadelphia, PA, 17th ed. (1985). For a brief review of methods for drug delivery, *see,* Langer, *Science* **249**:1527-1533 (1990).

[0319] The pharmaceutical compositions may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (*e.g.*, polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268 and 5,075,109.

[0320] Commonly, the pharmaceutical compositions are administered parenterally, *e.g.*, intravenously. Thus, the invention provides compositions for parenteral administration which comprise the compound dissolved or suspended in an acceptable carrier, preferably an aqueous carrier, *e.g.*, water, buffered water, saline, PBS and the like. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH adjusting and buffering agents, tonicity adjusting agents, wetting agents, detergents and the like.

[0321] These compositions may be sterilized by conventional sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile aqueous carrier prior to administration. The pH of the preparations typically will be between 3 and 11, more preferably from 5 to 9 and most preferably from 7 and 8.

[0322] In some embodiments the glycopeptides of the invention can be incorporated into liposomes formed from standard vesicle-forming lipids. A variety of methods are available for preparing liposomes, as described in, *e.g.*, Szoka *et al.*, *Ann. Rev. Biophys. Bioeng.* **9**: 467 (1980), U.S. Pat. Nos. 4,235,871, 4,501,728 and 4,837,028. The targeting of liposomes using a variety of targeting agents (*e.g.*, the sialyl galactosides of the invention) is well known in the art (*see, e.g.*, U.S. Patent Nos. 4,957,773 and 4,603,044).

[0323] Standard methods for coupling targeting agents to liposomes can be used. These methods generally involve incorporation into liposomes of lipid components, such as phosphatidylethanolamine, which can be activated for attachment of targeting agents, or derivatized lipophilic compounds, such as lipid-derivatized glycopeptides of the invention.

[0324] Targeting mechanisms generally require that the targeting agents be positioned on the surface of the liposome in such a manner that the target moieties are available for interaction with the target, for example, a cell surface receptor. The carbohydrates of the invention may be attached to a lipid molecule before the liposome is formed using methods known to those of skill in the art (*e.g.*, alkylation or acylation of a hydroxyl group present on the carbohydrate with a long chain alkyl halide or with a fatty acid, respectively). Alternatively, the liposome may be fashioned in such a way that a connector portion is first incorporated into the membrane at the time of forming the membrane. The connector portion must have a lipophilic portion, which is firmly embedded and anchored in the membrane. It must also have a reactive portion, which is chemically available on the aqueous surface of the liposome. The reactive portion is selected so that it will be chemically suitable to form a stable chemical bond with the targeting agent or carbohydrate, which is added later. In some cases it is possible to attach the target agent to the connector molecule directly, but in most instances it is more suitable to use a third molecule to act as a chemical bridge, thus linking the connector molecule which is in the membrane with the target agent or carbohydrate which is extended, three dimensionally, off of the vesicle surface.

[0325] The compounds prepared by the methods of the invention may also find use as diagnostic reagents. For example, labeled compounds can be used to locate areas of inflammation or tumor metastasis in a patient suspected of having an inflammation. For this use, the compounds can be labeled with ^{125}I, ^{14}C, or tritium.

[0326] The active ingredient used in the pharmaceutical compositions of the present invention is glycopegylated G-CSF and its derivatives having the biological properties of Follicle Stimulating Hormone to increase e.g., ovulation. Preferably, the G-CSF composition of the present invention is administered parenterally (e.g. IV, IM, SC or IP). Effective dosages are expected to vary considerably depending on the condition being treated and the route of administration but are expected to be in the range of about 0.1 (~7U) to 100 (~7000U) μg/kg body weight of the active material. Preferable doses for treatment of anemic conditions are about 50 to about 300

Units/kg three times a week. Because the present invention provides an G-CSF with an enhanced *in vivo* residence time, the stated dosages are optionally lowered when a composition of the invention is administered.

[0327] The following examples are provided to illustrate the conjugates, and methods and of the present invention, but not to limit the claimed invention.

EXAMPLES

EXAMPLE 1

GlycoPEGylation of G-CSF produced in CHO cells

a. Preparation of Asialo-Granulocyte-Colony Stimulation Factor (G-CSF)

[0328] G-CSF produced in CHO cells is dissolved at 2.5 mg/mL in 50 mM Tris 50 mM Tris-HCl pH 7.4, 0.15 M NaCl, 5 mM $CaCl_2$ and concentrated to 500 μL in a Centricon Plus 20 centrifugal filter. The solution is incubated with 300 mU/mL Neuraminidase II (*Vibrio cholerae*) for 16 hours at 32 °C. To monitor the reaction a small aliquot of the reaction is diluted with the appropriate buffer and a IEF gel performed. The reaction mixture is then added to prewashed N-(*p*-aminophenyl)oxamic acid-agarose conjugate (800 μL/mL reaction volume) and the washed beads gently rotated for 24 hours at 4 °C. The mixture is centrifuged at 10,000 rpm and the supernatant was collected. The beads are washed 3 times with Tris-EDTA buffer, once with 0.4 mL Tris-EDTA buffer and once with 0.2 mL of the Tris-EDTA buffer and all supernatants are pooled. The supernatant is dialyzed at 4 °C against 50 mM Tris –HCl pH 7.4, 1 M NaCl, 0.05% NaN_3 and then twice more against 50 mM Tris –HCl pH 7.4, 1 M NaCl, 0.05% NaN_3. The dialyzed solution is then concentrated using a Centricon Plus 20 centrifugal filter and stored at –20 °C. The conditions for the IEF gel were run according to the procedures and reagents provided by Invitrogen. Samples of native and desialylated G-CSF are dialyzed against water and analyzed by MALDI-TOF MS.

b. Preparation of G-CSF-(alpha2,3)-Sialyl-PEG

[0329] Desialylated G-CSF was dissolved at 2.5 mg/mL in 50 mM Tris-HCl, 0.15 M NaCl, 0.05% NaN_3, pH 7.2. The solution is incubated with 1 mM CMP-sialic acid-PEG and 0.1 U/mL of ST3Gal1 at 32°C for 2 days. To monitor the incorporation of sialic acid-PEG, a small aliquot of the reaction had CMP-SA-PEG-fluorescent ligand added; the label incorporated into the peptide is separated from the free label by gel filtration on a Toso Haas G3000SW analytical column using PBS buffer (pH 7.1). The fluorescent label incorporation into the peptide is quantitated using an in-line fluorescent detector. After 2 days, the reaction mixture is purified using a Toso Haas G3000SW preparative column using PBS buffer (pH 7.1) and collecting fractions based on UV absorption. The product of the reaction is analyzed using SDS-PAGE and IEF analysis according to the procedures and reagents supplied by Invitrogen. Samples of native and PEGylated G-CSF are dialyzed against water and analyzed by MALDI-TOF MS.

c. Preparation of G-CSF-(alpha2,8)-Sialyl-PEG

[0330] G-CSF produced in CHO cells, which contains an alpha2,3-sialylated O-linked glycan, is dissolved at 2.5 mg/mL in 50 mM Tris-HCl, 0.15 M NaCl, 0.05% NaN_3, pH 7.2. The solution is incubated with 1 mM CMP-sialic acid-PEG and 0.1 U/mL of CST-II at 32°C for 2 days. To monitor the incorporation of sialic acid-PEG, a small aliquot of the reaction has CMP-SA-PEG-fluorescent ligand added; the label incorporated into the peptide is separated from the free label by gel filtration on a Toso Haas G3000SW analytical column using PBS buffer (pH 7.1). The fluorescent label incorporation into the peptide is quantitated using an in-line fluorescent detector. After 2 days, the reaction mixture is purified using a Toso Haas G3000SW preparative column using PBS buffer (pH 7.1) and collecting fractions based on UV absorption. The product of the reaction is analyzed using SDS-PAGE and IEF analysis according to the procedures and reagents supplied by Invitrogen. Samples of native and PEGylated G-CSF are dialyzed against water and analyzed by MALDI-TOF MS.

d. Preparation of G-CSF-(alpha2,6)-Sialyl-PEG

[0331] G-CSF, containing only O-linked GalNAc, is dissolved at 2.5 mg/mL in 50 mM Tris-HCl, 0.15 M NaCl, 0.05% NaN_3, pH 7.2. The solution is incubated with 1 mM CMP-sialic acid-PEG and 0.1 U/mL of ST6GalNAcI or II at 32°C for 2 days. To monitor the incorporation of sialic acid-PEG, a small aliquot of the reaction has CMP-SA-PEG-fluorescent ligand added; the label incorporated into the peptide is separated from the free label by gel filtration on a Toso Haas G3000SW analytical column using PBS buffer (pH 7.1). The fluorescent label incorporation into the peptide is quantitated using an in-line fluorescent detector. After 2 days, the reaction mixture is purified using a Toso Haas G3000SW preparative column using PBS buffer (pH 7.1) and collecting fractions based on UV absorption. The product of the reaction is analyzed using SDS-PAGE and IEF analysis according to the procedures and reagents supplied by Invitrogen. Samples of native and PEGylated G-CSF are dialyzed against water and analyzed by MALDI-TOF MS.

[0332] G-CSF produced in CHO cells was treated with Arthrobacter sialidase and was then purified by size exclusion on Superdex 75 and was treated with ST3Gal1 or ST3 Gal2 and then with CMP-SA-PEG 20Kda. The resulting molecule was purified by ion exchange and gel filtration and analysis by SDS PAGE demonstrated that the PEGylation was complete. This is the first demonstration of glycoPEGylation of an O-linked glycan.

EXAMPLE 2

Recombinant GCSF - Expression, refolding and purification

- Harvest cells by centrifugation, discard supernatant. Results of growth on various media are shown in Figure 9.
- Resuspend cell pellet in 10mM Tris pH7.4, 75mM NaCl, 5mM EDTA - use 10ml/g (lysis buffer)
- Microlluidize cells (French press works as well)
- Centrifuge 30min, 4°C at 5,000RPM-discard supernatant
- Resuspend pellet in lysis buffer and centrifuge as above
- Wash IB's in 25mM Tris pH8, 100mM NaCl, 1%TX-100, 1% NaDOC, 5mM EDTA. Pellets are resuspended by pipetting and vortexing. Centrifuge 15min 4°C 5,000RPM. Repeat this step once more (total of two washes)
- Wash pellets two times in 25mM Tris pH8, 100mM NaCl, 5mM EDTA to remove detergents, centrifuge as above
- Resuspend pellets in dH2O to aliquot and centrifuge as above. Pellets are frozen at - 2OC
- IB's are resuspended at 20mg/ml in 6M guanidineHCl, 5mM EDTA, 100mM NaCl, 100mM Tris pH8, 10mM DTT using a pipettor, followed

by rotation for 2-4h at room temperature.

- Centrifuge solubilized IB's for lmin at room temperature at 14,000RPM. Save supernatant.
- Dilute supernatant 1:20 with refold buffer 50mM MES pH6, 240mM NaCl, 10mM
- KC1, 0.3mM lauryl maltoside, 0.055% PEG3350, 1mM GSH, O.1M GSSG, 0.5M arginine and refold on rotator overnight at 4°C.
- Transfer refold to Pierce snakeskin 7kDa MWCO for dialysis. Dialysis buffer 20mM NaOAc pH4, 50mM NaCI, 0.005% Tween-80, 0.1mM EDTA. Dialyze a total of 3 times versus at least a 200 fold excess at 4°C.
- After dialysis pass material through a 0.45μM filter.
- Equlibrate SP-sepharose column with the dialysis buffer and apply sample. Wash column with dialysis buffer and elute with dialysis buffer containing a salt gradient up to 1M NaCl. Protein typically is eluted at 300-400mM NaCl.
- Check material on SDS-PAGE (see *e.g.*, Figure 10).

EXAMPLE 3

The Two Enzyme Method in Two Pots

[0333] The following example illustrates the preparation of G-CSF-GalNAc-SA-PEG in two sequential steps wherein each intermediate product is purified before it is used in the next step.

a. Preparation of G-CSF-GalNAc (pH 6.2) from G-CSF and UDP-GalNAc using GalNAc-T2.

[0334] G-CSF (960 mcg) in 3.2 mL of packaged buffer was concentrated by utrafiltration using an UF filter (MWCO 5K) and then reconstituted with 1 mL of 25 mM MES buffer (pH 6.2, 0.005% NaN_3). UDP-GalNAc (6 mg, 9.24 mM), GalNAc-T2 (40 μL, 0.04 U), and 100 mM $MnCl_2$ (40 μL, 4 mM) were then added and the resulting solution was incubated at room temperature.

[0335] After 24 hrs, MALDI indicated the reaction was complete. The reaction mixture was directly subjected to HPLC purification using SEC (Superdex 75 and

Superdex 200) and an elution buffer comprising of PBS (phosphate buffered saline, pH 4.9 and 0.005% Tween 80). The collected peak of G-CSF-GalNAc was concentrated using a Centricon 5 KDa MWCO filter to about 150 µL and the volume adjusted to 1ml using PBS (phosphate buffered saline, pH 4.9 and 0.005% Tween 80). Final protein concentration 1 mg/mL (A_{280}), yield 100%. The sample was stored at 4 °C.

b. Preparation of G-CSF-GalNAc-SA-PEG using purified G-CSF-GalNAc, CMP-SA-PEG (20KDa) and mouse ST6GalNAc-TI (pH 6.2).

[0336] The G-CSF-GalNAc solution containing 1 mg of protein was buffer exchanged into 25 mM MES buffer (pH 6.2, 0.005% NaN_3) and CMP-SA-PEG (20KDa) (5 mg, 0.25 umol) was added. After dissolving, $MnCl_2$ (100 mcL, 100 mM solution) and ST6GalNAc-I (100 mcL, mouse enzyme) was added and the reaction mixture rocked slowly at 32 °C for three days. The reaction mixture was concentrated by ultrifiltration (MWCO 5K) and buffer exchanged with 25 mM NaOAc (pH 4.9) one time and then concentrated to 1 mL of total volume. The product was then purified using SP-sepharose (A: 25 mM NaOAc+0.005% tween-80 pH 4.5; B: 25 mM NaOAc+0.005% tween-80 pH 4.5+2M NaCl) at retention time 13—18 mins and SEC (Superdex 75; PBS-pH 7.2, 0.005% Tween 80) at retention time 8.6 mins (superdex 75, flow 1 ml/min) The desired fractions were collected, concentrated to 0.5 mL and stored at 4 °C.

EXAMPLE 4

<u>One Pot Method to Make G-CSF-GalNAc -SA-PEG with Simultaneous Addition of Enzymes</u>

[0337] The following example illustrates the preparation of G-CSF-GalNAc -SA-PEG in one pot using simultaneous addition of enzymes

1. One Pot process using mouse ST6GalNAc-I (pH 6.0).

[0338] G-CSF (960 µg of protein dissolved in 3.2 mL of the product formulation buffer) was concentrated by ultrafiltration (MWCO 5K) to 0.5 ml and reconstituted with 25 mM MES buffer (pH 6.0, 0.005% NaN_3) to a total volume of about 1 mL or a protein concentration of 1 mg/mL. UDP-GalNAc (6 mg, 9.21 µmol), GalNAc-T2 (80 μL, 80 mU), CMP-SA-PEG (20KDa) (6 mg, 0,3 μmol) and mouse enzyme ST6GalNAc-I (120 μL) and 100 mM $MnCl_2$(50 μL) were then added. The solution was rocked at 32°C for 48 hrs and purified using standard chromatography conditions on SP-sepharose. A total of 0.5 mg of protein (A_{280}) was obtained or about a 50% overall yield. The product structure was confirmed by analysis with both MALDI and SDS-PAGE.

