TARGETED HEPARAN SULFATASE COMPOUNDS

Heparan sulfatase (SGSH) constructs

With a His₉ tag at the C-terminus

1. SGSH-His₉

   Signal peptide [SGSH] His₉

2. Acm2-SGSH-His₉

   Signal peptide [Acm2][SGSH] His₉

3. SGSH-Acm2-His₉

   Signal peptide [SGSH][Acm2] His₉

4. Acm2-SGSH-Acm2-His₉

   Signal peptide [Acm2][SGSH][Acm2] His₉

Figure 1

Title: TARGETED HEPARAN SULFATASE COMPOUNDS

Abstract: The present invention is related to a compound that includes a lysosomal enzyme (e.g., heparan sulfatase) and a targeting moiety, for example, where compound is a fusion protein including heparan sulfatase and Angiopep-2. In certain embodiments, these compounds, owning to the presence of the targeting moiety can cross the blood-brain barrier or accumulate in the lysosome more effectively than the enzyme alone. The invention also features methods for treating lysosomal storage disorders (e.g., mucopolysaccharidosis Type IIa) using such compounds.
Declarations under Rule 4.17:

— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(H))

Published:

— with international search report (Art. 21(3))
TARGETED HEPARAN SULFATASE COMPOUNDS

Cross-Reference to Related Applications

This application claims benefit of U.S. Provisional Application Nos. 61/831,892, filed June 6, 2013, which is hereby incorporated by reference in its entirety.

Background of the Invention

The invention relates to compounds including a lysosomal enzyme (e.g., heparan N-sulfatase; SGSH) and a targeting moiety and the use of such conjugates in the treatment of disorders that result from a deficiency of such enzymes (e.g., Sanfilippo syndrome).

Lysosomal storage disorders are group of about 50 rare genetic disorders in which a subject has a defect in a lysosomal enzyme that is required for proper metabolism. These diseases typically result from autosomal or X-linked recessive genes. As a group, the incidence of these disorders is about 1:5000 to 1:10,000.

Sanfilippo syndrome or mucopolysaccharidosis III (MPS-III) results from a deficiency of enzymes required for the lysosomal degradation of heparan sulfate. The four types of MPS-III, Type A-D, are due to specific enzyme deficiencies. MPS-IIIa results from a deficiency of SGSH. MPS-IIIa is the most common of the four types, with an approximate incidence of 1 in 100,000 live births, accounting for approximately 60% of MPS-III cases. Those with the disorder are unable to break down and recycle these mucopolysaccharides, which are also known as glycosaminoglycans or GAG. This deficiency results in the buildup of GAG throughout the body, which has serious effects on the nervous system, joints, and various organ systems including heart, liver, and skin. There are also a number of physical symptoms, including coarse facial features, enlarged head and abdomen, and skin lesions. The life span of an affected child does not usually extend beyond the late teens to early twenties.

There is no cure for MPS-III. In addition to palliative measures, therapeutic approaches have included bone marrow grafts and enzyme replacement therapy.

Enzyme replacement therapy by intravenous administration of heparan sulfatase has been shown to have benefits; however, this approach does not improve the central nervous system deficits associated with MPS-III because the enzymes are not expected to cross the blood-brain barrier.
Methods for increasing delivery of the deficient enzymes (e.g., SGSH) to the brain have been and are being investigated, including intrathecal delivery. Intrathecal delivery, however, is a highly invasive technique.

Less invasive and more effective methods of treating MPS-III, that address the neurological disease symptoms, in addition to the other symptoms, would therefore be highly desirable.

**Summary of the Invention**

The present invention is directed to compounds that include a targeting moiety and a lysosomal enzyme (e.g., SGSH). These compounds are exemplified by SGSH-Angiopep-2 conjugates and fusion proteins which can be used to treat MPS-IIIa. Because these conjugates and fusion proteins are capable of crossing the BBB, they can treat not only the peripheral disease symptoms, but can also be effective in treating CNS symptoms. In addition, because targeting moieties such as Angiopep-2 are capable of targeting enzymes to the lysosomes, it is expected that these conjugates and fusion proteins are more effective than the enzymes by themselves.

Accordingly, in a first aspect, the invention features a compound including (a) a targeting moiety (e.g., a peptide or peptidic targeting moiety that may be less than 200, 150, 125, 100, 80, 60, 50, 40, 35, 30, 25, 24, 23, 22, 21, 20, or 19 amino acids) and (b) SGSH, an active fragment thereof, or an analog thereof, where the targeting moiety and the enzyme, fragment, or analog are joined by a linker. The lysosomal enzyme may be heparan sulfatase (SGSH), an SGSH fragment having SGSH activity, or an SGSH analog.

In the first aspect, the targeting moiety may include an amino acid sequence that is substantially identical to any of SEQ ID NOS:1-105 and 107-1 17 (e.g., Angiopep-2 (SEQ ID NO:97)). In other embodiments, the targeting moiety includes the formula Lys-Arg-X3-X4-X5-Lys (formula la), where X3 is Asn or Gin; X4 is Asn or Gin; and X5 is Phe, Tyr, or Trp, where the targeting moiety optionally includes one or more D-isomers of an amino acid recited in formula la. In other embodiments, the targeting moiety includes the formula Z1-Lys-Arg-X3-X4-X5-Lys-Z2 (formula lb), where X3 is Asn or Gin; X4 is Asn or Gin; X5 is Phe, Tyr, or Trp; Z1 is absent, Cys, Gly, Cys-Gly, Arg-Gly, Cys-Arg-Gly, Ser-Arg-Gly, Cys-Ser-Arg-Gly, Gly-Ser-Arg-Gly, Cys-Gly-Ser-Arg-Gly, Gly-Gly-Ser-Arg-Gly, Cys-Gly-Gly-Ser-Arg-Gly, Tyr-Gly-Gly-Ser-Arg-Gly, Cys-Tyr-Gly-Gly-
Ser-Arg-Gly, Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Cys-Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Phe-
Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Cys-Phe-Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Thr-Phe-Phe-
Tyr-Gly-Gly-Ser-Arg-Gly, or Cys-Thr-Phe-Phe-Tyr-Gly-Gly-Ser-Arg-Gly; and Z2 is
absent, Cys, Tyr, Tyr-Cys, Cys-Tyr, Thr-Glu-Glu-Tyr, or Thr-Glu-Glu-Tyr-Cys; and

where the targeting moiety optionally includes one or more D-isomers of an amino acid
recited in formula lb, Zl, or Z2. In other embodiments, the targeting moiety includes the
formula XI-X2-Asn-Asn-X5-X6 (formula Ha), where XI is Lys or D-Lys; X2 is Arg or
D-Arg; X5 is Phe or D-Phe; and X6 is Lys or D-Lys; and where at least one of XI, X2,
X5, or X6 is a D-amino acid. In other embodiments, the targeting moiety includes the
formula XI-X2-Asn-Asn-X5-X6-X7 (formula lib), where XI is Lys or D-Lys; X2 is Arg
or D-Arg; X5 is Phe or D-Phe; X6 is Lys or D-Lys; and X7 is Tyr or D-Tyr; and where at
least one of XI, X2, X5, X6, or X7 is a D-amino acid. In other embodiments, the
targeting moiety includes the formula Zl-Xl-X2-Asn-Asn-X5-X6-X7-Z2 (formula lie),
where XI is Lys or D-Lys; X2 is Arg or D-Arg; X5 is Phe or D-Phe; X6 is Lys or D-Lys;
X7 is Tyr or D-Tyr; Zl is absent, Cys, Gly, Cys-Gly, Arg-Gly, Cys-Arg-Gly, Ser-Arg-
Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Cys-Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Phe-Phe-Tyr-Gly-
Gly-Ser-Arg-Gly, or Cys-Thr-Phe-Phe-Tyr-Gly-Gly-Ser-Arg-Gly; and Z2 is absent, Cys, Tyr,
Tyr-Cys, Cys-Tyr, Thr-Glu-Glu-Tyr, or Thr-Glu-Glu-Tyr-Cys; where at least one of XI,
X2, X5, X6, or X7 is a D-amino acid; and where the polypeptide optionally includes one
or more D-isomers of an amino acid recited in Zl or Z2.

In the first aspect, the linker may be a covalent bond (e.g., a peptide bond) or one
or more amino acids. The compound may be a fusion protein (e.g., Angiopep-2-SGSH,
SGSH-Angiopep-2, or Angiopep-2-SGSH-Angiopep-2, or has the structure shown in
Figure 1). The compound may further include a second targeting moiety that is joined to
the compound by a second linker.

The invention also features a pharmaceutical composition including a compound
of the first aspect and a pharmaceutically acceptable carrier.

In another aspect, the invention features a method of treating or treating
prophylactically a subject having a lysosomal storage disorder (e.g., MPS-IIIa). The
method includes administering to the subject a compound of the first aspect or a pharmaceutical composition described herein. The lysosomal enzyme in the compound may be SGSH. The subject may be experiencing neurological symptoms (e.g., mental retardation). The method may be performed on or started on a subject that is less than six months, or 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, or 18 years of age. The subject may be an infant (e.g., less than 1 year old).

In certain embodiments, the targeting moiety is not an antibody (e.g., an antibody or an immunoglobulin that is specific for an endogenous BBB receptor such as the insulin receptor, the transferrin receptor, the leptin receptor, the lipoprotein receptor, and the IGF receptor).

In any of the above aspects, the targeting moiety may be substantially identical to any of the sequences of Table 1, or a fragment thereof. In certain embodiments, the peptide vector has a sequence of Angiopep-1 (SEQ ID NO:67), Angiopep-2 (SEQ ID NO:97) (An2), Angiopep-3 (SEQ ID NO: 107), Angiopep-4a (SEQ ID NO:108), Angiopep-4b (SEQ ID NO: 109), Angiopep-5 (SEQ ID NO: 110), Angiopep-6 (SEQ ID NO:111), Angiopep-7 (SEQ ID NO:112)) or reversed Angiopep-2 (SEQ ID NO: 117). The targeting moiety or compound may be efficiently transported into a particular cell type (e.g., any one, two, three, four, or five of liver, lung, kidney, spleen, and muscle) or may cross the mammalian BBB efficiently (e.g., Angiopep-1, -2, -3, -4a, -4b, -5, and -6).

In another embodiment, the targeting moiety or compound is able to enter a particular cell type (e.g., any one, two, three, four, or five of liver, lung, kidney, spleen, and muscle) but does not cross the BBB efficiently (e.g., a conjugate including Angiopep-7). The targeting moiety may be of any length, for example, at least 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 25, 35, 50, 75, 100, 200, or 500 amino acids, or any range between these numbers. In certain embodiments, the targeting moiety is less than 200, 150, 125, 100, 90, 80, 70, 60, 50, 40, 30, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, or 6 amino acids (e.g., 10 to 50 amino acids in length). The targeting moiety may be produced by recombinant genetic technology or chemical synthesis.
### Table 1: Exemplary targeting moieties

<table>
<thead>
<tr>
<th>SEQ ID</th>
<th>NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T F V Y G G C R A K R N N F K S A E D</td>
</tr>
<tr>
<td>2</td>
<td>T F Q Y G G C M G N G N N F V T E K E</td>
</tr>
<tr>
<td>3</td>
<td>P F F Y G G C G G N R N N F D T E E Y</td>
</tr>
<tr>
<td>4</td>
<td>S F Y Y G G C L G N K N N Y L R E E E</td>
</tr>
<tr>
<td>5</td>
<td>T F F Y G G C R A K R N N F K R A K Y</td>
</tr>
<tr>
<td>6</td>
<td>T F F Y G G C R G K R N N F K R A K Y</td>
</tr>
<tr>
<td>7</td>
<td>T F F Y G G C R A K K N N Y K R A K Y</td>
</tr>
<tr>
<td>8</td>
<td>T F F Y G G C R G K K N N F K R A K Y</td>
</tr>
<tr>
<td>9</td>
<td>T F Q Y G G C R A K R N N F K R A K Y</td>
</tr>
<tr>
<td>10</td>
<td>T F Q Y G G C R G K K N N F K R A K Y</td>
</tr>
<tr>
<td>11</td>
<td>T F F Y G G C L G K R N N F K R A K Y</td>
</tr>
<tr>
<td>12</td>
<td>T F F Y G G S L G K R N N F K R A K Y</td>
</tr>
<tr>
<td>13</td>
<td>P F F Y G G C G G K K N N F K R A K Y</td>
</tr>
<tr>
<td>14</td>
<td>T F F Y G G C R G K G N N Y K R A K Y</td>
</tr>
<tr>
<td>15</td>
<td>P F F Y G G C R G K R N N F L R A K Y</td>
</tr>
<tr>
<td>16</td>
<td>T F F Y G G C R G K R N N F K R E K Y</td>
</tr>
<tr>
<td>17</td>
<td>P F F Y G G C R A K K N N F K R A K E</td>
</tr>
<tr>
<td>18</td>
<td>T F F Y G G C R G K R N N F K R A K D</td>
</tr>
<tr>
<td>19</td>
<td>T F F Y G G C R A K R N N F D R A K Y</td>
</tr>
<tr>
<td>20</td>
<td>T F F Y G G C R G K K N N F K R A E Y</td>
</tr>
<tr>
<td>21</td>
<td>P F F Y G G C G A N R N N F K R A K Y</td>
</tr>
<tr>
<td>22</td>
<td>T F F Y G G C G G K K N N F K T A K Y</td>
</tr>
<tr>
<td>23</td>
<td>T F F Y G G C R G N R N N F L R A K Y</td>
</tr>
<tr>
<td>24</td>
<td>T F F Y G G C R G N R N N F K T A K Y</td>
</tr>
<tr>
<td>25</td>
<td>T F F Y G G S R G N R N N F K T A K Y</td>
</tr>
<tr>
<td>26</td>
<td>T F F Y G G C L G N G N N F K R A K Y</td>
</tr>
<tr>
<td>27</td>
<td>T F F Y G G C L G N R N N F L R A K Y</td>
</tr>
<tr>
<td>28</td>
<td>T F F Y G G C L G N R N N F K T A K Y</td>
</tr>
</tbody>
</table>
Polypeptides Nos. 5, 67, 76, and 91, include the sequences of SEQ ID NOS: 5, 67, 76, and 91, respectively, and are amidated at the C-terminus. Polypeptides Nos. 107, 109, and 110 include the sequences of SEQ ID NOS: 97, 109, and 110, respectively, and are acetylated at the N-terminus.

