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

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 April 2003 (24.04.2003)

PCT

(10) International Publication Number
WO 2003/032913 A2

(51) International Patent Classification[7]: A61K

(21) International Application Number:
PCT/US2002/032996

(22) International Filing Date: 16 October 2002 (16.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/329,650	16 October 2001 (16.10.2001)	US
10/136,841	30 April 2002 (30.04.2002)	US
10/136,639	30 April 2002 (30.04.2002)	US
60/384,452	29 May 2002 (29.05.2002)	US
60/386,019	5 June 2002 (05.06.2002)	US
60/408,816	6 September 2002 (06.09.2002)	US

(71) Applicant *(for all designated States except US)*: **SYMBIONTICS INC.** [US/US]; 4041 Forest Park Avenue, St. Louis, MO 63108 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **LEBOWITZ, Jonathan, H.** [US/US]; 1 Devondale Lane, St. Louis, MO 63131 (US).

(74) Agent: **WALLER, Patrick, R., H.**; Testa, Hurwitz & Thibeault, L.L.P., High Street Tower, 125 High Street, Boston, MA 02110 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

(48) Date of publication of this corrected version:
29 April 2004

(15) Information about Correction:
see PCT Gazette No. 18/2004 of 29 April 2004, Section II

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

(54) Title: METHODS AND COMPOSITIONS FOR TARGETING PROTEINS ACROSS THE BLOOD BRAIN BARRIER

(57) Abstract: Disclosed are methods and compositions for targeting therapeutic proteins to the brain. Methods and compositions of the invention involve associating an IGF moiety with a therapeutic protein in order to target the therapeutic protein to the brain. Soluble fusion proteins that include an IGF targeting moiety are transported to neural tissue in the brain from blood. Methods and compositions of the invention include therapeutic applications for treating lysosomal storage diseases. The invention also provides nucleic acids and cells for expressing IGF fusion proteins.

WO 2003/032913 A2

METHODS AND COMPOSITIONS FOR TARGETING PROTEINS ACROSS THE BLOOD BRAIN BARRIER

RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Serial Nos. 60/329,650, filed October 16, 2001; 10/136,639, filed April 30, 2002; 10/136,841, filed April 30, 2002; 60/384,452, filed May 29, 2002; 60/386,019, filed June 5, 2002; and 60/408,816, filed on September 6, 2002; the entire disclosures of which is incorporated herein by reference.

FIELD OF THE INVENTION

[0002] This invention provides a means for specifically delivering proteins to the brain. The ability to target proteins to the brain is of great utility in the treatment of neurological diseases. Methods and compositions of the invention are useful to target proteins to cells across the blood brain barrier, and in particular, to target proteins to the lysosomes of cells in the CNS, including neuronal cells, macrophage cells, and other cell types. Accordingly, the invention provides methods and compositions to deliver therapeutically useful proteins to treat lysosomal storage diseases ("LSDs") that affect the CNS.

BACKGROUND

[0003] The blood-brain barrier maintains a homeostatic environment in the central nervous system (CNS). The capillaries that supply the blood to the brain have tight junctions which block passage of most molecules through the capillary endothelial membranes. While the membranes do allow passage of lipid soluble materials, water soluble materials such as glucose, proteins and amino acids do not pass through the blood brain barrier. Mediated transport mechanisms exist to transport glucose and essential amino acids across the blood brain barrier. Active transport

mechanisms remove molecules which become in excess, such as potassium, from the brain. However, the blood brain barrier impedes the delivery of drugs to the CNS.

[0004] Many neurological diseases result from cellular defects in the CNS. In particular, many lysosomal storage diseases affect cells of the CNS and result in mild to serious neurological symptoms. Accordingly, the ability to deliver therapeutic compositions to the CNS is an important aspect of an effective treatment for many diseases, including many lysosomal storage diseases.

[0005] Methods have been designed to deliver needed drugs to the CNS such as direct delivery within the CNS by intrathecal delivery. However, methods are not available in the art to efficiently deliver drugs, and particularly protein-based drugs, from the blood stream to the CNS through the blood brain barrier.

[0006] Therefore, there is a need in the art for methods to deliver proteins to the brain parenchyma on the CNS side of the blood brain barrier, and in particular to deliver proteins to the lysosomes of cells in the CNS.

SUMMARY OF THE INVENTION

[0007] The present invention provides general methods and compositions for targeting compositions from the blood stream to the brain or CNS. According to the invention, an IGF moiety is used to target a molecule from the blood stream to the brain parenchyma on the other side of the blood brain barrier. Preferred molecules are therapeutic polypeptides.

[0008] Accordingly, the invention relates in one aspect to a protein including a therapeutic agent attached to an insulin-like growth factor (IGF) moiety or tag. In one embodiment, the protein is expressed as a fusion protein along with the IGF tag. In one embodiment, the IGF tag does not include IGF-II or a portion thereof. In another embodiment, the fusion protein also

includes a lysosomal targeting portion sufficiently duplicative of IGF-II such that the targeting portion binds the cation independent mannose-6-phosphate/IGF-II receptor to mediate uptake by a lysosome. In another embodiment, the fusion protein also comprises mannose-6-phosphate in order to target the protein to the lysosomes.

[0009] Preferred IGF moieties or tags are IGF-I or IGF-II tags. Most preferred IGF tags are IGF-I tags. In one aspect, the IGF tag is an intact IGF-I or IGF-II protein. Alternatively, an IGF tag is a portion of an IGF-I or IGF-II protein that is sufficient for targeting through the blood brain barrier. Preferred portions comprise at least one of the A, B, C, or D domains, or the C-terminal region or a portion thereof, of either IGF-I or IGF-II. In one embodiment, an IGF tag includes both an A and a B domain. According to the invention, the A and B domains provide core structural features of a preferred IGF moiety. The A and B domains may be linked by a linker peptide. Alternatively, the A and B domains may be provided as separate peptides that dimerize to form an IGF tag. Preferably, A and B domains from the same IGF protein are used. However, an A domain from IGF-I can be associated with a B domain from IGF-II. Similarly, an A domain from IGF-II can be associated with a B domain from IGF-I. Accordingly, compositions of the invention include chimeric IGF-I/IGF-II molecules. For example, an A domain from one IGF protein can be joined to the C and B domains of another IGF protein. Alternative combinations of A, B, and C domains are also useful. In further embodiments, the A domain of one IGF protein can be joined directly to the domain of another IGF protein, for example by using an amino acid bridge such as a two amino acid bridge.

[0010] A most preferred IGF moiety comprises an IGF-I portion selected from the group consisting of IGF-I fragments from about residue 1 to about residue 25, IGF-I fragments from about residue 25 to about residue 40, IGF-I fragments from about residue 40 to about residue 65, and IGF-I fragments from about residue 65 to about residue 70 of the IGF-I sequence shown in

Figure 1. Alternative preferred regions of IGF-I and IGF-II comprise regions of homology between IGF-I and IGF-II such as those shown in Figure 1 for human IGF-I and IGF-II. The sequences shown in Figure 1 relate to mature IGF-I and IGF-II proteins. Specific IGF variants described herein refer to the mature amino acid sequence numbering shown in Figure 1. In a further embodiment, an IGF tag comprises the C-terminal fragment of an IGF protein, for example the region C-terminal to the D domain shown in Figure 2. A preferred IGF tag includes an IGF-I C-terminal fragment. In addition, according to the invention, IGF tags include peptide tags with a sequence that is sufficiently duplicative of the IGF tags described herein to effectively target compositions of the invention to the brain parenchyma across the blood brain barrier. In some embodiments, an IGF tag includes at least one peptide sequence from an IGF-I protein and one from an IGF-II protein.

[0011] Most preferred IGF tags are based on human IGF proteins. However, IGF tags based on IGF proteins from other mammals, such as mouse, rabbit, monkey, and pig IGF proteins, are also useful according to the invention. Preferred IGF tags such as the IGF fragments, peptides, or domains described herein are between 1 and 100 amino acids long, more preferably between 10 and 50 amino acids long, and even more preferably about 25 amino acids long, and are sufficient for targeting associated peptides to the brain. Preferred IGF fragments, peptides, or domains are based on the mature IGF-I and IGF-II sequences.

[0012] IGF tags of the invention can be fused to a therapeutic peptide at its N-terminus, C-terminus, within the body of the therapeutic peptide, or a combination of the above. When an IGF moiety is fused to the N-terminus of a therapeutic protein, an IGF signal peptide is preferably included in the expression construct. However, an IGF signal peptide can also be included at the N-terminus when the IGF targeting moiety is located at the C-terminus or within the body of the therapeutic protein. In a preferred embodiment, the IGF tag is fused to the C-

terminal end of a peptide. In one embodiment, a first domain of an IGF tag is fused to a therapeutic peptide, and a second domain of the IGF tag is provided in a form that dimerizes with the first domain resulting in a protein that is targeted to the brain. For example, the therapeutic peptide can be fused to the A domain of an IGF protein, and dimerized with a B domain that is provided separately. Alternatively, the therapeutic peptide can be fused to the B domain of an IGF protein, and dimerized with an A domain that is provided separately.

[0013] The invention also relates to methods for identifying IGF-based peptide fragments that can reach neuronal tissue from blood and are useful to target an associated protein to the brain or CNS. According to the invention, the effectiveness of IGF-based tags can be assayed using methods described herein, such as localization assays based on radioactive labels or histochemical staining.

[0014] The invention also relates to a nucleic acid (e.g., a DNA molecule) encoding an IGF tag or a protein fused to an IGF tag, and to a cell (e.g., a cell cultured *in vitro* including a mammalian cell culture such as a CHO cell culture, and/or a unicellular organism such as *E. coli* or *Leishmania*) containing such a nucleic acid.

[0015] In another aspect, the invention relates to a method of producing a therapeutic agent for targeting across the blood brain barrier, and in particular to the lysosomes of cells in the CNS. The agent is produced by culturing a cell expressing a nucleic acid encoding a protein containing both a therapeutic agent and an IGF tag effective to target the protein across the blood brain barrier. The protein is then harvested (*e.g.* from the milieu about the cell, or by lysing the cell). The invention also relates to protein compositions described herein.

[0016] Accordingly, the invention relates in one aspect to a targeted therapeutic including a targeting moiety and a therapeutic agent that is therapeutically active and preferably active in a mammalian lysosome. "Therapeutically active," as used herein, encompasses at least

polypeptides or other molecules that provide an enzymatic activity to a cell or a compartment thereof that is deficient in that activity. "Therapeutically active" also encompasses other polypeptides or other molecules that are intended to ameliorate or to compensate for a biochemical deficiency in a cell, but does not encompass molecules that are primarily cytotoxic or cytostatic, such as chemotherapeutics.

[0017] In one embodiment, the targeting moiety is a means (*e.g.* a molecule) for binding the extracellular domain of the human cation-independent M6P receptor in an M6P-independent manner when the receptor is present in the plasma membrane of a target cell. In another embodiment, the targeting moiety is an unglycosylated lysosomal targeting domain that binds the extracellular domain of the human cation-independent M6P receptor. In either embodiment, the targeting moiety can include, for example, IGF-II; retinoic acid or a derivative thereof; a protein having an amino acid sequence at least 70% identical to a domain of urokinase-type plasminogen activator receptor; an antibody variable domain that recognizes the receptor; or variants thereof. In some embodiments, the targeting moiety binds to an IGF receptor (e.g., an IGF-I or an IGF-II receptor) with a submicromolar dissociation constant (*e.g.* less than 10^{-8} M, less than 10^{-9} M, less than 10^{-10} M, or between 10^{-7} M and 10^{-11} M) at or about pH 7.4 and with an dissociation constant at or about pH 5.5 of at least 10^{-6} M and at least ten times the dissociation constant at or about pH 7.4. In particular embodiments, the means for binding binds to the extracellular domain of a receptor at least 10-fold less avidly (*i.e.* with at least a ten-fold greater dissociation constant) at or about pH 5.5 than at or about pH 7.4; in one embodiment, the dissociation constant at or about pH 5.5 is at least 10^{-6} M. In a further embodiment, association of the targeted therapeutic with the means for binding is destabilized by a pH change from at or about pH 7.4 to at or about pH 5.5. In preferred embodiments, a targeting moiety retains IGF-II-like

binding affinity for the IGF-II receptor, but has reduced binding affinity for the IGF-I receptor (e.g., with a dissociation constant of 10^{-6}M or greater than 10^{-6}M).

[0018] The invention also relates to methods of treating a patient (*e.g.* a patient with a disorder in the CNS, and preferably a CNS disorder resulting from a lysosomal storage disorder) by administering, for example, a protein including a therapeutic agent effective in the mammalian CNS and an IGF tag to target the protein to the CNS. Preferably, the protein also comprises a lysosomal targeting portion such as those described in U.S. Serial No. 10/136,841, filed on April 30, 2002; attorney docket number SYM-007CP entitled "Targeted Therapuetic Proteins," filed on October 16, 2002; and atttorney docket number SYM-009 entitled "Targeted Therapeutic Proteins," filed on October 16, 2002,, or mannose-6-phosphate to target the protein to the lysosomes of deficient cells in the CNS. Similarly, the invention relates to methods of treating a patient by administering a nucleic acid encoding such a protein and/or by administering a cell (*e.g.* a human cell, or an organism such as *Leishmania*) containing a nucleic acid encoding such a protein.

[0019] This invention also provides methods for producing therapeutic proteins that are targeted to lysosomes and/or across the blood-brain barrier and that possess an extended half-life in circulation in a mammal. The methods include producing an underglycosylated therapeutic protein. As used herein, "underglycosylated" refers to a protein in which one or more carbohydrate structures that would normally be present if the protein were produced in a mammalian cell (such as a CHO cell) has been omitted, removed, modified, or masked, thereby extending the half-life of the protein in a mammal. Thus, a protein may be actually underglycosylated due to the absence of one or more carbohydrate structures, or functionally underglycosylated by modification or masking of one or more carbohydrate structures that promote clearance from circulation. For example, a structure could be masked (i) by the addition

of one or more additional moieties (*e.g.* carbohydrate groups, phosphate groups, alkyl groups, *etc.*) that interfere with recognition of the structure by a mannose or asialoglycoprotein receptor, (ii) by covalent or noncovalent association of the glycoprotein with a binding moiety, such as a lectin or an extracellular portion of a mannose or asialoglycoprotein receptor, that interferes with binding to those receptors *in vivo*, or (iii) any other modification to the polypeptide or carbohydrate portion of a glycoprotein to reduce its clearance from the blood by masking the presence of all or a portion of the carbohydrate structure.

[0020] In one embodiment, the therapeutic protein includes a peptide targeting moiety (*e.g.* IGF-I, IGF-II, or a portion thereof effective to bind a target receptor) that is produced in a host (*e.g.* bacteria or yeast) that does not glycosylate proteins as conventional mammalian cells (*e.g.* Chinese hamster ovary (CHO) cells) do. For example, proteins produced by the host cell may lack terminal mannose, fucose, and/or N-acetylglucosamine residues, which are recognized by the mannose receptor, or may be completely unglycosylated. In another embodiment, the therapeutic protein, which may be produced in mammalian cells or in other hosts, is treated chemically or enzymatically to remove one or more carbohydrate residues (*e.g.* one or more mannose, fucose, and/or N-acetylglucosamine residues) or to modify or mask one or more carbohydrate residues. Such a modification or masking may reduce binding of the therapeutic protein to the hepatic mannose and/or asialoglycoprotein receptors. In another embodiment, one or more potential glycosylation sites are removed by mutation of the nucleic acid encoding the targeted therapeutic protein, thereby reducing glycosylation of the protein when synthesized in a mammalian cell or other cell that glycosylates proteins.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021] Figure 1 shows a sequence alignment of mature human IGF-I (SEQ ID NO: 1) and IGF-II (SEQ ID NO: 2), indicating regions of homology and the A, B, C, and D domains.

[0022] Figure 2 is a two-dimensional representation of an IGF protein showing the signal sequence, the A, B, C, and D domains, and the C terminal sequence.

[0023] Figure 3 shows protein (Figure 3A, SEQ ID NO: 3) and nucleic acid (Figure 3B, SEQ ID NO: 4) sequences for human IGF-I mRNA.

[0024] Figure 4A depicts one form of a phosphorylated high mannose carbohydrate structure linked to a glycoprotein via an asparagine residue, and also depicts the structures of mannose and N-acetylglucosamine (GlcNAc). Figure 4B depicts a portion of the high mannose carbohydrate structure at a higher level of detail, and indicates positions vulnerable to cleavage by periodate treatment. The positions of the sugar residues within the carbohydrate structure are labeled with circled, capital letters A-H; phosphate groups are indicated with a circled capital P.

[0025] Figure 5 shows several types of underglycosylation.

DETAILED DESCRIPTION OF THE INVENTION

CNS targeting portion

[0026] According to the invention, an IGF moiety is useful for targeting a composition, preferably a protein composition, to the CNS, across the blood brain barrier. Preferably, an IGF tag is used to target a composition to the brain parenchyma. According to the invention, a composition may enter the CNS or brain parenchyma either directly across the blood-brain barrier (the BBB) or indirectly across the blood-cerebrospinal fluid barrier (the BCB). The BBB is formed by capillary endothelial cells and the BCB is formed by epithelial cells of the choroid plexus. Transport across either barrier typically involves transcytosis. According to the invention, a composition that is targeted across the BCB to the CSF can subsequently reach the

brain parenchyma. The CSF and brain parenchyma are separated by the ependyma, and diffusion or bulk flow can transport substances between these two compartments.

[0027] The invention exploits, in part, the recognition that [125I] IGF-I and IGF-II can be detected in the brain when infused into the carotid artery, and that IGF-I and analogs administered subcutaneously can be found in the cerebrospinal fluid. According to the invention, this suggests that both can traverse the BBB or BCB. According to the invention, the observed saturation of the transport process suggests that the process is carrier mediated. However, experimental analysis using a series of IGF-I analogs suggests that the IGF-I receptor, the IGF-II receptor, and IGF binding proteins-1,-3,-4, or 5 do not play a role in the blood brain barrier transport.

[0028] According to one aspect of the invention, preferred therapeutic compositions include a therapeutic peptide fused to an IGF tag. In one embodiment, the IGF tag does not include IGF-II or a portion thereof. Preferred therapeutic compositions include IGF-I tags that will direct LSD (lysosomal storage disease) proteins to which they are fused across the blood brain barrier. In this instance, the tag will not necessarily direct the protein to the lysosome of multiple cell types. However, by expressing such fusion proteins in mammalian cell culture systems, the invention exploits the endogenous M6P signal for lysosomal localization and uses the IGF-I tag to traverse the blood brain barrier. In preferred embodiments of the invention, a human IGF-I tag is used. In alternative embodiments, methods and compositions of the invention involve using allelic, species or other sequence variants of an IGF-I tag. Preferred sequence variants include mutations that lessen binding of the IGF tag to the IGF-I receptor and/or IGF binding proteins such as Leu^{60}-IGF-I, or Leu^{24} IGF-I which have diminished binding to the IGF-I receptor or Δ1-3 IGF-I which has diminished binding to IGF-binding proteins. Additional useful

sequence variants include IGF-I variants with amino acid replacements of Arg^{55} and Arg^{56}. Other mutant IGF protein tags with similar properties are also useful.

[0029] IGF-II based tags are also useful to target proteins to the brain. IGF-II has been reported to be transported across the blood brain barrier via transcytosis (Bickel *et al.* (2001) Adv. Drug Deliv. Rev. 46(1-3):247-79). According to the invention, preferred IGF-II-based tags target proteins to the brain and also target proteins to the lysosome via receptor binding in order to treat neurological symptoms associated with lysosomal storage diseases. Preferred variants of IGF-II have an amino acid replacement at Leu^{24}.

[0030] In another aspect of the invention, chimeric tags are used that include fragments of IGF-I and IGF-II, conferring preferred functional properties of each protein. In one embodiment, the retained portion of IGF-II includes regions of IGF-II known to be critical for binding to the IGF-II M6P receptor while the remainder of IGF-II would be substituted for the corresponding regions of IGF-I. This embodiment, is particularly useful where IGF-I is more active as a tag for traversing the blood brain barrier. In this embodiment, the tag has optimized activity for lysosomal targeting in addition to brain targeting. A recombinant form of this embodiment could be made in any expression system.

[0031] In a further aspect of the invention, a useful recombinant LSD protein includes any one of the different IGF-based lysosomal targeting tags described U.S. Serial No. 10/136,841, filed on April 30, 2002; attorney docket number SYM-007CP entitled "Targeted Therapuetic Proteins," filed on October 16, 2002; and atttorney docket number SYM-009 entitled "Targeted Therapeutic Proteins," filed on October 16, 2002.

[0032] In preferred embodiments, recombinant proteins of the invention including IGF-II tags are expressed in a mammalian expression system such as a CHO cell expression system.

According to the invention, the endogenous M6P signal added in the mammalian cell culture enhances the lysosomal targeting that may be provided by an IGF-II tag.

