# O-LINKED GLYCOSYLATION OF PEPTIDES

## CROSS-REFERENCES TO RELATED APPLICATIONS

**[0001]** The present application is related to U.S. Provisional Patent Application No. 60/535,284, filed January 8, 2004, U.S. Provisional Patent Application No. 60/544,411, filed February 12, 2004; U.S. Provisional Patent Application No. 60/546,631, filed February 20, 2004; U.S. Provisional Patent Application No. 60/555,813, filed March 23, 2004; U.S. Provisional Patent Application No. 60/570,891, filed May 12, 2004; each of which are incorporated herein by reference in their entirety for all purposes.

## Field of the Invention

**[0002]** The present invention relates to O-linked glycosylated glycopeptides, particularly therapeutic peptides and peptide mutants that include O-linked glycosylation sites not present in the wild-type peptide.

**[0003]** The administration of glycosylated and non-glycosylated peptides for engendering a particular physiological response is well known in the medicinal arts. For example, both purified and recombinant hGH are used for treating conditions and diseases due to hGH deficiency, *e.g.*, dwarfism in children, interferon has known antiviral activity and granulocyte colony stimulating factor stimulates the production of white blood cells.

**[0004]** A principal factor that has limited the use of therapeutic peptides is the difficulty inherent in engineering an expression system to express a peptide having the glycosylation pattern of the wild-type peptide. As is known in the art, improperly or incompletely glycosylated peptides can be immunogenic, leading to neutralization of the peptide and/or leading to the development of an allergic response. Other deficiencies of recombinantly produced glycopeptides include suboptimal potency and rapid clearance rates.

**[0005]** One approach to solving the problems inherent in the production of glycosylated peptide therapeutics has been to modify the peptides *in vitro* after they are expressed. Post-expression *in vitro* modification has been used to both modify of glycan structures and introduce of glycans at novel sites. A comprehensive toolbox of recombinant eukaryotic glycosyltransferases has become available, making *in vitro* enzymatic synthesis of mammalian glycoconjugates with custom designed glycosylation patterns and glycosyl

structures possible. See, for example, U.S. Patent No. 5,876,980; 6,030,815; 5,728,554; 5,922,577; and WO/9831826; US2003180835; and WO 03/031464.

**[0006]** In addition to manipulating the structure of glycosyl groups on polypeptides, glycopeptides can be prepared with one or more non-saccharide modifying groups, such as water soluble polymers. 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.

**[0007]** 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).

**[0008]** 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.

**[0009]** 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).

**[0010]** Carbohydrates are attached to glycopeptides in several ways of which N-linked to asparagine and mucin-type O-linked to serine and threonine are the most relevant for recombinant glycoprotein therapeuctics. Unfortunately, not all polypeptide comprise an N- or O-linked glycosylation site as part of their primary amino acid sequence. In other cases an existing glycosylation site may be inconvenient for the attachment of a modifying group (e.g., a water-soluble or water –insoluable polymers, therapeutic moieties, and or biomolecules) to the polypeptide, or attachment of such moieties at that site may cause an undesirable decrease in biological activity of the polypeptide. Thus there is a need in the art for methods that permit both the precise creation of glycosylation sites and the ability to precisely direct the modification of those sites.

## SUMMARY OF THE INVENTION

**[0011]** It is a discovery of the present invention that enzymatic glycoconjugation reactions can be specifically targeted to O-linked glycosylation sites and to glycosyl residues that are attached to O-linked glycosylation sites. The targeted O-linked glycosylation sites can be sites native to a wild-type peptide or, alternatively, they can be introduced into a peptide by mutation. Accordingly, the present invention provides polypeptides comprising mutated sites suitable for O-linked glycosylation and pharmaceutical compositions thereof. In addition, the present invention provides methods of making such polypeptides and using such polypeptides and/or pharmaceutical compositions thereof for therapeutic treatments.

**[0012]** Thus, in a first aspect, the invention provides an isolated polypeptide comprising a mutant peptide sequence, wherein the mutant peptide sequence encodes an O-linked glycosylation site that does not exist in the corresponding wild-type polypeptide.

**[0013]** In one embodiment, the isolated polypeptide is a G-CSF polypeptide.

**[0014]** In one embodiment, the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $M^1X_nTPLGP$ or $M^1B_oPZ_mX_nTPLGP$. In this embodiment, the superscript, 1, denotes the first position of the amino acid sequence of the wild-type G-CSF sequence (SEQ ID NO:3), the subscripts n and m are integers selected from 0 to 3, and at

least one of X and B is threonine or serine, and when more than one of X and B is threonine or serine, the identity of these moieties is independently selected. Also in this embodiment, Z is selected from glutamate, any uncharged amino acid or dipeptide combination including MQ, GQ, and MV. In another embodiment, the G-CSF polypeptide comprises a mutant peptide sequence selected from the sequences consisting of MVTPLGP, MQTPLGP, MIATPLGP, MATPLGP, MPTQGAMPLGP , MVQTPLGP, MQSTPLGP, MGQTPLGP, MAPTSSSPLGP, and MAPTPLGPA.

**[0015]** In another embodiment, the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $M^1TPXBO_rP$. In this embodiment the superscript, 1, denotes the first position of the amino acid sequence of the wild-type G-CSF sequence (SEQ ID NO:3), and the subscript r is an integer selected from 0 to 3, and at least one of X, B and O is threonine or serine, and when more than one of X, B and O is threonine or serine, the identity of these moieties is independently selected. In another embodiment, the G-CSF polypeptide comprises a mutant peptide sequence selected from the sequences consisting of: MTPTLGP, MTPTQLGP, MTPTSLGP, MTPTQGP, MTPTSSP, $M^1TPQTP$, $M^1TPTGP$, $M^1TPLTP$, $M^1TPNTGP$, MTPLGP, $M^1TPVTP$, $M^1TPMVTP$, and $MT^1P^2TQGL^3G^4P^5A^6S^7$.

**[0016]** In another embodiment, the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $LGX^{53}B_oLGI$, wherein the superscript denotes the position of the amino acid in the wild type G-CSF amino acid sequence, and X is histidine, serine, arginine, glutamic acid or tyrosine, and B is either threonine or serine, and o is an integer from 0 to 3. In another embodiment, the G-CSF polypeptide comprises a mutant peptide sequence selected from the sequences consisting of: LGHTLGI, LGSSLGI, LGYSLGI, LGESLGI, and LGSTLGI.

**[0017]** In another embodiment, the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $P^{129}Z_mJ_qO_rX_nPT$ wherein the superscript denotes the position of the amino acid in the wild type G-CSF amino acid sequence, and Z, J, O and X are independently selected from threonine or serine, and m, q, r, and n are integers independently selected from 0 to 3. In another embodiment, the G-CSF polypeptide comprises a mutant peptide sequence selected from the sequences consisting of: $P^{129}ATQPT$, $P^{129}TLGPT$, $P^{129}TQGPT$, $P^{129}TSSPT$, $P^{129}TQGAPT$, $P^{129}NTGPT$, PALQPTQT, $P^{129}ALTPT$, $P^{129}MVTPT$, $P^{129}ASSTPT$, $P^{129}TTQP$, $P^{129}NTLP$, $P^{129}TLQP$, $MAP^{129}ATQPTQGAM$, and $MP^{129}ATTQPTQGAM$.

**[0018]** In another embodiment, the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $PZ_mU_sJ_qP^{61}O_rX_nB_oC$ wherein the superscript denotes the position of the amino acid in the wild type G-CSF amino acid sequence, and at least one of Z, J, O, and U is

selected from threonine or serine, and when more than one of Z, J, O and U is threonine or serine, each is independently selected, X and B are any uncharged amino acid or glutamate, and m, s, q, r, n, and o are integers independently selected from 0 to 3. In another embodiment the G-CSF polypeptide comprises a mutant peptide sequence selected from the sequences consisting of: $P^{61}$TSSC, $P^{61}$TSSAC, LGIPTA $P^{61}$LSSC, LGIPTQ $P^{61}$LSSC, LGIPTQG $P^{61}$LSSC, LGIPQT $P^{61}$LSSC, LGIPTS $P^{61}$LSSC, LGIPTS $P^{61}$LSSC, LGIPTQP$^{61}$LSSC, LGTPWAP$^{61}$LSSC, LGTPFA $P^{61}$LSSC, $P^{61}$FTP, and SLGAP$^{58}$TAP$^{61}$LSS.

**[0019]** In another embodiment the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $Ø_aG_pJ_qO_rP^{175}X_nB_oZ_mU_s\Psi_t$ wherein the superscript denotes the position of the amino acid in SEQ ID NO:3, and at least one of Z, U, O, J, G, Ø, B and X is threonine or serine and when more than one of Z, U, O, J, G, Ø, B and X are threonine or serine, they are independently selected. Ø is optionally R, and G is optionally H. The symbol Ψ represents any uncharged amino acid residue or glutamate, and a, p, q, r, n, o, m, s, and t are integers independently selected from 0 to 3. In another embodiment the G-CSF polypeptide comprises a mutant peptide sequence selected from the sequences consisting of: RHLAQTP$^{175}$, RHLAGQTP$^{175}$, QP$^{175}$TQGAMP, RHLAQTP$^{175}$AM, QP$^{175}$TSSAP, QP$^{175}$TSSAP, QP$^{175}$TQGAMP, QP$^{175}$TQGAM, QP$^{175}$TQGA, QP$^{175}$TVM, QP$^{175}$NTGP, and QP$^{175}$QTLP.

**[0020]** In another embodiment the G-CSF polypeptide comprises a mutant peptide sequence selected from the sequences $P^{133}$TQTAMP$^{139}$, $P^{133}$TQGTMP, $P^{133}$TQGTNP, $P^{133}$TQGTLP, and PALQP$^{133}$TQTAMPA.

**[0021]** In another embodiment, the isolated polypeptide is an hGH polypeptide.

**[0022]** In one embodiment, the hGH polypeptide comprises a mutant peptide sequence with the formula of $P^{133}$JXBOZUK$^{140}$QTYS, wherein superscripts denote the position of the amino acid in (SEQ ID NO:20); and J is selected from threonine and arginine; X is selected from alanine, glutamine, isoleucine, and threonine; B is selected from glycine, alanine, leucine, valine, asparagine, glutamine, and threonine; O is selected from tyrosine, serine, alanine, and threonine; and Z is selected from isoleucine and methionine; and U is selected from phenylalanine and proline. In another embodiment, the hGH polypeptide comprises a mutant peptide sequence selected from the sequences consisting of: PTTGQIFK, PTTAQIFK, PTTLQIFK, PTTLYVFK, PTTVQIFK, PTTVSIFK, PTTNQIFK, PTTQQIFK, PTATQIFK, PTQGQIFK, PTQGAIFK, PTQGAMFK, PTIGQIFK, PTINQIFK, PTINTIFK, PTILQIFK, PTIVQIFK, PTIQQIFK, PTIAQIFK, $P^{133}$TTTQIFK$^{140}$QTYS, and $P^{133}$TQGAMPK$^{140}$QTYS.

**[0023]** In another embodiment, the hGH polypeptide comprises a mutant peptide sequence with the formula of $P^{133}$RTGQIPTQBYS wherein superscripts denote the position of the amino acid in SEQ ID NO:20; and B is selected from alanine and threonine. In another embodiment, the hGH polypeptide comprises a mutant peptide sequence selected from the sequences consisting of: PRTGQIPTQTYS and PRTGQIPTQAYS.

**[0024]** In another embodiment, the hGH polypeptide comprises a mutant peptide sequence with the formula of $L^{128}$XTBOP$^{133}$UTG wherein superscripts denote the position of the amino acid inSEQ ID NO:20; and X is selected from glutamic acid, valine and alanine; B is selcted from glutamine, glutamic acid, and glycine; O is selcted from serine and threonine; and U is selected from arginine, serine, alanine and leucine. In another embodiment, the hGH polypeptide comprises a mutant peptide sequence selected from the sequences consisting of: LETQSP$^{133}$RTG, LETQSP$^{133}$STG, LETQSP$^{133}$ATG, LETQSP$^{133}$LTG, LETETP$^{133}$R, LETETP$^{133}$A, LVTQSP$^{133}$RTG, LVTETP$^{133}$RTG, LVTETP$^{133}$ATG, and LATGSP$^{133}$RTG.

**[0025]** In another embodiment the hGH polypeptide comprises a mutant peptide sequence with the formula of $M^1$BPTX$_n$Z$_m$OPLSRL wherein the superscript 1, denotes the position of the amino acid in SEQ ID NO:19; and B is selected from phenylalanine, valine and alanine or a combination thereof; X is selected from glutamate, valine and proline Z is threonine; O is selected from leucine and isoleucine; and when X is proline, Z is threonine; and wherein n and m are integers selected from 0 and 2. In another embodiment, the hGH polypeptide comprises a mutant peptide sequence selected from the sequences consisting of: $M^1$FPTE IPLSRL, $M^1$FPTV LPLSRL, and $M^1$APTPTIPLSRL.

**[0026]** In still another embodiment the the hGH polypeptide comprises the following mutant peptide sequence: $M^1$VTPTIPLSRL.

**[0027]** In still another embodiment the hGH polypeptide comprises a mutant peptide sequence selected from $M^1$APTSSPTIPL$^7$SR$^9$ and DGSP$^{133}$NTGQIFK$^{140}$.

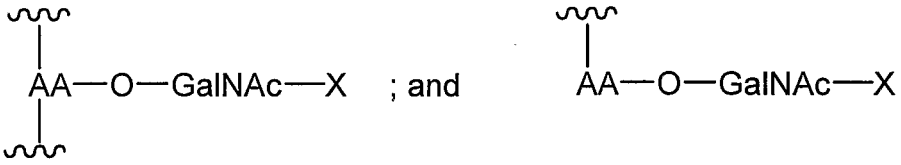
**[0028]** In another embodiment the isolated polypeptide is an IFN alpha polypeptide.

**[0029]** In one embodiment, the INF alpha polypeptide has a peptide sequence comprising a mutant amino acid sequence, and the peptide sequence corresponds to a region of INF alpha 2 having a sequence as shown in SEQ NO:22, and wherein the mutant amino acid sequence contains a mutation at a position corresponding to $T^{106}$ of INF alpha 2. In another embodiment the IFN alpha polypeptide is selected from the group consisting of IFN alpha, IFN alpha 4, IFN alpha 5, IFN alpha 6, IFN alpha 7, IFN alpha 8, IFN alpha 10, IFN alpha 14, IFN alpha 16, IFN alpha 17, and IFN alpha 21. In yet another embodiment, the IFN alpha polypeptide is an IFN alpha polypeptide comprising a mutant amino acid sequence selected

from the group consisting of: $^{99}$CVMQEERVTETPLMNADSIL$^{118}$, $^{99}$CVMQEEGVTETPLMNADSIL$^{118}$, and $^{99}$CVMQGVGVTETPLMNADSIL$^{118}$. In still another embodiment, the IFN alpha polypeptide is an IFN alpha 4 polypeptide comprising a mutant amino acid sequence selected from the group consisting of: $^{99}$CVIQEVGVTETPLMNVDSIL$^{118}$, and $^{99}$CVIQGVGVTETPLMKEDSIL$^{118}$. In another embodiment, the IFN alpha polypeptide is an IFN alpha 5 polypeptide comprising a mutant amino acid sequence selected from the group consisting of: $^{99}$CMMQEVGVTDTPLMNVDSIL$^{118}$, $^{99}$CMMQEVGVTETPLMNVDSIL$^{118}$ and $^{99}$CMMQGVGVTDTPLMNVDSIL$^{118}$. In an another embodiment, the IFN alpha polypeptide is an IFN alpha 6 polypeptide comprising a mutant amino acid sequence selected from the group consisting of: $^{99}$CVMQEVWVTGTPLMNEDSIL$^{118}$, $^{99}$CVMQEVGVTGTPLMNEDSIL$^{118}$, and $^{99}$CVMQGVGVTETPLMNEDSIL$^{118.}$ In yet an another embodiment, the IFN alpha polypeptide is an IFN alpha 7 polypeptide comprising a mutant amino acid sequence selected from the group consisting of: $^{99}$CVIQEVGVTETPLMNEDFIL$^{118}$, and $^{99}$CVIQGVGVTETPLMNEDFIL$^{118}$. In still another embodiment, the IFN alpha polypeptide is an IFN alpha 8 polypeptide comprising a mutant amino acid sequence selected from the group consisting of: $^{99}$CVMQEVGVTESPLMYEDSIL$^{118}$, and $^{99}$CVMQGVGVTESPLMYEDSIL$^{118}$. In another embodiment, the IFN alpha polypeptide is an IFN alpha 10 polypeptide comprising a mutant amino acid sequence selected from the group consisting of: $^{99}$CVIQEVGVTETPLMNEDSIL$^{118}$, and $^{99}$CVIQGVGVTETPLMNEDSIL$^{118}$. In another embodiment, the IFN alpha polypeptide is an IFN alpha 14 polypeptide comprising a mutant amino acid sequence selected from the group consisting of: $^{99}$CVIQEVGVTETPLMNEDSIL$^{118}$, and $^{99}$CVIQGVGVTETPLMNEDSIL$^{118}$. In another embodiment, the IFN alpha polypeptide is an IFN alpha 16 polypeptide comprising a mutant amino acid sequence selected from the group consisting of: $^{99}$CVTQEVGVTEIPLMNEDSIL$^{118}$, $^{99}$CVTQEVGVTETPLMNEDSIL$^{118}$, and $^{99}$CVTQGVGVTETPLMNEDSIL$^{118}$. In still another embodiment, the IFN alpha polypeptide is an IFN alpha 17 polypeptide comprising a mutant amino acid sequence selected from the group consisting of: $^{99}$CVIQEVGMTETPLMNEDSIL$^{118}$, $^{99}$CVIQEVGVTETPLMNEDSIL$^{118}$, and $^{99}$CVIQGVGMTETPLMNEDSIL$^{118}$. In one more embodiment, the IFN alpha polypeptide is an IFN alpha 21 polypeptide comprising a mutant amino acid sequence selected from the group consisting of: $^{99}$CVIQEVGVTETPLMNVDSIL$^{118}$, and $^{99}$CVIQGVGVTETPLMNVDSIL$^{118}$.

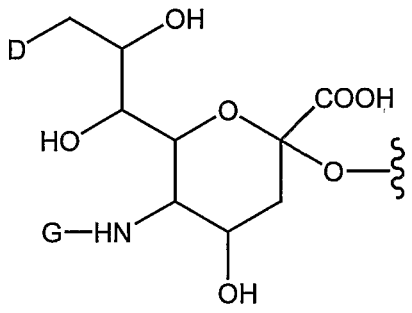
**[0030]** In a second aspect, the invention provides an isolated nucleic acid encoding a polypeptide comprising a mutant peptide sequence, wherein the mutant peptide sequence encodes an O-linked glycosylation site that does not exist in the corresponding wild-type polypeptide. In one embodiment the nucleic acid encoding a polypeptide comprising a mutant peptide sequence is comprised within an expression cassette. In another related embodiment, the present invention provides a cell comprises the nucleic acid of the present invention.

**[0031]** In a third aspect, the isolated polypeptide comprising a mutant peptide sequence, that encodes an O-linked glycosylation site that not existing in the corresponding wild-type polypeptide, has a formula selected from:



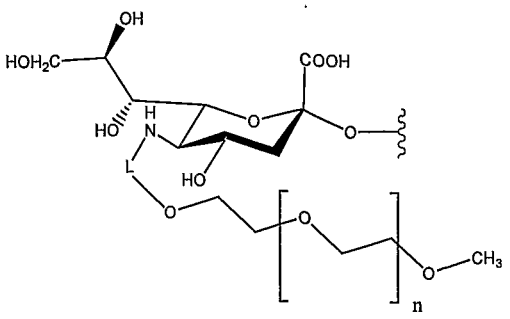

wherein AA is an amino acid side chain that comprises a hydroxyl moiety that is within the mutant polypeptide sequence; and X is a modifying group or a saccharyl moiety. In one embodiment X comprises a group selected from sialyl, galactosyl and Gal-Sia moieties, wherein at least one of said sialyl, galactosyl and Gal-Sia comprises a modifying group.

**[0032]** In another embodiment X comprises the moiety:



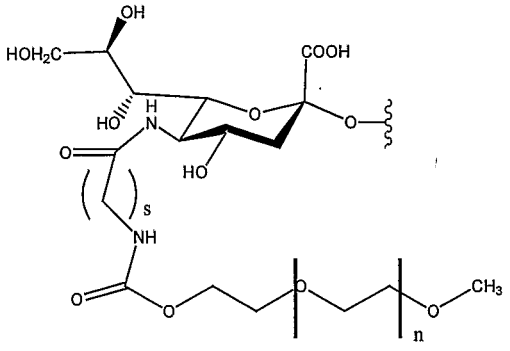

wherein D is a member selected from -OH and $R^1$-L-HN-;G is a member selected from $R^1$-L- and -C(O)($C_1$-$C_6$)alkyl; $R^1$ is a moiety comprising a member selected a moiety comprising a straight-chain or branched poly(ethylene glycol) residue; and L is a linker which is a member selected from a 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-.

**[0033]** In another embodiment X comprises the structure:



in which L is a substituted or unsubstituted alkyl or substituted or unsubstituted heteroalkyl group; and n is selected from the integers from 0 to about 500.

**[0034]** In another embodiment, X comprises the structure:



in which s is selected from the integers from 0 to 20.

**[0035]** In a fourth aspect the invention provides a method for making a glycoconjugate of an isolated polypeptide comprising a mutant peptide sequence encoding an O-linked glycosylation site that does not existing in the corresponding wild-type polypeptide, comprising the steps of:

(a) recombinantly producing the mutant polypeptide, and

(b) enzymatically glycosylating the mutant polypeptide with a modified sugar at said O-linked glycosylation site.

**[0036]** In a fifth aspect the invention provides a pharmaceutical composition of an isolated polypeptide comprising a mutant peptide sequence, wherein the mutant peptide sequence encodes an O-linked glycosylation site that does not exist in the corresponding wild-type polypeptide.

**[0037]** In one embodiment the pharmaceutical composition comprises an effective amount of a G-CSF polypeptide of the invention glycoconjugated with a modified sugar. In a related embodiment, the modified sugar is modified with a member selected from poly(ethylene glycol) and methoxy-poly(ethylene glycol) (m-PEG).

**[0038]** In another embodiment the pharmaceutical composition comprises an effective amount of an hGH polypeptide of the invention glycoconjugated with a modified sugar. In a related embodiment, the modified sugar is modified with a member selected from poly(ethylene glycol) and methoxy-poly(ethylene glycol) (m-PEG).

**[0039]** In another embodiment the pharmaceutical composition comprises an effective amount of an granulocyte macrophage colony stimulating factor polypeptide of the invention glycoconjugated with a modified sugar. In a related embodiment, the modified sugar is modified with a member selected from poly(ethylene glycol) and methoxy-poly(ethylene glycol) (m-PEG).

**[0040]** In another embodiment the pharmaceutical composition comprises an effective amount of an IFN alpha polypeptide of the invention glycoconjugated with a modified sugar. In a related embodiment, the modified sugar is modified with a member selected from poly(ethylene glycol) and methoxy-poly(ethylene glycol) (m-PEG).

**[0041]** In a sixth aspect the invention provides a method of providing therapy to a subject in need of said therapy, wherein the method comprises, administering to said subject an effective amount a pharmaceutical composition of the invention. In one embodiment, the therapy provided is G-CSF therapy. In another embodiment the therapy provided is granulocyte macrophage colony stimulating factor therapy. In another embodiment the therapy provided is interferon alpha therapy. In still another embodiment the therapy provided is Growth Hormone therapy.

**[0042]** Additional aspects, advantages and objects of the present invention will be apparent from the detailed description that follows.

## BRIEF DESCRIPTION OF THE DRAWINGS

**[0043]** **FIG. 1** is a plot of absorbance vs. GCSF concentration for unmodified G-CSF and glyco-PEG-ylated analogues in a NSF-60 cell proliferation assay.

**[0044]** **FIG. 2** is a plot of counts per minute (CPM) vs. time for a rat pharmacokinetic study using radioiodinated G-CSF and glycol—PEG-lated derivatives thereof.

**[0045]** **FIG. 3** is a plot of μg/mL G-CSF in blood vs. time (h) for a rat pharmacokinetic study using radioiodinated G-CSF and glycol—PEG-lated derivatives thereof.

**[0046]** **FIG. 4** is a plot showing the induction of white blood cells in mice using unmodified G-CSF and chemically- and glyco-PEG-ylated G-CSF.

**[0047]** **FIG. 5** is a plot of the results of an aggregation assay following radioiodination with the Bolton-Hunter reagent.

**[0048]** **FIG. 6** is a plot of the results of an accelerated stability study of glyco-PEG-ylated G-CSF.

**[0049]** **FIG. 7** is an expanded view of **FIG. 6**.

**[0050]** **FIG. 8** is a plot of the results of a rat IV pK Study using the Bolton Hunter radiolabeling process (precipitated plasma protein).

**[0051]** **FIG. 9** is a plot of the results of a rat IV pK Study using unlabeled G-CSF, chemically- and glyco-PEG-ylated G-CSF detected by ELISA.

**[0052]** **FIG. 10** showns representative sialyltransferases of use in the present invention.

## DETAILED DESCRIPTION OF THE INVENTION

### Abbreviations

**[0053]** PEG, poly(ethyleneglycol); m-PEG, methoxy-poly(ethylene glycol); PPG, poly(propyleneglycol); m-PPG, methoxy-poly(propylene glycol); Fuc, fucosyl; Gal, galactosyl; GalNAc, N-acetylgalactosaminyl; Glc, glucosyl; GlcNAc, N-acetylglucosaminyl; Man, mannosyl; ManAc, mannosaminyl acetate; Sia, sialic acid; and NeuAc, N-acetylneuraminyl.

### Definitions

**[0054]** 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.

**[0055]** 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).

**[0056]** 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.

**[0057]** The term "nucleic acid" or "polynucleotide" refers to deoxyribonucleic acids (DNA) or ribonucleic acids (RNA) and polymers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides that have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (*e.g.*, degenerate codon substitutions), alleles, orthologs, SNPs, and complementary sequences as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res.* **19**:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* **260**:2605-2608 (1985); and Rossolini *et al.*, *Mol. Cell. Probes* **8**:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, and mRNA encoded by a gene.

**[0058]** The term "gene" means the segment of DNA involved in producing a polypeptide chain. It may include regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

**[0059]** The term "isolated," when applied to a nucleic acid or protein, denotes that the nucleic acid or protein is essentially free of other cellular components with which it is associated in the natural state. It is preferably in a homogeneous state although it can be in

either a dry or aqueous solution. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein that is the predominant species present in a preparation is substantially purified. In particular, an isolated gene is separated from open reading frames that flank the gene and encode a protein other than the gene of interest. The term "purified" denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Particularly, it means that the nucleic acid or protein is at least 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure.

**[0060]** 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 having a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

**[0061]** There are various known methods in the art that permit the incorporation of an unnatural amino acid derivative or analog into a polypeptide chain in a site-specific manner, *see, e.g.*, WO 02/086075.

**[0062]** Amino acids may be referred to herein by either the commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

**[0063]** "Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, "conservatively modified variants" refers to those nucleic acids that encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of

functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein that encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid that encodes a polypeptide is implicit in each described sequence.

**[0064]** As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

**[0065]** The following eight groups each contain amino acids that are conservative substitutions for one another:

1) Alanine (A), Glycine (G);
2) Aspartic acid (D), Glutamic acid (E);
3) Asparagine (N), Glutamine (Q);
4) Arginine (R), Lysine (K);
5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V);
6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W);
7) Serine (S), Threonine (T); and
8) Cysteine (C), Methionine (M)

(*see, e.g.*, Creighton, *Proteins* (1984)).

**[0066]** Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical

Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

**[0067]** In the present application, amino acid residues are numbered according to their relative positions from the most N-terminal residue, which is numbered 1, in an unmodified wild-type polypeptide sequence.

**[0068]** "Peptide" refers to a polymer in which the monomers are amino acids and are joined together through amide bonds. Peptides of the present invention can vary in size, *e.g.*, from two amino acids to hundreds or thousands of amino acids, which alternatively is 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).

**[0069]** In the present application, amino acid residues are numbered according to their relative positions from the N-terminal, *e.g.*, the left most residue, which is numbered 1, in a peptide sequence.

**[0070]** The term "mutant polypeptide" or "mutein" refers to a form of a peptide that differs from its corresponding wild-type form or naturally existing form. A mutant peptide can contain one or more mutations, *e.g.*, replacement, insertion, deletion, etc. which result in the mutant peptide.

**[0071]** The term "peptide conjugate," refers to species of the invention in which a peptide is glycoconjugated with a modified sugar as set forth herein. In a representative example, the peptide is a mutant peptide having an O-linked glycosylation site not present in the wild-type peptide.

**[0072]** "Proximate a proline residue," as used herein refers to an amino acid that is less than about 10 amino acids removed from a proline residue, preferably, less than about 9, 8, 7,

6 or 5 amino acids removed from a proline residue, more preferably, less than about 4, 3, 2 or 1 residues removed from a proline residue. The amino acid "proximate a proline residue" may be on the C- or N-terminal side of the proline residue.

**[0073]** 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.

**[0074]** 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, water-soluble polymers, 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.

**[0075]** 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 water-soluble polymers 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). An exemplary polysaccharide is

poly(sialic acid). An exemplary poly(ether) is poly(ethylene glycol), *e.g.*, m-PEG. Poly(ethylene imine) is an exemplary polyamine, and poly(acrylic) acid is a representative poly(carboxylic acid).

**[0076]** The polymer backbone of the water-soluble polymer can be poly(ethylene glycol) (i.e. PEG). However, it should be understood that other related polymers are also suitable for use in the practice of this invention and that the use of the term PEG or poly(ethylene glycol) is intended to be inclusive and not exclusive in this respect. The term PEG includes poly(ethylene glycol) in any of its forms, including alkoxy PEG, difunctional PEG, multiarmed PEG, forked PEG, branched PEG, pendent PEG (i.e. PEG or related polymers having one or more functional groups pendent to the polymer backbone), or PEG with degradable linkages therein.

**[0077]** The polymer backbone can be linear or branched. Branched polymer backbones are generally known in the art. Typically, a branched polymer has a central branch core moiety and a plurality of linear polymer chains linked to the central branch core. PEG is commonly used in branched forms that can be prepared by addition of ethylene oxide to various polyols, such as glycerol, pentaerythritol and sorbitol. The central branch moiety can also be derived from several amino acids, such as lysine. The branched poly(ethylene glycol) can be represented in general form as $R(\text{-PEG-OH})_m$ in which R represents the core moiety, such as glycerol or pentaerythritol, and m represents the number of arms. Multi-armed PEG molecules, such as those described in U.S. Pat. No. 5,932,462, which is incorporated by reference herein in its entirety, can also be used as the polymer backbone.

**[0078]** Many other polymers are also suitable for the invention. Polymer backbones that are non-peptidic and water-soluble, with from 2 to about 300 termini, are particularly useful in the invention. Examples of suitable polymers include, but are not limited to, other poly(alkylene glycols), such as poly(propylene glycol) ("PPG"), copolymers of ethylene glycol and propylene glycol and the like, poly(oxyethylated polyol), poly(olefinic alcohol), poly(vinylpyrrolidone), poly(hydroxypropylmethacrylamide), poly(α-hydroxy acid), poly(vinyl alcohol), polyphosphazene, polyoxazoline, poly(N-acryloylmorpholine), such as described in U.S. Pat. No. 5,629,384, which is incorporated by reference herein in its entirety, and copolymers, terpolymers, and mixtures thereof. Although the molecular weight of each chain of the polymer backbone can vary, it is typically in the range of from about 100 Da to about 100,000 Da, often from about 6,000 Da to about 80,000 Da.

**[0079]** The term "glycoconjugation," as used herein, refers to the enzymatically mediated conjugation of a modified sugar species to an amino acid or glycosyl residue of a polypeptide, *e.g.*, a mutant human growth hormone of the present invention. A subgenus of "glycoconjugation" is "glycol-PEGylation," in which the modifying group of the modified sugar is poly(ethylene glycol), and alkyl derivative (*e.g.*, m-PEG) or reactive derivative (*e.g.*, $H_2N$-PEG, HOOC-PEG) thereof.

**[0080]** The terms "large-scale" and "industrial-scale" are used interchangeably and refer to a reaction cycle that produces at least about 250 mg, preferably at least about 500 mg, and more preferably at least about 1 gram of glycoconjugate at the completion of a single reaction cycle.

**[0081]** The term, "glycosyl linking group," as used herein refers to a glycosyl residue to which a modifying group (e.g., PEG moiety, therapeutic moiety, biomolecule) is covalently attached; the glycosyl linking group joins the modifying group to the remainder of the conjugate. 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. The glycosyl linking group can be a saccharide-derived structure that is degraded during formation of modifying group-modified sugar cassette (e.g., oxidation→Schiff base formation→reduction), or the glycosyl linking group may be intact. An "intact glycosyl linking group" refers to a linking group that is derived from a glycosyl moiety in which the saccharide monomer that links the modifying group and to the remainder of 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.

**[0082]** 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.

**[0083]** As used herein, "therapeutic moiety" means any agent useful for therapy including, but not limited to, antibiotics, anti-inflammatory agents, anti-tumor drugs, cytotoxins, and radioactive agents. "Therapeutic moiety" includes prodrugs of bioactive agents, constructs in which more than one therapeutic moiety is bound to a carrier, e.g, multivalent agents. Therapeutic moiety also includes proteins and constructs that include proteins. Exemplary proteins include, but are not limited to, Erythropoietin (EPO), Granulocyte Colony Stimulating Factor (GCSF), Granulocyte Macrophage Colony Stimulating Factor (GMCSF), Interferon (e.g., Interferon-α, -β, -γ), Interleukin (e.g., Interleukin II), serum proteins (e.g., Factors VII, VIIa, VIII, IX, and X), Human Chorionic Gonadotropin (HCG), Follicle Stimulating Hormone (FSH) and Lutenizing Hormone (LH) and antibody fusion proteins (e.g. Tumor Necrosis Factor Receptor ((TNFR)/Fc domain fusion protein)).

**[0084]** As used herein, "anti-tumor drug" means any agent useful to combat cancer including, but not limited to, cytotoxins and agents such as antimetabolites, alkylating agents, anthracyclines, antibiotics, antimitotic agents, procarbazine, hydroxyurea, asparaginase, corticosteroids, interferons and radioactive agents. Also encompassed within the scope of the term "anti-tumor drug," are conjugates of peptides with anti-tumor activity, *e.g.* TNF-α. Conjugates include, but are not limited to those formed between a therapeutic protein and a glycoprotein of the invention. A representative conjugate is that formed between PSGL-1 and TNF-α.

**[0085]** As used herein, "a cytotoxin or cytotoxic agent" means any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracinedione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Other toxins include, for example, ricin, CC-1065 and analogues, the duocarmycins. Still other toxins include diptheria toxin, and snake venom (*e.g.*, cobra venom).

**[0086]** As used herein, "a radioactive agent" includes any radioisotope that is effective in diagnosing or destroying a tumor. Examples include, but are not limited to, indium-111, cobalt-60. Additionally, naturally occurring radioactive elements such as uranium, radium,

and thorium, which typically represent mixtures of radioisotopes, are suitable examples of a radioactive agent. The metal ions are typically chelated with an organic chelating moiety.

**[0087]** Many useful chelating groups, crown ethers, cryptands and the like are known in the art and can be incorporated into the compounds of the invention (*e.g.*, EDTA, DTPA, DOTA, NTA, HDTA, *etc.* and their phosphonate analogs such as DTPP, EDTP, HDTP, NTP, *etc*). *See*, for example, Pitt *et al.*, "The Design of Chelating Agents for the Treatment of Iron Overload," In, INORGANIC CHEMISTRY IN BIOLOGY AND MEDICINE; Martell, Ed.; American Chemical Society, Washington, D.C., 1980, pp. 279-312; Lindoy, THE CHEMISTRY OF MACROCYCLIC LIGAND COMPLEXES; Cambridge University Press, Cambridge,1989; Dugas, BIOORGANIC CHEMISTRY; Springer-Verlag, New York, 1989, and references contained therein.

**[0088]** Additionally, a manifold of routes allowing the attachment of chelating agents, crown ethers and cyclodextrins to other molecules is available to those of skill in the art. *See*, for example, Meares *et al.*, "Properties of In Vivo Chelate-Tagged Proteins and Polypeptides." In, MODIFICATION OF PROTEINS: FOOD, NUTRITIONAL, AND PHARMACOLOGICAL ASPECTS;" Feeney, *et al.*, Eds., American Chemical Society, Washington, D.C., 1982, pp. 370-387; Kasina *et al.*, *Bioconjugate Chem.*, **9:** 108-117 (1998); Song *et al.*, *Bioconjugate Chem.*, **8:** 249-255 (1997).

**[0089]** 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.

**[0090]** As used herein, "administering" means oral administration, administration as a suppository, topical contact, intravenous, intraperitoneal, intramuscular, intralesional, or subcutaneous administration, administration by inhalation, 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), particularly by inhalation. 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.

**[0091]** 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.

**[0092]** 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).

**[0093]** 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.

**[0094]** 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%.

**[0095]** 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.

**[0096]** 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).

**[0097]** "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.

**[0098]** "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%.

**[0099]** 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.

**[0100]** "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.

**[0101]** 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.

**[0102]** 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 $-OCH_2-$.

**[0103]** 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".

**[0104]** 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.

**[0105]** 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.

**[0106]** 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 $-R'C(O)_2-$.

**[0107]** 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.

**[0108]** 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.

**[0109]** 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.

**[0110]** 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).

**[0111]** 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.

**[0112]** 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', -$CO_2$R', -CONR'R", -OC(O)NR'R", -NR"C(O)R', -NR'-C(O)NR"R"', -NR"$C(O)_2$R', -NR-C(NR'R"R"')=NR"", -NR-C(NR'R")=NR"', -S(O)R', -$S(O)_2$R', -$S(O)_2$NR'R", -$NRSO_2$R', -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).