In any of the above aspects, the targeting moiety may include an amino acid sequence having the formula:

\[
X_1\text{-}X_2\text{-}X_3\text{-}X_4\text{-}X_5\text{-}X_6\text{-}X_7\text{-}X_8\text{-}X_9\text{-}X_10\times\text{-}X_1\text{-}X_1\text{-}X_2\text{-}X_3\text{-}X_4\text{-}X_5\text{-}X_6\text{-}X_7\text{-}X_8\text{-}X_9\text{-}X_10\text{-}X_11\text{-}X_12\text{-}X_13\text{-}X_14\text{-}X_15\text{-}X_16\text{-}X_17\text{-}X_18\text{-}X_19
\]

where each of \(X_1\text{-}X_1\) (e.g., \(X_1\text{-}X_6\), \(X_8\), \(X_9\), \(X_11\text{-}X_14\), and \(X_16\text{-}X_19\)) is, independently, any amino acid (e.g., a naturally occurring amino acid such as Ala, Arg, Asn, Asp, Cys, Gin, Glu, Gly, His, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, and Val) or absent and at least one (e.g., 2 or 3) of \(X_1\), \(X_{10}\), and \(X_{15}\) is arginine. In some embodiments, \(X_7\) is Ser or Cys; or \(X_{10}\) and \(X_{15}\) each are independently Arg or Lys. In some embodiments, the residues from \(X_1\) through \(X_{19}\), inclusive, are substantially identical to any of the amino acid sequences of any one of SEQ ID NOS: 1-105 and 107-117 (e.g., Angiopep-1, Angiopep-2, Angiopep-3, Angiopep-4a, Angiopep-4b, Angiopep-5, Angiopep-6, and Angiopep-7). In some embodiments, at least one (e.g., 2, 3, 4, or 5) of the amino acids \(X_1\text{-}X_9\) is Arg. In some embodiments, the polypeptide has one or more additional cysteine residues at the N-terminal of the polypeptide, the C-terminal of the polypeptide, or both.

In any of the above aspects, the targeting moiety may include the amino acid sequence Lys-Arg-X3-X4-X5-Lys (formula 1a), where \(X_3\) is Asn or Gin; \(X_4\) is Asn or Gin; and \(X_5\) is Phe, Tyr, or Trp; where the polypeptide is optionally fewer than 200 amino acids in length (e.g., fewer than 150, 100, 75, 50, 45, 40, 35, 30, 25, 20, 19, 18, 17, 16, 15, 14, 12, 10, 11, 8, or 7 amino acids, or any range between these numbers); where the polypeptide optionally includes one or more D-isomers of an amino acid recited in formula 1a (e.g., a D-isomer of Lys, Arg, X3, X4, X5, or Lys); and where the polypeptide is not a peptide in Table 2.
In any of the above aspects, the targeting moiety may include the amino acid sequence Lys-Arg-X3-X4-X5-Lys (formula la), where X3 is Asn or Gin; X4 is Asn or Gin; and X5 is Phe, Tyr, or Trp; where the polypeptide is fewer than 19 amino acids in length (e.g., fewer than 18, 17, 16, 15, 14, 12, 10, 11, 8, or 7 amino acids, or any range between these numbers); and where the polypeptide optionally includes one or more D-isomers of an amino acid recited in formula la (e.g., a D-isomer of Lys, Arg, X3, X4, X5, or Lys).

In any of the above aspects, the targeting moiety may include the amino acid sequence of Zl-Lys-Arg-X3-X4-X5-Lys-Z2 (formula lb, where X3 is Asn or Gin; X4 is Asn or Gin; X5 is Phe, Tyr, or Trp; Zl is absent, Cys, Gly, Cys-Gly, Arg-Gly, Cys-Arg-Gly, Ser-Arg-Gly, Cys-Ser-Arg-Gly, Gly-Ser-Arg-Gly, Cys-Gly-Ser-Arg-Gly, Gly-Gly-Ser-Arg-Gly, Cys-Gly-Gly-Ser-Arg-Gly, Tyr-Gly-Gly-Ser-Arg-Gly, Cys-Tyr-Gly-Gly-Ser-Arg-Gly, Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Cys-Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Phe-Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Cys-Phe-Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Thr-Phe-Phe-Tyr-Gly-Gly-Ser-Arg-Gly, or Cys-Thr-Phe-Phe-Tyr-Gly-Gly-Ser-Arg-Gly; and Z2 is absent, Cys, Tyr, Tyr-Cys, Cys-Tyr, Thr-Glu-Glu-Tyr, or Thr-Glu-Glu-Tyr-Cys; and where the polypeptide optionally comprises one or more D-isomers of an amino acid recited in formula lb, Zl, or Z2.

In any of the above aspects, the targeting moiety may include the amino acid sequence Lys-Arg-Asn-Asn-Phe-Lys. In other embodiments, the targeting moiety has an amino acid sequence of Lys-Arg-Asn-Asn-Phe-Lys-Tyr. In still other embodiments, the targeting moiety has an amino acid sequence of Lys-Arg-Asn-Asn-Phe-Lys-Tyr-Cys.

In any of the above aspects, the targeting moiety may have the amino acid sequence of XI-X2-Asn-Asn-X5-X6 (formula IIa), where XI is Lys or D-Lys; X2 is Arg or D-Arg; X5 is Phe or D-Phe; and X6 is Lys or D-Lys; and where at least one (e.g., at least two, three, or four) of XI, X2, X5, or X6 is a D-amino acid.

In any of the above aspects, the targeting moiety may have the amino acid sequence of XI-X2-Asn-Asn-X5-X6-X7 (formula lib), where XI is Lys or D-Lys; X2 is Arg or D-Arg; X5 is Phe or D-Phe; X6 is Lys or D-Lys; and X7 is Tyr or D-Tyr; and where at least one (e.g., at least two, three, four, or five) of XI, X2, X5, X6, or X7 is a D-amino acid.
In any of the above aspects, the targeting moiety may have the amino acid sequence of Zl-X1-X2-Asn-Asn-X5-X6-X7-Z2 (formula lie), where X1 is Lys or D-Lys; X2 is Arg or D-Arg; X5 is Phe or D-Phe; X6 is Lys or D-Lys; X7 is Tyr or D-Tyr; Zl is absent, Cys, Gly, Cys-Gly, Arg-Gly, Cys-Arg-Gly, Ser-Arg-Gly, Cys-Ser-Arg-Gly, Gly-Ser-Arg-Gly, Cys-Gly-Ser-Arg-Gly, Gly-Gly-Ser-Arg-Gly, Cys-Tyr-Gly-Gly-Ser-Arg-Gly, or Cys-D-Phe-Tyr-Gly-Gly-Ser-Arg-Gly, or Cys-D-Lys-D-Arg-Asn-Asn-D-Phe-D-Lys-Thr-Glu-Tyr, or Thr-Glu-Glu-Tyr-Cys; where at least one of XI, X2, X5, X6, or X7 is a D-amino acid; and where the polypeptide optionally includes one or more D-isomers of an amino acid recited in Zl or Z2.

In any of the above aspects, the targeting moiety may have the amino acid sequence of Thr-Phe-Phe-Tyr-Gly-Gly-Ser-Arg-Gly-Lys-Arg-Asn-Asn-Phe-Lys-Thr-Glu-Glu-Tyr (An2), where any one or more amino acids are D-isomers. For example, the targeting moiety can have 1, 2, 3, 4, or 5 amino acids which are D-isomers. In a preferred embodiment, one or more or all of positions 8, 10, and 11 can be D-isomers. In yet another embodiment, one or more or all of positions 8, 10, 11, and 15 can have D-isomers.

(P5c); Lys-Arg-Asn-Phe-Lys-Tyr-Cys (P6); D-Lys-D-Arg-Asn-Asn-D-Phe-Lys-Tyr-Cys (P6a); D-Lys-D-Arg-Asn-D-Phe-D-Lys-Tyr-Cys (P6b); Thr-Phe-Phe-Tyr-Gly-Gly-Ser-D-Arg-Gly-D-Lys-D-Arg-Asn-Phe-D-Lys-Thr-Glu-Glu-Tyr; and D-Lys-D-Arg-Asn-D-Phe-D-Lys-D-Tyr-Cys (P6c); or a fragment thereof. In other embodiments, the targeting moiety has a sequence of one of the aforementioned peptides having from 0 to 5 (e.g., from 0 to 4, 0 to 3, 0 to 2, 0 to 1, 1 to 5, 1 to 4, 1 to 3, 1 to 2, 2 to 5, 2 to 4, 2 to 3, 3 to 5, 3 to 4, or 4 to 5) substitutions, deletions, or additions of amino acids.

In any of the above aspects, the polypeptide may be Phe-Tyr-Gly-Gly-Ser-Arg-Gly-Lys-Arg-Asn-Phe-Lys-Thr-Glu-Glu; Gly-Gly-Ser-Arg-Gly-Lys-Arg-Asn-Asn-Phe-Lys-Thr-Glu-Glu; Ser-Arg-Gly-Lys-Arg-Asn-Phe-Lys-Thr-Glu-Glu; Gly-Lys-Arg-Asn-Phe-Lys-Thr-Glu-Glu; Lys-Arg-Asn-Phe-Lys-Thr-Glu-Glu; or Lys-Arg-Asn-Phe-Lys, or a fragment thereof.

In any of the above aspects, the polypeptide may be Thr-Phe-Phe-Tyr-Gly-Gly-Ser-D-Arg-Gly-D-Lys-D-Arg-Asn-Phe-Lys-Thr-Glu-Glu-Tyr (3D-An2); Phe-Tyr-Gly-Gly-Ser-Arg-Gly-Lys-Arg-Asn-Asn-Phe-Lys-Thr-Glu-Glu-Tyr (PI); Phe-Tyr-Gly-Gly-Ser-Arg-Gly-D-Lys-D-Arg-Asn-D-Phe-Lys-Thr-Glu-Glu-Tyr-Cys (Pla); Phe-Tyr-Gly-Gly-Ser-Arg-Gly-D-Lys-D-Arg-Asn-D-Phe-D-Lys-Thr-Glu-Glu-Tyr-Cys (Plb); Phe-Tyr-Gly-Gly-Ser-Arg-Gly-D-Lys-D-Arg-Asn-D-Phe-D-Lys-Thr-Glu-Glu-Tyr-Cys (Pic); D-Phe-D-Tyr-Gly-Gly-Ser-D-Arg-Gly-D-Lys-D-Arg-Asn-D-Phe-D-Lys-Thr-Glu-D-Tyr-Cys (Plb) or a fragment thereof (e.g., deletion of 1 to 7 amino acids from the N-terminus of PI, Pla, Plb, Pic, or Pld); a deletion of 1 to 5 amino acids from the C-terminus of PI, Pla, Plb, Pic, or Pld; or deletions of 1 to 7 amino acids from the N-terminus of PI, Pla, Plb, Pic, or Pld and 1 to 5 amino acids from the C-terminus of PI, Pla, Plb, Pic, or Pld).

In any of the targeting moieties described herein, the moiety may include additions or deletions of 1, 2, 3, 4, or 5 amino acids (e.g., from 1 to 3 amino acids) may be made from an amino acid sequence described herein (e.g., from Lys-Arg-X3-X4-X5-Lys).

In any of the targeting moieties described herein, the moiety may have one or more additional cysteine residues at the N-terminal of the polypeptide, the C-terminal of the polypeptide, or both. In other embodiments, the targeting moiety may have one or more additional tyrosine residues at the N-terminal of the polypeptide, the C-terminal of
the polypeptide, or both. In yet further embodiments, the targeting moiety has the amino acid sequence Tyr-Cys and/or Cys-Tyr at the N-terminal of the polypeptide, the C-terminal of the polypeptide, or both.

In certain embodiments of any of the above aspects, the targeting moiety may be fewer than 15 amino acids in length (e.g., fewer than 10 amino acids in length).

In certain embodiments of any of the above aspects, the targeting moiety may have a C-terminus that is amidated. In other embodiments, the targeting moiety is efficiently transported across the BBB (e.g., is transported across the BBB more efficiently than Angiopep-2).

In certain embodiments of any of the above aspects, the fusion protein, targeting moiety, or lysosomal enzyme (e.g., SGSH), fragment, or analog is modified (e.g., as described herein). The fusion protein, targeting moiety, or lysosomal enzyme, fragment, or analog may be amidated, acetylated, or both. Such modifications may be at the amino or carboxy terminus of the polypeptide. The fusion protein, targeting moiety, or lysosomal enzyme, fragment, or analog may also include or be a peptidomimetic (e.g., those described herein) of any of the polypeptides described herein. The fusion protein, targeting moiety, or lysosomal enzyme, fragment, or analog may be in a multimeric form, for example, dimeric form (e.g., formed by disulfide bonding through cysteine residues).