[0033] According to the invention, useful minimal IGF tags and variant IGF tags can be identified based on known IGF-I and IGF-II sequences by testing minimal or variant IGF fragments in a CNS localization assay such as one described herein.

[0034] A preferred IGF tag is sufficiently duplicative of IGF-I to be targeted to the brain, but has reduced binding affinity for the IGF-I receptor thereby removing the mitogenic properties of IGF-I. However, a preferred IGF tag does bind to the IGF-II receptor in order to be targeted to lysosomes. Accordingly, in one embodiment, an IGF tag is based on the IGF-I sequence but includes two hydrophobic IGF-II residues at positions 54 and 55 instead of the IGF-I Arg residues at these positions.

Structure of IGF-II

[0035] NMR structures of IGF-II have been solved by two groups (see, *e.g.*, Protein Data Bank record 1IGL). The general features of the IGF-II structure are similar to IGF-I and insulin. The A and B domains of IGF-II correspond to the A and B chains of insulin. Secondary structural features include an alpha helix from residues 11-21 of the B region connected by a reverse turn in residues 22-25 to a short beta strand in residues 26-28. Residues 25-27 appear to form a small antiparallel beta sheet; residues 59-61 and residues 26-28 may also participate in intermolecular beta-sheet formation. In the A domain of IGF-II, alpha helices spanning residues 42-49 and 53-59 are arranged in an antiparallel configuration perpendicular to the B-domain helix. Hydrophobic clusters formed by two of the three disulfide bridges and conserved hydrophobic residues stabilize these secondary structure features. The N and C termini remain poorly defined as is the region between residues 31-40.

[0036] IGF-II binds to the IGF-II/M6P and IGF-I receptors with relatively high affinity and binds with lower affinity to the insulin receptor. IGF-II also interacts with a number if serum IGFBPs.

[0037] In a preferred embodiment, the lysosomal targeting portion is a protein, peptide, or other moiety that binds the cation independent M6P/IGF-II receptor in a mannose-6-phosphate-independent manner. Advantageously, this embodiment mimics the normal biological mechanism for uptake of LSD proteins, yet does so in a manner independent of mannose-6-phosphate.

[0038] For example, by fusing DNA encoding the mature IGF-II polypeptide to the 3' end of LSD gene cassettes, fusion proteins are created that can be taken up by a variety of cell types and transported to the lysosome. Alternatively, DNA encoding a precursor IGF-II polypeptide can be fused to the 3' end of an LSD gene cassette; the precursor includes a carboxyterminal portion that is cleaved in mammalian cells to yield the mature IGF-II polypeptide, but the IGF-II signal peptide is preferably omitted (or moved to the 5' end of the LSD gene cassette). This method has numerous advantages over methods involving glycosylation including simplicity and cost effectiveness, because once the protein is isolated, no further modifications need be made.

Binding to the IGF-II/M6P receptor

[0039] Substitution of IGF-II residues 48-50 (Phe Arg Ser) with the corresponding residues from insulin, (Thr Ser Ile), or substitution of residues 54-55 (Ala Leu) with the corresponding residues from IGF-I (Arg Arg) result in loss of binding to the IGF-II/ M6P receptor but retention of binding to the IGF-I and insulin receptors.

[0040] IGF-I and IGF-II share identical sequences and structures in the region of residues 48-50 yet have a 1000-fold difference in affinity for the IGF-II receptor. The NMR structure reveals a structural difference between IGF-I and IGF-II in the region of IGF-II residues 53-58

(IGF-I residues 54-59): the alpha-helix is better defined in IGF-II than in IGF-I and, unlike IGF-I, there is no bend in the backbone around residues 53 and 54. This structural difference correlates with the substitution of Ala 54 and Leu 55 in IGF-II with Arg 55 and Arg 56 in IGF-I. It is possible either that binding to the IGF-II receptor is disrupted directly by the presence of charged residues in this region or that changes in the structure engendered by the charged residues yield the changes in binding for the IGF-II receptor. In any case, substitution of uncharged residues for the two Arg residues in IGF-I resulted in higher affinities for the IGF-II receptor. Thus the presence of positively charged residues in these positions correlates with loss of binding to the IGF-II receptor.

[0041] IGF-II binds to repeat 11 of the cation-independent M6P receptor. Indeed, a minireceptor in which only repeat 11 is fused to the transmembrane and cytoplasmic domains of the cation-independent M6P receptor is capable of binding IGF-II (with an affinity approximately one tenth the affinity of the full length receptor) and mediating internalization of IGF-II and its delivery to lysosomes (Grimme *et al.* (2000) J. Biol. Chem. 275(43):33697-33703). The structure of domain 11 of the M6P receptor is known (Protein Data Base entries 1GP0 and 1GP3; Brown *et al.* (2002) EMBO J. 21(5):1054-1062). The putative IGF-II binding site is a hydrophobic pocket believed to interact with hydrophobic amino acids of IGF-II; candidate amino acids of IGF-II include leucine 8, phenylalanine 48, alanine 54, and leucine 55. Although repeat 11 is sufficient for IGF-II binding, constructs including larger portions of the cation-independent M6P receptor (*e.g.* repeats 10-13, or 1-15) generally bind IGF-II with greater affinity and with increased pH dependence (see, for example, Linnell *et al.* (2001) J. Biol. Chem. 276(26):23986-23991).

Binding to the IGF-I receptor

[0042] Substitution of IGF-II residues Tyr 27 with Leu, Leu 43 with Val or Ser 26 with Phe diminishes the affinity of IGF-II for the IGF-I receptor by 94-, 56-, and 4-fold respectively. Deletion of residues 1-7 of human IGF-II resulted in a 30-fold decrease in affinity for the human IGF-I receptor and a concomitant 12 fold increase in affinity for the rat IGF-II receptor. The NMR structure of IGF-II shows that Thr 7 is located near residues 48 Phe and 50 Ser as well as near the 9 Cys-47 Cys disulfide bridge. It is thought that interaction of Thr 7 with these residues can stabilize the flexible N-terminal hexapeptide required for IGF-I receptor binding. At the same time this interaction can modulate binding to the IGF-II receptor. Truncation of the C-terminus of IGF-II (residues 62-67) also appear to lower the affinity of IGF-II for the IGF-I receptor by 5 fold.

Deletion mutants of IGF-II

[0043] The binding surfaces for the IGF-I and cation-independent M6P receptors are on separate faces of IGF-II. Based on structural and mutational data, functional cation-independent M6P binding domains can be constructed that are substantially smaller than human IGF-II. For example, the amino terminal amino acids 1-7 and/or the carboxy terminal residues 62-67 can be deleted or replaced. Additionally, amino acids 29-40 can likely be eliminated or replaced without altering the folding of the remainder of the polypeptide or binding to the cation-independent M6P receptor. Thus, a targeting moiety including amino acids 8-28 and 41-61 can be constructed. These stretches of amino acids could perhaps be joined directly or separated by a linker. Alternatively, amino acids 8-28 and 41-61 can be provided on separate polypeptide chains. Comparable domains of insulin, which is homologous to IGF-II and has a tertiary structure closely related to the structure of IGF-II, have sufficient structural information to permit proper refolding into the appropriate tertiary structure, even when present in separate

polypeptide chains (Wang et al. (1991) Trends Biochem. Sci. 279-281). Thus, for example, amino acids 8-28, or a conservative substitution variant thereof, could be fused to a therapeutic agent; the resulting fusion protein could be admixed with amino acids 41-61, or a conservative substitution variant thereof, and administered to a patient. According to the invention, analogous fragments of IGF-I are useful in fusions with a therapeutic.

Binding to IGF Binding proteins

[0044] IGF-II and related constructs can be modified to diminish their affinity for IGFBPs, thereby increasing the bioavailability of the tagged proteins.

[0045] Substitution of IGF-II residue phenylalanine 26 with serine reduces binding to IGFBPs 1-5 by 5-75 fold. Replacement of IGF-II residues 48-50 with threonine-serine-isoleucine reduces binding by more than 100 fold to most of the IGFBPs. These residues are, however, also important for binding to the cation-independent mannose-6-phosphate receptor. The Y27L substitution that disrupts binding to the IGF-I receptor interferes with formation of the ternary complex with IGFBP3 and acid labile subunit; this ternary complex accounts for most of the IGF-II in the circulation. Deletion of the first six residues of IGF-II also interferes with IGFBP binding.

[0046] Studies on IGF-I interaction with IGFBPs revealed additionally that substitution of serine for phenylalanine 16 did not effect secondary structure but decreased IGFBP binding by between 40 and 300 fold. Changing glutamate 9 to lysine also resulted in a significant decrease in IGFBP binding. Furthermore, the double mutant lysine 9/ serine 16 exhibited the lowest affinity for IGFBPs. Although these mutations have not previously been tested in IGF-II, the conservation of sequence between this region of IGF-I and IGF-II suggests that a similar effect will be observed when the analogous mutations are made in IGF-II (glutamate 12 lysine/ phenylalanine 19 serine).

IGF homologs

[0047] The amino acid sequence of human IGF-I, IGF-II, or a portion thereof affecting transport into the brain, may be used as a reference sequence to determine whether a candidate sequence possesses sufficient amino acid similarity to have a reasonable expectation of success in the methods of the present invention. Preferably, variant sequences are at least 70% similar or 60% identical, more preferably at least 75% similar or 65% identical, and most preferably 80% similar or 70% identical to human IGF-I or IGF-II.

[0048] To determine whether a candidate peptide region has the requisite percentage similarity or identity to human IGF-I or IGF-II, the candidate amino acid sequence and human IGF-I or IGF-II are first aligned using the dynamic programming algorithm described in Smith and Waterman (1981) J. Mol. Biol. 147:195-197, in combination with the BLOSUM62 substitution matrix described in Figure 2 of Henikoff and Henikoff (1992) PNAS 89:10915-10919. For the present invention, an appropriate value for the gap insertion penalty is -12, and an appropriate value for the gap extension penalty is -4. Computer programs performing alignments using the algorithm of Smith-Waterman and the BLOSUM62 matrix, such as the GCG program suite (Oxford Molecular Group, Oxford, England), are commercially available and widely used by those skilled in the art.

[0049] Once the alignment between the candidate and reference sequence is made, a percent similarity score may be calculated. The individual amino acids of each sequence are compared sequentially according to their similarity to each other. If the value in the BLOSUM62 matrix corresponding to the two aligned amino acids is zero or a negative number, the pairwise similarity score is zero; otherwise the pairwise similarity score is 1.0. The raw similarity score is the sum of the pairwise similarity scores of the aligned amino acids. The raw score is then normalized by dividing it by the number of amino acids in the smaller of the candidate or

reference sequences. The normalized raw score is the percent similarity. Alternatively, to calculate a percent identity, the aligned amino acids of each sequence are again compared sequentially. If the amino acids are non-identical, the pairwise identity score is zero; otherwise the pairwise identity score is 1.0. The raw identity score is the sum of the identical aligned amino acids. The raw score is then normalized by dividing it by the number of amino acids in the smaller of the candidate or reference sequences. The normalized raw score is the percent identity. Insertions and deletions are ignored for the purposes of calculating percent similarity and identity. Accordingly, gap penalties are not used in this calculation, although they are used in the initial alignment.

IGF structural analogs

[0050] The known structures of human IGF proteins permit the design of IGF analogs using computer-assisted design principles such as those discussed in U.S. Patent Nos. 6,226,603 and 6,273,598. For example, the known atomic coordinates of IGF-II can be provided to a computer equipped with a conventional computer modeling program, such as INSIGHTII, DISCOVER, or DELPHI, commercially available from Biosym, Technologies Inc., or QUANTA, or CHARMM, commercially available from Molecular Simulations, Inc. These and other software programs allow analysis of molecular structures and simulations that predict the effect of molecular changes on structure and on intermolecular interactions. For example, the software can be used to identify modified analogs with the ability to form additional intermolecular hydrogen or ionic bonds, improving the affinity of the analog for the target receptor.

[0051] The software also permits the design of peptides and organic molecules with structural and chemical features that mimic the same features displayed on at least part of an IGF surface that is sufficient for targeting to the CNS. A preferred embodiment of the present invention relates to designing and producing a synthetic organic molecule having a framework

that carries chemically interactive moieties in a spatial relationship that mimics the spatial relationship of the chemical moieties disposed on the amino acid sidechains which are identified as associated with CNS targeting as described herein.

[0052] For example, upon identification of relevant chemical groups, the skilled artisan using a conventional computer program can design a small molecule having appropriate chemical moieties disposed upon a suitable carrier framework. Useful computer programs are described in, for example, Dixon (1992) Tibtech 10: 357-363; Tschinke et al. (1993) J. Med. Chem 36: 3863-3870; and Eisen el al. (1994) Proteins: Structure, Function, and Genetics 19: 199-221, the disclosures of which are incorporated herein by reference.

[0053] One particular computer program entitled "CAVEAT" searches a database, for example, the Cambridge Structural Database, for structures which have desired spatial orientations of chemical moieties (Bartlett et al. (1989) in "Molecular Recognition: Chemical and Biological Problems" (Roberts, S. M., ed) pp 182-196). The CAVEAT program has been used to design analogs of tendamistat, a 74 residue inhibitor of .alpha.-amylase, based on the orientation of selected amino acid side chains in the three-dimensional structure of tendamistat (Bartlett et al. (1989) supra).

[0054] Alternatively, upon identification of a series of analogs which target transport to the CNS, the skilled artisan may use a variety of computer programs which assist the skilled artisan to develop quantitative structure activity relationships (QSAR) and further to assist in the de novo design of additional analogs. Other useful computer programs are described in, for example, Connolly-Martin (1991) Methods in Enzymology 203:587-613; Dixon (1992) supra; and Waszkowycz et al. (1994) J. Med. Chenm. 37: 3994-4002.

Therapeutic agent

[0055] While methods and compositions of the invention are useful for producing and delivering any therapeutic agent to the CNS, the invention is particularly useful for gene products that overcome enzymatic defects associated with lysosomal storage diseases.

[0056] Preferred LSD genes are shown in Table 1. In a preferred embodiment, a wild-type LSD gene product is delivered to a patient suffering from a defect in the same LSD gene. In alternative embodiments, a functional sequence or species variant of the LSD gene is used. In further embodiments, a gene coding for a different enzyme that can rescue an LSD gene defect is used according to methods of the invention.

[0057] Table 1. Lysosomal Storage Diseases and associated enzyme defects

A. Glycogenosis Disorders		
Disease Name	Enzyme Defect	Substance Stored
Pompe Disease	Acid-a1, 4-Glucosidase	Glycogen α 1-4 linked Oligosaccharides
B. Glycolipidosis Disorders		
Disease Name	Enzyme Defect	Substance Stored
GM1 Gangliodsidosis	β-Galactosidase	GM_1 Ganliosides
Tay-Sachs Disease	β-Hexosaminidase A	GM_2 Ganglioside
GM2 Gangliosidosis: AB Variant	GM_2 Activator Protein	GM_2 Ganglioside
Sandhoff Disease	β-Hexosaminidase A&B	GM_2 Ganglioside
Fabry Disease	α-Galactosidase A	Globosides
Gaucher Disease	Glucocerebrosidase	Glucosylceramide
Metachromatic Leukodystrophy	Arylsulfatase A	Sulphatides
Krabbe Disease	Galactosylceramidase	Galactocerebroside
Niemann-Pick, Types A and B	Acid Sphingomyelinase	Sphingomyelin
Niemann-Pick, Type C	Cholesterol Esterification Defect	Sphingomyelin
Nieman-Pick, Type D	Unknown	Sphingomyelin
Farber Disease	Acid Ceramidase	Ceramide
Wolman Disease	Acid Lipase	Cholesteryl Esters

C. Mucopolysaccharide Disorders		
Disease Name	Enzyme Defect	Substance Stored
Hurler Syndrome (MPS IH)	α-L-Iduronidase	Heparan & Dermatan Sulfates
Scheie Syndrome (MPS IS)	α-L-Iduronidase	Heparan & Dermatan, Sulfates
Hurler-Scheie (MPS IH/S)	α-L-Iduronidase	Heparan & Dermatan Sulfates
Hunter Syndrome (MPS II)	Iduronate Sulfatase	Heparan & Dermatan Sulfates
Sanfilippo A (MPS IIIA)	Heparan N-Sulfatase	Heparan Sulfate
Sanfilippo B (MPS IIIB)	α-N-Acetylglucosaminidase	Heparan Sulfate
Sanfilippo C (MPS IIIC)	Acetyl-CoA-Glucosaminide Acetyltransferase	Heparan Sulfate
Sanfilippo D (MPS IIID)	N-Acetylglucosamine -6-Sulfatase	Heparan Sulfate
Morquio A (MPS IVA)	Galactosamine-6-Sulfatase	Keratan Sulfate
Morquio B (MPS IVB)	β-Galactosidase	Keratan Sulfate
Maroteaux-Lamy (MPS VI)	Arylsulfatase B	Dermatan Sulfate
Sly Syndrome (MPS VII)	β-Glucuronidase	

D. Oligosaccharide/Glycoprotein Disorders		
Disease Name	Enzyme Defect	Substance Stored
α–Mannosidosis	α–Mannosidase	Mannose/Oligosaccharides
β-Mannosidosis	β-Mannosidase	Mannose/Oligosaccharides
Fucosidosis	α-L-Fucosidase	Fucosyl Oligosaccharides
Asparylglucosaminuria	N-Aspartyl- β-Glucosaminidase	Asparylglucosamine Asparagines
Sialidosis (Mucolipidosis I)	α-Neuraminidase	Sialyloligosaccharides

Galactosialidosis (Goldberg Syndrome)	Lysosomal Protective Protein Deficiency	Sialyloligosaccharides
Schindler Disease	α-N-Acetyl-Galactosaminidase	
E. Lysosomal Enzyme Transport Disorders		
Disease Name	Enzyme Defect	Substance Stored
Mucolipidosis II (I-Cell Disease)	N-Acetylglucosamine-1- Phosphotransferase	Heparan Sulfate
Mucolipidosis III (Pseudo-Hurler Polydystrophy)	Same as ML II	

F. Lysosomal Membrane Transport Disorders		
Disease Name	Enzyme Defect	Substance Stored
Cystinosis	Cystine Transport Protein	Free Cystine
Salla Disease	Sialic Acid Transport Protein	Free Sialic Acid and Glucuronic Acid
Infantile Sialic Acid Storage Disease	Sialic Acid Transport Protein	Free Sialic Acid and Glucuronic Acid
G. Other		
Disease Name	Enzyme Defect	Substance Stored
Batten Disease (Juvenile Neuronal Ceroid Lipofuscinosis)	Unknown.	Lipofuscins
Infantile Neuronal Ceroid Lipofuscinosis	Palmitoyl-Protein Thioesterase	Lipofuscins
Mucolipidosis IV	Unknown	Gangliosides & Hyaluronic Acid
Prosaposin	Saposins A, B, C or D	

[0058] In one embodiment, the therapeutic agent is glucocerebrosidase, currently manufactured by Genzyme as an effective enzyme replacement therapy for Gaucher Disease. Currently, the enzyme is prepared with exposed mannose residues, which targets the protein specifically to cells of the macrophage lineage. Although the primary pathology in type 1

Gaucher patients are due to macrophage accumulating glucocerebroside, there may be therapeutic advantage to delivering glucocerebrosidase to other cell types. Targeting glucocerebrosidase to lysosomes using the present invention would target the agent to multiple cell types and may have a therapeutic advantage compared to other preparations.

Association between targeting portion and therapeutic portion

[0059] The therapeutic portion and the targeting portion of compositions of the invention are necessarily associated, directly or indirectly. In one embodiment, the therapeutic portion and the targeting portion are non-covalently associated. For example, the targeting portion could be biotinylated and bind an avidin moiety associated with the therapeutic portion. Alternatively, the targeting portion and the therapeutic portion could each be associated (*e.g.* as fusion proteins) with different subunits of a multimeric protein. In another embodiment, the targeting portion and the therapeutic portion are crosslinked to each other (*e.g.* using a chemical crosslinking agent).

[0060] In a preferred embodiment, the therapeutic portion is fused to the targeting portion as a fusion protein. The targeting portion may be at the amino-terminus of the fusion protein, the carboxy-terminus, or may be inserted within the sequence of the therapeutic portion at a position where the presence of the targeting portion does not unduly interfere with the therapeutic activity of the therapeutic portion.

[0061] Where the therapeutically active moiety is a heteromeric protein, one or more of the subunits may be associated with a targeting portion. Hexosaminidase A, for example, a lysosomal protein affected in Tay-Sachs disease, includes an alpha subunit and a beta subunit. Either the alpha subunit, or the beta subunit, or both may be associated with a targeting portion in accordance with the present invention. If, for example, the alpha subunit is associated with a targeting portion and is coexpressed with the beta subunit, an active complex is formed and targeted to the lysosome.