**[0113]** 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', -$CO_2$R', -CONR'R", -OC(O)NR'R", -NR"C(O)R', -NR'-C(O)NR"R"', -NR"$C(O)_2$R', -NR-C(NR'R"R"')=NR"", -NR-C(NR'R")=NR"', -S(O)R', -$S(O)_2$R', -$S(O)_2$NR'R", -$NRSO_2$R', -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.

**[0114]** 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)-, -$S(O)_2$-, -$S(O)_2$NR'- 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 –$S(O)_2$NR'-. The substituents R, R', R" and R'" are preferably independently selected from hydrogen or substituted or unsubstituted ($C_1$-$C_6$)alkyl.

**[0115]** As used herein, the term "heteroatom" is meant to include oxygen (O), nitrogen (N), sulfur (S) and silicon (Si).

**Introduction**

**[0116]** The present invention provides conjugates of glycopeptides in which a modified sugar moiety is attached either directly or indirectly (*e.g.*, through and intervening glycosyl residue) to an O-linked glycosylation site on the peptide. Also provided are methods for producing the conjugates of the invention.

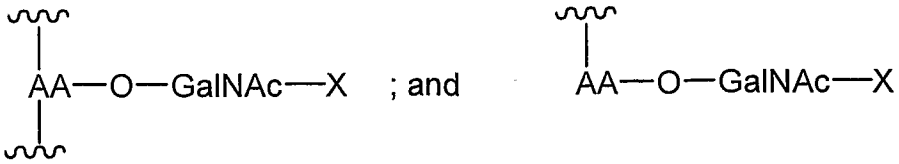
**[0117]** The O-linked glycosylation site is generally the hydroxy side chain of a natural (*e.g.*, serine, threonine) or unnatural (*e.g.*, 5-hydroxyproline or 5-hydroxylysine) amino acid. Exemplary O-linked saccharyl residues include N-acetylgalactosamine, galactose, mannose, GlcNAc, glucose, fucose or xylose.

**[0118]** The methods of the invention can be practiced on any peptide having an O-linked glycosylation site. For example, the methods are of use to produce O-linked glycoconjugates in which the glycosyl moiety is attached to an O-linked glycosylation site that is present in the wild type peptide. Accordingly, the present invention provides glycoconjugates of wild-type peptides that include an O-linked glycosylation site. Exemplary peptides according to this description include G-CSF, GM-CSF, IL-2 and interferon.

**[0119]** In exemplary embodiments the invention also provides novel mutant peptides that include one or more O-linked glycosylation sites that are not present in the corresponding wild-type peptide. In one embodiment the mutant polypeptide is a G-CSF polypeptide. In other exemplary embodiments the mutant polypeptide is an hGH polypeptide, an IFN alpha polypeptide or a GM-CSF polypeptide. Also provided are O-linked glycosylated versions of the mutant peptides, and methods of preparing O-linked glycosylated mutant peptides.

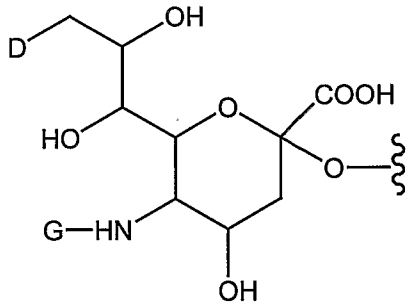
Additional methods include the elaboration, trimming back and/or modification of the O-linked glycosyl residue and glycosyl residues that are N-, rather than O-linked.

**[0120]** In an exemplary aspect, the invention provides a mutant peptide having the formula:



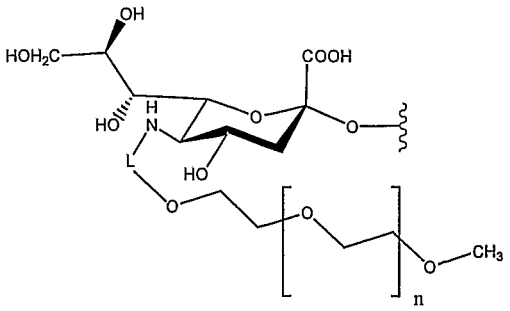

in which AA is an amino acid with a side chain that includes a hydroxyl moiety. Exemplary hydroxyamino acids are threonine and serine. The GalNAc moiety is linked to AA through the oxygen atom of the hydroxyl moiety. AA may be present in the wild type peptide or, alternatively, it is added or relocated by mutating the sequence of the wild type peptide. X is a modifying group, a saccharyl moiety, *e.g.*, sialyl, galactosyl and Gal-Sia groups, or a saccharyl moiety and a modifying group. In an exemplary embodiment, in which X is a saccharyl moiety, it includes a modifying group, as discussed herein. The glycosylated amino acid can be at the N- or C-peptide terminus or internal to the peptide sequence.

**[0121]** In an exemplary embodiment, X comprises a group selected from sialyl, galactosyl and Gal-Sia moieties, wherein at least one of said sialyl, galactosyl and Gal-Sia comprises a modifying group. In a further exemplary embodiment X comprises the moiety:

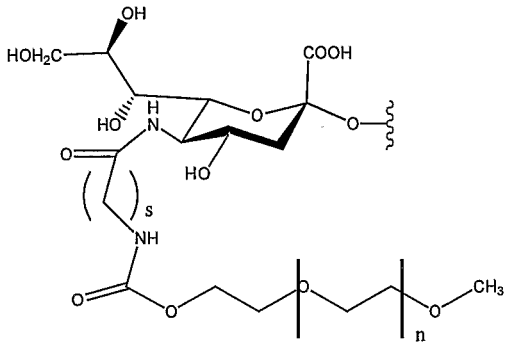



wherein D is a member selected from -OH and $R^1$-L-HN-;G is a member selected from $R^1$-L- and -C(O)($C_1$-$C_6$)alkyl; $R^1$ is a moiety comprising a member selected a moiety comprising a straight-chain or branched poly(ethylene glycol) residue; and L is a linker which is a member selected from a 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-.

**[0122]** In another exemplary embodiment X comprises the structure:



in which L is a substituted or unsubstituted alkyl or substituted or unsubstituted heteroalkyl group; and n is selected from the integers from 0 to about 2500. In yet another exemplary embodiment X comprises the structure:



in which s is selected from the integers from 0 to 20.

**[0123]** In another exemplary embodiment, AA is located within a proline-rich segment of the mutant peptide and/or it is proximate to a proline residue. Appropriate sequences forming O-linked glycosylation sites are readily determined by interrogating the enzymatic O-linked glycosylation of short peptides containing one or more putative O-linked glycosylation sites.

**[0124]** The conjugates of the invention are formed between peptides and diverse species such as water-soluble polymers, therapeutic moieties, diagnostic moieties, targeting moieties and the like. Also provided are conjugates that include two or more peptides linked together through a linker arm, *i.e.*, multifunctional conjugates; at least one peptide being O-glycosylated or including a mutant O-linked glycosylation site. The multi-functional conjugates of the invention can include two or more copies of the same peptide or a collection of diverse peptides with different structures, and/or properties. In exemplary conjugates according to this embodiment, the linker between the two peptides is attached to at least one of the peptides through an O-linked glycosyl residue, such as an O-linked glycosyl intact glycosyl linking group.

**[0125]** The conjugates of the invention are formed by the enzymatic attachment of a modified sugar to the glycosylated or unglycosylated peptide. The modified sugar is directly added to an O-linked glycosylation site, or to a glycosyl residue attached either directly or indirectly (e.g., through one or more glycosyl residue) to an O-linked glycosylation site. The invention also provides a conjugate of an O-linked glycosylated peptide in which a modified sugar is directly attached to an N-linked site, or to a glycosyl residue attached either directly or indirectly to an N-linked glycosylation site.

**[0126]** The modified sugar, when interposed between the peptide (or glycosyl residue) and the modifying group on the sugar becomes what is referred to herein 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 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 peptide backbone are also within the scope of the present invention.

**[0127]** 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 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.

**[0128]** The methods of the invention also provide conjugates of glycosylated and unglycosylated 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 to a peptide using an appropriate modified sugar can also be used to target a peptide to a particular tissue or cell surface receptor that is specific for the particular targeting agent. Moreover, there is provided a class of peptides that are specifically modified with a therapeutic moiety conjugated through a glycosyl linking group.

**O-Glycosylation**

**[0129]** The present invention provides O-linked glycosylated peptides, conjugates of these species and methods for forming O-linked glycosylated peptides that include a selected amino acid sequence ("an O-linked glycosylation site"). Of particular interest are mutant peptides that include an O-linked glycosylation site that is not present in the corresponding wild type peptide. The O-linked glycosylation site is a locus for attachment of a glycosyl residue that bears a modifying group.

**[0130]** Mucin-type O-linked glycosylation, one of the most abundant forms of protein glycosylation, is found on secreted and cell surface associated glycoproteins of all eukaryotic cells. There is great diversity in the structures created by O-linked glycosylation (hundreds of potential structures), which are produced by the catalytic activity of hundreds of glycosyltransferase enzymes that are resident in the Golgi complex. Diversity exists at the level of the glycan structure and in positions of attachment of O-glycans to protein backbones. Despite the high degree of potential diversity, it is clear that O-linked glycosylation is a highly regulated process that shows a high degree of conservation among multicellular organisms.

**[0131]** The first step in mucin-type O-linked glycosylation is catalysed by one or more members of a large family of UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases (GalNAc-transferases) (EC 2.4.1.41), which transfer GalNAc to serine and threonine acceptor sites (Hassan et al., *J. Biol. Chem.* **275**: 38197-38205 (2000)). To date twelve members of the mammalian GalNAc-transferase family have been identified and characterized (Schwientek et al., J. Biol. Chem. **277**: 22623-22638 (2002)), and several additional putative members of this gene family have been predicted from analysis of genome databases. The GalNAc-transferase isoforms have different kinetic properties and show differential expression patterns temporally and spatially, suggesting that they have distinct biological functions (Hassan et al., *J. Biol. Chem.* **275**: 38197-38205 (2000)). Sequence analysis of GalNAc-transferases have led to the hypothesis that these enzymes contain two distinct subunits: a central catalytic unit, and a C-terminal unit with sequence similarity to the plant

lectin ricin, designated the "lectin domain" (Hagen et al., J. Biol. Chem. **274**: 6797-6803 (1999); Hazes, *Protein Eng.* **10**: 1353-1356 (1997); Breton et al., *Curr. Opin. Struct. Biol.* **9**: 563-571 (1999)). Previous experiments involving site-specific mutagenesis of selected conserved residues confirmed that mutations in the catalytic domain eliminated catalytic activity. In contrast, mutations in the "lectin domain" had no significant effects on catalytic activity of the GalNAc-transferase isoform, GalNAc-T1 (Tenno *et al.*, *J. Biol. Chem.* 277(49): 47088-96 (2002)). Thus, the C-terminal "lectin domain" was believed not to be functional and not to play roles for the enzymatic functions of GalNAc-transferases (Hagen et al., J. Biol. Chem. **274**: 6797-6803 (1999)).

**[0132]** However, recent evidence demonstrates that some GalNAc-transferases exhibit unique activities with partially GalNAc-glycosylated glycopeptides. The catalytic actions of at least three GalNAc-transferase isoforms, GalNAc-T4, -T7, and –T10, selectively act on glycopeptides corresponding to mucin tandem repeat domains where only some of the clustered potential glycosylation sites have been GalNAc glycosylated by other GalNAc-transferases (Bennett et al., *FEBS Letters* **460**: 226-230 (1999); Ten Hagen et al., *J. Biol. Chem.* **276**: 17395-17404 (2001); Bennett et al., *J. Biol. Chem.* **273**: 30472-30481 (1998); Ten Hagen et al., *J. Biol. Chem.* **274**: 27867-27874 (1999)). GalNAc-T4 and -T7 recognize different GalNAc-glycosylated peptides and catalyse transfer of GalNAc to acceptor substrate sites in addition to those that were previously utilized. One of the functions of such GalNAc-transferase activities is predicted to represent a control step of the density of O-glycan occupancy in mucins and mucin-like glycoproteins with high density of O-linked glycosylation.

**[0133]** One example of this is the glycosylation of the cancer-associated mucin MUC1. MUC1 contains a tandem repeat O-linked glycosylated region of 20 residues (HGV**T**SAPDTRPAPG**S**TAPPA) with five potential O-linked glycosylation sites. GalNAc-T1, -T2, and -T3 can initiate glycosylation of the MUC1 tandem repeat and incorporate at only three sites (HGV<u>T</u>SAPDTRPAPG<u>ST</u>APPA, GalNAc attachment sites underlined). GalNAc-T4 is unique in that it is the only GalNAc-transferase isoform identified so far that can complete the O-linked glycan attachment to all five acceptor sites in the 20 amino acid tandem repeat sequence of the breast cancer associated mucin, MUC1. GalNAc-T4 transfers GalNAc to at least two sites not used by other GalNAc-transferase isoforms on the $GalNAc_4$TAP24 glycopeptide (<u>T</u>APPAHGV<u>T</u>SAPDTRPAPG<u>ST</u>APP, unique GalNAc-T4 attachment sites are in bold) (Bennett et al., *J. Biol. Chem.* **273**: 30472-

30481 (1998). An activity such as that exhibited by GalNAc-T4 appears to be required for production of the glycoform of MUC1 expressed by cancer cells where all potential sites are glycosylated (Muller et al., *J. Biol. Chem.* **274**: 18165-18172 (1999)). Normal MUC1 from lactating mammary glands has approximately 2.6 O-linked glycans per repeat (Muller et al., *J. Biol. Chem.* **272**: 24780-24793 (1997) and MUC1 derived from the cancer cell line T47D has 4.8 O-linked glycans per repeat (Muller et al., *J. Biol. Chem.* **274**: 18165-18172 (1999)). The cancer-associated form of MUC1 is therefore associated with higher density of O-linked glycan occupancy and this is accomplished by a GalNAc-transferase activity identical to or similar to that of GalNAc-T4.

**[0134]** Polypeptide GalNAc-transferases, which have not displayed apparent GalNAc-glycopeptide specificities, also appear to be modulated by their putative lectin domains (PCT WO 01/85215 A2). Recently, it was found that mutations in the GalNAc-T1 putative lectin domain, similarly to those previously analysed in GalNAc-T4 (Hassan *et al.*, *J. Biol. Chem.* **275**: 38197-38205 (2000)), modified the activity of the enzyme in a similar fashion as GalNAc-T4. Thus, while wild type GalNAc-T1 added multiple consecutive GalNAc residues to a peptide substrate with multiple acceptor sites, mutated GalNAc-T1 failed to add more than one GalNAc residue to the same substrate (Tenno *et al.*, *J. Biol. Chem.* 277(49): 47088-96 (2002)).

**[0135]** Since it has been demonstrated that mutations of GalNAc transferases can be utilized to produce glycosylation patterns that are distinct from those produced by the wild-type enzymes, it is within the scope of the present invention to utilize one or more mutant GalNAc transferase in preparing the O-linked glycosylated peptides of the invention.

**Mutant Peptides with O-linked Glycosylation Sites**

**[0136]** The peptides provided by the present invention include an amino acid sequence that is recognized as a GalNAc acceptor by one or more wild-type or mutant GalNac transferases. The amino acid sequence of the peptide is either the wild-type, for those peptides that include an O-linked glycosylation site, a mutant sequence in which a non-naturally ocurring O-linked glycosylation site is introduced, or a polypeptide comprising both naturally occuring and non-naturally occuring O-linked glycosylation sites. Exemplary peptides with which the present invention is practiced include granulocyte colony stimulating factor (G-CSF), *e.g.*, 175 and 178 amino acid wild types (with or without N-terminal methionine residues), interferon (*e.g.*, interferon alpha, e.g., interferon alpha 2b, or interferon alpha 2a), granulocyte macrophage

colony stimulating factor (GM-CSF), human growth hormone and interleukin (*e.g.*, interleukin 2). The emphasis of the following discussion on G-CSF, GM-CSF and IFN-α 2β is for clarity of illustration. Any number in the superscript of an amino acid indicates the amino acid position relative to the N-terminal methionine of the polypeptide. These numbers can be readily adjusted to reflect the absence of N-terminal methionine if the N-terminal of the polypeptide starts without a methionine. It is understood that the N-terminals of the exemplary peptides can start with or without a methionine. In addition, those of skill will understand that the strategy set forth herein for preparing O-linked glycoconjugated analogues of wild-type and mutant peptides is applicable to any peptide.

**[0137]** In an exemplary embodiment, the peptide is a biologically active G-CSF mutant that includes one or more mutation at a site selected from the N-terminus, adjacent to or encompassing $H^{53}$, $P^{61}$, $P^{129}$, $P^{133}$ and $P^{175}$. Biologically active G-CSF mutants of the present invention include any G-CSF polypeptide, in part or in whole, with one or more mutations that do not result in substantial or entire loss of its biological activity as it is measured by any suitable functional assays known to one skilled in the art. In one embodiment, mutations within the biologically active G-CSF mutants of the present invention are located within one or more O-linked glycosylation sites that do not naturally exist in wild type G-CSF. In another embodiment, mutations within the biologically active G-CSF mutants of the present invention reside within as well as outside of one or more O-linked glycosylation sites of the G-CSF mutants.

**[0138]** Representative wild type and mutant G-CSF polypeptides have sequences that are selected from:

**SEQ. ID NO. 1** (178 amino acid wild type)

mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklvseca tyklchpeel
vllghslgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp
elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr
aggvlvashl qsflevsyrv lrhlaqp;

**SEQ. ID NO. 2** (178 amino acid wild type without N-terminal methionine)

tplgpasslp qsfllkcleq vrkiqgdgaa lqeklvseca tyklchpeel
vllghslgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp

elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr
aggvlvashl qsflevsyrv lrhlaqp;

**SEQ. ID NO. 3** (175 amino acid wild type)

mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklca tyklchpeel
vllghslgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp
elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr
aggvlvashl qsflevsyrv lrhlaqp;

**SEQ. ID NO. 4** (175 amino acid wild type without N-terminal methionine)

mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklca tyklchpeel
vllghslgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp
elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr
aggvlvashl qsflevsyrv lrhlaqp;

**SEQ. ID NO. 5**

mvtplgpasslp qsfllkcleq vrkiqgdgaa lqeklca tyklchpeel
vllghslgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp
elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr
aggvlvashl qsflevsyrv lrhlaqp;

**SEQ. ID NO. 6**

mvtplgpasslp qsfllkcleq vrkiqgdgaa lqeklca tyklchpeel
vllghtlgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp
elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr
aggvlvashl qsflevsyrv lrhlaqp;

**SEQ. ID NO. 7**

mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklca tyklchpeel
vllghtlgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp
elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr
aggvlvashl qsflevsyrv lrhlaqp;

**SEQ. ID NO. 8**

mvtplgpasslp qsfllkcleq vrkiqgdgaa lqeklca tyklchpeel
vllgsslgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp

elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr aggvlvashl qsflevsyrv lrhlaqp;

**SEQ. ID NO. 9**

mqtplgpasslp qsfllkcleq vrkiqgdgaa lqeklca tyklchpeel vllghslgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr aggvlvashl qsflevsyrv lrhlaqp;

**SEQ. ID NO. 10**

mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklca tyklchpeel vllghslgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr aggvlvashl qsflevsyrv lrhlaqptqgamp; and

**SEQ. ID NO. 11**

mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklca tyklchpeel vllgsslgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr aggvlvashl qsflevsyrv lrhlaqp

**SEQ ID NO:12**

maitplgpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll ghslgipwap lsscpsqalq lagclsqlhs glflyqgllq alegispelg ptldtlqldv adfattiwqq meelgmapal qptqgampaf asafqrragg vlvashlqsf levsyrvlrh laqp

**SEQ ID NO:13**

mgvtetplgpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll ghslgipwap lsscpsqalq lagclsqlhs glflyqgllq alegispelg ptldtlqldv adfattiwqq meelgmapal qptqgampaf asafqrragg vlvashlqsf levsyrvlrh laqp

**SEQ ID NO:14**

maptplgpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll ghslgipwap lsscpsqalq lagclsqlhs glflyqgllq alegispelg

ptldtlqldv adfattiwqq meelgmapal qptqgampaf asafqrragg
vlvashlqsf levsyrvlrh laqp

**SEQ ID NO:15**

Mtptqglgpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll
ghslgipwap lsscpsqalq lagclsqlhs glflyqgllq alegispelg
ptldtlqldv adfattiwqq meelgmapal qptqgampaf asafqrragg
vlvashlqsf levsyrvlrh laqp

**SEQ ID NO:16**

mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll
ghslgipwap lsscpsqalq lagclsqlhs glflyqgllq alegispelg
ptldtlqldv adfattiwqq meelgmapatqptqgampaf asafqrragg
vlvashlqsf levsyrvlrh laqp

**SEQ ID NO:17**

Mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll
ghslgipftp lsscpsqalq lagclsqlhs glflyqgllq alegispelg
ptldtlqldv adfattiwqq meelgmapaL qptqgampaf asafqrragg
vlvashlqsf levsyrvlrh laqp

**SEQ ID NO:18**

mtplgpasslpqsfllkcleqvrkiqgdgaalqeklcatyklchpeelvllghslgi
pwaplsscpsqalqlagclsqlhsglflyqgllqalegispelgptldtlqldvadfa
ttiwqqmeelgmapalqptqtampafasafqrraggvlvashlqsflevsyrvlr
hlaqp.

**[0139]** In another exemplary embodiment, the peptide is a biologically active hGH mutant that includes one or more mutations at a site selected from the N-terminus or adjacent to or encompassing $P^{133}$. Biologically active hGH mutants of the present invention include any hGH polypeptide, in part or in whole, with one or more mutations that do not result in substantial or entire loss of its biological activity as it is measured by any suitable functional assays known to one skilled in the art. In one embodiment, mutations within the biologically active hGH mutants of the present invention are located within one or more O-linked glycosylation sites that do not naturally exist in wild type hGH. In another embodiment, mutations within the biologically active hGH mutants of the present invention reside within as well as outside of one or more O-linked glycosylation sites of the hGH mutants.

**[0140]** Representative wild type and mutant hGH polypeptides have sequences that are selected from:

**SEQ ID NO:19** (192 amino acid wild-type pituitary derived hGH comprising an N-terminal methionine)

mfptiplsrlfdnamlrahrlhqlafdtyqefeeayipkeqkysflqnpqtslcfse
siptpsnreetqqksnlellrislliqswlepvqflrsvfanslvygasdsnvydllk
dleegiqtlmgrledgsprtgqifkqtyskfdtnshnddallknygllycfrkdm
dkvetflrivqcrsvegscgf

**SEQ ID NO:20** (191 amino acid wild-type pituitary derived hGH lacking an N-Terminal methionine)

fptiplsrlfdnamlrahrlhqlafdtyqefeeayipkeqkysflqnpqtslcfsesi
ptpsnreetqqksnlellrislliqswlepvqflrsvfanslvygasdsnvydllkd
leegiqtlmgrledgsprtgqifkqtyskfdtnshnddallknygllycfrkdmd
kvetflrivqcrsvegscgf

**SEQ ID NO: 21** (wild type)

MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYI
PKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLE
LLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVY
DLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDT
NSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCR
SVEGSCGF

**[0141]** The following are representative mutant peptide sequences corresponding to the region underlined in the wild type SEQ ID NO:21: LEDGSP**T**TGQIFKQTYS, LEDGSP**TT**AQIFKQTYS, LEDGSP**TAT**QIFKQTYS, LEDGSP**TQGAM**FKQTYS, LEDGSP**TQ**GA**I**FKQTYS, LEDGSP**TQ**GQIFKQTYS, LEDGSP**TTLYV**FKQTYS, LEDGSP**TIN**TIFKQTYS, LEDGSP**TTVS**IFKQTYS, LEDGSPRTGQI**PT**QTYS, LEDGSPRTGQI**PT**QAYS, LEDGSP**TTL**QIFKQTYS, LE**TET**PRTGQIFKQTYS, L**VTET**PRTGQIFKQTYS, LE**TQ**SPRTGQIFKQTYS, L**VTQ**SPRTGQIFKQTYS, L**VTET**PATGQIFKQTYS, LEDGSP**TQGAMP**KQTYS, **and** LEDGSP**TTT**QIFKQTYS.

**[0142]** In another exemplary embodiment, the peptide is a biologically active IFN alpha mutant that includes one or more mutations at a site corresponding to $T^{106}$ of INF alpha 2, *e.g.,* adjacent to or encompassing an amino acid position in IFN alpha wild type, which

corresponds to or aligns with $T^{106}$ of INF alpha 2. Biologically active IFN alpha mutants of the present invention include any IFN alpha polypeptide, in part or in whole, with one or more mutations that do not result in substantial or entire loss of its biological activity as it is measured by any suitable functional assays known to one skilled in the art. In one embodiment, mutations within the biologically active IFN alpha mutants of the present invention are located within one or more O-linked glycosylation sites that do not naturally exist in wild type IFN alpha. In another embodiment, mutations within the biologically active IFN alpha mutants of the present invention reside within as well as outside of one or more O-linked glycosylation sites of the IFN alpha mutants.

**[0143]** A wild type and mutant IFN alpha polypeptide is shown below:

**SEQ ID NO:22** (from wild type IFN 2b)

$^{98}$CVIQGVGVTETPLMKEDSIL$^{117}$

**[0144]** Other appropriate O-linked glycosylation sequences for G-CSF and peptides other than G-CSF can be determined by preparing a polypeptide incorporating a putative O-linked glycosylation site and submitting that polypeptide to suitable O-linked glycosylation conditions, thereby confirming its ability to serve as an acceptor for a GalNac transferase. Moreover, as will be apparent to one of skill in the art, peptides that include one or more mutation are within the scope of the present invention. The mutations are designed to allow the adjustment of desirable properties of the peptides, e.g., activity and number and position of O- and/or N-linked glycosylation sites on the peptide.

**Acquisition of Peptide Coding Sequences**

***General Recombinant Technology***

**[0145]** This invention relies on routine techniques in the field of recombinant genetics. Basic texts disclosing the general methods of use in this invention include Sambrook and Russell, *Molecular Cloning, A Laboratory Manual* (3rd ed. 2001); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); and Ausubel *et al.*, eds., *Current Protocols in Molecular Biology* (1994).

**[0146]** For nucleic acids, sizes are given in either kilobases (kb) or base pairs (bp). These are estimates derived from agarose or acrylamide gel electrophoresis, from sequenced nucleic acids, or from published DNA sequences. For proteins, sizes are given in kilodaltons (kDa) or amino acid residue numbers. Proteins sizes are estimated from gel electrophoresis, from sequenced proteins, from derived amino acid sequences, or from published protein sequences.

**[0147]** Oligonucleotides that are not commercially available can be chemically synthesized, *e.g.*, according to the solid phase phosphoramidite triester method first described by Beaucage & Caruthers, *Tetrahedron Lett.* **22**: 1859-1862 (1981), using an automated synthesizer, as described in Van Devanter *et. al.*, *Nucleic Acids Res.* **12**: 6159-6168 (1984). Entire genes can also be chemically synthesized. Purification of oligonucleotides is performed using any art-recognized strategy, *e.g.*, native acrylamide gel electrophoresis or anion-exchange HPLC as described in Pearson & Reanier, *J. Chrom.* **255**: 137-149 (1983).

**[0148]** The sequence of the cloned wild-type peptide genes, polynucleotide encoding mutant peptides, and synthetic oligonucleotides can be verified after cloning using, *e.g.*, the chain termination method for sequencing double-stranded templates of Wallace *et al.*, *Gene* **16**: 21-26 (1981).

*Cloning and Subcloning of a Wild-Type Peptide Coding Sequence*

**[0149]** Numerous polynucleotide sequences encoding wild-type peptides have been determined and are available from a commercial supplier, e.g., human growth hormone, *e.g.*, GenBank Accession Nos. NM 000515, NM 002059, NM 022556, NM 022557, NM 022558, NM 022559, NM 022560, NM 022561, and NM 022562.

**[0150]** The rapid progress in the studies of human genome has made possible a cloning approach where a human DNA sequence database can be searched for any gene segment that has a certain percentage of sequence homology to a known nucleotide sequence, such as one encoding a previously identified peptide. Any DNA sequence so identified can be subsequently obtained by chemical synthesis and/or a polymerase chain reaction (PCR) technique such as overlap extension method. For a short sequence, completely *de novo* synthesis may be sufficient; whereas further isolation of full length coding sequence from a human cDNA or genomic library using a synthetic probe may be necessary to obtain a larger gene.

**[0151]** Alternatively, a nucleic acid sequence encoding a peptide can be isolated from a human cDNA or genomic DNA library using standard cloning techniques such as polymerase chain reaction (PCR), where homology-based primers can often be derived from a known nucleic acid sequence encoding a peptide. Most commonly used techniques for this purpose are described in standard texts, *e.g.*, Sambrook and Russell, *supra.*

**[0152]** cDNA libraries suitable for obtaining a coding sequence for a wild-type peptide may be commercially available or can be constructed. The general methods of isolating

mRNA, making cDNA by reverse transcription, ligating cDNA into a recombinant vector, transfecting into a recombinant host for propagation, screening, and cloning are well known (*see, e.g.*, Gubler and Hoffman, *Gene*, **25**: 263-269 (1983); Ausubel *et al.*, *supra*). Upon obtaining an amplified segment of nucleotide sequence by PCR, the segment can be further used as a probe to isolate the full-length polynucleotide sequence encoding the wild-type peptide from the cDNA library. A general description of appropriate procedures can be found in Sambrook and Russell, *supra*.

**[0153]** A similar procedure can be followed to obtain a full length sequence encoding a wild-type peptide, *e.g.*, any one of the GenBank Accession Nos mentioned above, from a human genomic library. Human genomic libraries are commercially available or can be constructed according to various art-recognized methods. In general, to construct a genomic library, the DNA is first extracted from an tissue where a peptide is likely found. The DNA is then either mechanically sheared or enzymatically digested to yield fragments of about 12-20 kb in length. The fragments are subsequently separated by gradient centrifugation from polynucleotide fragments of undesired sizes and are inserted in bacteriophage λ vectors. These vectors and phages are packaged *in vitro*. Recombinant phages are analyzed by plaque hybridization as described in Benton and Davis, *Science*, **196**: 180-182 (1977). Colony hybridization is carried out as described by Grunstein *et al.*, *Proc. Natl. Acad. Sci. USA*, **72**: 3961-3965 (1975).

**[0154]** Based on sequence homology, degenerate oligonucleotides can be designed as primer sets and PCR can be performed under suitable conditions (*see, e.g.*, White *et al.*, *PCR Protocols: Current Methods and Applications*, 1993; Griffin and Griffin, *PCR Technology*, CRC Press Inc. 1994) to amplify a segment of nucleotide sequence from a cDNA or genomic library. Using the amplified segment as a probe, the full-length nucleic acid encoding a wild-type peptide is obtained.

**[0155]** Upon acquiring a nucleic acid sequence encoding a wild-type peptide, the coding sequence can be subcloned into a vector, for instance, an expression vector, so that a recombinant wild-type peptide can be produced from the resulting construct. Further modifications to the wild-type peptide coding sequence, *e.g.*, nucleotide substitutions, may be subsequently made to alter the characteristics of the molecule.

*Introducing Mutations into a Peptide Sequence*

**[0156]** From an encoding polynucleotide sequence, the amino acid sequence of a wild-type peptide can be determined. Subsequently, this amino acid sequence may be modified to alter the protein's glycosylation pattern, by introducing additional glycosylation site(s) at various locations in the amino acid sequence.

**[0157]** Several types of protein glycosylation sites are well known in the art. For instance, in eukaryotes, N-linked glycosylation occurs on the asparagine of the consensus sequence Asn-$X_{aa}$-Ser/Thr, in which $X_{aa}$ is any amino acid except proline (Kornfeld et al., *Ann Rev Biochem* **54**:631-664 (1985); Kukuruzinska *et al.*, *Proc. Natl. Acad. Sci. USA* **84**:2145-2149 (1987); Herscovics *et al.*, *FASEB J* 7:540-550 (1993); and Orlean, *Saccharomyces* **Vol. 3** (1996)). O-linked glycosylation takes place at serine or threonine residues (Tanner *et al.*, *Biochim. Biophys. Acta.* **906**:81-91 (1987); and Hounsell *et al.*, *Glycoconj. J.* **13**:19-26 (1996)). Other glycosylation patterns are formed by linking glycosylphosphatidylinositol to the carboxyl-terminal carboxyl group of the protein (Takeda *et al.*, *Trends Biochem. Sci.* **20**:367-371 (1995); and Udenfriend *et al.*, *Ann. Rev. Biochem.* **64**:593-591 (1995). Based on this knowledge, suitable mutations can thus be introduced into a wild-type peptide sequence to form new glycosylation sites.

**[0158]** Although direct modification of an amino acid residue within a peptide polypeptide sequence may be suitable to introduce a new N-linked or O-linked glycosylation site, more frequently, introduction of a new glycosylation site is accomplished by mutating the polynucleotide sequence encoding a peptide. This can be achieved by using any of known mutagenesis methods, some of which are discussed below. Exemplary modifications to a G-CSF peptide include those illustrated in SEQ ID NO:5-18.

**[0159]** A variety of mutation-generating protocols are established and described in the art. *See, e.g.*, Zhang et al., *Proc. Natl. Acad. Sci. USA*, **94**: 4504-4509 (1997); and Stemmer, *Nature*, **370**: 389-391 (1994). The procedures can be used separately or in combination to produce variants of a set of nucleic acids, and hence variants of encoded polypeptides. Kits for mutagenesis, library construction, and other diversity-generating methods are commercially available.

**[0160]** Mutational methods of generating diversity include, for example, site-directed mutagenesis (Botstein and Shortle, *Science*, **229**: 1193-1201 (1985)), mutagenesis using uracil-containing templates (Kunkel, *Proc. Natl. Acad. Sci. USA*, **82**: 488-492 (1985)),

oligonucleotide-directed mutagenesis (Zoller and Smith, *Nucl. Acids Res.*, **10**: 6487-6500 (1982)), phosphorothioate-modified DNA mutagenesis (Taylor *et al.*, *Nucl. Acids Res.*, **13**: 8749-8764 and 8765-8787 (1985)), and mutagenesis using gapped duplex DNA (Kramer *et al.*, *Nucl. Acids Res.*, **12**: 9441-9456 (1984)).

**[0161]** Other methods for generating mutations include point mismatch repair (Kramer *et al.*, *Cell*, **38**: 879-887 (1984)), mutagenesis using repair-deficient host strains (Carter *et al.*, *Nucl. Acids Res.*, **13**: 4431-4443 (1985)), deletion mutagenesis (Eghtedarzadeh and Henikoff, *Nucl. Acids Res.*, **14**: 5115 (1986)), restriction-selection and restriction-purification (Wells *et al.*, *Phil. Trans. R. Soc. Lond. A*, **317**: 415-423 (1986)), mutagenesis by total gene synthesis (Nambiar *et al.*, *Science*, **223**: 1299-1301 (1984)), double-strand break repair (Mandecki, *Proc. Natl. Acad. Sci. USA*, **83**: 7177-7181 (1986)), mutagenesis by polynucleotide chain termination methods (U.S. Patent No. 5,965,408), and error-prone PCR (Leung *et al.*, *Biotechniques*, **1**: 11-15 (1989)).

*Modification of Nucleic Acids for Preferred Codon Usage in a Host Organism*

**[0162]** The polynucleotide sequence encoding a mutant peptide can be further altered to coincide with the preferred codon usage of a particular host. For example, the preferred codon usage of one strain of bacterial cells can be used to derive a polynucleotide that encodes a mutant peptide of the invention and includes the codons favored by this strain. The frequency of preferred codon usage exhibited by a host cell can be calculated by averaging frequency of preferred codon usage in a large number of genes expressed by the host cell (*e.g.*, calculation service is available from web site of the Kazusa DNA Research Institute, Japan). This analysis is preferably limited to genes that are highly expressed by the host cell. U.S. Patent No. 5,824,864, for example, provides the frequency of codon usage by highly expressed genes exhibited by dicotyledonous plants and monocotyledonous plants.

**[0163]** At the completion of modification, the mutant peptide coding sequences are verified by sequencing and are then subcloned into an appropriate expression vector for recombinant production in the same manner as the wild-type peptides.

***Expression and Purification of the Mutant Peptide***

**[0164]** Following sequence verification, the mutant peptide of the present invention can be produced using routine techniques in the field of recombinant genetics, relying on the polynucleotide sequences encoding the polypeptide disclosed herein.

*Expression Systems*

**[0165]** To obtain high-level expression of a nucleic acid encoding a mutant peptide of the present invention, one typically subclones a polynucleotide encoding the mutant peptide into an expression vector that contains a strong promoter to direct transcription, a transcription/translation terminator and a ribosome binding site for translational initiation. Suitable bacterial promoters are well known in the art and described, *e.g.*, in Sambrook and Russell, *supra,* and Ausubel *et al., supra.* Bacterial expression systems for expressing the wild-type or mutant peptide are available in, *e.g.*, *E. coli, Bacillus sp., Salmonella*, and *Caulobacter*. Kits for such expression systems are commercially available. Eukaryotic expression systems for mammalian cells, yeast, and insect cells are well known in the art and are also commercially available. In one embodiment, the eukaryotic expression vector is an adenoviral vector, an adeno-associated vector, or a retroviral vector.

**[0166]** The promoter used to direct expression of a heterologous nucleic acid depends on the particular application. The promoter is optionally positioned about the same distance from the heterologous transcription start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

**[0167]** In addition to the promoter, the expression vector typically includes a transcription unit or expression cassette that contains all the additional elements required for the expression of the mutant peptide in host cells. A typical expression cassette thus contains a promoter operably linked to the nucleic acid sequence encoding the mutant peptide and signals required for efficient polyadenylation of the transcript, ribosome binding sites, and translation termination. The nucleic acid sequence encoding the peptide is typically linked to a cleavable signal peptide sequence to promote secretion of the peptide by the transformed cell. Such signal peptides include, among others, the signal peptides from tissue plasminogen activator, insulin, and neuron growth factor, and juvenile hormone esterase of *Heliothis virescens*. Additional elements of the cassette may include enhancers and, if genomic DNA is used as the structural gene, introns with functional splice donor and acceptor sites.

