Title: PMSTI MUTANTS FOR CHEMOENZYMATIC SYNTHESIS OF SIALYL LEWIS X COMPOUNDS

Abstract: The present invention provides mutants of PMSTI for the preparation of sialyl-Lewis" oligosaccharides, and other sialosides with decreased sialidase activity.
PmST1 MUTANTS FOR CHEMOENZYMATIC SYNTHESIS OF SIALYL LEWIS X COMPOUNDS

CROSS-REFERENCES TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Application No. 61/515,702, filed August 5, 2011, which is incorporated in its entirety herein for all purposes.

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

[0002] This invention was made with Government support under Grant Nos. RO1 GM076360 awarded by National Institutes of Health. The Government has certain rights in this invention.

BACKGROUND OF THE INVENTION

[0003] Glycosyltransferase-catalyzed reactions have gained increasing attention and application for the synthesis of complex carbohydrates and glycoconjugates. Most mammalian glycosyltransferases suffer from no or low expression in E. coli systems and more restricted substrate specificity. In comparison, bacterial glycosyltransferases are generally easier to access using E. coli expression systems and have more promiscuous substrate flexibility. Nevertheless, despite the discovery of many bacterial glycosyltransferases which have promiscuities for both donor and acceptor substrates, the application of glycosyltransferases in the synthesis of carbohydrate-containing structures is limited by the availability and the substrate specificity of wild-type enzymes.

[0004] For example, sialyltransferases, the key enzymes that catalyze the transfer of a sialic acid residue from cytidine 5'-monophosphate-sialic acid (CMP-sialic acid) to an acceptor, have been commonly used for the synthesis of sialic acid-containing structures. Sialyl Lewis* [SLe*, Siaa2-3Gaipi -4(Fuca 1-3)GlcNAcpOR] is an important carbohydrate epitope involved in inflammation as well as adhesion and metastasis of cancer cells. It is a well-known tumor-
associated carbohydrate antigen and has been used as a candidate for cancer vaccine. The biosynthesis of SLe\(^x\) involves the formation of Siaoc2-3Gaip-4GlcNAcOR catalyzed by an oc2-3-sialyltransferase followed by an al-3-fucosyltransferase-catalyzed fucosylation. This biosynthetic sequence usually cannot be altered as common cc2-3-sialyltransferases do not use fucose-containing Lewis\(^x\) \([\text{Le}^x, \text{Gaip}^x(\text{Fucal}-3)\text{GlcNAcOR}]\) as a substrate.

[0005] As common terminal monosaccharides, sialic acids constitute a family of great structural diversity. So far, more than 50 structurally distinct sialic acid forms have been identified in nature. To obtain SLe\(^x\) with different sialic acid forms to elucidate the biological significance of naturally occurring sialic acid modifications, an efficient enzymatic approach is to use \(\text{Le}^x\) \([\text{Gaib}1-4(\text{Fucal}-3)\text{GlcNACbOR}]\) as a fucose-containing acceptor to add different sialic acid forms by a suitable a2-3-sialyltransferase. This process of introducing different forms of sialic acid onto the common fucosylated acceptor \(\text{Le}^x\) in the last step has significant advantages compared to the normal SLe\(^x\) biosynthetic pathway in which fucosylation is the last glycosylation process. It not only simplifies the synthetic scheme as a less number of reactions are needed, but also makes the purification process much easier as negatively charged SLe\(^x\) product is separated from neutral \(\text{Le}^x\) oligosaccharide instead of separating both negatively charged oligosaccharides SLe\(^x\) and non-fucosylated sialosides if fucosylation occurs in the last step.

[0006] We and others have demonstrated that a myxoma virus a2-3-sialyltransferase can use \(\text{Le}^x\) as an acceptor substrate for synthesizing SLe\(^x\). Nevertheless, the low expression level of the enzyme in \(\text{E. coli}\) (<0.1 mg L\(^{-1}\) culture) limits its application in preparative and large-scale synthesis of SLe\(^x\).

[0007] We have previously shown that a multifunctional a2-3-sialyltransferase from \textit{Pastenrella multocida} (PmSTI) has a good expression level in \(\text{E. coli}\) (100 mg L\(^{-1}\) culture) \((J. \text{Am. Chem. Soc. 2005, 127, 17618-17619.})\). It can use \(\text{Le}^x\) as an acceptor for the synthesis of SLe\(^x\) but the yields are poor (<20%) in spite of different conditions tested. What is needed, therefore, are a2-3-sialyltransferases having good a2-3-sialyltransferase activity with good expression levels, and lowered oc2-3-sialidase or donor substrate hydrolysis activity. Surprisingly, the present invention meets this and other needs.
BRIEF SUMMARY OF THE INVENTION

[0008] In some embodiments, the present invention provides an isolated glycosyltransferase, wherein the amino acid of the glycosyltransferase corresponding to position 120 of SEQ ID NO:1 is any amino acid other than M, the amino acid the glycosyltransferase corresponding to position 247 of SEQ ID NO:1 is any amino acid other than E, or the amino acid the glycosyltransferase corresponding to position 289 of SEQ ID NO:1 is any amino acid other than R. The glycosyltransferase of the present invention has decreased α2-3 sialidase or donor substrate hydrolysis activity compared to a control glycosyltransferase, wherein the amino acid of the control glycosyltransferase corresponding to position 120 of SEQ ID NO: 1 is M, the amino acid of the control glycosyltransferase corresponding to position 247 of SEQ ID NO:1 is E, and the amino acid of the control glycosyltransferase corresponding to position 289 of SEQ ID NO:1 is R. Finally, the glycosyltransferase of the present invention can be a member of the glycosyltransferase family 80 (GT80).

[0009] In some embodiments, the present invention provides a recombinant nucleic acid encoding an isolated glycosyltransferase of the present invention.

[0010] In some embodiments, the present invention provides a cell including a recombinant nucleic acid of the present invention.

[0011] In some embodiments, the present invention provides a method of preparing an oligosaccharide, the method including forming a reaction mixture including an acceptor sugar, a donor substrate of a sugar moiety and a nucleotide, and the glycosyltransferase of the present invention, under conditions sufficient to transfer the sugar moiety from the donor substrate to the acceptor sugar, thereby forming the oligosaccharide.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] Figure 1 shows the ternary crystal structure of PmST1 (PDB ID: 2IHZ) with bound CMP-3F(axial)-Neu5Ac and lactose (Figure 1A) and the structure of the modeled PmST1 double mutant E271F/R313Y of PmST1 wild type sequence SEQ ID NO: 13 (Figure 1B). The mutation sites are underlined. The mutant structure was obtained from automated homology modeling using Swiss-Model.
Figure 2 shows acceptor substrate specificity data for the a2-3-sialyltransferase activity of wild-type PmSTl (white columns) and its double mutant E271F/R313Y of PmSTl wild type sequence SEQ ID NO: 13 (black columns).

Figure 3 shows thermal stability data for the a2-3-sialyltransferase activity of wild-type PmSTl (white columns) and its double mutant E271F/R313Y of PmSTl wild type sequence SEQ ID NO: 13 (black columns).

Figure 4 shows HPLC-based time course studies of PmSTl-catalyzed a2-3-sialylation of Lewis* trisaccharide (1 mM) with periodical addition of sialyltransferase donor CMP-Neu5Ac (indicated by arrows). Numbers in parentheses represent the % consumption of CMP-Neu5Ac by capillary electrophoresis (CE) assays.

Figure 5 illustrates that water (in the donor hydrolysis reaction) competes with Lewis" (in PmSTl-catalyzed a2-3-sialylation reaction) for the consumption of CMP-Neu5Ac.

Figure 6 shows the SDS-PAGE analysis of the M144D mutant of PmSTl wild type sequence SEQ ID NO: 13. Lane 1: Protein marker; Lane 2: Whole cells before induction; Lane 3: Whole cells after induction; Lane 4: Cell lysate; Lane 5: Purified fraction.