2. *One pot process using chicken ST6GalNAc-I (pH 6.0).*

[0339] 14.4 mg of G-CSF; was concentrated to 3 mL final volume, buffer exchanged with 25 mM MES buffer (pH 6.0, 0.05% NaN_3, 0.004% Tween 80) and the volume was adjusted to 13 mL. The UDP-GalNAc (90 mg, 150 μmole), GalNAc-T2 (0.59 U), CMP-SA-PEG-20KDa (90 mg), chicken ST6GalNAc-I (0.44 U), and 100 mM $MnCl_2$ (600 mcL) were then added. The resulting mixture stood at room temperature for 60 hrs. The reaction mixture was then concentrated using a UF (MWCO 5K) and centrifugation. The residue (about 2 mL) was dissolved in 25 mM NaOAc buffer (pH 4.5) and concentrated again to 5 mL final volume. This sample was purified using SP-sepharose for about 10-23 min, SEC (Superdex 75, 17 min, flow rate 0.5 ml/min) and an additional SEC (Superdex 200, 23 min, flow rate 0.5 ml/min), to yield 3.6 mg (25% overall yield) of G-CSF-GalNAc-SA-PEG-20 KDa (A_{280} and BCA method).

EXAMPLE 5

<u>One Pot Method to Make G-CSF-GalNAc-Gal-SA-PEG with Sequential Addition of Enzymes</u>

[0340] The following example illustrates a method for making G-CSF-GalNAc-Gal-SA-PEG in one pot with sequential addition of enzymes.

1. *Starting from GalNAc-G-CSF*

a. Preparation of G-CSF-GalNAc (pH 6.2) from G-CSF and UDP-GalNAc using GalNAc-T2.

[0341] G-CSF (960 mcg) in 3.2 mL of packaged buffer was concentrated by utrafiltration using an UF filter (MWCO 5K) and then reconstituted with 1 mL of 25 mM MES buffer (pH 6.2, 0.005% NaN_3). UDP-GalNAc (6 mg, 9.24 mM), GalNAc-

T2 (40 μL, 0.04 U), and 100 mM $MnCl_2$ (40 μL, 4 mM) were then added and the resulting solution was incubated at room temperature.

b. Preparation of G-CSF-GalNAc-Gal-SA-PEG from G-CSF-GalNAc ; UDP-Galactose, SA-PEG-20Kdalton, and the Appropriate Enzymes

[0342] The UDP-Galactose (4 mg, 6.5 μmoles), core-1-Gal-T (320 μL, 160 mU), CMP-SA-PEG-20KDa (8 mg, 0.4 μmole), ST3Gal2 (80 μL, 0.07 mU) and 100 mM $MnCl_2$(80 μL) were directly added to the crude reaction mixture of the G-CSF-GalNAc (1.5 mg) in 1.5 ml 25 mM MES buffer (pH 6.0) from step a, above. The resulting mixture was incubated at 32°C for 60 hrs. The reaction mixture was centrifuged and the solution was concentrated using ultrafiltration (MWCO 5K) to 0.2 mL, and then redissolved with 25 mM NaOAc (pH 4.5) to a final volume of 1 mL. The product was purified using SP-sepharose (retention time of between 10-15 min), the peak fraction were concentrated using a spin filter (MWCO 5K) and the residue purified further using SEC (Superdex 75, retention time of 10.2 min). After concentration using a spin filter (MWCO 5K), the protein was diluted to 1 mL using formulation buffer with PBS, 2.5% mannitol, 0.005% polysorbate, pH 6.5 and formulated at a protein concentration of 850 mcg protein per mL (A_{280}). The overall yield was 55%.

EXAMPLE 6

<u>One Pot Method to Make G-CSF-GalNAc-Gal-SA-PEG with Simultaneous Addition of Enzymes</u>

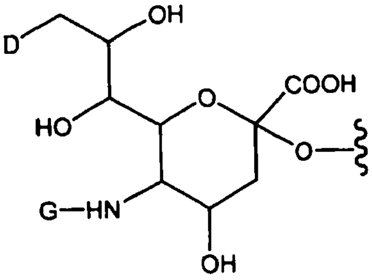
a. Starting from G-CSF.

[0343] G-CSF (960 mcg, 3.2 ml) was concentrated by ultrafiltration (MWCO 5K) and reconstituted with 25 mM Mes buffer (pH 6.0, 0.005% NaN_3). The total volume of the G-CSF solution was about 1 mg/ml. UDP-GalNAc (6 mg), GalNAc-T2 (80 μL, ~80 μU), UDP-Gal (6 mg), Core1 GalT (160 μL, 80 μU), CMP-SA-PEG(20K) (6 mg) and a 2,3-(O)-sialyltransferase (160 μL, 120 μU), 100 mM $MnCl_2$ (40 μL) were added. The resulting mixture was incubated at 32°C for 48 h. Purification was performed as described below using IEX and SEC. The resulting fraction containing the product were concentrated using ultrafiltration (MWCO 5K) and the volume was adjusted to about 1 mL with buffer. The protein concentration was determined to be 0.392 mg/ml by A280, giving an overall yield of 40% from G-CSF.

[0344] It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

WHAT IS CLAIMED IS:

1. A Granulocyte Colony Stimulating Factor peptide comprising the moiety:



wherein

D is a member selected from -OH and R^1-L-HN-;

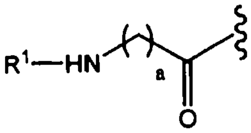
G is a member selected from R^1-L- and -C(O)(C_1-C_6)alkyl;

R^1 is a moiety comprising a member selected a moiety comprising a straight-chain or branched poly(ethylene glycol) residue; and

L is a linker which is a member selected from a bond, substituted or unsubstituted alkyl and substituted or unsubstituted heteroalkyl,

such that when D is OH, G is R^1-L-, and when G is –C(O)(C_1-C_6)alkyl, D is R^1-L-NH-.

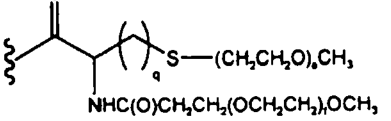
2. The peptide according to claim 1, wherein L-R^1 has the formula:



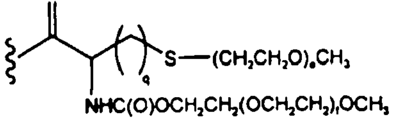
wherein

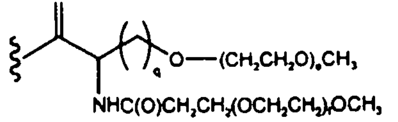
a is an integer from 0 to 20.

3. The peptide according to claim 1, wherein R^1 has a structure that is a member selected from:

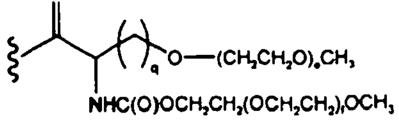


;





; and

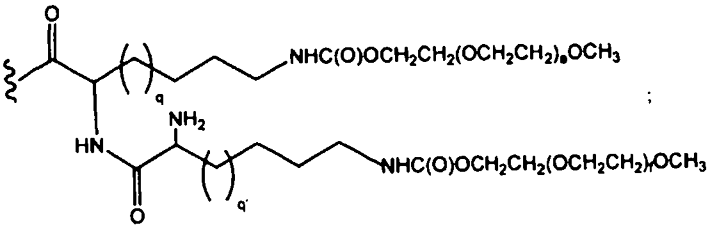


wherein

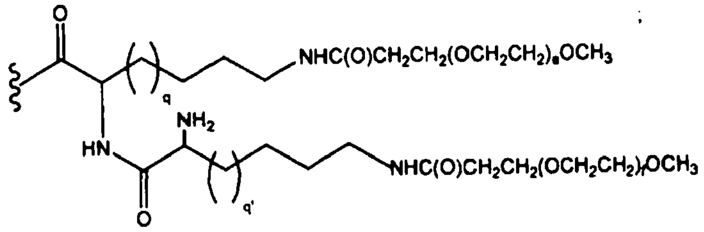
e and f are integers independently selected from 1 to 2500; and

q is an integer from 0 to 20.

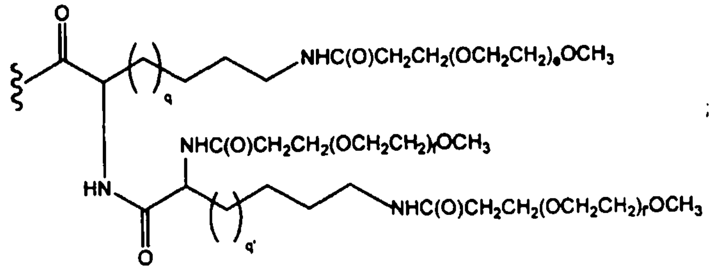
4. The peptide according to claim 1, wherein R^1 has a structure that is a member selected from:



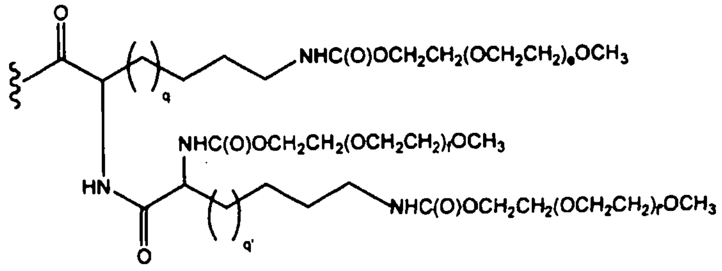
;



;



; and

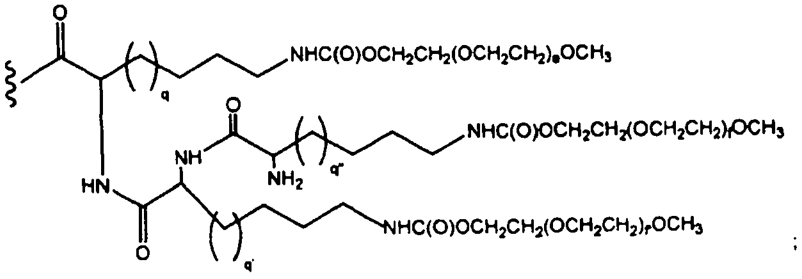


wherein

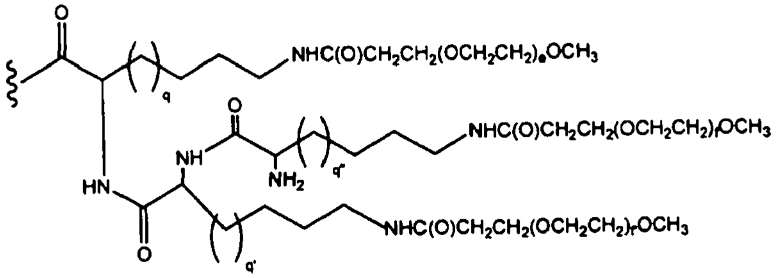
e, f and f' are integers independently selected from 1 to 2500; and

q and q' are integers independently selected from 1 to 20.

5. The peptide according to claim 1, wherein R^1 has a structure that is a member selected from:



; and



wherein

e, f and f' are integers independently selected from 1 to 2500; and

q, q' and q"are integers independently selected from 1 to 20.

6. The peptide according to claim 1, wherein R^1 has a structure that is a member selected from:

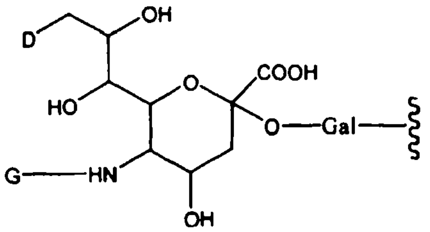
$\wr\!-\mathrm{C(O)CH_2CH_2(OCH_2CH_2)_eOCH_3}$; and

$\wr\!-\mathrm{C(O)OCH_2CH_2(OCH_2CH_2)_fOCH_3}$

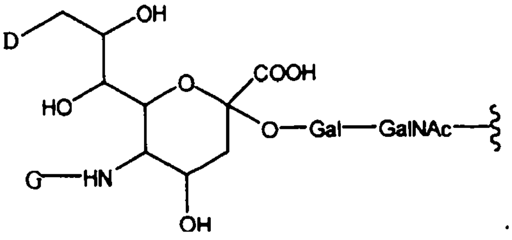
wherein

e and f are integers independently selected from 1 to 2500.

7. The G-CSF peptide according to claim 1, wherein said moiety has the formula:

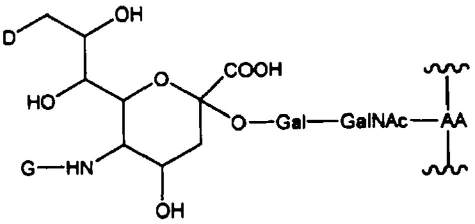


8. The G-CSF peptide according to claim 1, wherein said moiety has the formula:



.

9. The G-CSF peptide according to claim 1, wherein said moiety has the formula:



wherein

AA is an amino acid residue of said peptide.

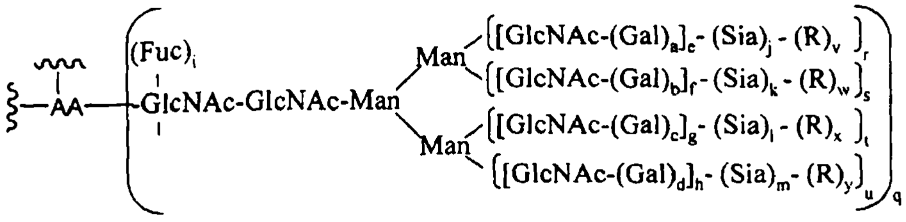
10. The G-CSF peptide according to claim 9, wherein said amino acid residue is a member selected from serine or threonine.

11. The G-CSF peptide according to claim 1, wherein said peptide has the amino acid sequence of SEQ. ID. NO:1.

12. The G-CSF peptide according to claim 11, wherein said amino acid residue is threonine at position 133 of SEQ. ID. NO:1.

13. The peptide according to claim 1, wherein said peptide has an amino acid sequence selected from SEQ. ID. NO:1 and SEQ ID NO:2.

14. The G-CSF peptide according to claim 1, wherein said moiety has the formula:



wherein

a, b, c, d, i, r, s, t, and u are integers independently selected from 0 and 1;

q is 1;

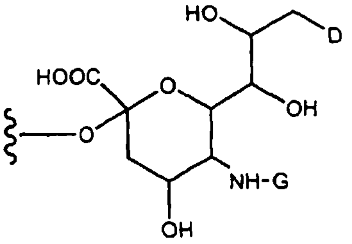
e, f, g, and h are members independently selected from the integers from 0 to 6;

j, k, l, and m are members independently selected from the integers from 0 and 100;

v, w, x, and y are independently selected from 0 and 1, and least one of v, w, x and y is 1;

AA is an amino acid residue of said G-CSF peptide;

Sia-(R) has the formula:



wherein

D is a member selected from -OH and R^1-L-HN-;

G is a member selected from R^1-L- and -C(O)(C$_1$-C$_6$)alkyl;

R^1 is a moiety comprising a member selected a straight-chain or branched poly(ethylene glycol) residue; and

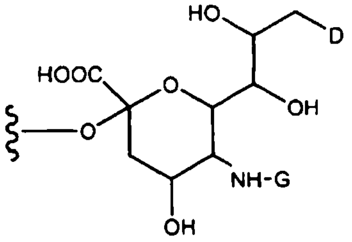
L is a linker which is a member selected from a bond, substituted or unsubstituted alkyl and substituted or unsubstituted heteroalkyl,

such that when D is OH, G is R^1-L-, and when G is –C(O)(C$_1$-C$_6$)alkyl, D is R^1-L-NH-.