In certain embodiments, the targeting moiety, lysosomal enzyme (e.g., SGSH), enzyme fragment, or enzyme analog has an amino acid sequence described herein with at least one amino acid substitution (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 substitutions), insertion, or deletion. The polypeptide may contain, for example, 1 to 12, 1 to 10, 1 to 5, or 1 to 3 amino acid substitutions, for example, 1 to 10 (e.g., to 9, 8, 7, 6, 5, 4, 3, 2) amino acid substitutions. The amino acid substitution(s) may be conservative or non-conservative. For example, the targeting moiety may have an arginine at one, two, or three of the positions corresponding to positions 1, 10, and 15 of the amino acid sequence of any of SEQ ID NO:1, Angiopep-1, Angiopep-2, Angiopep-3, Angiopep-4a, Angiopep-4b, Angiopep-5, Angiopep-6, and Angiopep-7.

In any of the above aspects, the compound may specifically exclude a polypeptide including or consisting of any of SEQ ID NOS:1-105 and 107-1 17 (e.g., Angiopep-1, Angiopep-2, Angiopep-3, Angiopep-4a, Angiopep-4b, Angiopep-5, Angiopep-6, and
Angiopep-7). In some embodiments, the polypeptides and conjugates of the invention exclude the polypeptides of SEQ ID NOS: 102, 103, 104, and 105.

In any of the above aspects, the linker (X) may be any linker known in the art or described herein. In particular embodiments, the linker is a covalent bond (e.g., a peptide bond), a chemical linking agent (e.g., those described herein), an amino acid or a peptide (e.g., 2, 3, 4, 5, 8, 10, or more amino acids).

In certain embodiments, the linker has the formula:

\[
\text{Y} - \text{N} - \text{Z}
\]

where \( n \) is an integer between 2 and 15 (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15); and either \( Y \) is a thiol on A and Z is a primary amine on B or \( Y \) is a thiol on B and Z is a primary amine on A. In certain embodiments, the linker is an N-Succinimidyl (acetyltthio)acetate (SATA) linker or a hydrazide linker. The linker may be conjugated to the enzyme (e.g., SGSH) or the targeting moiety (e.g., Angiopep-2), through a free amine, a cysteine side chain (e.g., of Angiopep-2-Cys or Cys-Angiopep-2), or through a glycosylation site.

In certain embodiments, the compound has the formula:

\[
\text{Enzyme-Lys-NH} - \text{Targeting moiety}
\]

where the "Lys-NH" group represents either a lysine present in the enzyme or an N-terminal or C-terminal lysine. In another example, the compound has the structure:

\[
\text{Enzyme} - \text{Targeting moiety}
\]

or

\[
\text{Targeting moiety} - \text{Enzyme}
\]
where each -NH- group represents a primary amino present on the targeting moiety and the enzyme, respectively. In particular embodiments, the enzyme may be SGSH or the targeting moiety may be Angiopep-2.

In certain embodiments, the compound is a fusion protein including the targeting moiety (e.g., Angiopep-2) and the lysosomal enzyme (e.g., SGSH), enzyme fragment, or enzyme analog.

In certain embodiments, the linker includes a click-chemistry reaction pair selected from the group consisting of a Huisgen 1,3-dipolar cycloaddition reaction between an alkynyl group and an azido group to form a triazole-containing linker; a Diels-Alder reaction between a diene having a 4π electron system (e.g., an optionally substituted 1,3-unsaturated compound, such as optionally substituted 1,3-butadiene, 1-methoxy-3-trimethylsilyloxy-1,3-butadiene, cyclopentadiene, cyclohexadiene, or furan) and a dienophile or heterodienophile having a 2π electron system (e.g., an optionally substituted alkenyl group or an optionally substituted alkynyl group); a ring opening reaction with a nucleophile and a strained heterocyclyl electrophile; and a splint ligation reaction with a phosphorothioate group and an iodo group; and a reductive amination reaction with an aldehyde group and an amino group. In one aspect of the invention, the linker is selected from the group consisting of monofluorocyclooctyne (MFCO), difluorocyclooctyne (DFCO), cyclooctyne (OCT), dibenzocyclooctyne (DIBO), biarylazacyclooctyne (BARAC), difluorobenzocyclooctyne (DIFBO), and bicyclo[6.1.0]nonyne (BCN). In another aspect, the linker is a maleimide group or an S-acetylthioacetate (SATA) group. The peptide targeting moiety is attached to the linker via an N-terminal azido group or a C-terminal azido group.

In one embodiment, the compound includes an Angiopep-2 joined to SGSH via a BCN linker. This compound can have the general structure

![formula III]

wherein R is:
where \( n \) is the number of Angiopep-2 moieties attached to SGSH via the linker and is between 1 to 6, and the NH group attached to SGSH represents the side chain primary amino group from a lysine in SGSH. The compound can also have the structure

![Formula IV](image)

(formula IV)

The compound can also have the structure

![Formula V](image)

(formula V)

In each of the above formulae, \( R \) is:

![Representation of R](image)

and the NH group attached to SGSH represents the side chain primary amino group from a lysine in SGSH.

In any of the aspects of the compounds of the invention, Angiopep-2 can be derivatized with an azide group at the N- or C-terminus of the polypeptide, such that the azide group can be reacted with an alkyn derivatized linker, in a click-chemistry reaction, to attach the Angiopep-2 to the linker. The invention also features a composition comprising a compound of formula III where an average value of \( n \) is between \( n \) is 1 to 6, e.g. 1 to 4, 3 to 6, 2 to 5, 1 to 3, 2 to 4, 3 to 5, or 4 to 6 (e.g., \( n \) is 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, or 6).
The compound with a BCN linker can also have the structure

(formula VI)

wherein R is:

where n is the number of Angiopep-2 moieties attached to SGSH via the linker and is between 1 to 6, and the NH group attached to SGSH represents the side chain primary amino group from a lysine in SGSH.

The invention features a composition including a compound of formula VI where an average value of n is between 1 to 6, e.g. 1 to 4, 3 to 6, 2 to 5, 1 to 3, 2 to 4, 3 to 5, or 4 to 6 (e.g., n is 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, or 6).

In one embodiment, the compound includes an Angiopep-2 joined to SGSH via a MFCO linker. The Angiopep-2 can be joined to the MFCO linker via the N-terminus amino group of Angiopep-2. The compound can have the structure

(formula VII)

wherein R is:
where \( n \) is the number of Angiopep-2 moieties attached to SGSH via the linker and is between 1 to 6, and the NH group attached to SGSH represents the side chain primary amino group from a lysine in SGSH.

The invention also features a composition including the compound of formula VII where the average value of \( n \) is between 1 and 6 (e.g., 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.4, 4.5, 5, 5.3, 5.5, or 6).

In one aspect of the invention, Angiopep-2 is joined to the MFCO linker via the side chain primary amino group of an amino acid (e.g., a lysine) at the C-terminus of Angiopep-2 and the compound has the structure

\[
\text{SGSH} \quad \begin{array}{c} \text{N} \\ \text{O} \\ \text{N} \end{array} \quad \begin{array}{c} \text{O} \\ \text{N} \end{array} \quad \begin{array}{c} \text{N} \end{array} \quad \begin{array}{c} \text{N} \end{array} \quad \begin{array}{c} \text{R} \end{array} \quad \text{N} \quad \text{N} \quad \text{n} \\
\text{(formula VIII)}
\]

wherein \( R \) is:

where \( n \) is the number of Angiopep-2 moieties attached to SGSH via the linker and is between 1 to 6, and the NH group attached to SGSH represents the side chain primary amino group from a lysine in SGSH. The invention features a composition including the compound of formula VIII where the average value of \( n \) is between 1 and 6 (e.g., 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.4, 4.5, 4.9, 5, 5.5, or 6).

In another embodiment of the invention, the compound includes Angiopep-2 joined to SGSH via a DBCO linker and has the structure

\[
\text{SGSH} \quad \begin{array}{c} \text{N} \\ \text{O} \\ \text{N} \end{array} \quad \begin{array}{c} \text{O} \\ \text{N} \end{array} \quad \begin{array}{c} \text{N} \end{array} \quad \begin{array}{c} \text{N} \end{array} \quad \begin{array}{c} \text{R} \end{array} \quad \text{N} \quad \text{N} \quad \text{n} \\
\text{(formula IX)}
\]
wherein R is:

where n is the number of Angiopep-2 moieties attached to SGSH via the linker and is between 1 to 6, and the NH group attached to SGSH represents the side chain primary amino group from a lysine in SGSH. The invention features a composition including the compound of formula IX where the average value of n is between 1 and 6 (e.g., 1, 1.3, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, or 6).

The invention also features a compound where Angiopep-2-Cys is joined to SGSH via a maleimide group and has the structure

\[
\text{(formula X)}
\]

where n is the number of Angiopep-2 moieties attached to SGSH via the linker and is between 1 to 6, wherein An₂Cys, the S moiety attached to An₂Cys represents the side chain sulfide on the cysteine in Angiopep-2-Cys, and the NH group attached to SGSH represents the side chain primary amino group from a lysine in SGSH. The invention features a composition including the compound of formula X where the average value of n is between 0.5 and 6 (e.g., 0.5, 0.8, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, or 6).

In an alternate embodiment, Cys-Angiopep-2 is joined to SGSH via a maleimide group and has the structure

\[
\text{(formula XI)}
\]

where n is the number of Angiopep-2 moieties attached to SGSH via the linker and is between 1 to 6, wherein Cys-An₂ is Cys-Angiopep-2, the S moiety attached to Cys-An₂ represents the side chain sulfide on the cysteine in Cys-Angiopep-2, and the NH group attached to SGSH represents the side chain primary amino group from a lysine in SGSH.
The invention features a composition including the compound of formula XI where the average value of n is between 0.5 and 6 (e.g., 0.5, 0.9, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, or 6).

In one aspect of the above embodiments, the linker can be a maleimide group functionalized with an alkyne group selected from the group consisting of monofluorocyclooctyne (MFCO), difluorocyclooctyne (DFCO), cyclooctyne (OCT), dibenzocyclooctyne (DIBO), biarylazacyclooctyne (BARAC), difluorobenzocyclooctyne (DIFBO), and bicyclo[6.1.0]nonyne (BCN) and the alkyne-functionalized maleimide is attached to an Angiopep-2 via an azido group attached to Angiopep-2.

In one embodiment of the invention, the compound includes Angiopep-2 joined to SGSH via an S-acetylthioacetate (SATA) group and has the structure

![Structure Formula XII](formula XII)

where n is the number of Angiopep-2 moieties attached to SGSH via the linker and is between 1-6, An2 is Angiopep-2, the NH group attached to An2 is the N-terminus amino group of Angiopep-2, and the NH group attached to SGSH represents the side chain primary amino group from a lysine in SGSH. The invention features a composition comprising the compound of formula XII where the average value of n is between 1 and 6 (e.g., 1, 1.5, 2, 2.5, 2.6, 3, 3.5, 4, 4.5, 5, 5.5, or 6).

The compounds described above can have 1, 2, 3, 4, 5, or more peptide targeting moieties attached to the enzyme via a linker, where the targeting moiety is Angiopep-2 and the enzyme is a lysosomal enzyme, e.g., SGSH.

The invention also features compositions that include the compounds that are represented by the above formulae, where the average number of Angiopep-2 moieties attached to each SGSH is between 1-6 (e.g., 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, or 6), preferably, between 1.5-5, more preferably between 2-4. In some aspects of the above composition, the average number of Angiopep-2 moieties attached to each SGSH can be about 2 (e.g., 1, 1.5, 2, 2.5, or 3). More preferably, the average number of Angiopep-2
moieties attached to each SGSH can be about 4 (e.g., 2, 2.5, 3, 3.5, 4, 4.5, or 5).
Alternatively, the average number of Angiopep-2 moieties attached to each SGSH can be about 6 (e.g., 3.5, 4, 4.5, 5, 5.5, 6, 6.5, or 7).

The invention features a composition that includes nanoparticles which are conjugated to any of the compounds described above. The invention also features a liposome formulation of any of the compounds featured above.

The invention features a pharmaceutical composition that includes any one of the compounds described above and a pharmaceutically acceptable carrier. The invention also features a method of treating or treating prophylactically a subject having a lysosomal storage disorder, where the method includes administering to a subject any of the above described compounds or compositions. In one aspect of the method, the lysosomal storage disorder is mucopolysaccharidosis Type IIia (MPS-IIIa) and the lysosomal enzyme is SGSH. In yet another aspect of the method, the subject has neurological symptoms. The subject can start treatment at under five years of age, preferably under three years of age. The subject can be an infant. The methods of the invention also include parenteral administration of the compounds and compositions of the invention.

By "subject" is meant a human or non-human animal (e.g., a mammal).

By "lysosomal enzyme" is meant any enzyme that is found in the lysosome in which a defect in that enzyme can lead to a lysosomal storage disorder.

By "lysosomal storage disorder" is meant any disease caused by a defect in a lysosomal enzyme. Approximately fifty such disorders have been identified.

By "targeting moiety" is meant a compound or molecule such as a polypeptide or a polypeptide mimetic that can be transported into a particular cell type (e.g., liver, lungs, kidney, spleen, or muscle), into particular cellular compartments (e.g., the lysosome), or across the BBB. In certain embodiments, the targeting moiety may bind to receptors present on brain endothelial cells and thereby be transported across the BBB by transcytosis. The targeting moiety may be a molecule for which high levels of transendothelial transport may be obtained, without affecting the cell or BBB integrity.

The targeting moiety may be a polypeptide or a peptidomimetic and may be naturally occurring or produced by chemical synthesis or recombinant genetic technology.
By "treating" a disease, disorder, or condition in a subject is meant reducing at least one symptom of the disease, disorder, or condition by administrating a therapeutic agent to the subject.

By "treating prophylactically" a disease, disorder, or condition in a subject is meant reducing the frequency of occurrence of or reducing the severity of a disease, disorder or condition by administering a therapeutic agent to the subject prior to the onset of disease symptoms.