Figure 7 shows the structural comparison between wild-type (WT) PmSTl and M144D mutant of PmSTl wild type sequence SEQ ID NO: 13 with bound CMP. Figure 7A shows the overall structure alignment of WT PmSTl and the PmSTl M144D mutant, both with CMP bound. Figure 7B shows the stereo view of the superposition near the active site for WT PmSTl and the M144D mutant with bound CMP-3F(o)-Neu5Ac (a donor substrate analog) and lactose acceptor. Figure 7C shows the active site of the ternary crystal structure of PmSTl (PDB ID: 2HZ) with bound CMP-3F(ax/a/-)-Neu5Ac and lactose.

Figure 8 shows $^{15}$N-$$^1$$H HSQC NMR spectra of $^{15}$N-labeled PmSTl (WT versus M144D mutant of PmSTl wild type sequence SEQ ID NO: 13; as well as apo versus CMP-bound).

Figure 9 shows the one-pot three-enzyme synthesis of sialyl Le$^x$$^3$$^b$Pro$^N_3$ (SLe$^x$$^3$$^b$Pro$^N_3$) containing different forms of sialic acids from Le$^x$$^3$$^b$Pro$^N_3$. Aldolase refers to Pasteurella.
multocida sialic acid aldolase, and NmCSS refers to Neisseria meningitidis CMP-sialic acid synthetase.

[0021] Figure 10 shows amino acid sequences alignment of GT80 sialyltransferases. Pp_Pst3-1 (GenBank accession number BAF63530), PspJPst3-2 (GenBank accession number BAF92025), Vsp_2,3ST (GenBank accession number BAF91160), P1ST6_JT-1 (GenBank accession number BAF91416), P1ST6_JT-2, (GenBank accession number BAI49484), Pd2,6ST (GenBank accession number BAA2516), Psp_pst6-1 (GenBank accession number BAF92026), Pm0188Ph (GenBank accession number DQ087233), and Hdo053P (GenBank accession number AAP95068).

DETAILED DESCRIPTION OF THE INVENTION

I. General

[0022] The present invention provides alpha2-3 sialyltransferase mutants of PmST1 having reduced alpha2-3 sialidase or donor substrate hydrolysis, useful for the preparation of oligosaccharides, and can be tolerant of fucosylated oligosaccharides. The mutations described herein can be incorporated into a variety of sialyltransferases to produce mutants having reduced sialidase or donor substrate activity.

II. Definitions

[0023] As used herein, the term "glycosyltransferase" refers to a polypeptide that catalyzes the formation of a glycoside or an oligosaccharide from a donor substrate and an acceptor or acceptor sugar. In general, a glycosyltransferase catalyzes the transfer of the monosaccharide moiety of the donor substrate to a hydroxy 1 group of the acceptor. The covalent linkage between the monosaccharide and the acceptor sugar can be a 1-4 linkage, a 1-3 linkage, a 1-6-linkage, a 1-2 linkage, a 2-3-linkage, a 2-6-linkage, a 2-8-linkage, or a 2-9-linkage. The linkage may be in the α- or β-configuration with respect to the anomeric carbon of the monosaccharide. Other types of linkages may be formed by the glycosyltransferases in the methods of the invention. Glycosyltransferases include, but are not limited to, sialyltransferases, heparosan synthases (HSs), glucosaminyltransferases, N-acetylglucosaminyltransferases, glucosyltransferases, glucuronyltransferases, N-acetylgalactosaminyltransferases, galactosyltransferases,
galacturonyltransferases, fucosyltransferases, mannosyltransferases, xylosyltransferases.
Sialyltransferases are enzymes ... 80 (GT80 using CAZy nomenclature), and includes beta-galactoside alpha-2,3-sialyltransferases that catalyze the following conversion: CMP-sialic acid + β-D-galactosyl-R = CMP + a-sialic acid-(2→3)-P-D-galactosyl-R, where the acceptor is GaipOR, where R is H, a monosaccharide, an oligosaccharide, a polysaccharide, a glycopeptide, a glycoprotein, or a glycolipid. GT80 family sialyltransferases also include galactoside or N-acetylgalactosaminide alpha-2,6-sialyltransferases that catalyze the following conversion: CMP-sialic acid + galactosyl/GalNAc-R = CMP + a-sialic acid-(2→3)^β-D-galactosyl/GalNAc-R, where the acceptor is GalOR or GalNAcOR, where R is H, serine or threonine on a peptide or protein, a monosaccharide, an oligosaccharide, a polysaccharide, a glycopeptide, a glycoprotein, or a glycolipid.

[0024] "Alpha2-3-sialidase" refers to an enzyme that catalyzes the hydrolysis of alpha2-3-glycosidic linkages of terminal sialic acids on oligosaccharides.

[0025] "Donor substrate hydrolysis" refers to hydrolysis of the nucleotide-sugar bond of the donor substrate.

[0026] An "amino acid" refers to any monomer unit that can be incorporated into a peptide, polypeptide, or protein. As used herein, the term "amino acid" includes the following twenty natural or genetically encoded alpha-amino acids: alanine (Ala or A), arginine (Arg or R), asparagine (Asn or N), aspartic acid (Asp or D), cysteine (Cys or C), glutamine (Gin or Q), glutamic acid (Glu or E), glycine (Gly or G), histidine (His or H), isoleucine (Ile or I), leucine (Leu or L), lysine (Lys or K), methionine (Met or M), phenylalanine (Phe or F), proline (Pro or P), serine (Ser or S), threonine (Thr or T), tryptophan (Trp or W), tyrosine (Tyr or Y), and valine (Val or V). In cases where "X" residues are undefined, these should be defined as "any amino acid." The structures of these twenty natural amino acids are shown in, e.g., Stryer et al., Biochemistry, 5th ed., Freeman and Company (2002), which is incorporated by reference.

Additional amino acids, such as selenocysteine and pyrolysine, can also be genetically coded for (Stadtman 1996) "Selenocysteine," Annu Rev Biochem, 65:83-100 and Ibba et al. (2002) "Genetic code: introducing pyrolysine," Curr Biol. 12(13):R464-R466, which are both
incorporated by reference). The term "amino acid" also includes unnatural amino acids, modified amino acids (e.g., having modified side chains and/or backbones), and amino acid analogs.

[0027] "Polypeptide," "peptide," and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. All three terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymers. As used herein, the terms encompass amino acid chains of any length, including full-length proteins, wherein the amino acid residues are linked by covalent peptide bonds.

[0028] The term "mutant," in the context of glycosyltransferases of the present invention, means a polypeptide, typically recombinant, that comprises one or more amino acid substitutions relative to a corresponding, naturally-occurring or unmodified glycosyltransferase, such as an alpha2-3 sialyltransferase.

[0029] In the context of glycosyltransferases, "corresponding to" another sequence (e.g., regions, fragments, nucleotide or amino acid positions, or the like) is based on the convention of numbering according to nucleotide or amino acid position number and then aligning the sequences in a manner that maximizes the percentage of sequence identity. Because not all positions within a given "corresponding region" need be identical, non-matching positions within a corresponding region may be regarded as "corresponding positions." Accordingly, as used herein, referral to an "amino acid of the glycosyltransferase corresponding to position [X]" of a specified glycosyltransferase refers to equivalent positions, based on alignment, in other glycosyltransferases and structural homologues and families. In some embodiments of the present invention, "correspondence" of amino acid positions are determined with respect to a region of the glycosyltransferase comprising one or more motifs of SEQ ID NO: 1, 13, 15, 17, 19, 21, 23, 25, 27 or 29. When a glycosyltransferase polypeptide sequence differs from SEQ ID NO: 1, 13, 15, 17, 19, 21, 23, 25, 27 or 29 (e.g., by changes in amino acids or addition or deletion of amino acids), it may be that a particular mutation associated with improved activity as discussed herein will not be in the same position number as it is in SEQ ID NO:1, 13, 15, 17, 19, 21, 23, 25, 27 or 29.
A nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation.

As used herein, "percent sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the sequence in the comparison window can comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

The terms "identical" or "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same. Sequences are "substantially identical" to each other if they have a specified percentage of nucleotides or amino acid residues that are the same (e.g., at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% identity over a specified region), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. These definitions also refer to the complement of a test sequence. Optionally, the identity exists over a region that is at least about 50 nucleotides in length, or more typically over a region that is 100 to 500 or 1000 or more nucleotides in length.