Expression Methods

[0062] Methods and compositions of the invention are useful in the context of many different expression systems. For example, a protein of the invention can be targeted to the CNS, and preferably taken up by lysosomes, whether it is expressed and isolated from *Leishmania*, baculovirus, yeast, bacteria, mammalian or other expression systems. Thus, the invention permits great flexibility in protein production. For example, if a protein to be produced includes one or more disulfide bonds, an appropriate expression system can be selected and modified, if appropriate, to further improve yield of properly folded protein. For example, one useful IGF targeting portion has three intramolecular disulfide bonds. Fusion proteins of the invention expressed in *E. coli* may be constructed to direct the protein to the periplasmic space. IGF tags, when fused to the C-terminus of another protein, can be secreted in an active form in the periplasm of *E. coli* (Wadensten, Ekebacke et al. 1991). To facilitate optimal folding of the IGF moiety, appropriate concentrations of reduced and oxidized glutathione are preferably added to the cellular milieu to promote disulfide bond formation. In the event that a fusion protein with disulfide bonds is incompletely soluble, any insoluble material is preferably treated with a chaotropic agent such as urea to solubilize denatured protein and refolded in a buffer having appropriate concentrations of reduced and oxidized glutathione, or other oxidizing and reducing agents, to facilitate formation of appropriate disulfide bonds (Smith, Cook et al. 1989). For example, IGF-I has been refolded using 6M guanidine-HCl and 0.1 M tris(2-carboxyethyl)phosphine reducing agent for denaturation and reduction of IGF-II (Yang, Wu et al. 1999). Refolding of proteins was accomplished in 0.1M Tris-HCl buffer (pH8.7) containing 1mM oxidized glutathione, 10 mM reduced glutathione, 0.2M KCl and 1mM EDTA.

[0063] Methods of the invention are also useful to target a protein directly to the CNS of a mammal without requiring a purification step. In one embodiment, an IGF fusion protein is

expressed in a symbiotic or parasitic organism that is administered to a host. The expressed IGF fusion protein is secreted by the organism into the blood stream and delivered across the blood brain barrier.

[0064] In some embodiments of the invention, CNS targeted proteins are delivered *in situ* via live *Leishmania* secreting the proteins into the lysosomes of infected macrophage. From this organelle, it leaves the cell and may be delivered across the blood brain barrier. Thus, the IGF tag and the therapeutic agent necessarily remain intact while the protein resides in the macrophage lysosome. Accordingly, when proteins designed for delivery to lysosomes in the CNS are expressed *in situ*, they are preferably modified to ensure compatibility with the lysosomal environment. In alternative embodiments, therapeutic proteins of the invention can be delivered by expression in *T. brucei* which can penetrate the BCB.

Nucleic acids and expression systems

[0065] Chimeric fusion proteins of the invention can be expressed in a variety of expression systems, including *in vitro* translation systems and intact cells. Since M6P modification is not a prerequisite for targeting, a variety of expression systems including yeast, baculovirus and even prokaryotic systems such as *E. coli* that do not glycosylate proteins are suitable for expression of targeted therapeutic proteins. In fact, an unglycosylated protein generally has improved bioavailability, since glycosylated proteins are rapidly cleared from the circulation through binding to the mannose receptor in hepatic sinusoidal endothelium.

[0066] Alternatively, production of chimeric targeted lysosomal enzymes in mammalian cell expression system produces proteins with multiple binding determinants for the cation-independent M6P receptor. Synergies between two or more cation-independent M6P receptor ligands (e.g. M6P and IGF-II, or M6P and retinoic acid) can be exploited: multivalent ligands have been demonstrated to enhance binding to the receptor by receptor crosslinking.

[0067] In general, gene cassettes encoding the chimeric therapeutic protein can be tailored for the particular expression system to incorporate necessary sequences for optimal expression including promoters, ribosomal binding sites, introns, or alterations in coding sequence to optimize codon usage. Because the protein is preferably secreted from the producing cell, a DNA encoding a signal peptide compatible with the expression system can be substituted for the endogenous signal peptide. For example, for expression of ß-glucuronidase and α-galactosidase A tagged with IGF-I or IGF-II in *Leishmania*, DNA cassettes encoding *Leishmania* signal peptides (GP63 or SAP) are inserted in place of the DNA encoding the endogenous signal peptide to achieve optimal expression. In mammalian expression systems the endogenous signal peptide may be employed but if the IGF-I or IGF-II tag is fused at the 5' end of the coding sequence, it could be desirable to use the IGF-I or IGF-II signal peptide.

[0068] CHO cells are a preferred mammalian host for the production of therapeutic proteins. The classic method for achieving high yield expression from CHO cells is to use a CHO cell line deficient in dihydrofolate reductase (DHFR), for example CHO line DUKX (O'Dell et al. (1998) Int. J. Biochem. Cell Biol. 30(7):767-71). This strain of CHO cells requires hypoxanthine and thymidine for growth. Co-transfection of the gene to be overexpressed with a DHFR gene cassette, on separate plasmids or on a single plasmid, permits selection for the DHFR gene and generally allows isolation of clones that also express the recombinant protein of choice. For example, plasmid pcDNA3 uses the cytomegalovirus (CMV) early region regulatory region promoter to drive expression of a gene of interest and pSV2DHFR to promote DHFR expression. Subsequent exposure of cells harboring the recombinant gene cassettes to incrementally increasing concentrations of the folate analog methotrexate leads to amplification of both the gene copy number of the DHFR gene and of the co-transfected gene.

[0069] A preferred plasmid for eukaryotic expression in this system contains the gene of interest placed downstream of a strong promoter such as CMV. An intron can be placed in the 3' flank of the gene cassette. A DHFR cassette can be driven by a second promoter from the same plasmid or from a separate plasmid. Additionally, it can be useful to incorporate into the plasmid an additional selectable marker such as neomycin phosphotransferase, which confers resistance to G418.

[0070] Another CHO expression system (Ulmasov et al. (2000) PNAS 97(26):14212-14217) relies on amplification of the gene of interest using G418 instead of the DHFR/methotrexate system described above. A pCXN vector with a slightly defective neomycin phosphotransferase driven by a weak promoter (see, e.g., Niwa et al. (1991) Gene 108:193-200) permits selection for transfectants with a high copy number (>300) in a single step.

[0071] Alternatively, recombinant protein can be produced in the human HEK 293 cell line using expression systems based on the Epstein-Barr Virus (EBV) replication system. This consists of the EBV replication origin oriP and the EBV ori binding protein, EBNA-1. Binding of EBNA-1 to oriP initiates replication and subsequent amplification of the extrachromosomal plasmid. This amplification in turn results in high levels of expression of gene cassettes housed within the plasmid. Plasmids containing oriP can be transfected into EBNA-1 transformed HEK 293 cells (commercially available from Invitrogen) or, alternatively, a plasmid such as pCEP4 (commercially available from Invitrogen) which drives expression of EBNA-1 and contains the EBV oriP can be employed.

[0072] In *E. coli*, the therapeutic proteins are preferably secreted into the periplasmic space. This can be achieved by substituting for the DNA encoding the endogenous signal peptide of the LSD protein a nucleic acid cassette encoding a bacterial signal peptide such as the ompA signal sequence. Expression can be driven by any of a number of strong inducible promoters such as

the lac, trp, or tac promoters. One suitable vector is pBAD/gIII (commercially available from Invitrogen) which uses the Gene III signal peptide and the araBAD promoter.

In vivo expression

[0073] A nucleic acid encoding a therapeutic protein, preferably a secreted therapeutic protein, can be advantageously provided directly to a patient suffering from a disease, or may be provided to a cell *ex vivo*, followed by administration of the living cell to the patient. *In vivo* gene therapy methods known in the art include providing purified DNA (e.g. as in a plasmid) , providing the DNA in a viral vector, or providing the DNA in a liposome or other vesicle (see, for example, U.S. Patent No. 5,827,703, disclosing lipid carriers for use in gene therapy, and U.S. Patent No. 6,281,010, providing adenoviral vectors useful in gene therapy).

[0074] Methods for treating disease by implanting a cell that has been modified to express a recombinant protein are also well known. See, for example, U.S. Patent No. 5,399,346, disclosing methods for introducing a nucleic acid into a primary human cell for introduction into a human. Although use of human cells for *ex vivo* therapy is preferred in some embodiments, other cells such as bacterial cells may be implanted in a patient's vasculature, continuously releasing a therapeutic agent. See, for example, U.S. Patent Nos. 4,309,776 and 5,704,910.

[0075] Methods of the invention are particularly useful for targeting a protein directly to a subcellular compartment without requiring a purification step. In one embodiment, an IGF-II fusion protein is expressed in a symbiotic or attenuated parasitic organism that is administered to a host. The expressed IGF-II fusion protein is secreted by the organism, taken up by host cells and targeted to their lysosomes.

[0076] In some embodiments of the invention, IGF fusion proteins such as GILT proteins are delivered *in situ* via live *Leishmania* secreting the proteins into the lysosomes of infected

macrophage. From this organelle, it leaves the cell and is taken up by adjacent cells not of the macrophage lineage. Thus, the IGF tag and the therapeutic agent necessarily remain intact while the protein resides in the macrophage lysosome. Accordingly, when GILT proteins are expressed *in situ*, they are preferably modified to ensure compatibility with the lysosomal environment. Human ß-glucuronidase (human "GUS"), an exemplary therapeutic portion, normally undergoes a C-terminal peptide cleavage either in the lysosome or during transport to the lysosome (e.g. between residues 633 and 634 in GUS). Thus, in embodiments where a GUS-GILT construct is to be expressed by *Leishmania* in a macrophage lysosome, human GUS is preferably modified to render the protein resistant to cleavage, or the residues following residue 633 are preferably simply omitted from a GILT fusion protein. Similarly, any IGF tag of the invention, is preferably modified to increase its resistance to proteolysis, or a minimal binding peptide (e.g. as identified by phage display or yeast two hybrid) is substituted for the wildtype IGF moiety.

Underglycosylation

[0077] Targeted therapeutic proteins are preferably underglycosylated: one or more carbohydrate structures that would normally be present if the protein were produced in a mammalian cell is preferably omitted, removed, modified, or masked, extending the half-life of the protein in a mammal. Underglycosylation can be achieved in many ways, several of which are diagrammed in Figure 5. As shown in Figure 5, a protein may be actually underglycosylated, actually lacking one or more of the carbohydrate structures, or functionally underglycosylated through modification or masking of one or more of the carbohydrate structures. A protein may be actually underglycosylated when synthesized, as discussed in Example 12, and may be completely unglycosylated (as when synthesized in *E. coli*), partially unglycosylated (as when

synthesized in a mammalian system after disruption of one or more glycosylation sites by site-directed mutagenesis), or may have a non-mammalian glycosylation pattern. Actual underglycosylation can also be achieved by deglycosylation of a protein after synthesis. As discussed in Example 12, deglycosylation can be through chemical or enzymatic treatments, and may lead to complete deglycosylation or, if only a portion of the carbohydrate structure is removed, partial deglycosylation.

Administration

[0078] The targeted therapeutics produced according to the present invention can be administered to a mammalian host by any route. Thus, as appropriate, administration can be oral or parenteral, including intravenous and intraperitoneal routes of administration. In addition, administration can be by periodic injections of a bolus of the therapeutic or can be made more continuous by intravenous or intraperitoneal administration from a reservoir which is external (e.g., an i.v. bag). In certain embodiments, the therapeutics of the instant invention can be pharmaceutical-grade. That is, certain embodiments comply with standards of purity and quality control required for administration to humans. Veterinary applications are also within the intended meaning as used herein.

[0079] The formulations, both for veterinary and for human medical use, of the therapeutics according to the present invention typically include such therapeutics in association with a pharmaceutically acceptable carrier therefor and optionally other ingredient(s). The carrier(s) can be "acceptable" in the sense of being compatible with the other ingredients of the formulations and not deleterious to the recipient thereof. Pharmaceutically acceptable carriers, in this regard, are intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with

pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds (identified according to the invention and/or known in the art) also can be incorporated into the compositions. The formulations can conveniently be presented in dosage unit form and can be prepared by any of the methods well known in the art of pharmacy/microbiology. In general, some formulations are prepared by bringing the therapeutic into association with a liquid carrier or a finely divided solid carrier or both, and then, if necessary, shaping the product into the desired formulation.

[0080] A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include oral or parenteral, e.g., intravenous, intradermal, inhalation, transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. Ph can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide.

[0081] Useful solutions for oral or parenteral administration can be prepared by any of the methods well known in the pharmaceutical art, described, for example, in Remington's Pharmaceutical Sciences, (Gennaro, A., ed.), Mack Pub., 1990. Formulations for parenteral administration also can include glycocholate for buccal administration, methoxysalicylate for

rectal administration, or cutric acid for vaginal administration. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic. Suppositories for rectal administration also can be prepared by mixing the drug with a non-irritating excipient such as cocoa butter, other glycerides, or other compositions that are solid at room temperature and liquid at body temperatures. Formulations also can include, for example, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, hydrogenated naphthalenes, and the like. Formulations for direct administration can include glycerol and other compositions of high viscosity. Other potentially useful parenteral carriers for these therapeutics include ethylene-vinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation administration can contain as excipients, for example, lactose, or can be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or oily solutions for administration in the form of nasal drops, or as a gel to be applied intranasally. Retention enemas also can be used for rectal delivery.

[0082] Formulations of the present invention suitable for oral administration can be in the form of discrete units such as capsules, gelatin capsules, sachets, tablets, troches, or lozenges, each containing a predetermined amount of the drug; in the form of a powder or granules; in the form of a solution or a suspension in an aqueous liquid or non-aqueous liquid; or in the form of an oil-in-water emulsion or a water-in-oil emulsion. The therapeutic can also be administered in the form of a bolus, electuary or paste. A tablet can be made by compressing or moulding the drug optionally with one or more accessory ingredients. Compressed tablets can be prepared by compressing, in a suitable machine, the drug in a free-flowing form such as a powder or granules, optionally mixed by a binder, lubricant, inert diluent, surface active or dispersing agent.

Molded tablets can be made by molding, in a suitable machine, a mixture of the powdered drug and suitable carrier moistened with an inert liquid diluent.

[0083] Oral compositions generally include an inert diluent or an edible carrier. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients. Oral compositions prepared using a fluid carrier for use as a mouthwash include the compound in the fluid carrier and are applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose; a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

[0084] Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor ELTM (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition can be sterile and can be fluid to the extent that easy syringability exists. It can be stable under the conditions of manufacture and storage and can be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the

required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, and sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

[0085] Sterile injectable solutions can be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation include vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[0086] Formulations suitable for intra-articular administration can be in the form of a sterile aqueous preparation of the therapeutic which can be in microcrystalline form, for example, in the form of an aqueous microcrystalline suspension. Liposomal formulations or biodegradable polymer systems can also be used to present the therapeutic for both intra-articular and ophthalmic administration.

[0087] Formulations suitable for topical administration, including eye treatment, include liquid or semi-liquid preparations such as liniments, lotions, gels, applicants, oil-in-water or water-in-oil emulsions such as creams, ointments or pasts; or solutions or suspensions such as drops. Formulations for topical administration to the skin surface can be prepared by dispersing

the therapeutic with a dermatologically acceptable carrier such as a lotion, cream, ointment or soap. In some embodiments, useful are carriers capable of forming a film or layer over the skin to localize application and inhibit removal. Where adhesion to a tissue surface is desired the composition can include the therapeutic dispersed in a fibrinogen-thrombin composition or other bioadhesive. The therapeutic then can be painted, sprayed or otherwise applied to the desired tissue surface. For topical administration to internal tissue surfaces, the agent can be dispersed in a liquid tissue adhesive or other substance known to enhance adsorption to a tissue surface. For example, hydroxypropylcellulose or fibrinogen/thrombin solutions can be used to advantage. Alternatively, tissue-coating solutions, such as pectin-containing formulations can be used.

[0088] For inhalation treatments, such as for asthma, inhalation of powder (self-propelling or spray formulations) dispensed with a spray can, a nebulizer, or an atomizer can be used. Such formulations can be in the form of a finely comminuted powder for pulmonary administration from a powder inhalation device or self-propelling powder-dispensing formulations. In the case of self-propelling solution and spray formulations, the effect can be achieved either by choice of a valve having the desired spray characteristics (i.e., being capable of producing a spray having the desired particle size) or by incorporating the active ingredient as a suspended powder in controlled particle size. For administration by inhalation, the therapeutics also can be delivered in the form of an aerosol spray from a pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer. Nasal drops also can be used.

[0089] Systemic administration also can be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants generally are known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and filsidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or

suppositories. For transdermal administration, the therapeutics typically are formulated into ointments, salves, gels, or creams as generally known in the art.

[0090] In one embodiment, the therapeutics are prepared with carriers that will protect against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials also can be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Pat. No. 4,522,811. Microsomes and microparticles also can be used.

[0091] Oral or parenteral compositions can be formulated in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

[0092] Generally, the therapeutics identified according to the invention can be formulated for parenteral or oral administration to humans or other mammals, for example, in therapeutically effective amounts, e.g., amounts which provide appropriate concentrations of the drug to target tissue for a time sufficient to induce the desired effect. Additionally, the therapeutics of the

present invention can be administered alone or in combination with other molecules known to have a beneficial effect on the particular disease or indication of interest. By way of example only, useful cofactors include symptom-alleviating cofactors, including antiseptics, antibiotics, antiviral and antifungal agents and analgesics and anesthetics.

[0093] The effective concentration of the therapeutics identified according to the invention that is to be delivered in a therapeutic composition will vary depending upon a number of factors, including the final desired dosage of the drug to be administered and the route of administration. The preferred dosage to be administered also is likely to depend on such variables as the type and extent of disease or indication to be treated, the overall health status of the particular patient, the relative biological efficacy of the therapeutic delivered, the formulation of the therapeutic, the presence and types of excipients in the formulation, and the route of administration. In some embodiments, the therapeutics of this invention can be provided to an individual using typical dose units deduced from the earlier-described mammalian studies using non-human primates and rodents. As described above, a dosage unit refers to a unitary, i.e. a single dose which is capable of being administered to a patient, and which can be readily handled and packed, remaining as a physically and biologically stable unit dose comprising either the therapeutic as such or a mixture of it with solid or liquid pharmaceutical diluents or carriers.

[0094] In certain embodiments, organisms are engineered to produce the therapeutics identified according to the invention. These organisms can release the therapeutic for harvesting or can be introduced directly to a patient. In another series of embodiments, cells can be utilized to serve as a carrier of the therapeutics identified according to the invention.

[0095] Therapeutics of the invention also include the "prodrug" derivatives. The term prodrug refers to a pharmacologically inactive (or partially inactive) derivative of a parent molecule that requires biotransformation, either spontaneous or enzymatic, within the organism

to release or activate the active component. Prodrugs are variations or derivatives of the therapeutics of the invention which have groups cleavable under metabolic conditions. Prodrugs become the therapeutics of the invention which are pharmaceutically active *in vivo*, when they undergo solvolysis under physiological conditions or undergo enzymatic degradation. Prodrug of this invention can be called single, double, triple, and so on, depending on the number of biotransformation steps required to release or activate the active drug component within the organism, and indicating the number of functionalities present in a precursor-type form. Prodrug forms often offer advantages of solubility, tissue compatibility, or delayed release in the mammalian organism (see, Bundgard, Design of Prodrugs, pp. 7-9, 21-24, Elsevier, Amsterdam 1985 and Silverman, The Organic Chemistry of Drug Design and Drug Action, pp. 352-401, Academic Press, San Diego, Calif., 1992). Moreover, the prodrug derivatives according to this invention can be combined with other features to enhance bioavailability.

EXAMPLES

EXAMPLE 1. Fusion protein expressing constructs.

[0096] Nucleic acid constructs for expressing therapeutic protein fusions of the invention can be made recombinantly according to methods known in the art. For example, oligonucleotides complementary to genes encoding the different components described herein can be used to make synthetic genes or to amplify the natural genes and construct gene fusions. In preferred embodiments, proteins of the invention are expressed from a recombinant gene comprising a signal sequence. Examples of useful nucleic acids include nucleic acids that encode IGF targeting moieties of the invention. Such nucleic acids can be based on the sequences of IGF-1 shown in Figure 3.

EXAMPLE 2. Expression and purification methods.

[0097] Expression product can also be isolated from serum free media using other protozoa, including other *Leishmania* species. In general, the expression strain is grown in medium with serum, diluted into serum free medium, and allowed to grow for several generations, preferably 2-5 generations, before the expression product is isolated. For example, production of secreted recombinant LSD proteins can be isolated from *Leishmania mexicana* promastigotes that are cultured initially in 50 mL 1X M199 medium in a 75 cm2 flask at 27° C. When the cell density reaches 1-3x 10^7/mL, the culture is used to inoculate 1.2 L of M199 media. When the density of this culture reaches about 5x10^6/mL, the cells were harvested by centrifugation, resuspended in 180 mL of the supernatant and used to inoculate 12 L of "Zima" medium in a 16 L spinner flask. The initial cell density of this culture is typically about 5x 10^5 /mL. This culture is expanded to a cell density of about 1.0 – 1.7 x 10e^7 cells/mL. When this cell density is reached, the cells are separated from the culture medium by centrifugation and the supernatant is filtered at 4°C through a 0.2 μ filter to remove residual promastigotes. The filtered media was concentrated from 12.0 L to 500 mL using a tangential flow filtration device (MILLIPORE Prep/Scale-TFF cartridge).