15. The peptide according to claim 14, wherein said amino acid residue is an asparagine residue.

16. The peptide according to claim 1, wherein said peptide is a bioactive Granulocyte Colony Stimulating Factor peptide.

17. A method of making a G-CSF peptide conjugate comprising the moiety:



wherein

D is a member selected from -OH and R^1-L-HN-;

G is a member selected from R^1-L- and -C(O)(C$_1$-C$_6$)alkyl;

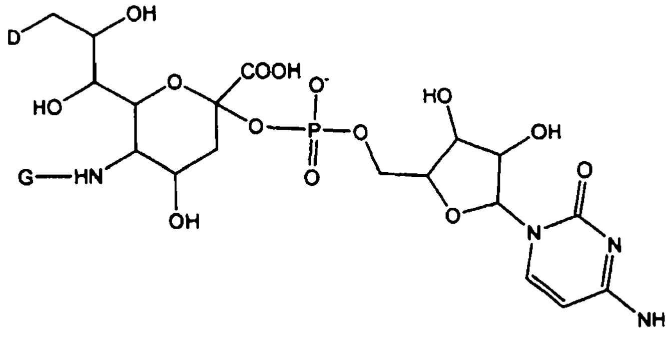
R^1 is a moiety comprising a member selected a straight-chain or branched poly(ethylene glycol) residue; and

L is a linker which is a member selected from a bond, substituted or unsubstituted alkyl and substituted or unsubstituted heteroalkyl,

such that when D is OH, G is R^1-L-, and when G is –C(O)(C$_1$-C$_6$)alkyl, D is R^1-L-NH-,

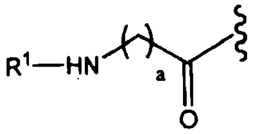
said method comprising:

(a) contacting a substrate G-CSF peptide with a PEG-sialic acid donor moiety having the formula:



and an enzyme that transfers said PEG-sialic acid onto an amino acid or glycosyl residue of said G-CSF peptide, under conditions appropriate for the transfer.

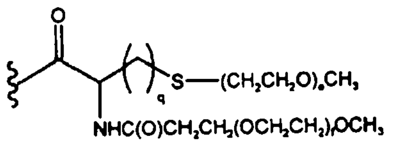
18. The method according to claim 17, wherein L-R^1 has the formula:



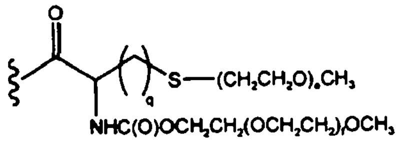
wherein

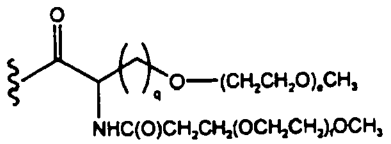
a is an integer from 0 to 20.

19. The method according to claim 17, wherein R^1 has a structure that is a member selected from:

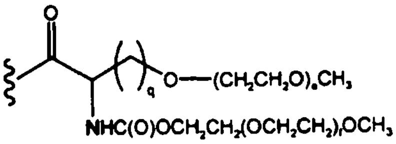


;





; and

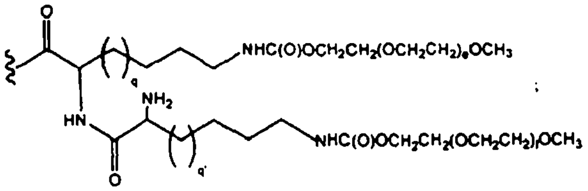


wherein

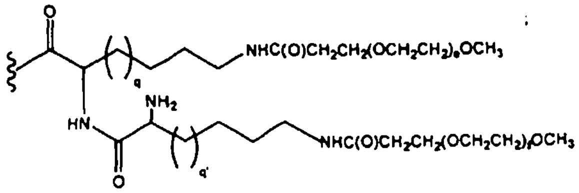
e and f are integers independently selected from 1 to 2500; and

q is an integer from 0 to 20.

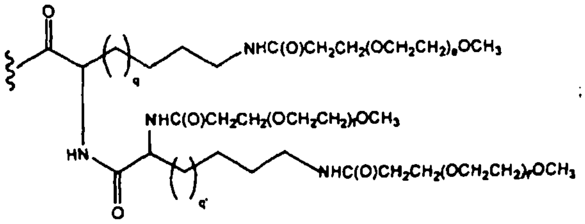
20. The method according to claim 17, wherein R^1 has a structure that is a member selected from:



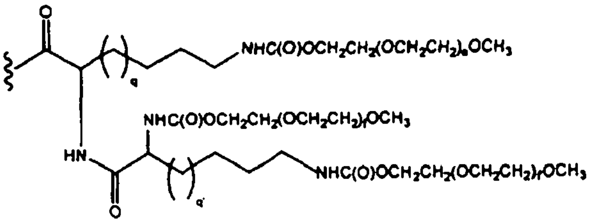
;



;



; and

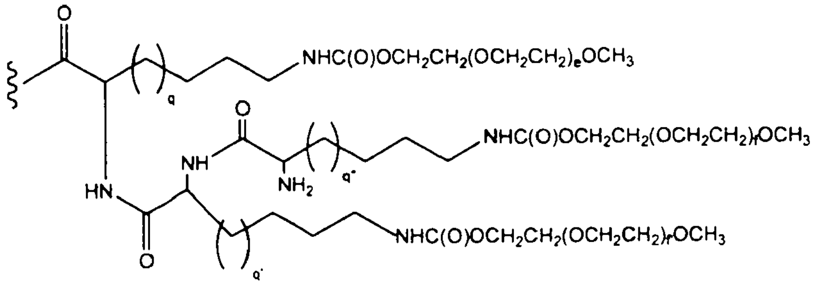


wherein

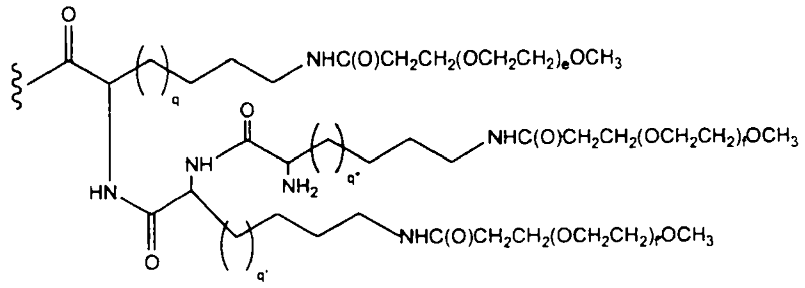
e, f and f' are integers independently selected from 1 to 2500; and

q and q' are integers independently selected from 1 to 20.

21. The method according to claim 17, wherein R^1 has a structure that is a member selected from:



; and



wherein

e, f and f' are integers independently selected from 1 to 2500; and

q, q' and q"are integers independently selected from 1 to 20.

22. The method according to claim 17, wherein R^1 has a structure that is a member selected from:

$\wr\!-C(O)CH_2CH_2(OCH_2CH_2)_eOCH_3$; and

$\wr\!-C(O)OCH_2CH_2(OCH_2CH_2)_fOCH_3$

wherein

e and f are integers independently selected from 1 to 2500.

23. The method of claim **17**, further comprising, prior to step (a):
(b) expressing said substrate Granulocyte Colony Stimulating Factor peptide in a suitable host.

24. The method of claim **17**, wherein said host is selected from an insect cell and a mammalian cell.

25. Use of a peptide according to claim 1 in the manufacture of a preparation for stimulating inflammatory leukocyte production in a mammal.

26. Use of a peptide according to claim 1, in the manufacture of a preparation for treating infection in a subject in need thereof.

27. A pharmaceutical formulation comprising the Granulocyte Colony Stimulating Factor peptide according to claim 1, and a pharmaceutically acceptable carrier.

28. A method of refolding an insoluble recombinant granulocyte colony stimulating factor (GCSF) protein, the method comprising the steps of:
(a) solubilizing the GCSF protein; and
(b) contacting the soluble GCSF protein with a buffer comprising a redox couple to refold the GCSF protein, wherein the refolded GCSF protein is biologically active.

29. A peptide according to any one of claims 1 to 16, substantially as herein described with reference to and as illustrated in any of the examples.

30. A method according to any one of claims 17 to 24, substantially as herein described with reference to and as illustrated in any of the examples.

31. Use according to claim 25 or claim 26, substantially as herein described with reference to and as illustrated in any of the examples.

32. A formulation according to claim 27, substantially as herein described with reference to and as illustrated in any of the examples.

33. A method according to claim 28, substantially as herein described with reference to and as illustrated in any of the examples.