The terms "similarity" or "percent similarity," in the context of two or more polypeptide sequences, refer to two or more sequences or subsequences that have a specified percentage of amino acid residues that are either the same or similar as defined by a conservative amino acid substitutions (e.g., 60% similarity, optionally 65%, 70%, 75%, 80%, 85%, 90%, or 95% similar over a specified region), when compared and aligned for maximum correspondence.
over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Sequences are "substantially similar" to each other if they are at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, or at least 55% similar to each other. Optionally, this similarly exists over a region that is at least about 50 amino acids in length, or more typically over a region that is at least about 100 to 500 or 1000 or more amino acids in length.

[0034] For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters are commonly used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities or similarities for the test sequences relative to the reference sequence, based on the program parameters.

[0035] A "comparison window," as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well known in the art. Optimal alignment of sequences for comparison can be conducted, for example, by the local homology algorithm of Smith and Waterman (Adv. Appl. Math. 2:482, 1970), by the homology alignment algorithm of Needleman and Wunsch (J. Mol. Biol. 48:443, 1970), by the search for similarity method of Pearson and Lipman (Proc. Natl. Acad. Sci. USA 85:2444, 1988), by computerized implementations of these algorithms (e.g., GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), or by manual alignment and visual inspection (see, e.g., Ausubel et al., Current Protocols in Molecular Biology (1995 supplement)).

[0036] Algorithms suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (Nuc. Acids Res. 25:3389-402, 1977), and Altschul et al. (J. Mol. Biol. 215:403-10, 1990), respectively. Software for performing BLAST analyses is publicly available through the National Center for
Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. \( T \) is referred to as the neighborhood word score threshold (Altschul et al, supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters \( M \) (reward score for a pair of matching residues; always >0) and \( N \) (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity \( X \) from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters \( W, T, \) and \( X \) determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength \( (W) \) of 11, an expectation \( (E) \) or 10, \( M=-5, N=-4 \) and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation \( (E) \) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff, Proc. Natl. Acad. Sci. USA 89:10915, 1989) alignments \( (B) \) of 50, expectation \( (E) \) of 10, \( M=5, N=-4 \), and a comparison of both strands.

[0037] The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin and Altschul, Proc. Natl. Acad. Sci. USA 90:5873-87, 1993). One measure of similarity provided by the BLAST algorithm is the smallest sum probability \( (P(N)) \), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, typically less than about 0.01, and more typically less than about 0.001.

[0038] "Recombinant," as used herein, refers to an amino acid sequence or a nucleotide sequence that has been intentionally modified by recombinant methods. By the term
"recombinant nucleic acid" herein is meant a nucleic acid, originally formed in vitro, in general, by the manipulation of a nucleic acid by endonucleases, in a form not normally found in nature. Thus an isolated, mutant glycosyltransferase nucleic acid, in a linear form, or an expression vector formed in vitro by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell, it will replicate non-recombinantly, i.e., using the in vivo cellular machinery of the host cell rather than in vitro manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered recombinant for the purposes of the invention. A "recombinant protein" is a protein made using recombinant techniques, i.e., through the expression of a recombinant nucleic acid as depicted above.

[0039] The term "vector" refers to a piece of DNA, typically double-stranded, which may have inserted into it a piece of foreign DNA. The vector may be, for example, of plasmid origin. Vectors contain "replicon" polynucleotide sequences that facilitate the autonomous replication of the vector in a host cell. Foreign DNA is defined as heterologous DNA, which is DNA not naturally found in the host cell, which, for example, replicates the vector molecule, encodes a selectable or screenable marker, or encodes a transgene. The vector is used to transport the foreign or heterologous DNA into a suitable host cell. Once in the host cell, the vector can replicate independently of or coincidental with the chromosomal DNA, and several copies of the vector and its inserted DNA can be generated. In addition, the vector can also contain the necessary elements that permit transcription of the inserted DNA into an mRNA molecule or otherwise cause replication of the inserted DNA into multiple copies of RNA. Some expression vectors additionally contain sequence elements adjacent to the inserted DNA that increase the half-life of the expressed mRNA and/or allow translation of the mRNA into a protein molecule. Many molecules of mRNA and polypeptide encoded by the inserted DNA can thus be rapidly synthesized.

[0040] The term "nucleotide," in addition to referring to the naturally occurring ribonucleotide or deoxyribonucleotide monomers, shall herein be understood to refer to related structural variants thereof, including derivatives and analogs, that are functionally equivalent with respect
to the particular context in which the nucleotide is being used (e.g., hybridization to a complementary base), unless the context clearly indicates otherwise.

[0041] The term "nucleic acid" or "polynucleotide" refers to a polymer that can be corresponded to a ribose nucleic acid (RNA) or deoxyribose nucleic acid (DNA) polymer, or an analog thereof. This includes polymers of nucleotides such as RNA and DNA, as well as synthetic forms, modified (e.g., chemically or biochemically modified) forms thereof, and mixed polymers (e.g., including both RNA and DNA subunits). Exemplary modifications include methylation, substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, and the like), pendent moieties (e.g., polypeptides), intercalators (e.g., acridine, psoralen, and the like), chelators, alkylators, and modified linkages (e.g., alpha anomic nucleic acids and the like). Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Typically, the nucleotide monomers are linked via phosphodiester bonds, although synthetic forms of nucleic acids can comprise other linkages (e.g., peptide nucleic acids as described in Nielsen et al. (Science 254:1497-1500, 1991). A nucleic acid can be or can include, e.g., a chromosome or chromosomal segment, a vector (e.g., an expression vector), an expression cassette, a naked DNA or RNA polymer, the product of a polymerase chain reaction (PCR), an oligonucleotide, a probe, and a primer. A nucleic acid can be, e.g., single-stranded, double-stranded, or triple-stranded and is not limited to any particular length. Unless otherwise indicated, a particular nucleic acid sequence comprises or encodes complementary sequences, in addition to any sequence explicitly indicated.

[0042] As used herein, the term "oligosaccharide" refers to a compound containing at least two sugars covalently linked together. Oligosaccharides include disaccharides, trisaccharides, tetrasaccharides, pentasaccharides, hexasaccharides, heptasaccharides, octasaccharides, and the like. Covalent linkages generally consist of glycosidic linkages (i.e., C-O-C bonds) formed from the hydroxyl groups of adjacent sugars. Linkages can occur between the 1-carbon and the 4-carbon of adjacent sugars (i.e., a 1-4 linkage), the 1-carbon and the 3-carbon of adjacent sugars (i.e., a 1-3 linkage), the 1-carbon and the 6-carbon of adjacent sugars (i.e., a 1-6 linkage), or the 1-carbon and the 2-carbon of adjacent sugars (i.e., a 1-2 linkage). A sugar can be linked within
an oligosaccharide such that the anomeric carbon is in the α- or β-configuration. The oligosaccharides prepared according to the methods of the invention can also include linkages between carbon atoms other than the 1-, 2-, 3-, 4-, and 6-carbons.

[0043] "Acceptor sugar" refers a sugar that accepts the sugar being added. For example, the acceptor sugar can be an oligosaccharide, such as a fucosylated oligosaccharide, that accepts a sialic acid or analog thereof.

[0044] "Donor substrate" refers to a compound having a nucleotide and the sugar that is added to the acceptor, where the sugar and nucleotide are covalently bound together. The sugar can be sialic acid or analogs thereof. The nucleotide can be any suitable nucleotide such as cytidine monophosphate (CMP).

[0045] "Sialic acid aldolase" refers to an aldolase that prepares sialic acid using pyruvate and N-acetyl mannose (ManNAc).