By a polypeptide which is "efficiently transported across the BBB" is meant a polypeptide that is able to cross the BBB at least as efficiently as Angiopep-6 (i.e., greater than 38.5% that of Angiopep-1 (250 nM) in the in situ brain perfusion assay described in U.S. Patent Application No. 11/807,597, filed May 29, 2007, hereby incorporated by reference). Accordingly, a polypeptide which is "not efficiently transported across the BBB" is transported to the brain at lower levels (e.g., transported less efficiently than Angiopep-6).

By a polypeptide or compound which is "efficiently transported to a particular cell type" is meant that the polypeptide or compound is able to accumulate (e.g., either due to increased transport into the cell, decreased efflux from the cell, or a combination thereof) in that cell type to at least a 10% (e.g., 25%, 50%, 100%, 200%, 500%, 1,000%, 5,000%, or 10,000%) greater extent than either a control substance, or, in the case of a conjugate, as compared to the unconjugated agent. Such activities are described in detail in International Application Publication No. WO 2007/009229, hereby incorporated by reference.

By "substantial identity" or "substantially identical" is meant a polypeptide or polynucleotide sequence that has the same polypeptide or polynucleotide sequence, respectively, as a reference sequence, or has a specified percentage of amino acid residues or nucleotides, respectively, that are the same at the corresponding location within a reference sequence when the two sequences are optimally aligned. For example, an amino acid sequence that is "substantially identical" to a reference sequence has at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the reference amino acid sequence. For polypeptides, the length of comparison sequences will generally be at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 50, 75, 90, 100, 150, 200, 250, 300, or 350 contiguous amino acids.
(e.g., a full-length sequence). For nucleic acids, the length of comparison sequences will generally be at least 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides (e.g., the full-length nucleotide sequence). Sequence identity may be measured using sequence analysis software on the default setting (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). Such software may match similar sequences by assigning degrees of homology to various substitutions, deletions, and other modifications.

Other features and advantages of the invention will be apparent from the following Detailed Description, the drawings, and the claims.

**Brief Description of the Drawings**

*Figure 1* is an image illustrating three SGS-H-AngioPep-2 constructs.

*Figure 2* is an image illustrating the SDS-PAGE analysis of purified SGS-H-AngioPep2 constructs.

*Figure 3* is an image illustrating the reaction used to determine the SGS-H enzymatic activity of the constructs.

**Detailed Description**

The present invention relates to compounds that include a lysosomal enzyme (e.g., SGS-H) and a targeting moiety (e.g., Angiopep-2) joined by a linker (e.g., a peptide bond). The targeting moiety is capable of transporting the enzyme to the lysosome and/or across the BBB. Such compounds are exemplified by Angiopep-2-SGSH conjugates and fusion proteins. These proteins maintain SGS-H enzymatic activity both in an enzymatic assay and in a cellular model of MPS-IIIa. Because targeting moieties such as Angiopep-2 are capable of transporting proteins across the BBB, these conjugates are expected to have not only peripheral activity, but have activity in the central nervous system (CNS). In addition, targeting moieties such as Angiopep-2 are taken up by cells by receptor mediated transport mechanism (such as LRP-1) into lysosomes.

Accordingly, we believe that these targeting moieties can increase enzyme concentrations in the lysosome, thus resulting in more effective therapy, particular in tissues and organs that express the LRP-1 receptor, such as liver, kidney, and spleen.
These features overcome some of the biggest disadvantages of current therapeutic approaches because intravenous administration of SGSH by itself does not treat CNS disease symptoms. In contrast to physical methods for bypassing the BBB, such intrathecal or intracranial administration, which are highly invasive and thus generally an unattractive solution to the problem of CNS delivery, the present invention allows for noninvasive brain delivery. In addition, improved transport of the therapeutic to the lysosomes may allow for reduced dosing or reduced frequency of dosing, as compared to standard enzyme replacement therapy.

Lysosomal storage disorders

Lysosomal storage disorders are a group of disorders in which the metabolism of lipids, glycoproteins, or mucopolysaccharides is disrupted based on enzyme dysfunction. This dysfunction leads to cellular buildup of the substance that cannot be properly metabolized. Symptoms vary from disease to disease, but problems in the organ systems (liver, heart, lung, spleen), bones, as well as neurological problems are present in many of these diseases. Typically, these diseases are caused by rare genetic defects in the relevant enzymes. Most of these diseases are inherited in autosomal recessive fashion.

Lysosomal enzymes

The present invention may use any lysosomal enzyme known in the art that is useful for treating a lysosomal storage disorder. The compounds of the present invention are exemplified by heparan sulfatase (SGSH). The compounds may include SGSH, a fragment of SGSH that retains enzymatic activity, or an SGSH analog, which may include amino acid sequences substantially identical (e.g., at least 70, 80, 85, 90, 95, 96, 97, 98, or 99% identical) to the human SGSH sequence and retains enzymatic activity.

To test whether particular fragment or analog has enzymatic activity, the skilled artisan can use any appropriate assay. Assays for measuring SGSH activity, for example, are known in art, including those described in Bielicki et al., Biochem. J. 329:145-150, 1998 and Yogalingam et al, J. Biol. Chem. 271:27259-27265, 1996. Using any of these assays, the skilled artisan would be able to determine whether a particular SGSH fragment or analog has enzymatic activity.
In certain embodiments, an enzyme fragment (e.g., an SGSH fragment) is used. SGSH fragments may be at least 50, 100, 150, 200, 250, 300, 350, 400, 450, or 500 amino in length. In certain embodiments, the enzyme may be modified, e.g., using any of the polypeptide modifications described herein.

Targeting moieties

The compounds of the invention can feature any of targeting moieties described herein, for example, any of the peptides described in Table 1 (e.g., Angiopep-1, Angiopep-2, or reversed Angiopep-2), or a fragment or analog thereof. In certain embodiments, the polypeptide may have at least 35%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, or even 100% identity to a polypeptide described herein. The polypeptide may have one or more (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15) substitutions relative to one of the sequences described herein. Other modifications are described in greater detail below.

The invention also features fragments of these polypeptides (e.g., a functional fragment). In certain embodiments, the fragments are capable of efficiently being transported to or accumulating in a particular cell type (e.g., liver, eye, lung, kidney, or spleen) or are efficiently transported across the BBB. Truncations of the polypeptide may be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or more amino acids from either the N-terminus of the polypeptide, the C-terminus of the polypeptide, or a combination thereof. Other fragments include sequences where internal portions of the polypeptide are deleted.

Additional polypeptides may be identified by using one of the assays or methods described herein. For example, a candidate polypeptide may be produced by conventional peptide synthesis, conjugated with paclitaxel and administered to a laboratory animal. A biologically-active polypeptide conjugate may be identified, for example, based on its ability to increase survival of an animal injected with tumor cells and treated with the conjugate as compared to a control which has not been treated with a conjugate (e.g., treated with the unconjugated agent). For example, a biologically active polypeptide may be identified based on its location in the parenchyma in an in situ cerebral perfusion assay.

Assays to determine accumulation in other tissues may be performed as well. Labelled conjugates of a polypeptide can be administered to an animal, and accumulation
in different organs can be measured. For example, a polypeptide conjugated to a
detectable label (e.g., a near-IR fluorescence spectroscopy label such as Cy5.5) allows
live in vivo visualization. Such a polypeptide can be administered to an animal, and the
presence of the polypeptide in an organ can be detected, thus allowing determination of
the rate and amount of accumulation of the polypeptide in the desired organ. In other
embodiments, the polypeptide can be labelled with a radioactive isotope (e.g., ^125I). The
polypeptide is then administered to an animal. After a period of time, the animal is
sacrificed and the organs are extracted. The amount of radioisotope in each organ can
then be measured using any means known in the art. By comparing the amount of a
labeled candidate polypeptide in a particular organ relative to the amount of a labeled
control polypeptide, the ability of the candidate polypeptide to access and accumulate in a
particular tissue can be ascertained. Appropriate negative controls include any peptide or
polypeptide known not to be efficiently transported into a particular cell type (e.g., a
peptide related to Angiopep that does not cross the BBB, or any other peptide).

Additional sequences are described in U.S. Patent No. 5,807,980 (e.g., SEQ ID
NO: 102 therein), 5,780,265 (e.g., SEQ ID NO: 103), 5,1 18,668 (e.g., SEQ ID NO: 105).
An exemplary nucleotide sequence encoding an aprotinin analog atgagaccag atttctgcct
cgagccgccg tacactgggc cctgcaaagc cgagccgccgcctgaaacgctctgacttctct acaatgcaa ggccaggctg
tctagacct tctatacgg cgctgacgta gctaagcgta atcggccgaa gactgcgtgct acacgtcga
SEQ ID NO: 106; Genbank accession No. X04666). Other examples of aprotinin analogs
may be found by performing a protein BLAST (Genbank:
www.ncbi.nlm.nih.gov/BLAST/) using the synthetic aprotinin sequence (or portion thereof) disclosed in International Application No. PCT/CA2004/00001 1. Exemplary
aprotinin analogs are also found under accession Nos. CAA37967 (GL58005) and
140521 8C (GL3604747).

**Modified polypeptides**

The fusion proteins, targeting moieties, and lysosomal enzymes, fragments, or
analogs used in the invention may have a modified amino acid sequence. In certain
embodiments, the modification does not destroy significantly a desired biological activity
(e.g., ability to cross the BBB or enzymatic activity). The modification may reduce (e.g.,
by at least 5%, 10%, 20%, 25%, 35%, 50%, 60%, 70%, 75%, 80%, 90%, or 95%), may
have no effect, or may increase (e.g., by at least 5%, 10%, 25%, 50%, 100%, 200%, 500%, or 1000%) the biological activity of the original polypeptide. The modified peptide vector or polypeptide therapeutic may have or may optimize a characteristic of a polypeptide, such as in vivo stability, bioavailability, toxicity, immunological activity, immunological identity, and conjugation properties.

Modifications include those by natural processes, such as posttranslational processing, or by chemical modification techniques known in the art. Modifications may occur anywhere in a polypeptide including the polypeptide backbone, the amino acid side chains and the amino- or carboxy-terminus. The same type of modification may be present in the same or varying degrees at several sites in a given polypeptide, and a polypeptide may contain more than one type of modification. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslational natural processes or may be made synthetically. Other modifications include pegylation, acetylation, acylation, addition of acetomidomethyl (Acm) group, ADP-ribosylation, alkylation, amidation, biotinylation, carbamoylation, carboxymethylation, esterification, covalent attachment to flavin, covalent attachment to a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of drug, covalent attachment of a marker (e.g., fluorescent or radioactive), covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation and ubiquitination.

A modified polypeptide can also include an amino acid insertion, deletion, or substitution, either conservative or non-conservative (e.g., D-amino acids, desamino acids) in the polypeptide sequence (e.g., where such changes do not substantially alter the biological activity of the polypeptide). In particular, the addition of one or more cysteine residues to the amino or carboxy terminus of any of the polypeptides of the invention can facilitate conjugation of these polypeptides by, e.g., disulfide bonding. For example,
Angiopep-1 (SEQ ID NO:67), Angiopep-2 (SEQ ID NO:97), or Angiopep-7 (SEQ ID NO:1 12) can be modified to include a single cysteine residue at the amino-terminus (SEQ ID NOS: 71, 113, and 115, respectively) or a single cysteine residue at the carboxy-terminus (SEQ ID NOS: 72, 114, and 116, respectively). Amino acid substitutions can be conservative (i.e., wherein a residue is replaced by another of the same general type or group) or non-conservative (i.e., wherein a residue is replaced by an amino acid of another type). In addition, a non-naturally occurring amino acid can be substituted for a naturally occurring amino acid (i.e., non-naturally occurring conservative amino acid substitution or a non-naturally occurring non-conservative amino acid substitution).

Polypeptides made synthetically can include substitutions of amino acids not naturally encoded by DNA (e.g., non-naturally occurring or unnatural amino acid).

Examples of non-naturally occurring amino acids include D-amino acids, an amino acid having an acetylaminomethyl group attached to a sulfur atom of a cysteine, a pegylated amino acid, the omega amino acids of the formula NH₂(CH₂)ₙCOOH wherein n is 2-6, neutral nonpolar amino acids, such as sarcosine, t-butyl alanine, t-butyl glycine, N-methyl isoleucine, and norleucine. Phenylglycine may substitute for Trp, Tyr, or Phe; citrulline and methionine sulfoxide are neutral nonpolar, cysteic acid is acidic, and ornithine is basic. Proline may be substituted with hydroxyproline and retain the conformation conferring properties.

Analogs may be generated by substitutional mutagenesis and retain the biological activity of the original polypeptide. Examples of substitutions identified as "conservative substitutions" are shown in Table 2. If such substitutions result in a change not desired, then other type of substitutions, denominated "exemplary substitutions" in Table 2, or as further described herein in reference to amino acid classes, are introduced and the products screened.

Substantial modifications in function or immunological identity are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side chain properties:
(1) hydrophobic: norleucine, methionine (Met), Alanine (Ala), Valine (Val), Leucine (Leu), Isoleucine (He), Histidine (His), Tryptophan (Trp), Tyrosine (Tyr), Phenylalanine (Phe),

(2) neutral hydrophilic: Cysteine (Cys), Serine (Ser), Threonine (Thr)

(3) acidic/negatively charged: Aspartic acid (Asp), Glutamic acid (Glu)

(4) basic: Asparagine (Asn), Glutamine (Gin), Histidine (His), Lysine (Lys), Arginine (Arg)

(5) residues that influence chain orientation: Glycine (Gly), Proline (Pro);

(6) aromatic: Tryptophan (Trp), Tyrosine (Tyr), Phenylalanine (Phe), Histidine (His),

(7) polar: Ser, Thr, Asn, Gin

(8) basic positively charged: Arg, Lys, His, and;

(9) charged: Asp, Glu, Arg, Lys, His

Other amino acid substitutions are listed in Table 2.