FIGURE 1

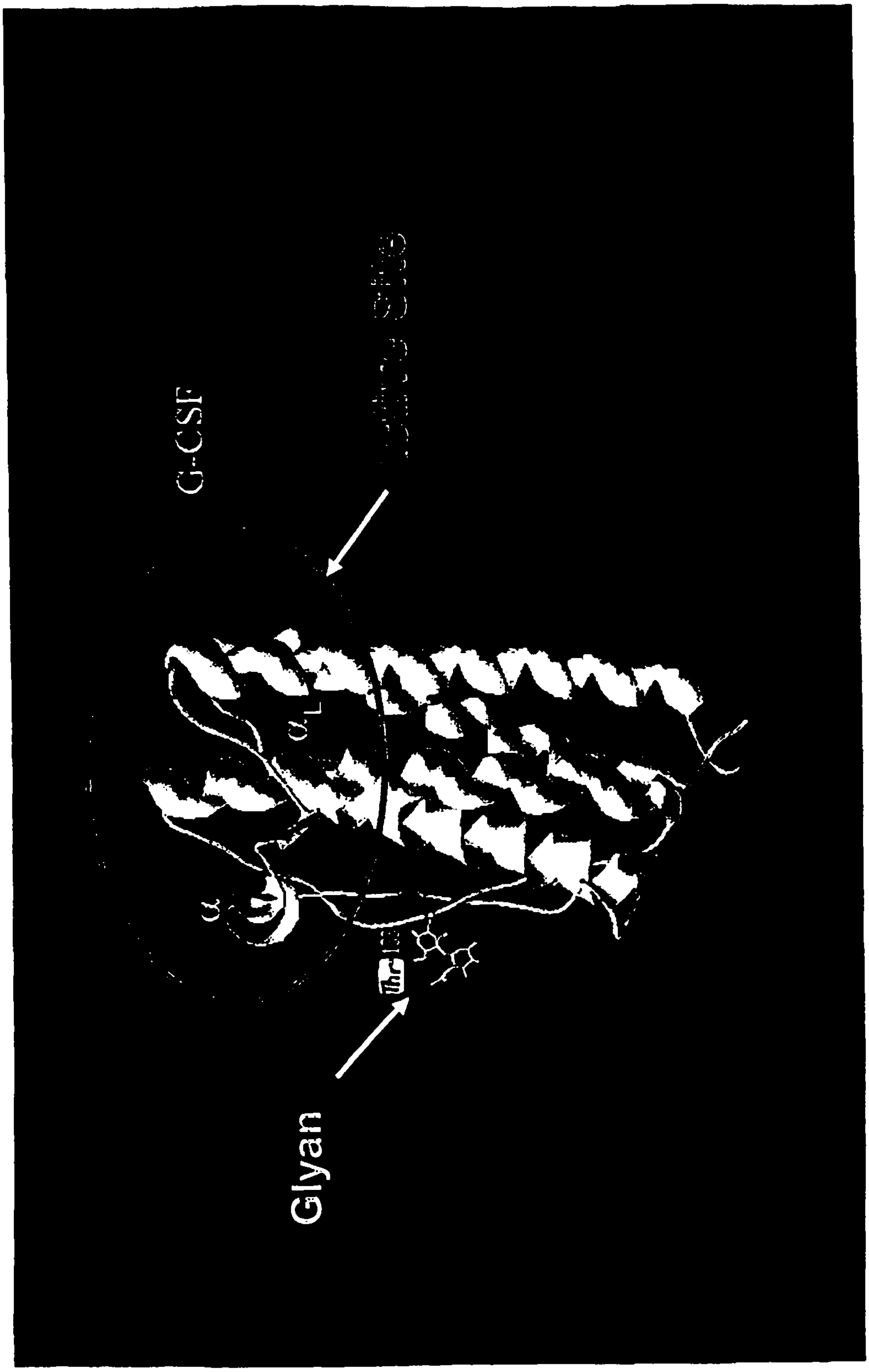


FIGURE 2

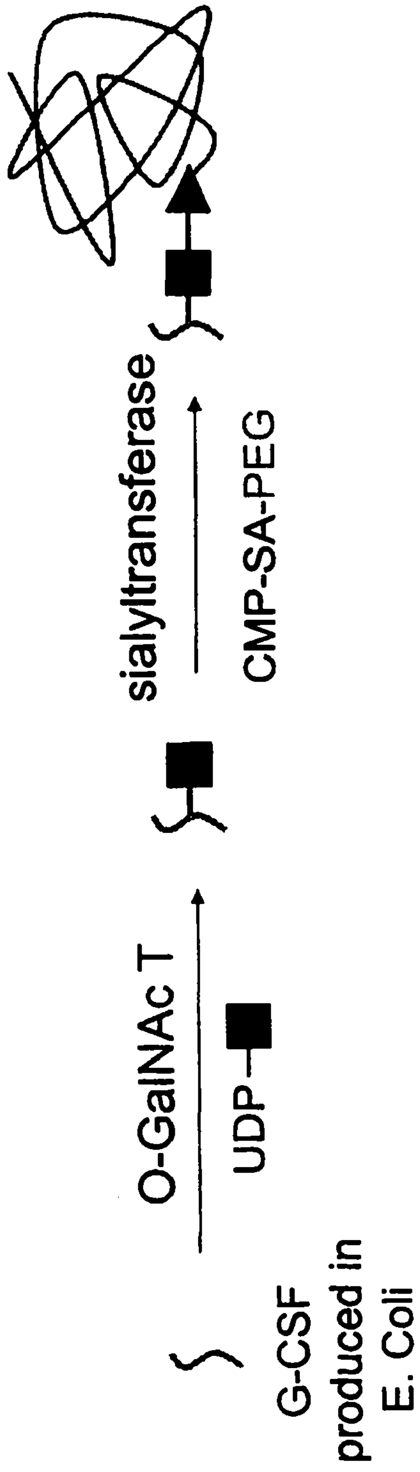

G-CSF produced in E. Coli
O-GalNAc T
UDP—■
sialyltransferase
CMP-SA-PEG

FIGURE 3

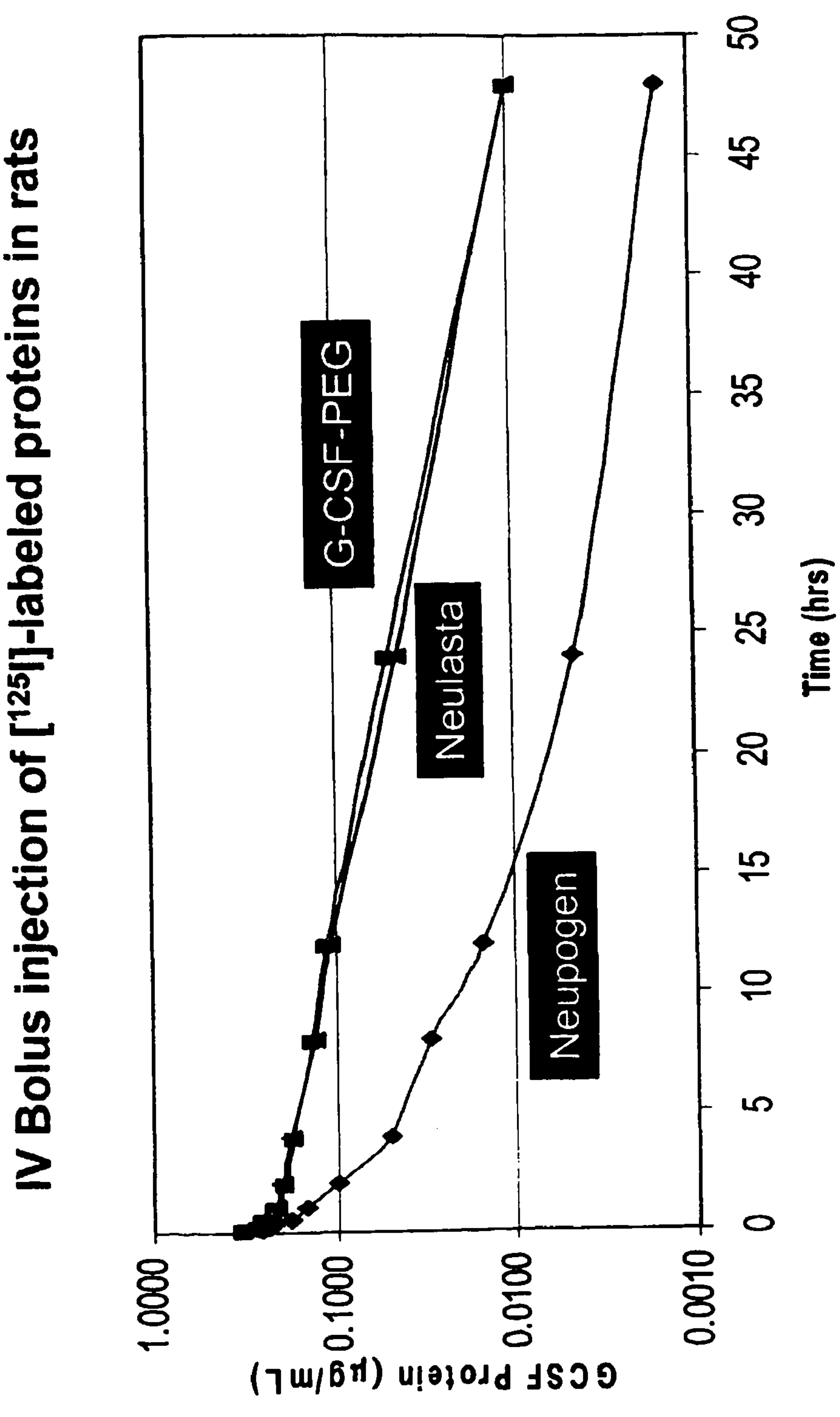

IV Bolus injection of [^{125}I]-labeled proteins in rats
G-CSF-PEG
Neulasta
Neupogen
1.0000
0.1000
0.0100
0.0010
GCSF Protein (µg/mL)
0
5
10
15
20
25
30
35
40
45
50
Time (hrs)

FIGURE 4

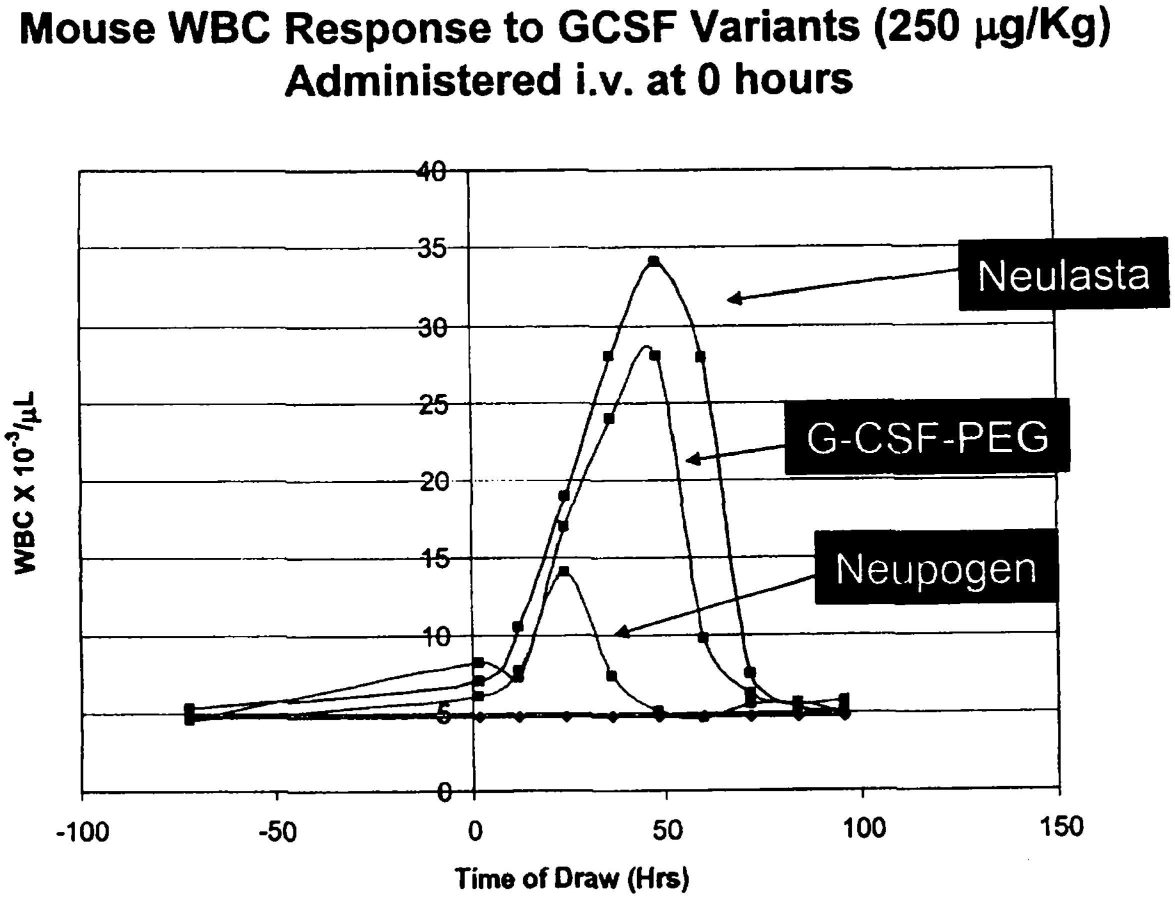

Mouse WBC Response to GCSF Variants (250 μg/Kg)
Administered i.v. at 0 hours
WBC X 10⁻³/μL
40
35
30
25
20
15
10
5
0
-100
-50
0
50
100
150
Time of Draw (Hrs)
Neulasta
G-CSF-PEG
Neupogen

FIGURE 5

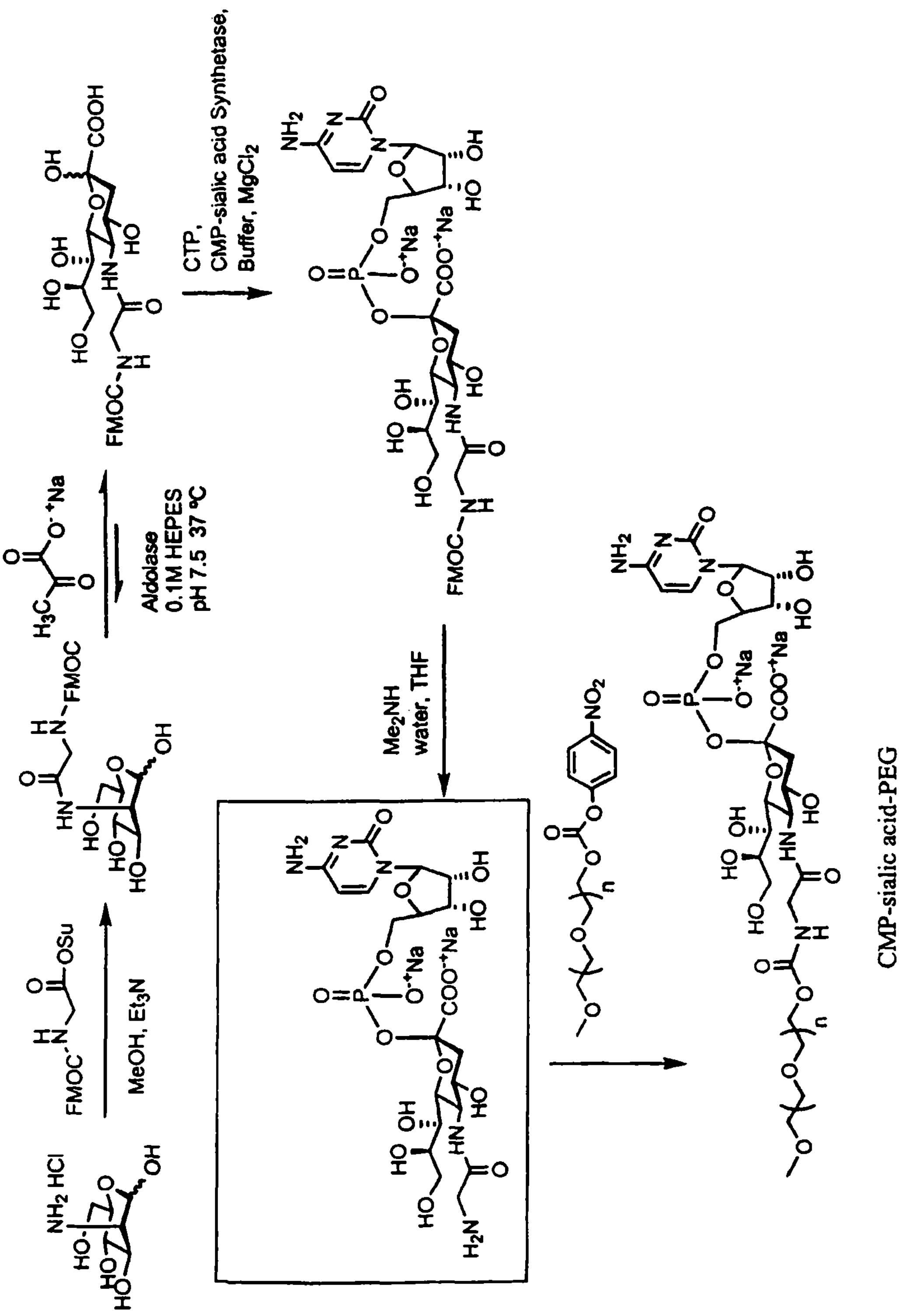

NH2 HCl
FMOC-N(H)-CH2-C(O)-OSu
MeOH, Et3N
FMOC
H3C-C(O)-C(O)-O⁻ +Na
Aldolase
0.1M HEPES
pH 7.5 37 °C
FMOC-N(H)
COOH
CTP,
CMP-sialic acid Synthetase,
Buffer, MgCl2
NH2
COO⁻+Na
O⁻+Na
Me2NH
water, THF
H2N
NO2
n
CMP-sialic acid-PEG

FIGURE 6

175 amino acid variant

MTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCA
TYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLAGCLS
QLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFAT
TIWQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVL
VASHLQSFLEVSYRVLRHLAQP **(SEQ ID NO: 1)**.

174 amino acid variant

TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATY
KLCHPEELVLLGHSLGIPWAPLSSCPSQALQLAGCLSQL
HSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTI
WQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLV
ASHLQSFLEVSYRVLRHLAQP **(SEQ ID NO: 2)**.

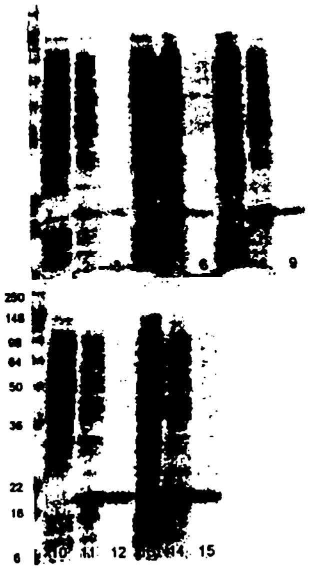
FIGURE 7

Testing growth conditions – washing IBs

37°C, 1mM (final) IPTG				IB yields (~0.7-1g/L)
	Media	Buffered pH	Time	wet IB pellet (mg)
1.	LB	n/a	5hr	120
2.	TB	7.4 (K-PO$_4$)	5hr	230
3.	TB	6 (MES)	5hr	170
4.	LB	n/a	o/n	190
5.	TB	7.4 (K-PO$_4$)	o/n	250

Cell samples, loaded equally by ug
(better expression of overnights more apparent)

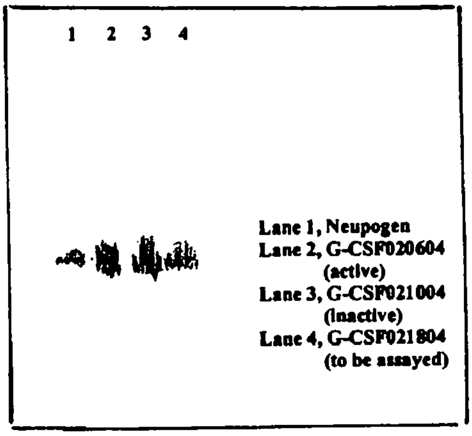
1. #1 lysate, supe
2. #1 lysate, pellet
3. #1 final IB
4. #2 lysate, supe
5. #2 lysate, pellet
6. #2 final IB
7. #3 lysate, supe
8. #3 lysate, pellet
9. #3 final IB
10. #4 lysate, supe
11. #4 lysate, pellet
12. #4 final IB
13. #5 lysate, supe
14. #5 lysate, pellet
15. #5 final IB



Purified IB's are shown in lanes 3, 6, 9, 12 and 15. Note IB lanes loaded at approximately ½ the amount of the preceding lane (based on predicted pellet weight).

FIGURE 8

Western Blot Analysis of Refolded G-CSF Native Polyacrylamide Gel Electrophoresis



Lane 1, Neupogen
Lane 2, G-CSF020604 (active)
Lane 3, G-CSF021004 (inactive)
Lane 4, G-CSF021804 (to be assayed)

G-CSF020604 and G-CSF021004 came from the same batch of refolding reaction. The only difference was that G-CSF021004 was saved from the flowthough of G-CSF020604 by adjusting pH and reloading onto SP Sepharose. G-CSF021804 was a separate refolding batch.

After SP-sepharose you see only one band by Coomassie stain or Western.