III. Glycosyltransferases

[0046] The present invention includes a variety of sialyltransferases with reduced sialidase and/or donor substrate hydrolysis activity. Sialyltransferases are one class of glycosyltransferases, enzymes that catalyze the transfer of a sugar from a nucleotide-sugar complex (donor substrate) to an acceptor, a mono, di or oligosaccharide. Sialyltransferases catalyze the transfer of N-acetylneuraminic acid, and analogs thereof, from a sialic acid-nucleotide complex, the donor substrate, to the terminal sugar of the acceptor which can be a monosaccharide, an oligosaccharide, a glycolipid, a glycopeptide, or a glycoprotein. Representative sialyltransferases include, but are not limited to, sialyltransferases in family EC 2.4.99, such as beta-galactosamide alpha-2,6-sialyltransferase (EC 2.4.99.1), alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase (EC 2.4.99.3), beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4), N-acetyllactosaminide alpha-2,3-sialyltransferase (EC 2.4.99.6), alpha-N-acetyl-neuraminidase alpha-2,8-sialyltransferase (EC 2.4.99.8); lactosyleceramide alpha-2,3-sialyltransferase (EC 2.4.99.9). The sialyltransferases of the present invention also include those of the CAZy GT80 family, or EC 2.4.99.4, drawn to alpha2-3 and alpha2-6 sialyltransferases, as well as sialyltransferases in the GT29, GT30, GT38, GT42, GT52, and GT73 families. Representative GT80 sialyltransferases include, but are not limited to, PmSTl,

[0047] The glycosyltransferases of the present invention include those having decreased a2-3 sialidase or donor substrate hydrolysis activity compared to a control glycosyltransferase. a2-3 sialidase activity refers to the back reaction starting from the product oligosaccharide, cleaving the glycosidic bond between the sugar from the donor substrate and the acceptor, resulting in the donor substrate and the acceptor.

[0048] In some embodiments, the glycosyltransferase can be an a2-3-sialyltransferase. The a2-3-sialyltransferases of the present invention can include sialyltransferases of Pasteurella multocida. In some embodiments, the glycosyltransferases of the present invention can have a motif in the sialyltransferase domain including at least one of sialyltransferase motif A (YDDGS, corresponding to positions 139-143 of PmST1 wild type, SEQ ID NO: 13) and sialyltransferase motif B (KGH, corresponding to positions 309-311 of PmST1 wild type, SEQ ID NO: 13).

[0049] The glycosyltransferases of the present invention can include a polypeptide having any suitable percent identity to the control sequence. For example, the glycosyltransferases of the present invention can include a polypeptide having a percent sequence identity to the control glycosyltransferase sequence of at least 20, 30, 40, 50, 55, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98 or at least 99%. In some embodiments, percent sequence identity can be at least 80%. In some embodiments, percent sequence identity can be at least 90%. In some embodiments, percent sequence identity can be at least 95%. In some embodiments, the glycosyltransferase includes a polypeptide sequence having at least 80% sequence identity to SEQ ID NO: 1.

[0050] In some embodiments, the isolated glycosyltransferase includes a polypeptide sequence of SEQ ID NO: 3 (M120D), SEQ ID NO: 5 (M120H), SEQ ID NO: 7 (E247F), SEQ ID NO: 9 (R289Y) or SEQ ID NO: 11 (E247F/R289Y).

[0051] The precise length of glycosyltransferases can vary, so the precise amino acid positions corresponding to each mutation can vary depending on the particular control glycosyltransferase used. Amino acid and nucleic acid sequence alignment programs are readily available (see, e.g.,
those referred to *supra*) and, given the particular motifs identified herein, serve to assist in the identification of the exact amino acids (and corresponding codons) for modification in accordance with the present invention. The positions of several mutations are shown in the table below for the PmST1 wild type sequence (SEQ ID NO: 13) and the A24PmST1 (SEQ ID NO: 1) sequence.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>PmST1 wild type (SEQ ID NO: 13)</th>
<th>A24PmST1 (SEQ ID NO: 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M144D</td>
<td>M120D</td>
</tr>
<tr>
<td>2</td>
<td>M144H</td>
<td>M120H</td>
</tr>
<tr>
<td>3</td>
<td>E271F</td>
<td>E247F</td>
</tr>
<tr>
<td>4</td>
<td>R313Y</td>
<td>R289Y</td>
</tr>
<tr>
<td>5</td>
<td>E271F/R313Y</td>
<td>E247F/R289Y</td>
</tr>
</tbody>
</table>

The above table illustrates "correspondence" of an amino acid position to a different sequence. For example, amino acid position 144 in the PmST1 wild type sequence (SEQ ID NO: 13) corresponds to position 120 of the A24PmST1 sequence (SEQ ID NO: 1).

The control glycosyltransferases of the present invention includes any suitable glycosyltransferase or sialyltransferase. The glycosyltransferases of the present invention includes mutants corresponding to any position of PmST1 wild type sequence (SEQ ID NO: 13) and A24PmST1 (SEQ ID NO: 1) (see *Biochemistry* 2006, 45(7), 2139, and 2007, 46(21), 6288). For example, the glycosyltransferases of the present invention include, but are not limited to, mutants at least one of positions 120, 247 and 289 of A24PmST1 (SEQ ID NO: 1). Other glycosyltransferases include mutants at least one of positions 144, 271 and 313 of PmST1 wild type sequence (SEQ ID NO: 13). The mutants can include any suitable amino acid other than the native amino acid. For example, the amino acid can be V, I, L, M, F, W, P, S, T, A, G, C, Y, N, Q, D, E, K, R, or H. In some embodiments, the control glycosyltransferase can be the PmST1 wild type sequence (SEQ ID NO: 13) or the A24PmST1 (SEQ ID NO: 1). In some embodiments, the control glycosyltransferase can be A24PmST1 (SEQ ID NO: 1).
NO:1 is any amino acid other than M, the amino acid the glycosyltransferase corresponding to position 247 of SEQ ID NO:1 is any amino acid other than E, or the amino acid the glycosyltransferase corresponding to position 289 of SEQ ID NO:1 is any amino acid other than R. The glycosyltransferase of the present invention has decreased α2-3 sialidase or donor substrate hydrolysis activity compared to a control glycosyltransferase, wherein the amino acid of the control glycosyltransferase corresponding to position 120 of SEQ ID NO:1 is M, the amino acid of the control glycosyltransferase corresponding to position 247 of SEQ ID NO:1 is E, and the amino acid of the control glycosyltransferase corresponding to position 289 of SEQ ID NO:1 is R. Finally, the glycosyltransferase of the present invention can be a member of the glycosyltransferase family 80 (GT80).

[0055] In some embodiments, the isolated glycosyltransferase has decreased α2-3 sialidase activity, and includes at least one of the amino acid corresponding to position 247 of SEQ ID NO:1 is any amino acid other than E, and the amino acid corresponding to position 289 of SEQ ID NO:1 is any amino acid other than R. Decreased α2-3 sialidase activity can be measured by the ratio of α2-3 sialidase activity for the control glycosyltransferase to the α2-3 sialidase activity of the isolated glycosyltransferase. The ratio can be at least 2:1, 3:1, 4:1, 5:1, 10:1, 20:1, 30:1, 40:1, 50:1, 100:1, 200:1, 300:1, 400:1, 500:1 or at least 1000:1. In some embodiments, the ratio is at least 5:1. In some embodiments, the ratio is at least 10:1. In some embodiments, the ratio is at least 100:1. In some embodiments, the ratio is at least 1000:1.

[0056] In some embodiments, the isolated glycosyltransferase having decreased α2-3 sialidase activity includes the amino acid corresponding to position 247 of SEQ ID NO:1 is any amino acid other than E, and the amino acid corresponding to position 289 of SEQ ID NO:1 is any amino acid other than R.

[0057] In some embodiments, the isolated glycosyltransferase having decreased α2-3 sialidase activity includes the amino acid corresponding to position 117 of SEQ ID NO:1 is D or E. In some embodiments, the isolated glycosyltransferase having decreased α2-3 sialidase activity includes the amino acid corresponding to position 117 of SEQ ID NO:1 is A, G, V, L or I. In some embodiments, the isolated glycosyltransferase having decreased α2-3 sialidase activity includes the amino acid corresponding to position 287 of SEQ ID NO:1 is H, K, R, W or F.
Other glycosyltransferases of the present invention have decreased donor substrate hydrolysis activity. Decreased donor substrate hydrolysis activity can be measured by the ratio of donor substrate hydrolysis activity for the control glycosyltransferase to the donor substrate hydrolysis activity of the isolated glycosyltransferase. The ratio can be at least 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1 or at least 10:1. In some embodiments, the isolated glycosyltransferase has decreased donor substrate hydrolysis activity, wherein the amino acid corresponding to position 289 of SEQ ID NO:1 can be any amino acid other than M. In some embodiments, the ratio of donor substrate hydrolysis activity for the control α2-3 sialidase to the donor substrate hydrolysis activity of the isolated glycosyltransferase is at least 2:1.