**[0168]** In addition to a promoter sequence, the expression cassette should also contain a transcription termination region downstream of the structural gene to provide for efficient termination. The termination region may be obtained from the same gene as the promoter sequence or may be obtained from different genes.

**[0169]** The particular expression vector used to transport the genetic information into the cell is not particularly critical. Any of the conventional vectors used for expression in eukaryotic or prokaryotic cells may be used. Standard bacterial expression vectors include plasmids such as pBR322-based plasmids, pSKF, pET23D, and fusion expression systems such as GST and LacZ. Epitope tags can also be added to recombinant proteins to provide convenient methods of isolation, *e.g.*, c-myc.

**[0170]** Expression vectors containing regulatory elements from eukaryotic viruses are typically used in eukaryotic expression vectors, *e.g.*, SV40 vectors, papilloma virus vectors, and vectors derived from Epstein-Barr virus. Other exemplary eukaryotic vectors include pMSG, pAV009/$A^+$, pMTO10/$A^+$, pMAMneo-5, baculovirus pDSVE, and any other vector allowing expression of proteins under the direction of the SV40 early promoter, SV40 later promoter, metallothionein promoter, murine mammary tumor virus promoter, Rous sarcoma virus promoter, polyhedrin promoter, or other promoters shown effective for expression in eukaryotic cells.

**[0171]** In some exemplary embodiments the expression vector is chosen from pCWin1, pCWin2, pCWin2/MBP, pCWin2-MBP-SBD ($pMS_{39}$), and pCWin2-MBP-MCS-SBD ($pMXS_{39}$) as disclosed in co-owned U.S. Patent application filed April 9, 2004 which is incorporated herein by reference.

**[0172]** Some expression systems have markers that provide gene amplification such as thymidine kinase, hygromycin B phosphotransferase, and dihydrofolate reductase. Alternatively, high yield expression systems not involving gene amplification are also suitable, such as a baculovirus vector in insect cells, with a polynucleotide sequence encoding the mutant peptide under the direction of the polyhedrin promoter or other strong baculovirus promoters.

**[0173]** The elements that are typically included in expression vectors also include a replicon that functions in *E. coli*, a gene encoding antibiotic resistance to permit selection of bacteria that harbor recombinant plasmids, and unique restriction sites in nonessential regions of the plasmid to allow insertion of eukaryotic sequences. The particular antibiotic resistance gene chosen is not critical, any of the many resistance genes known in the art are suitable. The prokaryotic sequences are optionally chosen such that they do not interfere with the replication of the DNA in eukaryotic cells, if necessary.

**[0174]** When periplasmic expression of a recombinant protein (*e.g.*, a hgh mutant of the present invention) is desired, the expression vector further comprises a sequence encoding a secretion signal, such as the E. coli OppA (Periplasmic Oligopeptide Binding Protein) secretion signal or a modified version thereof, which is directly connected to 5' of the coding sequence of the protein to be expressed. This signal sequence directs the recombinant protein produced in cytoplasm through the cell membrane into the periplasmic space. The expression vector may further comprise a coding sequence for signal peptidase 1, which is capable of enzymatically cleaving the signal sequence when the recombinant protein is entering the periplasmic space. More detailed description for periplasmic production of a recombinant protein can be found in, *e.g.*, Gray *et al.*, *Gene* **39**: 247-254 (1985), U.S. Patent Nos. 6,160,089 and 6,436,674.

**[0175]** As discussed above, a person skilled in the art will recognize that various conservative substitutions can be made to any wild-type or mutant peptide or its coding sequence while still retaining the biological activity of the peptide. Moreover, modifications of a polynucleotide coding sequence may also be made to accommodate preferred codon usage in a particular expression host without altering the resulting amino acid sequence.

*Transfection Methods*

**[0176]** Standard transfection methods are used to produce bacterial, mammalian, yeast or insect cell lines that express large quantities of the mutant peptide, which are then purified using standard techniques (*see, e.g.*, Colley *et al.*, *J. Biol. Chem.* **264**: 17619-17622 (1989); *Guide to Protein Purification*, in *Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990)). Transformation of eukaryotic and prokaryotic cells are performed according to standard techniques (*see, e.g.*, Morrison, *J. Bact.* **132**: 349-351 (1977); Clark-Curtiss & Curtiss, *Methods in Enzymology* **101**: 347-362 (Wu *et al.*, eds, 1983).

**[0177]** Any of the well-known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, polybrene, protoplast fusion, electroporation, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA, or other foreign genetic material into a host cell (*see, e.g.*, Sambrook and Russell, *supra*). It is only necessary that the particular genetic engineering procedure used be capable of successfully introducing at least one gene into the host cell capable of expressing the mutant peptide.

*Detection of Expression of Mutant Peptide in Host Cells*

**[0178]** After the expression vector is introduced into appropriate host cells, the transfected cells are cultured under conditions favoring expression of the mutant peptide. The cells are then screened for the expression of the recombinant polypeptide, which is subsequently recovered from the culture using standard techniques (*see, e.g.*, Scopes, *Protein Purification: Principles and Practice* (1982); U.S. Patent No. 4,673,641; Ausubel *et al., supra*; and Sambrook and Russell, *supra*).

**[0179]** Several general methods for screening gene expression are well known among those skilled in the art. First, gene expression can be detected at the nucleic acid level. A variety of methods of specific DNA and RNA measurement using nucleic acid hybridization techniques are commonly used (*e.g.*, Sambrook and Russell, *supra*). Some methods involve an electrophoretic separation (*e.g.*, Southern blot for detecting DNA and Northern blot for detecting RNA), but detection of DNA or RNA can be carried out without electrophoresis as well (such as by dot blot). The presence of nucleic acid encoding a mutant peptide in transfected cells can also be detected by PCR or RT-PCR using sequence-specific primers.

**[0180]** Second, gene expression can be detected at the polypeptide level. Various immunological assays are routinely used by those skilled in the art to measure the level of a gene product, particularly using polyclonal or monoclonal antibodies that react specifically with a mutant peptide of the present invention, such as a polypeptide having the amino acid sequence of SEQ ID NO:1-7, (*e.g.*, Harlow and Lane, *Antibodies, A Laboratory Manual*, Chapter 14, Cold Spring Harbor, 1988; Kohler and Milstein, *Nature*, **256**: 495-497 (1975)). Such techniques require antibody preparation by selecting antibodies with high specificity against the mutant peptide or an antigenic portion thereof. The methods of raising polyclonal and monoclonal antibodies are well established and their descriptions can be found in the literature, *see, e.g.*, Harlow and Lane, *supra*; Kohler and Milstein, *Eur. J. Immunol.*, **6**: 511-519 (1976). More detailed descriptions of preparing antibody against the mutant peptide of the present invention and conducting immunological assays detecting the mutant peptide are provided in a later section.

*Purification of Recombinantly Produced Mutant Peptide*

**[0181]** Once the expression of a recombinant mutant peptide in transfected host cells is confirmed, the host cells are then cultured in an appropriate scale for the purpose of purifying the recombinant polypeptide.

*1. Purification of Recombinantly Produced Mutant Peptide from Bacteria*

**[0182]** When the mutant peptides of the present invention are produced recombinantly by transformed bacteria in large amounts, typically after promoter induction, although expression can be constitutive, the proteins may form insoluble aggregates. There are several protocols that are suitable for purification of protein inclusion bodies. For example, purification of aggregate proteins (hereinafter referred to as inclusion bodies) typically involves the extraction, separation and/or purification of inclusion bodies by disruption of bacterial cells, *e.g.*, by incubation in a buffer of about 100-150 μg/ml lysozyme and 0.1% Nonidet P40, a non-ionic detergent. The cell suspension can be ground using a Polytron grinder (Brinkman Instruments, Westbury, NY). Alternatively, the cells can be sonicated on ice. Alternate methods of lysing bacteria are described in Ausubel *et al.* and Sambrook and Russell, both *supra*, and will be apparent to those of skill in the art.

**[0183]** The cell suspension is generally centrifuged and the pellet containing the inclusion bodies resuspended in buffer which does not dissolve but washes the inclusion bodies, *e.g.*, 20 mM Tris-HCl (pH 7.2), 1 mM EDTA, 150 mM NaCl and 2% Triton-X 100, a non-ionic detergent. It may be necessary to repeat the wash step to remove as much cellular debris as possible. The remaining pellet of inclusion bodies may be resuspended in an appropriate buffer (*e.g.*, 20 mM sodium phosphate, pH 6.8, 150 mM NaCl). Other appropriate buffers will be apparent to those of skill in the art.

**[0184]** Following the washing step, the inclusion bodies are solubilized by the addition of a solvent that is both a strong hydrogen acceptor and a strong hydrogen donor (or a combination of solvents each having one of these properties). The proteins that formed the inclusion bodies may then be renatured by dilution or dialysis with a compatible buffer. Suitable solvents include, but are not limited to, urea (from about 4 M to about 8 M), formamide (at least about 80%, volume/volume basis), and guanidine hydrochloride (from about 4 M to about 8 M). Some solvents that are capable of solubilizing aggregate-forming proteins, such as SDS (sodium dodecyl sulfate) and 70% formic acid, may be inappropriate for use in this procedure due to the possibility of irreversible denaturation of the proteins, accompanied by a lack of immunogenicity and/or activity. Although guanidine hydrochloride and similar agents are denaturants, this denaturation is not irreversible and renaturation may occur upon removal (by dialysis, for example) or dilution of the denaturant, allowing re-formation of the immunologically and/or biologically active protein of interest. After solubilization, the protein can be separated from other bacterial proteins by standard

separation techniques. For further description of purifying recombinant peptide from bacterial inclusion body, *see, e.g.*, Patra *et al.*, *Protein Expression and Purification* **18**: 182-190 (2000).

**[0185]** Alternatively, it is possible to purify recombinant polypeptides, *e.g.*, a mutant peptide, from bacterial periplasm. Where the recombinant protein is exported into the periplasm of the bacteria, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to those of skill in the art (*see e.g.*, Ausubel *et al.*, *supra*). To isolate recombinant proteins from the periplasm, the bacterial cells are centrifuged to form a pellet. The pellet is resuspended in a buffer containing 20% sucrose. To lyse the cells, the bacteria are centrifuged and the pellet is resuspended in ice-cold 5 mM $MgSO_4$ and kept in an ice bath for approximately 10 minutes. The cell suspension is centrifuged and the supernatant decanted and saved. The recombinant proteins present in the supernatant can be separated from the host proteins by standard separation techniques well known to those of skill in the art.

*2. Standard Protein Separation Techniques for Purification*

**[0186]** When a recombinant polypeptide, *e.g.*, the mutant peptide of the present invention, is expressed in host cells in a soluble form, its purification can follow the standard protein purification procedure described below.

*i. Solubility Fractionation*

**[0187]** Often as an initial step, and if the protein mixture is complex, an initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest, *e.g.*, a mutant peptide of the present invention. The preferred salt is ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the amount of water in the protein mixture. Proteins then precipitate on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol is to add saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This will precipitate the most hydrophobic proteins. The precipitate is discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if necessary, through either dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold

ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

*ii. Size Differential Filtration*

**[0188]** Based on a calculated molecular weight, a protein of greater and lesser size can be isolated using ultrafiltration through membranes of different pore sizes (for example, Amicon or Millipore membranes). As a first step, the protein mixture is ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of a protein of interest, *e.g.*, a mutant peptide. The retentate of the ultrafiltration is then ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed as described below.

*iii. Column Chromatography*

**[0189]** The proteins of interest (such as the mutant peptide of the present invention) can also be separated from other proteins on the basis of their size, net surface charge, hydrophobicity, or affinity for ligands. In addition, antibodies raised against peptide can be conjugated to column matrices and the peptide immunopurified. All of these methods are well known in the art.

**[0190]** It will be apparent to one of skill that chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (*e.g.*, Pharmacia Biotech).

***Immunoassays for Detection of Mutant Peptide Expression***

**[0191]** To confirm the production of a recombinant mutant peptide, immunological assays may be useful to detect in a sample the expression of the polypeptide. Immunological assays are also useful for quantifying the expression level of the recombinant hormone. Antibodies against a mutant peptide are necessary for carrying out these immunological assays.

*Production of Antibodies against Mutant Peptide*

**[0192]** Methods for producing polyclonal and monoclonal antibodies that react specifically with an immunogen of interest are known to those of skill in the art (*see, e.g.,* Coligan, *Current Protocols in Immunology* Wiley/Greene, NY, 1991; Harlow and Lane, *Antibodies: A Laboratory Manual* Cold Spring Harbor Press, NY, 1989; Stites *et al.* (eds.) *Basic and Clinical Immunology* (4th ed.) Lange Medical Publications, Los Altos, CA, and references

cited therein; Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed.) Academic Press, New York, NY, 1986; and Kohler and Milstein *Nature* **256**: 495-497, 1975). Such techniques include antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors (*see*, Huse et al., *Science* **246**: 1275-1281, 1989; and Ward et al., *Nature* **341**: 544-546, 1989).

**[0193]** In order to produce antisera containing antibodies with desired specificity, the polypeptide of interest (*e.g.*, a mutant peptide of the present invention) or an antigenic fragment thereof can be used to immunize suitable animals, *e.g.*, mice, rabbits, or primates. A standard adjuvant, such as Freund's adjuvant, can be used in accordance with a standard immunization protocol. Alternatively, a synthetic antigenic peptide derived from that particular polypeptide can be conjugated to a carrier protein and subsequently used as an immunogen.

**[0194]** The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the antigen of interest. When appropriately high titers of antibody to the antigen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich antibodies specifically reactive to the antigen and purification of the antibodies can be performed subsequently, *see*, Harlow and Lane, *supra*, and the general descriptions of protein purification provided above.

**[0195]** Monoclonal antibodies are obtained using various techniques familiar to those of skill in the art. Typically, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (*see*, Kohler and Milstein, *Eur. J. Immunol.* **6**:511-519, 1976). Alternative methods of immortalization include, *e.g.*, transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and the yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host.

**[0196]** Additionally, monoclonal antibodies may also be recombinantly produced upon identification of nucleic acid sequences encoding an antibody with desired specificity or a binding fragment of such antibody by screening a human B cell cDNA library according to the general protocol outlined by Huse *et al.*, *supra*. The general principles and methods of

recombinant polypeptide production discussed above are applicable for antibody production by recombinant methods.

**[0197]** When desired, antibodies capable of specifically recognizing a mutant peptide of the present invention can be tested for their cross-reactivity against the wild-type peptide and thus distinguished from the antibodies against the wild-type protein. For instance, antisera obtained from an animal immunized with a mutant peptide can be run through a column on which a wild-type peptide is immobilized. The portion of the antisera that passes through the column recognizes only the mutant peptide and not the wild-type peptide. Similarly, monoclonal antibodies against a mutant peptide can also be screened for their exclusivity in recognizing only the mutant but not the wild-type peptide.

**[0198]** Polyclonal or monoclonal antibodies that specifically recognize only the mutant peptide of the present invention but not the wild-type peptide are useful for isolating the mutant protein from the wild-type protein, for example, by incubating a sample with a mutant peptide-specific polyclonal or monoclonal antibody immobilized on a solid support.

**Immunoassays for Detecting Mutant Peptide Expression**

**[0199]** Once antibodies specific for a mutant peptide of the present invention are available, the amount of the polypeptide in a sample, *e.g.*, a cell lysate, can be measured by a variety of immunoassay methods providing qualitative and quantitative results to a skilled artisan. For a review of immunological and immunoassay procedures in general *see*, *e.g.*, Stites, *supra;* U.S. Patent Nos. 4,366,241; 4,376,110; 4,517,288; and 4,837,168.

*Labeling in Immunoassays*

**[0200]** Immunoassays often utilize a labeling agent to specifically bind to and label the binding complex formed by the antibody and the target protein. The labeling agent may itself be one of the moieties comprising the antibody/target protein complex, or may be a third moiety, such as another antibody, that specifically binds to the antibody/target protein complex. A label may be detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Examples include, but are not limited to, magnetic beads (*e.g.*, Dynabeads$^{TM}$), fluorescent dyes (*e.g.*, fluorescein isothiocyanate, Texas red, rhodamine, and the like), radiolabels (*e.g.*, $^{3}H$, $^{125}I$, $^{35}S$, $^{14}C$, or $^{32}P$), enzymes (*e.g.*, horse radish peroxidase, alkaline phosphatase, and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic (*e.g.*, polystyrene, polypropylene, latex, *etc.*) beads.

**[0201]** In some cases, the labeling agent is a second antibody bearing a detectable label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second antibody can be modified with a detectable moiety, such as biotin, to which a third labeled molecule can specifically bind, such as enzyme-labeled streptavidin.

**[0202]** Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G, can also be used as the label agents. These proteins are normal constituents of the cell walls of streptococcal bacteria. They exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a variety of species (*see*, generally, Kronval, et al. *J. Immunol.*, **111**: 1401-1406 (1973); and Akerstrom, et al., *J. Immunol.*, **135**: 2589-2542 (1985)).

*Immunoassay Formats*

**[0203]** Immunoassays for detecting a target protein of interest (*e.g.*, a mutant human growth hormone) from samples may be either competitive or noncompetitive. Noncompetitive immunoassays are assays in which the amount of captured target protein is directly measured. In one preferred "sandwich" assay, for example, the antibody specific for the target protein can be bound directly to a solid substrate where the antibody is immobilized. It then captures the target protein in test samples. The antibody/target protein complex thus immobilized is then bound by a labeling agent, such as a second or third antibody bearing a label, as described above.

**[0204]** In competitive assays, the amount of target protein in a sample is measured indirectly by measuring the amount of an added (exogenous) target protein displaced (or competed away) from an antibody specific for the target protein by the target protein present in the sample. In a typical example of such an assay, the antibody is immobilized and the exogenous target protein is labeled. Since the amount of the exogenous target protein bound to the antibody is inversely proportional to the concentration of the target protein present in the sample, the target protein level in the sample can thus be determined based on the amount of exogenous target protein bound to the antibody and thus immobilized.

**[0205]** In some cases, western blot (immunoblot) analysis is used to detect and quantify the presence of a mutant peptide in the samples. The technique generally comprises separating sample proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support (such as a nitrocellulose filter, a nylon filter, or a

derivatized nylon filter) and incubating the samples with the antibodies that specifically bind the target protein. These antibodies may be directly labeled or alternatively may be subsequently detected using labeled antibodies (*e.g.*, labeled sheep anti-mouse antibodies) that specifically bind to the antibodies against a mutant peptide.

**[0206]** Other assay formats include liposome immunoassays (LIA), which use liposomes designed to bind specific molecules (*e.g.*, antibodies) and release encapsulated reagents or markers. The released chemicals are then detected according to standard techniques (*see*, Monroe *et al.*, *Amer. Clin. Prod. Rev.*, **5**: 34-41 (1986)).

**The Conjugates**

**[0207]** In a representative aspect, the present invention provides a glycoconjugate between a peptide and a selected modifying group, in which the modifying group is conjugated to the peptide through a glycosyl linking group, e.g., an intact glycosyl linking group. The glycosyl linking group is directly bound to an O-linked glycosylation site on the peptide or, alternatively, it is bound to an O-linked glycosylation site through one or more additional glycosyl residues. Methods of preparing the conjugates are set forth herein and in U.S. Patent No. 5,876,980; 6,030,815; 5,728,554; 5,922,577; WO 98/31826; US2003180835; and WO 03/031464.

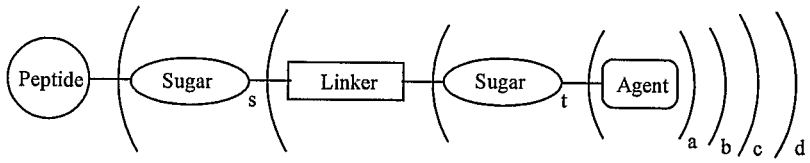
**[0208]** Exemplary peptides include an O-linked GalNAc residue that is bound to the O-linked glycosylation site through the action of a GalNAc transferase. The GalNAc itself may be the intact glycosyl linking group. The GalNAc may also be further elaborated by, for example, a Gal or Sia residue, either of which can act as the intact glycosyl linking group. In representative embodiments, the O-linked saccharyl residue is GalNAc-X , GalNAc-Gal-Sia-X , or GalNAc-Gal-Gal-Sia-X, in which X is a modifying group.

**[0209]** In an exemplary embodiment, the peptide is a mutant peptide that includes an O-linked glycosylation site not present in the wild-type peptide. The peptide is preferably O-glycosylated at the mutated site with a GalNAc residue. The discussion immediately preceding regarding the structure of the saccharyl moiety is relevant here as well.

**[0210]** The link between the peptide and the selected moiety includes an intact glycosyl linking group interposed between the peptide and the modifying 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 peptide. The saccharide component of the

modified sugar, when interposed between the 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.

**[0211]** 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 a therapeutic agent, a bioactive agent, a detectable lable, water-soluble moiety or the like. The "agent" can be a peptide, e.g, enzyme, antibody, anitgen, etc. 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." The identity of the peptide is without limitation.

**[0212]** In an exemplary embodiment, the selected moiety is a water-soluble polymer, e.g., PEG, m-PEG, PPG, m-PPG, etc. The water-soluble polymer is covalently attached to the peptide via a glycosyl linking group. The glycosyl linking group is covalently attached to either an amino acid residue or a glycosyl residue of the 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 (e.g., GalNAc) is attached to an amino acid residue (e.g., Thr or Ser).

**[0213]** In an exemplary embodiment, the protein is an interferon. The interferons are antiviral glycoproteins that, in humans, are secreted by human primary fibroblasts after induction with virus or double-stranded RNA. Interferons are of interest as therapeutics, e.g, antiviral agents (e.g., hepatitis B and C), antitumor agents (e.g., hepatocellular carcinoma) and in the treatment of multiple sclerosis. For references relevant to interferon-α, *see*, Asano, *et al.*, *Eur. J. Cancer*, **27(Suppl 4)**:S21-S25 (1991); Nagy, *et al.*, *Anticancer Research*, **8(3)**:467-470 (1988); Dron, *et al.*, *J. Biol. Regul. Homeost. Agents*, **3(1)**:13-19 (1989); Habib, *et al.*, *Am. Surg.*, **67(3)**:257-260 (3/2001); and Sugyiama, *et al.*, *Eur. J. Biochem.*, **217**:921-927 (1993). For references discussing intefereon-β, see, e.g., Yu, *et al.*, *J. Neuroimmunol.*, **64(1)**:91-100 (1996); Schmidt, J., *J. Neurosci. Res.*, **65(1)**:59-67 (2001); Wender, *et al.*, *Folia Neuropathol.*, **39(2)**:91-93 (2001); Martin, *et al.*, *Springer Semin. Immunopathol.*, **18(1)**:1-24

(1996); Takane, *et al.*, *J. Pharmacol. Exp. Ther.*, **294(2)**:746-752 (2000); Sburlati, *et al.*, *Biotechnol. Prog.*, **14**:189-192 (1998); Dodd, *et al.*, *Biochimica et Biophysica Acta*, **787**:183-187 (1984); Edelbaum, *et al.*, *J. Interferon Res.*, **12**:449-453 (1992); Conradt, *et al.*, *J. Biol. Chem.*, **262(30)**:14600-14605 (1987); Civas, *et al.*, *Eur. J. Biochem.*, **173**:311-316 (1988); Demolder, *et al.*, *J. Biotechnol.*, **32**:179-189 (1994); Sedmak, *et al.*, *J. Interferon Res.*, **9(Suppl 1)**:S61-S65 (1989); Kagawa, *et al.*, *J. Biol. Chem.*, **263(33)**:17508-17515 (1988); Hershenson, *et al.*, U.S. Patent No. 4,894,330; Jayaram, *et al.*, *J. Interferon Res.*, **3(2)**:177-180 (1983); Menge, *et al.*, *Develop. Biol. Standard.*, **66**:391-401 (1987); Vonk, *et al.*, *J. Interferon Res.*, **3(2)**:169-175 (1983); and Adolf, *et al.*, *J. Interferon Res.*, **10**:255-267 (1990).

**[0214]** In an exemplary interferon conjugate, interferon alpha, e.g., interferon alpha 2b and 2a, is conjugated to a water soluble polymer through an intact glycosyl linker.

**[0215]** In a further exemplary embodiment, the invention provides a conjugate of human granulocyte colony stimulating factor (G-CSF). G-CSF is a glycoprotein that stimulates proliferation, differentiation and activation of neutropoietic progenitor cells into functionally mature neutrophils. Injected G-CSF is rapidly cleared from the body. See, for example, Nohynek, et al., *Cancer Chemother. Pharmacol.*, 39:259-266 (1997); Lord, et al., *Clinical Cancer Research*, 7(7):2085-2090 (07/2001); Rotondaro, et al., *Molecular Biotechnology*, 11(2):117-128 (1999); and Bönig, et al., *Bone Marrow Transplantation*, 28: 259-264 (2001).

**[0216]** The present invention encompasses a method for the modification of GM-CSF. GM-CSF is well known in the art as a cytokine produced by activated T-cells, macrophages, endothelial cells, and stromal fibroblasts. GM-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. Further GM-CSF is a macrophage-activating factor and promotes the differentiation of Lagerhans cells into dendritic cells. Like G-CSF, GM-CSF also has clinical applications in bone marrow replacement following chemotherapy

**[0217]** In addition to providing conjugates that are formed through an enzymatically added intact 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 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 peptide through an intact glycosyl linking group. In another 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 of the peptide to which the glycosyl linking group is attached has the same structure.

**[0218]** Also provided is a peptide conjugate having a population of water-soluble polymer moieties covalently bound thereto through a glycosyl linking group. In another embodiment, essentially every member of the population of water soluble polymer moieties is bound to an amino acid residue of the peptide via an intact glycosyl linking group, and each amino acid residue having an intact glycosyl linking group attached thereto has the same structure.

**[0219]** The present invention also provides conjugates analogous to those described above in which the peptide is conjugated to a therapeutic moiety, diagnostic moiety, targeting moiety, toxin moiety or the like via a glycosyl linking group. Each of the above-recited moieties can be a small molecule, natural polymer (e.g., polypeptide) or synthetic polymer.

**[0220]** In a still further embodiment, the invention provides conjugates that localize selectively in a particular tissue due to the presence of a targeting agent as a component of the conjugate. In an exemplary embodiment, the targeting agent is a protein. Exemplary proteins include transferrin (brain, blood pool), HS-glycoprotein (bone, brain, blood pool), antibodies (brain, tissue with antibody-specific antigen, blood pool), coagulation factors V-XII (damaged tissue, clots, cancer, blood pool), serum proteins, e.g., α-acid glycoprotein, fetuin, α-fetal protein (brain, blood pool), β2-glycoprotein (liver, atherosclerosis plaques, brain, blood pool), G-CSF, GM-CSF, M-CSF, and EPO (immune stimulation, cancers, blood pool, red blood cell overproduction, neuroprotection), albumin (increase in half-life), IL-2 and IFN-α.

**[0221]** In an exemplary targeted conjugate, interferon alpha 2β (IFN-α 2β) is conjugated to transferrin via a bifunctional linker that includes an intact glycosyl linking group at each terminus of the PEG moiety (Scheme 1). For example, one terminus of the PEG linker is functionalized with an intact sialic acid linker that is attached to transferrin and the other is functionalized with an intact O-linked GalNAc linker that is attached to IFN-α 2β.

**[0222]** The conjugates of the invention can include 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. Also included within the invention are conjugates in which more than one selected moiety is attached to a peptide via a multivalent linking group.

**The Methods**

**[0223]** 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. Additionally, the invention provides methods for targeting conjugates of the invention to a particular tissue or region of the body.

**[0224]** Thus, the invention provides a method of forming a covalent conjugate between a selected moiety and a peptide. 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 peptide. The polymer, therapeutic moiety or biomolecule is conjugated to the 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 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 appropriate to form a covalent bond between the modified sugar and the peptide. The sugar moiety of the modified sugar is preferably selected from nucleotide sugars, activated sugars and sugars, which are neither nucleotides nor activated.

**[0225]** The acceptor peptide (O-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 peptide can be either a full-length protein or a fragment. Moreover, the peptide can be a wild type or mutated peptide. In an exemplary embodiment, the peptide includes a mutation that adds one or more N- or O-linked glycosylation sites to the peptide sequence.

**[0226]** In an exemplary embodiment, the peptide 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, GalNAc is added to a serine or threonine and the galactosylated peptide is sialylated with a sialic acid-modifying group cassette using ST6Gal-1. Alternatively, the 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.

**[0227]** 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.

**[0228]** In the conjugates of the invention, the Sia-modifying group cassette can be linked to the Gal in an α-2,6, or α-2,3 linkage.

**[0229]** 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.

**[0230]** The method of the invention also provides for modification of incompletely glycosylated 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.

**[0231]** 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.

**[0232]** Moreover, in addition to peptides, the methods of the present invention can be practiced with other biological structures (*e.g.*, glycolipids, lipids, sphingoids, ceramides, whole cells, and the like, containing an O-linked glycosylation site).

**[0233]** 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.

**[0234]** 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.

**[0235]** 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 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.

**[0236]** 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).

**[0237]** 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.

**[0238]** 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).

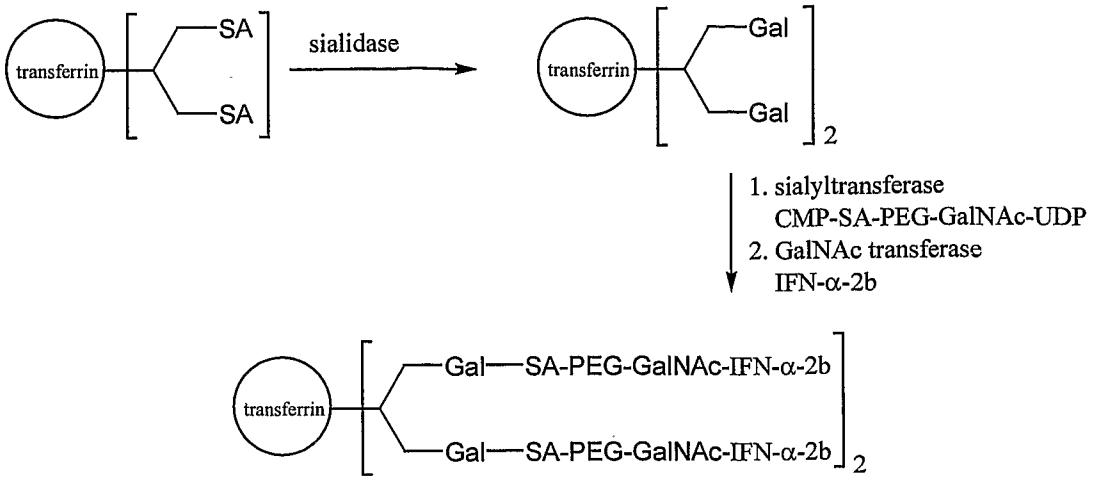
**[0239]** In one embodiment, the invention provides a method for linking two or more peptides through a linking group. The linking group is of any useful structure and may be selected from straight- and branched-chain structures. Preferably, each terminus of the linker, which is attached to a peptide, includes a modified sugar (i.e., a nascent intact glycosyl linking group).

**[0240]** In an exemplary method of the invention, two peptides are linked together via a linker moiety that includes a PEG linker. The construct conforms to the general structure set forth in the cartoon above. As described herein, the construct of the invention includes two intact glycosyl linking groups (i.e., s + t = 1). The focus on a PEG linker that includes two glycosyl groups is for purposes of clarity and should not be interpreted as limiting the identity of linker arms of use in this embodiment of the invention.

**[0241]** Thus, a PEG moiety is functionalized at a first terminus with a first glycosyl unit and at a second terminus with a second glycosyl unit. The first and second glycosyl units are preferably substrates for different transferases, allowing orthogonal attachment of the first and second peptides to the first and second glycosylunits, respectively. In practice, the $(glycosyl)^1$-PEG-$(glycosyl)^2$ linker is contacted with the first peptide and a first transferase for which the first glycosyl unit is a substrate, thereby forming $(peptide)^1$-$(glycosyl)^1$-PEG-$(glycosyl)^2$. Transferase and/or unreacted peptide is then optionally removed from the reaction mixture. The second peptide and a second transferase for which the second glycosyl unit is a substrate are added to the $(peptide)^1$-$(glycosyl)^1$-PEG-$(glycosyl)^2$ conjugate, forming $(peptide)^1$-$(glycosyl)^1$-PEG-$(glycosyl)^2$-$(peptide)^2$ ; at least one of the glycosyl residues is either directly or indirectly O-linked. Those of skill in the art will appreciate that the method outlined above is also applicable to forming conjugates between more than two peptides by, for example, the use of a branched PEG, dendrimer, poly(amino acid), polsaccharide or the like

**[0242]** In an exemplary embodiment, interferon alpha 2β (IFN-α 2β) is conjugated to transferrin via a bifunctional linker that includes an intact glycosyl linking group at each terminus of the PEG moiety (Scheme 1). The IFN conjugate has an *in vivo* half-life that is increased over that of IFN alone by virtue of the greater molecular sized of the conjugate. Moreover, the conjugation of IFN to transferrin serves to selectively target the conjugate to the brain. For example, one terminus of the PEG linker is functionalized with a CMP sialic acid and the other is functionalized with an UDP GalNAc. The linker is combined with IFN in the presence of a GalNAc transferase, resulting in the attachment of the GalNAc of the linker arm to a serine and/or threonine residue on the IFN.

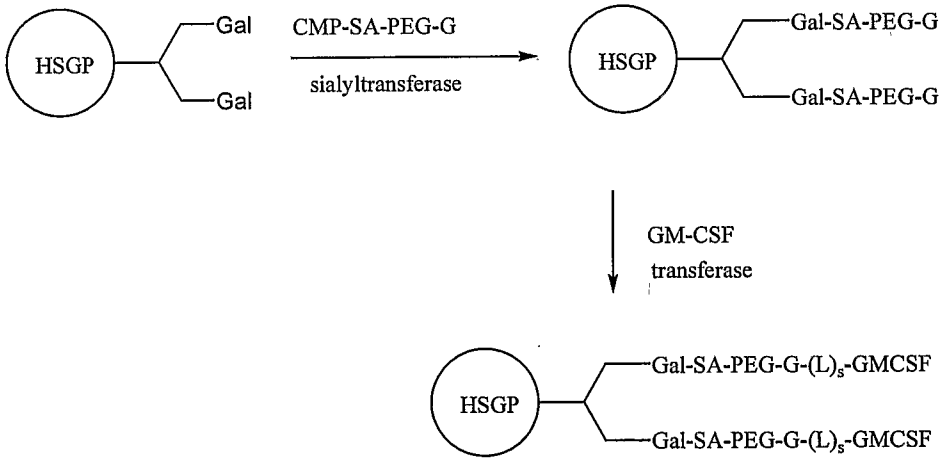
**Scheme 1**




**[0243]** The processes described above can be carried through as many cycles as desired, and is not limited to forming a conjugate between two peptides with a single linker. Moreover, those of skill in the art will appreciate that the reactions functionalizing the intact glycosyl linking groups at the termini of the PEG (or other) linker with the peptide can occur simultaneously in the same reaction vessel, or they can be carried out in a step-wise fashion. When the reactions are carried out in a step-wise manner, the conjugate produced at each step is optionally purified from one or more reaction components (e.g., enzymes, peptides).

**[0244]** A still further exemplary embodiment is set forth in Scheme 2. Scheme 2 shows a method of preparing a conjugate that targets a selected protein, e.g., GM-CSF, to bone and increases the circulatory half-life of the selected protein.

**Scheme 2**



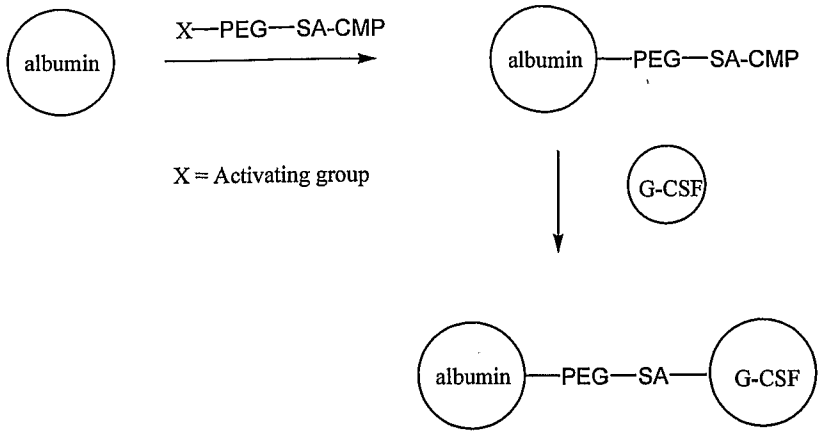
in which G is a glycosyl residue on an activated sugar moiety (e.g., sugar nucleotide), which is converted to an intact glycosyl linker group in the conjugate. When s is greater than 0, L is a saccharyl linking group such as GalNAc, or GalNAc-Gal.