FIGURE 9

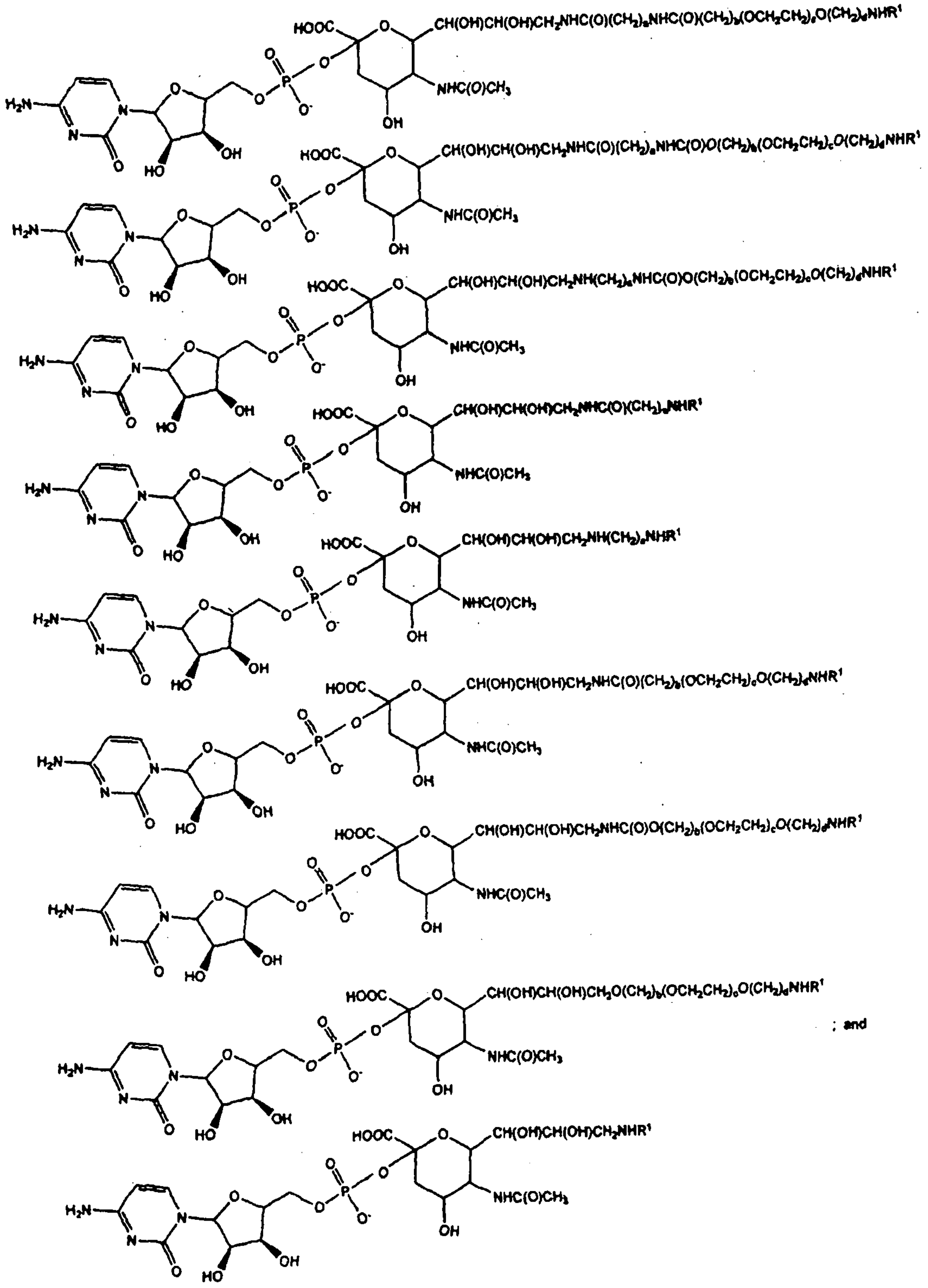
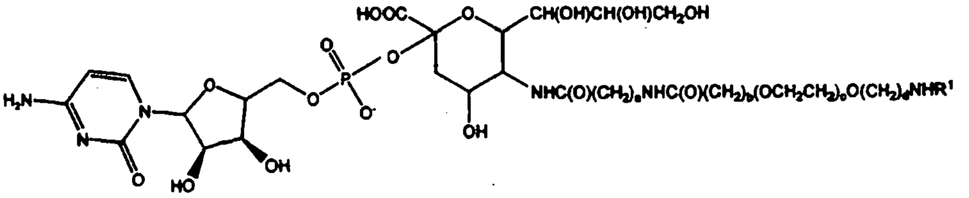
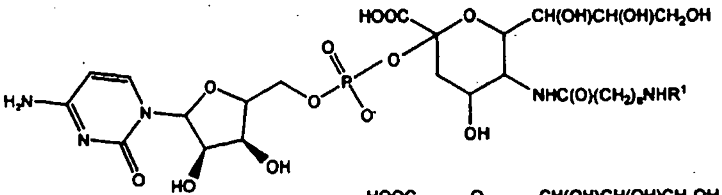

HOOC
CH(OH)CH(OH)CH2NHC(O)(CH2)aNHC(O)(CH2)b(OCH2CH2)cO(CH2)dNHR1
NHC(O)CH3
OH
H2N
N
O
HO
OH
P
O-
CH(OH)CH(OH)CH2NHC(O)(CH2)aNHC(O)O(CH2)b(OCH2CH2)cO(CH2)dNHR1
CH(OH)CH(OH)CH2NH(CH2)aNHC(O)O(CH2)b(OCH2CH2)cO(CH2)dNHR1
CH(OH)CH(OH)CH2NHC(O)(CH2)aNHR1
CH(OH)CH(OH)CH2NH(CH2)aNHR1
CH(OH)CH(OH)CH2NHC(O)(CH2)b(OCH2CH2)cO(CH2)dNHR1
CH(OH)CH(OH)CH2NHC(O)O(CH2)b(OCH2CH2)cO(CH2)dNHR1
CH(OH)CH(OH)CH2O(CH2)b(OCH2CH2)cO(CH2)dNHR1
; and
CH(OH)CH(OH)CH2NHR1

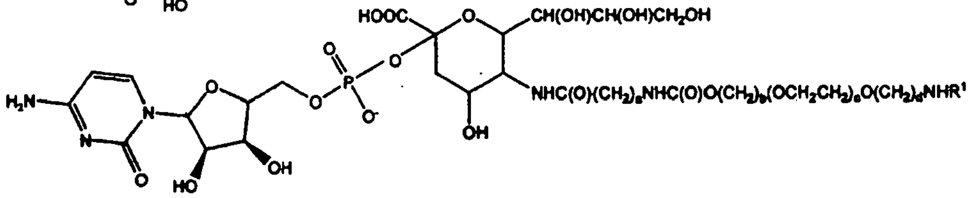
FIGURE 10



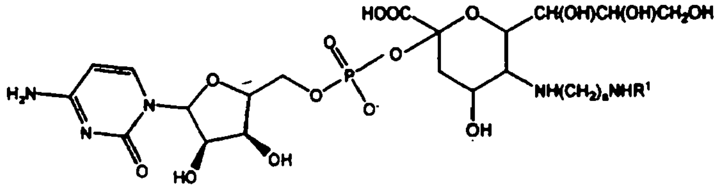
;



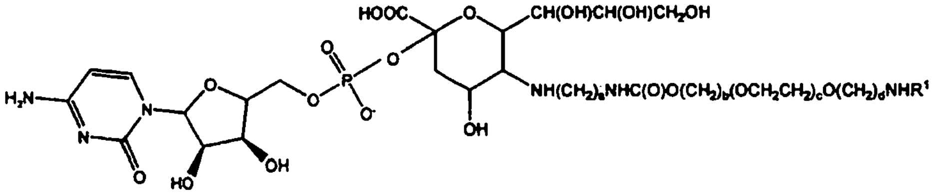
;



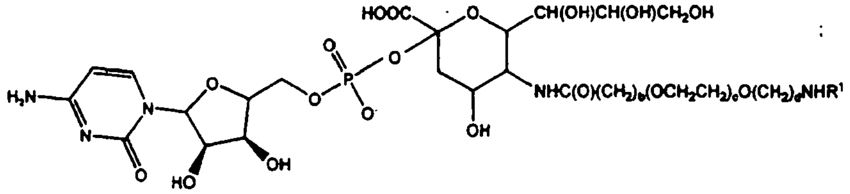
;



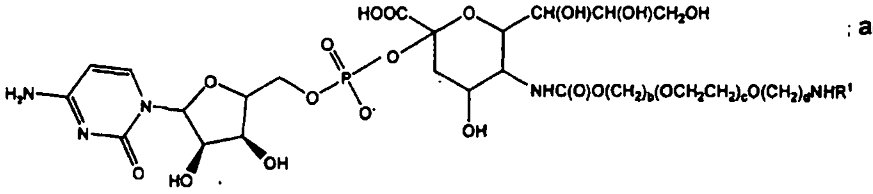
;



;



;



; and

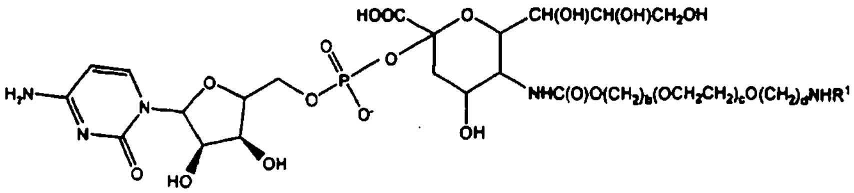


FIGURE 11A

Protein	Organism	EC#	GenBank / GenPept	SwissProt	PDB / 3D
At1g08280	*Arabidopsis thaliana*	n.d.	AC011438 [illegible] BT004583 [illegible] NC_003070 [illegible]	[illegible]	
At1g08660/F22O13.14	*Arabidopsis thaliana*	n.d.	AC003981 AAF99778.1 AY064135 AAL36042.1 AY124807 AAM70516.1 NC_003070 NP_172342.1 NM_180609 NP_850940.1	Q8VZJ0 Q9FRR9	
At3g48820/T21J18_90	*Arabidopsis thaliana*	n.d.	AY080589 [illegible] AY133816 [illegible] AL132963 CAB87[illegible] NM_114741 NP_[illegible]	Q8R[illegible]00 Q9M301	
α-2,3-sialyltransferase (ST3GAL-IV)	*Bos taurus*	n.d.	AJ584673 CAE48298.1		
α-2,3-sialyltransferase (St3Gal-V)	*Bos taurus*	n.d.	AJ585768 CAE51392.1		
α-2,6-sialyltransferase (Siat7b)	*Bos taurus*	n.d.	AJ620651 CAF05850.1		
α-2,8-sialyltransferase (SIAT8A)	*Bos taurus*	2.4.99.8	AJ699418 CAG27880.1		
α-2,8-sialyltransferase (Siat8D)	*Bos taurus*	n.d.	AJ699421 CAG27883.1		
α-2,8-sialyltransferase ST8Siα-III (Siat8C)	*Bos taurus*	n.d.	AJ704563 CAG28696.1		
CMP α-2,6-sialyltransferase (ST6Gal I)	*Bos taurus*	2.4.99.1	Y15111 CAA75385.1 NM_177517 NP_803483.1	O18974	
sialyltransferase 8 (fragment)	*Bos taurus*	n.d.	AF450088 AAL47018.1	Q8WN13	
sialyltransferase ST3Gal-II (Siat4B)	*Bos taurus*	n.d.	AJ748841 CAG44450.1		
sialyltransferase ST3Gal-III (Siat6)	*Bos taurus*	n.d.	AJ748842 CAG44451.1		
sialyltransferase ST3Gal-VI (Siat10)	*Bos taurus*	n.d.	AJ748843 CAG44452.1		
ST3Gal[illegible]	*Bos taurus*	n.d.	AJ305086 CAC24698.1	Q9BEG[illegible]	
St6GalNAc-VI	*Bos taurus*	n.d.	AJ620949 CAF06586.1		
CDS4	*Branchiostoma floridae*	n.d.	AF391289 AAM18873.1	Q8T77[illegible]	
polysialyltransferase (PST) (fragment) ST8Sia IV	*Cercopithecus aethiops*	2.4.99.-	AF210729 AAF17105.1	Q9TT09	
polysialyltransferase (STX) (fragment) ST8Sia II	*Cercopithecus aethiops*	2.4.99.-	AF210318 AAF17104.1	Q9TT10	
α-2,3-sialyltransferase ST3Gal I (Siat4)	*Ciona intestinalis*	n.d.	AJ626815 CAF25173.1		
α-2,3-sialyltransferase ST3Gal I (Siat4)	*Ciona savignyi*	n.d.	AJ626814 CAF25172.1		
α-2,8-polysialyltransferase ST8Sia IV	*Cricetulus griseus*	2.4.99.-	- AAE28634 Z46801 CAA86822.1	Q64690	
Gal β-1,3/4-GlcNAc α-2,3-sialyltransferase St3Gal I	*Cricetulus griseus*	n.d.	AY266675 AAP22942.1	Q80W[illegible]	
Gal β-1,3/4-GlcNAc α-2,3-sialyltransferase St3Gal II (fragment)	*Cricetulus griseus*	n.d.	AY266676 AAP22943.1	Q80WK9	
α-2,3-sialyltransferase ST3Gal I (Siat4)	*Danio rerio*	n.d.	AJ783740 CAH04017.1		
α-2,3-sialyltransferase ST3Gal II (Siat5)	*Danio rerio*	n.d.	AJ783741 CAH04018.1		
α-2,3-sialyltransferase ST3Gal III (Siat6)	*Danio rerio*	n.d.	AJ626821 CAF25179.1		
α-2,3-sialyltransferase ST3Gal IV (Siat4c)	*Danio rerio*	n.d.	AJ744809 CAG32845.1		

FIGURE 11B

Protein	Organism	EC#	GenBank / GenPept	SwissProt	PDB / 3D
α-2,3-sialyltransferase ST3Gal V-r (Siat5-related)	*Danio rerio*	n.d.	AJ783742 [illegible]		
α-2,6-sialyltransferase ST6Gal I (Siat1)	*Danio rerio*	n.d.	AJ744801 CAG32837.1		
α-2,6-sialyltransferase ST6GalNAc II (Siat7B)	*Danio rerio*	n.d.	AJ634459 CAG25680.1		
α-2,6-sialyltransferase ST6GalNAc V (Siat7E) (fragment)	*Danio rerio*	n.d.	AJ646874 CAG26703.1		
α-2,6-sialyltransferase ST6GalNAc VI (Siat7F) (fragment)	*Danio rerio*	n.d.	AJ646883 [illegible]		
α-2,8-sialyltransferase ST8Sia I (Siat 8A) (fragment)	*Danio rerio*	n.d.	AJ715535 CAG29374.1		
α-2,8-sialyltransferase ST8Sia III (Siat 8C) (fragment)	*Danio rerio*	n.d.	AJ715543 CAG29382.1		
α-2,8-sialyltransferase ST8Sia IV (Siat 8D) (fragment)	*Danio rerio*	n.d.	AJ715545 CAG29384.1		
α-2,8-sialyltransferase ST8Sia V (Siat 8E) (fragment)	*Danio rerio*	n.d.	AJ715546 CAG29385.1		
α-2,8-sialyltransferase ST8Sia VI (Siat 8F) (fragment)	*Danio rerio*	n.d.	AJ715551 CAG29390.1		
β-galactosamide α-2,6-sialyltransferase II (ST6Gal II)	*Danio rerio*	n.d.	AJ627627 CAF29495.1		
N-glycan α-2,8-sialyltransferase	*Danio rerio*	n.d.	BC050483 AAH50483.1 AY055462 AAL17875.1 NM_153662 NP_705948.1	Q7ZU51 Q8QH83	
ST3Gal III-related (siat6r)	*Danio rerio*	n.d.	BC053179 AAH53179.1 AJ626820 CAF25178.1 NM_200355 NP_956649.1	Q7T3B9	
St3Gal-V	*Danio rerio*	n.d.	AJ619960 CAF04061.1		
st6GalNAc-VI	*Danio rerio*	n.d.	BC060932 AAH60932.1 AJ620947 CAF06584.1		
α-2,6-sialyltransferase (CG4871) ST6Gal I	*Drosophila melanogaster*	2.4.99.1	AE003465 AAF47256.1 AF218237 AAG13185.1 AF397532 AAK92126.1 AE003465 AAM70791.1 NM_079129 NP_523853.1 NM_166684 NP_726474.1	Q9GU23 Q9W121	
α-2,3-sialyltransferase (ST3Gal-VI)	*Gallus gallus*	n.d.	AJ585767 CAE51391.1 AJ627204 CAF25503.1		
α-2,3-sialyltransferase ST3Gal I	*Gallus gallus*	2.4.99.4	X80503 CAA56666.1 NM_205217 NP_990548.1	Q11200	
α-2,3-sialyltransferase ST3Gal IV (fragment)	*Gallus gallus*	2.4.99.-	AF035250 AAC14163.1	O73724	
α-2,3-sialytransferase (ST3GAL-II)	*Gallus gallus*	n.d.	AJ585761 CAE51385.2		
α-2,6-sialyltransferase (Siat7b)	*Gallus gallus*	n.d.	AJ620653 CAF05852.1		
α-2,6-sialyltransferase ST6Gal I	*Gallus gallus*	2.4.99.1	X75558 CAA53235.1 NM_205241 NP_990572.1	Q92182	
α-2,6-sialyltransferase ST6GalNAc I	*Gallus gallus*	2.4.99.3	AAE68028.1 AAE68029.1 X74946 CAA52902.1 NM_205240 [illegible]	Q92183	
α-2,6-sialyltransferase ST6GalNAc II	*Gallus gallus*	2.4.99.-	X77775 AAE68030.1 NM_205233 CAA54813.1 NP_990564.1	Q92184	
α-2,6-sialyltransferase ST6GalNAc III (SIAT7C)	*Gallus gallus*	n.d.	AJ634455 CAG25677.1		