In some embodiments, the amino acid corresponding to position 120 of SEQ ID NO: 1 can be any amino acid of V, I, L, F, W, P, S, T, A, G, C, Y, N, Q, D, E, K, R, or H. In some embodiments, the amino acid corresponding to position 120 of SEQ ID NO: 1 can be any amino acid of D, E, H, K or R. In some embodiments, the amino acid corresponding to position 120 of SEQ ID NO: 1 can be any amino acid of D or H. In some embodiments, the amino acid corresponding to position 120 of SEQ ID NO: 1 can be amino acid D. In some embodiments, the amino acid corresponding to position 120 of SEQ ID NO: 1 can be amino acid H.

In some embodiments, the amino acid corresponding to position 247 of SEQ ID NO: 1 can be any amino acid of V, I, L, M, F, W, P, S, T, A, G, C, Y, N, Q, D, K, R, or H. In some embodiments, the amino acid corresponding to position 247 of SEQ ID NO: 1 can be any amino acid of F, Y or W. In some embodiments, the amino acid corresponding to position 247 of SEQ ID NO: 1 can be amino acid F.

In some embodiments, the amino acid corresponding to position 289 of SEQ ID NO: 1 can be any amino acid of V, I, L, M, F, W, P, S, T, A, G, C, Y, N, Q, D, E, K, and H. In some embodiments, the amino acid corresponding to position 289 of SEQ ID NO: 1 can be any amino acid of Y, F or W. In some embodiments, the amino acid corresponding to position 289 of SEQ ID NO: 1 can be amino acid Y.

The glycosyltransferases of the present invention can have one or more mutations. In some embodiments, the glycosyltransferase includes the amino acid corresponding to position 247 of SEQ ID NO: 1 can be any amino acid of F, Y or W, and the amino acid corresponding to position 289 of SEQ ID NO: 1 can be any amino acid of Y, F or W. In some embodiments, the
amino acid corresponding to position 247 of SEQ ID NO:1 can be amino acid F, and the amino acid corresponding to position 289 of SEQ ID NO:1 can be amino acid Y.

[0063] In some embodiments, the isolated glycosyltransferase can be the amino acid corresponding to position 120 of SEQ ID NO:1 is D, E, H, K or R, the amino acid corresponding to position 247 of SEQ ID NO:1 is F, Y or W, or the amino acid corresponding to position 289 of SEQ ID NO:1 is Y, F or W. In some embodiments, the isolated glycosyltransferase can be the amino acid corresponding to position 120 of SEQ ID NO:1 is D or H, the amino acid corresponding to position 247 of SEQ ID NO:1 is F, or the amino acid corresponding to position 289 of SEQ ID NO:1 is Y.

[0064] The glycosyltransferases of the present invention can be constructed by mutating the DNA sequences that encode the corresponding unmodified glycosyltransferase (e.g., a wild-type glycosyltransferase or a corresponding variant from which the glycosyltransferase of the invention is derived), such as by using techniques commonly referred to as site-directed mutagenesis. Nucleic acid molecules encoding the unmodified form of the glycosyltransferase can be mutated by a variety of techniques well-known to one of ordinary skill in the art. (See, e.g., PCR Strategies (M. A. Innis, D. H. Gelfand, and J. J. Sninsky eds., 1995, Academic Press, San Diego, CA) at Chapter 14; PCR Protocols: A Guide to Methods and Applications (M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White eds., Academic Press, NY, 1990).

[0065] By way of non-limiting example, the two primer system, utilized in the Transformer Site-Directed Mutagenesis kit from Clontech, may be employed for introducing site-directed mutants into a polynucleotide encoding an unmodified form of the glycosyltransferase. Following denaturation of the target plasmid in this system, two primers are simultaneously annealed to the plasmid; one of these primers contains the desired site-directed mutation, the other contains a mutation at another point in the plasmid resulting in elimination of a restriction site. Second strand synthesis is then carried out, tightly linking these two mutations, and the resulting plasmids are transformed into a mutS strain of E. coli. Plasmid DNA is isolated from the transformed bacteria, restricted with the relevant restriction enzyme (thereby linearizing the unmutated plasmids), and then retransformed into E. coli. This system allows for generation of mutations directly in an expression plasmid, without the necessity of subcloning or generation of single-stranded phagemids. The tight linkage of the two mutations and the subsequent
linearization of unmutated plasmids result in high mutation efficiency and allow minimal screening. Following synthesis of the initial restriction site primer, this method requires the use of only one new primer type per mutation site. Rather than prepare each positional mutant separately, a set of "designed degenerate" oligonucleotide primers can be synthesized in order to introduce all of the desired mutations at a given site simultaneously. Transformants can be screened by sequencing the plasmid DNA through the mutagenized region to identify and sort mutant clones. Each mutant DNA can then be restricted and analyzed by electrophoresis, such as for example, on a Mutation Detection Enhancement gel (Mallinckrodt Baker, Inc., Phillipsburg, NJ) to confirm that no other alterations in the sequence have occurred (by band shift comparison to the unmutagenized control). Alternatively, the entire DNA region can be sequenced to confirm that no additional mutational events have occurred outside of the targeted region.

[0066] Verified mutant duplexes in pET (or other) overexpression vectors can be employed to transform E. coli such as, e.g., strain E. coli BL21 (DE3) pLysS, for high level production of the mutant protein, and purification by standard protocols. The method of FAB-MS mapping, for example, can be employed to rapidly check the fidelity of mutant expression. This technique provides for sequencing segments throughout the whole protein and provides the necessary confidence in the sequence assignment. In a mapping experiment of this type, protein is digested with a protease (the choice will depend on the specific region to be modified since this segment is of prime interest and the remaining map should be identical to the map of unmutated protein). The set of cleavage fragments is fractionated by, for example, microbore HPLC (reversed phase or ion exchange, again depending on the specific region to be modified) to provide several peptides in each fraction, and the molecular weights of the peptides are determined by standard methods, such as FAB-MS. The determined mass of each fragment are then compared to the molecular weights of peptides expected from the digestion of the predicted sequence, and the correctness of the sequence quickly ascertained. Since this mutagenesis approach to protein modification is directed, sequencing of the altered peptide should not be necessary if the MS data agrees with prediction. If necessary to verify a changed residue, CAD-tandem MS/MS can be employed to sequence the peptides of the mixture in question, or the target peptide can be purified for subtractive Edman degradation or carboxypeptidase Y digestion depending on the location of the modification.
Recombinant Nucleic Acids

[0067] Mutant glycosyltransferases with at least one amino acid substituted can be generated in various ways. In the case of amino acids located close together in the polypeptide chain, they may be mutated simultaneously using one oligonucleotide that codes for all of the desired amino acid substitutions. If however, the amino acids are located some distance from each other (separated by more than ten amino acids, for example) it is more difficult to generate a single oligonucleotide that encodes all of the desired changes. Instead, one of two alternative methods may be employed. In the first method, a separate oligonucleotide is generated for each amino acid to be substituted. The oligonucleotides are then annealed to the single-stranded DNA simultaneously, and the second strand of DNA that is synthesized from the template will encode all of the desired amino acid substitutions. An alternative method involves two or more rounds of mutagenesis to produce the desired mutant. The first round is as described for the single mutants: DNA encoding the unmodified glycosyltransferase is used for the template, an oligonucleotide encoding the first desired amino acid substitution(s) is annealed to this template, and the heteroduplex DNA molecule is then generated. The second round of mutagenesis utilizes the mutated DNA produced in the first round of mutagenesis as the template. Thus, this template already contains one or more mutations. The oligonucleotide encoding the additional desired amino acid substitution(s) is then annealed to this template, and the resulting strand of DNA now encodes mutations from both the first and second rounds of mutagenesis. This resultant DNA can be used as a template in a third round of mutagenesis, and so on.