**[0245]** The use of reactive derivatives of PEG (or other linkers) to attach one or more peptide moieties to the linker is within the scope of the present invention. The invention is not limited by the identity of the reactive PEG analogue. Many activated derivatives of poly(ethyleneglycol) are available commercially and in the literature. It is well within the abilities of one of skill to choose, and synthesize if necessary, an appropriate activated PEG derivative with which to prepare a substrate useful in the present invention. *See*, Abuchowski *et al. Cancer Biochem. Biophys.*, **7**: 175-186 (1984); Abuchowski *et al.*, *J. Biol. Chem.*, **252**: 3582-3586 (1977); Jackson *et al.*, *Anal. Biochem.*, **165**: 114-127 (1987); Koide *et al.*, *Biochem Biophys. Res. Commun.*, **111**: 659-667 (1983)), tresylate (Nilsson *et al.*, *Methods Enzymol.*, **104**: 56-69 (1984); Delgado *et al.*, *Biotechnol. Appl. Biochem.*, **12**: 119-128 (1990)); N-hydroxysuccinimide derived active esters (Buckmann *et al.*, *Makromol. Chem.*, **182**: 1379-1384 (1981); Joppich *et al.*, *Makromol. Chem.*, **180**: 1381-1384 (1979); Abuchowski *et al.*, *Cancer Biochem. Biophys.*, **7**: 175-186 (1984); Katre*et al. Proc. Natl. Acad. Sci. U.S.A.*, **84**: 1487-1491 (1987); Kitamura *et al.*, *Cancer Res.*, **51**: 4310-4315 (1991); Boccu *et al.*, *Z. Naturforsch.*, **38C**: 94-99 (1983), carbonates (Zalipsky *et al.*, POLY(ETHYLENE GLYCOL) CHEMISTRY: BIOTECHNICAL AND BIOMEDICAL APPLICATIONS, Harris, Ed., Plenum Press, New York, 1992, pp. 347-370; Zalipsky *et al.*, *Biotechnol. Appl. Biochem.*, **15**: 100-114 (1992); Veronese *et al.*, *Appl. Biochem. Biotech.*, **11**: 141-152 (1985)), imidazolyl formates (Beauchamp *et al.*, *Anal. Biochem.*, **131**: 25-33 (1983); Berger *et al.*, *Blood*, **71**: 1641-1647 (1988)), 4-dithiopyridines (Woghiren *et al.*, *Bioconjugate Chem.*, **4**: 314-318 (1993)), isocyanates (Byun *et al.*, *ASAIO Journal*, M649-M-653 (1992)) and epoxides (U.S. Pat. No. 4,806,595, issued to Noishiki *et al.*, (1989). Other linking groups include the urethane linkage between amino groups and activated PEG. *See*, Veronese, *et al.*, *Appl. Biochem. Biotechnol.*, **11**: 141-152 (1985).

**[0246]** In another exemplary embodiment in which a reactive PEG derivative is utilized, the invention provides a method for extending the blood-circulation half-life of a selected peptide, in essence targeting the peptide to the blood pool, by conjugating the peptide to a synthetic or natural polymer of a size sufficient to retard the filtration of the protein by the glomerulus (e.g., albumin). *See*, Scheme 3. This embodiment of the invention is illustrated

in Scheme in which G-CSF is conjugated to albumin via a PEG linker using a combination of chemical and enzymatic modification.

**Scheme 3**




**[0247]** Thus, as shown in Scheme 3, a residue (e.g., amino acid side chain) of albumin is modified with a reactive PEG derivative, such as X-PEG-(CMP-sialic acid), in which X is an activating group (e.g, active ester, isothiocyanate, etc). The PEG derivative and G-CSF are combined and contacted with a transferase for which CMP-sialic acid is a substrate. In a further illustrative embodiment, an ε-amine of lysine is reacted with the N-hydroxysuccinimide ester of the PEG-linker to form the albumin conjugate. The CMP-sialic acid of the linker is enzymatically conjugated to an appropriate residue on GCSF, e.g, Gal, or GalNAc thereby forming the conjugate. Those of skill will appreciate that the above-described method is not limited to the reaction partners set forth. Moreover, the method can be practiced to form conjugates that include more than two protein moieties by, for example, utilizing a branched linker having more than two termini.

**Modified Sugars**

**[0248]** Modified glycosyl donor species ("modified sugars") are preferably selected from modified sugar nucleotides, activated modified sugars and modified sugars that are simple saccharides that are neither nucleotides nor activated. Any desired carbohydrate structure can be added to a peptide using the methods of the invention. Typically, the structure will be a monosaccharide, but the present invention is not limited to the use of modified monosaccharide sugars; oligosaccharides and polysaccharides are useful as well.

**[0249]** The modifying group is attached to a sugar moiety by enzymatic means, chemical means or a combination thereof, thereby producing a modified sugar. The sugars are

substituted at any position that allows for the attachment of the modifying moiety, yet which still allows the sugar to function as a substrate for the enzyme used to ligate the modified sugar to the peptide. In another embodiment, when sialic acid is the sugar, the sialic acid is substituted with the modifying group at either the 9-position on the pyruvyl side chain or at the 5-position on the amine moiety that is normally acetylated in sialic acid.

**[0250]** In certain embodiments of the present invention, a modified sugar nucleotide is utilized to add the modified sugar to the peptide. Exemplary sugar nucleotides that are used in the present invention in their modified form include nucleotide mono-, di- or triphosphates or analogs thereof. In another embodiment, the modified sugar nucleotide is selected from a UDP-glycoside, CMP-glycoside, or a GDP-glycoside. Even more preferably, the modified sugar nucleotide is selected from an UDP-galactose, UDP-galactosamine, UDP-glucose, UDP-glucosamine, GDP-mannose, GDP-fucose, CMP-sialic acid, or CMP-NeuAc. N-acetylamine derivatives of the sugar nucletides are also of use in the method of the invention.

**[0251]** The invention also provides methods for synthesizing a modified peptide using a modified sugar, *e.g.*, modified-galactose, -fucose, -GalNAc and –sialic acid. When a modified sialic acid is used, either a sialyltransferase or a trans-sialidase (for α2,3-linked sialic acid only) can be used in these methods.

**[0252]** In other embodiments, the modified sugar is an activated sugar. Activated modified sugars, which are useful in the present invention are typically glycosides which have been synthetically altered to include an activated leaving group. As used herein, the term "activated leaving group" refers to those moieties, which are easily displaced in enzyme-regulated nucleophilic substitution reactions. Many activated sugars are known in the art. See, for example, Vocadlo et al., In CARBOHYDRATE CHEMISTRY AND BIOLOGY, Vol. 2, Ernst *et al.* Ed., Wiley-VCH Verlag: Weinheim, Germany, 2000; Kodama *et al.*, *Tetrahedron Lett.* **34**: 6419 (1993); Lougheed, *et al., J. Biol. Chem.* **274**: 37717 (1999)).

**[0253]** Examples of activating groups (leaving groups) include fluoro, chloro, bromo, tosylate ester, mesylate ester, triflate ester and the like. Preferred activated leaving groups, for use in the present invention, are those that do not significantly sterically encumber the enzymatic transfer of the glycoside to the acceptor. Accordingly, preferred embodiments of activated glycoside derivatives include glycosyl fluorides and glycosyl mesylates, with glycosyl fluorides being particularly preferred. Among the glycosyl fluorides, α-galactosyl fluoride, α-mannosyl fluoride, α-glucosyl fluoride, α-fucosyl fluoride, α-xylosyl fluoride, α-

sialyl fluoride, α-N-acetylglucosaminyl fluoride, α-N-acetylgalactosaminyl fluoride, β-galactosyl fluoride, β-mannosyl fluoride, β-glucosyl fluoride, β-fucosyl fluoride, β-xylosyl fluoride, β-sialyl fluoride, β-N-acetylglucosaminyl fluoride and β-N-acetylgalactosaminyl fluoride are most preferred.

**[0254]** By way of illustration, glycosyl fluorides can be prepared from the free sugar by first acetylating the sugar and then treating it with HF/pyridine. This generates the thermodynamically most stable anomer of the protected (acetylated) glycosyl fluoride (*i.e.*, the α-glycosyl fluoride). If the less stable anomer (*i.e.*, the β-glycosyl fluoride) is desired, it can be prepared by converting the peracetylated sugar with HBr/HOAc or with HCl to generate the anomeric bromide or chloride. This intermediate is reacted with a fluoride salt such as silver fluoride to generate the glycosyl fluoride. Acetylated glycosyl fluorides may be deprotected by reaction with mild (catalytic) base in methanol (*e.g.* NaOMe/MeOH). In addition, many glycosyl fluorides are commercially available.

**[0255]** Other activated glycosyl derivatives can be prepared using conventional methods known to those of skill in the art. For example, glycosyl mesylates can be prepared by treatment of the fully benzylated hemiacetal form of the sugar with mesyl chloride, followed by catalytic hydrogenation to remove the benzyl groups.

**[0256]** In a further exemplary embodiment, the modified sugar is an oligosaccharide having an antennary structure. In another embodiment, one or more of the termini of the antennae bear the modifying moiety. When more than one modifying moiety is attached to an oligosaccharide having an antennary structure, the oligosaccharide is useful to "amplify" the modifying moiety; each oligosaccharide unit conjugated to the peptide attaches multiple copies of the modifying group to the peptide. The general structure of a typical conjugate of the invention as set forth in the drawing above, encompasses multivalent species resulting from preparing a conjugate of the invention utilizing an antennary structure. Many antennary saccharide structures are known in the art, and the present method can be practiced with them without limitation.

**[0257]** Exemplary modifying groups are discussed below. The modifying groups can be selected for their ability to impart to a peptide one or more desirable property. Exemplary properties include, but are not limited to, enhanced pharmacokinetics, enhanced pharmacodynamics, improved biodistribution, providing a polyvalent species, improved water solubility, enhanced or diminished lipophilicity, and tissue targeting.

*Water-Soluble Polymers*

**[0258]** 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.

**[0259]** 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)).

**[0260]** 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."

**[0261]** 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, M*acronol. 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).

**[0262]** 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.

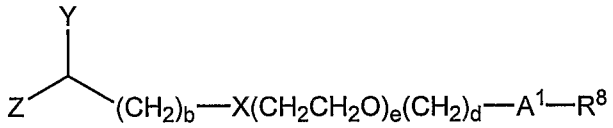
The active ester is used to form conjugates with a biologically active agent such as a protein or peptide.

**[0263]** 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.

**[0264]** 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.

**[0265]** 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.

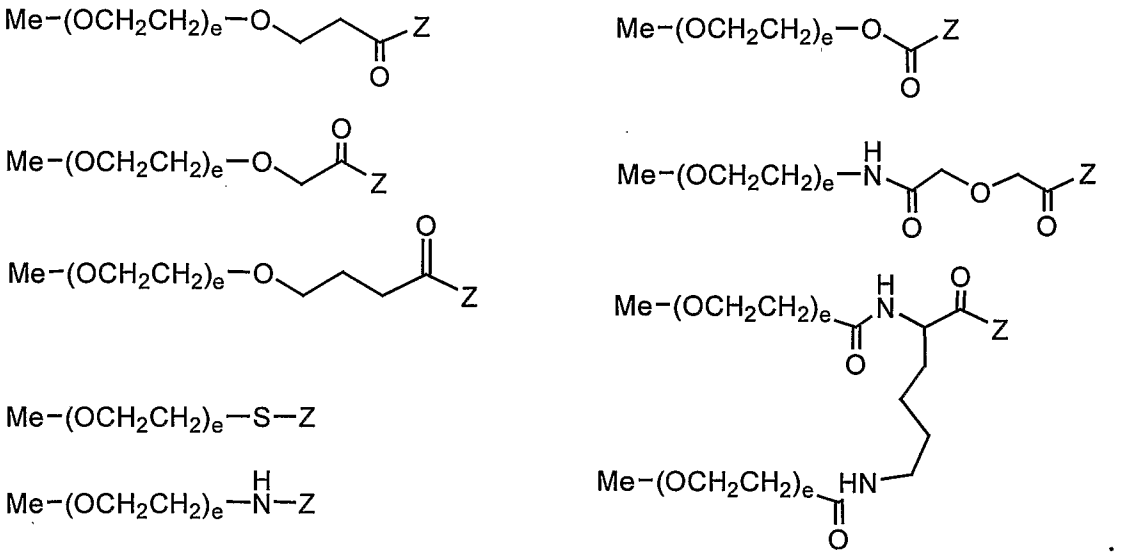
**[0266]** Exemplary poly(ethylene glycol) molecules of use in the invention include, but are not limited to, those having the formula:



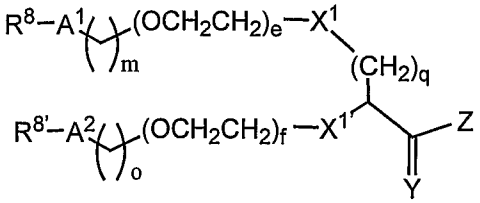

in which $R^8$ is H, OH, $NH_2$, substituted or unsubstituted alkyl, substituted or unsubstituted aryl, substituted or unsubstituted heteroaryl, substituted or unsubstituted heterocycloalkyl, substituted or unsubstituted heteroalkyl, *e.g.*, acetal, OHC-, $H_2N$-$(CH_2)_q$-, HS-$(CH_2)_q$, or -$(CH_2)_qC(Y)Z^1$. 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-$R^9$, the alcohol portion of activated esters; -$(CH_2)_pC(Y^1)V$, or -$(CH_2)_pU(CH_2)_sC(Y^1)_v$. The symbol Y represents H(2), =O, =S, =N-$R^{10}$. The symbols X, Y, $Y^1$, $A^1$, and U independently

represent the moieties O, S, N-$R^{11}$. The symbol V represents OH, $NH_2$, halogen, S-$R^{12}$, 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.

**[0267]** In other exemplary embodiments, the poly(ethylene glycol) molecule is selected from the following:

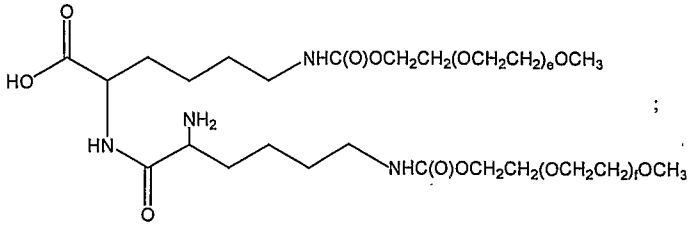



**[0268]** The poly(ethylene glycol) useful in forming the conjugate of the invention is either linear or branched. Branched poly(ethylene glycol) molecules suitable for use in the invention include, but are not limited to, those described by the following formula:

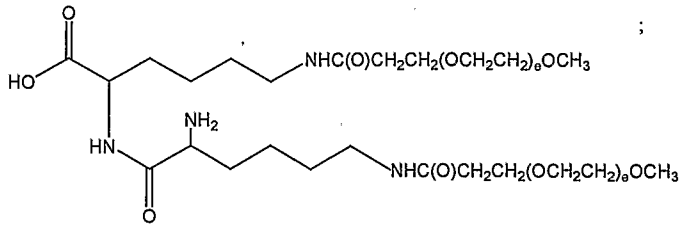



in which $R^8$ and $R^{8'}$ are members independently selected from the groups defined for $R^8$, above. $A^1$ and $A^2$ are members independently selected from the groups defined for $A^1$, above. The indices e, f, o, and q are as described above. Z and Y are as described above. $X^1$ and $X^{1'}$ are 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.

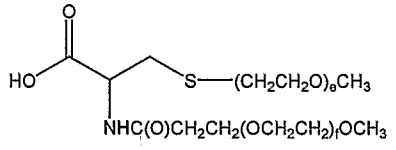
**[0269]** In other exemplary embodiments, the branched PEG is based upon a cysteine, serine or di-lysine core. Thus, further exemplary branched PEGs include:



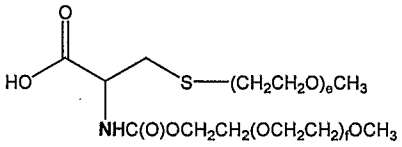

;

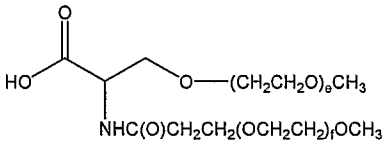



;

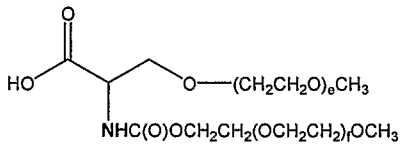


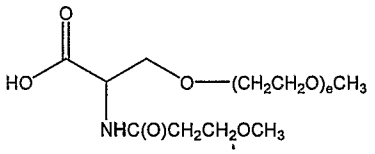

;



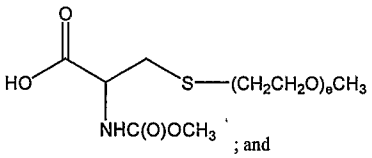




;

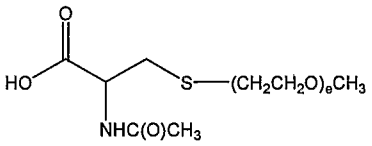


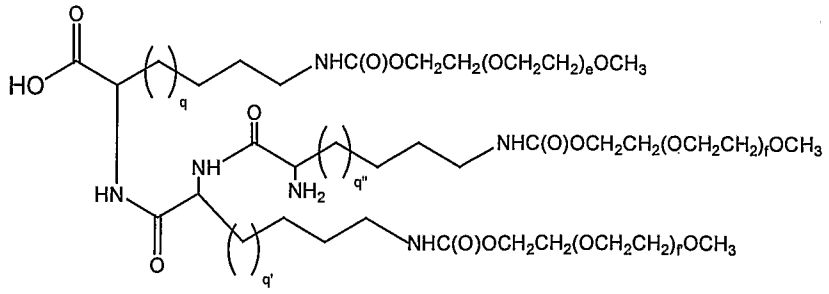

;



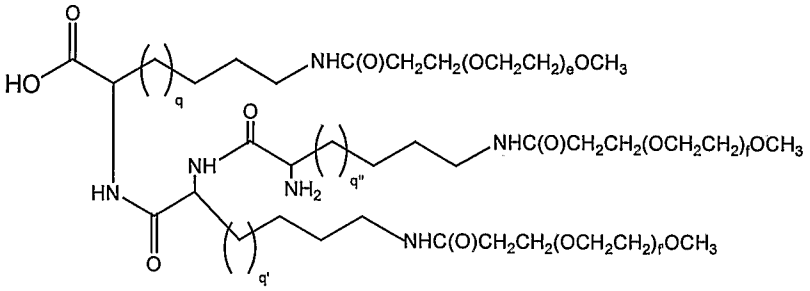

; and




**[0270]** 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

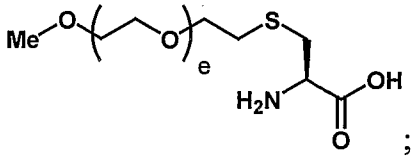



in which e, f and f' are independently selected integers from 1 to 2500; and q, q' and q" are independently selected integers from 1 to 20.

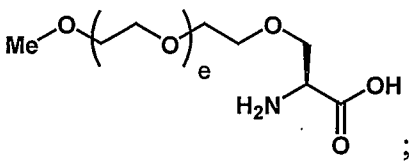
**[0271]** In exemplary embodiments of the invention, the PEG is m-PEG (5 kD, 10 kD, or 20kD). An exemplary branched PEG species is a serine- or cysteine-(m-PEG)$_2$ in which the m-PEG is a 20 kD m-PEG.

**[0272]** 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.

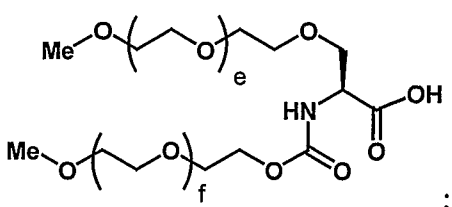
**[0273]** Specific embodiments according to the invention include:




;

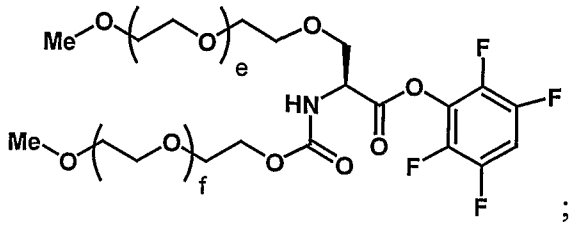



; and

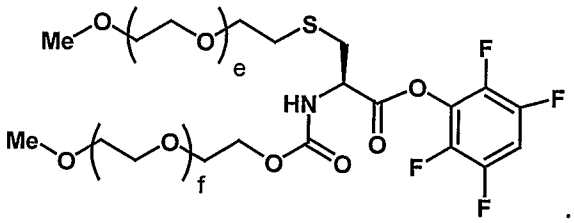



;

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

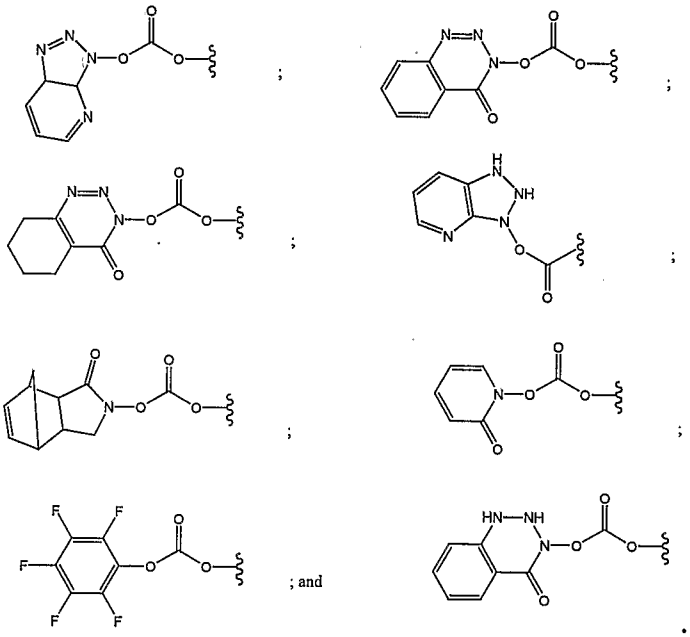


; and



.

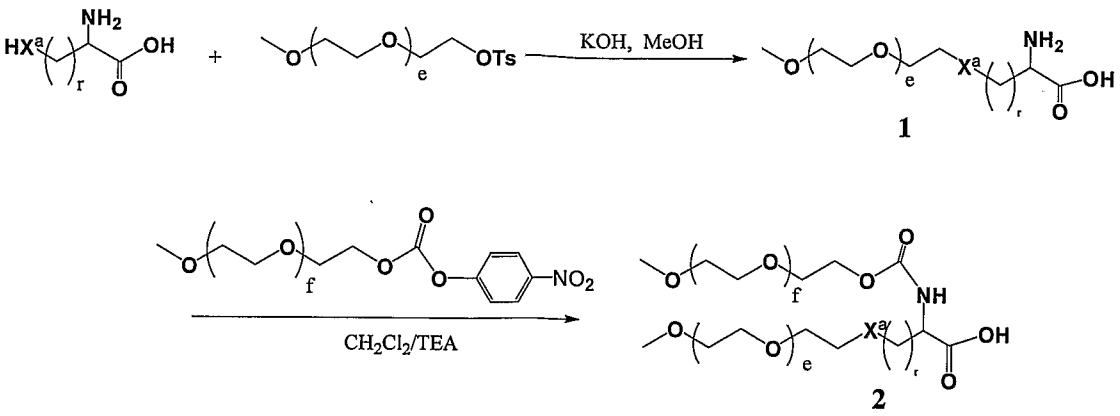
**[0274]** Other activating, or leaving groups, appropriate for activating linear PEGs of use in preparing the compounds set forth herein include, but are not limited to the species:



**[0275]** PEG molecules that are activated with these and other species and methods of making the activated PEGs are set forth in WO 04/083259.

**[0276]** 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.

**[0277]** 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.

**[0278]** 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.

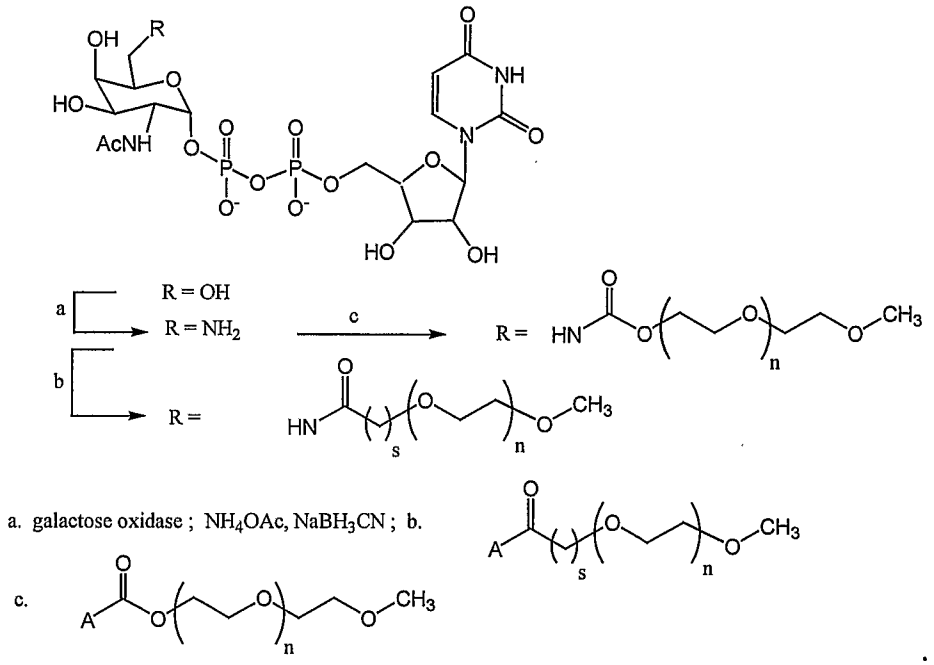
**[0279]** In an exemplary embodiment, the modifying group is a PEG moiety, however, any modifying group, *e.g.*, water-soluble polymer, water-insoluble polymer, therapeutic moiety, etc., can be incorporated in a glycosyl moiety through an appropriate linkage. The modified sugar is formed by enzymatic 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*

**[0280]** 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.

**[0281]** In one embodiment in which the saccharide core is galactose or glucose, $R^5$ is NHC(O)Y.

**[0282]** 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.

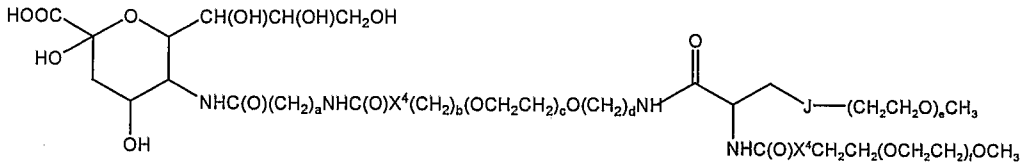



**[0283]** In the scheme above, the index n represents an integer from 1 to 2500, preferably from 10 to 1500, and more preferably from 10 to 1200. The symbol "A" represents an activating group, *e.g.*, a halo, a component of an activated ester (*e.g.*, a N-hydroxysuccinimide ester), a component of a carbonate (*e.g.*, p-nitrophenyl carbonate) and

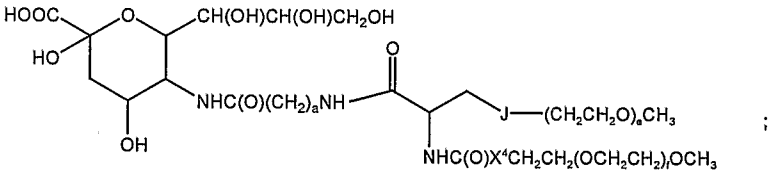
the like. Those of skill in the art will appreciate that other PEG-amide nucleotide sugars are readily prepared by this and analogous methods.

**[0284]** In other exemplary embodiments, the amide moiety is replaced by a group such as a urethane or a urea.

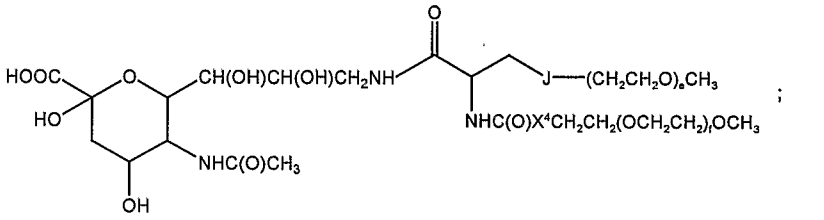
**[0285]** In still further embodiments, $R^1$ is a branched PEG, for example, one of those species set forth above. Illustrative compounds according to this embodiment include:



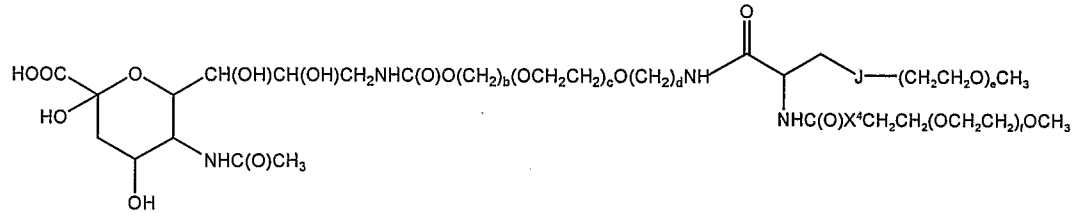

;



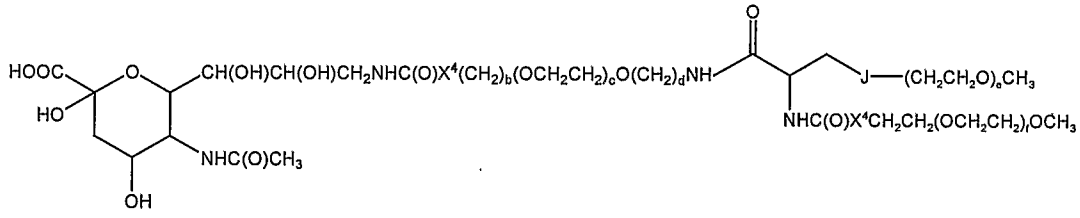

;




;

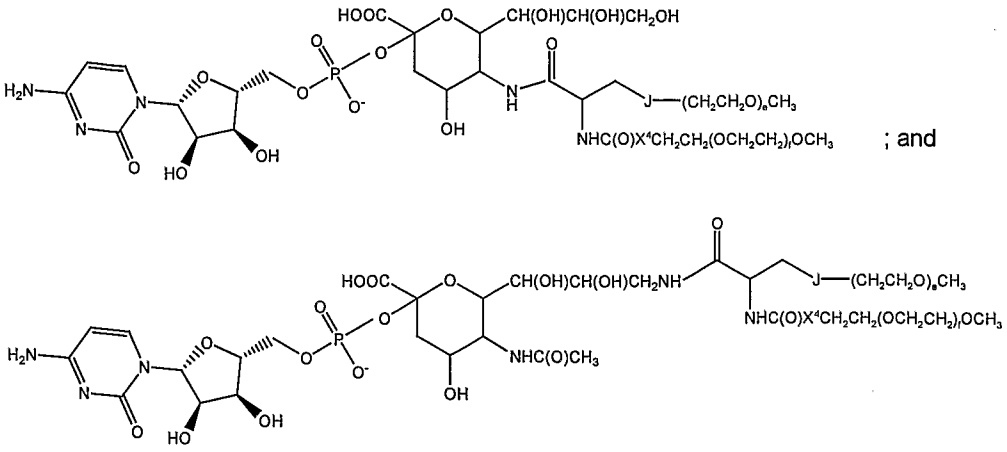



; and




in which $X^4$ is a bond or O, and J is S or O.

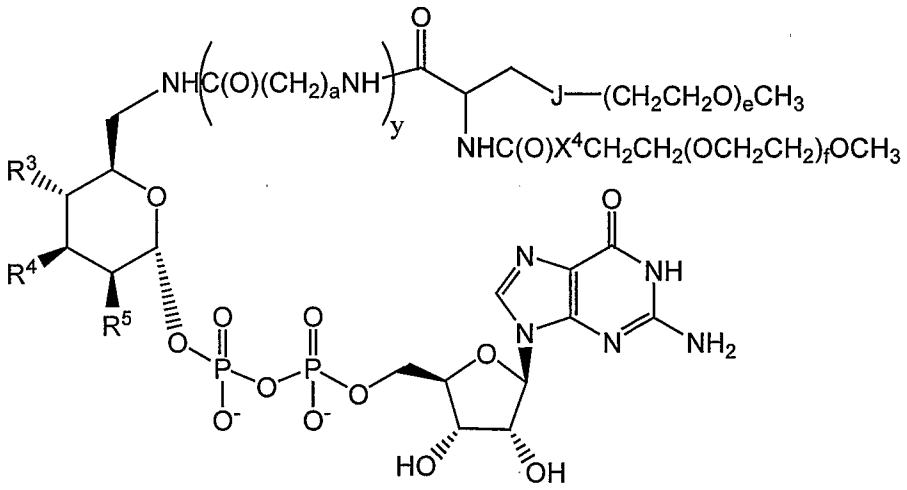
**[0286]** Moreover, as discussed above, the present invention provides peptide conjugates that are formed using nucleotide sugars that are modified with a water-soluble polymer, which is either straight-chain or branched. For example, compounds having the formula shown below are within the scope of the present invention:




; and

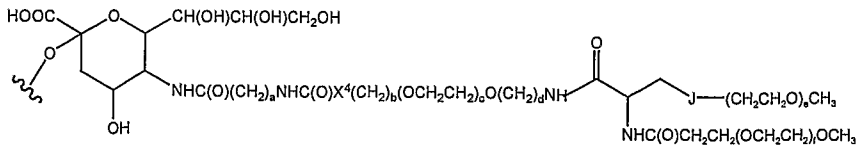
in which $X^4$ is O or a bond, and Jis S or O.

**[0287]** 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:

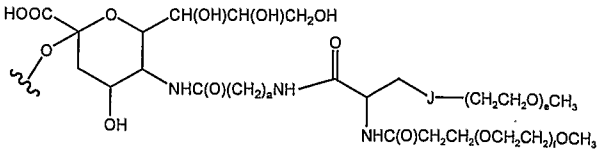



in which $X^4$ is a bond or O, J is S or O, and y is 0 or 1.

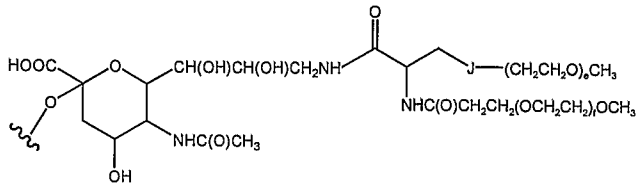
**[0288]** Also provided are conjugates of peptides and glycopeptides, lipids and glycolipids that include the compositions of the invention. For example, the invention provides conjugates having the following formulae:



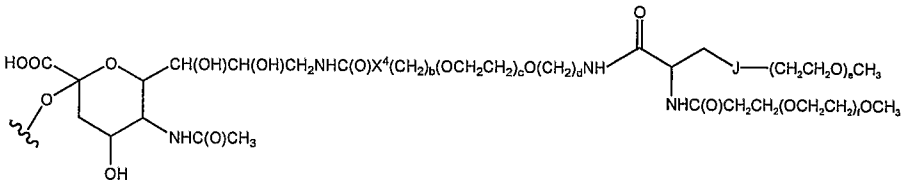
;



;



; and



wherein J s S or O.

***Water-insoluble polymers***

**[0289]** 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.

**[0290]** 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.

**[0291]** 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.

**[0292]** 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.

**[0293]** 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.

**[0294]** 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.

**[0295]** 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.

**[0296]** 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.

**[0297]** 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.

**[0298]** 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).

**[0299]** 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.

**[0300]** 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.

**[0301]** 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.

**[0302]** 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.

**[0303]** 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.

**[0304]** 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.

**[0305]** 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.

**[0306]** 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).

**[0307]** In another 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.

**[0308]** 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.

**[0309]** 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.

**[0310]** 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.

**[0311]** 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 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.

**[0312]** 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%.

***Biomolecules***

**[0313]** In another embodiment, the modified sugar bears a biomolecule. In still further embodiments, the biomolecule is a functional protein, enzyme, antigen, antibody, peptide, nucleic acid (*e.g.*, single nucleotides or nucleosides, oligonucleotides, polynucleotides and single- and higher-stranded nucleic acids), lectin, receptor or a combination thereof.

**[0314]** Preferred biomolecules are essentially non-fluorescent, or emit such a minimal amount of fluorescence that they are inappropriate for use as a fluorescent marker in an assay. Moreover, it is generally preferred to use biomolecules that are not sugars. An exception to this preference is the use of an otherwise naturally occurring sugar that is modified by covalent attachment of another entity (*e.g.*, PEG, biomolecule, therapeutic moiety, diagnostic moiety, *etc.*). In an exemplary embodiment, a sugar moiety, which is a biomolecule, is conjugated to a linker arm and the sugar-linker arm cassette is subsequently conjugated to a peptide via a method of the invention.

**[0315]** Biomolecules useful in practicing the present invention can be derived from any source. The biomolecules can be isolated from natural sources or they can be produced by synthetic methods. Peptides can be natural peptides or mutated peptides. Mutations can be effected by chemical mutagenesis, site-directed mutagenesis or other means of inducing mutations known to those of skill in the art. Peptides useful in practicing the instant invention include, for example, enzymes, antigens, antibodies and receptors. Antibodies can

be either polyclonal or monoclonal; either intact or fragments. The peptides are optionally the products of a program of directed evolution

**[0316]** Both naturally derived and synthetic peptides and nucleic acids are of use in conjunction with the present invention; these molecules can be attached to a sugar residue component or a crosslinking agent by any available reactive group. For example, peptides can be attached through a reactive amine, carboxyl, sulfhydryl, or hydroxyl group. The reactive group can reside at a peptide terminus or at a site internal to the peptide chain. Nucleic acids can be attached through a reactive group on a base (*e.g.*, exocyclic amine) or an available hydroxyl group on a sugar moiety (*e.g.*, 3'- or 5'-hydroxyl). The peptide and nucleic acid chains can be further derivatized at one or more sites to allow for the attachment of appropriate reactive groups onto the chain. *See,* Chrisey *et al. Nucleic Acids Res.* **24:** 3031-3039 (1996).

**[0317]** In a further embodiment, the biomolecule is selected to direct the peptide modified by the methods of the invention to a specific tissue, thereby enhancing the delivery of the peptide to that tissue relative to the amount of underivatized peptide that is delivered to the tissue. In a still further embodiment, the amount of derivatized peptide delivered to a specific tissue within a selected time period is enhanced by derivatization by at least about 20%, more preferably, at least about 40%, and more preferably still, at least about 100%. Presently, preferred biomolecules for targeting applications include antibodies, hormones and ligands for cell-surface receptors.

**[0318]** In still a further exemplary embodiment, there is provided as conjugate with biotin. Thus, for example, a selectively biotinylated peptide is elaborated by the attachment of an avidin or streptavidin moiety bearing one or more modifying groups.

***Therapeutic Moieties***

**[0319]** In another embodiment, the modified sugar includes a therapeutic moiety. Those of skill in the art will appreciate that there is overlap between the category of therapeutic moieties and biomolecules; many biomolecules have therapeutic properties or potential.