FIGURE 11C

Protein	Organism	EC#	GenBank / GenPept		SwissProt	PDB / 3D
(fragment)						
α-2,6-sialyltransferase ST6GalNAc V (SIAT7E) (fragment)	*Gallus gallus*	n.d.	AJ646877	CAG26706.1		
α-2,8-sialyltransferase (GD3 Synthase) ST8Sia I	*Gallus gallus*	2.4.99.-	U73176	AAC[illegible]	P79783	
α-2,8-sialyltransferase (SIAT8B)	*Gallus gallus*	n.d.	AJ699419	CAG27881.1		
α-2,8-sialyltransferase (SIAT8C)	*Gallus gallus*	n.d.	AJ699420	CAG27882.1		
α-2,8-sialyltransferase (SIAT8F)	*Gallus gallus*	n.d.	AJ699424	CAG27886.1		
α-2,8-syalyltransferase ST8Siα-V (SIAT8C)	*Gallus gallus*	n.d.	AJ704564	CAG28697.1		
β-galactosamide α-2,6-sialyltransferase II (ST6Gal II)	*Gallus gallus*	n.d.	AJ627629	CAF29497.1		
GM3 synthase (SIAT9)	*Gallus gallus*	2.4.99.9	AY515255	AAS83519.1		
polysialyltransferase ST8Sia IV	*Gallus gallus*	2.4.99.-	AF008194	AAB95120.1	O42399	
α-2,3-sialyltransferase ST3Gal I	*Homo sapiens*	2.4.99.4	L29555	AAA36612.1	Q11201	
			AF059321	AAC17874.1	O60677	
			L13972	AAC37574.1	Q9UN51	
			AF155238	AAD39238.1		
			AF186191	AAG29876.1		
			BC018357	AAH18357.1		
			NM_003033	NP_003024.1		
			NM_173344	NP_775479.1		
α-2,3-sialyltransferase ST3Gal II	*Homo sapiens*	2.4.99.4	U63090	AAB40389.1	Q16842	
			BC036777	AAH36777.1	O00654	
			X96667	CAA65447.1		
			NM_006927	NP_008858.1		
α-2,3-sialyltransferase ST3Gal III (SiaT6)	*Homo sapiens*	2.4.99.6	L23768	AAA35778.1	Q11203	
			BC050380	AAH50380.1	Q86UR6	
			AF425851	AAO13859.1	Q86UR7	
			AF425852	AAO13860.1	Q86UR8	
			AF425853	AAO13861.1	Q86UR9	
			AF425854	AAO13862.1	Q86US0	
			AF425855	AAO13863.1	Q86US1	
			AF425856	AAO13864.1	Q86US2	
			AF425857	AAO13865.1	Q8IX43	
			AF425858	AAO13866.1	Q8IX44	
			AF425859	AAO13867.1	Q8IX45	
			AF425860	AAO13868.1	Q8IX46	
			AF425861	AAO13869.1	Q8IX47	
			AF425862	AAO13870.1	Q8IX48	
			AF425863	AAO13871.1	Q8IX49	
			AF425864	AAO13872.1	Q8IX50	
			AF425865	AAO13873.1	Q8IX51	
			AF425866	AAO13874.1	Q8IX52	
			AF425867	AAO13875.1	Q8IX53	
			AY167992	AAO38806.1	Q8IX54	
			AY167993	AAO38807.1	Q8IX55	
			AY167994	AAO38808.1	Q8IX56	
			AY167995	AAO38809.1	Q8IX57	
			AY167996	AAO38810.1	Q8IX58	
			AY167997	AAO38811.1		
			AY167998	AAO38812.1		
			NM_006279	NP_006270.1		
			NM_174964	NP_777624.1		
			NM_174965	NP_777625.1		
			NM_174966	NP_777626.1		
			NM_174967	NP_777627.1		
			NM_174969	NP_777629.1		
			NM_174970	NP_777630.1		
			NM_174972	NP_777632.1		

FIGURE 11D

Protein	Organism	EC#	GenBank / GenPept	SwissProt	PDB / 3D
α-2,3-sialyltransferase ST3Gal IV	*Homo sapiens*	2.4.99.-	L23767 AAA16460.1 AF035249 AAC14162.1 BC010645 AAH10645.1 AY040826 AAK93790.1 AF516602 AAM66431.1 AF516603 AAM66432.1 AF516604 AAM66433.1 AF525084 AAM81378.1 X74570 CAA52662.1 CR456858 CAG33139.1 NM_006278 NP_006269.1	Q11206 O60497 Q96QQ9 Q8N6A6 Q8N6A7 Q8NFD3 Q8NFG7	
α-2,3-sialyltransferase ST3Gal VI	*Homo sapiens*	2.4.99.4	AF119391 AAD39131.1 BC023312 AAH23312.1 AB022918 BAA77609.1 AX877828 CAE89895.1 AX886023 CAF00161.1 NM_006100 NP_006091.1	Q9Y274 [illegible]	
α-2,6-sialyltransferase (ST6Gal II ; KIAA1877)	*Homo sapiens*	n.d.	BC008680 AAH08680.1 AB058780 BAB47506.1 AB059555 BAC24793.1 AJ512141 CAD54408.1 AX795193 CAE48260.1 AX795193 CAE48261.1 NM_032528 NP_115917.1	Q86Y44 Q8IUG7 Q96HE4 Q96JF0	
α-2,6-sialyltransferase (ST6GALNAC III)	*Homo sapiens*	n.d.	BC059363 AAH59363.1 AY358540 AAQ88904.1 AK091215 BAC03611.1 AJ507291 CAD45371.1 NM_152996 NP_694541.1	Q8N259 Q8NDV1	
α-2,6-sialyltransferase (ST6GalNAc V)	*Homo sapiens*	n.d.	BC001201 AAH01201.1 AK056241 BAB71127.1 AL035409 CAB72344.1 AJ507292 CAD45372.1 NM_030965 NP_112227.1	Q9BVH7	
α-2,6-sialyltransferase (SThM) ST6GalNAc II	*Homo sapiens*	2.4.99.-	U14550 AAA52228.1 BC040455 AAH40455.1 AJ251053 CAB61434.1 NM_006456 NP_006447.1	Q9UJ37 Q12971	
α-2,6-sialyltransferase ST6Gal I	*Homo sapiens*	2.4.99.1	BC031476 AAH31476.1 BC040009 AAH40009.1 A17362 CAA01327.1 A23699 CAA01686.1 X17247 CAA35111.1 X54363 CAA38246.1 X62822 CAA44634.1 NM_003032 NP_003023.1 NM_173216 NP_775323.1	P15907	
α-2,6-sialyltransferase ST6GalNAc I	*Homo sapiens*	2.4.99.3	BC022462 AAH22462.1 AY096001 AAM22800.1 AY358918 AAQ89277.1 AK000113 BAA90953.1 Y11339 CAA72179.2 NM_018414 NP_060884.1	Q8TBJ6 Q9NSC7 Q9NXQ7	
α-2,8-polysialyltransferase ST8Sia IV	*Homo sapiens*	2.4.99.-	L41680 AAC41775.1 BC027866 AAH27866.1 BC053657 AAH53657.1 NM_005668 NP_005659.1	Q8N1F4 Q92187 Q92693	
α-2,8-sialyltransferase (GD3 synthase) ST8Sia I	*Homo sapiens*	2.4.99.8	L32867 AAA62366.1 L43494 AAC37586.1 BC046158 AAH46158.1 [illegible] AAQ53140.1 AY569975 AAS75783.1 D26360 BAA05391.1 X77922 CAA54891.1 NM_003034 NP_003025.1	Q86X71 Q92185 Q93064 [illegible]	
α-2,8-sialyltransferase	*Homo sapiens*	2.4.99.-	L29556 AAA36613.1	Q92186	

FIGURE 11E

Protein	Organism	EC#	GenBank / GenPept	SwissProt	PDB / 3D
ST8Sia II			U82762 AAB51242.1 U33551 **AAC24458.1** BC069584 AAH69584.1 NM_006011 NP_006002.1	Q92470 Q92746	
α-2,8-sialyltransferase ST8Sia III	*Homo sapiens*	2.4.99.-	AF004668 [illegible] AF003092 [illegible] NM_015879 [illegible]	[illegible]	
α-2,8-sialyltransferase ST8Sia V	*Homo sapiens*	2.4.99.-	U91641 **AAC51727.1** CR457037 CAG33318.1 NM_013305 NP_037437.1	O15466	
ENSR00000020221 (fragment)		n.d.	AC023295 - [illegible]	[illegible]	
lactosylceramide α-2,3-sialyltransferase (ST3Gal V)	*Homo sapiens*	2.4.99.9	AF105026 **AAD14634.1** AF119415 AAF66146.1 BC065936 AAH65936.1 AY152815 AAO16866.1 AAP65066 AAP65066.1 AY359105 AAQ89463.1 AB018356 BAA33950.1 AX876536 CAE89320.1 NM_003896 NP_003887.2	Q9UNP4 O94902	
N-acetylgalactosaminide α-2,6-sialyltransferase (ST6GalNAc VI)	*Homo sapiens*	2.4.99.-	BC006564 AAH06564.1 BC007802 AAH07802.1 BC016299 AAH16299.1 AY358672 AAQ89035.1 AB035173 BAA87035.1 AK023900 [illegible] [illegible] [illegible] [illegible] [illegible] NP_038471.2	Q969X2 Q9H8A2 Q9ULB8	
N-acetylgalactosaminide α-2,6-sialyltransferase IV (ST6GalNAc IV)	*Homo sapiens*	2.4.99.-	AF127142 AAF00102.1 BC036705 AAH36705.1 - AAP63349.1 AB035172 **BAA87034.1** AK000600 BAA91281.1 Y17461 CAB44354.1 AJ271734 CAC07404.1 AX061620 CAC24981.1 AX068265 CAC27250.1 AX969252 CAF14360.1 NM_014403 NP_055218.3 NM_175039 NP_778204.1	Q9H4F1 Q9NWU6 Q9UKU1 Q9ULB9 Q9Y3G3 Q9Y3G4	
ST8SIA-VI (fragment)	*Homo sapiens*	n.d.	AJ621583 CAF21722.1 XM_291725 XP_291725.2	[illegible]	
unnamed protein product	*Homo sapiens*	n.d.	AK021929 **BAB13940.1** AX881696 CAE91353.1	Q9HAA9	
Gal β-1,3/4-GlcNAc α-2,3-sialyltransferase (ST3Gal III)	*Mesocricetus auratus*	2.4.99.6	AJ245699 **CAB53394.1**	Q9QXF6	

FIGURE 11F

Protein		Organism	EC#	GenBank / GenPept		SwissProt	PDB / 3D
Gal β-1,3/4-GlcNAc α-2,3-sialyltransferase (ST3Gal IV)		*Mesocricetus auratus*	2.4.99.6	AJ245700	CAB53395.1	Q9QXF5	
GD3 synthase (fragment) ST8Sia I		*Mesocricetus auratus*	n.d.	AF141657	AAD3387[illegible]	Q9WUL1	
				[illegible]	[illegible]	[illegible]	
polysialyltransferase (ST8Sia IV)		*Mesocricetus auratus*	2.4.99.-	AJ245701	CAB53396.1	Q9QXF4	
α-2,3-sialyltransferase ST3Gal I	*St3gal1*	*Mus musculus*	2.4.99.4	AF214028	AAF6097[illegible]	[illegible]751	
				AK031344	BAC27356.1	Q11202	
				AK078469	BAC37290.1	Q9JL30	
				X73523	CAA5191[illegible]		
				NM_009177	NP_03320[illegible]		
α-2,3-sialyltransferase ST3Gal II	*St3gal2*	*Mus musculus*	2.4.99.4	BC015264	AAH15264.1	Q11204	
				BC066064	AAH66064.1	Q8BPL0	
				AK034554	BAC28752.1	Q8BSA0	
				AK034863	BAC28859.1	Q8BSE9	
				AK053827	BAC35543.1	Q91WH6	
				X76989	CAA54294.1		
				NM_009179	NP_033205.1		
				NM_178048	NP_835149.1		
α-2,3-sialyltransferase ST3Gal III	*St3gal3*	*Mus musculus*	2.4.99.-	BC006710	AAH06710.1	P97325	
				AK005053	BAB23779.1	Q922X5	
				AK013016	BAB28598.1	Q9CZ48	
				X84234	CAA59013.1	Q9DBB6	
				NM_009176	NP_033202.2		
α-2,3-sialyltransferase ST3Gal IV	*St3gal4*	*Mus musculus*	2.4.99.4	BC011121	AAH11121.1	P97354	
				BC050773	AAH50773.1	Q61325	
				D28941	BAA06068.1	Q91Y74	
				AK008543	BAB25732.1	Q921R5	
				AB061305	BAB47508.1	Q9CVE8	
				X95809	CAA65076.1		
				NM_009178	NP_033204.2		
α-2,3-sialyltransferase ST3Gal VI	*St3gal6*	*Mus musculus*	2.4.99.4	AF119390	AAD39130.1	Q80UR7	
				BC052338	AAH52338.1	Q8BLV1	
				AB063326	BAB79494.1	Q8VIB3	
				AK033562	BAC28360.1	Q9WVG2	
				AK041173	BAC30851.1		
				NM_018784	NP_061254		
α-2,6-sialyltransferase ST6GalNAc II	*St6galnac2*	*Mus musculus*	2.4.99.-	NM_009180	6677963	P70277	
				BC010208	AAH10208.1	Q9DC24	
				AB027198	BAB00637.1	Q9JJM5	
				AK004613	BAB23410.1		
				X93999	CAA63821.1		
				X94000	CAA63822.1		
				NM_009180	NP_033206.2		
α-2,6-sialyltransferase ST6Gal I	*St6gal1*	*Mus musculus*	2.4.99.1	[illegible]	[illegible]68031.1	Q64685	
				BC027833	AAH27833.1	Q8BM62	
				D16106	BAA03680.1	Q8K1L1	
				AK034[illegible]	BAC28828.1		
				AK084124	BAC39120.1		
				NM_145933	NP_666045.1		
α-2,6-sialyltransferase ST6Gal II	*St6gal2*	*Mus musculus*	n.d.	AK082566	BAC38534.1	Q8BUU4	
				AB095093	BAC87752.1		
				AK129462	BAC98272.1		
				NM_172829	NP_766417.1		
α-2,6-sialyltransferase ST6GalNAc I	*St6galnac1*	*Mus musculus*	2.4.99.3	Y11274	CAA72137.1	Q9QZ39	
				NM_011371	NP_035501.1	Q9JJP5	
α-2,6-sialyltransferase ST6GalNAc III	*St6galnac3*	*Mus musculus*	n.d.	BC058387	AAH58387.1	Q9WUV2	
				AK034804	BAC28836.1	Q9JHP5	
				Y11342	CAA72181.2		
				Y11343	CAB95031.1		
				NM_011372	NP_035502		
α-2,6-sialyltransferase ST6GalNAc IV	*St6galnac4*	*Mus musculus*	2.4.99.7	BC056451	AAH56451.1	Q8C3J2	
				AK085730	BAC39523.1	Q9JHP2	
				AJ007310	CAA07446.1	Q9R2B6	
				Y15779	CAB43507.1	O88725	