[0068] Accordingly, also provided are recombinant nucleic acids, optionally isolated, encoding any of the glycosyltransferases of the present invention (e.g., glycosyltransferases comprising any of SEQ ID NOs:4, 6, 8, 10 and 12). Using a nucleic acid of the present invention, encoding a glycosyltransferase of the invention, a variety of vectors can be made. Any vector containing replicon and control sequences that are derived from a species compatible with the host cell can be used in the practice of the invention. Generally, expression vectors include transcriptional and translational regulatory nucleic acid regions operably linked to the nucleic acid encoding the mutant glycosyltransferase. The term "control sequences" refers to DNA sequences necessary
for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. In addition, the vector may contain a Positive Retroregulatory Element (PRE) to enhance the half-life of the transcribed mRNA (see Gelfand et al. U.S. Patent No. 4,666,848). The transcriptional and translational regulatory nucleic acid regions will generally be appropriate to the host cell used to express the glycosyltransferase. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells. In general, the transcriptional and translational regulatory sequences may include, e.g., promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In typical embodiments, the regulatory sequences include a promoter and transcriptional start and stop sequences. Vectors also typically include a polylinker region containing several restriction sites for insertion of foreign DNA. In certain embodiments, "fusion flags" are used to facilitate purification and, if desired, subsequent removal of tag/flag sequence, e.g., "His-Tag". However, these are generally unnecessary when purifying an thermoactive and/or thermostable protein from a mesophilic host (e.g., E. coli) where a "heat-step" may be employed. The construction of suitable vectors containing DNA encoding replication sequences, regulatory sequences, phenotypic selection genes, and the mutant glycosyltransferase of interest are prepared using standard recombinant DNA procedures. Isolated plasmids, viral vectors, and DNA fragments are cleaved, tailored, and ligated together in a specific order to generate the desired vectors, as is well-known in the art (see, e.g., Sambrook et al, Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press, New York, NY, 2nd ed. 1989)). In some embodiments, the present invention provides a recombinant nucleic acid encoding an isolated glycosyltransferase of the present invention.

**Host Cells**

[0069] In certain embodiments, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used. Suitable selection genes can include, for example, genes coding for ampicillin and/or tetracycline resistance, which enables cells transformed with these vectors to grow in the presence of these antibiotics.
[0070] In one aspect of the present invention, a nucleic acid encoding a glycosyltransferase of the invention is introduced into a cell, either alone or in combination with a vector. By "introduced into" or grammatical equivalents herein is meant that the nucleic acids enter the cells in a manner suitable for subsequent integration, amplification, and/or expression of the nucleic acid. The method of introduction is largely dictated by the targeted cell type. Exemplary methods include CaP0 4 precipitation, liposome fusion, LIPOFECT®N, electroporation, viral infection, and the like.

[0071] In some embodiments, prokaryotes are used as host cells for the initial cloning steps of the present invention. Other host cells include, but are not limited to, eukaryotic (e.g., mammalian, plant and insect cells), or prokaryotic (bacterial) cells. Exemplary host cells include, but are not limited to, _Escherichia coli, Saccharomyces cerevisiae, Pichia pastoris, Sf9_ insect cells, and CHO cells. They are particularly useful for rapid production of large amounts of DNA, for production of single-stranded DNA templates used for site-directed mutagenesis, for screening many mutants simultaneously, and for DNA sequencing of the mutants generated. Suitable prokaryotic host cells include _E. coli_ K12 strain 94 (ATCC No. 31,446), _E. coli_ strain W3110 (ATCC No. 27,325), _E. coli_ K12 strain DG116 (ATCC No. 53,606), _E. coli_ X1776 (ATCC No. 31,537), and _E. coli_ B; however many other strains of _E. coli_, such as HB101, JM101, NM522, NM538, NM539, and many other species and genera of prokaryotes including bacilli such as _Bacillus subtilis_, other enterobacteriaceae such as _Salmonella typhimurium_ or _Serratia marcesans_, and various _Pseudomonas_ species can all be used as hosts. Prokaryotic host cells or other host cells with rigid cell walls are typically transformed using the calcium chloride method as described in section 1.82 of Sambrook _et al._, _supra_. Alternatively, electroporation can be used for transformation of these cells. Prokaryote transformation techniques are set forth in, for example Dower, in _Genetic Engineering, Principles and Methods_ 12:275-296 (Plenum Publishing Corp., 1990); Hanahan _et al._, _Meth. Enzymol._ 204:63, 1991. Plasmids typically used for transformation of _E. coli_ include pBR322, pUC18, pUC19, pUC118, pUC1 19, and Bluescript M13, all of which are described in sections 1.12-1.20 of Sambrook _et al._, _supra_. However, many other suitable vectors are available as well.

[0072] In some embodiments, the glycosyltransferases of the present invention are produced by culturing a host cell transformed with an expression vector containing a nucleic acid encoding
the glycosyltransferase, under the appropriate conditions to induce or cause expression of the glycosyltransferase. Methods of culturing transformed host cells under conditions suitable for protein expression are well-known in the art (see, e.g., Sambrook et al., supra). Suitable host cells for production of the glycosyltransferases from lambda pl promoters-containing plasmid vectors include E. coli strain DG16 (ATCC No. 53606) (see US Pat. No. 5,079,352 and Lawyer, F.C. et al, *PCR Methods and Applications* 2:275-87, 1993, which are both incorporated herein by reference). Following expression, the glycosyltransferase can be harvested and purified. Methods for purifying the thermostable glycosyltransferase are described in, for example, Lawyer et al., supra. In some embodiments, the present invention provides a cell including a recombinant nucleic acid of the present invention. In some embodiments, the cell can be prokaryotes, eukaryotes, mammalian, plant, bacteria or insect cells.

**IV. Methods of Making Oligosaccharides**

[0073] The glycosyltransferases of the present invention can be used to prepare oligosaccharides, specifically to add N-acetylneuraminic acid (Neu5Ac), other sialic acids, and analogs thereof, to a monosaccharide, an oligosaccharide, a glycolipid, a glycopeptide, or a glycoprotein. As shown in Figure 5, the glycosyltransferase PmST1, catalyzes the addition of CMP-Neu5Ac to a fucosylated oligosaccharide by transferring the Neu5Ac to the oligosaccharide.

[0074] In some embodiments, the present invention provides a method of preparing an oligosaccharide, the method including forming a reaction mixture including an acceptor sugar, a donor substrate containing a sugar moiety and a nucleotide, and the glycosyltransferase of the present invention, under conditions sufficient to transfer the sugar moiety from the donor substrate to the acceptor sugar, thereby forming the oligosaccharide.

[0075] The acceptor sugar can be any suitable oligosaccharide, glycolipid, glycopeptide, or glycoprotein. When the acceptor sugar is an oligosaccharide, any suitable oligosaccharide can be used. For example, the acceptor sugar can be Galpl-4GlcNAcPOR, wherein R can be H, a sugar or an oligosaccharide. Alternatively, the acceptor sugar can be fucosylated, such as Gaipi-4(Fucα1-3)O1cNAcβOR (LewisβOR or LeβOR) wherein R can be H, a sugar or an oligosaccharide.
[0076] The donor substrate includes a nucleotide and sugar. Any nucleotide can be used, include, but are not limited to, adenine, guanine, cytosine, uracil and thymine nucleotides with one, two or three phosphate groups. In some embodiments, the nucleotide can be cytidine monophosphate (CMP). The sugar can be any suitable sugar. When the glycosyltransferase is a sialyltransferase, the sugar can be N-acetylneuraminic acid or Neu5Ac, other sialic acids and analogs thereof. Sialic acid is a general term for N- and O-substituted derivatives of neuraminic acid, and includes, but is not limited to, N-acetyl (Neu5Ac) or N-glycolyl (Neu5Gc) substitutions, as well as Osubstitutions including acetyl, lactyl, methyl, sulfate and phosphate, among others. In some embodiments, the sialic acid can be a compound of the formula:

\[
\text{R}^1\text{R}^2\text{R}^3\text{R}^4\text{O}\text{CO}^+\text{HO}\text{OH}.
\]

wherein \(\text{R}^1\) can be H, OH, N3, NHC(0)Me, NHC(0)CH2OH, NHC(0)CH2N3, NHC(O)CH2C=CH2, NHC(0)CH2F, NHC(0)CH2NHCBz, NHC(0)CH2OC(0)Me, or NHC(0)CH2OBn; and \(\text{R}^2\), \(\text{R}^3\), and \(\text{R}^4\) can be independently selected from H, OH, N3, OMe, F, OSO3, OPO3H, or OC(0)Me. In some embodiments, the donor substrate can be CMP-Neu5Ac.