**[0320]** The therapeutic moieties can be agents already accepted for clinical use or they can be drugs whose use is experimental, or whose activity or mechanism of action is under investigation. The therapeutic moieties can have a proven action in a given disease state or can be only hypothesized to show desirable action in a given disease state. In another embodiment, the therapeutic moieties are compounds, which are being screened for their

ability to interact with a tissue of choice. Therapeutic moieties, which are useful in practicing the instant invention include drugs from a broad range of drug classes having a variety of pharmacological activities. Preferred therapeutic moieties are essentially non-fluorescent, or emit such a minimal amount of fluorescence that they are inappropriate for use as a fluorescent marker in an assay. Moreover, it is generally preferred to use therapeutic moieties that are not sugars. An exception to this preference is the use of a sugar that is modified by covalent attachment of another entity, such as a PEG, biomolecule, therapeutic moiety, diagnostic moiety and the like. In another exemplary embodiment, a therapeutic sugar moiety is conjugated to a linker arm and the sugar-linker arm cassette is subsequently conjugated to a peptide via a method of the invention.

**[0321]** Methods of conjugating therapeutic and diagnostic agents to various other species are well known to those of skill in the art. *See,* for example 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.

**[0322]** In an exemplary embodiment, the therapeutic moiety is attached to the modified sugar via a linkage that is cleaved under selected conditions. Exemplary conditions include, but are not limited to, a selected pH (e.g., stomach, intestine, endocytotic vacuole), the presence of an active enzyme (e.g, esterase, reductase, oxidase), light, heat and the like. Many cleavable groups are known in the art. *See*, for example, Jung *et al.*, *Biochem. Biophys. Acta,* **761**: 152-162 (1983); Joshi *et al.*, *J. Biol. Chem.*, **265**: 14518-14525 (1990); Zarling *et al.*, *J. Immunol.*, **124**: 913-920 (1980); Bouizar *et al.*, *Eur. J. Biochem.*, **155**: 141-147 (1986); Park *et al.*, *J. Biol. Chem.*, **261**: 205-210 (1986); Browning *et al.*, *J. Immunol.*, **143**: 1859-1867 (1989).

**[0323]** Classes of useful therapeutic moieties include, for example, non-steroidal anti-inflammatory drugs (NSAIDS). The NSAIDS can, for example, be selected from the following categories: (*e.g.*, propionic acid derivatives, acetic acid derivatives, fenamic acid derivatives, biphenylcarboxylic acid derivatives and oxicams); steroidal anti-inflammatory drugs including hydrocortisone and the like; antihistaminic drugs (*e.g.*, chlorpheniramine, triprolidine); antitussive drugs (*e.g.*, dextromethorphan, codeine, caramiphen and carbetapentane); antipruritic drugs (*e.g.*, methdilazine and trimeprazine); anticholinergic drugs (*e.g.*, scopolamine, atropine, homatropine, levodopa); anti-emetic and antinauseant drugs (*e.g.*, cyclizine, meclizine, chlorpromazine, buclizine); anorexic drugs (*e.g.*,

benzphetamine, phentermine, chlorphentermine, fenfluramine); central stimulant drugs (*e.g.*, amphetamine, methamphetamine, dextroamphetamine and methylphenidate); antiarrhythmic drugs (*e.g.*, propanolol, procainamide, disopyramide, quinidine, encainide); β-adrenergic blocker drugs (*e.g.*, metoprolol, acebutolol, betaxolol, labetalol and timolol); cardiotonic drugs (*e.g.*, milrinone, amrinone and dobutamine); antihypertensive drugs (*e.g.*, enalapril, clonidine, hydralazine, minoxidil, guanadrel, guanethidine);diuretic drugs (*e.g.*, amiloride and hydrochlorothiazide); vasodilator drugs (*e.g.*, diltiazem, amiodarone, isoxsuprine, nylidrin, tolazoline and verapamil); vasoconstrictor drugs (*e.g.*, dihydroergotamine, ergotamine and methylsergide); antiulcer drugs (*e.g.*, ranitidine and cimetidine); anesthetic drugs (*e.g.*, lidocaine, bupivacaine, chloroprocaine, dibucaine); antidepressant drugs (*e.g.*, imipramine, desipramine, amitryptiline, nortryptiline); tranquilizer and sedative drugs (*e.g.*, chlordiazepoxide, benacytyzine, benzquinamide, flurazepam, hydroxyzine, loxapine and promazine); antipsychotic drugs (*e.g.*, chlorprothixene, fluphenazine, haloperidol, molindone, thioridazine and trifluoperazine); antimicrobial drugs (antibacterial, antifungal, antiprotozoal and antiviral drugs).

**[0324]** Antimicrobial drugs which are preferred for incorporation into the present composition include, for example, pharmaceutically acceptable salts of β-lactam drugs, quinolone drugs, ciprofloxacin, norfloxacin, tetracycline, erythromycin, amikacin, triclosan, doxycycline, capreomycin, chlorhexidine, chlortetracycline, oxytetracycline, clindamycin, ethambutol, hexamidine isothionate, metronidazole, pentamidine, gentamycin, kanamycin, lineomycin, methacycline, methenamine, minocycline, neomycin, netilmycin, paromomycin, streptomycin, tobramycin, miconazole and amantadine.

**[0325]** Other drug moieties of use in practicing the present invention include antineoplastic drugs (*e.g.*, antiandrogens (*e.g.*, leuprolide or flutamide), cytocidal agents (*e.g.*, adriamycin, doxorubicin, taxol, cyclophosphamide, busulfan, cisplatin, β-2-interferon) anti-estrogens (*e.g.*, tamoxifen), antimetabolites (*e.g.*, fluorouracil, methotrexate, mercaptopurine, thioguanine). Also included within this class are radioisotope-based agents for both diagnosis and therapy, and conjugated toxins, such as ricin, geldanamycin, mytansin, CC-1065, the duocarmycins, Chlicheamycin and related structures and analogues thereof.

**[0326]** The therapeutic moiety can also be a hormone (*e.g.*, medroxyprogesterone, estradiol, leuprolide, megestrol, octreotide or somatostatin); muscle relaxant drugs (*e.g.*, cinnamedrine, cyclobenzaprine, flavoxate, orphenadrine, papaverine, mebeverine, idaverine,

ritodrine, diphenoxylate, dantrolene and azumolen); antispasmodic drugs; bone-active drugs (*e.g.*, diphosphonate and phosphonoalkylphosphinate drug compounds); endocrine modulating drugs (*e.g.*, contraceptives (*e.g.*, ethinodiol, ethinyl estradiol, norethindrone, mestranol, desogestrel, medroxyprogesterone), modulators of diabetes (*e.g.*, glyburide or chlorpropamide), anabolics, such as testolactone or stanozolol, androgens (*e.g.*, methyltestosterone, testosterone or fluoxymesterone), antidiuretics (*e.g.*, desmopressin) and calcitonins).

**[0327]** Also of use in the present invention are estrogens (*e.g.*, diethylstilbesterol), glucocorticoids (*e.g.*, triamcinolone, betamethasone, etc.) and progestogens, such as norethindrone, ethynodiol, norethindrone, levonorgestrel; thyroid agents (*e.g.*, liothyronine or levothyroxine) or anti-thyroid agents (*e.g.*, methimazole); antihyperprolactinemic drugs (*e.g.*, cabergoline); hormone suppressors (*e.g.*, danazol or goserelin), oxytocics (*e.g.*, methylergonovine or oxytocin) and prostaglandins, such as mioprostol, alprostadil or dinoprostone, can also be employed.

**[0328]** Other useful modifying groups include immunomodulating drugs (*e.g.*, antihistamines, mast cell stabilizers, such as lodoxamide and/or cromolyn, steroids (*e.g.*, triamcinolone, beclomethazone, cortisone, dexamethasone, prednisolone, methylprednisolone, beclomethasone, or clobetasol), histamine H2 antagonists (*e.g.*, famotidine, cimetidine, ranitidine), immunosuppressants (*e.g.*, azathioprine, cyclosporin), etc. Groups with anti-inflammatory activity, such as sulindac, etodolac, ketoprofen and ketorolac, are also of use. Other drugs of use in conjunction with the present invention will be apparent to those of skill in the art.

***Preparation of Modified Sugars***

**[0329]** In general, the sugar moiety and the modifying group 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.

**[0330]** 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.

**[0331]** 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.

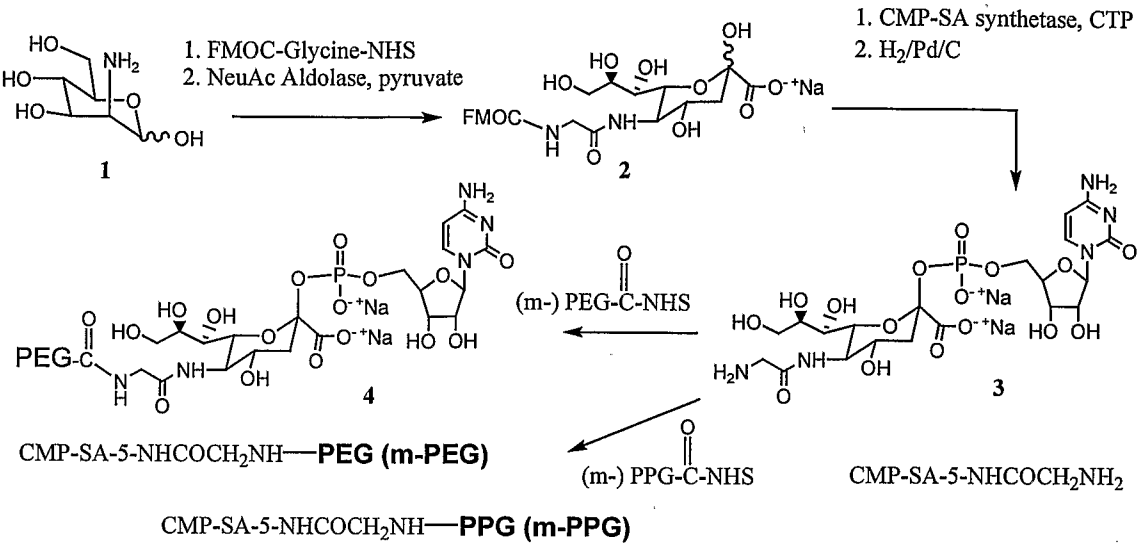
**[0332]** In the discussion that follows, a number of specific examples of modified sugars that are useful in practicing the present invention are set forth. In the exemplary embodiments, a sialic acid derivative is utilized as the sugar nucleus to which the modifying group is attached. The focus of 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)).

**[0333]** In an exemplary embodiment, the peptide that is modified by a method of the invention is a glycopeptide that is produced in prokaryotic cells (*e.g.*, *E. coli*), eukaryotic cells including yeast and 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 glyco-PEG-ylated, glyco-PPG-ylated or otherwise modified with a modified sialic acid.

**[0334]** 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

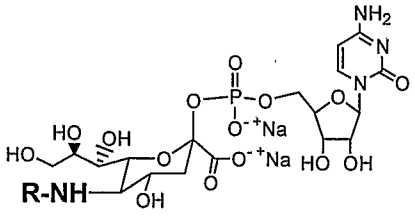
adduct is utilized as a locus of PEG or PPG attachment by reacting compound **3** with an activated (m-) PEG or (m-) PPG derivative (*e.g.*, PEG-C(O)NHS, PPG-C(O)NHS), producing **4** or **5**, respectively.

**Scheme 4**

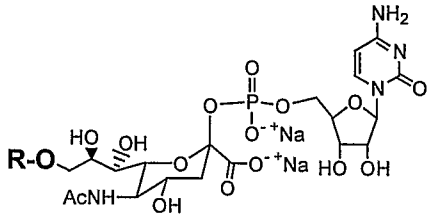



**[0335]** Table 2 sets forth representative examples of sugar monophosphates that are derivatized with a PEG or PPG moiety. Certain of the compounds of Table 2 are prepared by the method of Scheme 4. Other derivatives are prepared by art-recognized methods. *See*, for example, Keppler *et al.*, *Glycobiology* **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.

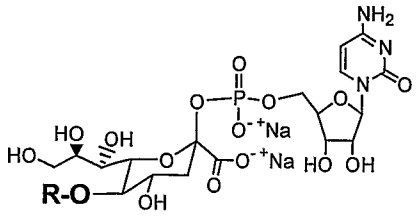
**Table 2**



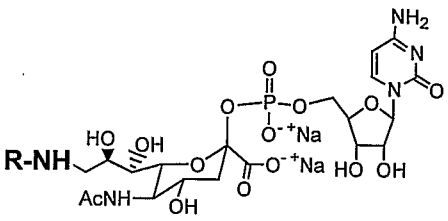

CMP-SA-5-NH-R



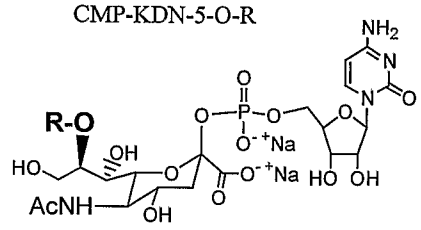

CMP-NeuAc-9-O-R



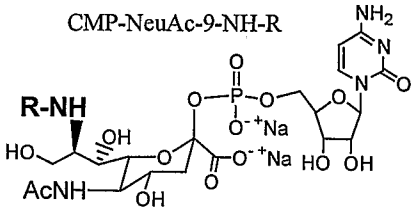
CMP-KDN-5-O-R



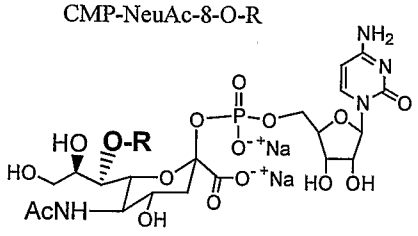
CMP-NeuAc-9-NH-R



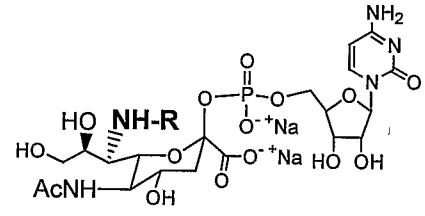
CMP-NeuAc-8-O-R



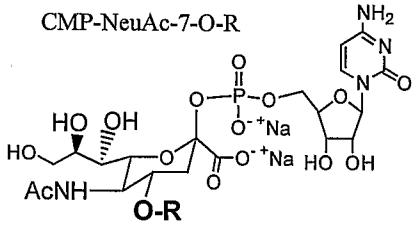
CMP-NeuAc-8-NH-R



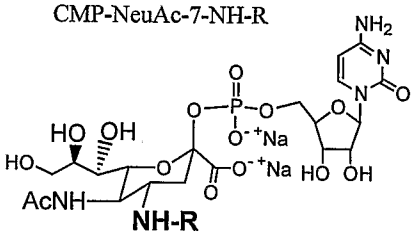
CMP-NeuAc-7-O-R



CMP-NeuAc-7-NH-R

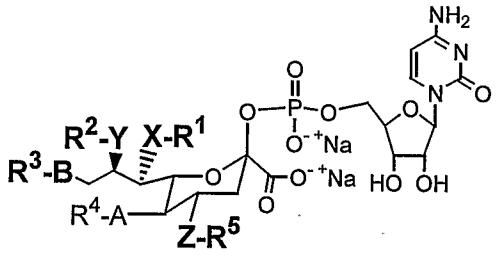


CMP-NeuAc-4-O-R



CMP-NeuAc-4-NH-R

**[0336]** 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 Formula I:



(I)

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 water-soluble polymer, therapeutic moiety, biomolecule or other moiety. Alternatively, these symbols represent a linker that is bound to a water-soluble polymer, therapeutic moiety, biomolecule or other moiety.

**[0337]** Exemplary moieties attached to the conjugates disclosed herein include, but are not limited to, PEG derivatives (*e.g.*, alkyl-PEG, acyl-PEG, acyl-alkyl-PEG, alkyl-acyl-PEG carbamoyl-PEG, aryl-PEG), PPG derivatives (*e.g.*, alkyl-PPG, 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).

***Cross-linking Groups***

**[0338]** Preparation of the modified sugar for use in the methods of the present invention includes attachment of a modifying group to a sugar residue and forming a stable adduct, which is a substrate for a glycosyltransferase. The sugar and modifying group can be coupled by a zero- or higher-order cross-linking agent. 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.

**[0339]** An exemplary strategy involves incorporation of a protected sulfhydryl onto the sugar using the heterobifunctional crosslinker SPDP (n-succinimidyl-3-(2-pyridyldithio)propionate and then deprotecting the sulfhydryl for formation of a disulfide bond with another sulfhydryl on the modifying group.

**[0340]** If SPDP detrimentally affects the ability of the modified sugar to act as a glycosyltransferase substrate, one of an array of other crosslinkers such as 2-iminothiolane or N-succinimidyl S-acetylthioacetate (SATA) is used to form a disulfide bond. 2-iminothiolane reacts with primary amines, instantly incorporating an unprotected sulfhydryl onto the amine-containing molecule. SATA also reacts with primary amines, but incorporates a protected sulfhydryl, which is later deacetaylated using hydroxylamine to produce a free sulfhydryl. In each case, the incorporated sulfhydryl is free to react with other sulfhydryls or protected sulfhydryl, like SPDP, forming the required disulfide bond.

**[0341]** The above-described strategy is exemplary, and not limiting, of linkers of use in the invention. Other crosslinkers are available that can be used in different strategies for crosslinking the modifying group to the peptide. For example, TPCH(S-(2-thiopyridyl)-L-cysteine hydrazide and TPMPH ((S-(2-thiopyridyl) mercapto-propionohydrazide) react with carbohydrate moieties that have been previously oxidized by mild periodate treatment, thus forming a hydrazone bond between the hydrazide portion of the crosslinker and the periodate generated aldehydes. TPCH and TPMPH introduce a 2-pyridylthione protected sulfhydryl group onto the sugar, which can be deprotected with DTT and then subsequently used for conjugation, such as forming disulfide bonds between components.

**[0342]** If disulfide bonding is found unsuitable for producing stable modified sugars, other crosslinkers may be used that incorporate more stable bonds between components. The heterobifunctional crosslinkers GMBS (N-gama-malimidobutyryloxy)succinimide) and SMCC (succinimidyl 4-(N-maleimido-methyl)cyclohexane) react with primary amines, thus introducing a maleimide group onto the component. The maleimide group can subsequently react with sulfhydryls on the other component, which can be introduced by previously mentioned crosslinkers, thus forming a stable thioether bond between the components. If steric hindrance between components interferes with either component's activity or the ability of the modified sugar to act as a glycosyltransferase substrate, crosslinkers can be used which

introduce long spacer arms between components and include derivatives of some of the previously mentioned crosslinkers (*i.e.*, SPDP). Thus, there is an abundance of suitable crosslinkers, which are useful; each of which is selected depending on the effects it has on optimal peptide conjugate and modified sugar production.

**[0343]** 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.

*i. Preferred Specific Sites in Crosslinking Reagents*

*1. Amino-Reactive Groups*

**[0344]** In one embodiment, the sites on the cross-linker are amino-reactive groups. Useful non-limiting examples of amino-reactive groups include N-hydroxysuccinimide (NHS) esters, imidoesters, isocyanates, acylhalides, arylazides, p-nitrophenyl esters, aldehydes, and sulfonyl chlorides.

**[0345]** NHS esters react preferentially with the primary (including aromatic) amino groups of a modified sugar component. The imidazole groups of histidines are known to compete

with primary amines for reaction, but the reaction products are unstable and readily hydrolyzed. The reaction involves the nucleophilic attack of an amine on the acid carboxyl of an NHS ester to form an amide, releasing the N-hydroxysuccinimide. Thus, the positive charge of the original amino group is lost.

**[0346]** Imidoesters are the most specific acylating reagents for reaction with the amine groups of the modified sugar components. At a pH between 7 and 10, imidoesters react only with primary amines. Primary amines attack imidates nucleophilically to produce an intermediate that breaks down to amidine at high pH or to a new imidate at low pH. The new imidate can react with another primary amine, thus crosslinking two amino groups, a case of a putatively monofunctional imidate reacting bifunctionally. The principal product of reaction with primary amines is an amidine that is a stronger base than the original amine. The positive charge of the original amino group is therefore retained.

**[0347]** Isocyanates (and isothiocyanates) react with the primary amines of the modified sugar components to form stable bonds. Their reactions with sulfhydryl, imidazole, and tyrosyl groups give relatively unstable products.

**[0348]** Acylazides are also used as amino-specific reagents in which nucleophilic amines of the affinity component attack acidic carboxyl groups under slightly alkaline conditions, *e.g.* pH 8.5.

**[0349]** Arylhalides such as 1,5-difluoro-2,4-dinitrobenzene react preferentially with the amino groups and tyrosine phenolic groups of modified sugar components, but also with sulfhydryl and imidazole groups.

**[0350]** p-Nitrophenyl esters of mono- and dicarboxylic acids are also useful amino-reactive groups. Although the reagent specificity is not very high, α- and ε-amino groups appear to react most rapidly.

**[0351]** Aldehydes such as glutaraldehyde react with primary amines of modified sugar. Although unstable Schiff bases are formed upon reaction of the amino groups with the aldehydes of the aldehydes, glutaraldehyde is capable of modifying the modified sugar with stable crosslinks. At pH 6-8, the pH of typical crosslinking conditions, the cyclic polymers undergo a dehydration to form α-β unsaturated aldehyde polymers. Schiff bases, however, are stable, when conjugated to another double bond. The resonant interaction of both double bonds prevents hydrolysis of the Schiff linkage. Furthermore, amines at high local

concentrations can attack the ethylenic double bond to form a stable Michael addition product.

**[0352]** Aromatic sulfonyl chlorides react with a variety of sites of the modified sugar components, but reaction with the amino groups is the most important, resulting in a stable sulfonamide linkage.

*2. Sulfhydryl-Reactive Groups*

**[0353]** In another embodiment, the sites are sulfhydryl-reactive groups. Useful, non-limiting examples of sulfhydryl-reactive groups include maleimides, alkyl halides, pyridyl disulfides, and thiophthalimides.

**[0354]** Maleimides react preferentially with the sulfhydryl group of the modified sugar components to form stable thioether bonds. They also react at a much slower rate with primary amino groups and the imidazole groups of histidines. However, at pH 7 the maleimide group can be considered a sulfhydryl-specific group, since at this pH the reaction rate of simple thiols is 1000-fold greater than that of the corresponding amine.

**[0355]** Alkyl halides react with sulfhydryl groups, sulfides, imidazoles, and amino groups. At neutral to slightly alkaline pH, however, alkyl halides react primarily with sulfhydryl groups to form stable thioether bonds. At higher pH, reaction with amino groups is favored.

**[0356]** Pyridyl disulfides react with free sulfhydryls via disulfide exchange to give mixed disulfides. As a result, pyridyl disulfides are the most specific sulfhydryl-reactive groups.

**[0357]** Thiophthalimides react with free sulfhydryl groups to form disulfides.

*3. Carboxyl-Reactive Residue*

**[0358]** In another embodiment, carbodiimides soluble in both water and organic solvent, are used as carboxyl-reactive reagents. These compounds react with free carboxyl groups forming a pseudourea that can then couple to available amines yielding an amide linkage teach how to modify a carboxyl group with carbodiimde (Yamada *et al., Biochemistry* **20**: 4836-4842, 1981).

*ii. Preferred Nonspecific Sites in Crosslinking Reagents*

**[0359]** In addition to the use of site-specific reactive moieties, the present invention contemplates the use of non-specific reactive groups to link the sugar to the modifying group.

**[0360]** Exemplary non-specific cross-linkers include photoactivatable groups, completely inert in the dark, which are converted to reactive species upon absorption of a photon of appropriate energy. In one embodiment, photoactivatable groups are selected from precursors of nitrenes generated upon heating or photolysis of azides. Electron-deficient nitrenes are extremely reactive and can react with a variety of chemical bonds including N-H, O-H, C-H, and C=C. Although three types of azides (aryl, alkyl, and acyl derivatives) may be employed, arylazides are presently. The reactivity of arylazides upon photolysis is better with N-H and O-H than C-H bonds. Electron-deficient arylnitrenes rapidly ring-expand to form dehydroazepines, which tend to react with nucleophiles, rather than form C-H insertion products. The reactivity of arylazides can be increased by the presence of electron-withdrawing substituents such as nitro or hydroxyl groups in the ring. Such substituents push the absorption maximum of arylazides to longer wavelength. Unsubstituted arylazides have an absorption maximum in the range of 260-280 nm, while hydroxy and nitroarylazides absorb significant light beyond 305 nm. Therefore, hydroxy and nitroarylazides are most preferable since they allow to employ less harmful photolysis conditions for the affinity component than unsubstituted arylazides.

**[0361]** In another preferred embodiment, photoactivatable groups are selected from fluorinated arylazides. The photolysis products of fluorinated arylazides are arylnitrenes, all of which undergo the characteristic reactions of this group, including C-H bond insertion, with high efficiency (Keana *et al., J. Org. Chem.* **55**: 3640-3647, 1990).

**[0362]** In another embodiment, photoactivatable groups are selected from benzophenone residues. Benzophenone reagents generally give higher crosslinking yields than arylazide reagents.

**[0363]** In another embodiment, photoactivatable groups are selected from diazo compounds, which form an electron-deficient carbene upon photolysis. These carbenes undergo a variety of reactions including insertion into C-H bonds, addition to double bonds (including aromatic systems), hydrogen attraction and coordination to nucleophilic centers to give carbon ions.

**[0364]** In still another embodiment, photoactivatable groups are selected from diazopyruvates. For example, the p-nitrophenyl ester of p-nitrophenyl diazopyruvate reacts with aliphatic amines to give diazopyruvic acid amides that undergo ultraviolet photolysis to

form aldehydes. The photolyzed diazopyruvate-modified affinity component will react like formaldehyde or glutaraldehyde forming crosslinks.

*iii. Homobifunctional Reagents*

*1. Homobifunctional crosslinkers reactive with primary amines*

**[0365]** Synthesis, properties, and applications of amine-reactive cross-linkers are commercially described in the literature (for reviews of crosslinking procedures and reagents, *see above*). Many reagents are available (*e.g.*, Pierce Chemical Company, Rockford, Ill.; Sigma Chemical Company, St. Louis, Mo.; Molecular Probes, Inc., Eugene, OR.).

**[0366]** Preferred, non-limiting examples of homobifunctional NHS esters include disuccinimidyl glutarate (DSG), disuccinimidyl suberate (DSS), bis(sulfosuccinimidyl) suberate (BS), disuccinimidyl tartarate (DST), disulfosuccinimidyl tartarate (sulfo-DST), bis-2-(succinimidooxycarbonyloxy)ethylsulfone (BSOCOES), bis-2-(sulfosuccinimidooxy-carbonyloxy)ethylsulfone (sulfo-BSOCOES), ethylene glycolbis(succinimidylsuccinate) (EGS), ethylene glycolbis(sulfosuccinimidylsuccinate) (sulfo-EGS), dithiobis(succinimidyl-propionate (DSP), and dithiobis(sulfosuccinimidylpropionate (sulfo-DSP). Preferred, non-limiting examples of homobifunctional imidoesters include dimethyl malonimidate (DMM), dimethyl succinimidate (DMSC), dimethyl adipimidate (DMA), dimethyl pimelimidate (DMP), dimethyl suberimidate (DMS), dimethyl-3,3'-oxydipropionimidate (DODP), dimethyl-3,3'-(methylenedioxy)dipropionimidate (DMDP), dimethyl-,3'-(dimethylenedioxy)dipropionimidate (DDDP), dimethyl-3,3'-(tetramethylenedioxy)-dipropionimidate (DTDP), and dimethyl-3,3'-dithiobispropionimidate (DTBP).

**[0367]** Preferred, non-limiting examples of homobifunctional isothiocyanates include: p-phenylenediisothiocyanate (DITC), and 4,4'-diisothiocyano-2,2'-disulfonic acid stilbene (DIDS).

**[0368]** Preferred, non-limiting examples of homobifunctional isocyanates include xylene-diisocyanate, toluene-2,4-diisocyanate, toluene-2-isocyanate-4-isothiocyanate, 3-methoxydiphenylmethane-4,4'-diisocyanate, 2,2'-dicarboxy-4,4'-azophenyldiisocyanate, and hexamethylenediisocyanate.

**[0369]** Preferred, non-limiting examples of homobifunctional arylhalides include 1,5-difluoro-2,4-dinitrobenzene (DFDNB), and 4,4'-difluoro-3,3'-dinitrophenyl-sulfone.

**[0370]** Preferred, non-limiting examples of homobifunctional aliphatic aldehyde reagents include glyoxal, malondialdehyde, and glutaraldehyde.

**[0371]** Preferred, non-limiting examples of homobifunctional acylating reagents include nitrophenyl esters of dicarboxylic acids.

**[0372]** Preferred, non-limiting examples of homobifunctional aromatic sulfonyl chlorides include phenol-2,4-disulfonyl chloride, and α-naphthol-2,4-disulfonyl chloride.

**[0373]** Preferred, non-limiting examples of additional amino-reactive homobifunctional reagents include erythritolbiscarbonate which reacts with amines to give biscarbamates.

*2. Homobifunctional Crosslinkers Reactive with Free Sulfhydryl Groups*

**[0374]** Synthesis, properties, and applications of such reagents are described in the literature (for reviews of crosslinking procedures and reagents, *see above*). Many of the reagents are commercially available (*e.g.*, Pierce Chemical Company, Rockford, Ill.; Sigma Chemical Company, St. Louis, Mo.; Molecular Probes, Inc., Eugene, OR).

**[0375]** Preferred, non-limiting examples of homobifunctional maleimides include bismaleimidohexane (BMH), N,N'-(1,3-phenylene) bismaleimide, N,N'-(1,2-phenylene)bismaleimide, azophenyldimaleimide, and bis(N-maleimidomethyl)ether.

**[0376]** Preferred, non-limiting examples of homobifunctional pyridyl disulfides include 1,4-di-3'-(2'-pyridyldithio)propionamidobutane (DPDPB).

**[0377]** Preferred, non-limiting examples of homobifunctional alkyl halides include 2,2'-dicarboxy-4,4'-diiodoacetamidoazobenzene, α,α'-diiodo-p-xylenesulfonic acid, α, α'-dibromo-p-xylenesulfonic acid, N,N'-bis(b-bromoethyl)benzylamine, N,N'-di(bromoacetyl)phenylthydrazine, and 1,2-di(bromoacetyl)amino-3-phenylpropane.

*3. Homobifunctional Photoactivatable Crosslinkers*

**[0378]** Synthesis, properties, and applications of such reagents are described in the literature (for reviews of crosslinking procedures and reagents, *see above*). Some of the reagents are commercially available (*e.g.*, Pierce Chemical Company, Rockford, Ill.; Sigma Chemical Company, St. Louis, Mo.; Molecular Probes, Inc., Eugene, OR).

**[0379]** Preferred, non-limiting examples of homobifunctional photoactivatable crosslinker include bis-β-(4-azidosalicylamido)ethyldisulfide (BASED), di-N-(2-nitro-4-azidophenyl)-cystamine-S,S-dioxide (DNCO), and 4,4'-dithiobisphenylazide.

*iv. HeteroBifunctional Reagents*

*1. Amino-Reactive HeteroBifunctional Reagents with a Pyridyl Disulfide Moiety*

**[0380]** Synthesis, properties, and applications of such reagents are described in the literature (for reviews of crosslinking procedures and reagents, *see above*). Many of the reagents are commercially available (*e.g.*, Pierce Chemical Company, Rockford, Ill.; Sigma Chemical Company, St. Louis, Mo.; Molecular Probes, Inc., Eugene, OR).

**[0381]** Preferred, non-limiting examples of hetero-bifunctional reagents with a pyridyl disulfide moiety and an amino-reactive NHS ester include N-succinimidyl-3-(2-pyridyldithio)propionate (SPDP), succinimidyl 6-3-(2-pyridyldithio)propionamidohexanoate (LC-SPDP), sulfosuccinimidyl 6-3-(2-pyridyldithio)propionamidohexanoate (sulfo-LCSPDP), 4-succinimidyloxycarbonyl-α-methyl-α-(2-pyridyldithio)toluene (SMPT), and sulfosuccinimidyl 6-α-methyl-α-(2-pyridyldithio)toluamidohexanoate (sulfo-LC-SMPT).

*2. Amino-Reactive HeteroBifunctional Reagents with a Maleimide Moiety*

**[0382]** Synthesis, properties, and applications of such reagents are described in the literature. Preferred, non-limiting examples of hetero-bifunctional reagents with a maleimide moiety and an amino-reactive NHS ester include succinimidyl maleimidylacetate (AMAS), succinimidyl 3-maleimidylpropionate (BMPS), N- γ-maleimidobutyryloxysuccinimide ester (GMBS)N-γ-maleimidobutyryloxysulfo succinimide ester (sulfo-GMBS) succinimidyl 6-maleimidylhexanoate (EMCS), succinimidyl 3-maleimidylbenzoate (SMB), m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), m-maleimidobenzoyl-N-hydroxysulfosuccinimide ester (sulfo-MBS), succinimidyl 4-(N-maleimidomethyl)-cyclohexane-1-carboxylate (SMCC), sulfosuccinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate (sulfo-SMCC), succinimidyl 4-(p-maleimidophenyl)butyrate (SMPB), and sulfosuccinimidyl 4-(p-maleimidophenyl)butyrate (sulfo-SMPB).

*3. Amino-Reactive HeteroBifunctional Reagents with an Alkyl Halide Moiety*

**[0383]** Synthesis, properties, and applications of such reagents are described in the literature Preferred, non-limiting examples of hetero-bifunctional reagents with an alkyl halide moiety and an amino-reactive NHS ester include N-succinimidyl-(4-iodoacetyl)aminobenzoate (SIAB), sulfosuccinimidyl-(4-iodoacetyl)aminobenzoate (sulfo-SIAB), succinimidyl-6-(iodoacetyl)aminohexanoate (SIAX), succinimidyl-6-(6-((iodoacetyl)-amino)hexanoylamino)hexanoate (SIAXX), succinimidyl-6-(((4-(iodoacetyl)-amino)-

methyl)-cyclohexane-1-carbonyl)aminohexanoate (SIACX), and succinimidyl-4((iodoacetyl)-amino)methylcyclohexane-1-carboxylate (SIAC).

**[0384]** An example of a hetero-bifunctional reagent with an amino-reactive NHS ester and an alkyl dihalide moiety is N-hydroxysuccinimidyl 2,3-dibromopropionate (SDBP). SDBP introduces intramolecular crosslinks to the affinity component by conjugating its amino groups. The reactivity of the dibromopropionyl moiety towards primary amine groups is controlled by the reaction temperature (McKenzie *et al., Protein Chem.* **7**: 581-592 (1988)).

**[0385]** Preferred, non-limiting examples of hetero-bifunctional reagents with an alkyl halide moiety and an amino-reactive p-nitrophenyl ester moiety include p-nitrophenyl iodoacetate (NPIA).

**[0386]** Other cross-linking agents are known to those of skill in the art. *See*, for example, Pomato *et al.*, U.S. Patent No. 5,965,106. It is within the abilities of one of skill in the art to choose an appropriate cross-linking agent for a particular application.

*v. Cleavable Linker Groups*

**[0387]** In yet a further embodiment, the linker group is provided with a group that can be cleaved to release the modifying group from the sugar residue. Many cleaveable groups are known in the art. *See,* for example, Jung *et al.*, *Biochem. Biophys. Acta* **761**: 152-162 (1983); Joshi *et al.*, *J. Biol. Chem.* **265**: 14518-14525 (1990); Zarling *et al.*, *J. Immunol.* **124**: 913-920 (1980); Bouizar *et al.*, *Eur. J. Biochem.* **155**: 141-147 (1986); Park *et al.*, *J. Biol. Chem.* **261**: 205-210 (1986); Browning *et al.*, *J. Immunol.* **143**: 1859-1867 (1989). Moreover a broad range of cleavable, bifunctional (both homo- and hetero-bifunctional) linker groups is commercially available from suppliers such as Pierce.

**[0388]** Exemplary cleaveable moieties can be cleaved using light, heat or reagents such as thiols, hydroxylamine, bases, periodate and the like. Moreover, certain preferred groups are cleaved *in vivo* in response to being endocytized (*e.g.*, cis-aconityl; *see,* Shen *et al., Biochem. Biophys. Res. Commun.* **102**: 1048 (1991)). Preferred cleaveable groups comprise a cleaveable moiety which is a member selected from the group consisting of disulfide, ester, imide, carbonate, nitrobenzyl, phenacyl and benzoin groups.

**Conjugation of Modified Sugars to Peptides**

**[0389]** The 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 acceptor is consumed. The considerations discussed below, while set forth in the context of a sialyltransferase, are generally applicable to other glycosyltransferase reactions.

**[0390]** 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), and U.S. Pat. Nos. 5,352,670, 5,374,541, and 5,545,553.

**[0391]** 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.

**[0392]** In another embodiment, each of the first and second enzyme is a glycosyltransferase. In another embodiment, one enzyme is an endoglycosidase. In an additional 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 peptide at any point either before or after the addition of the modified sugar to the peptide.

**[0393]** The O-linked glycosyl moieties of the conjugates of the invention are generally originate with a GalNAc moiety that is attached to the peptide. Any member of the family of GalNAc transferases can be used to bind a GalNAc moiety to the peptide (Hassan H, Bennett EP, Mandel U, Hollingsworth MA, and Clausen H (2000). Control of Mucin-Type O-Glycosylation: O-Glycan Occupancy is Directed by Substrate Specificities of Polypeptide GalNAc-Transferases. (Eds. Ernst, Hart, and Sinay). Wiley-VCH chapter "Carbohydrates in Chemistry and Biology - a Comprehension Handbook", 273-292). The GalNAc moiety itself can be the intact glycosyl linker. Alternatively, the saccharyl residue is built out using one more enzyme and one or more appropriate glycosyl substrate for the enzyme, the modified sugar being added to the built out glycosyl moiety.

**[0394]** 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 cleave 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.

**[0395]** 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.

**[0396]** In another 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.

**[0397]** 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 30 °C. In another exemplary embodiment, one or more components of the present method are conducted at an elevated temperature using a thermophilic enzyme.