**Table 2: Amino acid substitutions**

<table>
<thead>
<tr>
<th>Original residue</th>
<th>Exemplary substitution</th>
<th>Conservative substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala (A)</td>
<td>Val, Leu, Ile</td>
<td>Val</td>
</tr>
<tr>
<td>Arg (R)</td>
<td>Lys, Gin, Asn</td>
<td>Lys</td>
</tr>
<tr>
<td>Asn (N)</td>
<td>Gln, His, Lys, Arg</td>
<td>Gln</td>
</tr>
<tr>
<td>Asp (D)</td>
<td>Glu</td>
<td>Glu</td>
</tr>
<tr>
<td>Cys (C)</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td>Gln (Q)</td>
<td>Asn</td>
<td>Asn</td>
</tr>
<tr>
<td>Glu (E)</td>
<td>Asp</td>
<td>Asp</td>
</tr>
<tr>
<td>Gly (G)</td>
<td>Pro</td>
<td>Pro</td>
</tr>
<tr>
<td>His (H)</td>
<td>Asn, Gin, Lys, Arg</td>
<td>Arg</td>
</tr>
<tr>
<td>Ile (I)</td>
<td>Leu, Val, Met, Ala, Phe, norleucine</td>
<td>Leu</td>
</tr>
<tr>
<td>Leu (L)</td>
<td>Norleucine, Ile, Val, Met, Ala, Phe</td>
<td>Ile</td>
</tr>
<tr>
<td>Lys (K)</td>
<td>Arg, Gin, Asn</td>
<td>Arg</td>
</tr>
<tr>
<td>Met (M)</td>
<td>Leu, Phe, Ile</td>
<td>Leu</td>
</tr>
<tr>
<td>Phe (F)</td>
<td>Leu, Val, Ile, Ala</td>
<td>Leu</td>
</tr>
<tr>
<td>Pro (P)</td>
<td>Gly</td>
<td>Gly</td>
</tr>
<tr>
<td>Ser (S)</td>
<td>Thr</td>
<td>Thr</td>
</tr>
<tr>
<td>Thr (T)</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td>Trp (W)</td>
<td>Tyr</td>
<td>Tyr</td>
</tr>
<tr>
<td>Tyr (Y)</td>
<td>Trp, Phe, Thr, Ser</td>
<td>Phe</td>
</tr>
<tr>
<td>Val (V)</td>
<td>Ile, Leu, Met, Phe, Ala, norleucine</td>
<td>Leu</td>
</tr>
</tbody>
</table>
Polypeptide derivatives and peptidomimetics

In addition to polypeptides consisting of naturally occurring amino acids, peptidomimetics or polypeptide analogs are also encompassed by the present invention and can form the fusion proteins, targeting moieties, or lysosomal enzymes, enzyme fragments, or enzyme analogs used in the compounds of the invention. Polypeptide analogs are commonly used in the pharmaceutical industry as non-peptide drugs with properties analogous to those of the template polypeptide. The non-peptide compounds are termed "peptide mimetics" or peptidomimetics (Fauchere et al., Infect. Immun. 54:283-287, 1986 and Evans et al, J. Med. Chem. 30:1229-1239, 1987). Peptide mimetics that are structurally related to therapeutically useful peptides or polypeptides may be used to produce an equivalent or enhanced therapeutic or prophylactic effect.

Generally, peptidomimetics are structurally similar to the paradigm polypeptide (i.e., a polypeptide that has a biological or pharmacological activity) such as naturally-occurring receptor-binding polypeptides, but have one or more peptide linkages optionally replaced by linkages such as -CH₂NH-, -CH₂S-, -CH₂CH₂-, -CH=CH- (cis and trans), -CH₂SO-, -CH(OH)CH₂-, -COCH₂- etc., by methods well known in the art (Spatola, Peptide Backbone Modifications, Vega Data, 1:267, 1983; Spatola et al., Life Sci. 38:1243-1249, 1986; Hudson et al, Int. J. Pept. Res. 14:177-185, 1979; and Weinstein, 1983, Chemistry and Biochemistry, of Amino Acids, Peptides and Proteins, Weinstein eds, Marcel Dekker, New York). Such polypeptide mimetics may have significant advantages over naturally occurring polypeptides including more economical production, greater chemical stability, enhanced pharmacological properties (e.g., half-life, absorption, potency, efficiency), reduced antigenicity, and others.

While the targeting moieties described herein may efficiently cross the BBB or target particular cell types (e.g., those described herein), their effectiveness may be reduced by the presence of proteases. Likewise, the effectiveness of the lysosomal enzymes, enzyme fragments, or enzyme analogs used in the compounds of the invention may be similarly reduced. Serum proteases have specific substrate requirements, including L-amino acids and peptide bonds for cleavage. Furthermore, exopeptidases, which represent the most prominent component of the protease activity in serum, usually act on the first peptide bond of the polypeptide and require a free N-terminus (Powell et
al., Pharm. Res. 10:1268-1273, 1993). In light of this, it is often advantageous to use modified versions of polypeptides. The modified polypeptides retain the structural characteristics of the original L-amino acid polypeptides, but advantageously are not readily susceptible to cleavage by protease and/or exopeptidases.

Systematic substitution of one or more amino acids of a consensus sequence with D-amino acid of the same type (e.g., an enantiomer; D-lysine in place of L-lysine) may be used to generate more stable polypeptides. Thus, a polypeptide derivative or peptidomimetic as described herein may be all L-, all D-, or mixed D, L polypeptides. The presence of an N-terminal or C-terminal D-amino acid increases the in vivo stability of a polypeptide because peptidases cannot utilize a D-amino acid as a substrate (Powell et al., Pharm. Res. 10:1268-1273, 1993). Reverse-D polypeptides are polypeptides containing D-amino acids, arranged in a reverse sequence relative to a polypeptide containing L-amino acids. Thus, the C-terminal residue of an L-amino acid polypeptide becomes N-terminal for the D-amino acid polypeptide, and so forth. Reverse D-polypeptides retain the same tertiary conformation and therefore the same activity, as the L-amino acid polypeptides, but are more stable to enzymatic degradation in vitro and in vivo, and thus have greater therapeutic efficacy than the original polypeptide (Brady and Dodson, Nature 368:692-693, 1994 and Jameson et al., Nature 368:744-746, 1994). In addition to reverse-D-polypeptides, constrained polypeptides comprising a consensus sequence or a substantially identical consensus sequence variation may be generated by methods well known in the art (Rizo et al., Ann. Rev. Biochem. 61:387-418, 1992). For example, constrained polypeptides may be generated by adding cysteine residues capable of forming disulfide bridges and, thereby, resulting in a cyclic polypeptide. Cyclic polypeptides have no free N- or C-termini. Accordingly, they are not susceptible to proteolysis by exopeptidases, although they are, of course, susceptible to endopeptidases, which do not cleave at polypeptide termini. The amino acid sequences of the polypeptides with N-terminal or C-terminal D-amino acids and of the cyclic polypeptides are usually identical to the sequences of the polypeptides to which they correspond, except for the presence of N-terminal or C-terminal D-amino acid residue, or their circular structure, respectively.

A cyclic derivative containing an intramolecular disulfide bond may be prepared by conventional solid phase synthesis while incorporating suitable S-protected cysteine or
homocysteine residues at the positions selected for cyclization such as the amino and carboxy termini (Sah et al., *J. Pharm. Pharmacol.* 48:197, 1996). Following completion of the chain assembly, cyclization can be performed either (1) by selective removal of the S-protecting group with a consequent on-support oxidation of the corresponding two free SH-functions, to form a S-S bonds, followed by conventional removal of the product from the support and appropriate purification procedure or (2) by removal of the polypeptide from the support along with complete side chain de-protection, followed by oxidation of the free SH-functions in highly dilute aqueous solution.

The cyclic derivative containing an intramolecular amide bond may be prepared by conventional solid phase synthesis while incorporating suitable amino and carboxyl side chain protected amino acid derivatives, at the position selected for cyclization. The cyclic derivatives containing intramolecular -S-alkyl bonds can be prepared by conventional solid phase chemistry while incorporating an amino acid residue with a suitable amino-protected side chain, and a suitable S-protected cysteine or homocysteine residue at the position selected for cyclization.

Another effective approach to confer resistance to peptidases acting on the N-terminal or C-terminal residues of a polypeptide is to add chemical groups at the polypeptide termini, such that the modified polypeptide is no longer a substrate for the peptidase. One such chemical modification is glycosylation of the polypeptides at either or both termini. Certain chemical modifications, in particular N-terminal glycosylation, have been shown to increase the stability of polypeptides in human serum (Powell et al., *Pharm. Res.* 10:1268-1273, 1993). Other chemical modifications which enhance serum stability include, but are not limited to, the addition of an N-terminal alkyl group, consisting of a lower alkyl of from one to twenty carbons, such as an acetyl group, and/or the addition of a C-terminal amide or substituted amide group. In particular, the present invention includes modified polypeptides consisting of polypeptides bearing an N-terminal acetyl group and/or a C-terminal amide group.

Also included by the present invention are other types of polypeptide derivatives containing additional chemical moieties not normally part of the polypeptide, provided that the derivative retains the desired functional activity of the polypeptide. Examples of such derivatives include (1) N-acyl derivatives of the amino terminal or of another free amino group, wherein the acyl group may be an alkanoyl group (e.g., acetyl, hexanoyl,
octanoyl) an aroyl group (e.g., benzoyl) or a blocking group such as F-moc (fluorenylmethyl-O-CO-); (2) esters of the carboxy terminal or of another free carboxy or hydroxyl group; (3) amide of the carboxy-terminal or of another free carboxyl group produced by reaction with ammonia or with a suitable amine; (4) phosphorylated derivatives; (5) derivatives conjugated to an antibody or other biological ligand and other types of derivatives.

Longer polypeptide sequences which result from the addition of additional amino acid residues to the polypeptides described herein are also encompassed in the present invention. Such longer polypeptide sequences can be expected to have the same biological activity and specificity (e.g., cell tropism) as the polypeptides described above. While polypeptides having a substantial number of additional amino acids are not excluded, it is recognized that some large polypeptides may assume a configuration that masks the effective sequence, thereby preventing binding to a target (e.g., a member of the LRP receptor family). These derivatives could act as competitive antagonists. Thus, while the present invention encompasses polypeptides or derivatives of the polypeptides described herein having an extension, desirably the extension does not destroy the cell targeting activity or enzymatic activity of the compound.

Other derivatives included in the present invention are dual polypeptides consisting of two of the same, or two different polypeptides, as described herein, covalently linked to one another either directly or through a spacer, such as by a short stretch of alanine residues or by a putative site for proteolysis (e.g., by cathepsin, see e.g., U.S. Patent No. 5,126,249 and European Patent No. 495 049). Multimers of the polypeptides described herein consist of a polymer of molecules formed from the same or different polypeptides or derivatives thereof.

The present invention also encompasses polypeptide derivatives that are chimeric or fusion proteins containing a polypeptide described herein, or fragment thereof, linked at its amino- or carboxy-terminal end, or both, to an amino acid sequence of a different protein. Such a chimeric or fusion protein may be produced by recombinant expression of a nucleic acid encoding the protein. For example, a chimeric or fusion protein may contain at least 6 amino acids shared with one of the described polypeptides which desirably results in a chimeric or fusion protein that has an equivalent or greater functional activity.
**Assays to identify peptidomimetics**

As described above, non-peptidyl compounds generated to replicate the backbone geometry and pharmacophore display (peptidomimetics) of the polypeptides described herein often possess attributes of greater metabolic stability, higher potency, longer duration of action, and better bioavailability.


Once a polypeptide as described herein is identified, it can be isolated and purified by any number of standard methods including, but not limited to, differential solubility (e.g., precipitation), centrifugation, chromatography (e.g., affinity, ion exchange, and size exclusion), or by any other standard techniques used for the purification of peptides, peptidomimetics, or proteins. The functional properties of an identified polypeptide of interest may be evaluated using any functional assay known in the art. Desirably, assays for evaluating downstream receptor function in intracellular signaling are used (e.g., cell proliferation).
For example, the peptidomimetics compounds of the present invention may be obtained using the following three-phase process: (1) scanning the polypeptides described herein to identify regions of secondary structure necessary for targeting the particular cell types described herein; (2) using conformationally constrained dipeptide surrogates to refine the backbone geometry and provide organic platforms corresponding to these surrogates; and (3) using the best organic platforms to display organic pharmocophores in libraries of candidates designed to mimic the desired activity of the native polypeptide. In more detail the three phases are as follows. In phase 1, the lead candidate polypeptides are scanned and their structure abridged to identify the requirements for their activity. A series of polypeptide analogs of the original are synthesized. In phase 2, the best polypeptide analogs are investigated using the conformationally constrained dipeptide surrogates. Indolizidin-2-one, indolizidin-9-one and quinolizidinone amino acids (I²aa, I³aa and Qaa respectively) are used as platforms for studying backbone geometry of the best peptide candidates. These and related platforms (reviewed in Halab et al., *Biopolymers* 55:101-122, 2000 and Hanessian et al., *Tetrahedron* 53:12789-12854, 1997) may be introduced at specific regions of the polypeptide to orient the pharmocophores in different directions. Biological evaluation of these analogs identifies improved lead polypeptides that mimic the geometric requirements for activity. In phase 3, the platforms from the most active lead polypeptides are used to display organic surrogates of the pharmocophores responsible for activity of the native peptide. The pharmocophores and scaffolds are combined in a parallel synthesis format. Derivation of polypeptides and the above phases can be accomplished by other means using methods known in the art.

Structure function relationships determined from the polypeptides, polypeptide derivatives, peptidomimetics or other small molecules described herein may be used to refine and prepare analogous molecular structures having similar or better properties. Accordingly, the compounds of the present invention also include molecules that share the structure, polarity, charge characteristics, and side chain properties of the polypeptides described herein.