Other donor substrates are useful in the methods of the present invention. In other embodiments, the sialic acid can be a compound of the formula:

\[
\text{R}^1\text{R}^2\text{OH}\text{OH}\text{CO}^+\text{HO}\text{OH}.
\]

[0077] Any glycosyltransferase of the present invention can be used in the methods of the present invention. In some embodiments, the glycosyltransferase can include a polypeptide sequence such as SEQ ID NO: 3 (M120D), SEQ ID NO: 5 (M120H), SEQ ID NO: 7 (E247F), SEQ ID NO: 9 (R289Y) or SEQ ID NO: 11 (E247F/R289Y). In some embodiments, the glycosyltransferase can include a polypeptide sequence such as SEQ ID NO: 3 (M120D) or SEQ ID NO: 5 (M120H). In some embodiments, the glycosyltransferase can include a polypeptide sequence such as SEQ ID NO: 7 (E247F), SEQ ID NO: 9 (R289Y) or SEQ ID NO: 11 (E247F/R289Y). The glycosyltransferases can be, for example, purified, secreted by a cell present in the reaction mixture, or can catalyze the reaction within a cell expressing the glycosyltransferase.
In another aspect of the present invention, reaction mixtures are provided comprising the glycosyltransferases as described herein. The reaction mixtures can further comprise reagents for use in glycosylation techniques. For example, in certain embodiments, the reaction mixtures comprise a buffer, salts (e.g., Mn$^{2+}$, Mg$^{2+}$), and labels (e.g., fluorophores).

The donor substrate can be prepared prior to preparation of the oligosaccharide, or prepared in situ immediately prior to preparation of the oligosaccharide. In some embodiments, the method of the present invention also includes forming a reaction mixture including a CMP-sialic acid synthetase, cytidine triphosphate, and N-acetylmuramic acid (Neu5Ac) or a Neu5Ac analog, under conditions suitable to form the CMP-Neu5Ac or CMP-Neu5Ac analog. In some embodiments, the step of forming the donor substrate and the step of forming the oligosaccharide are performed in one pot.

In some embodiments, the sugar is prepared separately prior to use in the methods of the present invention. Alternatively, the sugar can be prepared in situ immediately prior to use in the methods of the present invention. In some embodiments, the method also includes forming a reaction mixture including a sialic acid aldolase, pyruvic acid or derivatives thereof, and N-acetylmannosamine or derivatives thereof, under conditions suitable to form the Neu5Ac or Neu5Ac analog. In some embodiments, the step of forming the sugar, the step of forming the donor substrate and the step of forming the oligosaccharide are performed in one pot.

The oligosaccharide prepared by the method of the present invention can be any suitable oligosaccharide, glycolipid or glycoprotein. For example, the oligosaccharide can be an a2-3-linked sialyloligosaccharide. In some embodiments, the oligosaccharide can be a fucosylated oligosaccharide. In some embodiments, the oligosaccharide can be Neu5Aca2-3Galpl-4(Fucal-3)GlcNAcpOR (Sia-Lewis$^\text{a}$OR or SLe$^\text{a}$OR) wherein R can be H, a monosaccharide, or an oligosaccharide. In some embodiments, the oligosaccharide can be Neu5Aca2-3Gaipi-4GlcNAcpOR, wherein R can H, a monosaccharide, or an oligosaccharide.
V. Examples

Example 1. Decreasing the sialidase activity of multifunctional *Pasteurella multocida* alpha2-3-sialyltransferase 1 (PmST1) by site-directed mutagenesis

Methods

5 [0082] Materials. *Escherichia coli* BL21 (DE3) was from Invitrogen (Carlsbad, CA, USA). Ni²⁺-NTA agarose (nickel-nitrilotriacetic acid agarose) and QIAprep spin miniprep kit were from Qiagen (Valencia, CA, USA). Bicinchoninic acid (BCA) protein assay kit was from Pierce Biotechnology, Inc. (Rockford, IL). QuikChange Multi Site-Directed Mutagenesis Kit was from Agilent Technologies company/Stratagene (Santa Clara, CA).

10 [0083] Site-directed mutagenesis. Site-directed mutagenesis was carried out using the QuikChange multi-site-directed mutagenesis kit from Stratagene according to the manufacturer's protocol. The primers used were
5’ACCGGCACGCACACTTGTTGGAATACCGATGTGC3’ for E271F and
5’ATCTACTTTAAAGGGCATCCTTATGGTGAAATTAATGACTAC3’ for R313Y. The sites of mutations are underlined.

[0084] Protein expression and purification. The plasmids containing the mutant genes were transformed into *E. coli* BL21 (DE3). The *E. coli* cells were cultured in LB-rich media (10 g L⁻¹ tryptone, 5 g L⁻¹ yeast extract, and 10 g L⁻¹ NaCl) supplemented with ampicillin (100 µg ml⁻¹). Overexpression of the mutant proteins was achieved by adding 0.1 mM of isopropyl-1-thio-p-D-galactopyranoside (IPTG) to the *E. coli* culture when its OD₆₀₀nm reached 0.8. The incubation of the induced culture was performed at 37 °C for 3 h with vigorous shaking at 250 rpm in a C25KC incubator shaker (New Brunswick Scientific, Edison, NJ).

[0085] His₆-tagged mutant proteins were purified from the cell lysate. To obtain the cell lysate, the cell pellet harvested by centrifugation at 4000 rpm for 2 h was resuspended in 20 ml of lysozyme (50 µg ml⁻¹ and DNasel (3 µg ml⁻¹) were then added to the resuspended cells followed by shaking at 37 °C for 60 min. The cell lysate was obtained as the supernatant after centrifugation at 11,000 rpm for 20 min. Purification of His₆-tagged proteins from the lysate was achieved using an ÄKTA FPLC system (GE Healthcare)
equipped with a HisTrap™ FF 5 mL column. The column was pre-equilibrated with 8 column volumes of the binding buffer (5 mM imidazole, 0.5 M NaCl, 50 mM Tris-HCl pH 7.5) prior to lysate loading. After the sample loading, the column was washed with 8 column volumes of the binding and washing buffer (40 mM imidazole, 0.5 M NaCl, 50 mM Tris-HCl pH 7.5). Protein elution was carried out with 8 column volumes of the elute buffer (200 mM imidazole, 0.5 M NaCl, 50 mM Tris-HCl pH 7.5). The fractions containing the purified enzyme were collected and stored at 4 °C.

[0086] **Kinetic assays.** The kinetic assays for the sialidase activity were performed in duplicate in a total volume of 10 μL in MES buffer (100 mM, pH 5.5) containing different concentrations of Neu5Acα2-3LacMU (0.4, 1.0, 2.0, 4.0, 10.0, 20.0, 40.0, and 60.0 mM) and the mutant proteins (2.5 mg mL⁻¹ ofD141A, 1.6 mg mL⁻¹ ofE271F, 1 mg mL⁻¹ ofR313Y, and 3.2 mg mL⁻¹ ofE271F/R313Y). All reactions were allowed to proceed at 37 °C for 60 min (D141A), 1 min (E271F), 25 min (R313Y), and 20 min (E271F/R313Y). The apparent kinetic parameters were obtained by fitting the experimental data (the average values of duplicate assay results) into the Michaelis-Menten equation using Grafit 5.0.

[0087] To obtain the apparent kinetic parameters of LacβMU as the acceptor for the α2-3-sialyltransferase activity, the kinetic assays were performed in duplicate in reaction mixtures of 10 pL containing Tris-HCl buffer (100 mM, pH 8.5), a fixed concentration of CMP-Neu5Ac (1 mM), different concentrations of LacβMU (0.2, 0.5, 1.0, 2.0, 5.0, and 9.0 mM) and the mutant proteins (2 pg mL⁻¹ ofE271F, 2 pg mL⁻¹ ofR313Y, and 1.6 μg mL⁻¹ ofE271F/R313Y). All reactions were allowed to proceed at 37 °C for 5 min (E271F), 7 min (R313Y), and 10 min (E271F/R313Y). The apparent kinetic parameters were obtained by fitting the experimental data (the average values of duplicate assay results) into the Michaelis-Menten equation using Grafit 5.0.