**[0398]** 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.

**[0399]** The present invention also provides for the industrial-scale production of modified peptides. As used herein, an industrial scale generally produces at least about 250 mg, preferably at least about 500 mg, and more preferably at least about 1 gram of finished, purified conjugate, preferably after a single reaction cycle, *i.e.*, the conjugate is not a combination the reaction products from identical, consecutively iterated synthesis cycles.

**[0400]** 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 (m-) 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 water-soluble polymers, therapeutic moieties, and biomolecules.

**[0401]** An enzymatic approach can be used for the selective introduction of (m-) PEG-ylated or (m-) PPG-ylated 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 peptide backbone, onto existing sugar residues of a glycopeptide or onto sugar residues that have been added to a peptide.

**[0402]** An acceptor for the sialyltransferase is present on the 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 GalNAc, 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)).

**[0403]** 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 peptide to include an acceptor by methods known to those of skill in the art. In an exemplary embodiment, a GalNAc residue is added to an O-linked glycosylation site by the action of a GalNAc transferase. Hassan H, Bennett EP, Mandel U, Hollingsworth MA, and Clausen H (2000). Control of Mucin-Type O-Glycosylation: O-Glycan Occupancy is Directed by Substrate Specificities of Polypeptide GalNAc-Transferases. (Eds. Ernst, Hart, and Sinay). Wiley-VCH chapter "Carbohydrates in Chemistry and Biology - a Comprehension Handbook", 273-292.

**[0404]** In an exemplary embodiment, the galactosyl acceptor is assembled by attaching a galactose residue to an appropriate acceptor linked to the peptide, *e.g.*, a GalNAc. The method includes incubating the 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.

**[0405]** 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.

**[0406]** In the discussion that follows, the method of the invention is exemplified by the use of modified sugars having a water-soluble polymer 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.

**[0407]** In an exemplary embodiment, an O-linked carbohydrate residue is "trimmed" prior to the addition of the modified sugar. For example a GalNAc-Gal residue is trimmed back to

GalNAc. A modified sugar bearing a water-soluble polymer is conjugated to one or more of the sugar residues exposed by the "trimming." In one example, a glycopeptide is "trimmed" and a water-soluble polymer is added to the resulting O-side chain amino acid or glycopeptide glycan via a saccharyl moiety, *e.g.*, Sia, Gal or GalNAc moiety conjugated to the water-soluble polymer. The modified saccharyl moiety is attached to an acceptor site on the "trimmed" glycopeptide. Alternatively, an unmodified saccharyl moiety, *e.g.*, Gal can be added the terminus of the O-linked glycan.

**[0408]** In another exemplary embodiment, a water-soluble polymer is added to a GalNAc residue via a modified sugar having a galactose residue. Alternatively, an unmodified Gal can be added to the terminal GalNAc residue.

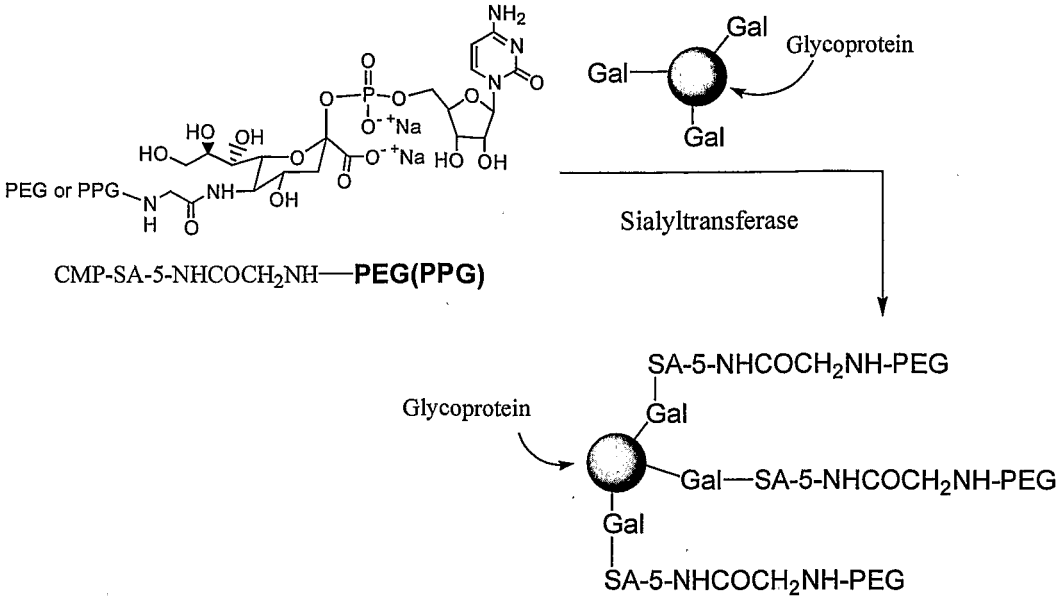
**[0409]** In yet a further example, a water-soluble polymer is added onto a Gal residue using a modified sialic acid.

**[0410]** In another exemplary embodiment, an O-linked glycosyl residue is "trimmed back" to the GalNAc attached to the amino acid. In one example, a water-soluble polymer is added via a Gal modified with the polymer. Alternatively, an unmodified Gal is added to the GalNAc, followed by a Gal with an attached water-soluble polymer. In yet another embodiment, one or more unmodified Gal residue is added to the GalNAc, followed by a sialic acid moiety modified with a water-soluble polymer.

**[0411]** The exemplary embodiments discussed above provide an illustration of the power of the methods set forth herein. Using the methods of the invention, 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.

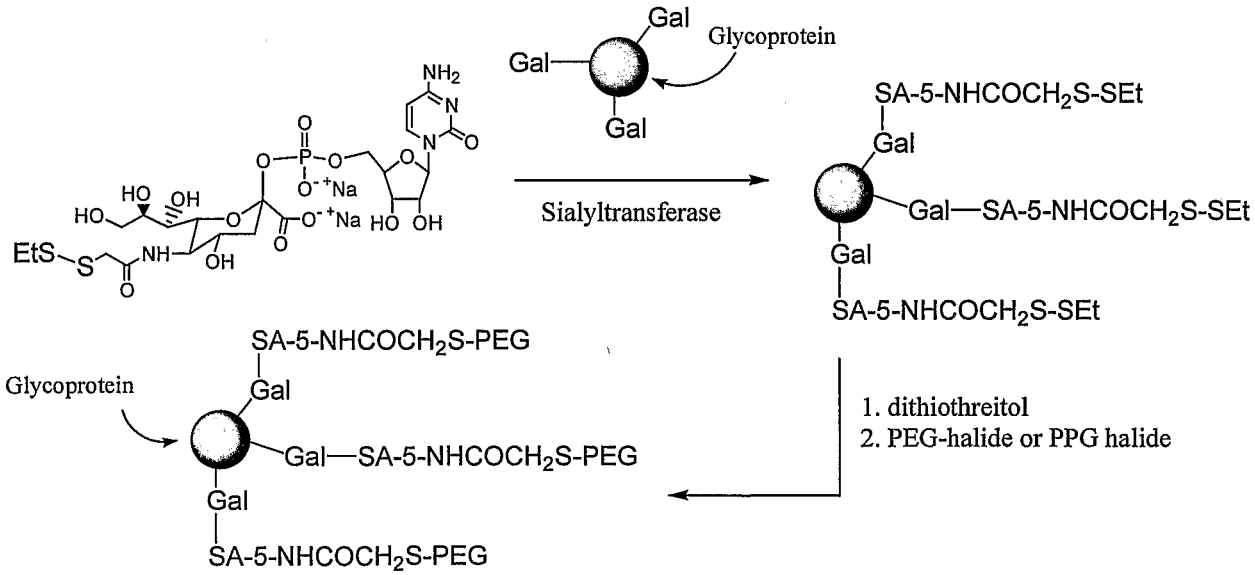
**[0412]** In an exemplary embodiment, the water-soluble polymer is added to a terminal Gal residue using a polymer modified sialic acid. An appropriate sialyltransferase is used to add a modified sialic acid. The approach is summarized in Scheme 5.

**Scheme 5**



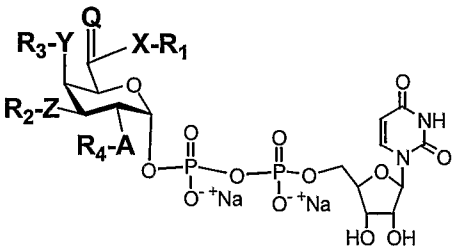

**[0413]** In yet a further approach, summarized in Scheme 6, a masked reactive functionality is present on the sialic acid. The masked reactive group is preferably unaffected by the conditions used to attach the modified sialic acid to the peptide. After the covalent attachment of the modified sialic acid to the peptide, the mask is removed and the peptide is conjugated with an agent such as PEG, PPG, a therapeutic moiety, biomolecule or other agent. 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 6**

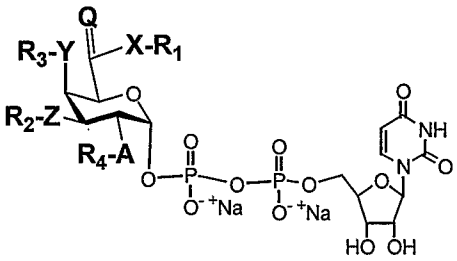


**[0414]** Any modified sugar can be used with its appropriate glycosyltransferase, depending on the terminal sugars of the oligosaccharide side chains of the glycopeptide (Table 3). As discussed above, the terminal sugar of the glycopeptide required for introduction of the PEG-ylated or PPGylated structure 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).

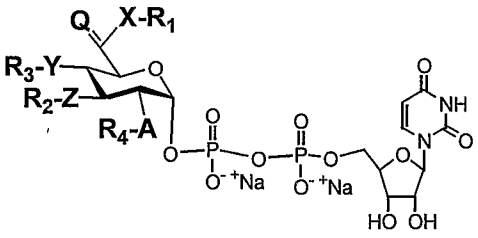
**Table 3**



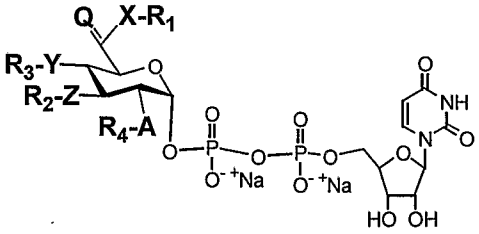
UDP-galactose-derivatives



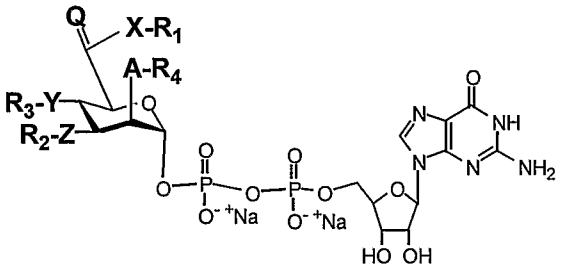
UDP-galactosamine-derivatives
(when A = NH, $R_4$ may be acetyl)



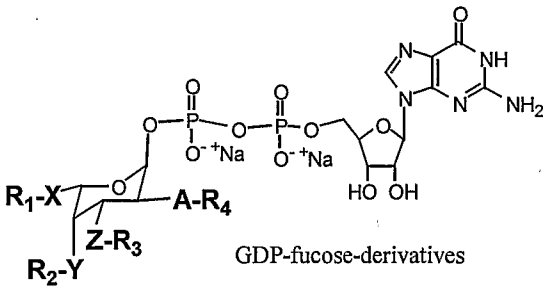
UDP-Glucose-derivatives



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



GDP-Mannose-derivatives

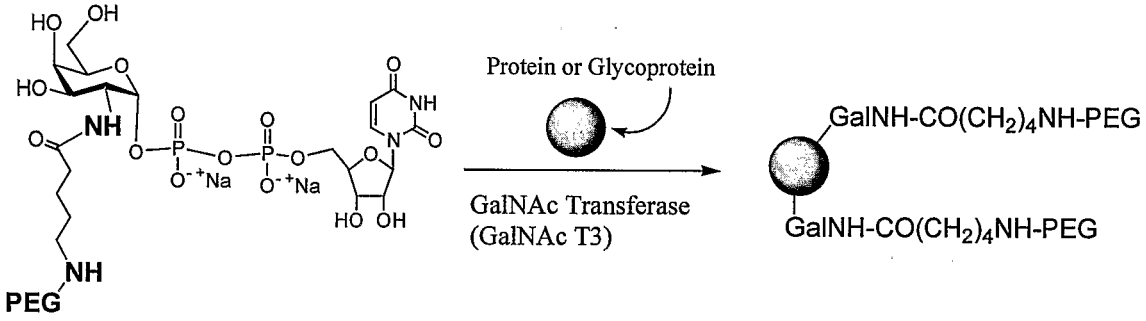


GDP-fucose-derivatives

| | |
|---|---|
| X = O, NH, S, $CH_2$, N-($R_1$-$_5$)$_2$.<br>Y = X; Z = X; A = X; B = X.<br><br>Q = $H_2$, O, S, NH, N-R.<br><br>R, $R_1$-$_4$ = H, Linker-M, M.<br><br>M = Ligand of interest | Ligand of interest = acyl-PEG, acyl-PPG, alkyl-PEG, acyl-alkyl-PEG, acyl-alkyl-PEG, carbamoyl-PEG, carbamoyl-PPG, PEG, PPG, acyl-aryl-PEG, acyl-aryl-PPG, aryl-PEG, aryl-PPG, Mannose-$_6$-phosphate, heparin, heparan, SLex, Mannose, FGF, VFGF, protein, chondroitin, keratan, dermatan, albumin, integrins, peptides, etc. |

**[0415]** In an alternative embodiment, the modified sugar is added directly to the peptide backbone using a glycosyltransferase known to transfer sugar residues to the O-linked glycosylation site on the peptide backbone. This exemplary embodiment is set forth in Scheme 7. Exemplary glycosyltransferases useful in practicing the present invention include, but are not limited to, GalNAc transferases (GalNAc T1-20), 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 7**




**[0416]** 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 peptide. In another example, an enzymatic reaction is utilized to "cap" (*e.g.*, sialylate) 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 peptide. Further elaboration of the modified sugar-peptide conjugate is within the scope of the invention.

**[0417]** In another exemplary embodiment, the glycopeptide is conjugated to a targeting agent, *e.g.*, transferrin (to deliver the peptide across the blood-brain barrier, and to endosomes), carnitine (to deliver the peptide to muscle cells; *see*, for example, LeBorgne *et al.*, *Biochem. Pharmacol.* **59**: 1357-63 (2000), and phosphonates, *e.g.*, bisphosphonate (to target the peptide to bone and other calciferous tissues; *see*, for example, Modern Drug Discovery, August 2002, page 10). Other agents useful for targeting are apparent to those of skill in the art. For example, glucose, glutamine and IGF are also useful to target muscle.

**[0418]** The targeting moiety and therapeutic peptide are conjugated by any method discussed herein or otherwise known in the art. Those of skill will appreciate that peptides in addition to those set forth above can also be derivatized as set forth herein. Exemplary peptides are set forth in the Appendix attached to copending, commonly owned US Provisional Patent Application No. 60/328,523 filed October 10, 2001.

**[0419]** In an exemplary embodiment, the targeting agent and the therapeutic peptide are coupled via a linker moiety. In this embodiment, at least one of the therapeutic peptide or the targeting agent is coupled to the linker moiety via an intact glycosyl linking group according to a method of the invention. In an exemplary embodiment, the linker moiety includes a poly(ether) such as poly(ethylene glycol). In another exemplary embodiment, the linker moiety includes at least one bond that is degraded *in vivo*, releasing the therapeutic peptide from the targeting agent, following delivery of the conjugate to the targeted tissue or region of the body.

**[0420]** In yet another exemplary embodiment, the *in vivo* distribution of the therapeutic moiety is altered via altering a glycoform on the therapeutic moiety without conjugating the therapeutic peptide to a targeting moiety. For example, the therapeutic peptide can be shunted away from uptake by the reticuloendothelial system by capping a terminal galactose moiety of a glycosyl group with sialic acid (or a derivative thereof).

*i. Enzymes*

*1. Glycosyltransferases*

**[0421]** Glycosyltransferases catalyze the addition of activated sugars (donor NDP-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.

**[0422]** 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.

**[0423]** 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.

**[0424]** 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.

**[0425]** 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.

**[0426]** 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.

**[0427]** 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 *rhl* operon of *Pseudomonas aeruginosa.*

**[0428]** 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).

*a) Fucosyltransferases*

**[0429]** 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.

**[0430]** 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.

*b) Galactosyltransferases*

**[0431]** 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.

**[0432]** 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)).

**[0433]** Also suitable in the practice of the invention are r soluble forms of α1, 3-galactosyltransferase such as that reported by Cho,S.K. and Cummings,R.D. (1997) J. Biol. Chem., 272, 13622-13628.

*c) Sialyltransferases*

**[0434]** 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, 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)).

**[0435]** 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 5).

**Table 5: 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-<br>NeuAcα2,3Galβ1,3GlcNAc- | 1 |
| ST3Gal IV | Mammalian | NeuAcα2,3Galβ1,4GlcNAc-<br>NeuAcα2,3Galβ1,3GlcNAc- | 1 |
| ST6Gal II | Mammalian | NeuAcα2,6Galβ1,4GlcNAc | |
| ST6Gal II | photobacterium | NeuAcα2,6Galβ1,4GlcNAc- | 2 |
| ST3Gal V | *N. meningitides*<br>*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)

**[0436]** 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 another embodiment, the claimed sialylation methods use a rat ST3Gal III.

**[0437]** Other exemplary sialyltransferases of use in the present invention include those isolated from Campylobacter jejuni, including the α(2,3). *See*, e.g, WO99/49051.

**[0438]** Sialyltransferases other those listed in Table 5, 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-$\alpha_1$ 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-$\alpha_1$ 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 (as illustrated for ST3Gal III in this disclosure). Other exemplary sialyltransferases are shown in Figure 10.

*d) GalNAc transferases*

**[0439]** 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, $\alpha(1,3)$ N-acetylgalactosaminyltransferase, $\beta(1,4)$ 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)).

**[0440]** Production of proteins such as the enzyme GalNAc $T_{I\text{-}XX}$ 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.

*2. Sulfotransferases*

**[0441]** 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).

*3. Cell-Bound Glycosyltransferases*

**[0442]** 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).

**[0443]** 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.

**[0444]** 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.

*4. Fusion Proteins*

**[0445]** 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.)

*5. Immobilized Enzymes*

**[0446]** In addition to cell-bound enzymes, 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.

**Purification of Peptide Conjugates**

**[0447]** 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.

**[0448]** 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, SP-Sepharose, 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.

**[0449]** 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, *e.g.*, SP Sepharose. Additionally, the

modified glycoprotein may be purified by affinity chromatography. HPLC may also be employed for one or more purification steps.

**[0450]** 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.

**[0451]** Within another embodiment, supernatants from systems which sproduce 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.

**[0452]** 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.

**[0453]** 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**

**[0454]** Polypeptides modified at various O-linked glycosylation site according to the method of the present invention have a broad range of pharmaceutical applications. For example, modified erythropoietin (EPO) may be used for treating general anemia, aplastic anemia, chemo-induced injury (such as injury to bone marrow), chronic renal failure,

nephritis, and thalassemia. Modified EPO may be further used for treating neurological disorders such as brain/spine injury, multiple sclerosis, and Alzheimer's disease.

**[0455]** A second example is interferon-α (IFN-α), which may be used for treating AIDS and hepatitis B or C, viral infections caused by a variety of viruses such as human papilloma virus (HBV), coronavirus, human immunodeficiency virus (HIV), herpes simplex virus (HSV), and varicella-zoster virus (VZV), cancers such as hairy cell leukemia, AIDS-related Kaposi's sarcoma, malignant melanoma, follicular non-Hodgkins lymphoma, Philladephia chromosome (Ph)-positive, chronic phase myelogenous leukemia (CML), renal cancer, myeloma, chronic myelogenous leukemia, cancers of the head and neck, bone cancers, as well as cervical dysplasia and disorders of the central nervous system (CNS) such as multiple sclerosis. In addition, IFN-α modified according to the methods of the present invention is useful for treating an assortment of other diseases and conditions such as Sjogren's symdrome (an autoimmune disease), Behcet's disease (an autoimmune inflammatory disease), fibromyalgia (a musculoskeletal pain/fatigue disorder), aphthous ulcer (canker sores), chronic fatigue syndrome, and pulmonary fibrosis.

**[0456]** Another example is interferon-β, which is useful for treating CNS disorders such as multiple sclerosis (either relapsing/remitting or chronic progressive), AIDS and hepatitis B or C, viral infections caused by a variety of viruses such as human papilloma virus (HBV), human immunodeficiency virus (HIV), herpes simplex virus (HSV), and varicella-zoster virus (VZV), otological infections, musculoskeletal infections, as well as cancers including breast cancer, brain cancer, colorectal cancer, non-small cell lung cancer, head and neck cancer, basal cell cancer, cervical dysplasia, melanoma, skin cancer, and liver cancer. IFN-β modified according to the methods of the present invention is also used in treating other diseases and conditions such as transplant rejection (*e.g.*, bone marrow transplant), Huntington's chorea, colitis, brain inflammation, pulmonary fibrosis, macular degeneration, hepatic cirrhosis, and keratoconjunctivitis.

**[0457]** Granulocyte colony stimulating factor (G-CSF) is a further example. G-CSF modified according to the methods of the present invention may be used as an adjunct in chemotherapy for treating cancers, and to prevent or alleviate conditions or complications associated with certain medical procedures, *e.g.*, chemo-induced bone marrow injury; leucopenia (general); chemo-induced febrile neutropenia; neutropenia associated with bone marrow transplants; and severe, chronic neutropenia. Modified G-CSF may also be used for

transplantation; peripheral blood cell mobilization; mobilization of peripheral blood progenitor cells for collection in patients who will receive myeloablative or myelosuppressive chemotherapy; and reduction in duration of neutropenia, fever, antibiotic use, hospitalization following induction/consolidation treatment for acute myeloid leukemia (AML). Other condictions or disorders may be treated with modified G-CSF include asthma and allergic rhinitis.

**[0458]** As one additional example, human growth hormone (hGH) modified according to the methods of the present invention may be used to treat growth-related conditions such as dwarfism, short-stature in children and adults, cachexia/muscle wasting, general muscular atrophy, and sex chromosome abnormality (*e.g.*, Turner's Syndrome). Other conditions may be treated using modified hGH include: short-bowel syndrome, lipodystrophy, osteoporosis, uraemaia, burns, female infertility, bone regeneration, general diabetes, type II diabetes, osteo-arthritis, chronic obstructive pulmonary disease (COPD), and insomia. Moreover, modified hGH may also be used to promote various processes, *e.g.*, general tissue regeneration, bone regeneration, and wound healing, or as a vaccine adjunct.

**[0459]** Thus, 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, water-soluble polymer, therapeutic moiety or biomolecule and a glycosylated or non-glycosylated peptide. The polymer, therapeutic moiety or biomolecule is conjugated to the peptide via an intact glycosyl linking group interposed between and covalently linked to both the peptide and the polymer, therapeutic moiety or biomolecule.

**[0460]** 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).

**[0461]** 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 matrises, such as 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.

**[0462]** Commonly, the pharmaceutical compositions are administered subcutaneously or 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 also contain detergents such as Tween 20 and Tween 80; stablizers such as mannitol, sorbitol, sucrose, and trehalose; and preservatives such as EDTA and m-cresol. 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.

**[0463]** 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.

**[0464]** 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).

**[0465]** 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.

**[0466]** 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.

**[0467]** 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.

**[0468]** 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

#### *1.1a Preparation of Interferon alpha-2β-GalNAc (pH 6.2)*

**[0469]** Interferon alpha-2β was reconstituted by adding 200 μL water to 4 mg of IFN alpha-2β. When the solid was dissolved, 1.92 mL reaction buffer (20 mM MES, pH 6.2, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM $MnCl_2$, 0.05% polysorbate, and 0.05% $NaN_3$), was added. UDP-GalNAc (4.16 mg; 3 mM) and GalNAc T2 (80 mU; 80 μL) were then added and the reaction mixture was incubated at 32 °C with slow rotary movement. The reaction was monitored using MALDI analysis and was essentially complete after 72 h

. Once complete, the reaction mixture was submitted for peptide mapping, and analysis of site occupancy.

*1.1b Preparation of Interferon alpha-2β-GalNAc (pH 7.4).*

**[0470]** The interferon alpha 2β was reconstituted as described by the manufacturer. Water, 50 µL, was added to 50 µg of IFN alpha-2β. When the solid was dissolved, the reaction buffer (20 mM MES, pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM $MnCl_2$, 0.05% polysorbate, and 0.05% $NaN_3$.), 50 µL was added. The UDP-GalNAc (100 µg; 3 mM) and GalNAc T2 (8 mU; 8 µL) were then added and the reaction mixture incubated at 32 °C under a slow rotary movement. The reaction was monitored using MALDI analysis and was found to be complete within about 48 to 72 h

*1.2 Preparation of Interferon-alpha-2β-GalNAc-SA-PEG-20kilodalton using CMP-SA-PEG and ST6GalNAcI*

**[0471]** The IFN-alpha-2β-GalNAc (1.0 mL, ~2 mg, 0.1 µmole) from *1.1* (above) was buffer exchanged (2x) using a 5 kilodalton MWCO Filter Centricon cartridge and a second buffer (20 mM MES, pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM $MnCl_2$, 0.05% polysorbate, and 0.05% $NaN_3$). The IFN-alpha-2β-GalNAc was reconstituted from the spin cartridge using the second buffer, 1.0 mL, and both CMP-SA-PEG-20kilodalton (10 mg, 0.5 micromoles) and ST6GalNAc1 (200 µL) were added to the reaction mixture. The reaction was incubated at 32 °C for 96 h with slow rotary movement. The product, IFN-alpha-2β-GalNAc-SA-PEG-20kilodalton was purified using SP Sepharose and SEC (Superdex 75) chromatography. The addition of sialic acid-PEG was verified using MALDI analysis .

*1.3. Preparation of Interferon-alpha-2β-GalNAc-Gal-SA-PEG-20kilodalton using CMP-SA-PEG, core-1-β1,3-galactosyl-transferase, and ST3Gal2*

**[0472]** The IFN-alpha-2β-GalNAc (1.0 mL, ~2 mg, 0.1 µmole) from the addition of GalNAc described above (pH 6.2) was buffer exchanged (2x) using a 5 kilodalton MWCO Filter Centricon cartridge and a second buffer (20 mM MES, pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM $MnCl_2$, 0.05% polysorbate, and 0.05% $NaN_3$). The IFN-alpha-2β-GalNAc was reconstituted from the spin cartridge using 1.0 mL of the second buffer, containing CMP-SA-PEG-20kilodalton (10 mg, 0.5 micromoles), UDP-Galactose (1.8 mg, 3 mM), core-1-β1,3-galactosyl-transferase (200 mU on resin) and ST3Gal2 (200 mU, α2,3-(O)-sialytransferase). The reaction mixture was incubated at 32 °C for 96 h with slow rotary movement. The

product, IFN-alpha-2β-GalNAc-Gal-SA-PEG-20kilodalton, was purified by SP Sepharose and SEC (Superdex 75) chromatography. The addition of sialic acid-PEG was verified using MALDI analysis.

### *1.9 Protein Concentration Assay*

**[0473]** Protein concentration was determined using a spectrophotometer at a fixed absorbance of 280 nm with 1 cm path length of cell. Triplicate readings were measured for a tested sample with water and buffer as controls. Protein concentration was determined using extinction coefficient at 0.799 mL/mg protein.

### *1.10 Formulation of Final Product*

**[0474]** The formulation buffer contained pyrogen-free PBS, pH 6.5, 2.5% mannitol, and 0.05% Polysorbate 80 that was degassed by vacuum and sterile filtered (0.2 μm).

**[0475]** Endotoxin was removed using a Detoxi-Gel™ equilibrated with 5 column beds of the formulation buffer (PBS, pH 6.5, 2.5% mannitol, and 0.05% Polysorbate 80). The flow rate was controlled by gravity at ~ 0.3 mL/min. Product samples were applied onto the gel, and the product eluted using the formulation buffer. The volume of the collected product was adjusted with additional formulation buffer to provide a protein concentration of about 100 μg/mL.

**[0476]** The peptide formulations were sterile filtered (0.2 μ) and the effluent was dispensed as 1 mL aliquots into 2.0 mL pyrogen-free vials. In addition, aliquots were taken for endotoxin and protein analysis. All products were stored at 4 °C.

### *1.13 Pharmacokinetic Study*

**[0477]** The pharmacokinetic analysis was performed using radioiodinated protein. After administration of the labeled interferons by IV tail vein injections into the rats, the clearance rate was measured as the reduction in radioactivity in blood drawn at specific intervals over 72 h. Each time point is a measure of at least five rats.

### *1.14 Results*

**[0478]** The reaction rate of GalNAc-T2 was measured at two pH's, a neutral pH (7.4) and a slightly acidic pH (6.2). Glycosylation with GalNAc proceeded sucessfully at both pH 6.2

and pH 7.4. As can be seen in the MALDI analysis of the reaction progress, the reaction rate was faster at pH 7.4 than at pH 6.2.

**[0479]** GalNAc-T2 and GalNAc were added to interferon alpha-2β quantitatively at either pH 6.2 or pH 7.4. The reaction was followed by MALDI. During the enzymatic reaction, a new interferon alpha mass ion formed (IFN-alpha-2b 19,281 Da and IFN-alpha-2β-GalNAc, 19,485 Da).

**[0480]** The product of the reaction at pH 6.2, IFN-alpha-2b-GalNAc, was submitted to analysis to determine the position of substitution of the GalNAc on the protein. Peptide mapping and site occupancy mapping were used for this purpose. Peptide mapping using TIC of LC-MS/MS and a GluC digest of IFN-alpha-2b produced a peptide fragment of mass 1018.69. MS/MS peptide amino acid sequencing of the peptide mass ion of 1018.69 containing the GalNAc indicated that sugar was attached to $T^{106}$.

**[0481]** The sialyl-PEGylation of IFN-alpha-2b-GalNAc was examined using ST6GalNAc-1 and CMP-SA-PEG-20 kilodalton. The reaction of IFN-alpha-2b -GalNAc produced the PEG-ylated protein, which was visible by SDS PAGE. In general, the reaction proceeded at 32 °C for 96 h. The reaction was monitored by SDS PAGE. SDS PAGE indicated that about 70% of the IFN-alpha-2b -GalNAc was converted to IFN-alpha-2b -GalNAc-SA-PEG-20 kilodalton. The MALDI analysis of the new band indicated a mass ion of 41,500 Daltons, the mass of IFN-alpha-2b -GalNAc-SA-PEG-20 kilodalton.

**[0482]** The glycoform of PEG-ylated interferon alpha-2b containing the GalNAc-Gal-SA-PEG structure was also produced. The reaction was performed using the conditions described above. The desired product was detected by SDS PAGE. A one pot, two-step reaction was used to produce the desired product, beginning with IFN-alpha-2β-GalNAc with core-1-β3-galactosyltransferase-1, ST3Gal2, UDP-galactose and CMP-SA-PEG-20 kilodalton. The reaction was incubated at 32 °C for 96 h. The reaction was monitored by SDS PAGE. After 24 h, the reaction was about 70% complete. The MALDI of the product indicated a mass ion of 41,900 Da, which originates from the desired IFN-alpha-2β-GalNAc-Gal-SA-PEG-20 kilodalton product.

**[0483]** Both glycoforms of the PEG-ylated interferon alpha-2b products were purified using a two-step process. In the first step, ion-exchange chromatography was performed using SP Sepharose. This procedure removed unreacted PEG materials and provided some separation of other proteins. The ion exchange step was followed by separation on SEC. A

Superdex 75 column was used to remove remaining smaller proteins including the glycosyltransferases and unPEG-ylated interferon alpha. Both PEG-ylated glycoforms of interferon alpha were purified to greater than 90% as shown by SDS PAGE).

**[0484]** The antiviral data indicates that PEG-ylated glycoforms **A** and **B** retain their antiviral effects).

**[0485]** The radioiodinated PEG-ylated proteins were injected into rats via their tail veins, the AUC for both proteins was 5-7 fold greater un-PEG-ylated interferon alpha-2β.

**[0486]** Glycoform **A** (IFN-alpha-2β-GalNAc-SA-PEG-20kilodalton) and **B** (IFN-alpha-2β-GalNAc-Gal-SA-PEG-20kilodalton) were both bioactive.

**EXAMPLE 2**

***2.1 Preparation of G-CSF-GalNAc (pH 6.2)***

**[0487]** 960 μg of G-CSF in 3.2 mL of buffer was concentrated by utrafiltration using a UF filter (5 kilodalton) and 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 for 48 hours. After 48 hours, MALDI indicated the reaction was complete (shift of the mass ion from 18800 to 19023 mass units). The reaction mixture was purified by HPLC using SEC (Superdex 75 and Superdex 200). The column was eluted using phosphate buffered saline, pH 4.9 and 0.005% Tween 80. The peak corresponding to G-CSF-GalNAc was collected and concentrated to about 150 μL using a Centricon 5 kilodalton filter and the volume was adjusted to 1 mL using PBS (phosphate buffered saline, pH 4.9 and 0.005% Tween 80); protein concentration was 1 mg/mL $A_{280}$).

***2.2 Preparation of G-CSF-GalNAc-Gal (pH 6.0)***

**[0488]** G-CSF-GalNAc (100 μg) was added to a 100 μL of a solution containing 25 mM MES buffer, pH 6.0, 1.5 mM UDP-GalNAc, 10 mM $MgCl_2$ and 80 mU GalNAc-T2. The CMP-SA-PEG-20 kilodalton (0.5 mg, 0.025 μmole), UDP-galactose 75 μg (0.125 μmole), core-1-Gal-T 20 μL (10 mU) were then added and the solution which was slowly rocked at 32 °C for 24 hours. MALDI indicated complete conversion of G-CSF-GalNAc into G-CCSF-GalNAc-Gal.

### *2.3 Preparation of G-CSF-GalNAc-SA-PEG-20 kilodalton (C).*

*2.3a Sequential Process (pH 6.2).*

**[0489]** A G-CSF-GalNAc solution containing 1 mg of protein was buffer exchanged into 25 mM MES buffer (pH 6.2, 0.005% $NaN_3$) then 5 mg, (0.25 μmole) CMP-SA-PEG (20kilodalton) was added. Finally, 100 μL, of a 100 mM $MnCl_2$ solution and ST6GalNAc-I (100 μL) were added and the reaction mixture was rocked slowly at 32 °C. Aliquots were taken at time points (24, 48 and 72 h) and analyzed by SDS-PAGE. After 24 h, no further reaction was observed. The reaction mixture was concentrated by spin filtration (5 kilodalton), buffer exchanged against 25 mM NaOAc (pH 4.9) and concentrated to 1 mL. The product was purified using ion exchange (SP-Sepharose, 25 mM NaOAc, pH 4.9) and SEC (Superdex 75; PBS-pH 7.2, 0.005% tween 80, 1 ml/min). The desired fraction was collected, concentrated to 0.5 mL and stored at 4 °C.

*2.3b One Pot process using ST6GalNAc-I (pH 6.0)*

**[0490]** 960 μg of G-CSF protein dissolved in 3.2 mL of product formulation buffer was concentrated by spin filtration (5 kilodalton) to 0.5 mL and reconstituted in 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. Following reconstitution UDP-GalNAc (6 mg, 9.21 μmol), GalNAc-T2 (80 μL, 80 mU), CMP-SA-PEG-20 kilodalton (6 mg, 0,3 μmol ) and mouse enzyme ST6GalNAc-I (120 μL)were added. The solution was rocked at 32 °C for 48 hours. Following the reaction the productwas purified using standard chromatography conditions on SP-Sepharose and SEC as described above. A total of 0.5 mg of protein ($A_{280}$) was obtained, for about a 50% overall yield. The product structure was confirmed by analysis with both MALDI and SDS-PAGE

### *2.4 Preparation of G-CSF-GalNAc-Gal-SA-PEG-20 kilodalton (D)*

*2.4a Starting from G-CSF-GalNAc*

**[0491]** UDP-galactose (4 mg, 6.5 μmole), core-1-Gal-$T_1$ (320 μL, 160 mU), CMP-SA-PEG-20 kilodalton (8 mg, 0.4 μmole), ST3Gal2 (80 μL, 0.07 mU) and 80 μL of 100 mM $MnCl_2$ were directly added to the crude 1.5 mL of reaction mixture of the G-CSF-GalNAc (1.5 mg) in 25 mM MES buffer (pH 6.0) from Example 2.1 (above). The resulting mixture was incubated at 32 °C for 60 hours, however, the reaction was complete after 24 h. The reaction mixture was centrifuged and the solution was concentrated to 0.2 mL using ultrafiltration (5 kilodalton) and then redissolved in 25 mM NaOAc (pH 4.5) to a final

volume of 1 mL. The product was purified using SP-Sepharose, the peak fractions were concentrated using a spin filter (5 kilodalton) and the residue purified further using SEC (Superdex 75). After concentration using a spin filter (5 kilodalton), the protein was diluted to 1 mL using formulation buffer consisting of PBS, 2.5% mannitol, 0.005% polysorbate, pH 6.5, and formulated at a protein concentration of 850 μg protein per mL ($A_{280}$). The overall yield was 55%. The MALDI analysis is shown in **FIG 28**.

*2.4b Starting from G-CSF*

**[0492]** 960 μg, of G-CSF (3.2 mL) was concentrated by spin filter (5 kilodalton) and reconstituted with 25 mM MES buffer (pH 6.0, 0.005% $NaN_3$). The total volume of the G-CSF solution was adjusted to about 1 mg/mL and UDP-GalNAc (6 mg), GalNAc-T2 (80 μL), UDP-galactose (6 mg), core-1-Gal-$T_1$ (160 μL, 80 μU), CMP-SA-PEG (20 kilodalton) (6 mg), ST3Gal-2 (160 μL, 120 μU) and $MnCl_2$ (40 μL of a 100 mM solution) were added. The resulting mixture was incubated at 32 °C for 48 h.