[0098] Preferred growth media for this method are M199 and "Zima" growth media. However, other serum containing and serum free media are also useful. M199 growth media is as follows: (1L batch) = 200 mL 5X M199 (with phenol pH indicator) mixed at 5X + 637 mL H_2O, 50.0 mL FBS, 50.0 mL EF, 20.0 mL of 50g/mL SAT, 2.0 mL of 0.25% hemin in 50% triethanolamine , 10 mL of 10mM adenine in 50mM Hepes pH 7.5, 40.0 mL of 1M Hepes pH 7.5, 1mL of 0.1% biotin in 95% ethanol, 10.0 mL of penicillin/streptomycin. All serums used are inactivated by heat. The final volume = 1 L and is filter sterilized. "Zima" modified M199 media is as follows: (20.0 L batch) = 217.8g M199 powder (–)phenol red + 7.0g sodium bicarbonate, 200.0 mL of 10mM adenine in 50mM Hepes pH 7.5, 800.0 mL 0f Hepes free acid

pH 7.5, 20.0 mL 0.1% biotin in 95% ethanol, 200.0 mL penicillin/streptomycin, 2780.0 mL H20 Final volume = 20.0 L and is filter sterilized.

[0099] According to one aspect of the invention, LSD proteins secreted from *Leishmania* and containing carbohydrate with terminal mannose residues can be purified as follows. For example, recombinant ß-glucuronidase from *Leishmania mexicana* containing plasmsid pXSAP0-GUS was grown in M199 culture medium with a small amount of serum proteins. When the culture reached a density of > 1.0 x 10^7 promastigotes/mL the *L. mexicana* were removed by centrifugation, 10 min at 500 x g. The harvested culture medium was passed through a 0.2 µm filter to remove particulates before being loaded directly onto a Concanavalin A (ConA)-agarose column (4% cross-linked beaded agarose, Sigma). The ConA-agarose column was pretreated with 1 M NaCl, 20 mM Tris pH 7.4, 5 mM each of $CaCl_2$, $MgCl_2$ and $MnCl_2$ and then equilibrated with 5 volumes of column buffer (20 mM Tris pH 7.4, 1 mM $CaCl_2$, and 1 mM $MnCl_2$). A total of 179,800 units (nmol/hr) of GUS activity (in 2 L) in culture medium was loaded onto a 22 mL ConA agarose column. No activity was detectable in the flow through or wash. The GUS activity was eluted with column buffer containing 200 mM methyl mannopyranoside. Eluted fractions containing the activity peak were pooled and concentrated: 143900 units of GUS activity were recovered from the column (80% recovery of activity loaded onto the column). This demonstrates that the recombinant ß-GUS secreted from *L. mexicana* possesses carbohydrate with terminal mannose residues and further points out the potential for using the interaction of mannose with ConA as the basis for an affinity purification step. Accordingly, the presence of high mannose carbohydrate can serve as the basis of an affinity step in the purification of recombinant LSD proteins using lectin affinity chromatography.

EXAMPLE 3. Assays for crossing the blood brain barrier.

[0100] According to the invention, a useful model system to determine whether a protein, particularly an LSD protein, tagged with an IGF tag crosses the blood-brain barrier, is the MPSVII mouse model (Wolfe and Sands (1996) Protocols for Gene Transfer in Neuroscience: Towards Gene Therapy of Neurological Disorders Chapter 20: 263-273). For example, recombinant human ß-glucuronidase fused to an IGF tag can be produced in any convenient expression system such as *Leishmania*, yeast, mammalian, bacculovirus and other expression systems. *L. mexicana* expressing and secreting ß-GUS is grown at 26°C in 100 ml Standard Promastigote medium (M199 with 40 mM HEPES, pH 7.5, 0.1 mM adenine, 0.0005% hemin, 0.0001% biotin, 5% fetal bovine serum, 5% embryonic fluid, 50 units/ml penicillin, 50 µg/ml streptomycin and 50µg/ml nourseothricin). After reaching a density of approximately 5 x 10^6 promastigotes/ml, the promastigotes is collected by centrifugation for 10 min. at 1000 x g at room temperature; these promastigotes were used to inoculate 1 liter of low protein medium (M199 supplemented with 0.1 mM adenine, 0.0001% biotin, 50 units/ml penicillin and 50 µg/ml streptomycin) at room temperature. The 1 liter cultures are contained in 2 liter capped flasks with a sterile stir bar so that the cultures could be incubated at 26°C with gentle stirring. The 1 liter cultures are aerated twice a day by moving them into a laminar flow hood, removing the caps and swirling vigorously before replacing the caps. When the cultures reach a density of 2-3 x 10^7 promastigotes/ml, the cultures are centrifuged as previously described except the promastigote pellet is discarded and the media decanted into sterile flasks. The addition of 434 g $(NH_4)_2SO_4$ per liter precipitates active GUS protein from the medium; the salted out medium is stored at 4°C overnight. Precipitated proteins are harvested either by centrifugation at 10,500 x g for 30 min. or filtration through Gelman Supor-800 membrane; the proteins are resuspended in 10 mM Tris pH 8, 1 mM $CaCl_2$ and stored at -80°C until dialysis. The crude preparations from

several liters of medium are thawed, pooled, placed in dialysis tubing (Spectra/Por -7, MWCO 25,000), and dialyzed overnight against two 1 liter volumes of DMEM with bicarbonate (Dulbecco's Modified Eagle's Medium). The ammonium sulfate fraction is further purified on a ConA column.

[0101] GUS minus mice generated by heterozygous matings of B6.C-H-2^{bml}/ByBIR-gus^{mps}/+ mice are used to assess the effectiveness of GUS-IGF fusion proteins or derivatives in enzyme replacement therapy. Two formats are used. In one format, 3-4 animals are given a single injection of 20,000U of enzyme in 100 µl enzyme dilution buffer (150 mM NaCl, 10 mM Tris, pH7.5). Mice are killed 72-96 hours later to assess the efficacy of the therapy. In a second format, mice are given weekly injections of 20,000 units over 3-4 weeks and are killed 1 week after the final injection. Histochemical and histopathologic analysis of liver, spleen and brain are carried out by published methods. In the absence of therapy, cells (*e.g.* macrophages and Kupffer cells) of GUS minus mice develop large intracellular storage compartments resulting from the buildup of waste products in the lysosomes. It is anticipated that in cells in mice treated with GUS fusion constructs of the invention, the size of these compartments will be visibly reduced or the compartments will shrink until they are no longer visible with a light microscope.

[0102] According to the invention, newborn mice do not possess a complete blood brain barrier. However, by day 15 the blood brain barrier is formed to the point that ß–glucuronidase no longer can be detected in the brain. Accordingly, the above experiments are preferably performed on mice that are at day 15 or greater.

[0103] According to one embodiment of the invention, experiments first assess the ability of complete IGF-I and IGF-II tags to direct proteins across the blood brain barrier. Next, specific mutant versions of the proteins that disrupt receptor or IGF binding protein binding are assayed. For domain swaps, the B domain of IGF-II (residues 1-28 of the mature protein)

contains only two differences from IGF-I that could conceivably alter transport across the blood brain barrier G11 and T16. Altering these residues in IGF-II would is essentially a domain B swap. Another swap of regions between residues 28 and 41 of IGF-II and the corresponding region of IGF-I can also be tested. This essentially swaps the C domains of the two proteins which contains the most divergent regions of the two proteins. An alternative swap switches the C-terminal 15 residues with the corresponding region of IGF-I. These three chimeras provide an essentially complete picture of how any differences in uptake across the blood brain barrier between IGF-I and IGF-II correlate with sequence/structural differences between the two proteins.

EXAMPLE 4. Assays for protein accumulation in the brain or CNS.

[0104] Radioactive assays can be used to monitor the accumulation of protein product in the brain. For example, the uptake and accumulation of a radioactively labeled protein in the brain parenchyma can be assayed as disclosed in Reinhardt and Bondy (1994) Endocrinology 135:1753-1761.

[0105] Enzyme assays can also be used to monitor the accumulation of protein product in the brain. Enzyme assays are particularly useful when the therapeutic protein moiety is an enzyme for which there is an assay that is applicable for histochemical staining. Useful enzyme assays for lysosomal storage disease proteins include assays disclosed in Sly at al. (2001) P.N.A.S. 98(5): 2205-2210, and in Wolfe and Sands (1996) Protocols for Gene Transfer in Neuroscience: Towards Gene Therapy of Neurological Disorders Chapter 20: 263-273.

EXAMPLE 5. *In vivo* therapy.

[0106] GUS minus mice generated by heterozygous matings of B6.C-H-2^{bml}/ByBIR-gus^{mps}/+ mice (Birkenmeier, Davisson et al. 1989) are used to assess the effectiveness of compositions of the invention in enzyme replacement therapy. Two formats are used. In one format, 3-4 animals are given a single injection of 20,000U of enzyme in 100 µl enzyme dilution buffer (150 mM NaCl, 10 mM Tris, pH7.5). Mice are killed 72-96 hours later to assess the efficacy of the therapy. In a second format, mice are given weekly injections of 20,000 units over 3-4 weeks and are killed 1 week after the final injection. Histochemical and histopathologic analysis of liver, spleen and brain are carried out by published methods (Birkenmeier, Barker et al. 1991; Sands, Vogler et al. 1994; Daly, Vogler et al. 1999). In the absence of therapy, cells (*e.g.* macrophages and Kupffer cells) of GUS minus mice develop large intracellular storage compartments resulting from the buildup of waste products in the lysosomes. It is anticipated that in cells in mice treated with compositions of the invention, the size of these compartments will be visibly reduced or the compartments will shrink until they are no longer visible with a light microscope.

[0107] Similarly, humans with lysosomal storage diseases will be treated using constructs targeting an appropriate therapeutic portion to their CNS and in particular to lysosomes within the CNS. In some instances, treatment will take the form of regular (*e.g.* weekly) injections of a fusion protein of the invention. In other instances, treatment will be achieved through administration of a nucleic acid to permit persistent *in vivo* expression of the fusion protein, or through administration of a cell (*e.g.* a human cell, or a unicellular organism) expressing the fusion protein in the patient. For example, a protein the invention may be expressed *in situ* using a *Leishmania* vector as described in U.S. Patent No. 6,020,144, issued February 1, 2000; and PCT Serial No. PCT/US01/44935, filed November 30, 2001.

EXAMPLE 6. Enzyme assays.

[0108] Assays for GUS activity are performed as described in Example 3 and/or as described below.

[0109] Glass assay tubes are numbered in triplicate, and 100 µL of 2x GUS reaction mix are added to each tube. 2x GUS reaction mix is prepared by adding 100 mg of 4-methylumbelliferyl-ß-D glucuronide to 14.2 mL 200 mM sodium acetate, pH adjusted to 4.8 with acetic acid. Up to 100 µL of sample are added to each tube; water is added to a final reaction volume of 200 µL. The reaction tubes are covered with parafilm and incubated in a 37°C water bath for 1-2 hours. The reaction is stopped by addition of 1.8 mL of stop buffer (prepared by dissolving 10.6 g of Na_2CO_3 and 12.01 g of glycine in a final volume of 500 mL of water, adjusting the pH to 10.5 and filter-sterilizing into a repeat-dispensor). A fluorimeter is then calibrated using 2 mL of stop solution as a blank, and the fluorescence is read from the remaining samples. A standard curve is prepared using 1, 2, 5, 10, and 20 µL of a 166 µM 4-methylumbelliferone standard in a final volume of 2 mL stop buffer.

[0110] The 4-methylumbelliferone standard solution is prepared by dissolving 2.5 mg 4-methylumbelliferone in 1 mL ethanol and adding 99 mL of sterile water, giving a concentration of approximately 200 nmol/mL. The precise concentration is determined spectrophotometrically. The extinction coefficient at 360 nm is 19,000 cm^{-1} M^{-1}. For example, 100 µL is added to 900 µL of stop buffer, and the absorbance at 360 nm is read. If the reading is 0.337, then the concentration of the standard solution is 0.337 x 10 (dilution)/19,000 = 177 µM, which can then be diluted to 166 µM by addition of an appropriate amount of sterile water.

EXAMPLE 7. Protein production in mammalian cells.

CHO cells

[0111] In one example, GUS-GILTΔ1-7 (an IGF-II targeting portion with a deleton of the N-terminal 7 amino acids) and GUSΔC18-GILTΔ1-7 (a fusion protein in which GUSΔC18 is fused to the N-terminal of GILTΔ1-7 were expressed in CHO cells using the system of Ulmasov *et al.* (2000) PNAS 97(26):14212-14217. Appropriate gene cassettes were inserted into the *Eco* RI site of the pCXN vector, which was electroporated into CHO cells at 50 μF and 1,200 V in a 0.4-cm cuvette. Selection of colonies and amplification was mediated by 400 μg/mL G418 for 2-3 weeks. The CHO cells were propagated in MEM media supplemented with 15% FBS, 1.2 mM glutamine, 50 μg/mL proline, and 1 mM pyruvate. For enzyme production cells were plated in multifloor flasks in MEM. Once cells reached confluence, collection medium (Weymouth medium supplemented with 2% FBS, 1.2 mM glutamine, and 1 mM pyruvate) was applied to the cells. Medium containing the secreted recombinant enzyme was collected every 24-72 hours. A typical level of secretion for one GUS-GILTΔ1-7 cell line was 4000-5000 units/mL/24 hours.

[0112] A number of GUSΔC18-GILTΔ1-7 CHO lines were assayed for the amount of secreted enzyme produced. The six highest producers secreted between 8600 and 14900 units/mL/24 hours. The highest producing line was selected for collection of protein.

HEK 293 cells

[0113] GUS-GILT cassettes were cloned into pCEP4 (Invitrogen) for expression in HEK 293 cells. Cassettes used included wild-type GUS-GILT; GUS-GILTΔ1-7; GUS-GILTY27L; GUSΔC18-GILTΔ1-7; GILTY27L, and GUS-GILTF19S/E12K.

[0114] HEK 293 cells were cultured to 50-80% confluency in 12-well plates containing DMEM medium with 4 mM glutamine and 10% FBS. Cells were transfected with pCEP-GUS-GILT DNA plasmids using FuGENE 6 (Roche) as described by the manufacturer. 0.5 μg DNA and 2 μL of FuGENE 6 were added per well. Cells were removed from wells 2-3 days post-

transfection using trypsin, then cultured in T25 cm^2 culture flasks containing the above DMEM medium with 100 μg/mL hygromycin to select for a stable population of transfected cells. Media containing hygromycin were changed every 2-3 days. The cultures were expanded to T75 cm^2 culture flasks within 1-2 weeks. For enzyme production cells were plated in multifloor flasks in DMEM. Once cells reached confluence, collection medium (Weymouth medium supplemented with 2% FBS, 1.2 mM glutamine, and 1 mM pyruvate) was applied to the cells. This medium has been optimized for CHO cells, not for 293 cells; accordingly, levels of secretion with the HEK 293 lines may prove to be significantly higher in alternate media.

[0115] Levels of secreted enzyme are shown in Table 2.

[0116] Table 2.

Cell line	Recombinant Protein	Units/mL/24 hours
HEK293 2-1	GUS-GILT	3151
HEK293 2-2	GUSΔC18-GILTΔ1-7	10367
HEK293 2-3	GUS-GILTΔ1-7	186
HEK293 4-4	GILTY27L	3814
HEK293 3-5	GUS-GILTF19S/E12K	13223
HEK293 3-6	GILTY27L	7948
CHO 15	GUSΔC18-GILTΔ1-7	18020

EXAMPLE 8. Purification of GUS fusion proteins.

[0117] Chromatography, including conventional chromatography and affinity chromatography, can be used to purify GUS fusion proteins such as GUS-GILT fusion proteins and other fusion proteins of the invention.

Conventional chromatography

[0118] One procedure for purifying GUS fusion proteins produced in *Leishmania* is described in Example 2. An alternative procedure is described in the following paragraph.

[0119] Culture supernatants from *Leishmania mexicana* cell lines expressing GUS-GILT fusions were harvested, centrifuged, and passed through a 0.2 μ filter to remove cell debris. The supernatants were concentrated using a tangential ultrafilter with a 100,000 molecular weight cutoff and stored at -80°C. Concentrated supernatants were loaded directly onto a column containing Concanavalin A (ConA) immobilized to beaded agarose. The column was washed with ConA column buffer (50 mM Tris pH 7.4, 1mM $CaCl_2$, 1mM $MnCl_2$) before mannosylated proteins including GUS-GILT fusions were eluted using a gradient of 0-0.2M methyl-α-D-pyranoside in the ConA column buffer. Fractions containing glucuronidase activity were pooled, concentrated, and the buffer exchanged to SP column buffer (25 mM sodium phosphate pH 6, 20 mM NaCl, 1 mM EDTA) in preparation for the next column. The concentrated fractions were loaded onto an SP fast flow column equilibrated in the same buffer, and the column was washed with additional SP column buffer. The GUS-GILT fusions were eluted from the column in two steps: 1) a gradient of 0-0.15 M glucuronic acid in 25 mM sodium phosphate pH 6 and 10% glycerol, followed by 0.2 M glucuronic acid, 25 mM sodium phosphate pH 6, 10% glycerol. Fractions containing glucuronidase activity were pooled, and the buffer exchanged to 20 mM potassium phosphate pH 7.4. These pooled fractions were loaded onto an HA-ultrogel column equilibrated with the same buffer. The GUS-GILT fusion proteins were eluted with an increasing gradient of phosphate buffer, from 145-340 mM potassium phosphate pH 7.4. The fractions containing glucuronidase activity were pooled, concentrated, and stored at -80°C in 20 mM Tris pH 8 with 25% glycerol.

[0120] A conventional chromatography method for purifying GUS-GILT fusion proteins produced in mammalian cells is described in the following paragraphs.

[0121] Mammalian cells overexpressing a GUS-GILT fusion protein are grown to confluency in Nunc Triple Flasks, then fed with serum-free medium (Waymouth MB 752/1)

supplemented with 2% fetal bovine serum to collect enzyme for purification. The medium is harvested and the flasks are refed at 24 hour intervals. Medium from several flasks is pooled and centrifuged at 5000 x g for 20 minutes at 4°C to remove detached cells, *etc.* The supernatant is removed and aliquots are taken for a β-GUS assay. The medium can now be used directly for purification or frozen at -20°C for later use.

[0122] 1 L of secretion medium is thawed at 37°C (if frozen), filtered through a 0.2 μ filter, and transferred to a 4L beaker. The volume of the medium is diluted 4-fold by addition of 3 L of dd water to reduce the salt concentration; the pH of the diluted medium is adjusted to 9.0 using 1 M Tris base. 50 mL of DEAE-Sephacel pre-equilibrated with 10 mM Tris pH 9.0 is added to the diluted medium and stirred slowly with a large stirring bar at 4°C for 2 hours. (A small aliquot can be removed, microfuged, and the supernatant assayed to monitor binding.) When binding is complete, the resin is collected on a fritted glass funnel and washed with 750 mL of 10 mM Tris pH 9.0 in several batches. The resin is transferred to a 2.5 cm column and washed with an additional 750 mL of the same buffer at a flow rate of 120 mL/hour. The DEAE column is eluted with a linear gradient of 0-0.4 M NaCl in 10 mM Tris pH 9.0. The fractions containing the GUS-GILT fusion proteins are detected by 4-methylumbelliferyl-ß-D glucuronide assay, pooled, and loaded onto a 600 mL column of Sephacryl S-200 equilibrated with 25 mM Tris pH 8, 1 mM β-glycerol phosphate, 0.15 M sodium chloride and eluted with the same buffer.