In summary, based on the disclosure herein, those skilled in the art can develop peptides and peptidomimetics screening assays which are useful for identifying compounds for targeting an agent to particular cell types (e.g., those described herein).
The assays of this invention may be developed for low-throughput, high-throughput, or ultra-high throughput screening formats. Assays of the present invention include assays amenable to automation.

5 **Linkers**

The lysosomal enzyme (e.g., SGSH), enzyme fragment, or enzyme analog may be bound to the targeting moiety either directly (e.g., through a covalent bond such as a peptide bond) or may be bound through a linker. Linkers include chemical linking agents (e.g., cleavable linkers) and peptides.

10 In some embodiments, the linker is a chemical linking agent. The lysosomal enzyme (e.g., SGSH), enzyme fragment, or enzyme analog and targeting moiety may be conjugated through sulfhydryl groups, amino groups (amines), and/or carbohydrates or any appropriate reactive group. Homobifunctional and heterobifunctional cross-linkers (conjugation agents) are available from many commercial sources. Regions available for cross-linking may be found on the polypeptides of the present invention. The cross-linker may comprise a flexible arm, e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 carbon atoms. Exemplary cross-linkers include BS3 ([bis(sulfosuccinimidyl)suberate]; BS3 is a homobifunctional N-hydroxysuccinimide ester that targets accessible primary amines), NHS/EDC (N-hydroxysuccinimide and N-ethyl-

20 -(dimethylaminopropyl)carbodiimide; NHS/EDC allows for the conjugation of primary amine groups with carboxyl groups), sulfo-EMCS ([N-e-Maleimidocaproic acidjhydrazide; sulfo-EMCS are heterobifunctional reactive groups (maleimide and NHS-ester) that are reactive toward sulfhydryl and amino groups), hydrazide (most proteins contain exposed carbohydrates and hydrazide is a useful reagent for linking carboxyl groups to primary amines), and SATA (N-succinimidyl-S-acetylthioacetate; SATA is reactive towards amines and adds protected sulfhydryls groups).

To form covalent bonds, one can use as a chemically reactive group a wide variety of active carboxyl groups (e.g., esters) where the hydroxyl moiety is physiologically acceptable at the levels required to modify the peptide. Particular agents include N-

25 hydroxysuccinimide (NHS), N-hydroxy-sulfosuccinimide (sulfo-NHS), maleimide-

30 benzoyl-succinimide (MBS), gamma-maleimido-butryrolyoxy succinimide ester (GMBS),
maleimido propionic acid (MPA) maleimido hexanoic acid (MHA), and maleimido undecanoic acid (MUA).

Primary amines are the principal targets for NHS esters. Accessible a-amine groups present on the N-termini of proteins and the e-amine of lysine react with NHS esters. An amide bond is formed when the NHS ester conjugation reaction reacts with primary amines releasing N-hydroxysuccinimide. These succinimide-containing reactive groups are herein referred to as succinimidyl groups. In certain embodiments of the invention, the functional group on the protein will be a thiol group and the chemically reactive group will be a maleimido-containing group such as gamma-maleimido-butrylamide (GMBA or MPA). Such maleimido-containing groups are referred to herein as maleido groups.

The maleimido group is most selective for sulphydryl groups on peptides when the pH of the reaction mixture is 6.5-7.4. At pH 7.0, the rate of reaction of maleimido groups with sulphydryls (e.g., thiol groups on proteins such as serum albumin or IgG) is 1000-fold faster than with amines. Thus, a stable thioether linkage between the maleimido group and the sulphydryl can be formed.

In other embodiments, the linker includes at least one amino acid (e.g., a peptide of at least 2, 3, 4, 5, 6, 7, 10, 15, 20, 25, 40, or 50 amino acids). In certain embodiments, the linker is a single amino acid (e.g., any naturally occurring amino acid such as Cys).

In other embodiments, a glycine-rich peptide such as a peptide having the sequence [Gly-Gly-Gly-Gly-Ser]ₙ where n is 1, 2, 3, 4, 5 or 6 is used, as described in U.S. Patent No. 7,271,149. In other embodiments, a serine-rich peptide linker is used, as described in U.S. Patent No. 5,525,491. Serine rich peptide linkers include those of the formula [X-X-X-Gly]ₙ, where up to two of the X are Thr, and the remaining X are Ser, and y is 1 to 5 (e.g., Ser-Ser-Ser-Ser-Gly, where y is greater than 1). In some cases, the linker is a single amino acid (e.g., any amino acid, such as Gly or Cys). Other linkers include rigid linkers (e.g., PAPAP and (PT)ₙP, where n is 2, 3, 4, 5, 6, or 7) and a-helical linkers (e.g., A(EAAAK)ₙA, where n is 1, 2, 3, 4, or 5).

Examples of suitable linkers are succinic acid, Lys, Glu, and Asp, or a dipeptide such as Gly-Lys. When the linker is succinic acid, one carboxyl group thereof may form an amide bond with an amino group of the amino acid residue, and the other carboxyl group thereof may, for example, form an amide bond with an amino group of the peptide
or substituent. When the linker is Lys, Glu, or Asp, the carboxyl group thereof may form an amide bond with an amino group of the amino acid residue, and the amino group thereof may, for example, form an amide bond with a carboxyl group of the substituent. When Lys is used as the linker, a further linker may be inserted between the ε-amino group of Lys and the substituent. In one particular embodiment, the further linker is succinic acid which, e.g., forms an amide bond with the ε-amino group of Lys and with an amino group present in the substituent. In one embodiment, the further linker is Glu or Asp (e.g., which forms an amide bond with the ε-amino group of Lys and another amide bond with a carboxyl group present in the substituent), that is, the substituent is an Nε-acylated lysine residue.

**Click-chemistry linkers**

In particular embodiments, the linker is formed by the reaction between a click-chemistry reaction pair. By click-chemistry reaction pair is meant a pair of reactive groups that participates in a modular reaction with high yield and a high thermodynamic gain, thus producing a click-chemistry linker. In this embodiment, one of the reactive groups is attached to the enzyme moiety and the other reactive group is attached to the polypeptide. Exemplary reactions and click-chemistry pairs include a Huisgen 1,3-dipolar cycloaddition reaction between an alkynyl group and an azido group to form a triazole-containing linker; a Diels-Alder reaction between a diene having a 4π electron system (e.g., an optionally substituted 1,3-unsaturated compound, such as optionally substituted 1,3-butadiene, 1-methoxy-3-trimethylsilyloxy-1,3-butadiene, cyclopentadiene, cyclohexadiene, or furan) and a dienophile or heterodienophile having a 2π electron system (e.g., an optionally substituted alkenyl group or an optionally substituted alkynyl group); a ring opening reaction with a nucleophile and a strained heterocyclcycl electrophile; a splint ligation reaction with a phosphorothioate group and an iodo group; and a reductive amination reaction with an aldehyde group and an amino group (Kolb et al, *Angew. Chem. Int. Ed.*, 40:2004-2021 (2001); Van der Eycken et al, *QSAR Comb. Set*, 26:1 115-1326 (2007)).

In particular embodiments of the invention, the polypeptide is linked to the enzyme moiety by means of a triazole-containing linker formed by the reaction between a alkynyl group and an azido group click-chemistry pair. In such cases, the azido group
may be attached to the polypeptide and the alkynyl group may be attached to the enzyme moiety. Alternatively, the azido group may be attached to the enzyme moiety and the alkynyl group may be attached to the polypeptide. In certain embodiments, the reaction between an azido group and the alkynyl group is uncatalyzed, and in other embodiments the reaction is catalyzed by a copper(I) catalyst (e.g., copper(I) iodide), a copper(II) catalyst in the presence of a reducing agent (e.g., copper(II) sulfate or copper(II) acetate with sodium ascorbate), or a ruthenium-containing catalyst (e.g., Cp*RuCl(PPh₃)₂ or Cp*RuCl(COD)). Exemplary linkers include monofluorocyclooctyne (MFCO), difluorocyclooctyne (DFCO), cyclooctyne (OCT), dibenzocyclooctyne (DIBO), biarylazacyclooctyne (BARAC), difluorobenzocyclooctyne (DIFBO), and bicyclo[6.1.0]nonyne (BCN).

Treatment of lysosomal storage disorders

The present invention also features methods for treatment of lysosomal storage disorders such as MPS-IIIa. MPS-IIIa is characterized by cellular accumulation of glycosaminoglycans (GAG) which results from the inability of the individual to break down these products.

In certain embodiments, treatment is performed on a subject who has been diagnosed with a mutation in the SGSH gene, but does not yet have disease symptoms (e.g., an infant or subject under the age of 2). In other embodiments, treatment is performed on an individual who has at least one MPS-IIIa symptom (e.g., any of those described herein).

Initial MPS-IIIa symptoms begin to manifest themselves young children and these symptoms generally include slowing of development and/or behavioral problems, followed by progressive intellectual decline resulting in severe dementia and progressive motor disease. Of the MPS diseases, MPS-III presents with the mildest physical abnormalities, as the clinical features of the disease are mainly neurological. The methods of the invention may involve treatment of subjects having any of the symptoms described herein. Diagnosis is general confirmed by assay of the enzyme levels in tissue samples and gene sequencing.
Treatment may be performed in a subject of any age, starting from infancy to adulthood. Subjects may begin treatment at birth, six months, or 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, or 18 years of age.

5 Administration and dosage

The present invention also features pharmaceutical compositions that contain a therapeutically effective amount of a compound of the invention. The composition can be formulated for use in a variety of drug delivery systems. One or more physiologically acceptable excipients or carriers can also be included in the composition for proper formulation. Suitable formulations for use in the present invention are found in *Remington’s Pharmaceutical Sciences*, Mack Publishing Company, Philadelphia, PA, 17th ed., 1985. For a brief review of methods for drug delivery, see, e.g., Langer (*Science* 249:1527-1533, 1990).

The pharmaceutical compositions are intended for parenteral, intranasal, topical, oral, or local administration, such as by a transdermal means, for prophylactic and/or therapeutic treatment. The pharmaceutical compositions can be administered parenterally (e.g., by intravenous, intramuscular, or subcutaneous injection), or by oral ingestion, or by topical application or intraarticular injection at areas affected by the vascular or cancer condition. Additional routes of administration include intravascular, intra-arterial, intratumor, intraperitoneal, intraventricular, intraepidural, as well as nasal, ophthalmic, intrascleral, intraorbital, rectal, topical, or aerosol inhalation administration. Sustained release administration is also specifically included in the invention, by such means as depot injections or erodible implants or components. Thus, the invention provides compositions for parenteral administration that include the above mention agents dissolved or suspended in an acceptable carrier, preferably an aqueous carrier, e.g., water, buffered water, saline, PBS, and the like. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH adjusting and buffering agents, tonicity adjusting agents, wetting agents, detergents and the like. The invention also provides compositions for oral delivery, which may contain inert ingredients such as binders or fillers for the formulation of a tablet, a capsule, and the like. Furthermore, this invention provides
compositions for local administration, which may contain inert ingredients such as solvents or emulsifiers for the formulation of a cream, an ointment, and the like.

These compositions may be sterilized by conventional sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile aqueous carrier prior to administration. The pH of the preparations typically will be between 3 and 11, more preferably between 5 and 9 or between 6 and 8, and most preferably between 7 and 8, such as 7 to 7.5. The resulting compositions in solid form may be packaged in multiple single dose units, each containing a fixed amount of the above-mentioned agent or agents, such as in a sealed package of tablets or capsules. The composition in solid form can also be packaged in a container for a flexible quantity, such as in a squeezable tube designed for a topically applicable cream or ointment.

The compositions containing an effective amount can be administered for prophylactic or therapeutic treatments. In prophylactic applications, compositions can be administered to a subject diagnosed as having mutation associated with a lysosomal storage disorder (e.g., a mutation in the SGSH gene). Compositions of the invention can be administered to the subject (e.g., a human) in an amount sufficient to delay, reduce, or preferably prevent the onset of the disorder. In therapeutic applications, compositions are administered to a subject (e.g., a human) already suffering from a lysosomal storage disorder (e.g., MPS-IIIa) in an amount sufficient to cure or at least partially arrest the symptoms of the disorder and its complications. An amount adequate to accomplish this purpose is defined as a "therapeutically effective amount," an amount of a compound sufficient to substantially improve at least one symptom associated with the disease or a medical condition. For example, in the treatment of a lysosomal storage disease, an agent or compound that decreases, prevents, delays, suppresses, or arrests any symptom of the disease or condition would be therapeutically effective. A therapeutically effective amount of an agent or compound is not required to cure a disease or condition but will provide a treatment for a disease or condition such that the onset of the disease or condition is delayed, hindered, or prevented, or the disease or condition symptoms are ameliorated, or the term of the disease or condition is changed or, for example, is less severe or recovery is accelerated in an individual.
Amounts effective for this use may depend on the severity of the disease or condition and the weight and general state of the subject. Idursulfase is recommended for weekly intravenous administration of 0.5 mg/kg. A compound of the invention may, for example, be administered at an equivalent dosage (i.e., accounting for the additional molecular weight of the fusion protein vs. heparan sulfatase) and frequency. The compound may be administered at an iduronase equivalent dose, e.g., 0.01, 0.05, 0.1, 0.5, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75, 1.0, 1.25, 1.5, 2.0, 2.5, 3.0, 4.0, or 5 mg/kg weekly, twice weekly, every other day, daily, or twice daily. The therapeutically effective amount of the compositions of the invention and used in the methods of this invention applied to mammals (e.g., humans) can be determined by the ordinarily-skilled artisan with consideration of individual differences in age, weight, and the condition of the mammal. Because certain compounds of the invention exhibit an enhanced ability to cross the BBB and to enter lysosomes, the dosage of the compounds of the invention can be lower than (e.g., less than or equal to about 90%, 75%, 50%, 40%, 30%, 20%, 15%, 12%, 10%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, or 0.1% of) the equivalent dose of required for a therapeutic effect of the unconjugated agent. The agents of the invention are administered to a subject (e.g. a mammal, such as a human) in an effective amount, which is an amount that produces a desirable result in a treated subject (e.g., reduction of GAG accumulation). Therapeutically effective amounts can also be determined empirically by those of skill in the art.