FIGURE 11G

Protein		Organism	EC#	GenBank / GenPept		SwissProt	PDB / 3D
[illegible]				Y15780	CA[illegible]	[illegible]HP0	
				Y19055	CA[illegible]	[illegible]QUP9	
				Y19057	CA[illegible]	Q9R2B5	
				NM_011373	NP_[illegible]		
α-2,8-sialyltransferase (GD3 synthase) ST8Sia I	*St8sia1*	*Mus musculus*	2.4.99.8	L38677	AAA91869.1	Q64468	
				BC024821	AAH24821.1	Q64687	
				AK046188	BAC32625.1	Q8BL76	
				AK052444	BAC34994.1	Q8BWI0	
				X84235	CAA59014.1	Q8K1C1	
				AJ401102	CAC20706.1	Q9EPK0	
				NM_011374	NP_035504.1		
α-2,8-sialyltransferase (ST8Sia VI)	*St8sia6*	*Mus musculus*	n.d.	AB059554	BAC01265.1	Q8BI43	
				AK085105	BAC3[illegible]	Q8K4T1	
				NM_145838	NP_665837.1		
α-2,8-sialyltransferase ST8Sia II	*St8sia2*	*Mus musculus*	2.4.99.-	X83562	CAA58548.1	O35696	
				X99646	CAA67965.1		
				X99647	CAA67965.1		
				X99648	CAA67965.1		
				X99649	CAA67965.1		
				X99650	CAA67965.1		
				X99651	CAA67965.1		
				NM_009181	NP_033207.1		
α-2,8-sialyltransferase ST8Sia IV	*St8sia4*	*Mus musculus*	2.4.99.8	BC060112	AAH60112.1	Q64692	
				AK003690	BAB22941.1	Q8BY70	
				AK041723	BAC31044.1		
				AJ223956	CAA11685.1		
				X86000	CAA59992.1		
				Y09484	CAA70692.1		
				NM_009183	NP_033209.1		
α-2,8-sialyltransferase ST8Sia V	*St8sia5*	*Mus musculus*	2.4.99.-	BC034855	AAH34855.1	P70126	
				AK078670	BAC37354.1	P70127	
				X98014	CAA66642.1	P70128	
				X98014	CAA66643.1	Q8BJW0	
				X98014	CAA66644.1	Q8JZQ3	
				NM_013666	NP_038694.1		
				NM_153124	NP_694764.1		
				NM_177416	NP_803135.1		
α-2,8-sialytransferase ST8Sia III	*St8sia3*	*Mus musculus*	2.4.99.-	BC075645	AAH75645.1	Q64689	
				AK015874	BAB30012.1	Q9CUJ6	
				X80502	CAA56665.1		
				NM_009182	NP_033208.1		
GD1 synthase (ST6GalNAc V)	*St6galnac5*	*Mus musculus*	n.d.	BC055737	AAH55737.1	Q8CAM7	
				AB030836	BAA85747.1	Q8CBX1	
				AB028840	BAA89292.1	Q9QYJ1	
				AK034387	BAC28693.1	Q9R0K6	
				AK038434	BAC29997.1		
				AK042683	BAC31331.1		
				NM_012028	NP_036158.2		
GM3 synthase (α-2,3-sialyltransferase) ST3Gal V	*St3gal5*	*Mus musculus*	2.4.99.9	AF119416	AAF66147.1	O88829	
					AAP65063.1	Q9CZ65	
				AB018048	BAA33491.1	Q9QWF9	
				AB013302	BAA76467.1		
				AK012961	BAB28571.1		
				Y15003	CAA75236.1		
				NM_011375	NP_035505.1		
N-acetylgalactosaminide α-2,6-sialyltransferase (ST6GalNAc VI)	*St6galnac6*	*Mus musculus*	2.4.99.-	BC036985	AAH36985.1	Q8CDC3	
				AB035174	BAA87036.1	Q8JZW3	
				AB035123	BAA95940.1	Q9JM95	
				AK030648	BAC27064.1	Q9R0G9	
				NM_016973	NP_058669.1		
M138L		*Myxoma virus*	n.d.	U46578	AAD00069.1		
				AF170726	AAE61323.1		
				NC_001132	AAE61326.1		
					AAF15026.1		
					NP_051852.1		
α-2,3-sialyltransferase		*Oncorhynchus*	n.d.	AJ585760	CAE51384.1		

FIGURE 11H

Protein	Organism	EC#	GenBank / GenPept	SwissProt	PDB / 3D
(St3Gal-I)	*mykiss*				
α-2,6-sialyltransferase (Siat1)	*Oncorhynchus mykiss*	n.d.	AJ620649 CAF05[illegible]		
α-2,8-polysialyltransferase IV (ST8Sia IV)	*Oncorhynchus mykiss*	n.d.	AB094402 BAC77411.1	Q7T2X5	
GalNAc α-2,6-sialyltransferase (RtST6GalNAc)	*Oncorhynchus mykiss*	n.d.	AB097943 BAC77520[illegible]	Q7T2X4	
α-2,3-sialyltransferase ST3Gal IV	*Oryctolagus cuniculus*	2.4.99.-	AF121967 AAF28871.1	Q9N257	
OJ1217_F02.7	*Oryza sativa (japonica cultivar-group)*	n.d.	AP004084 BAD07616.1		
OSJNBa0043L24.2 or OSJNBb0002J11.9	*Oryza sativa (japonica cultivar-group)*	n.d.	AL731626 CAD41185.1 AL662969 CAE04714.1		
P0683f02.18 or P0489B03.1	*Oryza sativa (japonica cultivar-group)*	n.d.	AP003289 BAB63715.1 AP003794 BAB90552.1		
α-2,6-sialyltransferase ST6GalNAc V (Siat7E) (fragment)	*Oryzias latipes*	n.d.	AJ646876 CAG26705.1		
α-2,3-sialyltransferase ST3Gal I (Siat4)	*Pan troglodytes*	n.d.	AJ744803 CAG32839.1		
α-2,3-sialyltransferase ST3Gal II (Siat5)	*Pan troglodytes*	n.d.	AJ744804 CAG32840.1		
α-2,3-sialyltransferase ST3Gal III (Siat6)	*Pan troglodytes*	n.d.	AJ626819 CAF25177.1		
α-2,3-sialyltransferase ST3Gal IV (Siat4c)	*Pan troglodytes*	n.d.	AJ626824 CAF25182.1		
α-2,3-sialyltransferase ST3Gal VI (Siat10)	*Pan troglodytes*	n.d.	AJ744808 CAG32844.1		
α-2,6-sialyltransferase (Sia7A)	*Pan troglodytes*	n.d.	AJ748740 CAG38615.1		
α-2,6-sialyltransferase (Sia7B)	*Pan troglodytes*	n.d.	AJ748741 CAG38616.1		
α-2,6-sialyltransferase ST6GalNAc III (Siat7C)	*Pan troglodytes*	n.d.	AJ634454 CAG25676.1		
α-2,6-sialyltransferase ST6GalNAc IV (Siat7D) (fragment)	*Pan troglodytes*	n.d.	AJ646870 CAG26699.1		
α-2,6-sialyltransferase ST6GalNAc V (Siat7E)	*Pan troglodytes*	n.d.	AJ646875 CAG26704.1		
α-2,6-sialyltransferase ST6GalNAc VI (Siat7F) (fragment)	*Pan troglodytes*	n.d.	AJ646884 CAG26711.1		
α-2,8-sialyltransferase 8A (Siat8A)	*Pan troglodytes*	2.4.99.8	AJ697658 CAG26896.1		
α-2,8-sialyltransferase 8B (Siat8B)	*Pan troglodytes*	n.d.	AJ697659 CAG26897.1		
α-2,8-sialyltransferase 8C (Siat8C)	*Pan troglodytes*	n.d.	AJ697660 CAG26898.1		
α-2,8-sialyltransferase 8D (Siat8D)	*Pan troglodytes*	n.d.	AJ697661 CAG26899.1		
α-2,8-sialyltransferase 8E (Siat8E)	*Pan troglodytes*	n.d.	AJ697662 CAG26900.1		
α-2,8-sialyltransferase 8F (Siat8F)	*Pan troglodytes*	n.d.	AJ697663 CAG26901.1		
β-galactosamide α-2,6-sialyltransferase I (ST6Gal I; Siat1)	*Pan troglodytes*	2.4.99.1	AJ627624 CAF29492.1		
β-galactosamide α-2,6-sialyltransferase II (ST6Gal II)	*Pan troglodytes*	n.d.	AJ627625 CAF29493.1		
GM3 synthase ST3Gal V	*Pan troglodytes*	n.d.	AJ744807 CAG32843.1		

FIGURE 11I

Protein	Organism	EC#	GenBank / GenPept	SwissProt	PDB / 3D
(Siat9)					
S13[illegible]	[illegible]	n.d.	NC_001266 NP_[illegible]		
α-2,3-sialyltransferase ST3Gal III	*Rattus norvegicus*	2.4.99.6	M97754 AAA42146.1 NM_031697 NP_113885.1	Q02734	
α-2,3-sialyltransferase ST3Gal IV (Siat4c)	*Rattus norvegicus*	n.d.	AJ626825 CAE2[illegible]		
α-2,3-sialyltransferase ST3Gal VI	*Rattus norvegicus*	n.d.	AJ626743 CAF25053.1		
α-2,6-sialyltransferase ST3Gal II	*Rattus norvegicus*	2.4.99.-	X76988 CAA54293.1 NM_031695 NP_1[illegible].1	Q11205	
α-2,6-sialyltransferase ST6Gal I	*Rattus norvegicus*	2.4.99.1	M18769 AAA41196.1 M83143 AAB07233.1	P13721	
α-2,6-sialyltransferase ST6GalNAc I (Siat7A)	*Rattus norvegicus*	n.d.	AJ634458 CAG25[illegible]		
α-2,6-sialyltransferase ST6GalNAc II (Siat7B)	*Rattus norvegicus*	n.d.	AJ634457 CAG25679.1		
α-2,6-sialyltransferase ST6GalNAc III	*Rattus norvegicus*	2.4.99.-	L29554 AAC42086.1 BC072501 AAH7250[illegible] NM_019123 NP_0619[illegible]	Q64686	
α-2,6-sialyltransferase ST6GalNAc IV (Siat7D) (fragment)	*Rattus norvegicus*	n.d.	AJ646871 CAG26700.1		
α-2,6-sialyltransferase ST6GalNAc V (Siat7E)	*Rattus norvegicus*	n.d.	AJ646872 CAG26701.1		
α-2,6-sialyltransferase ST6GalNAc VI (Siat7F) (fragment)	*Rattus norvegicus*	n.d.	AJ646881 CAG26710.1		
α-2,8-sialyltransferase (GD3 synthase) ST8Sia I	*Rattus norvegicus*	2.4.99.-	U53883 AAC27541.1 D45255 BAA08213.1	P70564 P97713	
α-2,8-sialyltransferase (SIAT8E)	*Rattus norvegicus*	n.d.	AJ699422 CAG27884.1		
α-2,8-sialyltransferase (SIAT8F)	*Rattus norvegicus*	n.d.	AJ699423 CAG27885.1		
α-2,8-sialyltransferase ST8Sia II	*Rattus norvegicus*	2.4.99.-	L13445 AAA42147.1 NM_057156 NP_476497.1	Q07977 Q64688	
α-2,8-sialyltransferase ST8Sia III	*Rattus norvegicus*	2.4.99.-	U55938 AAB50061.1 NM_013029 NP_037161.1	P97877	
α-2,8-sialyltransferase ST8Sia IV	*Rattus norvegicus*	2.4.99.-	U90215 AAB49989.1	O08563	
β-galactosamide α-2,6-sialyltransferase II (ST6Gal II)	*Rattus norvegicus*	n.d.	AJ627628 CAF29494.1		
GM3 synthase ST3Gal V	*Rattus norvegicus*	n.d.	AB018049 BAA33492.1 NM_031337 NP_112627.1	O88830	
sialyltransferase ST3Gal-I (Siat4A)	*Rattus norvegicus*	n.d.	AJ748840 CAG[illegible]		
α-2,3-sialyltransferase (St3Gal-II)	*Silurana tropicalis*	n.d.	AJ585763 CAE51387.1		
α-2,6-sialyltransferase (Siat7b)	*Silurana tropicalis*	n.d.	AJ620650 CAF05849.1		
α-2,6-sialyltransferase (St6galnac)	*Strongylocentrotus purpuratus*	n.d.	AJ699425 CAG27887.1		
α-2,3-sialyltransferase (ST3GAL-III)	*Sus scrofa*	n.d.	AJ585765 CAE51389.1		
α-2,3-sialyltransferase (ST3GAL-IV)	*Sus scrofa*	n.d.	AJ584674 CAE48299.1		
α-2,3-sialyltransferase ST3Gal I	*Sus scrofa*	2.4.99.4	M97753 AAA31125.1	Q0274[illegible]	
α-2,6-sialyltransferase (fragment) ST6Gal I	*Sus scrofa*	2.4.99.1	AF136746 AAD33059.1	Q9XSG8	
β-galactosamide α-2,6-sialyltransferase (ST6GalNAc-V)	*Sus scrofa*	n.d.	AJ620948 CAF06585.2		
sialyltransferase (fragment) ST6Gal I	*sus scrofa*	n.d.	AF041031 AAC15633.1	O62717	