[0123] The fractions containing the GUS-GILT fusion proteins are pooled and dialyzed with 3 x 4L of 25 mM sodium acetate pH 5.5, 1 mM β-glycerol phosphate, 0.025% sodium azide. The dialyzed enzyme is loaded at a flow rate of 36 mL/hour onto a 15 mL column of CM-Sepharose equilibrated with 25 mM sodium acetate pH 5.5, 1 mM β-glycerol phosphate, 0.025% sodium azide. It is then washed with 10 column volumes of this same buffer. The CM column is eluted with a linear gradient of 0-0.3 M sodium chloride in the equilibration buffer. The

fractions containing the GUS-GILT fusion proteins are pooled and loaded onto a 2.4 x 70 cm (Bed volume = 317 mL) column of Sephacryl S-300 equilibrated with 10 mM Tris pH 7.5, 1 mM β-glycerol phosphate, 0.15 M NaCl at a flow rate of 48 mL/hour. The fractions containing the fusion proteins are pooled; the pool is assayed for GUS activity and for protein concentration to determine specific activity. Aliquots are run on SDS-PAGE followed by Coomassie or silver staining to confirm purity. If a higher concentration of enzyme is required, Amicon Ultrafiltration Units with an XM-50 membrane (50,000 molecular weight cutoff) or Centricon C-30 units (30,000 molecular weight cutoff) can be used to concentrate the fusion protein. The fusion protein is stored at -80°C in the 10 mM Tris pH 7.5, 1 mM sodium β-glycerol phosphate, 0.15 M NaCl buffer.

Affinity chromatography

[0124] Affinity chromatography conditions are essentially as described in Islam *et al.* (1993) J. Biol. Chem. 268(30):22627-22633. Conditioned medium from mammalian cells overexpressing a GUS-GILT fusion protein (collected and centrifuged as described above for conventional chromatography) is filtered through a 0.22 μ filter. Sodium chloride (crystalline) is added to a final concentration of 0.5M, and sodium azide is added to a final concentration of 0.025% by adding 1/400 volume of a 10% stock solution. The medium is applied to a 5 mL column of anti-human β-glucuronidase-Affigel 10 (pre-equilibrated with Antibody Sepharose Wash Buffer: 10 mM Tris pH 7.5, 10 mM potassium phosphate, 0.5 M NaCl, 0.025% sodium azide) at a rate of 25 mL/hour at 4°C. Fractions are collected and monitored for any GUS activity in the flow-through. The column is washed at 36 mL/hour with 10-20 column volumes of Antibody Sepharose Wash Buffer. Fractions are collected and monitored for GUS activity. The column is eluted at 36 mL/hour with 50 mL of 10 mM sodium phosphate pH 5.0 + 3.5 M

$MgCl_2$. 4 mL fractions are collected and assayed for GUS activity. Fractions containing the fusion protein are pooled, diluted with an equal volume of P6 buffer (25 mM Tris pH 7.5, 1 mM β-glycerol phosphate, 0.15 mM NaCl, 0.025% sodium azide) and desalted over a BioGel P6 column (pre-equilibrated with P6 buffer) to remove the $MgCl_2$ and to change the buffer to P6 buffer for storage. The fusion protein is eluted with P6 buffer, fractions containing GUS activity are pooled, and the pooled fractions assayed for GUS activity and for protein. An SDS-PAGE gel stained with Coomassie Blue or silver stain is used to confirm purity. The fusion protein is stored frozen at -80°C in P6 buffer for long-term stability.

EXAMPLE 9. Uptake experiments on mammalian-produced proteins.

[0125] Culture supernatants from HEK293 cell lines or CHO cell lines producing GUS or GUS-GILT constructs were harvested through a 0.2 μm filter to remove cells GM 4668 fibroblasts were cultured in 12-well tissue culture plates in DMEM supplemented with 15% (v/v) fetal calf serum at 37°C in 5% CO_2. Cells were washed once with uptake medium (DMEM + 2% BSA (Sigma A-7030)) at 37°C. Fibroblasts were then cultured (3-21 hours) with 1000-4000 units of enzyme per mL of uptake medium. In some experiments, competitors for uptake were added. Mannose-6-phosphate (Calbiochem 444100) was added to some media at concentrations from 2-8 mM and pure recombinant IGF-II (Cell Sciences OU100) was added to some media at 2.86 mM, representing a 10-100 fold molar excess depending on the quantity of input enzyme. Uptake was typically measured in triplicate wells.

[0126] After incubation, the media were removed from the wells and assayed in duplicate for GUS activity. Wells were washed five times with 1 mL of 37°C phosphate-buffered saline, then incubated for 15 minutes at room temperature in 0.2 mL of lysis buffer (10 mM Tris, pH 7.5, 100 mM NaCl, 5 mM EDTA, and 1% NP-40). Cell lysates were transferred to microfuge tubes and spun at 13,000 rpm for 5 minutes to remove cell debris. Two 10 μL aliquots of lysate

were assayed for GUS activity using a standard fluorometric assay. Three 10 μL aliquots of lysate were assayed for protein concentration (Pierce Micro BCA protein assay, Pierce, IL).

[0127] An initial experiment compared uptake of CHO-produced GUS-GILTΔ1-7 with CHO-produced GUSΔC18-GILTΔ1-7. As shown in Table 3, the GUSΔC18-GILTΔ1-7 protein, which was engineered to eliminate a potential protease cleavage site, has significantly higher levels of uptake levels that can be inhibited by IGF-II and by M6P. In contrast, the uptake of a recombinant GUS produced in mammalian cells lacking the IGF-II tag was unaffected by the presence of excess IGF-II but was completely abolished by excess M6P. In this experiment, uptake was performed for 18 hours.

[0128] Table 3.

Enzyme	Input units	Uptake (units/mg)	+IGF-II (units/mg)	% IGF-II inhibition	+M6P (units/mg)	% M6P inhibition
CHO GUS-GILTΔ1-7	982	310±27	84±20	73	223±36	28
CHO GUSΔC18-GILTΔ1-7	1045	704±226	258±50	63	412±79	41
CHO GUS	732	352±30	336±77	5	1±0.2	99.7

[0129] A subsequent experiment assessed the uptake of CHO- and HEK293-produced enzymes by human fibroblasts from MPSVII patients. In this experiment, uptake was for 21 hours.

[0130] TABLE 4.

Enzyme	Input units	Uptake (units/mg)	+IGF-II Uptake (units/mg)	% IGF-II inhibition
CHO GUSΔC18-GILTΔ1-7	2812	4081±1037	1007±132	75
HEK GUS-GILT	2116	1432±196		
HEK GUSΔC18-GILTΔ1-7	3021	5192±320	1207±128	77
HEK GUS-GILTY27L	3512	1514±203		
HEK GUS-GILTF19SE12K	3211	4227±371	388±96	90.8
HEK GUS-GILTF19S	3169	4733±393	439±60	90.7

[0131] A further experiment assessed the uptake of selected enzymes in the presence of IGF-II, 8mM M6P, or both inhibitors. Uptake was measured for a period of 22.5 hours.

[0132] TABLE 5.

Enzyme	Input units	Uptake (units/ mg)	+IGF-II (units/ mg)	% IGF-II inhibition	+M6P (units/ mg)	%M6P inhibition	+IGF-II +M6P (units/ mg)	%IGF-II+M6P inhibition
CHO GUSΔC18-GILTΔ1-7	1023	1580± 150	473±27	70	639±61	60	0±1	100
HEK GUS-GILTF19S E12K	880	1227± 76	22±2	98.2	846±61	31	0±3	100
HEK GUS-GILTF19S	912	1594± 236	217±17	86	952±96	60	15±2	99.06

[0133] The experiments described above show that CHO and HEK293 production systems are essentially equivalent in their ability to secrete functional recombinant proteins. The experiments also show that the presence of excess IGF-II diminishes uptake of tagged proteins by 70-90+%, but does not markedly affect uptake of untagged protein (4.5%), indicating specific IGF-II-mediated uptake of the mammalian-produced protein. Unlike *Leishmania*-produced proteins, the enzymes produced in mammalian cells are expected to contain M6P. The presence of two ligands on these proteins capable of directing uptake through the M6P/IGF-II receptor implies that neither excess IGF-II nor excess M6P should completely abolish uptake. Furthermore, since the two ligands bind to discrete locations on the receptor, binding to the receptor via one ligand should not be markedly affected by the presence of an excess of the other competitor.

[0134] Binding of GUS-GILT proteins to the M6P/IGF-II receptor on fibroblasts are measured and the rate of uptake is assessed similar to published methods (York *et al.* (1999) J. Biol. Chem. 274(2):1164-71). GM4668 fibroblasts cultured in 12 well culture dishes as described above are washed in ice-cold media minus serum containing 1% BSA. Ligand, (either

GUS, GUS-GILT or GUS-ΔGILT, or control proteins) is added to cells in cold media minus serum plus 1% BSA. Upon addition of ligand, the plates are incubated on ice for 30 minutes. After 30 minutes, ligand is removed and cells are washed quickly 5 times with ice cold media. Wells for the 0 time point receive 1 ml ice cold stripping buffer (0.2 M Acetic acid, pH 3.5, 0.5M NaCl). The plate is then floated in a 37° water bath and 0.5 ml prewarmed media is added to initiate uptake. At every stopping point, 1 ml of stripping buffer is added. When the experiment is over, aliquots of the stripping buffer are saved for fluorometric assay of ß-glucuronidase activity as described in Example 3. Cells are then lysed as described above and the lysate assayed for ß-glucuronidase activity. Alternatively, immunological methods can be used to test the lysate for the presence of the targeted therapeutic protein.

[0135] It is expected that GUS-GILT is rapidly taken up by fibroblasts in a matter of minutes once the temperature is shifted to 37°C (York *et al.* (1999) J. Biol. Chem. 274(2):1164-71) and that the enzyme activity persists in the cells for many hours.

EXAMPLE 10. *In vivo* therapy

[0136] Initially, GUS minus mice can be used to assess the effectiveness of GUS-GILT and derivatives thereof in enzyme replacement therapy. GUS minus mice are generated by heterozygous matings of B6.C-*H*-2^{bm1}/ByBIR-gus^{mps}/+ mice as described by Birkenmeier *et al.* (1989) J. Clin. Invest 83(4):1258-6. Preferably, the mice are tolerant to human β-GUS. The mice may carry a transgene with a defective copy of human β-GUS to induce immunotolerance to the human protein (Sly *et al.* (2001) PNAS 98:2205-2210). Alternatively, human β-GUS (*e.g.* as a GUS-GILT protein) can be administered to newborn mice to induce immunotolerance. However, because the blood-brain barrier is not formed until about day 15 in mice, it is simpler

to determine whether GILT-GUS crosses the blood-brain barrier when initiating injections in mice older than 15 days; transgenic mice are therefore preferable.

[0137] The initial experiment is to determine the tissue distribution of the targeted therapeutic protein. At least three mice receive a CHO-produced GILT-tagged β-GUS protein referred to herein as GUSΔC18-GILTΔ1-7, in which GUSΔ18, a β-GUS protein omitting the last eighteen amino acids of the protein, is fused to the N-terminus of Δ1-7 GILT, an IGF-II protein missing the first seven amino acids of the mature protein. Other mice receive either β-GUS, a buffer control, or a GUSΔC18-GILTΔ1-7 protein treated with periodate and sodium borohydride as described in Example 12. Generally, preferred doses are in the range of 0.5-7 mg/kg body weight. In one example, the enzyme dose is 1 mg/kg body weight administered intravenously, and the enzyme concentration is about 1-3 mg/mL. In addition, at least three mice receive a dose of 5 mg/kg body weight of GUSΔC18-GILTΔ1-7 protein treated with periodate and sodium borohydride. After 24 hours, the mice are sacrificed and the following organs and tissues are isolated: liver, spleen, kidney, brain, lung, muscle, heart, bone, and blood. Portions of each tissue are homogenized and the β-GUS enzyme activity per mg protein is determined as described in Sly *et al.* (2001) PNAS 98:2205-2210. Portions of the tissues are prepared for histochemistry and/or histopathology carried out by published methods (see, *e.g.*, Vogler *et al.* (1990) Am J. Pathol. 136:207-217).

[0138] Further experiments include multiple injection protocols in which the mice receive weekly injections at a dose of 1 mg/kg body weight. In addition, measurement of the half-life of the periodate-modified enzyme is determined in comparison with untreated enzyme as described in Example 12.

[0139] Two other assay formats can be used. In one format, 3-4 animals are given a single injection of 20,000U of enzyme in 100 µl enzyme dilution buffer (150 mM NaCl, 10 mM

Tris, pH7.5). Mice are killed 72-96 hours later to assess the efficacy of the therapy. In a second format, mice are given weekly injections of 20,000 units over 3-4 weeks and are killed 1 week after the final injection. Histochemical and histopathologic analysis of liver, spleen and brain are carried out by published methods (Birkenmeier *et al.* (1991) Blood 78(11):3081-92; Sands *et al.* (1994) J. Clin. Invest 93(6):2324-31; Daly *et al.* (1999) Proc. Natl. Acad. Sci. USA 96(5):2296-300). In the absence of therapy, cells (*e.g.* macrophages and Kupffer cells) of GUS minus mice develop large intracellular storage compartments resulting from the buildup of waste products in the lysosomes. It is anticipated that in cells in mice treated with GUS-GILT constructs, the size of these compartments will be visibly reduced or the compartments will shrink until they are no longer visible with a light microscope.

[0140] Similarly, humans with lysosomal storage diseases will be treated using constructs targeting an appropriate therapeutic portion to their lysosomes. In some instances, treatment will take the form of regular (*e.g.* weekly) injections of a GILT protein. In other instances, treatment will be achieved through administration of a nucleic acid to permit persistent *in vivo* expression of a GILT protein, or through administration of a cell (*e.g.* a human cell, or a unicellular organism) expressing the GILT protein in the patient. For example, the GILT protein can be expressed *in situ* using a *Leishmania* vector as described in U.S. Patent No. 6,020,144, issued February 1, 2000; U.S. Provisional Application No. 60/250,446, filed November 30, 2001; U.S. Provisional Application No. 60/290,281, filed May 11, 2001; and PCT Serial No. PCT/US01/44935, filed November 30, 2001.

[0141] Targeted therapeutic proteins of the invention can also be administered, and their effects monitored, using methods (enzyme assays, histochemical assays, neurological assays, survival assays, reproduction assays, *etc.*) previously described for use with GUS. See, for example, Vogler *et al.* (1993) Pediatric Res. 34(6):837-840; Sands *et al.* (1994) J. Clin. Invest.

93:2324-2331; Sands *et al.* (1997) J. Clin. Invest. 99:1596-1605; O'Connor *et al.* (1998) J. Clin. Invest. 101:1394-1400; and Soper *et al.* (1999) 45(2):180-186.

EXAMPLE 11.

[0142] The objective of these experiments is to evaluate the efficacy of GILT-modified alpha-galactosidase A (α-GAL A) as an enzyme replacement therapy for Fabry's disease.

[0143] Fabry's disease is a lysosomal storage disease resulting from insufficient activity of α-GAL A, the enzyme responsible for removing the terminal galactose from GL-3 and other neutral sphingolipids. The diminished enzymatic activity occurs due to a variety of missense and nonsense mutations in the x-linked gene. Accumulation of GL-3 is most prevalent in lysosomes of vascular endothelial cells of the heart, liver, kidneys, skin and brain but also occurs in other cells and tissues. GL-3 buildup in the vascular endothelial cells ultimately leads to heart disease and kidney failure.

[0144] Enzyme replacement therapy is an effective treatment for Fabry's disease, and its success depends on the ability of the therapeutic enzyme to be taken up by the lysosomes of cells in which GL-3 accumulates. The Genzyme product, Fabrazyme, is recombinant α-GAL A produced in DUKX B11 CHO cells that has been approved for treatment of Fabry's patients in Europe due to its demonstrated efficacy.

[0145] The ability of Fabrazyme to be taken up by cells and transported to the lysosome is due to the presence of mannose 6-phosphate (M6P) on its N-linked carbohydrate. Fabrazyme is delivered to lysosomes through binding to the mannose-6-phosphate/IGF-II receptor (M6P/IGF-IIr), present on the cell surface of most cell types, and subsequent receptor mediated endocytosis. Fabrazyme reportedly has three N-linked glycosylation sites at ASN residues 108, 161, and 184. The predominant carbohydrates at these positions are fucosylated biantennary

bisialylated complex, monophosphorylated mannose-7 oligomannose, and biphosphorylated mannose-7 oligomannose, respectively.

[0146] The glycosylation independent lysosomal targeting (GILT) technology of the present invention directly targets therapeutic proteins to the lysosome via a different interaction with the M6P/IGF-Iir. A targeting ligand is derived from mature human IGF-II, which also binds with high affinity to the M6P/IGF-Iir. In current applications, the IGF-II tag is provided as a c-terminal fusion to the therapeutic protein, although other configurations are feasible including cross-linking. The competency of GILT-modified enzymes for uptake into cells has been established using GILT-modified ß-glucuronidase, which is efficiently taken up by fibroblasts in a process that is competed with excess IGF-II. Advantages of the GILT modification are increased binding to the M6P/IGF-II receptor, enhanced uptake into lysosomes of target cells, altered or improved pharmacokinetics, and expanded, altered or improved range of tissue distribution. The improved range of tissue distributions could include delivery of GILT-modified α-GAL A across the blood-brain barrier since IGF proteins demonstrably cross the blood-brain barrier.

[0147] Another advantage of the GILT system is the ability to produce uptake-competent proteins in non-mammalian expression systems where M6P modifications do not occur. In certain embodiments, GILT-modified protein will be produced primarily in CHO cells. In certain others, the GILT tag will be placed at the c-terminus of α-GAL A although the invention is not so limited.

EXAMPLE 12. Underglycosylated therapeutic proteins.

[0148] The efficacy of a targeted therapeutic can be increased by extending the serum half-life of the targeted therapeutic. Hepatic mannose receptors and asialoglycoprotein receptors

eliminate glycoproteins from the circulation by recognizing specific carbohydrate structures (Lee *et al.* (2002) Science 295(5561):1898-1901; Ishibashi *et al.* (1994) J. Biol. Chem. 269(45):27803-6). In some embodiments, the present invention permits targeting of a therapeutic to lysosomes and/or across the blood brain barrier in a manner dependent not on a carbohydrate, but on a polypeptide or an analog thereof. Actual underglycosylation of these proteins is expected to greatly increase their half-life in the circulation, by minimizing their removal from the circulation by the mannose and asialoglycoprotein receptors. Similarly, functional deglycosylation (*e.g.* by modifying the carbohydrate residues on the therapeutic protein, as by periodate/ sodium borohydride treatment) achieves similar effects by interfering with recognition of the carbohydrate by one or more clearance pathways. Nevertheless, because targeting of the protein relies, in most embodiments, on protein-receptor interactions rather than carbohydrate-receptor interactions, modification or elimination of glycosylation should not adversely affect targeting of the protein to the lysosome and/or across the blood brain barrier.

[0149] Any fusion protein of the invention, for exampe, any IGF-I tagged protein; or any lysosomal enzyme using a peptide targeting signal such as IGF-II can be chemically or enzymatically deglycosylated or modified to produce a therapeutic with the desirable properties of specific lysosomal targeting plus long serum half-life. In the case of some lysosomal storage diseases where it might be important to deliver the therapeutic to macrophage or related cell types via mannose receptor, fully glycosylated therapeutics can be used in combination with underglycosylate targeted therapeutics to achieve targeting to the broadest variety of cell types.

Proteins underglycosylated when synthesized

[0150] In some cases it will be preferable to produce the targeted therapeutic protein initially in a system that does not produce a fully glycosylated protein. For example, a targeted

therapeutic protein can be produced in *E. coli*, thereby generating a completely unglycosylated protein. Alternatively, an unglycosylated protein is produced in mammalian cells treated with tunicamycin, an inhibitor of Dol-PP-GlcNAc formation. If, however, a particular targeted therapeutic does not fold correctly in the absence of glycosylation, it is preferably produced initially as a glycosylated protein, and subsequently deglycosylated or rendered functionally underglycosylated.

[0151] Underglycosylated targeted therapeutic proteins can also by prepared by engineering a gene encoding the targeted therapeutic protein so that an amino acid that normally serves as an acceptor for glycosylation is changed to a different amino acid. For example, an asparagine residue that serves as an acceptor for N-linked glycosylation can be changed to a glutamine residue, or another residue that is not a glycosylation acceptor. This conservative change is most likely to have a minimal impact on enzyme structure while eliminating glycosylation at the site. Alternatively, other amino acids in the vicinity of the glycosylation acceptor can be modified, disrupting a recognition motif for glycosylation enzymes without necessarily changing the amino acid that would normally be glycosylated.

[0152] In the case of GUS, removal of any one of 4 potential glycosylation sites lessens the amount of glycosylation while retaining ample enzyme activity (Shipley *et al.* (1993) J. Biol. Chem. 268(16):12193-8). Removal of some sets of two glycosylation sites from GUS still permits significant enzyme activity. Removal of all four glycosylation sites eliminates enzyme activity, as does treatment of cells with tunicamycin, but deglycosylation of purified enzyme results in enzymatically active material. Therefore, loss of activity associated with removal of the glycosylation sites is likely due to incorrect folding of the enzyme.

[0153] Other enzymes, however, fold correctly even in the absence of glycosylation. For example, bacterial β-glucuronidase is naturally unglycosylated, and can be targeted to a

mammalian lysosome and/or across the blood brain barrier using the targeting moieties of the present invention. Such enzymes can be synthesized in an unglycosylated state, rather than, for example, synthesizing them as glycosylated proteins and subsequently deglycosylating them.