Single or multiple administrations of the compositions of the invention including an effective amount can be carried out with dose levels and pattern being selected by the treating physician. The dose and administration schedule can be determined and adjusted based on the severity of the disease or condition in the subject, which may be monitored throughout the course of treatment according to the methods commonly practiced by clinicians or those described herein.

The compounds of the present invention may be used in combination with either conventional methods of treatment or therapy or may be used separately from conventional methods of treatment or therapy.

When the compounds of this invention are administered in combination therapies with other agents, they may be administered sequentially or concurrently to an individual. Alternatively, pharmaceutical compositions according to the present invention may be
comprised of a combination of a compound of the present invention in association with a pharmaceutically acceptable excipient, as described herein, and another therapeutic or prophylactic agent known in the art.

The following examples are intended to illustrate, rather than limit, the invention.

**Example 1**

**Design of SGSH-Angiopep-2 fusion proteins**

A series of SGSH-Angiopep-2 constructs were designed. The SGSH (NM_000199) Human cDNA clone was obtained from OriGene (cat# RC208270). Three basic configurations were used: an N-terminal fusion (An2-SGSH and An2-SGSH-His), a C-terminal fusion (SGSH-An2 and SGSH-An2-His), and an N- and C-terminal fusion (An2-SGSH-An2 and An2-SGSH-An2-His), both with and without an 8x His tag (Figure 1). A control without Angiopep-2 was also generated (SGSH and SGSH-His). SDS-PAGE analysis of the three constructs in shown in Figure 2.

**Example 2**

**Enzyme Activity of SGSH-conjugates**

The three SGSH-conjugates were tested for enzyme activity (Table 2). As shown in Table 2, all three of the SGSH-conjugates displayed enzymatic activity. The N-terminal fusion displayed the highest enzymatic activity of the three conjugates and was comparable to the control enzyme.

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>62871</td>
<td>R</td>
<td>P</td>
<td>R</td>
</tr>
<tr>
<td>An2-SGSH</td>
<td>64548</td>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>SGSH-An2</td>
<td>63159</td>
<td>R</td>
<td>P</td>
<td>R</td>
</tr>
</tbody>
</table>
SGSH enzyme assay

rhSGSH enzyme activity was determined using a two-steps fluorimetric assay using 4-Methylumbelliferyl 2-sulfamino-2-deoxy-a-D-glucopyranoside sodium salt as the substrate (Santa Cruz Biotechnology, cat # sc-206911). In the first step rhSGSH enzyme was incubated with the substrate at 37°C for 17h (Figure 3). In the second step an excess of alpha-glucosidase (Sigma-Aldrich) was added and incubated at 37°C for 24h to generate glucosamine and 4-methylumbelliferone (4MU). The reaction was stopped by the addition of carbonate buffer, pH 10.7 containing 0.025% Triton X-100 and the fluorescence of 4MU was determined (365 nm excitation; 450 nm emission).

Other embodiments

All patents, patent applications, and publications mentioned in this specification are herein incorporated by reference, to the same extent as if each independent patent, patent application, or publication was specifically and individually indicated to be incorporated by reference.

What is claimed is:
CLAIMS

1. A compound comprising (a) a peptide or peptidic targeting moiety less than 150 amino acids and (b) a heparan sulfatase, an active fragment thereof, or an analog thereof, wherein said targeting moiety and said enzyme, fragment, or analog are joined by a linker.

2. The compound of claim 1, wherein said targeting moiety comprises an amino acid sequence that is at least 70% identical to any of SEQ ID NOS: 97, 1-96, 98-105 and 107-117.

3. The compound of claim 2, wherein said targeting moiety comprises the sequence of Angiopep-2 (SEQ ID NO: 97).

4. The compound of claim 3, wherein said targeting moiety optionally comprises one or more D-isomers of an amino acid recited in SEQ ID NO: 97.

5. The compound of claim 1, wherein said targeting moiety comprises the formula Lys-Arg-X3-X4-X5-Lys (formula la),
   wherein:
   X3 is Asn or Gin;
   X4 is Asn or Gin; and
   X5 is Phe, Tyr, or Trp;
   wherein said targeting moiety optionally comprises one or more D-isomers of an amino acid recited in formula la.

6. The compound of claim 1, wherein said targeting moiety comprises the formula Z1-Lys-Arg-X3-X4-X5-Lys-Z2 (formula lb),
   wherein:
   X3 is Asn or Gin;
   X4 is Asn or Gin;
   X5 is Phe, Tyr, or Trp;
Z1 is absent, Cys, Gly, Cys-Gly, Arg-Gly, Cys-Arg-Gly, Ser-Arg-Gly, Cys-Ser-
Ser-Arg-Gly, Tyr-Gly-Gly-Ser-Arg-Gly, Cys-Tyr-Gly-Gly-Ser-Arg-Gly, Phe-Tyr-Gly-
Gly-Ser-Arg-Gly, Cys-Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Phe-Phe-Tyr-Gly-Gly-Ser-Arg-
Gly, Cys-Phe-Phe-Tyr-Gly-Gly-Ser-Arg-Gly, Thr-Phe-Phe-Tyr-Gly-Gly-Ser-Arg-Gly, or
Cys-Thr-Phe-Phe-Tyr-Gly-Gly-Ser-Arg-Gly; and

Z2 is absent, Cys, Tyr, Tyr-Cys, Cys-Tyr, Thr-Glu-Glu-Tyr, or Thr-Glu-Glu-Tyr-
Cys; and

wherein said targeting moiety optionally comprises one or more D-isomers of an
amino acid recited in formula lb, Z1, or Z2.

7. The compound of claim 6, wherein said targeting moiety comprises at least
three D-isomers of an amino acid recited in formula lb, Z1, or Z2.

8. The compound of claim 7, wherein said targeting moiety has the formula Thr-
Phe-Phe-Tyr-Gly-Gly-Ser-D-Arg-Gly-D-Lys-D-Arg-Asn-Asn-Phe-D-Lys-Thr-Glu-Glu-
Tyr.

9. The compound of claim 7, wherein said targeting moiety has the formula Thr-
Phe-Phe-Tyr-Gly-Gly-Ser-D-Arg-Gly-D-Lys-D-Arg-Asn-Asn-Phe-D-Lys-Thr-Glu-Glu-
Tyr.

10. The compound of claim 1, wherein said targeting moiety comprises the
formula XI-X2-Asn-Asn-X5-X6 (formula Ha),

wherein:

XI is Lys or D-Lys;
X2 is Arg or D-Arg;
X5 is Phe or D-Phe; and
X6 is Lys or D-Lys; and

wherein at least one of XI, X2, X5, or X6 is a D-amino acid.
11. The compound of claim 1, wherein said targeting moiety comprises the formula X1-X2-Asn-Asn-X5-X6-X7 (formula lib), wherein:
   X1 is Lys or D-Lys;
   X2 is Arg or D-Arg;
   X5 is Phe or D-Phe;
   X6 is Lys or D-Lys; and
   X7 is Tyr or D-Tyr; and
wherein at least one of X1, X2, X5, X6, or X7 is a D-amino acid.

12. The compound of claim 1, wherein said targeting moiety comprises the formula Z1-X1-X2-Asn-Asn-X5-X6-X7-Z2 (formula lie), wherein:
   X1 is Lys or D-Lys;
   X2 is Arg or D-Arg;
   X5 is Phe or D-Phe;
   X6 is Lys or D-Lys;
   X7 is Tyr or D-Tyr;
   Z2 is absent, Cys, Tyr, Tyr-Cys, Cys-Tyr, Thr-Glu-Glu-Tyr, or Thr-Glu-Glu-Tyr-Cys;
wherein at least one of X1, X2, X5, X6, or X7 is a D-amino acid; and
wherein said targeting moiety optionally comprises one or more D-isomers of an amino acid recited in Z1 or Z2.

13. The compound of any of claims 1-12, wherein said linker is a covalent bond or one or more amino acids.
14. The compound of claim 13, wherein said covalent bond is a peptide bond.

15. The compound of claim 14, wherein said compound is a fusion protein.

16. The compound of any of claims 1-12, wherein said linker is a chemical conjugate.

17. The compound of claim 16, wherein said compound has the structure:

```
Enzyme-Lys-NH
```

wherein the "Lys-NH" group represents either a lysine present in the enzyme or an N-terminal or C-terminal lysine.

18. The compound of claim 17, wherein said compound has the structure:

```
Enzyme-Lys-NH
```

19. The compound of claim 16, wherein said compound has the structure:

```
Enzyme
```

or

```
Enzyme
```

wherein each -NH- group represents a primary amino present on the targeting moiety and the enzyme, respectively.
20. The compound of claim 19, wherein said compound has the structure:

![Enzyme](image)

or

![Enzyme](image)

21. The compound of claim 16, wherein said linker is conjugated through a glycosylation site.

22. The compound of claim 21, wherein said linker is a hydrazide or a hydrazide derivative.

23. The compound of any of claims 1-22, wherein said compound further comprises a second targeting moiety, said second targeting moiety being joined to said compound by a second linker.

24. A composition comprising one or more nanoparticles, wherein said nanoparticle is conjugated to any one of the compounds of claims 1-23.

25. A composition comprising a liposome formulation of any one of the compounds of claims 1-23.

26. A pharmaceutical composition comprising a compound of any of claims 1-23 and a pharmaceutically acceptable carrier.

27. A method of treating or treating prophylactically a subject having mucopolysaccharidosis I/IIa, said method comprising administering to said subject a compound of any of claims 1-26.
28. The method of claim 27, wherein said subject has neurological symptoms.

29. The method of claim 27, wherein said subject starts treatment under five years of age.

30. The method of claim 29, wherein said subject starts treatment under three years of age.

31. The method of claim 30, wherein said subject is an infant.

32. The method of claim 27, wherein said administering comprises parenteral administration.
Heparan sulfatase (SGSH) constructs

With a His$_8$ tag at the C-terminus

1. SGSH-His$_8$

<table>
<thead>
<tr>
<th>Signal peptide</th>
<th>SGSH</th>
<th>8x His</th>
</tr>
</thead>
</table>

2. An2-SGSH-His$_8$

<table>
<thead>
<tr>
<th>Signal peptide</th>
<th>Angiopep-2</th>
<th>SGSH</th>
<th>8x His</th>
</tr>
</thead>
</table>

3. SGSH-An2-His$_8$

<table>
<thead>
<tr>
<th>Signal peptide</th>
<th>SGSH</th>
<th>Angiopep-2</th>
<th>8x His</th>
</tr>
</thead>
</table>

4. An2-SGSH-An2-His$_8$

<table>
<thead>
<tr>
<th>Signal peptide</th>
<th>Angiopep-2</th>
<th>SGSH</th>
<th>Angiopep-2</th>
<th>8x His</th>
</tr>
</thead>
</table>

Figure 1
SDS-PAGE analysis of purified heparan sulfatase (SGSH) enzymes

Figure 2
rhSGSH Enzymatic Activity Assay

Step 1
SGSH

Step 2 (24h)
α-glucosidase

D-glucosamine
+ 4MU

Figure 3
INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2014/050523

A. CLASSIFICATION OF SUBJECT MATTER

- IPC: C12N 9/16 (2006.01), C07K 14/81 (2006.01), C07K 19/00 (2006.01), C07K 7/06 (2006.01), C07K 7/08 (2006.01), C12N 9/96 (2006.01)

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

- IPC: C12N 9/16 (2006.01), C07K 14/81 (2006.01), C07K 19/00 (2006.01), C07K 7/06 (2006.01), C07K 7/08 (2006.01), C12N 9/96 (2006.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used)

- Canadian Patent Database, TotalPatent, Scopus, Medline, BIOSIS; Search terms: heparan sulfatase, MPS-III, lysosome, target, blood-brain barrier, peptide, conjugate, fusion protein, angiopep, aprotinin, linkers, D-amino acid, click chemistry, chimeric, triazole

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X,P WO 2013/078564 A2 (BOIVIN, D. ET AL.) 6 June 2013 (06-06-2013) see the whole document</td>
<td>1-32</td>
</tr>
<tr>
<td>X WO 2004/108071 A2 (VERMA, I. ET AL.) 16 December 2004 (16-12-2004) see the whole document, especially abstract, page 5, line 28-page 6, line 18, page 16, lines 28-32, page 21, lines 17-26, page 22, lines 19-26 and claims</td>
<td>1, 2, 5, 6, 13-16 and 23-32</td>
</tr>
<tr>
<td>X WO 2005/002515 A2 (ZANKEL, T. ET AL.) 13 January 2005 (13-01-2005) see the whole document, especially claims 1-59</td>
<td>1, 2, 5, 6, 13-16 and 23-32</td>
</tr>
<tr>
<td>X EP 2 333 074 A1 (STEINFELD, R.) 15 June 2011 (15-06-2011) see the whole document, especially abstract, paragraphs [0007], [0008] and [0035]-[0037], Table 8 and claims</td>
<td>1, 13-16 and 23-32</td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:
  - "A" document defining the general state of the art which is not considered to be of particular relevance
  - "E" earlier application or patent but published on or after the international filing date
  - "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
  - "O" document referring to an oral disclosure, use, exhibition or other means of communication published prior to the international filing date but later than the priority date claimed

Further documents are listed in the continuation of Box C.