***2.5 SP Sepharose HPLC Chromatography***

**[0493]** The SP Sepharose was performed as described in Example 1.4.

***2.6 Size Exclusion Chromatography***

**[0494]** SEC was performed as described in Example 1.5. The purified samples were stored at 4 °C.

***2.6a Hydrophobic interaction chromatography (HIC)***

Follwing the first step of chromatographic chromatography HIC can be used as a second purification step to remove contaminants other thn un-Pegylated G-CSF. Thus, a method is available for the purification of glycopegylated G-CSF that has been through an initial purificatio on a gel permeation column.

***2.7 SDS PAGE Analysis***

**[0495]** The SDS PAGE was performed as set forth in Example 1.6.

***2.8 MALDI Analysis***

**[0496]** MALDI analysis was performed as described in Example 1.7.

*2.9 Peptide Mapping Analysis*

**[0497]** Protein mapping analysis was performed as illustrated in Example 1.8

*2.10 Protein Concentration Assay*

**[0498]** Protein concentration was determined as described in Example 1.9.

*2.11 Product Formulation*

**[0499]** The product was formulated as set forth in Example 1.10

*2.12 Endotoxin Determination*

**[0500]** Endotoxin was determined as set forth in Example 1.11.

*2.13 Cell proliferation assay*

**[0501]** A G-CSF proliferation assay with a NFS-60 cell line and a Tf-1 cell line were performed according to standard procedures. The cells were plated into a 96 well plate at 25000 cell/ml in the presence of different concentrations of G-CSF (51 nM, 25.5 nM, 12.75 nM, 3.2 nM, 1.6 nM, 0.8 nM, 0 nM), a chemically PEG-ylated G-CSF analogue, and PEGylated G-CSF **C** from Example 2.3 (above), and PEGylated G-CSF **D** from Example 2.4 (above). The cells were incubated at 37 °C for 48 hours. A colorimetric MTT assay was used to determine the cell viability.

*2.14 In Vivo Activity: White Blood Cell (WBC) Production in the Rat*

**[0502]** Two doses of drug (50 μg/kg, 250 μg/kg) were examined for each of **C**, G-CSF and a chemically PEG-ylated G-CSF using mice. Blood was drawn at time points of 2 hour, 12 hour, 24 hour, 36 hour, 48 hour, 60 hour, 72 hour, 84 hour and 96 hour, and the WBC and neutrophil counts were measured (**FIG. 4**).

*2.15 Accelerated Stability Study*

**[0503]** An accelerated stability study of PEGylated G-CSF, **C,** from Example 2.3, and PEGylated G-CSF, **D,** from Example 2.4 was performed using a buffer at pH 8.0 heated to 40 °C. 72 μg of PEGylated G-CSF **C,** was diluted to 8 mL with formulation buffer (PBS, 2.5% mannitol, 0.005% polysorbate 80). 1 mg of PEGylated G-CSF **D,** was diluted with 16 mL of formulation buffer. Both solutions were adjusted to pH 8.0 with NaOH and the resulting solution was sterile filtered into pyrogen-free tubes. The samples were slowly rotated at 40

°C and aliquots (0.8 mL) were taken at timepoints of 0 hour, 72 hours and 168 hours. Analysis was performed using SEC (Superdex 200) as described above (**FIG. 6** and **FIG. 7**).

### *2.16 Protein Radiolabeling*

**[0504]** G-CSF was radiolabeled using the Bolton Hunter reagent. This reaction was performed at pH 7.4 for 15 minutes and was followed by a SEC (Superdex 200) purification. Once purified, the formulation buffer pH was adjusted to 5.0 and the protein concentration was determined by $A_{280}$.

### *2.17 ELISA Assay*

**[0505]** An Elisa assay was utilized to quantify the G-CSF derivatives in rat plasma. The pharmacokinetic results are shown in **FIG. 9**.

### *2.18 Pharmacokinetic Study*

**[0506]** Two pharmacokinetic studies were performed. For the first pharmacokinetic study proteins were radiolabeled and administered by IV tail vein injections into rats. Clearance rate was measured as the reduction in radioactivity in blood drawn at specific intervals over 48 hours. Each time point was a measure of at least five rats.

**[0507]** Specifically,10 μg of G-CSF derivative was injected per animal (~1 μg of labeled protein and 9 μg of unlabeled protein). In addition to the blood being drawn and counted as described above, plasma was also collected and the protein acid was precipitated. The protein pellets were then also counted for radioactivity. The data from these studies is shown is **FIG. 2**, **FIG. 3** and **FIG. 8**.

**[0508]** In the second pharmacokinetic study the unlabeled G-CSF derivatives (30 μg per animal) were administered by IV tail vein injections into rats. Blood samples were drawn at the time points indicated and the samples analzed by the G-CSF ELISA assay. The data is shown in **FIG. 9**.

### *2.19 Results*

**[0509]** Human GalNAc T2 transferred GalNAc to G-CSF expressed in *E. coli* \ using UDP-GalNAc as the donor. Depending on the pH of the reaction buffer, one or two GalNAc moities were added to G-CSF as determined by MALDI. Addition of the second GalNAc proceeded slowly amounting to about 10-15% of the total product. One GalNAc could be

selectively added to G-CSF, in conversion yields of over 90%, by adjusting the pH of the reaction solution to 6.0-6.2. Addition of the second GalNAc occurred when the reaction was performed at a pH between about 7.2 and 7.4. Both $Co^{+2}$ and $Mn^{+2}$ are useful divalent metal ions in the reaction. Peptide mapping of the reaction products indicated that the predominant product of the reaction was addition of GalNAc to threonine-133, the natural site of O-linked glycosylation in mammalian systems \. The second GalNAc was observed in the amino terminal peptide fragment of G-CSF and is postulated to occur at threonine-2.

**[0510]** The reaction of G-CSF-GalNAc with ST6GalNAc-1 (chicken or mouse) and CMP-SA-PEG-20 kilodalton provided the product G-CSF-GalNAc-SA-PEG-20 kilodalton, which was verified by MALDI\, with conversion yields of about 50% as determined by SDS-PAGE \. The G-CSF-GalNAc could also be further elongated using core-1-Gal-T and UDP-galactose to provide complete conversion to G-CSF-GalNAc-Gal\. Glyco-PEG-ylation of this intermediate with ST3Gal2 and CMP-SA-PEG-20 kilodalton then provided the product G-CSF-GalNAc-Gal-SA-PEG-20 kilodalton in overall yields of about 50% \. These reactions were performed either sequentially in one pot or simultaneously in one pot starting from G-CSF or its glycosylated intermediates. In these studies, little or no difference was observed in overall yield by using either approach.

**[0511]** The products of the glycosylation or glyco-PEG-ylation reactions were purified using a combination of ion exchange and SEC. The ion exchange step removes the unreacted G-CSF or its glycosylated intermediates (GalNAc or GalNAc-Gal) as well as any unreacted CMP-SA-PEG-20 kilodalton\. The SEC step removed remaining unreacted G-CSF and other protein contaminants from the glycosyltransferases used in the process\. The G-CSF's containing the GalNAc-SA-PEG-20 kilodalton or the GalNAc-Gal-SA-PEG-20 kilodalton had identical properties and retention times using these purification methods. The final products had typical profiles as shown in.

**[0512]** Once purified, the PEG-ylated proteins were formulated in a PBS buffer containing 2.5% mannitol and 0.005% Tween 80. Initially, pH 6.5 was used in the formulation but aggregation of the glyco-PEG-ylated protein was a concern (see below) so the formulation buffer pH was lowered to 5.0. Literature reports have indicated that G-CSF aggregation is prevented by maintaining a solution pH between 4-5. Endotoxin was removed using an endotoxin removal cartridge using sterile technique. Protein concentrations were typically adjusted to concentrations between 100 μg/mL to 1 mg/mL as required for biological studies.

Endotoxin calculations were typically below 3EU/ml by this process. The formulated products are stored at 4.

**[0513]** The products were tested in an *in vitro* cell proliferation assay using NSF-60 cells sensitive to G-CSF. It was observed that both the GalNAc-SA-PEG-20 kilodalton and GalNAc-Gal-SA-20 kilodalton products were effective at initiating cell proliferation (**FIG. 1**).

**[0514]** An accelerated stability study was performed on a chemically PEG-ylate G-CSF and **C** (G-CSF-GalNAc-SA-PEG-20 kilodalton). The formulation buffer pH was adjusted to 8.0 and the temperature was raised to 40 °C. Samples were taken of each protein at times 0, 72 and 168 h (**FIG. 6**and **FIG. 7**). Chemically PEG-ylated G-CSF was observed to aggregate entirely under these conditions within 168 h. SEC using a Superdex 200 chromatography was used to separate the aggregates. Although the glycoconjugate G-CSF-GalNAc-SA-PEG-20 kilodalton also formed aggregates that were separable using SEC, the aggregation occurred at a much slower rate.

**[0515]** The glyco-PEG-ylated G-CSF was radioiodinated using the Bolton Hunter reagent. A cold labeling study was also performed prior to the actual radiolabeling to determine the extent of aggregation and to establish a methodology for removing any aggregates formed. Use of the Bolton Hunter reagent (cold) did provide some aggregates as shown in **FIG. 5**. SEC using a Superdex 200 column removed the aggregates and provided the monomeric, labeled material. Similar results were obtained using $^{125}I$ labeled reagent. The use of the formulation minimized aggregation on storage. Protein content was measured by measuring the absorbance at $A_{280}$.

**[0516]** The results of the rat pK study incorporating G-CSF, chemically PEG-ylated G-CSF and the PEG-G-CSF conjugate labeled with the Bolton Hunter reagent are shown in **FIG. 3**. In this study, blood and protein precipitated from plasma were counted for radioactivity after IV administration of 10 μg of G-CSF conjugate per rat. The data from both blood and plasma protein clearly indicate that the PEG conjugate and Chemically PEG-ylated G-CSF have identical clearance rates (**FIG. 3** and **FIG. 8**).

**[0517]** The ability of the G-CSF derivatives to initiate WBC production was then examined in a mouse model. Each test compound was injected IV as a single bolus and the induction of WBC and neutrophils was monitored over time. Chemically PEG-ylated G-CSF was the most potent protein tested when administered at 250 μg/kg. The PEG conjugate (G-CSF-

GalNAc-SA-PEG-20 kilodalton) induced WBC production to almost the same degree as Chemically PEG-ylated G-CSF at 250 μg/kg, and far greater than G-CSF at a similar concentration.

**EXAMPLE 3**

**[0518]** This example discloses amino acid sequence mutations that introduce changes introduce O-linked glycosylation sites, *i.e.*, serine or threonine residues, into a preferably proline-containing site in the 175 amino acid wild-type sequence of G-CSF or any modified version thereof. As a reference the 175 amino acid wild-type G-CSF sequence is shown below:

MTPLGPASSLP QSFLLKCLEQ VRKIQGDGAA LQEKLCA
TYKLCHPEEL VLLGHSLGIP WAPLSSCPSQ ALQLAGCLSQ
LHSGLFLYQG LLQALEGISP ELGPTLDTLQ LDVADFATTI
WQQMEELGMA PALQPTQGAM PAFASAFQRR AGGVLVASHL
QSFLEVSYRV LRHLAQP **(SEQ ID NO:2)**

*3.1 N-terminal Mutations*

**[0519]** In the N-terminal mutants, the N-terminus of a wild-type G-CSF, $M^1$TPLGPA (SEQ ID NO:), is replaced with either $M^1X_n$TPLGPA or $M^1B_oPZ_mX_n$TPLGPA. Wherein n, o and m are integers slected from 0 to 3, and at least one of X, B and O is Thr or Ser. When more than one of X, B and O is Thr or Ser, the identity of these moieties is independently selected. Where they appear, superscripts denote the position of the amino acid in the wild-type starting sequence.

**[0520]** Preferred examples include:

**M**$^1$VTPL$^4$GPA (SEQ ID NO:)
**M**$^1$QTPL$^4$GPA (SEQ ID NO:)
**M**$^1$ATPL$^4$GPA (SEQ ID NO:)
**M**$^1$PTQGAMPL$^4$GPA (SEQ ID NO:)
**M**$^1$VQTPL$^4$GPA (SEQ ID NO:)
**M**$^1$QSTPL$^4$GPA (SEQ ID NO:)
**M**$^1$GQTPL$^4$GPA (SEQ ID NO:)
**M**$^1$APTSSSPL$^4$GPA (SEQ ID NO:)

$\mathbf{M^1}$APTPL$^4$GPA (SEQ ID NO:)

### *3.2 Internal Mutation Site 1*

**[0521]** In these mutants, the N-terminus of a wild-type GCSF, **$\mathbf{M^1}$TPLGP** (SEQ ID NO:8), is replaced with **$\mathbf{M^1TP}X_nB_oO_r\mathbf{P}$**. Wherein n, o and r are integers slected from 0 to 3, and at least one of X, B and O is Thr or Ser. When more than one of X, B and O is Thr or Ser, the identity of these moieties is independently selected. Where they appear, superscripts denote the position of the amino acid in the wild-type starting sequence.

**[0522]** Preferred mutations include:

**$\mathbf{M^1}$**TPTLGP (SEQ ID NO:8)
**$\mathbf{M^1}$**TPTQLGP (SEQ ID NO:8)
**$\mathbf{M^1}$**TPTSLGP (SEQ ID NO:8)
**$\mathbf{M^1}$**TPTQGP (SEQ ID NO:8)
**$\mathbf{M^1}$**TPTSSP (SEQ ID NO:8)
M$^1$TPQTP (SEQ ID NO:8)
M$^1$TPTGP (SEQ ID NO:8)
M$^1$TPLTP (SEQ ID NO:8)
M$^1$TPNTGP (SEQ ID NO:8)
M$^1$TPVTP (SEQ ID NO:8)
M$^1$TPMVTP (SEQ ID NO:8)
MT$^1$P$^2$TQGL$^3$G$^4$P$^5$A$^6$S$^7$ (SEQ ID NO:8)

### *3.3 Internal Mutation Site 2*

**[0523]** This mutation is made for the purpose of maintaining G-CSF activity. In these mutants, the amino acid sequence containing H$^{53}$, **LGH$^{53}$SLGI** (SEQ ID NO: ) is mutated to **LG**H$^{53}$B$_o$**LGI**, where Θ is H, S, R,E or Y, and B is either Thr or Ser.

**[0524]** Preferred examples include:

LGHTLGI
LGSSLGI
LGYSLGI
LGESLGI
LGSTLGI

*3.4 Internal Mutation Site 3*

**[0525]** In this type of mutant, the amino acid sequence encompassing $P^{129}$, **$P^{129}$ALQPT (SEQ ID NO: )**, is mutated to **$P^{129}Z_mJ_qO_rX_nPT$**, wherein Z, J, O and X are independently selected from Thr or Ser, **and m, q, r, and n are integers slected from 0 to 3.**

**[0526]** Preferred examples include:

$P^{129}$TLGPT

$P^{129}$TQGPT

$P^{129}$TSSPT

$P^{129}$TQGAPT

$P^{129}$NTGPT

$P^{129}$ALTPT

$P^{129}$MVTPT

$P^{129}$ASSTPT

$P^{129}$TTQP

$P^{129}$NTLP

$P^{129}$TLQP

MAP$^{129}$ATQPTQGAM

MP$^{129}$ATTQPTQGAM

*3.5 Internal Mutation Site 4*

**[0527]** In this type of mutant, the amino acid sequence surrounding $P^{61}$, LGIPWAP$^{61}$LSSC (SEQ ID NO:), is replaced with $PZ_mU_sJ_qP^{61}O_rX_nB_oC$, wherein m, s, q, r, n, and o are integers slected from 0 to 3, and at least one of Z, J, O, X, B and U is selected as either Thr or Ser. When more than one of Z, J, O X, B and U is Thr or Ser, each is independently selected

**[0528]** Preferred examples include:

$P^{61}$TSSC

$P^{61}$TSSAC

LGIPTA $P^{61}$LSSC

LGIPTQ $P^{61}$LSSC

LGIPTQG $P^{61}$LSSC

LGIPQT $P^{61}$LSSC

LGIPTS $P^{61}$LSSC

LGIPTS $P^{61}$LSSC

LGIPTQ$P^{61}$LSSC

LGTPWA$P^{61}$LSSC

LGTPFA $P^{61}$LSSC

$P^{61}$FTP

SLGA$P^{58}$TA$P^{61}$LSS

*3.6 C-terminal Mutations*

**[0529]** In this type of mutant, the amino acid sequence at the C-terminus of a wild-type G-CSF, RHLAQ$P^{175}$ (SEQ ID NO: ) is replaced with $Ø_aG_pJ_qO_rP^{175}X_nB_oZ_mU_sΨ_t$, wherein a, p, q, r, n, o, m, s, and t are integers slected from 0 to 3, and at least one of Z, U, O, J, G, Ø, B and X is Thr or Ser and when more than one of Z, U, O, J, G, Ø, B and X are Thr or Ser, they are independently selected. Ø is optionally R, and G is optionally H. The symbol Ψ represents any uncharged amino acid residue or E (glutamate).

**[0530]** Preferred examples include:

RHLAQT$P^{175}$

RHLAGQT$P^{175}$

Q$P^{175}$TQGAMP

RHLAQT$P^{175}$AM

Q$P^{175}$TSSAP

Q$P^{175}$TSSAP

Q$P^{175}$TQGAMP

Q$P^{175}$TQGAM

Q$P^{175}$TQGA

Q$P^{175}$TVM

Q$P^{175}$NTGP

Q$P^{175}$QTLP

*3.7 Internal Mutations surrounding $P^{133}$*

**[0531]** Additional G-CSF mutants include those with internal mutations surrounding the amino acid $P^{133}$. Examples include:

$P^{133}TQTAMP^{139}$

$P^{133}TQGTMP$

$P^{133}TQGTNP$

$P^{133}TQGTLP$

$PALQP^{133}TQTAMPA$

**EXAMPLE 4**

**[0532]** Mutations in the amino acid sequence of granulocyte colony stimulating factor (G-CSF) can introduce additional sites for O-linked glycosylation, such that the protein may be modified at these sites using the method of the present invention. This example sets forth selected representative mutants of the invention.

*4.1 G-CSF (wild type 178 aa variant)*

mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklvseca tyklchpeel vllghslgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr aggvlvashl qsflevsyrv lrhlaqp (SEQ ID NO:1)

*4.2 G-CSF (wild type 175 aa variant)*

mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklca tyklchpeel vllghslgip waplsscpsq alqlagclsq lhsglflyqg llqalegisp elgptldtlq ldvadfatti wqqmeelgma palqptqgam pafasafqrr aggvlvashl qsflevsyrv lrhlaqp **(SEQ ID NO:3)**

*4.9 G-CSF Mutant 1 (Amino Terminal mutation)*

miatplgpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll ghslgipwap lsscpsqalq lagclsqlhs glflyqgllq alegispelg ptldtlqldv adfattiwqq meelgmapal qptqgampaf asafqrragg vlvashlqsf levsyrvlrh laqp

*4.10G-CSF Mutant 2 (Amino Terminal mutation)*

mgvtetplgpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll ghslgipwap lsscpsqalq lagclsqlhs glflyqgllq alegispelg ptldtlqldv adfattiwqq meelgmapal qptqgampaf asafqrragg vlvashlqsf levsyrvlrh laqp

*4.11 G-CSF Mutant 3 (Amino Terminal mutation)*

maptplgpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll ghslgipwap lsscpsqalq lagclsqlhs glflyqgllq alegispelg ptldtlqldv adfattiwqq meelgmapal qptqgampaf asafqrragg vlvashlqsf levsyrvlrh laqp

*4.12 G-CSF Mutant 4 (Site 1)*

**mtp$^{3}$tqgl**gpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll ghslgipwap lsscpsqalq lagclsqlhs glflyqgllq alegispelg ptldtlqldv adfattiwqq meelgmapal qptqgampaf asafqrragg vlvashlqsf levsyrvlrh laqp

*4.13 G-CSF Mutant 5 (Site 3)*

Mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll ghslgipwap lsscpsqalq lagclsqlhs glflyqgllq alegispelg ptldtlqldv adfattiwqq meelgmap$^{129}$**at qp**tqgampaf asafqrragg vlvashlqsf levsyrvlrh laqp

*4.14 G-CSF Mutant 6 (Site 4)*

Mtplgpasslp qsfllkcleq vrkiqgdgaa lqeklcatyk lchpeelvll ghslgip$^{58}$ftp lsscpsqalq lagclsqlhs glflyqgllq alegispelg ptldtlqldv adfattiwqq meelgmapaL qptqgampaf asafqrragg vlvashlqsf levsyrvlrh laqp

## EXAMPLE 5

### GlycoPEGylation of G-CSF produced in CHO cells

***5a. Preparation of Asialo-Granulocyte-Colony Stimulation Factor (G-CSF)***

**[0533]** G-CSF produced in CHO cells was 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 was 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 was diluted with the appropriate buffer and a IEF gel performed. The reaction mixture was 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 was centrifuged at 10,000 rpm and the supernatant was collected. The beads were 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 were pooled. The supernatant was 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 was 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 were dialyzed against water and analyzed by MALDI-TOF MS.

***5b. Preparation of G-CSF-(alpha2,3)-Sialyl-PEG***

**[0534]** 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 was 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 was 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 was quantitated using an in-line fluorescent detector. After 2 days, the reaction mixture was 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 was analyzed using SDS-PAGE and IEF analysis according to the procedures and reagents supplied by Invitrogen. Samples of native and PEGylated G-CSF were dialyzed against water and analyzed by MALDI-TOF MS.

***5c. Preparation of G-CSF-(alpha2,8)-Sialyl-PEG***

**[0535]** G-CSF produced in CHO cells, which contains an alpha 2,3-sialylated O-linked glycan, 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 was 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 was 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 was quantitated using an in-line fluorescent detector. After 2 days, the reaction mixture was 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 was analyzed using SDS-PAGE and IEF analysis according to the procedures and reagents supplied by

Invitrogen. Samples of native and PEGylated G-CSF were dialyzed against water and analyzed by MALDI-TOF MS.

### *5d. Preparation of G-CSF-(alpha 2,6)-Sialyl-PEG*

**[0536]** G-CSF, containing only O-linked GalNAc, 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 was 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 was 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 was quantitated using an in-line fluorescent detector. After 2 days, the reaction mixture was 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 was analyzed using SDS-PAGE and IEF analysis according to the procedures and reagents supplied by Invitrogen. Samples of native and PEGylated G-CSF were dialyzed against water and analyzed by MALDI-TOF MS.

**[0537]** 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 was the first demonstration of glycoPEGylation of an O-linked glycan.

## EXAMPLE 6

### *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 NaC1, 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 7**

**The Two Enzyme Method in Two Pots**

**[0538]** 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.

*7a. Preparation of G-CSF-GalNAc (pH 6.2) from G-CSF and UDP-GalNAc using GalNAc-T2.*

**[0539]** G-CSF (960 μg) 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.

**[0540]** 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.

*7b. Preparation of G-CSF-GalNAc-SA-PEG using purified G-CSF-GalNAc, CMP-SA-PEG (20KDa) and mouse ST6GalNAc-TI (pH 6.2).*

**[0541]** 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 8**

**One Pot Method to Make G-CSF-GalNAc -SA-PEG with Simultaneous Addition of Enzymes**

**[0542]** The following example illustrates the preparation of G-CSF-GalNAc -SA-PEG in one pot using simultaneous addition of enzymes

***8a. One Pot process using mouse ST6GalNAc-I (pH 6.0).***

**[0543]** 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.

***8b. One pot process using chicken ST6GalNAc-I (pH 6.0).***

**[0544]** 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 9**

**One Pot Method to Make G-CSF-GalNAc-Gal-SA-PEG with Sequential Addition of Enzymes**

**[0545]** The following example illustrates a method for making G-CSF-GalNAc-Gal-SA-PEG in one pot with sequential addition of enzymes.

**9.1** ***Starting from GalNAc-G-CSF***

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

**[0546]** 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***

**[0547]** 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 10**

**One Pot Method to Make G-CSF-GalNAc-Gal-SA-PEG with Simultaneous Addition of Enzymes**

***a. Starting from G-CSF.***

**[0548]** 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.

**EXAMPLE 11**

**[0549]** The following Example illustrates an alternative enzymatic method to obtain large quantities of GlycoPEGylated G-CSF.

**[0550]** Granulocyte Colony Stimulating Factor (G-CSF) protein was expressed in *E. coli* and refolded from inclusion bodies as disclosed in Example X (above).

***11a. Priming the reaction by addition of GalNAc:***

**[0551]** GalNAc-ylation of G-CSF was carried out at 33 °C in 50 mM Bis-Tris pH 6.5 buffer containing 1 mM $MnCl_2$ using refolded GalNAcT2 in the presence UDP-GalNAc. This step primes the reaction enabling both GalNAc transferase and sialyltransferase to work together in subsequent steps to very efficiently produce maximum amount of GCSF-PEG in a short period of time.

***11b. PEGylation process:***

**[0552]** PEGylation was started 2 (+/- 1) hour after GalNAc-ylation by directly adding CMP-SA-PEG (20K) and ST6GalNAcI (chicken or human) to the priming reaction. This step produces substrate (GCSF-O-GalNAc) for the sialyltransferases to drive the reaction faster in a shorter period of time than can be achieved in a two step reaction wherein the GCSF-O-GalNAc is first purified from the UDP-GalNAc and other reaction components (*see e.g.*, Example X, above). Furthermore the primed one pot reaction produces a higher yield of product than does a one pot reaction in which all components are added simultaneously.

**[0553]** Indeed, comparison of several types of one pot reactions shows that when all the components were added simultaneously and incubated for 23 hours, the GCSF-PEG produced was 77 %. In contrast, when addition of all the enzymes required for the PEGylation reaction was preceded by the 2 hr GalNAc-ylation step described above, product yield was 85 %. Therefore, the sequential addition of reaction components resulted in a 10 % higher yield than was obtained when all reaction components are added simultaneously.

**EXAMPLE 12**

**[0554]** This Example describes the results of O-linked GalNAc-ylation of six mutant G-CSF proteins.

***12.1. GalNAcsylation of mutant G-CSF protein:***

**[0555]** All the sequences of mutant G-CSF proteins are listed below. Having these proteins, O-linked glycosylation was examined. Under the same condition for glycosylation of native G-CSF, GalNAc-$T_2$ (BV) was used *in vitro* with UDP-GalNAc in 25 mM MES buffer ( pH 6.0 ). MALDI was used to monitor the reaction. Measurement of increasing molecular weight of proteins provided GalNAc addition number. For one addition of GalNAc, increased molecular weight should be 203 Da. Based on MALDI results, we found that mutant G-CSF-2, -3, -4, accepted one GalNAc; and mutant G-CSF-5 some addition was also observed, and mutant G-CSF-1 accepted two GalNAcs, forming MAPT-G-CSF$(GalNAc)_2$ ( Molecular weight increasing from 18965 to 19369 Daltons).

**Table X. GalNAc addition of Mutant G-CSF ( MW measured by MALDI)**

| **Peptide** | **MW**(Intact material) | MW (GalNAc- | Number of GalNAc |
|---|---|---|---|

| | | adduct) | addition |
|---|---|---|---|
| MutantG-CSF-1 (MAPT-G-CSF) | 18965 | 19369 | 2 |
| MutantG-CSF-2 | 18766 | 19029 | 1 |
| MutantG-CSF-3 | 18822 | 19026 | 1 |
| MutantG-CSF-4 | 19369 | 19574 | 1 |
| MutantG-CSF-5 | 18957 | 18853 | 1 |
| MutantG-CSF-6 | | | NT |
| Native G-CSF | 18800 | 19023 | 1 |

**[0556]** Peptide mapping and N-terminal analysis were used for determination of glycosylation sites of MAPT-G-CSF-$(GalNAc)_2$. In the Glu C-digested peptide mapping a G-1+GalNAc peak was found, indicating one GalNAc was added at G-1 sequence. N-terminal Edman degdation analysis suggested the normal T was lost indicting that GalNAc was added onto T residue.

*12.2 GlycoPEGylation of mutant G-CSF sequences*

**a. GlycoPEGylation of mutant G-CSF sequence and buffer impact on the glycoPEGylation of MAPT-G-CSF**

**[0557]** An examination of glycoPEGylation (20K) of 5 mutants was undertaken. GlycoPEGylation was performed using three enzyme/ three nucleotides system. (UDP-GalNAc/GalNAc-$T_2$/UDP-Gal/Core GalT/CMP-SA-PEG/O-sialyltransferase) in 25 mM MES buffer (pH 6.0). All mutants can be monoglycoPEGylated. No appreciable diPEGYlation in this condition was detected by SDS-PAGE gel by Comassie Blue Stain.

**[0558]** Since MAPT-G-CSF accept two GalNAcs, this mutant should receive two PEGs in theory. Accordingly, we examined the buffer impact on the PEGylation of MAPT-G-CSF as a starting material. Four different buffers ( 1. 1M MES buffer; 2. 25 mM MES buffer (pH 6.0); 3. 50 mM Bis-tris buffer(pH 6.0); 4. 1M HEPS buffer (pH 7.4) were investigated for this reaction. It was found that MAPT-G-CSF can be PEGylated in all of the buffer system tested. However, monoPEGylation product was still a major one. In case 1M MES and 1 M HEPS buffer were used, some diPEGYlation product was formed, indicating that high concentration buffer improves the glycoPEGylation .

**b. Comparision of GlycoPEGylation efficiency by forming MAPT-G-CSF(GalNAc-SA-PEG)$_2$ and MAPT-G-CSF(GalNAc-Gal-SA-PEG)$_2$**

**[0559]** In order to see glycoPEGylation efficiency of Muant G-CSF-1 catalyzed by different enzymes, two enzymes ( $St_6$GalNAcI and O-siayltransferase) were examined for sialylPEGylation. Accordingly, MAPT-G-CSF was converted into MAPT-G-CSF(GalNAc)$_2$ and MAPT-G-CSF(GalNAc-Gal)$_2$ for siaylPEGylation. The former was treated with CMP-SA-lys-PEG(20K)/ St6GalNAc I and the latter was treated with CMP-SA-PEG(20K)/O-sialyltransferase. Both reactions were performed in 25 mM MES buffer(pH 6.0) and 1mg/ml protein concentration. The PEGylation efficiency can be seen in SDS-Page gel. It appeared that two enzymes were pretty similar in glycoPEGylation of this protein using CMP-SA-Lys-PEG (20KDa) under the condition tested.

**c. High protein concentration led to formation of MAPT-G-CSF((GalNAc-SA-PEG(20KDa))$_2$ as a major product.**

**[0560]** After examining the impact of enzyme and buffer on glycoPEGylation, as described above, the influence of protein concentration on the PEGylation by combining with a factor of high buffer concentration using $ST_6$GalNAcI as GlycoPEGylation enzyme. So we applied UDP-GalNAc/GalNAc-$T_2$ and CMP-SA-PEG(20KDa)/St6GalNAcI for glycoPEGylation of MAPT-G-CSF using 8~10 mg/ml protein concentration for reaction in 1 M MES buffer(pH 6.0). The result suggested that under this condition, the desired diPEGylation product became the major. Over 90% conversion was also achieved by applying more CMP-SA-PEG (20K) and enzyme. PEGylated G-CSF product, MAPT-G-CSF((GalNAc-SA-PEG(20KDa))$_2$ was purified by combining SP-Sepharose and SEC purification on Supderdex 200.

***12.3. Cell proliferation activity of MAPT-G-CSF-(GalNAc-SA-PEG)$_2$***

**[0561]** Cell proliferation assay of MAPT-G-CSF-(GalNAc-SA-PEG)$_2$ with NFS-60 cell line and Tf-1 cell line were performed. The assay was performed using protein concentration between 0 ng/ml to 1000 ng/ml. MAPT-G-CSF(GalNAc-SA-PEG(20K))$_2$ was active in this assay.

**12.4 Experimental Details**

*12.4a General procedure of GalNAcsylation of Mutant G-CSF*

**[0562]** Certain volume of mutant G-CSF solution (for 100 ug protein) was buffer exchanged with MES buffer ( 25 mM + 0.005% $NaN_3$, pH 6.0 ). The final volume was adjusted to 100

ug/100 ul. To this solution was added 5 ul 100 mM $MnCl_2$ and GalNAc-$T_2$ ( 1 mU). The resulting mixture was rocked at rt for a period of time required for MALDI or QTOF analysis.

*12.4b Preparation of MAPTP-G-CSF-$(GalNAc)_2$)*

**[0563]** MAPTP-G-CSF 5.4 mg (KJ-675-159, 0.18 mg/ml, 0.053 umol) was exchanged with MES buffer ( 25 mM + 0.005% $NaN_3$, pH 6.0). The final volume was adjusted to 5.4 ml. To this solution, UDP-GalNAc (5 mg, 0.15 umol), 100 mM $MnCl_2$ 0.25 ml and GalNAc-$T_2$ (1.0U/ml, 50 ul ) were added. The resulting mixture was rocked at 32°C for 24h. $M^+$ (MALDI): 19364 (MAPT-G-CSF-$(GalNAc)_2$ verse 18951 (MVPTP-G-CSF).

*12.4c General procedure of glycoPEGylation of mutant G-CSF sequences by one-pot reaction)*

**[0564]** Mutant G-CSF 100 ug( Mutant G-CSF-1,2,3,4,5) was mixted with UDP-GalNAc ( 0.6 mg, 0.923 umol), GalNAc-$T_2$ ( 20 ul, 8 mU), UDP-Gal( 0.6 mg, 0.923 umol), Core 1 Gal T( 20 ul, 10 mU), CMP-SA-PEG(20K) (1 mg, 0.05 umol), St3GalII( 20 ul, 28 mU), 100 mM MnCl2 3 ul in 100 ul 25 mM MES buffer( pH 6.0+ 0.005% $NaN_3$). The resulting mixture was rocked at rt for 24h. GlycoPEGylation was followed by SDS-PAGE.

*12.4d Comparision of mutant G-CSF-1 glycoPEGyaltion(20KDa) in various buffer system*

**[0565]** $GalNAc_2$-MATP-G-CSF ( 54 ug ) was buffer exchanged to the following four buffer system(1. 1 M MES buffer(pH 6.0); 2. 25 mM MES buffer(pH 6.0); 3. 50 mM Bis-Tris buffer (pH 6.5); 4. 1M HEPS buffer (pH 7.4). Then CMP-SA-PEG (20K) ( 216 ug ) ST6GalNAcI( BV, 1U/mL, 2.5 ul), 100 mM $MnCl_2$ 2.5 ul were added. The resulting mixture was rocked at rt for 24 h. SDS-PAGE gel was used to follow the reaction.

*12.4e Comparision of GlycoPEGylation of MAPT-G-CSF by using $ST_6GalNAc_I$ and O-sialyltransferase (Wang787-29 and 787-40)*

*12.4e1 Using St6GalNAc I*

**[0566]** First step: 30 ml KJ-675-159 solution ( 0.18 mg/ml, 5.4 mg protein in total ) was concentrated by ultrifiltration (MWCO 5K) at 3500 g, and then buffer exchanged with 25 mM MES buffer( pH 6.0). Final volume was adjusted to 5.4 ml in a plastic tube. GalNAc-$T_2$ (1.0U /ml, 50 ul) was added, followed by addition of 0.25 mL $MnCl_2$. The resulting mixture was rocked at 32°C for 24 h. MALDI suggested that the reaction went to completion. The reaction mixture was concentrated by UF(MWCO 5K) and diluted with 25 mM MES buffer to 5 ml, then CMP-SA-PEG(20K) ( 2×25mg), $ST_6GalNAc_I$ ( BV, 1U/ml) , 100 mM $MnCl_2$

0.25 ml were added. The resulting mixture was rocked at 32°C overnight. SDS-PAGE was used for the reaction.

*12.4e2 Using O-silyltransferase($St_3Gal_{II}$):*

**[0567]** 200 ug $GalNAc_2$-MATP-G-CSF in 200 ul 25 mM MES buffer (pH 6.0) was mixed with UDP-Gal 0.6 mg and core GalT (0.2U/ml, 10 ul) and 10 ul 100 mM $MnCl_2$. The resulting mixture was rocked at 32°C for 24h. The reaction mixture was concentrated by UF (MWCO 5K) and diluted with 25 mM MES buffer to 200 ul. CMP-SA-PEG (800 ug ), $St_3GalII$ (1.0U/ml, 10 ul ) , 10 ul 100 mM $MnCl_2$ were added. The resulting mixture was rocked at rt for 24 h. The resulting mixture was rocked at 32°C overnight. SDS-PAGE gel was used to follow the reaction.

*12.4f MAPTP-G-CSF-(GalNAc-SA-PEG(20K)$_2$ from glycoPEGylation of MAPT-G-CSF-(GalNAc)$_2$ (Wang 787-42)*

**[0568]** MAPTP-G-CSF solution (540 ug) was concentrated and exchanged with 1M MES buffer (pH 6.0) and adjusted to 50 ul. Then UDP-GalNAc (100 ug, 0.15 umol, 5 eq), $GalNAcT_2$ (5.0 U/ml, 5 ul) and 100 mM $MnCl_2$ (5 ul ) was added. The resulting mixture was rocked at RT overnight. Then CMP-SA-PEG (20K) (2.16 mg, 0.108 umol) and $St_6GalNAcI$ (1.0 U/ml, 50 ul) were added. The solution was rocked at rt for 60h.. Additional CMP-SA-PEG(20K) (2.16 mg , 0.108 umol) and St6GalNAcI (1.0U/ml, 50 ul) were added, followed by slow rotation at rt for 24 h. Reaction mixture was exchanged with buffer A (25 mM NaOAc, 0.005% polysorbate 80, pH 4.5), then purified on an Amersham SP-FF (5 mL) column with an isocratic elution of 100% A for 10 minutes followed by a linear gradient of 100% A to 20 % B over 20 minutes at a flow rate of 3 mL $min^{-1}$, where B = 25 mM NaOAc, 2 M NaCl 0.005% polysorbate 80, pH 4.5. The peak at retention time 17 mins was pooled and concentrated to 0.5 ml, which was further purified on an Amersham HiLoad Superdex 200 (16 x 600 mm, 34 μm) with phosphate buffered saline, pH 5.0, 0.005% Tween80, at a flow rate of 0.4 mL $min^{-1}$. Product fractions at retention time 160 mins was pooled, concentrated to provide 30 ug of MAPT-G-CSF(GalNAc-SA-PEG(20K))$_2$( BCA). The yield was not optimized.