Deglycosylation

[0154] If the targeted therapeutic is produced in a mammalian cell culture system, it is preferably secreted into the growth medium, which can be harvested, permitting subsequent purification of the targeted therapeutic by, for example, chromatographic purification protocols, such as those involving ion exchange, gel filtration, hydrophobic chromatography, ConA chromatography, affinity chromatography or immunoaffinity chromatography.

[0155] Chemical deglycosylation of glycoproteins can be achieved in a number of ways, including treatment with trifluoromethane sulfonic acid (TFMS), or treatment with hydrogen fluoride (HF).

[0156] Chemical deglycosylation by TFMS (Sojar *et al.* (1989) J. Biol. Chem. 264(5):2552-9; Sojar *et al.* (1987) Methods Enzymol. 138:341-50): 1 mg GILT-GUS is dried under vacuum overnight. The dried protein is treated with 150 µl TFMS at 0°C for 0.5-2 hours under nitrogen with occasional shaking. The reaction mix is cooled to below -20°C in a dry ice-ethanol bath and the reaction is neutralized by the gradual addition of a prechilled (-20°C) solution of 60% pyridine in water. The neutralized reaction mix is then dialyzed at 4°C against several changes of NH_4HCO_3 at pH 7.0. Chemical deglycosylation with TFMS can result in modifications to the treated protein including methylation, succinimide formation and isomerization of aspartate residues (Douglass *et al.* (2001) J. Protein Chem. 20(7):571-6).

[0157] Chemical deglycosylation by HF (Sojar *et al.* (1987) Methods Enzymol. 138:341-50): The reaction is carried out in a closed reaction system such as can be obtained from Peninsula

Laboratories, Inc. 10 mg GILT-GUS is vacuum dried and placed in a reaction vessel which is then connected to the HF apparatus. After the entire HF line is evacuated, 10 mL anhydrous HF is distilled over from the reservoir with stirring of the reaction vessel. The reaction is continued for 1-2 hours at 0°C. Afterwards, a water aspirator removes the HF over 15-30 minutes. Remaining traces of HF are removed under high vacuum. The reaction mixture is dissolved in 2 mL 0.2M NaOH to neutralize any remaining HF and the pH is readjusted to 7.5 with cold 0.2M HCl.

[0158] Enzymatic deglycosylation (Thotakura et al. (1987) Methods Enzymol. 138:350-9): N-linked carbohydrates can be removed completely from glycoproteins using protein N-glycosidase (PNGase) A or F. In one embodiment, a glycoprotein is denatured prior to treatment with a glycosidase to facilitate action of the enzyme on the glycoprotein; the glycoprotein is subsequently refolded as discussed in the "In vitro refolding" section above. In another embodiment, excess glycosidase is used to treat a native glycoprotein to promote effective deglycosylation.

[0159] In the case of a targeted therapeutic protein that is actually underglycosylated, it is possible that the reduced glycosylation will reveal protease-sensitive sites on the targeted therapeutic protein, which will diminish the half-life of the protein. N-linked glycosylation is known to protect a subset of lysosomal enzymes from proteolysis (Kundra *et al.* (1999) J. Biol. Chem. 274(43):31039-46). Such protease-sensitive sites are preferably engineered out of the protein (*e.g.* by site-directed mutagenesis). As discussed below, the risk of revealing either a protease-sensitive site or a potential epitope can be minimized by incomplete deglycosylation or by modifying the carbohydrate structure rather than omitting the carbohydrate altogether.

Modification of carbohydrate structure or partial deglycosylation

[0160] In some embodiments, the therapeutic protein is partially deglycosylated. For example, the therapeutic protein can be treated with an endoglycosidase such as endoglycosidase H, which cleaves N-linked high mannose carbohydrate but not complex type carbohydrate leaving a single GlcNAc residue linked to the asparagine. A therapeutic protein treated in this way will lack high mannose carbohydrate, reducing interaction with the hepatic mannose receptor. Even though this receptor recognizes terminal GlcNAc, the probability of a productive interaction with the single GlcNAc on the protein surface is not as great as with an intact high mannose structure. If the therapeutic protein is produced in mammalian cells, any complex carbohydrate present on the protein will remains unaffected by the endoH treatment and may be terminally sialylated, thereby diminishing interactions with hepatic carbohydrate recognizing receptors. Such a protein is therefore likely to have increased half-life. At the same time, steric hinderance by the remaining carbohydrate should shield potential epitopes on the protein surface from the immune system and diminish access of proteases to the protein surface (*e.g.* in the protease-rich lysosomal environment).

[0161] In other embodiments, glycosylation of a therapeutic protein is modified, *e.g.* by oxidation, reduction, dehydration, substitution, esterification, alkylation, sialylation, carbon-carbon bond cleavage, or the like, to reduce clearance of the therapeutic protein from the blood. In some preferred embodiments, the therapeutic protein is not sialylated. For example, treatment with periodate and sodium borohydride is effective to modify the carbohydrate structure of most glycoproteins. Periodate treatment oxidizes vicinal diols, cleaving the carbon-carbon bond and replacing the hydroxyl groups with aldehyde groups; borohydride reduces the aldehydes to hydroxyls. Many sugar residues include vicinal diols and, therefore, are cleaved by this treatment. As shown in Figure 4A, a protein may be glycosylated on an asparagine residue with a high mannose carbohydrate that includes N-acetylglucosamine residues near the asparagine

and mannose residues elsewhere in the structure. As shown in Figure 4B, the terminal mannose residues have three consecutive carbons with hydroxyl groups; both of the carbon-carbon bonds involved are cleaved by periodate treatment. Some nonterminal mannose residues also include a vicinal diol, which would similarly be cleaved. Nevertheless, while this treatment converts cyclic carbohydrates into linear carbohydrates, it does not completely remove the carbohydrate, minimizing risks of exposing potentially protease-sensitive or antigenic polypeptide sites.

[0162] The half-life of lysosomal enzyme β-glucuronidase is known to increase more than tenfold after sequential treatment with periodate and sodium borohydride (Houba *et al.* (1996) Bioconjug. Chem. 7(5):606-11; Stahl *et al* (1976) PNAS 73:4045-4049; Achord *et al.* (1977) Pediat. Res. 11:816-822; Achord *et al.* (1978) Cell 15(1):269-78). Similarly, ricin has been treated with a mixture of periodate and sodium cyanoborohydride (Thorpe et al. (1985) Eur. J. Biochem. 147:197). After injection into rats, the fraction of ricin adsorbed by the liver decreased from 40% (untreated ricin) to 20% (modified ricin) of the injected dose with chemical treatment. In contrast the amount of ricin in the blood increased from 20% (untreated ricin) to 45% (treated ricin). Thus, the treated ricin enjoyed a wider tissue distribution and longer half-life in the circulation.

[0163] A β-glucuronidase construct (or other glycoprotein) coupled to a targeting moiety of the invention when deglycosylated or modified by sequential treatment with periodate and sodium borohydride should enjoy a similar (*e.g.* more than twofold, more than fourfold, or more than tenfold) increase in halflife while still retaining a high affinity for the cation-independent M6P receptor, permitting targeting of the construct to the lysosome of all cell types that possess this receptor. The construct is also predicted to cross the blood brain barrier efficiently. In contrast, if a β-glucuronidase preparation that relies on M6P for lysosomal targeting is deglycosylated or treated with periodate and sodium borohydride, it will enjoy an elevated serum

half-life but will be unable to target the lysosome since the M6P targeting signal will have been modified by the treatment.

[0164] Carbohydrate modification by sequential treatment with periodate and sodium borohydride can be performed as follows: Purified GILT-GUS is incubated with 40 mM $NaIO_4$ in 50 mM sodium acetate pH 4.5 for 2 hours at 4°C. The reaction is stopped by addition of excess ethylene glycol and unreacted reagents are removed by passing the reaction mix over Sephadex G-25M equilibrated with PBS pH 7.5. This treatment is followed by incubation with 40 mM $NaBH_4$ in PBS at pH 7.5 and 37°C for three hours and then for one hour at 4°C. Passing the reaction mixture over a Sephadex G-25M column eluted with PBS at pH 7.5 terminates the reaction.

[0165] Another protocol for periodate and sodium borohydride treatment is described in Hickman *et al.* (1974) BBRC 57:55-61. The purified protein is dialyzed into 0.01M sodium phosphate pH 6.0, 0.15 M NaCl. Sodium periodate is added to a final concentration of 0.01M and the reaction proceeds at 4°C in the dark for at least six hours. Treatment of β-hexosaminidase with periodate under these conditions is sufficient to prevent uptake of the protein by fibroblasts; uptake is normally dependent on M6P moieties on the β-hexosaminidase with the M6P receptor on the fibroblast cell surface. Thus, periodate oxidation modifies M6P sufficiently to abolish its ability to interact with the M6P receptor.

[0166] Alternatively, the carbohydrate can be modified by treatment with periodate and cyanoborohydride in a one step reaction as disclosed in Thorpe *et al.* (1985) Eur. J. Biochem. 147:197-206.

[0167] The presence of carbohydrate in a partially deglycosylated protein or a protein with a modified glycosylation pattern should shield potential polypeptide epitopes that might be uncovered by complete absence of glycosylation. In the event that a therapeutic protein does

provoke an immune response, immunosuppressive therapies can be used in conjunction with the therapeutic protein (Brooks (1999) Molecular Genetics and Metabolism 68:268-275). For example, it has been reported that about 15% of Gaucher disease patients treated with alglucerase developed immune responses (Beutler, *et al.*, in The Metabolic and Molecular Bases of Inherited Disease, 8th ed. (2001), Scriver *et al.*, eds., pp. 3635-3668). Fortunately, many (82/142) of the patients that produced antibody against alglucerase became tolerized by the normal treatment regimen (Rosenberg *et al.*, (1999) Blood 93:2081-2088). Thus, to benefit the small minority of patients who may develop an immune response, a patient receiving a therapeutic protein also receives an immunosuppressive therapy in some embodiments of the invention.

Testing

[0168] To verify that a protein is underglycosylated, it can be tested by exposure to ConA. An underglycosylated protein is expected to demonstrate reduced binding to ConA-sepharose when compared to the corresponding fully glycosylated protein.

[0169] An actually underglycosylated protein can also be resolved by SDS-PAGE and compared to the corresponding fully-glycosylated protein. For example, chemically deglycosylated GUS-GILT can be compared to untreated (glycosylated) GUS-GILT and to enzymatically deglycosylated GUS-GILT prepared with PNGase A. The underglycosylated protein is expected to have a greater mobility in SDS-PAGE when compared to the fully glycosylated protein.

[0170] Underglycosylated targeted therapeutic proteins display uptake that is dependent on the targeting domain. Underglycosylated proteins should display reduced uptake (and, preferably, substantially no uptake) that is dependent on mannose or M6P. These properties can be experimentally verified in cell uptake experiments.

[0171] For example, a GUS-GILT protein synthesized in mammalian cells and subsequently treated with periodate and borohydride can be tested for functional deglycosylation by testing M6P-dependent and mannose-dependent uptake. To demonstrate that M6P-dependent uptake has been reduced, uptake assays are performed using GM4668 fibroblasts. In the absence of competitor, treated and untreated enzyme will each display significant uptake. The presence of excess IGF-II substantially reduces uptake of treated and untreated enzyme, although untreated enzyme retains residual uptake via a M6P-dependent pathway. Excess M6P reduces the uptake of untreated enzyme, but is substantially less effective at reducing the uptake of functionally deglycosylated protein. For treated and untreated enzymes, the simultaneous presence of both competitors should substantially abolish uptake.

[0172] Uptake assays to assess mannose-dependent uptake are performed using J774-E cells, a mouse macrophage-like cell line bearing mannose receptors but few, if any, M6P receptors (Diment *et al.* (1987) J. Leukocyte Biol. 42:485-490). The cells are cultured in DMEM, low glucose, supplemented with 10% FBS, 4 mM glutamine, and antibiotic, antimycotic solution (Sigma, A-5955). Uptake assays with these cells are performed in a manner identical to assays performed with fibroblasts. In the presence of excess M6P and IGF-II, which will eliminate uptake due to any residual M6P/IGF-II receptor, fully glycosylated enzyme will display significant uptake due to interaction with the mannose receptor. Underglycosylated enzyme is expected to display substantially reduced uptake under these conditions. The mannose receptor-dependent uptake of fully glycosylated enzyme can be competed by the addition of excess (100 μg/mL) mannan.

[0173] Pharmacokinetics of deglycosylated GUS-GILT can be determined by giving intravenous injections of 20,000 enzyme units to groups of three MPSVII mice per timepoint. For each timepoint 50 μL of blood is assayed for enzyme activity.

EQUIVALENTS

[0174] The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The foregoing embodiments are therefore to be considered in all respects illustrative rather than limiting on the invention described herein. Scope of the invention is thus indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are intended to be embraced therein.

INCORPORATION BY REFERENCE

[0175] The disclosure of each of the patent documents and scientific publications disclosed herein, and U.S.S.N. 60/250,446 filed November 30, 2000; U.S.S.N. 60/250,444 filed November 30, 2000; U.S.S.N. 60/290,281 filed May 11, 2001; U.S.S.N. 60/287,531, filed April 30, 2001; U.S.S.N. 60/304,609, filed July 10, 2001; U.S.S.N. 60/329,461, filed October 15, 2001, a U.S.S.N. 60/351,276, filed January 23, 2002; U.S.SN. 10/136,841, filed on April 30, 2002; attorney docket number SYM-007CP entitled "Targeted Therapuetic Proteins," filed on October 16, 2002; atttorney docket number SYM-009 entitled "Targeted Therapeutic Proteins," filed on October 16, 2002; and PCT Serial No. PCT/US01/44935, filed November 30, 2001; are incorporated by reference into this application in their entirety.

CLAIMS

1. A method for targeting a polypeptide to the brain, the method comprising the step of providing a polypeptide in association with an IGF moiety to a mammal, thereby resulting in accumulation of the polypeptide in the brain of said mammal.

2. A method for producing a CNS-targeted polypeptide, the method comprising the step of expressing a polypeptide in association with an IGF moiety, thereby to produce a polypeptide with a CNS-targeting tag.

3. A composition comprising a polypeptide in association with an IGF moiety, wherein said IGF moiety targets the composition to the brain.

4. The composition of claim 3 wherein said polypeptide and IGF moiety are expressed as a fusion protein.

5. The composition of claim 3 or 4, wherein said IGF moiety is an IGF-I targeting moiety.

6. A therapeutic fusion protein comprising:
a therapeutic domain; and,
a targeting domain that enhances localization of the therapeutic domain to the CNS of a patient where the therapeutic domain is therapeutically active.

7. The therapeutic fusion protein of claim 6, wherein the targeting domain further binds to an extracellular domain of a receptor on an exterior surface of a cell in the CNS and, upon internalization of the receptor, permits localization of the therapeutic domain to a subcellular compartment where the therapeutic domain is therapeutically acitve.

8. The therapeutic fusion protein of claim 6, wherein said targeting domain comprises an IGF-I and an IGF-II targeting portion.

9. The therapeutic fusion protein of claim 6, wherein the subcellular compartment is selected from the group consisting of a lysosome, an endosome, endoplasmic reticulum, and golgi complex.

10. The therapeutic fusion protein of claim 9, wherein the subcellular compartment is a lysosome.

11. The therapeutic fusion protein of claim 7, wherein the receptor undergoes continuous endocytosis.

12. The therapeutic fusion protein of claim 6, wherein the therapeutic domain has a therapeutic enzymatic activity.

13. The therapeutic fusion protein of claim 12, wherein a cellular or subcellular deficiency in the enzymatic activity is associated with a human disease.

14. The therapeutic fusion protein of claim 13, wherein the human disease is a lysosomal storage disease.

15. A nucleic acid encoding the therapeutic fusion protein of claim 6.

16. A cell comprising the nucleic acid of claim 15.

17. A method of producing a therapeutic fusion protein, the method comprising the step of providing to the cell of claim 16 conditions permitting expression of the therapeutic fusion protein.

18 The method of claim 17, comprising culturing the cell *in vitro*.

19. The method of claim 17, comprising maintaining the cell inside a mammalian body.

20. The method of claim 17, further comprising harvesting the therapeutic fusion protein.

21. A method of treating a patient, the method comprising administering to the patient the therapeutic fusion protein of claim 6.

22. A method of treating a patient, the method comprising administering to the patient the nucleic acid of claim 15.

23. A method of treating a patient, the method comprising administering to the patient the cell of claim 16.

24. A method of treating a patient, the method comprising:

(i) synthesizing a targeted therapeutic comprising a therapeutic agent that is therapeutically active in mammalian lysosome and a targeting moiety that results in accumulation of therapeutic agent in mammalian CNS; and

(ii) administering the targeted therapeutic to the patient.

25. The method of claim 1 or 2, wherein said IGF moiety is an IGF-I targeting moiety.

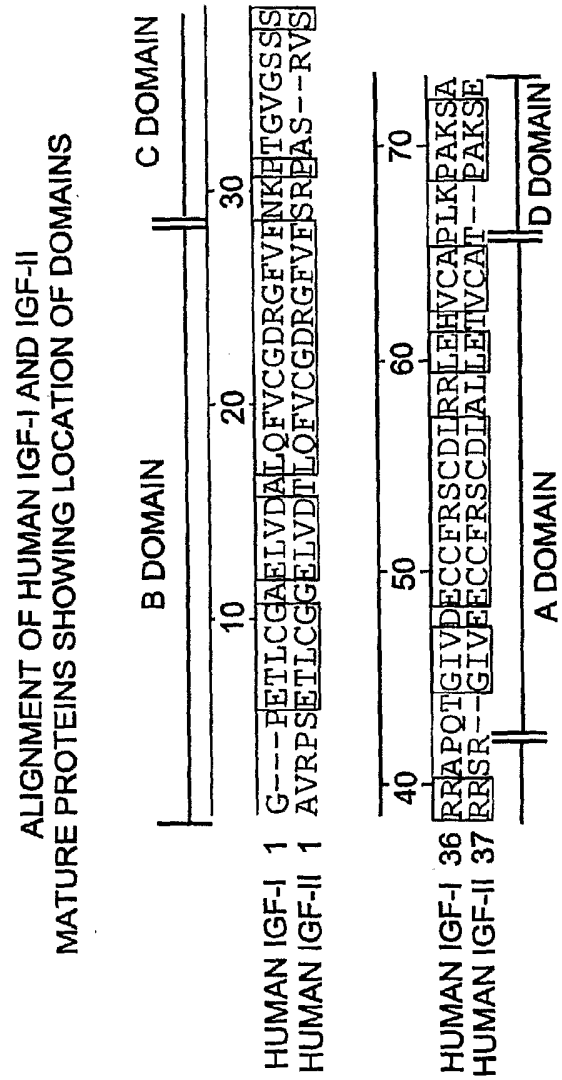
ALIGNMENT OF HUMAN IGF-I AND IGF-II
MATURE PROTEINS SHOWING LOCATION OF DOMAINS
B DOMAIN
C DOMAIN
A DOMAIN
D DOMAIN
HUMAN IGF-I 1 G---PETLCGAELVDALQFVCGDRGFYFNKPTGVGSS
HUMAN IGF-II 1 AVRPSETLCGGELVDTLQFVCGDRGFYFSRPASRVS
HUMAN IGF-I 36 RRAPQTGIVDECCFRSCDLRRLEMYCAPLKPAKSA
HUMAN IGF-II 37 RRSR--GIVEECCFRSCDLALLETYCAT--PAKSE