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document of the same patent family

Date of the actual completion of the international search

02 July 2014 (02-07-2014)

Date of mailing of the international search report

01 August 2014 (01-08-2014)

Name and mailing address of the ISA/CA

Canadian Intellectual Property Office

Place du Portage 1, C1 14 - 1st Floor, Box PCT

50 Victoria Street

Gatineau, Quebec K1A 0C9

Facsimile No.: 001-819-953-2476

Authorized officer

Sandra Hurley (819) 934-7934
<table>
<thead>
<tr>
<th>Category*</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>WO 03/102583 A1 (LEBOWITZ, J., ET AL.) 11 December 2003 (11-12-2003) see the whole document, especially abstract, paragraphs [0006], [0007], [0012], [0053] and [0076], Table 1 and claims</td>
<td>1. 13-16 and 23-32</td>
</tr>
<tr>
<td>X</td>
<td>WO 2006/108052 A2 (STEFANO, J.) 12 October 2006 (12-10-2006) see the whole document, especially abstract, paragraphs [0010], [0012], [0034], [0035] and [0073] and Table 2</td>
<td>1. 16, 21 and 23-32</td>
</tr>
<tr>
<td>Y</td>
<td>WO 2010/063 122 A1 (CASTAIGNE, J.-P., ET AL.) 10 June 2010 (10-06-2010) see the whole document</td>
<td>3, 4, 7-12, 17-20 and 22</td>
</tr>
<tr>
<td>Y</td>
<td>HEIN, C., ET AL.: &quot;Click Chemistry: a powerful tool for pharmaceutical sciences&quot; PHARMACEUTICAL RESEARCH, October 2008 (10-2008) vol. 25, no. 10, pages 2216-2230, ISSN:0724-8741 see the whole document</td>
<td>17 and 18</td>
</tr>
</tbody>
</table>
1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:

   a. (means)
      - [ ] on paper
      - [x] in electronic form

   b. (time)
      - [ ] in the international application as filed
      - [ ] together with the international application in electronic form
      - [x] subsequently to this Authority for the purposes of search

2. In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:
This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ✔ Claim Nos.: 27-32  
   because they relate to subject matter not required to be searched by this Authority, namely:
   Although claims 27-32 of the present application are directed to a method of medical treatment, which this Authority is not obliged to search under Rule 39.1(iv) of the PCT, the search has been carried out on the basis of the use of the alleged inventive compounds specified therein in the treatment of mucopolysaccharidosis Ilia (MPS-IIIa).

2. ❌ Claim Nos.:  
   because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ❌ Claim Nos.:  
   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

This International Searching Authority found multiple inventions in this international application, as follows:

1. ❌ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ❌ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ❌ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claim Nos.:

4. ❌ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim Nos.:
<table>
<thead>
<tr>
<th>Patent Document</th>
<th>Publication Date</th>
<th>Patent Family</th>
<th>Publication Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>WO201 3078564A2</td>
<td>06 June 2013</td>
<td>WO201 3078564A3</td>
<td>06 September 2013</td>
</tr>
<tr>
<td></td>
<td>(06-06-2013)</td>
<td>CA284741 8A1</td>
<td>07 March 2013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US201 3059801A</td>
<td>07 March 2013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO201 3033602A2</td>
<td>07 March 2013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO201 3033602A3</td>
<td>08 May 2014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA280003235A1</td>
<td>06 June 2013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO201 3078562A3</td>
<td>06 September 2013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO201 3078562A3</td>
<td>06 September 2013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU2003236670A1</td>
<td>04 January 2005 (04-01-2005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU2003236670A8</td>
<td>04 January 2005 (04-01-2005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU2004253471A2</td>
<td>13 January 2005 (13-01-2005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU2004253471B2</td>
<td>13 May 2010 (13-05-2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA2525236A1</td>
<td>13 January 2005 (13-01-2005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP1 638605A2</td>
<td>29 March 2006 (29-03-2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP1 638605A4</td>
<td>12 August 2009 (12-08-2009)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP1 638605B1</td>
<td>08 January 2014 (08-01-2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP2007526227A</td>
<td>13 September 2007 (13-09-2007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP201 1207889A</td>
<td>20 October 2011 (20-10-2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US2006029609A1</td>
<td>09 February 2006 (09-02-2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US7569544B2</td>
<td>04 August 2009 (04-08-2009)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US201 0028370A1</td>
<td>04 February 2010 (04-02-2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US7829537B2</td>
<td>09 November 2010 (09-11-2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US86091 03B2</td>
<td>17 December 2013 (17-12-2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US2005026823A1</td>
<td>03 February 2005 (03-02-2005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP251 331 3A1</td>
<td>24 October 2012 (24-10-2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US201 2308544A1</td>
<td>06 December 2012 (06-12-2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO201 1073199A1</td>
<td>23 June 2011 (23-06-2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AT384736T</td>
<td>15 February 2008 (15-02-2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU2003237314B2</td>
<td>04 March 2010 (04-03-2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU20082200706A1</td>
<td>06 March 2008 (06-03-2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU2008230016A1</td>
<td>13 November 2008 (13-11-2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU201 0214643A1</td>
<td>16 September 2010 (16-09-2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA2445577A1</td>
<td>07 November 2002 (07-11-2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA2445577C</td>
<td>03 July 2012 (03-07-2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA2487815A1</td>
<td>11 December 2003 (11-12-2003)</td>
</tr>
</tbody>
</table>

CONTINUED ON NEXT PAGE
<table>
<thead>
<tr>
<th>Publication Number</th>
<th>Date</th>
<th>Patent Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE6022481 6D1</td>
<td>13 March 2008</td>
<td>(13-03-2008)</td>
</tr>
<tr>
<td>EP1 453631 6T2</td>
<td>22 January 2009</td>
<td>(22-01-2009)</td>
</tr>
<tr>
<td>EP1 453631 6A4</td>
<td>18 August 2004</td>
<td>(18-08-2004)</td>
</tr>
<tr>
<td>EP1 446007A4</td>
<td>07 December 2005</td>
<td>(07-12-2005)</td>
</tr>
<tr>
<td>EP1 974752A1</td>
<td>01 October 2008</td>
<td>(01-10-2008)</td>
</tr>
<tr>
<td>EP1 974752B1</td>
<td>26 September 2012</td>
<td>(26-09-2012)</td>
</tr>
<tr>
<td>ES2300439T3</td>
<td>16 June 2008</td>
<td>(16-06-2008)</td>
</tr>
<tr>
<td>HK1 124549A1</td>
<td>03 May 2013</td>
<td>(03-05-2013)</td>
</tr>
<tr>
<td>IL1 58623A</td>
<td>22 September 2009</td>
<td>(22-09-2009)</td>
</tr>
<tr>
<td>IL1 58623D0</td>
<td>12 May 2004</td>
<td>(12-05-2004)</td>
</tr>
<tr>
<td>IL1 61352D0</td>
<td>27 September 2004</td>
<td>(27-09-2004)</td>
</tr>
<tr>
<td>IL1 97348D0</td>
<td>24 December 2009</td>
<td>(24-12-2009)</td>
</tr>
<tr>
<td>JP2004535791A</td>
<td>02 December 2004</td>
<td>(02-12-2004)</td>
</tr>
<tr>
<td>JP4641 705B2</td>
<td>02 March 2011</td>
<td>(02-03-2011)</td>
</tr>
<tr>
<td>JP2006501 15B5</td>
<td>12 January 2006</td>
<td>(12-01-2006)</td>
</tr>
<tr>
<td>JP4828121 B2</td>
<td>30 November 2011</td>
<td>(30-11-2011)</td>
</tr>
<tr>
<td>JP2005506340A</td>
<td>03 March 2005</td>
<td>(03-03-2005)</td>
</tr>
<tr>
<td>JP2009203241A</td>
<td>10 September 2009</td>
<td>(10-09-2009)</td>
</tr>
<tr>
<td>JP201 00431 03A</td>
<td>25 February 2010</td>
<td>(25-02-2010)</td>
</tr>
<tr>
<td>US739681 B2</td>
<td>08 July 2008</td>
<td>(08-07-2008)</td>
</tr>
<tr>
<td>US2005261 805A1</td>
<td>22 December 2005</td>
<td>(22-12-2005)</td>
</tr>
<tr>
<td>US7629309B2</td>
<td>08 December 2009</td>
<td>(08-12-2009)</td>
</tr>
<tr>
<td>US7858576B2</td>
<td>28 December 2010</td>
<td>(28-12-2010)</td>
</tr>
<tr>
<td>US2008241 118A1</td>
<td>02 October 2008</td>
<td>(02-10-2008)</td>
</tr>
<tr>
<td>US201 0075906A1</td>
<td>25 March 2010</td>
<td>(25-03-2010)</td>
</tr>
<tr>
<td>US82071 14B2</td>
<td>26 June 2012</td>
<td>(26-06-2012)</td>
</tr>
<tr>
<td>US201 2141 452A1</td>
<td>07 June 2012</td>
<td>(07-06-2012)</td>
</tr>
<tr>
<td>US201 2295327A1</td>
<td>22 November 2012</td>
<td>(22-11-2012)</td>
</tr>
<tr>
<td>US2004050309A1</td>
<td>08 January 2004</td>
<td>(08-01-2004)</td>
</tr>
<tr>
<td>US2009203575A1</td>
<td>13 August 2009</td>
<td>(13-08-2009)</td>
</tr>
<tr>
<td>US201 4056867A1</td>
<td>27 February 2014</td>
<td>(27-02-2014)</td>
</tr>
<tr>
<td>WO2008751 0A2</td>
<td>07 November 2002</td>
<td>(07-11-2002)</td>
</tr>
<tr>
<td>WO2008751 0A3</td>
<td>06 May 2004</td>
<td>(06-05-2004)</td>
</tr>
<tr>
<td>WO2003291 3A9</td>
<td>29 April 2004</td>
<td>(29-04-2004)</td>
</tr>
<tr>
<td>WO2003291 3A3</td>
<td>12 August 2004</td>
<td>(12-08-2004)</td>
</tr>
</tbody>
</table>

WO20061 08052A2  12 October 2006  (12-10-2006)
WO20061 08052A3  30 November 2006  (30-11-2006)
EP1 877099B1  19 September 2012  (19-09-2012)
JP2008535842A  04 September 2008  (04-09-2008)
JP51 3781 4B2  06 February 2013  (06-02-2013)
US200622834B1  12 October 2006  (12-10-2006)
US8124073B2  28 February 2012  (28-02-2012)
US201 2141 507A1  07 June 2012  (07-06-2012)
<table>
<thead>
<tr>
<th>Application Number</th>
<th>Date</th>
<th>Filing Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>WO200063122A1</td>
<td>10 June 2010</td>
<td>(10-06-2010)</td>
</tr>
<tr>
<td>AU2009322043A1</td>
<td>07 July 2011</td>
<td>(07-07-2011)</td>
</tr>
<tr>
<td>AU2009322044A1</td>
<td>07 July 2011</td>
<td>(07-07-2011)</td>
</tr>
<tr>
<td>AU2009322045A1</td>
<td>07 July 2011</td>
<td>(07-07-2011)</td>
</tr>
<tr>
<td>CA2745499A1</td>
<td>10 June 2010</td>
<td>(10-06-2010)</td>
</tr>
<tr>
<td>CA2745524A1</td>
<td>10 June 2010</td>
<td>(10-06-2010)</td>
</tr>
<tr>
<td>CA2745527A1</td>
<td>10 June 2010</td>
<td>(10-06-2010)</td>
</tr>
<tr>
<td>CN102272163A</td>
<td>07 December 2011</td>
<td>(07-12-2011)</td>
</tr>
<tr>
<td>CN102307904A</td>
<td>04 January 2012</td>
<td>(04-01-2012)</td>
</tr>
<tr>
<td>CN102348723A</td>
<td>08 February 2012</td>
<td>(08-02-2012)</td>
</tr>
<tr>
<td>CN103665170A</td>
<td>26 March 2014</td>
<td>(26-03-2014)</td>
</tr>
<tr>
<td>EP2370471A1</td>
<td>05 October 2011</td>
<td>(05-10-2011)</td>
</tr>
<tr>
<td>EP2370471A4</td>
<td>01 May 2013</td>
<td>(01-05-2013)</td>
</tr>
<tr>
<td>EP2370472A1</td>
<td>05 October 2011</td>
<td>(05-10-2011)</td>
</tr>
<tr>
<td>EP2370472A4</td>
<td>24 April 2013</td>
<td>(24-04-2013)</td>
</tr>
<tr>
<td>JP2012510796A</td>
<td>17 May 2012</td>
<td>(17-05-2012)</td>
</tr>
<tr>
<td>JP2012510797A</td>
<td>17 May 2012</td>
<td>(17-05-2012)</td>
</tr>
<tr>
<td>MX201005963A</td>
<td>01 September 2011</td>
<td>(01-09-2011)</td>
</tr>
<tr>
<td>MX201005964A</td>
<td>01 September 2011</td>
<td>(01-09-2011)</td>
</tr>
<tr>
<td>MX201005965A</td>
<td>01 September 2011</td>
<td>(01-09-2011)</td>
</tr>
<tr>
<td>RU2011125366A</td>
<td>10 January 2013</td>
<td>(10-01-2013)</td>
</tr>
<tr>
<td>RU20111125367A</td>
<td>10 January 2013</td>
<td>(10-01-2013)</td>
</tr>
<tr>
<td>US2010256055A1</td>
<td>07 October 2010</td>
<td>(07-10-2010)</td>
</tr>
<tr>
<td>US2011288009A1</td>
<td>24 November 2011</td>
<td>(24-11-2011)</td>
</tr>
<tr>
<td>US2011288011A1</td>
<td>24 November 2011</td>
<td>(24-11-2011)</td>
</tr>
<tr>
<td>WO200063123A1</td>
<td>10 June 2010</td>
<td>(10-06-2010)</td>
</tr>
<tr>
<td>WO200063124A1</td>
<td>10 June 2010</td>
<td>(10-06-2010)</td>
</tr>
</tbody>
</table>