FIGURE 11J

Protein	Organism	EC#	GenBank / GenPept	SwissProt	PDB / 3D
ST6GALNAC-V	*Sus scrofa*	n.d.	AJ620948 C[illegible]		
α-2,3-sialyltransferase (Siat5-r)	*Takifugu rubripes*	n.d.	AJ744805 CAG32841.1		
α-2,3-sialyltransferase ST3Gal I (Siat4)	*Takifugu rubripes*	n.d.	AJ626816 CAF25[illegible]		
α-2,3-sialyltransferase ST3Gal II (Siat5) (fragment)	*Takifugu rubripes*	n.d.	AJ626817 CAF25175.1		
α-2,3-sialyltransferase ST3Gal III (Siat6)	*Takifugu rubripes*	n.d.	AJ626818 CAF25[illegible]		
α-2,6-sialyltransferase ST6Gal I (Siat1)	*Takifugu rubripes*	n.d.	AJ744800 CAG32836.1		
α-2,6-sialyltransferase ST6GalNAc II (Siat7B)	*Takifugu rubripes*	n.d.	AJ634460 CAG25[illegible]		
α-2,6-sialyltransferase ST6GalNAc II B (Siat7B-related)	*Takifugu rubripes*	n.d.	AJ634461 CAG25682.1		
α-2,6-sialyltransferase ST6GalNAc III (Siat7C) (fragment)	*Takifugu rubripes*	n.d.	AJ634456 CAG25678.1		
α-2,6-sialyltransferase ST6GalNAc IV (siat7D) (fragment)	*Takifugu rubripes*	2.4.99.3	Y17466 CAB44338.1 AJ646869 CAG26698.1	Q9W6U6	
α-2,6-sialyltransferase ST6GalNAc V (Siat7E) (fragment)	*Takifugu rubripes*	n.d.	AJ646873 CAG26702.1		
α-2,6-sialyltransferase ST6GalNAc VI (Siat7F) (fragment)	*Takifugu rubripes*	n.d.	AJ646880 CAG26709.1		
α-2,8-sialyltransferase ST8Sia I (Siat 8A) (fragment)	*Takifugu rubripes*	n.d.	AJ715534 CAG29373.1		
α-2,8-sialyltransferase ST8Sia II (Siat 8B) (fragment)	*Takifugu rubripes*	n.d.	AJ715538 CAG29377.1		
α-2,8-sialyltransferase ST8Sia III (Siat 8C) (fragment)	*Takifugu rubripes*	n.d.	AJ715541 CAG29380.1		
α-2,8-sialyltransferase ST8Sia IIIr (Siat 8Cr)	*Takifugu rubripes*	n.d.	AJ715542 CAG29381.1		
α-2,8-sialyltransferase ST8Sia V (Siat 8E) (fragment)	*Takifugu rubripes*	n.d.	AJ715547 CAG29386.1		
α-2,8-sialyltransferase ST8Sia VI (Siat 8F) (fragment)	*Takifugu rubripes*	n.d.	AJ715549 CAG29388.1		
α-2,8-sialyltransferase ST8Sia VIr (Siat 8Fr)	*Takifugu rubripes*	n.d.	AJ715550 CAG29389.1		
α-2,3-sialyltransferase (Siat5-r)	*Tetraodon nigroviridis*	n.d.	AJ744806 CAG32842.1		
α-2,3-sialyltransferase ST3Gal I (Siat4)	*Tetraodon nigroviridis*	n.d.	AJ744802 CAG32838.1		
α-2,3-sialyltransferase ST3Gal III (Siat6)	*Tetraodon nigroviridis*	n.d.	AJ626822 CAF25180.1		
α-2,6-sialyltransferase ST6GalNAc II (Siat7B)	*Tetraodon nigroviridis*	n.d.	AJ634462 CAG25683.1		
α-2,6-sialyltransferase ST6GalNAc V (Siat7E) (fragment)	*Tetraodon nigroviridis*	n.d.	AJ646879 CAG26708.1		
α-2,8-sialyltransferase ST8Sia I (Siat 8A) (fragment)	*Tetraodon nigroviridis*	n.d.	AJ715536 CAG29375.1		
α-2,8-sialyltransferase ST8Sia II (Siat 8B) (fragment)	*Tetraodon nigroviridis*	n.d.	AJ715537 CAG29376.1		
α-2,8-sialyltransferase	*Tetraodon*	n.d.	AJ715539 CAG29378.1		

FIGURE 11K

Protein	Organism	EC#	GenBank / GenPept	SwissProt	PDB / 3D
ST8Sia [illegible] (fragment)	*nigroviridis*		[illegible]		
α-2,8-sialyltransferase ST8Sia IIIr (Siat 8Cr) (fragment)	*Tetraodon nigroviridis*	n.d.	AJ715540 CAG29379.1		
α-2,8-sialyltransferase ST8Sia V (Siat 8E) (fragment)	*Tetraodon nigroviridis*	n.d.	AJ715548 CAG293[illegible]		
α-2,3-sialyltransferase (St3Gal-II)	*Xenopus laevis*	n.d.	AJ585762 CAE51386.1		
α-2,3-sialyltransferase (St3Gal-VI)	*Xenopus laevis*	n.d.	AJ585766 CAE5[illegible]		
α-2,3-sialyltransferase St3Gal-III (Siat6)	*Xenopus laevis*	n.d.	AJ585764 CAE51388.1 AJ626823 CAF25181.1		
α-2,8-polysialyltransferase	*Xenopus laevis*	2.4.99.-	AB007468 BAA3261[illegible]	O93234	
α-2,8-sialyltransferase ST8Siα-I (Siat8A;GD3 synthase)	*Xenopus laevis*	n.d.	AY272056 AAQ16162.1 AY272057 AAQ16163.1 AJ704562 CAG28695.1		
Unknown (protein for MGC:81265)	*Xenopus laevis*	n.d.	BC068760 AAH68760.1		
α-2,3-sialyltransferase (3Gal-VI)	*Xenopus tropicalis*	n.d.	AJ626744 CAF25054.1		
α-2,3-sialyltransferase (Siat4c)	*Xenopus tropicalis*	n.d.	AJ622908 CAF22058.1		
α-2,6-sialyltransferase ST6GalNAc V (Siat7E) (fragment)	*Xenopus tropicalis*	n.d.	AJ646878 CAG26707.1		
α-2,8-sialyltransferase ST8Sia III (Siat 8C) (fragment)	*Xenopus tropicalis*	n.d.	AJ715544 CAG29383.1		
β-galactosamide α-2,6-sialyltransferase II (ST6Gal II)	*Xenopus tropicalis*	n.d.	AJ627628 CAF29496.1		
sialytransferase St8SiaI	*Xenopus tropicalis*	n.d.	AY652775 AAT67042		
poly-α-2,8-sialosyl sialyltransferase (NeuS)	*Escherichia coli K1*	2.4.-.-	M76370 AAA24213.1 X60598 CAA43053.1	Q57269 [illegible]	
polysialyltransferase	*Escherichia coli K92*	2.4.-.-	M88479 AAA24215.1	Q47404	
α-2,8 polysialyltransferase SiaD	*Neisseria meningitidis B1940,*	2.4.-.-	M95053 AAA20478.1 X78068 CAA54985.1	Q51281 Q5[illegible]145	
SynE	*Neisseria meningitidis FAM18*	n.d.	U75650 AAB53842.1	O06435	
polysialyltransferase (SiaD)(fragment)	*Neisseria meningitidis M1019*	n.d.	AY234192 AAO85290.1		
SiaD (fragment)	*Neisseria meningitidis M209*	n.d.	AY281046 AAP34769.1		
SiaD (fragment)	*Neisseria meningitidis M3045*	n.d.	AY2810[illegible] AAP3[illegible]		
polysialyltransferase (SiaD)(fragment)	*Neisseria meningitidis M3315*	n.d.	AY234191 AAO85289.1		
SiaD (fragment)	*Neisseria meningitidis M3515*	n.d.	AY281047 AAP34770.1		
polysialyltransferase (SiaD)(fragment)	*Neisseria meningitidis M4211*	n.d.	AY234190 AAO85288.1		
SiaD (fragment)	*Neisseria meningitidis M4642*	n.d.	AY281048 AAP3477[illegible]		
polysialyltransferase (SiaD)(fragment)	*Neisseria meningitidis M5177*	n.d.	AY234193 AAO85291.1		
SiaD	*Neisseria meningitidis M5178*	n.d.	AY281043 AAP3[illegible]		
SiaD (fragment)	*Neisseria meningitidis M980*	n.d.	AY281045 AAP34768.1		
NMB0067	*Neisseria meningitidis MC58*	n.d.	NC_003112 NP_273131		

FIGURE 11L

Protein	Organism	EC#	GenBank / GenPept	SwissProt	PDB / 3D
Lst	*Aeromonas punctata Sch3*	n.d.	AF126256 AAS6662[illegible]		
ORF2	*Haemophilus influenzae A2*	n.d.	M94855 **AAA24979.1**		
HI1699	*Haemophilus influenzae Rd*	n.d.	U32842 AAC2334[illegible] NC_000907 NP_[illegible]	[illegible]48211	
α-2,3-sialyltransferase	*Neisseria gonorrhoeae F62*	2.4.99.4	U60664 **AAC44539.1** AAE67205.1	P72074	
α-2,3-sialyltransferase	*Neisseria meningitidis 126E, NRCC 4010*	2.4.99.4	U60662 AAC4454[illegible]	[illegible]	
α-2,3-sialyltransferase	*Neisseria meningitidis 406Y, NRCC 4030*	2.4.99.4	U60661 **AAC44543.1**		
α-2,3-sialyltransferase (NMB0922)	*Neisseria meningitidis MC58*	2.4.99.4	U60660 AAC44541.1 AE002443 AAF41330.1 NC_003112 **NP_273962.1**	P72097	
NMA1118	*Neisseria meningitidis Z2491*	n.d.	AL162755 CAB84380.1 NC_003116 NP_283887.1	**Q9JUV5**	
PM0508	*Pasteurella multocida PM70*	n.d.	AE006086 **AAK02592.1** NC_002663 NP_245445.1	Q9CNC4	
WaaH	*Salmonella enterica SARB25*	n.d.	AF519787 **AAM82550.1**	Q8KS93	
WaaH	*Salmonella enterica SARB3*	n.d.	AF519788 **AAM82551.1**	Q8KS92	
WaaH	*Salmonella enterica SARB39*	n.d.	AF519789 **AAM82552.1**		
WaaH	*Salmonella enterica SARB53*	n.d.	AF519790 **AAM82553.1**		
WaaH	*Salmonella enterica SARB57*	n.d.	AF519791 **AAM82554.1**	Q8KS91	
WaaH	*Salmonella enterica SARB71*	n.d.	AF519793 **AAM82556.1**	Q8KS89	
WaaH	*Salmonella enterica SARB8*	n.d.	AF519792 **AAM82555.1**	Q8KS90	
WaaH	*Salmonella enterica SARC10V*	n.d.	AF519779 **AAM88840.1**	Q8KS99	
WaaH (fragment)	*Salmonella enterica SARC12*	n.d.	AF519781 **AAM88842.1**		
WaaH (fragment)	*Salmonella enterica SARC13I*	n.d.	AF519782 **AAM88843.1**	Q8KS98	
WaaH (fragment)	*Salmonella enterica SARC14I*	n.d.	AF519783 **AAM88844.1**	Q8KS97	
WaaH	*Salmonella enterica SARC15II*	n.d.	AF519784 **AAM88845.1**	Q8KS9[illegible]	
WaaH	*Salmonella enterica SARC16II*	n.d.	AF519785 **AAM88846.1**	Q8KS95	

FIGURE 11M

Protein	Organism	EC#	GenBank / GenPept	SwissProt	PDB / 3D
WaaH (fragment)	*Salmonella enterica SARC3I*	n.d.	AF519772 **AAM88[illegible]**	Q8KSA4	
WaaH (fragment)	*Salmonella enterica SARC4I*	n.d.	AF519773 **AAM88835.1**	Q8KSA3	
WaaH	*Salmonella enterica SARC5IIa*	n.d.	AF519774 **AAM8883[illegible]**	[illegible]	
WaaH	*Salmonella enterica SARC6IIa*	n.d.	AF519775 **AAM88837.1**	Q8KSA2	
WaaH	*Salmonella enterica SARC8*	n.d.	AF519777 **AAM88838.1**	Q8KSA1	
WaaH	*Salmonella enterica SARC9V*	n.d.	AF519778 **AAM88839.1**	Q8KSA0	
UDP-glucose : α-1,2-glucosyltransferase (WaaH)	*Salmonella enterica subsp. arizonae SARC 5*	2.4.1.-	AF511116 **AAM48166.1**		
bifunctional α-2,3/-2,8-sialyltransferase (Cst-II)	*Campylobacter jejuni ATCC 43449*	n.d.	AF401529 **AAL06004.1**	Q93CZ5	
Cst	*Campylobacter jejuni 81-176*	n.d.	AF305571 **AAL09368.1**		
α-2,3-sialyltransferase (Cst-III)	*Campylobacter jejuni ATCC 43429*	2.4.99.-	AY044156 **AAK73183.1**		
α-2,3-sialyltransferase (Cst-III)	*Campylobacter jejuni ATCC 43430*	2.4.99.-	AF400047 **AAK85419.1**		
α-2,3-sialyltransferase (Cst-II)	*Campylobacter jejuni ATCC 43432*	2.4.99.-	AF215659 **AAG43979.1**	Q9F0M9	
α-2,3/8-sialyltransferase (CstII)	*Campylobacter jejuni ATCC 43438*	n.d.	AF400048 **AAK91725.1**	Q93MQ0	
α-2,3-sialyltransferase cst-II	*Campylobacter jejuni ATCC 43446*	2.4.99.-	AF167344 **AAF34137.1**		
α-2,3-sialyltransferase (Cst-II)	*Campylobacter jejuni ATCC 43456*	2.4.99.-	AF401528 **AAL05990.1**	Q93D05	
α-2,3-/α-2,8-sialyltransferase (CstII)	*Campylobacter jejuni ATCC 43460*	2.4.99.-	AY044868 **AAK96001.1**	Q938X6	
α-2,3/8-sialyltransferase (Cst-II)	*Campylobacter jejuni ATCC 700297*	n.d.	AF216647 **AAL36462.1**		
ORF	*Campylobacter jejuni GB11*	n.d.	AY422[illegible] **AAR82875.1**		
α-2,3-sialyltransferase cstIII	*Campylobacter jejuni MSC57360*	2.4.99.-	AF195055 **AAG29922.1**		
α-2,3-sialyltransferase cstIII Cj1140	*Campylobacter jejuni NCTC 11168*	2.4.99.-	AL139077 **CAB73395.1** NC_002163 **NP_282288.1**	Q9PNF4	
α-2,3/α-2,8-sialyltransferase II (cstII)	*Campylobacter jejuni O:10*	n.d.	- **AAO96669.1** AX934427 **CAF04167.1**		
α-2,3/α-2,8-sialyltransferase II (CstII)	*Campylobacter jejuni O:19*	n.d.	AX934431 **CAF04169.1**		
α-2,3/α-2,8-sialyltransferase II (CstII)	*Campylobacter jejuni O:36*	n.d.	AX934436 **CAF04171.1**		
α-2,3/α-2,8-sialyltransferase II (CstII)	*Campylobacter jejuni O:4*	n.d.	AX934434 **CAF04170.1**		
α-2,3/α-2,8-sialyltransferase II (CstII)	*Campylobacter jejuni O:41*	n.d.	- **AAO96670.1** - **AAT17967.1** AX934429 **CAF04168.1**		
α-2,3-sialyltransferase cst-I	*Campylobacter jejuni OH4384*	2.4.99.-	AF130466 **AAF13495.1** [illegible] **AAS36261.1**	Q9RGF1	
bifunctional α-2,3/-2,8-sialyltransferase (Cst-II)	*Campylobacter jejuni OH4384*	2.4.99.-	AF130984 **AAF31771.1** AX934425 **CAF04166.1**		1RO7 C 1RO8 A

FIGURE 11N

Protein	Organism	EC#	GenBank / GenPept	SwissProt	PDB / 3D
HI0352 (fragment)	*Haemophilus influenzae Rd*	n.d.	U32720 AAC2201[illegible] X57315 CAA4056[illegible] NC_000907 NP_43851[illegible]	[illegible]324	
PM1174	*Pasteurella multocida PM70*	n.d.	AE006157 AAK03258.1 NC_002663 NP_246111.1	Q9CLP3	
Sequence 10 from patent US 6503744	*Unknown.*	n.d.	-AAO9[illegible]		
Sequence 10 from patent US 6699705	*Unknown.*	n.d.	-AAT17969.1		
Sequence 12 from patent US 6699705	*Unknown.*	n.d.	-AAT17970.1		
Sequence 2 from patent US 6709834	*Unknown.*	n.d.	-AAT23232.1		
Sequence 3 from patent US 6503744	*Unknown.*	n.d.	-AAO96668.1		
Sequence 3 from patent US 6699705	*Unknown.*	n.d.	-AAT17965.1		
Sequence 34 from patent US 6503744	*Unknown.*	n.d.	-AAO96684.1		
Sequence 35 from patent US 6503744 (fragment)	*Unknown.*	n.d.	-AAO96685.1 -AAS36262.1		
Sequence 48 from patent US 6699705	*Unknown.*	n.d.	-AAT17988.1		
Sequence 5 from patent US 6699705	*Unknown.*	n.d.	-AAT17966.1		
Sequence 9 from patent US 6503744	*Unknown.*	n.d.	-AAO96671.1		