*12.4 g Sequences of G-CSF mutants*

**Mutant G-CSF-1:**

MAPTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSL
GIPWAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVAD

FATTIWQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVL RHLAQP **(SEQ ID NO: 9)**

**Mutant G-CSF-2:**

MTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP WAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFAT TIWQQMEELGMAPATQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHL AQP **(SEQ ID NO: )**

**Mutant G-CSF-3:**

MTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP WAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFAT TIWQQMEELGMAPALQPTQTAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHL AQP **(SEQ ID NO: )**

**Mutant G-CSF-4 (C-terminal tag):**

MTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP WAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFAT TIWQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHL AQPTQGAMP **(SEQ ID NO:8 )**

**Mutant G-CSF-5 ( N-terminal MIATP):**

MIATPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLG IPWAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADF ATTIWQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLR HLAQP **(SEQ ID NO:10 )**

**Mutant G-CSF-6 ( 177 Mer):**

MTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLVSECATYKLCHPEELVLLGHS LGIPWAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVA DFATTIWQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRV LRHLAQP **(SEQ ID NO:1)**

**Human recombinant G-CSF expressed in E coli:**

MTPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP
WAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDYADFAT
TIWQQMEELGMAPALQPT[134]QGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLR
HLAQP **(SEQ ID NO:2)**

**EXAMPLE 13**

**[0569]** The following Example illustrates preparation of a GlycoPEGylated hGH protein The wild-type hGH has no natural glycosylation site, therefore a *de novo* O-glycosylation site was engineered into a mutant hGH protein which was then be glycosylated with a GalNAc transferase and sialylPEGylated at the mutant site. Five mutant hGH proteins were designed to incorporate an O-glycosylation site at either the amino terminus or in the loop region of the protein molecule. The five mutant proteins were produced and each was tested for hGH activity in a Nb2-11 cell proliferation assay.

*13.1 Mutant hGH Amino Acid Sequences*:

**192 amino acid Wild-type pituitary derived hGH comprising an N-Terminal methionine**

MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP
QTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDS
NVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYC
FRKDMDKVETFLRIVQCRSVEGSCGF **(SEQ ID NO:)**

**191 amino acid Wild-type pituitary derived hGH lacking an N-Terminal methionine**

FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQ
TSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSN
VYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCF
RKDMDKVETFLRIVQCRSVEGSCGF **(SEQ ID NO:)**

**MVTP mutant:**

(M)VTPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFL QNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGA SDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGL LYCFRKDMDKVETFLRIVQCRSVEGSCGF **(SEQ ID NO: )**

**PTQGAMP** mutant:

MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP QTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDS NVYDLLKDLEEGIQTLMGRLEDGSPTQGAMPKQTYSKFDTNSHNDDALLKNYGLLY CFRKDMDKVETFLRIVQCRSVEGSCGF **(SEQ ID NO: )**

**TTT mutant:**

MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP QTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDS NVYDLLKDLEEGIQTLMGRLEDGSPTTTQIFKQTYSKFDTNSHNDDALLKNYGLLYC FRKDMDKVETFLRIVQCRSVEGSCGF **(SEQ ID NO: )**

**MAPT mutant:**

MAPTSSPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSF LQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYG ASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYG LLYCFRKDMDKVETFLRIVQCRSVEGSCGF **(SEQ ID NO: )**

**NTG mutant:**

MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP QTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDS NVYDLLKDLEEGIQTLMGRLEDGSPNTGQIFKQTYSKFDTNSHNDDALLKNYGLLY CFRKDMDKVETFLRIVQCRSVEGSCGF

**[0570]** The four hGH mutants were tested for the ability to act as substrates for glycosyltransferase GalNAcT2. Of the four hGH mutants, two were found to be glycosylated by GalNAcT2 by MALDI-MS analysis.

### *13.2 Preparation of hGH-(TTT)-GalNAc-SA-PEG-30KDa.*

**[0571]** For the TTT mutant, GalNAc addition gave rise to a complex mixture of unglycosylated, and 1-GalNAc and 2-GalNAc species. Peptide mapping experiments (trypsin digest) showed that the two GalNAc's were added to the T12 peptide (L129-K141) containing the TTT mutation. The (M)VTP mutant showed only a trace of GalNAc added by MALDI-MS.

**[0572]** The hGH-TTT-mutant (4.0 mL, 6.0 mg, 0.27 micromoles) was buffer exchanged twice with 15 mL of Washing Buffer (20 mM HEPES, 150 mM NaCl, 0.02% $NaN_3$, pH 7.4) and once with Reaction Buffer (20 mM HEPES, 150 mM NaCl, 5 mM $MnCl_2$, 5 mM $MgCl_2$, 0.02% $NaN_3$, pH 7.4) then concentrated to 2.0 mL using a Centricon centrifugal filter, 5 KDa MWCO.

**[0573]** The hGH-TTT mutant was combined with UDP-GalNAc (1.38 micromoles, 0.90 mg) and GalNAc-T2 (0.12 mL, 120 mU). The reaction was incubated at 32°C with gentle shaking for 19 hours. The reaction was analyzed by MALDI-MS and partial addition of GalNAc to the hGH-TTT mutant was observed (approximately 40%). CMP-SA-PEG-30K (16 mg, 0.533 micromoles) and ST6GalNAc1 (0.375 mL, 375 mU) were added to the reaction mixture to bring the total volume to 2.85 mL. The reaction was incubated at 32°C with gentle shaking for 22 h. The reaction was monitored by SDS PAGE at 0 h and 22 h. The extent of reaction was determined by SDS-PAGE gel. The product, hGH-(TTT)-GalNAc-SA-PEG-30 KDa, was purified using SP Sepharose and analyzed by SDS-PAGE. Very low yield of the desired hGH-(TTT)-GalNAc-SA-PEG-30 KDa was observed.

### *13.3 Preparation of hGH-(PTQGAMP)-GalNAc-SA-PEG-30KDa.*

**[0574]** The PTQGAMP mutant was was readily glycosylated with UDP-GalNAc and GalNAc T2, then GlycoPEGylated using CMP-SA-PEG-30KDa and ST6GalNAc1 on 10 mg scale to yield 1.45 mg of purified hGH-(PTQGAMP)-GalNAc-SA-PEG-30KDa. Peptide mapping experiments (trypsin digest) located the GalNAc on the trypsin T12 peptide (L129-K141) containing the PTQGAMP mutation.

**[0575]** The hGH-PTQGAMP-mutant (4.55 mL, 10.0 mg, 0.45 micromoles) was buffer exchanged twice with 15 mL of Washing Buffer (20 mM HEPES, 150 mM NaCl, 0.02% $NaN_3$, pH 7.4) and once with Reaction Buffer (20 mM HEPES, 150 mM NaCl, 5 mM $MnCl_2$,

5 mM $MgCl_2$, 0.02% $NaN_3$, pH 7.4) then concentrated to 3 mL using a Centricon centrifugal filter, 5 KDa MWCO.

**[0576]** The hGH-PTQGAMP mutant was combined with UDP-GalNAc (2.26 micromoles, 1.47 mg) and GalNAc-T2 (0.1 mL, 100 mU). The reaction was incubated at 32°C with gentle shaking for 22 hours. The reaction was analyzed by MALDI-MS and complete addition of GalNAc to the hGH-PTQGAMP mutant was observed. CMP-SA-PEG-30K (27 mg, 0.9 micromoles) and ST6GalNAc1 (0.350 mL, 350 mU) were added to the reaction mixture to bring the total volume to 3.4 mL. The reaction was incubated at 32°C with gentle shaking for 24 h. The reaction was monitored by SDS PAGE at 0 hours and 16.5 hours. The extent of reaction was determined by SDS-PAGE gel. The product, hGH-(PTQGAMP)-GalNAc-SA-PEG-30 KDa, was purified using SP Sepharoseand SEC (Superdex 200) chromatographyand then formulated. The final product was analyzed by MALDI, peptide map and SDS-PAGE (silver stain). Protein was determined by BCA vs. BSA standard. The overall isolated yield (1.45 mg) was 12.5 % based on protein.

## EXAMPLE 14

**[0577]** This example sets forth the preparation of a GM-CSF PEG glycoconjugate of the invention.

*14.1 Preparation of (PEG(20K)-SA-Gal-GalNAc)2-GM-CSF and PEG(20k)-SA-Gal-GalNAc-GM-CSF*

**[0578]** GM-CSF (1 mg) was dissolved in 25 mM MES buffer (1 mL) (pH 6.0, 0.005% $NaN_3$), then UDP-GalNAc (1 mg), GalNAc-$T_2$ (200 μL, 0.38 U/mL, 0.076 U), 100 mM $MnCl_2$ (80 μL) were added. The resulting mixture was incubated at room temperature for 72 h. MALDI indicated GalNAc2-GM-CSF was formed.

**[0579]** UDP-Gal (6 mg, 9.8 mmol ), core-1-Gal-$T_1$ (0.5 U/mL, 80 μL), CMP-SA-PEG (20 kilodalton) (6 mg, 0.3 μmol), α-(O)-sialyltransferase (1 U/mL, 120 μL), 100 mM $MnCl_2$ (50 μL) were added. The resulting mixture was slowly rotated at 32° C for 48 h. The reaction mixture was centrifuged at 2 rpm for 5 min. The protein solution was taken. The remain resin was mixed with 1 mL 25 mM MES buffer (pH 6.0) and vibrated for 30 sec. The suspension was concentrated in again; the protein solutions were combined and concentrated to 200 mcL. HPLC Purification provided glyco-PEG-ylated GM-CSF.

## EXAMPLE 15

**[0580]** An O-linked glycosylation site similar to that of interferon alpha-2 can be incorporated into any interferon alpha protein at the same relative position. This can be performed by aligning the amino acid sequence of interest with the IFN-alpha-2b sequence (10-20 amino acids long) and modifying the amino acid sequence to incorporate the glycosylation site. Mutation with any amino acid, deletion or insertion can be used to create the site. Exemplary mutants maintain as high an homology as possible with the IFN-alpha-2 sequence in this region with an emphasis on the T at position 106 (shown below in bold). An example of how this is performed is shown below.

**Alignments of Interferon alpha's in the NCBI Protein Database**

| GI# | AA# | AA Sequence | Name |
|---|---|---|---|
| IFN-a-2β | 1 | CVIQGVGV**T**ETPLMKEDSIL 20 | **(SEQ ID NO:X)** |
| 124449 | 98 | ..... 117 | IFN-alpha 2 (a,b,c) |
| 20178265 | 99 | ....E...E.....N..... 118 | IFN-alpha 14 |
| 124453 | 99 | ....E...E.....N..... 118 | IFN-alpha 10 |
| 585316 | 99 | ....E..ME.....N..... 118 | IFN-alpha 17 |
| 124442 | 99 | ....E...E.....N..F.. 118 | IFN-alpha 7 |
| 124438 | 99 | ....E...E.....NV.... 118 | IFN-alpha 4 |
| 417188 | 99 | ..M.E...I.S...Y..... 118 | IFN-alpha 8 |
| 20178289 | 99 | ....E...E.....NV.... 118 | IFN-alpha 21 |
| 124457 | 99 | .MM.E...ED....NV.... 118 | IFN-alpha 5 |
| 124463 | 99 | ..T.E...E.IA..N..... 118 | IFN-alpha 16 |
| 124460 | 99 | ..M.E.W.GG....N..... 118 | IFN-alpha 6 |
| 124455 | 99 | ..M.EER.G.....NA.... 118 | IFN-alpha 1/13 |

**[0581]** Glycosylation/Glyco-PEG-ylation occurs at $T^{106}$ (IFN-alpha-2). Protein numbering begins with the first amino acid after removal of the protein leader sequence of the naturally expressed pre-pro form.

**[0582]** Interferon alpha mutations to introduce O-Linked Glycosylation Sites in IFN-alpha's that lack this site.

| GI# | AA# | AA Sequence | Name |
|---|---|---|---|
| IFN-a-2β | 1 | CVIQGVGV**T**ETPLMKEDSIL 20 | **(SEQ ID NO:X)** |
| 124449 | 98 | ..... 117 | IFN-alpha 2 (a,b,c) |
| 20178265 | 99 | ....E...**T**.....N..... 118 | IFN-alpha 14 ($E^{107}T$) |
| 20178265 | 99 | ....**G**...**T**.....N..... 118 | IFN-alpha 14 ($E^{103}G$; $E^{107}T$) |
| 124453 | 99 | ....E...**T**.....N..... 118 | IFN-alpha 10 ($E^{107}T$) |
| 124453 | 99 | ....**G**...**T**.....N..... 118 | IFN-alpha 10 ($E^{103}G$; $E^{107}T$) |
| 585316 | 99 | ....E..M**T**.....N..... 118 | IFN-alpha 17 ($E^{107}T$) |
| 585316 | 99 | ....E..**VT**.....N..... 118 | IFN-alpha 17 ($ME^{107}VT$) |
| 585316 | 99 | ....**G**..M**T**.....N..... 118 | IFN-alpha 17 ($E^{103}G$; $E^{107}T$) |
| 124442 | 99 | ....E...**T**.....N..F.. 118 | IFN-alpha 7 ($E^{107}T$) |
| 124442 | 99 | ....**G**...**T**.....N..F.. 118 | IFN-alpha 7 ($E^{103}G$; $E^{107}T$) |
| 124438 | 99 | ....E...**T**.....NV.... 118 | IFN-alpha 4 ($E^{107}T$) |
| 124438 | 99 | ....**G**...**T**.....NV.... 118 | IFN-alpha 4 ($E^{103}G$; $E^{107}T$) |
| 417188 | 99 | ..M.E...**T**.S...Y..... 118 | IFN-alpha 8 ($I^{107}T$) |
| 417188 | 99 | ..M.**G**...**T**.S...Y..... 118 | IFN-alpha 8 ($E^{103}G$; $I^{107}T$) |
| 20178289 | 99 | ....E...**T**.....NV.... 118 | IFN-alpha 21 ($E^{107}T$) |

20178289 99 ....**G**...**T**.....NV.... 118 IFN-alpha 21 ($E^{103}G$; $E^{107}T$)

124457 99 .MM.E...**T**D....NV.... 118 IFN-alpha 5 ($E^{107}T$)

124457 99 .MM.E...**TE**....NV.... 118 IFN-alpha 5 ($ED^{108}TE$)

124457 99 .MM.**G**...**T**D....NV.... 118 IFN-alpha 5 ($E^{103}G$; $E^{107}T$)

124463 99 ..T.E...**T**.I**P**..N..... 118 IFN-alpha 16 ($E^{107}T$;$A^{110}P$)

124463 99 ..T.E...**T**.**TP**..N..... 118 IFN-alpha 16 ($E^{107}T$;$IA^{110}TP$)

124463 99 ..T.**G**...**T**.**TP**..N..... 118 IFN-alpha 16 ($E^{103}G$;$E^{107}T$;$IA^{110}TP$)

124460 99 ..M.E.W.**T**G....N..... 118 IFN-alpha 6 ($G^{107}T$)

124460 99 ..M.E.**G**.**T**G....N..... 118 IFN-alpha 6 ($W^{105}G$;$G^{107}T$)

124460 99 ..M.**G**.**G**.**TE**....N..... 118 IFN-alpha 6 (E103G;$W^{105}G$;$GG^{108}TE$)

124455 99 ..M.EER.**T**.....NA.... 118 IFN-alpha 1/13 ($G^{107}T$)

124455 99 ..M.EE**G**.**T**.....NA.... 118 IFN-alpha 1/13 ($R^{105}G$;$G^{107}T$)

124455 99 ..M.**GVG**.**T**.....NA.... 118 IFN-alpha 1/13 ($EER^{105}GVG$;$G^{107}T$)

The GI numbers in the above table, except the first number 124449, refer to those of the unmodified wild-type proteins.

**[0583]** The O-linked glycosylation site can be created in any interferon alpha isoform by placing a T or S at the appropriate amino acid site as shown above. The substitution is T as shown in the above table. The amino acid sequences between the various interferon alpha forms are similar. Any amino acid mutation, insertion, deletion can be made in this region as

long as the T or S is at the appropriate position for glycosylation/glyco-PEG-ylation relative to $P^{109}$ (IFN-alpha-2) in the alignment sequence shown above.

**[0584]** While this invention has been disclosed with reference to specific embodiments, it is apparent that other embodiments and variations of this invention may be devised by others skilled in the art without departing from the true spirit and scope of the invention.

**[0585]** All patents, patent applications, and other publications cited in this application are incorporated by reference in the entirety.

## WHAT IS CLAIMED IS:

1. An isolated polypeptide comprising a mutant peptide sequence, wherein the mutant peptide sequence encodes an O-linked glycosylation site that does not exist in a wild-type polypeptide corresponding to the isolated polypeptide.

2. The polypeptide of claim 1, wherein the polypeptide is a G-CSF polypeptide.

3. The polypeptide of claim 2, wherein the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $M^1X_nTPLGP$ or $M^1B_oPZ_mX_nTPLGP$, and wherein

the superscript denotes the position of the amino acid in the wild-type G-CSF amino acid sequence (SEQ ID NO:3), the subscripts n and m are integers selected from 0 to 3, and

at least one of X and B is Thr or Ser, and

when more than one of X and B is Thr or Ser, the identity of these moieties is independently selected, and

Z is selected from glutamate, or any uncharged amino acid.

4. The mutant G-CSF polypeptide of claim 3, wherein the mutant peptide sequence is selected from the sequences consisting of MVTPLGP, MQTPLGP, MIATPLGP), MATPLGP, MPTQGAMPLGP , MVQTPLGP, MQSTPLGP, MGQTPLGP, MAPTSSSPLGP, and MAPTPLGPA.

5. The polypeptide of claim 2, wherein the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $M^1TPX_nB_oO_rP$

wherein

the superscript denotes the position of the amino acid in SEQ ID NO:3, and

the subscripts n, o, and r are integers selected from 0 to 3, and

at least one of X, B and O is Thr or Ser, and

when more than one of X, B and O is Thr or Ser, the identity of these moieties is independently selected.

6. The polypeptide of claim 5, wherein the mutant peptide sequence is selected from the sequences consisting of: MTPTLGP, MTPTQLGP, MTPTSLGP, MTPTQGP, MTPTSSP, $M^1$TPQTP, $M^1$TPTGP, $M^1$TPLTP, $M^1$TPNTGP, MTPLGP (G-CSF mut #4), $M^1$TPVTP, $M^1$TPMVTP, and $MT^1P^2TQGL^3G^4P^5A^6S^7$.

7. The polypeptide of claim 2, wherein the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $LGX^{53}B_oLGI$

wherein

the superscript denotes the position of the amino acid in the wild type G-CSF amino acid sequence (SEQ ID NO: 3), and

X is histidine, serine, arginine, glutamic acid or tyrosine, and

B is either threonine or serine, and

o is an integer from 0 to 3.

8. The polypeptide of claim 7, wherein the mutant peptide sequence is selected from the sequences consisting of: LGHTLGI, LGSSLGI, LGYSLGI, LGESLGI, and LGSTLGI.

9. The polypeptide of claim 2, wherein the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $P^{129}Z_mJ_qO_rX_nPT$

wherein

the superscript denotes the position of the amino acid in the wild type G-CSF amino acid sequence (SEQ ID NO. 3),

Z, J, O and X are independently selected from Thr or Ser, and

m, q, r, and n are integers independently selected from 0 to 3..

10. The polypeptide of claim 9, wherein the mutant peptide sequence is selected from the sequences consisting of: $P^{129}$ATQPT, $P^{129}$TLGPT, $P^{129}$TQGPT, $P^{129}$TSSPT, $P^{129}$TQGAPT, $P^{129}$NTGPT, PALQPTQT, $P^{129}$ALTPT, $P^{129}$MVTPT, $P^{129}$ASSTPT, $P^{129}$TTQP, $P^{129}$NTLP, $P^{129}$TLQP, MAP$^{129}$ATQPTQGAM, and MP$^{129}$ATTQPTQGAM.

11. The polypeptide of claim 2, wherein the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $PZ_mU_sJ_qP^{61}O_rX_nB_oC$

wherein

the superscript denotes the position of the amino acid in the wild type G-CSF amino acid sequence (SEQ ID NO. 3),
at least one of Z, J, O, and U is selected from threonine or serine, and
when more than one of Z, J, O and U is threonine or serine, each is independently selected, and
m, s, q, r, n, and o are integers independently selected from 0 to 3.

12. The polypeptide of claim 11, wherein the mutant peptide sequence is selected from the sequences consisting of: $P^{61}$TSSC, $P^{61}$TSSAC, LGIPTA $P^{61}$LSSC, LGIPTQ $P^{61}$LSSC, LGIPTQG $P^{61}$LSSC, LGIPQT $P^{61}$LSSC, LGIPTS $P^{61}$LSSC, LGIPTS $P^{61}$LSSC, LGIPTQP$^{61}$LSSC, LGTPWAP$^{61}$LSSC, LGTPFA $P^{61}$LSSC, $P^{61}$FTP, and SLGAP$^{58}$TAP$^{61}$LSS.

13. The polypeptide of claim 2, wherein the G-CSF polypeptide comprises a mutant peptide sequence with the formula of $Ø_aG_pJ_qO_rP^{175}X_nB_oZ_mU_s\Psi_t$
wherein
the superscript denotes the position of the amino acid in the wild type G-CSF amino acid sequence (SEQ ID NO. 3),
at least one of Z, U, O, J, G, Ø, B and X is threonine or serine, and when more than one of Z, U, O, J, G, Ø, B and X are threonine or serine, they are independently selected; Ø is optionally R, and G is optionally H; the symbol Ψ represents any uncharged amino acid residue or glutamate and
a, p, q, r, n, o, m, s, and t are integers independently selected from 0 to 3..

14. The polypeptide of claim 13, wherein the mutant peptide sequence is selected from the sequences consisting of: RHLAQTP$^{175}$, RHLAGQTP$^{175}$, QP$^{175}$TQGAMP, RHLAQTP$^{175}$AM, QP$^{175}$TSSAP, QP$^{175}$TSSAP, QP$^{175}$TQGAMP, QP$^{175}$TQGAM, QP$^{175}$TQGA, QP$^{175}$TVM, QP$^{175}$NTGP, and QP$^{175}$QTLP.

15. The polypeptide of claim 2, comprises a mutant peptide sequence selected from the sequences P$^{133}$TQTAMP$^{139}$, P$^{133}$TQGTMP, P$^{133}$TQGTNP, P$^{133}$TQGTLP, and PALQP$^{133}$TQTAMPA.

16. The polypeptide of claim 1, wherein the polypeptide is an hGH polypeptide.

17. The polypeptide of claim 16, wherein the mutant peptide sequence comprises a sequence selected from: $M^{1}APTSSPTIPL^{7}SR^{9}$ and $DGSP^{133}NTGQIFK^{140}$

18. The polypeptide of claim 15, wherein the hGH polypeptide comprises a mutant peptide sequence with a formula of $P^{133}JXBOZUK^{140}QTYS$, and

wherein

the superscript denotes the position of the amino acid in the wild type hGH amino acid sequence (SEQ ID NO: 20), and

J is selected from threonine and arginine;

X is selected from alanine, glutamine, isoleucine, and threonine;

B is selected from glycine, alanine, leucine, valine, asparagine, glutamine, and threonine;

O is selected from tyrosine, serine, alanine, and threonine;

Z is selected from isoleucine and methionine; and

U is selcted from phenylalanine and proline.

19. The polypeptide of claim **18**, wherein the mutant peptide sequence is selected from the group consisting of PTTGQIFK, PTTAQIFK, PTTLQIFK, PTTLYVFK, PTTVQIFK, PTTVSIFK, PTTNQIFK, PTTQQIFK, PTATQIFK, PTQGQIFK, PTQGAIFK, PTQGAMFK, PTIGQIFK, PTINQIFK, PTINTIFK, PTILQIFK, PTIVQIFK, PTIQQIFK, PTIAQIFK, $P^{133}TTTQIFK^{140}QTYS$, and $P^{133}TQGAMPK^{140}QTYS$.

20. The polypeptide of claim 15, wherein the hGH polypeptide comprises a mutant peptide sequence with a formula of $P^{133}RTGQIPTQBYS$

wherein

the superscript denotes the position of the amino acid in the wild type hGH amino acid sequence (SEQ ID NO:20), and

B is selected from alanine and threonine.

21. The polypeptide of claim 20, wherein the mutant peptide sequence is selected from the group consisting of PRTGQIPTQTYS and PRTGQIPTQAYS.

22. The polypeptide of claim 16, wherein the hGH polypeptide comprises a mutant peptide sequence with a formula of $L^{128}XTBOP^{133}UTG$

wherein

superscripts denote the position of the amino acid in the wild-type hGH amino acid sequence; and wherein

X is selected from glutamic acid, valine and alanine;

B is selcted from glutamine, glutamic acid, and glycine;

O is selcted from serine and threonine; and

U is selected from arginine, serine, alanine and leucine

23. The mutant hGH polypeptide of claim 22, wherein the mutant peptide sequence **is** selected from the group consisting of: LETQSP$^{133}$RTG, LETQSP$^{133}$STG, LETQSP$^{133}$ATG, LETQSP$^{133}$LTG, LETETP$^{133}$R, LETETP$^{133}$A, LVTQSP$^{133}$RTG, LVTETP$^{133}$RTG, LVTETP$^{133}$ATG, and LATGSP$^{133}$RTG.

24. The polypeptide of claim 16, wherein the hGH polypeptide comprises a mutant peptide sequence with a formula of $M^1BPTX_nZ_mOPLSRL$

wherein

wherein the superscript denotes the position of the amino acid in the wild type hGH amino acid sequence (SEQ ID NO:19); and

B is selected from phenylalanine, valine and alanine or a combination thereof;

X is selected from glutamate, valine and proline

Z is threonine;

O is selected from leucine and isoleucine; and

when X is proline, Z is threonine; and

wherein

n and m are integers selected from 0 and 2.

25. The polypeptide of claim 24, wherein the mutant peptide sequence is selected from the group consisting of $M^1$FPTE IPLSRL, $M^1$FPTV LPLSRL, and $M^1$APTPTIPLSRL.

**26.** The polypeptide of claim 24, wherein the mutant peptide sequence **is** $M^1$VTPTIPLSRL, wherein the superscript 1, denotes the first position amino acid in the wild type hGH amino acid sequence (SEQ ID NO:19)

27. The polypeptide of claim 15, wherein the mutant peptide sequence **is** selected from the group consisting of: LEDGSPTTGQIFKQTYS,

LEDGSPTTAQIFKQTYS, LEDGSPTATQIFKQTYS, LEDGSPTQGAMFKQTYS, LEDGSPTQGAIFKQTYS, LEDGSPTQGQIFKQTYS, LEDGSPTTLYVFKQTYS, LEDGSPTINTIFKQTYS, LEDGSPTTVSIFKQTYS, LEDGSPRTGQIPTQTYS, LEDGSPRTGQIPTQAYS, LEDGSPTTLQIFKQTYS, LETETPRTGQIFKQTYS, LVTETPRTGQIFKQTYS, LETQSPRTGQIFKQTYS, LVTQSPRTGQIFKQTYS, LVTETPATGQIFKQTYS, LEDGSPTQGAMPKQTYS, and LEDGSPTTTQIFKQTYS.

28. The polypeptide of claim 1, wherein the polypeptide is an IFN alpha polypeptide.

29. The polypeptide of claim **28**, wherein wherein the INF alpha polypeptide has a peptide sequence comprising a mutant amino acid sequence, and the peptide sequence corresponds to a region of INF alpha 2 having a sequence as shown in SEQ NO:22, and wherein the mutant amino acid sequence contains a mutation to a threonine or serine amino acid at a position corresponding to $T^{106}$ of INF alpha 2.

30. The polypeptide of claim **29,** wherein the IFN alpha polypeptide is selected from the group consisting of IFN alpha, IFN alpha 4, IFN alpha 5, IFN alpha 6, IFN alpha 7, IFN alpha 8, IFN alpha 10, IFN alpha 14, IFN alpha 16, IFN alpha 17, and IFN alpha 21.

31. The polypeptide of claim **30,** wherein the IFN alpha polypeptide is an IFN alpha polypeptide comprising a mutant amino acid sequence selected from the group consisting of:

$^{99}$CVMQEERVTETPLMNADSIL$^{118}$, $^{99}$CVMQEEGVTETPLMNADSIL$^{118}$, and $^{99}$CVMQGVGVTETPLMNADSIL$^{118}$.

32. The polypeptide of claim **30,** wherein the IFN alpha polypeptide is an IFN alpha 4 polypeptide comprising a mutant amino acid sequence selected from the group consisting of:

$^{99}$CVIQEVGVTETPLMNVDSIL$^{118}$, and $^{99}$CVIQGVGVTETPLMKEDSIL$^{118}$.

33. The polypeptide of claim **30,** wherein the IFN alpha polypeptide is an IFN alpha 5 polypeptide comprising a mutant amino acid sequence selected from the group consisting of:

$^{99}$CMMQEVGVTDTPLMNVDSIL$^{118}$, $^{99}$CMMQEVGVTETPLMNVDSIL$^{118}$ and $^{99}$CMMQGVGVTDTPLMNVDSIL$^{118}$.

34. The polypeptide of claim **30,** wherein the IFN alpha polypeptide is an IFN alpha 6 polypeptide comprising a mutant amino acid sequence selected from the group consisting of:

$^{99}$CVMQEVWVTGTPLMNEDSIL$^{118}$, $^{99}$CVMQEVGVTGTPLMNEDSIL$^{118}$, and $^{99}$CVMQGVGVTETPLMNEDSIL$^{118.}$

35. The polypeptide of claim **30,** wherein the IFN alpha polypeptide is an IFN alpha 7 polypeptide comprising a mutant amino acid sequence selected from the group consisting of:

$^{99}$CVIQEVGVTETPLMNEDFIL$^{118}$, and $^{99}$CVIQGVGVTETPLMNEDFIL$^{118}$.

36. The polypeptide of claim **30,** wherein the IFN alpha polypeptide is an IFN alpha 8 polypeptide comprising a mutant amino acid sequence selected from the group consisting of:

$^{99}$CVMQEVGVTESPLMYEDSIL$^{118}$, and $^{99}$CVMQGVGVTESPLMYEDSIL$^{118}$.

37. The polypeptide of claim **30,** wherein the IFN alpha polypeptide is an IFN alpha 10 polypeptide comprising a mutant amino acid sequence selected from the group consisting of:

$^{99}$CVIQEVGVTETPLMNEDSIL$^{118}$, and $^{99}$CVIQGVGVTETPLMNEDSIL$^{118}$.

38. The polypeptide of claim **30,** wherein the IFN alpha polypeptide is an IFN alpha 14 polypeptide comprising a mutant amino acid sequence selected from the group consisting of:

$^{99}$CVIQEVGVTETPLMNEDSIL$^{118}$, and $^{99}$CVIQGVGVTETPLMNEDSIL$^{118}$.

39. The polypeptide of claim **30,** wherein the IFN alpha polypeptide is an IFN alpha 16 polypeptide comprising a mutant amino acid sequence selected from the group consisting of:

$^{99}$CVTQEVGVTEIPLMNEDSIL$^{118}$, $^{99}$CVTQEVGVTETPLMNEDSIL$^{118}$, and $^{99}$CVTQGVGVTETPLMNEDSIL$^{118}$.

40. The polypeptide of claim **30,** wherein the IFN alpha polypeptide is an IFN alpha 17 polypeptide comprising a mutant amino acid sequence selected from the group consisting of:

$^{99}$CVIQEVGMTETPLMNEDSIL$^{118}$, $^{99}$CVIQEVGVTETPLMNEDSIL$^{118}$, and $^{99}$CVIQGVGMTETPLMNEDSIL$^{118}$.

41. The polypeptide of claim **30,** wherein the IFN alpha polypeptide is an IFN alpha 21 polypeptide comprising a mutant amino acid sequence selected from the group consisting of:

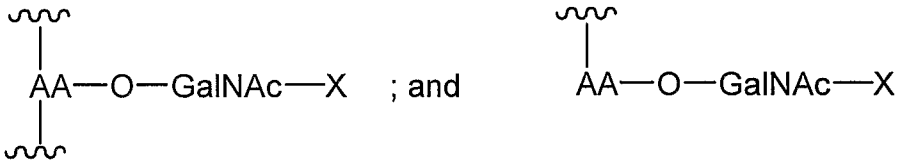
$^{99}$CVIQEVGVTETPLMNVDSIL$^{118}$, and $^{99}$CVIQGVGVTETPLMNVDSIL$^{118}$.

42. An isolated nucleic acid encoding the polypeptide of claim 1**.**

43. An expression cassette comprising the nucleic acid of claim **42.**

**44.** A cell comprising the nucleic acid of claim **42.**

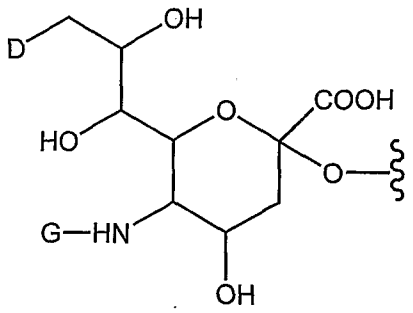
45. The polypeptide of claim 1**,** having a formula selected from:




wherein AA is an amino acid a side chain that comprises a hydroxyl moiety that is within the mutant peptide sequence; and X a modifying group or a saccharyl moiety.

46. The polypeptide according to claim **45,** wherein X comprises a group selected from sialyl, galactosyl and Gal-Sia moieties, wherein at least one of said sialyl, galactosyl and Gal-Sia comprises a modifying group.

47. The polypeptide according to claim **45,** wherein X comprises the moiety:




wherein

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

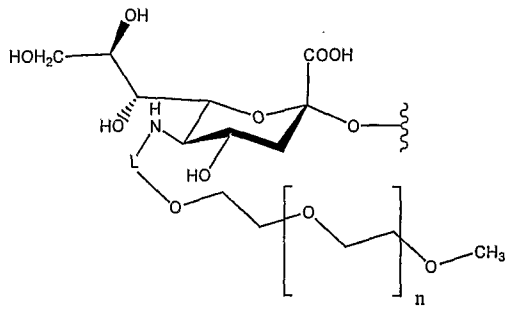
G is a member selected from $R^1$-L- and -C(O)($C_1$-$C_6$)alkyl;

$R^1$ 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,

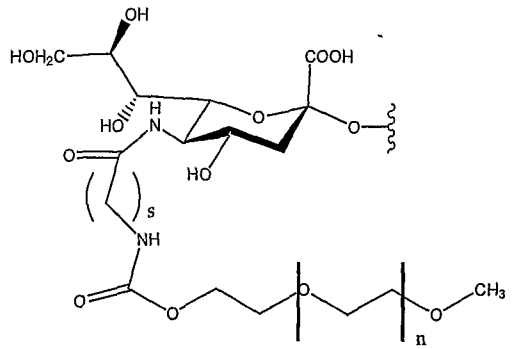
such that when D is OH, G is $R^1$-L-, and when G is –C(O)($C_1$-$C_6$)alkyl, D is $R^1$-L-NH-.

48. The polypeptide according to claim **45**, wherein X comprises the structure:




in which L is a substituted or unsubstituted alkyl or substituted or unsubstituted heteroalkyl group; and n is selected from the integers from 0 to about 500.

49. The polypeptide according to claim **45,** wherein X comprises the structure:



wherein s is selected from the integers from 0 to 20.

50. A method for making a glycoconjugate of the polypeptide of claim 1, comprising the steps of:

(a) recombinantly producing the polypeptide, and

(b) enzymatically glycosylating the polypeptide with a modified sugar at said O-linked glycosylation site.

51. A pharmaceutical composition of a granulocyte colony stimulating factor (G-CSF) comprising: an effective amount of the polypeptide of claim 2, wherein said polypeptide is glycoconjugated with a modified sugar.

52. The pharmaceutical composition according to claim **51**, wherein said modified sugar is modified with a member selected from poly(ethylene glycol) and methoxy-poly(ethylene glycol) (m-PEG).

53. A pharmaceutical composition of human Growth Hormone (hGH) comprising an effective amount of the polypeptide of claim 16, wherein said polypeptide is glycoconjugated with a modified sugar.

**54.** The pharmaceutical composition according to claim **53**, wherein said modified sugar is modified with a member selected from poly(ethylene glycol) and methoxy-poly(ethylene glycol) (m-PEG).

55. A pharmaceutical composition of a granulocyte macrophage colony stimulating factor (GM-CSF) comprising an effective amount of GM-CSF polypeptide comprising a mutant peptide sequence, wherein the mutant sequence comprises an O-linked glycosylation site that does not exist in **a** wild-type GM-CSF polypeptide, and wherein said polypeptidepeptide is glycoconjugated with a modified sugar.

56. The pharmaceutical composition according to claim **55**, wherein said modified sugar is modified with a member selected from poly(ethylene glycol) and methoxy-poly(ethylene glycol) (m-PEG).

57. A pharmaceutical composition of an interferon alpha-2b comprising an effective amount of the polypeptide of claim **28**, wherein said polypeptide is glycoconjugated with a modified sugar.

58. The pharmaceutical composition according to claim **57**, wherein said modified sugar is modified with a member selected from poly(ethylene glycol) and methoxy-poly(ethylene glycol) (m-PEG).

59. A method of providing G-CSF therapy to a subject in need of said therapy, said method comprising, administering to said subject an effective amount the pharmaceutical composition of claim **51**.

60. A method of providing granulocyte macrophage colony stimulating factor therapy to a subject in need of said therapy, said method comprising:

administering to said subject an effective amount the pharmaceutical composition of claim **55**.

61. A method of providing interferon therapy to a subject in need of said therapy, said method comprising:

administering to said subject an effective amount the pharmaceutical composition of claim **57**.

62. A method of providing Growth Hormone therapy to a subject in need of said therapy, said method comprising:

administering to said subject an effective amount the pharmaceutical composition of claim **53**.