FIG. 1

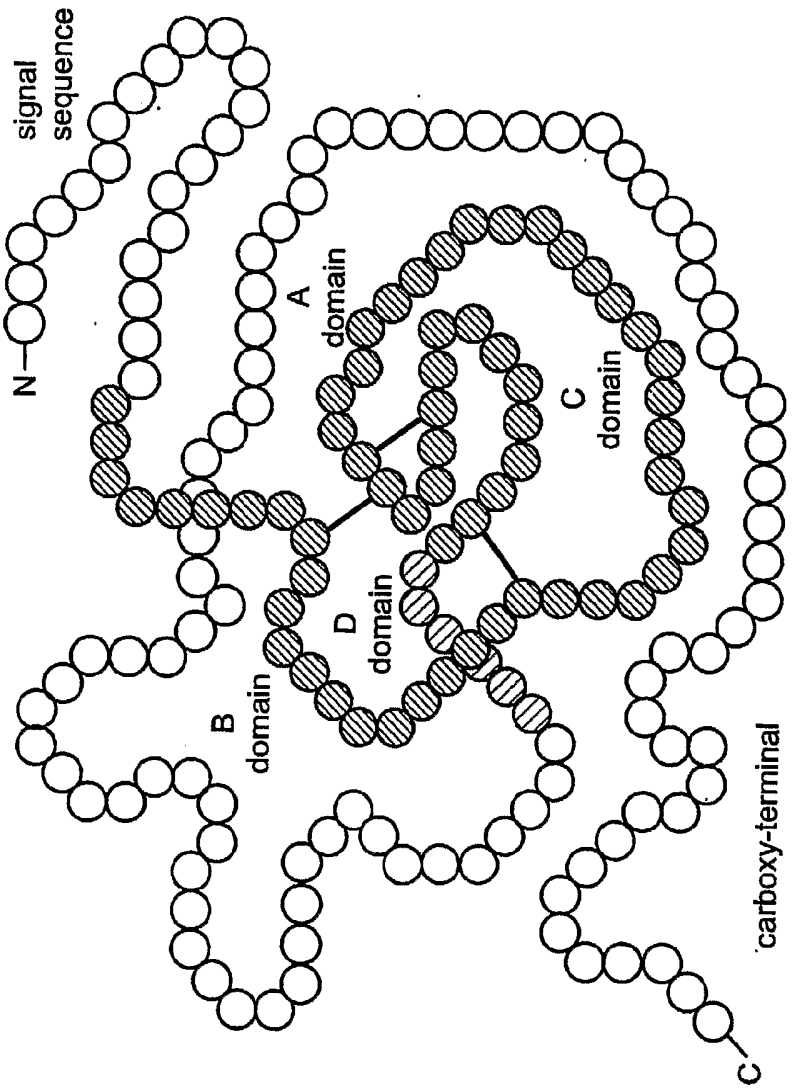
signal sequence
N
A domain
B domain
C domain
D domain
carboxy-terminal sequence (E peptide)
C

FIG. 2

MGKISSLPTQLFKCCFCDFLKVKMHTMSSSHLFYLALCLLTFTS
SATAGPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDL
RRLEMYCAPLKPAKSARSVRAQRHTDMPKTQKEVHLKNASRGSAGNKNYRM

FIG. 3A

FIG. 3B-1
FIG. 3B-2
FIG. 3B-3
FIG. 3B-4
FIG. 3B-5

FIG. 3B

1 tcactgtcac tgctaaattc agagcagatt agagcctgcg caatggaata aagtcctcaa
61 aattgaaatg tgacattgct ctcaacatct cccatctctc tggatttcct ttgcttcat
121 tattcctgct aaccaattca tttcagact ttgtacttca gaagcaatgg gaaaaatcag
181 cagtcttcca acccaattat ttaagtgctg cttttgtgat ttcttgaagg tgaagatgca
241 caccatgtcc tcctcgcatc tcttctacct ggcgctgtgc ctgctcacct tcaccagctc
301 tgccacggct ggaccggaga cgctctgcgg ggctgagctg gtggatgctc ttcagttcgt
361 gtgtggagac aggggctttt atttcaacaa gcccacaggg tatggctcca gcagtcggag
421 ggcgcctcag acaggcatcg tggatgagtg ctgcttccgg agctgtgatc taaggaggct
481 ggagatgtat tgcgcacccc tcaagcctgc caagtcagct cgctctgtcc gtgcccagcg
541 ccacaccgac atgcccaaga cccagaagga agtacatttg aagaacgcaa gtagagggag
601 tgcaggaaac aagaactaca ggatgtagga agaccctcct gaggagtgaa gagtgacatg
661 ccaccgcagg atcctttgct ctgcacgagt tacctgttaa actttggaac acctaccaaa
721 aaataagttt gataacattt aaaagatggg cgttcccccc aatgaaatac acaagtaaac
781 attccaacat tgtctttagg agtgatttgc accttgcaaa aatggtcctg gagttggtag
841 attgctgttg atcttttatc aataatgttc tatagaaaag aaaaaaaaat atatatatat
901 atatatctta gtccctgcct ctcaagagcc acaaatgcat gggtgttgta tagatccagt
961 tgcactaaat tcctctctga atcttggctg ctggagccat tcattcagca accttgtcta
1021 agtggtttat gaattgtttc cttatttgca cttctttcta cacaactcgg gctgtttgtt
1081 ttacagtgtc tgataatctt gttagtctat acccaccacc tcccttcata accttatat
1141 ttgccgaatt tggcctcctc aaaagcagca gcaagtcgtc aagaagcaca ccaattctaa
1201 cccacaagat tccatctgtg gcatttgtac caaatataag ttggatgcat tttatttag
1261 acacaaaggct ttatttttcc acatcatgct tacaaaaaag aataatgcaa atagttgcaa
1321 ctttgaggcc aatcatttt aggcatatgt tttaaacata gaaagtttct tcaacttcaaa
1381 agagttcctt caaatgatga gttaatgtgc aacctaatta gtaacttcc tctttttatt
1441 ttttccatat agagcactat gtaaatttag catatcaatt atacaggata tatcaaacag
1501 tatgtaaaac tctgttttt agtataatgg tgctattttg tagtttgtta tatgaaagag
1561 tctggccaaa acggtaatac gtgaaagcaa aacaataggg gaagcctgga gccaaagatg

FIG. 3B-1

1621 acacaagggg aagggtactg aaaacaccat ccatttggga aagaaggcaa agtccccccca
1681 gttatgcctt ccaagaggaa cttcagacac aaaagtccac tgatgcaaat tggactggcg
1741 agtccagaga ggaaactgtg gnatggaaaa agcagaaggc taggaatttt agcagtcctg
1801 gtttctttt ctcatggaag aaatgaacat ctgccagctg tgtcatggac tcaccactgt
1861 gtgaccttgg gcaagtcact tcacctctct gtgcctcagt ttcctcatct gcaaaatggg
1921 ggcaatatgt catctaccta cctcaaaggg gtggtataag gtttaaaaag ataaagattc
1981 agatttttt accctgggtt gctgtaaggg tgcaacatca gggcgcttga gttgctgaga
2041 tgcaaggaat tctataaata acccattcat agcatagcta gagattggtg aattgaatgc
2101 tcctgacatc tcagttcttg tcagtgaagc tatccaaata actggccaac tagttgttaa
2161 aagctaacag ctcaatctct taaaacactt ttcaaaatat gtgggaagca tttgatttc
2221 aatttgattt tgaattctgc atttggtttt atgaatacaa agataagtga aaagagagaa
2281 aggaaaagaa aaaggagaaa aacaaagaga ttttctaccag tgaaaggggga attaattact
2341 ctttgttagc actcactgac tcttctatgc agttactaca tatctagtaa aaccttgttt
2401 aatactataa ataatattct attcatttg aaaaacacaa tgattccttc tttctaggc
2461 aatataagga aagtgatcca aaatttgaaa tattaaaata atatctaata aaaagtcaca
2521 aagttatctt ctttaacaaa ctttactctt attcttagct gtatatacat ttttttaaaa
2581 agtttgttaa aatatgcttg actagagttt cagttgaaag gcaaaaaactt ccatcacaac
2641 aagaaatttc ccatgcctgc tcagaagggt agcccctagc tctctgtgaa tgtgtttat
2701 ccattcaact gaaaattggt atcaagaaag tccactggtt agtgtactag tccatcatag
2761 cctagaaaat gatccctatc tgcagatcaa gattttctca ttagaacaat gaattatcca
2821 gcattcagat ctttctagtc accttagaac tttttggtta aaagtaccca ggcttgatta
2881 ttcatgcaa attctatatt ttacattctt ggaaagtcta tatgaaaaac aaaaataaca
2941 tcttcagttt ttctcccact gggtcacctc aaggatcaga ggccaggaaa aaaaaaaaag
3001 actcctgga tctctgaata tatgcaaaaa gaaggcccca tttagtggag ccagcaatcc
3061 tgttcagtca acaagtattt taactctcag tccaacatta tttgaattga gcacctcaag
3121 catgcttagc aatgttctaa tcactatgga cagatgtaaa agaaactata catcattttt
3181 gccctctgcc tgttttccag acatacaggt tctgtggaat aagatactgg actcctcttc

FIG. 3B-2

3241 ccaagatggc acttctttt atttcttgtc cccagtgtgt accttttaaa attattccct
3301 ctcaacaaaa ctttataggc agtcttctgc agacttaaca tgttttctgt catagttaga
3361 tgtgataatt ctaagagtgt ctatgactta tttccttcac ttaattctat ccacagtcaa
3421 aaatcccccca aggaggaaag ctgaaagatg caactgccaa tattatcttt cttaactttt
3481 tccaacacat aatcctctcc aactggatta taaataaatt gaaaataact cattatacca
3541 attcactatt ttatttttta atgaattaaa actagaaaac aaattgatgc aaaccctgga
3601 agtcagttga ttactatata ctacagcaga atgactcaga tttcatagaa aggagcaacc
3661 aaaatgtcac aaccaaaact ttacaaggctt tgcttcagaa ttagattgct ttataattct
3721 tgaatgaggc aatttcaaga tatttgtaaa agaacagtaa acattggtaa gaatgagctt
3781 tcaactcata ggcttatttc caatttaatt gaccatactg gatacttagg tcaaatttct
3841 gttctctctt gcccaaataa tattaaagta ttatttgaac tttttaagat gaggcagttc
3901 ccctgaaaaa gttaatgcag ctctccatca gaatccactc ttctagggat atgaaaatct
3961 cttaacaccc accctacata cacagacaca cacacacaca cacacacaca cacacacaca
4021 cacacattca ccctaaggat ccaatggaat actgaaaaga aatcacttcc ttgaaaattt
4081 tattaaaaaa caaacaaaca aacaaaaagc ctgtccaccc ttgagaatcc ttcctctcct
4141 tggaacgtca atgtttgtgt agatgaaacc atctcatgct ctgtggctcc agggtttctg
4201 ttactatttt atgcacttgg gagaaggctt agaataaaag atgtagcaca ttttgctttc
4261 ccatttattg tttggccagc tatgccaatg tggtgctatt gtttctttaa gaaagtactt
4321 gactaaaaaa aaaagaaaaa aagaaaaaaa agaaagcata gacatatttt tttaaagtat
4381 aaaaacaaca attctataga tagatggctt aataaaataag cattaggtct atctagccac
4441 caccaccttt caacttttta tcactcacaa gtagtgtact gttcaccaaa ttgtgaattt
4501 gggggtgcag gggcaggagt tggaaatttt ttaaagttag aaggctccat tgttttgttg
4561 gctctcaaac ttagcaaaat tagcaatata ttatccaatc ttctgaactt gatcaagagc
4621 atggagaata aacgcgggaa aaaagatctt ataggcaaat agaagaattt aaaagataag
4681 taagttcctt attgattttt gtgcactctg ctctaaaaca gatattcagc aagtggagaa
4741 aataagaaca aagagaaaaa atacatagat ttacctgcaa aaaatagctt ctgccaaatc

FIG. 3B-3

4801 cccttgggt attctttggc atttactggt ttatagaaga cattctccct tcacccagac
4861 atctcaaaga gcagtagctc tcatgaaaag caatcactga tctcatttgg gaaatgttgg
4921 aaagtatttc cttatgagat gggggttatc tactgataaa gaaagaattt atgagaaatt
4981 gttgaaagag atggctaaca atctgtgaag atttttgtt tcttggtttt gttttttttt
5041 ttttttttac tttatacagt ctttatgaat ttcttaatgt tcaaaatgac ttggttcttt
5101 tcttctttt tttatatcag aatgaggaat aataagttaa acccacatag actctttaaa
5161 actataggct agatagaaat gtatgtttga cttgttgaag ctataatcag actatttaaa
5221 atgttttgct attttaatc ttaaaagatt gtgctaattt attagagcag aacctgtttg
5281 gctctccctca gaagaaagaa tctttccatt caaatcacat ggctttccac caatatttc
5341 aaaagataaa tctgatttat gcaatggcat catttatttt aaaacagaag aattgtgaaa
5401 gtttatgccc ctcccttgca aagaccataa agtccagatc tggtaggggg gcaacaacaa
5461 aaggaaaaatg ttgttgattc ttggttttgg atttttgttt gttttcaatg ctagtgttta
5521 atcctgtagt acatatttgc ttattgctat tttaatatatt tataagacct tcctgttagg
5581 tattagaaag tgatacatag atatctttttt tgtgtaattt ctatttaaaa aagagagaag
5641 actgtcagaa gctttaagtg catatggtac aggataaaga tatcaattta aataaccaat
5701 tcctatctgg aacaatgctt ttgttttta aagaaacctc tcacagataa gacagaggcc
5761 caggggattt ttgaagctgt ctttattctg cccccatccc aacccagccc ttattatttt
5821 agtatctgcc tcagaatttt atagagggct gaccaagctg aaactctaga attaaaggaa
5881 cctcactgaa aacatatatt tcacgtgttc cctctctttt ttttccttt tgtgagatgg
5941 ggtctcgcac tgtcccccag gctggagtgc agtggcatga tctcggctca ctgcaacctc
6001 cacctcctgg gtttaagcga ttctcctgcc tcagcctcct gagtagctgg gattacaggc
6061 acccaccact atgcccggct aatttttgg attttaata gagacggggt ttaccatgt
6121 tggccaggtt ggactcaaac tcctgacctt gtgatttgcc cgcctcagcc tcccaaattg
6181 ctgggattac aggcatgagc caccacaccc tgcccatgtg ttcctctta atgtatgatt
6241 acatggatct taaacatgat ccttctctcc tcattcttca actatctttg atgggtctt
6301 tcaaggggaa aaaaatccaa gctttttaa agtaaaaaaa aaaaaagaga ggacacaaaa

FIG. 3B-4

```
6361 ccaaatgtta ctgctcaact gaaatatgag ttaagatgga gacagagttt ctcctaataa
6421 ccggagctga attacctttc actttcaaaa acatgacctt ccacaatcct tagaatctgc
6481 ctttttttat attactgagg cctaaaagta aacattactc atttatttt gcccaaaatg
6541 cactgatgta aagtaggaaa aataaaaaca gagctctaaa atcccttca agccacccat
6601 tgacccccact caccaactca tagcaaagtc acttctgtta atcccttaat ctgattttgt
6661 ttggatattt atcttgtacc cgctgctaaa cacactgcag gagggactct gaaacctcaa
6721 gctgtctact tacatctttt atctgtgtct gtgtatcatg aaaatgtcta ttcaaaatat
6781 caaaaccttt caaatatcac gcagcttata ttcagtttac ataaaaggccc caaataccat
6841 gtcagatctt tttggtaaaa gagttaatga actatgagaa ttgggattac atcatgtatt
6901 ttgcctcatg tatttttatc acacttatag gccaagtgtg ataaaataaac ttacagacac
6961 tgaattaatt tcccctgcta ctttgaaacc agaaaaataat gactggccat tcgttacatc
7021 tgtcttagtt gaaaagcata ttttttatta aattaattct gattgtattt gaaattatta
7081 ttcaattcac ttatggcaga ggaatatcaa tcctaatgac ttctaaaaat gtaactaatt
7141 gaatcattat cttacattta ctgtttaata agcatatttt gaaaatgtat ggctagagtg
7201 tcataataaa atggtatatc tttctttagt aattacaaaa aaaaaaaaaa aaaaaaaaaa //
```

FIG. 3B-5

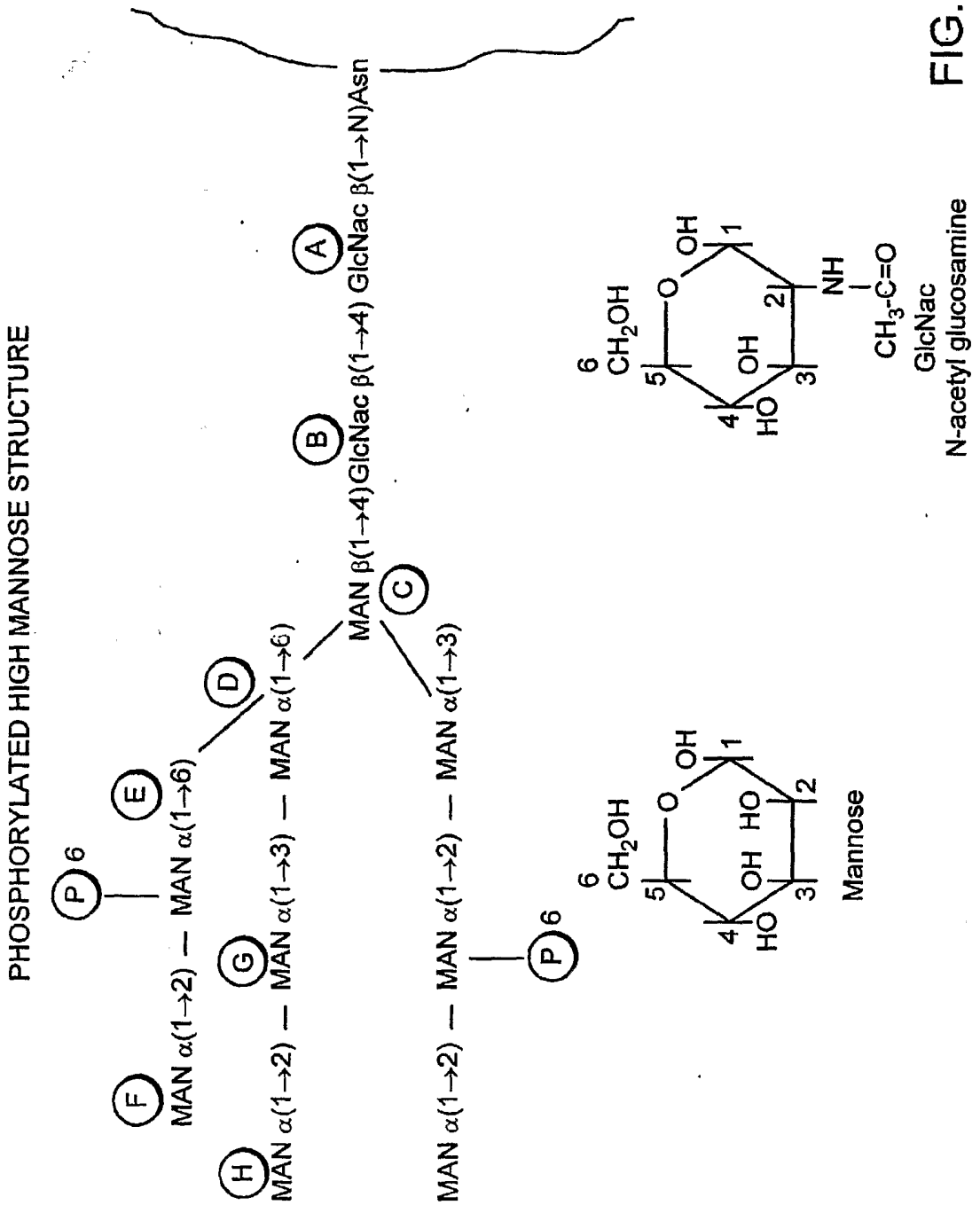
PHOSPHORYLATED HIGH MANNOSE STRUCTURE
P 6
E
MAN α(1→2) — MAN α(1→6)
D
F
MAN α(1→2) — MAN α(1→3) — MAN α(1→6)
G
H
MAN α(1→2) — MAN α(1→2) — MAN α(1→3)
C
MAN β(1→4)GlcNac β(1→4) GlcNac β(1→N)Asn
B
A
P 6
6 CH2OH
OH
NH
CH3-C=O
GlcNac
N-acetyl glucosamine
Mannose

FIG. 4A

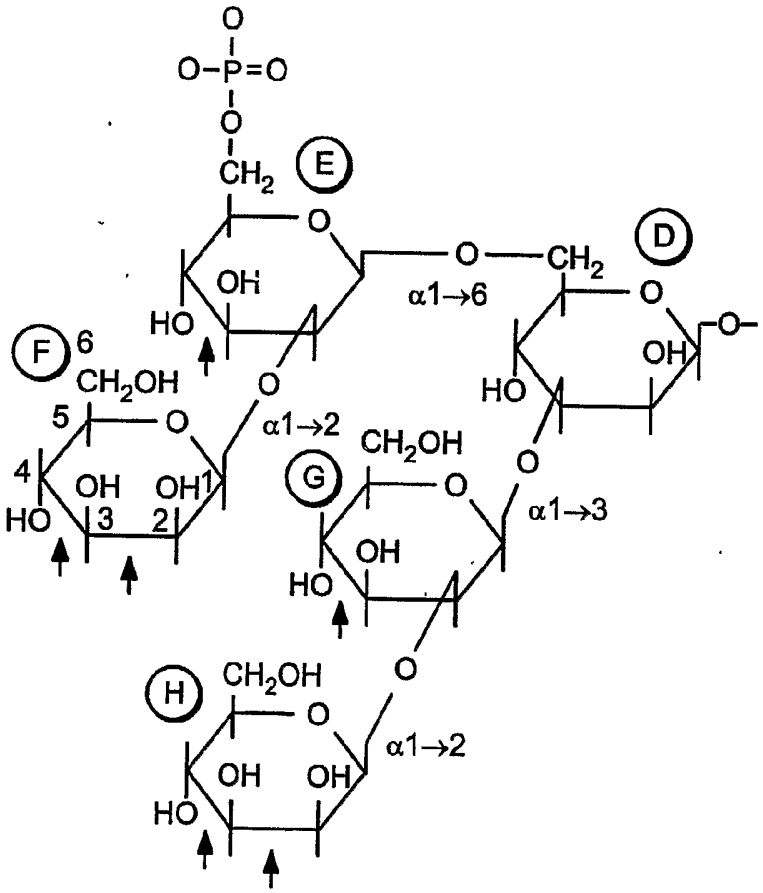
O
O–P=O
O
CH2
E
O
O
CH2
D
OH
α1→6
HO
O
F
6
CH2OH
5
O
α1→2
CH2OH
4
OH
OH
1
G
O
HO
3
2
α1→3
OH
HO
O
H
CH2OH
O
α1→2
OH
OH
HO

FIG. 4B

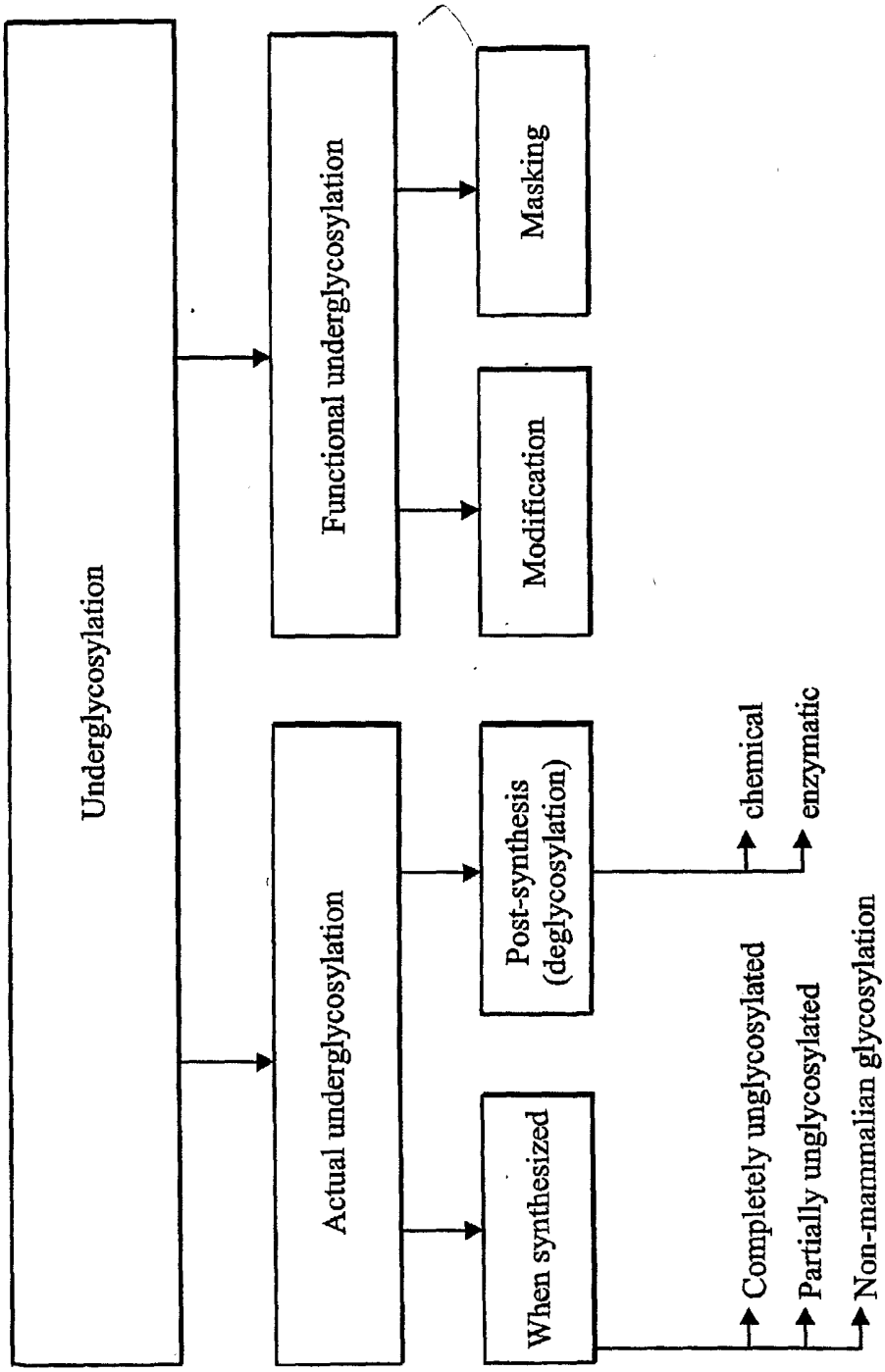
Underglycosylation
Functional underglycosylation
Masking
Modification
Actual underglycosylation
Post-synthesis (deglycosylation)
chemical
enzymatic
When synthesized
Completely unglycosylated
Partially unglycosylated
Non-mammalian glycosylation

FIG. 5

SEQUENCE LISTING

<110> LeBowitz, Jonathan

<120> METHODS AND COMPOSITIONS FOR TARGETING PROTEINS ACROSS THE BLOOD BRAIN BARRIER

<130> SYM-008PC1

<140> PCT/US 02/32996
<141> 2002-10-16

<150> US 60/329,650
<151> 2001-10-16

<150> US 10/136,841
<151> 2002-04-30

<150> US 10/136,639
<151> 2002-04-30

<150> US 60/384,452
<151> 2002-05-29

<150> US 60/386,019
<151> 2002-06-05

<150> US 60/408,816
<151> 2002-09-06

<160> 4

<170> PatentIn version 3.1

<210> 1
<211> 70
<212> PRT
<213> Homo sapiens

<400> 1

Gly Pro Glu Thr Leu Cys Gly Ala Glu Leu Val Asp Ala Leu Gln Phe
1 5 10 15

Val Cys Gly Asp Arg Gly Phe Tyr Phe Asn Lys Pro Thr Gly Tyr Gly
20 25 30

Ser Ser Ser Arg Arg Ala Pro Gln Thr Gly Ile Val Asp Glu Cys Cys
35 40 45

Phe Arg Ser Cys Asp Leu Arg Arg Leu Glu Met Tyr Cys Ala Pro Leu
50 55 60

Lys Pro Ala Lys Ser Ala
65 70

<210> 2
<211> 67
<212> PRT
<213> Homo sapiens

<400> 2

Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr
1 5 10 15

Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala
20 25 30

Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe
35 40 45

Arg Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala
50 55 60

Lys Ser Glu
65

<210> 3
<211> 153
<212> PRT
<213> Homo sapiens

<400> 3

Met Gly Lys Ile Ser Ser Leu Pro Thr Gln Leu Phe Lys Cys Cys Phe
1 5 10 15

Cys Asp Phe Leu Lys Val Lys Met His Thr Met Ser Ser Ser His Leu
20 25 30

Phe Tyr Leu Ala Leu Cys Leu Leu Thr Phe Thr Ser Ser Ala Thr Ala
35 40 45

Gly Pro Glu Thr Leu Cys Gly Ala Glu Leu Val Asp Ala Leu Gln Phe
50 55 60

Val Cys Gly Asp Arg Gly Phe Tyr Phe Asn Lys Pro Thr Gly Tyr Gly
65 70 75 80

Ser Ser Ser Arg Arg Ala Pro Gln Thr Gly Ile Val Asp Glu Cys Cys
85 90 95

Phe Arg Ser Cys Asp Leu Arg Arg Leu Glu Met Tyr Cys Ala Pro Leu
100 105 110

Lys Pro Ala Lys Ser Ala Arg Ser Val Arg Ala Gln Arg His Thr Asp
115 120 125

Met Pro Lys Thr Gln Lys Glu Val His Leu Lys Asn Ala Ser Arg Gly
130 135 140

Ser Ala Gly Asn Lys Asn Tyr Arg Met
145 150

<210> 4
<211> 7260
<212> DNA
<213> Homo sapiens

<400> 4
tcactgtcac tgctaaattc agagcagatt agagcctgcg caatggaata aagtcctcaa 60

aattgaaatg tgacattgct ctcaacatct cccatctctc tggatttcct tttgcttcat 120

tattcctgct aaccaattca ttttcagact ttgtacttca gaagcaatgg gaaaaatcag 180

cagtcttcca acccaattat ttaagtgctg cttttgtgat ttcttgaagg tgaagatgca 240

caccatgtcc tcctcgcatc tcttctacct ggcgctgtgc ctgctcacct tcaccagctc 300

tgccacggct ggaccggaga cgctctgcgg ggctgagctg gtggatgctc ttcagttcgt 360

gtgtggagac aggggctttt atttcaacaa gcccacaggg tatggctcca gcagtcggag 420

ggcgcctcag acaggcatcg tggatgagtg ctgcttccgg agctgtgatc taaggaggct 480

ggagatgtat tgcgcacccc tcaagcctgc caagtcagct cgctctgtcc gtgcccagcg 540

ccacaccgac atgcccaaga cccagaagga agtacatttg aagaacgcaa gtagagggag 600

tgcaggaaac aagaactaca ggatgtagga agaccctcct gaggagtgaa gagtgacatg 660

ccaccgcagg atcctttgct ctgcacgagt tacctgttaa actttggaac acctaccaaa 720

aaataagttt gataacattt aaaagatggg cgtttccccc aatgaaatac acaagtaaac 780

attccaacat tgtctttagg agtgatttgc accttgcaaa aatggtcctg gagttggtag 840

attgctgttg atcttttatc aataatgttc tatagaaaag aaaaaaaaat atatatatat 900

atatatctta gtccctgcct ctcaagagcc acaaatgcat gggtgttgta tagatccagt 960

tgcactaaat tcctctctga atcttggctg ctggagccat tcattcagca accttgtcta 1020

agtggtttat gaattgtttc cttatttgca cttctttcta cacaactcgg gctgtttgtt 1080

ttacagtgtc tgataatctt gttagtctat acccaccacc tcccttcata acctttatat 1140

ttgccgaatt tggcctcctc aaaagcagca gcaagtcgtc aagaagcaca ccaattctaa 1200

cccacaagat tccatctgtg gcatttgtac caaatataag ttggatgcat tttattttag 1260

acacaaagct ttatttttcc acatcatgct [illegible] 1320

ctttgaggcc aatcattttt aggcatatgt tttaaacata gaaagtttct tcaactcaaa 1380

agagttcctt caaatgatga gttaatgtgc aacctaatta gtaactttcc tcttttttatt 1440

ttttccatat agagcactat gtaaatttag catatcaatt atacaggata tatcaaacag 1500

tatgtaaaac tctgtttttt agtataatgg tgctattttg tagtttgtta tatgaaagag 1560

tctggccaaa acggtaatac gtgaaagcaa aacaataggg gaagcctgga gccaaagatg 1620

acacaagggg aagggtactg aaaacaccat ccatttggga aagaaggcaa agtcccccca 1680

gttatgcctt ccaagaggaa cttcagacac aaaagtccac tgatgcaaat tggactggcg 1740

agtccagaga ggaaactgtg gaatggaaaa agcagaaggc taggaatttt agcagtcctg 1800

gtttcttttt ctcatggaag aaatgaacat ctgccagctg tgtcatggac tcaccactgt 1860

gtgaccttgg gcaagtcact. tcacctctct gtgcctcagt ttcctcatct gcaaaatggg 1920

ggcaatatgt catctaccta cctcaaaggg gtggtataag gtttaaaaag ataaagattc 1980

agatttttt accctgggtt gctgtaaggg tgcaacatca gggcgcttga gttgctgaga 2040

tgcaaggaat tctataaata acccattcat agcatagcta gagattggtg aattgaatgc 2100

tcctgacatc tcagttcttg tcagtgaagc tatccaaata actggccaac tagttgttaa 2160

aagctaacag ctcaatctct taaaacactt ttcaaaatat gtgggaagca tttgattttc 2220

aatttgattt tgaattctgc atttggtttt atgaatacaa agataagtga aaagagagaa 2280

aggaaaagaa aaaggagaaa aacaaagaga tttctaccag tgaaagggga attaattact 2340

ctttgttagc actcactgac tcttctatgc agttactaca tatctagtaa aaccttgttt 2400

aatactataa ataatattct attcattttg aaaaacacaa tgattccttc ttttctaggc 2460

aatataagga aagtgatcca aaatttgaaa tattaaaata atatctaata aaaagtcaca 2520

aagttatctt ctttaacaaa ctttactctt attcttagct gtatatacat ttttttaaaa 2580

agtttgttaa aatatgcttg actagagttt cagttgaaag gcaaaaactt ccatcacaac 2640

aagaaatttc ccatgcctgc tcagaagggt agcccctagc tctctgtgaa tgtgttttat 2700

ccattcaact gaaaattggt atcaagaaag tccactggtt agtgtactag tccatcatag 2760

cctagaaaat gatccctatc tgcagatcaa gattttctca ttagaacaat gaattatcca 2820

gcattcagat ctttctagtc accttagaac tttttggtta aaagtaccca ggcttgatta 2880

tttcatgcaa attctatatt ttacattctt ggaaagtcta tatgaaaaac aaaaataaca 2940

tcttcagttt ttctcccact gggtcacctc aaggatcaga ggccaggaaa aaaaaaaaag 3000

actccctgga tctctgaata tatgcaaaaa gaaggcccca tttagtggag ccagcaatcc 3060

tgttcagtca acaagtattt taactctcag tccaacatta tttgaattga gcacctcaag 3120

catgcttagc aatgttctaa tcactatgga cagatgtaaa agaaactata catcattttt 3180

gccctctgcc tgttttccag acatacaggt tctgtggaat aagatactgg actcctcttc 3240

ccaagatggc acttcttttt atttcttgtc cccagtgtgt acctttttaaa attattccct 3300

ctcaacaaaa ctttataggc agtcttctgc agacttaaca tgttttctgt catagttaga 3360

tgtgataatt ctaagagtgt ctatgactta tttccttcac ttaattctat ccacagtcaa 3420

aaatccccca aggaggaaag ctgaaagatg caactgccaa tattatcttt cttaactttt 3480

tccaacacat aatcctctcc aactggatta taaataaatt gaaaataact cattatacca 3540

attcactatt ttatttttta atgaattaaa actagaaaac aaattgatgc aaaccctgga 3600

agtcagttga ttactatata ctacagcaga atgactcaga tttcatagaa aggagcaacc 3660

aaaatgtcac aaccaaaact ttacaagctt tgcttcagaa ttagattgct ttataattct 3720

tgaatgaggc aatttcaaga tatttgtaaa agaacagtaa acattggtaa gaatgagctt 3780

tcaactcata ggcttatttc caatttaatt gaccatactg gatacttagg tcaaatttct 3840

gttctctctt gcccaaataa tattaaagta ttatttgaac tttttaagat gaggcagttc 3900

ccctgaaaaa gttaatgcag ctctccatca gaatccactc ttctagggat atgaaaatct 3960

cttaacaccc accctacata cacagacaca cacacacaca cacacacaca cacacacaca 4020

cacacattca ccctaaggat ccaatggaat actgaaaaga aatcacttcc ttgaaaattt 4080

tattaaaaaa caaacaaaca aacaaaaagc ctgtccaccc ttgagaatcc ttcctctcct 4140

tggaacgtca atgtttgtgt agatgaaacc atctcatgct ctgtggctcc agggtttctg 4200

ttactatttt atgcacttgg gagaaggctt agaataaaag atgtagcaca ttttgctttc 4260

ccatttattg tttggccagc tatgccaatg tggtgctatt gtttctttaa gaaagtactt 4320

gactaaaaaa aaaagaaaaa aagaaaaaaa agaaagcata gacatatttt tttaaagtat 4380

aaaaacaaca attctataga tagatggctt aataaaatag cattaggtct atctagccac 4440

caccaccttt caacttttta tcactcacaa gtagtgtact gttcaccaaa ttgtgaattt 4500

gggggtgcag gggcaggagt tggaaatttt ttaaagttag aaggctccat tgttttgttg 4560

gctctcaaac ttagcaaaat tagcaatata ttatccaatc ttctgaactt gatcaagagc 4620

atggagaata aacgcgggaa aaaagatctt ataggcaaat agaagaattt aaaagataag 4680

taagttcctt attgattttt gtgcactctg ctctaaaaca gatattcagc aagtggagaa 4740

aataagaaca aagagaaaaa atacatagat ttacctgcaa aaaatagctt ctgccaaatc 4800

ccccttgggt attctttggc atttactggt ttatagaaga cattctccct tcacccagac 4860

atctcaaaga gcagtagctc tcatgaaaag caatcactga tctcatttgg gaaatgttgg 4920

aaagtatttc cttatgagat gggggttatc tactgataaa gaaagaattt atgagaaatt 4980

gttgaaagag atggctaaca atctgtgaag atttttttgtt tcttggtttt gtttttttt 5040

tttttttac tttatacagt ctttatgaat ttcttaatgt tcaaaatgac ttggttcttt 5100

tcttcttttt tttatatcag aatgaggaat aataagttaa acccacatag actctttaaa 5160

actataggct agatagaaat gtatgtttga cttgttgaag ctataatcag actatttaaa 5220

atgttttgct atttttaatc ttaaaagatt gtgctaattt attagagcag aacctgtttg 5280

gctctcctca gaagaaagaa tctttccatt caaatcacat ggctttccac caatattttc 5340

aaaagataaa tctgatttat gcaatggcat catttatttt aaaacagaag aattgtgaaa 5400

gtttatgccc ctcccttgca aagaccataa agtccagatc tggtaggggg gcaacaacaa 5460

aaggaaaatg ttgttgattc ttggttttgg attttgtttt gttttcaatg ctagtgttta 5520

atcctgtagt acatatttgc ttattgctat tttaatattt tataagacct tcctgttagg 5580

tattagaaag tgatacatag atatcttttt tgtgtaattt ctatttaaaa aagagagaag 5640

actgtcagaa gctttaagtg catatggtac aggataaaga tatcaattta aataaccaat 5700

tcctatctgg aacaatgctt ttgtttttta aagaaacctc tcacagataa gacagaggcc 5760

caggggattt ttgaagctgt ctttattctg cccccatccc aacccagccc ttattatttt 5820

agtatctgcc tcagaatttt atagagggct gaccaagctg aaactctaga attaaaggaa 5880

cctcactgaa aacatatatt tcacgtgttc cctctctttt ttttcctttt tgtgagatgg 5940

ggtctcgcac tgtcccccag gctggagtgc agtggcatga tctcggctca ctgcaacctc 6000

cacctcctgg gtttaagcga ttctcctgcc tcagcctcct gagtagctgg gattacaggc 6060

acccaccact atgcccggct aatttttttgg atttttaata gagacggggt tttaccatgt 6120

tggccaggtt ggactcaaac tcctgacctt gtgatttgcc cgcctcagcc tcccaaattg 6180

ctgggattac aggcatgagc caccacaccc tgcccatgtg ttccctctta atgtatgatt 6240

acatggatct taaacatgat ccttctctcc tcattcttca actatctttg atggggtctt 6300

tcaaggggaa aaaaatccaa gctttttttaa agtaaaaaaa aaaaaagaga ggacacaaaa 6360

ccaaatgtta ctgctcaact gaaatatgag ttaagatgga gacagagttt ctcctaataa 6420

ccggagctga attacctttc actttcaaaa acatgacctt ccacaatcct tagaatctgc 6480

ctttttttat attactgagg cctaaaagta aacattactc attttatttt gcccaaaatg 6540

cactgatgta aagtaggaaa aataaaaaca gagctctaaa atccctttca agccacccat 6600

tgaccccact caccaactca tagcaaagtc acttctgtta atcccttaat ctgattttgt 6660

ttggatattt atcttgtacc cgctgctaaa cacactgcag gagggactct gaaacctcaa 6720

gctgtctact tacatctttt atctgtgtct gtgtatcatg aaaatgtcta ttcaaaatat 6780

caaaaccttt caaatatcac gcagcttata ttcagtttac ataaaggccc caaataccat 6840

gtcagatctt tttggtaaaa gagttaatga actatgagaa ttgggattac atcatgtatt 6900

ttgcctcatg tatttttatc acacttatag gccaagtgtg ataaataaac ttacagacac 6960

tgaattaatt tcccctgcta ctttgaaacc agaaaataat gactggccat tcgttacatc 7020

tgtcttagtt gaaaagcata ttttttatta aattaattct gattgtattt gaaattatta 7080

ttcaattcac ttatggcaga ggaatatcaa tcctaatgac ttctaaaaat gtaactaatt 7140

gaatcattat cttacattta ctgtttaata agcatatttt gaaaatgtat ggctagagtg 7200

tcataataaa atggtatatc tttctttagt aattacaaaa aaaaaaaaaa aaaaaaaaaa 7260