[0088] To obtain the apparent kinetic parameters of CMP-Neu5Ac as the donor for the α2-3-sialyltransferase activity, the kinetic assays were performed in duplicate in reaction mixtures of 10 pL containing Tris-HCl buffer (100 mM, pH 8.5), a fixed concentration of LacβMU (1 mM), different concentrations of CMP-Neu5Ac (0.1, 0.2, 0.5, 1.0, 2.0, 5.0, 10.0 and 20.0 mM) and the mutant proteins (2 pg mL⁻¹ ofE271F, 2 pg mL⁻¹ ofR313Y, and 1.6 μg mL⁻¹ ofE271F/R313Y).
All reactions were allowed to proceed at 37 °C for 2 min (E271F), 7 min (R313Y), and 5 min (E271F/R313Y). The apparent kinetic parameters were obtained by fitting the experimental data (the average values of duplicate assay results) into the Michaelis-Menten equation using Grafit 5.0.

[0089] All the sialidase and a2-3-sialyltransferase assays were performed in an HPLC system. Reactions were stopped by adding 10 µL of ethanol. After necessary dilutions were performed to adjust the concentrations of the fluorescent-labeled compounds, the samples were then kept on ice until aliquots of 8 µL were injected and analyzed by a Shimadzu LC-6AD system equipped with a membrane on-line degasser, a temperature control unit, and a fluorescence detector (Shimadzu RF-10AXL). A reverse-phase Premier C18 column (250 x 4.6 mm i.d., 5 µm particle size, Shimadzu) protected with a C18 guard column cartridge was used. The mobile phase was 25% acetonitrile. The fluorescent compounds LacβMU and Neu5Aca2-3LacpMU were detected by excitation at 325 nm and emission at 372 nm.

[0090] Acceptor substrate specificity assays by HPLC. Assays were performed in duplicate in 20 µL of Tris-HCl buffer (100 mM, pH 8.5) containing CMP-Neu5Ac (1 mM), a fluorescent acceptor (1 mM), MgCl₂ (20 mM), and an enzyme (2 DOC mL⁻¹, wild-type PmSTl or E271R/R313Y mutant). Reactions were allowed to proceed for 5 min at 37 °C. The 4-methylumbelliferone (MU)-labeled fluorescent acceptors and the products formed were detected with excitation at 325 nm and emission at 372 nm. The 9-fluorenylmethylcarbamate (Fmoc)-labeled fluorescent acceptors and the products formed were detected with excitation at 262 nm and emission at 313 nm. The 2-aminobenzoic acid (2AA)-labeled fluorescent acceptors and the products formed were detected with excitation at 315 nm and emission at 400 nm.

[0091] Stability studies by HPLC. Thermal stability studies were carried out by incubating wild-type PmSTl or E271F/R313Y mutant solution (20 µg mL⁻¹) at 37 °C. Samples were withdrawn at various time intervals for enzyme activity assays.

Results

[0092] Kinetics of the a2-3-sialidase activity. [0093] To test the involvement of D141 and H31 in the a2-3-sialidase activity of PmSTl, the ct2-3-sialidase activity of two previously obtained PmSTl mutants, D141A and H31A, were evaluated using a fluorescent a2-3-sialidase assay.
sialoside, Neu5Aca2-3LacpMU, as the substrate. The a2-3-sialidase activity of H31 A mutant was too low to obtain the kinetic data. For the a2-3-sialidase activity of D141 A mutant, the $K_m$ value (15 ± 1 mM) was about the same as the wild-type PmSTl (24 mM), but its catalytic efficiency was about 7,300-fold lower than that of the wild-type PmSTl mainly due to a much slower turnover number of the D141A mutant (Table 1). These data indicated that both D141 and H31 1 are important for the a2-3-sialidase activity of PmSTl.

Table 1: Apparent kinetic data for the a2-3-sialidase activity of wild-type PmSTl (WT) and PmSTl mutants.

<table>
<thead>
<tr>
<th></th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_m$ (s$^{-1}$ mM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT$^4$</td>
<td>24</td>
<td>2.3 × 10$^2$</td>
<td>9.5</td>
</tr>
<tr>
<td>D141A</td>
<td>15 ± 1</td>
<td>(1.9 ± 0.1) × 10$^2$</td>
<td>1.3 × 10$^3$</td>
</tr>
<tr>
<td>E271F</td>
<td>5.7 ± 0.9</td>
<td>0.92 ± 0.04</td>
<td>0.16</td>
</tr>
<tr>
<td>R313Y</td>
<td>5 ± 1</td>
<td>0.18 ± 0.01</td>
<td>3.6 × 10$^3$</td>
</tr>
<tr>
<td>E271F/R313Y</td>
<td>(5.4 ± 0.6) × 10$^2$</td>
<td>0.83 ± 0.08</td>
<td>1.5 × 10$^3$</td>
</tr>
</tbody>
</table>

[0094] Kinetics of the a2-3-sialidase activity of the mutants E271F, R313Y, and E271F/R313Y. The designed PmSTl mutants E271F, R313Y, and E271F/R313Y were expressed in E. coli using the same expression condition as the wild-type PmSTl (100 mg L$^{-1}$ culture) and achieved a compatible level of expression (90 mg L$^{-1}$ culture). Similar to the wild-type PmSTl, one-step N$i^{2+}$-column purification was sufficient to provide pure protein (>99%) of the mutants.

[0095] The kinetic assays for the a2-3-sialidase activity of the mutants E271F, R313Y, and E271F/R313Y using a fluorescent 4-methylumbelliferyl sialoside, Neu5Aca2-3LacpMU, as the substrate (Table 1) indicated that E271F mutation decreased the a2-3-sialidase activity of PmSTl about 59-fold which was mainly caused by a 250-fold decrease in the turnover number despite of a 4.2-fold decrease in the $k_{cat}$ value. As expected, the R313Y mutation at a site close to the critical H31 1 residue for the a2-3-sialidase activity of PmSTl caused a 2,639-fold decrease in the catalytic efficiency ($k_{cat}/K_m = 0.0036$ s$^{-1}$ mM$^{-1}$) compared to the wild-type PmSTl ($k_{cat}/K_m = 9.5$ s$^{-1}$ mM$^{-1}$) mainly due to a (1,278-fold) decreased $k_{cat}$ value and a 2-fold increased $K_m$ value.
The E271F/R313Y double mutant had the lowest α2-3-sialidase activity (\(k_{\text{cat}}K_m = 0.0015 \text{ s}^{-1} \text{ mM}^{-1}\)) which was a 6,333-fold decrease compared to the wild-type PmSTl due to a 22.5-fold increase in the \(K_m\) value and a 277-fold decrease in the \(k_{\text{cat}}\) value.

Kinetics of the α2-3-sialyltransferase activity of the mutants E271F, R313Y, and E271F/R313Y. Kinetic assays (Table 2) for the α2-3-sialyltransferase activity of mutants E271F, R313Y, and E271F/R313Y using LacpMU as the fluorescent acceptor and CMP-Neu5Ac as the donor indicated that either E271F or R313Y mutation did not cause significant changes on either the \(K_m\) or the \(k_{\text{cat}}\) value, leading to quite consistent catalytic efficiencies (\(k_{\text{cat}}/K_m = 28-39 \text{ s}^{-1} \text{ mM}^{-1}\)) compared to the wild-type PmSTl (\(k_{\text{cat}}/K_m = 34 \text{ s}^{-1} \text{ mM}^{-1}\)).

Table 2: Apparent kinetic data for the α2-3-sialyltransferase activity of wild-type PmSTl (WT) and PmSTl mutants.

<table>
<thead>
<tr>
<th>Enzymes</th>
<th>(K_m) (mM)</th>
<th>(k_{\text{cat}}) (s(^{-1}))</th>
<th>(k_{\text{cat}}/K_m) (s(^{-1}) mM(^{-1}))</th>
<th>(K_m) (mM)</th>
<th>(k_{\text{cat}}) (s(^{-1}))</th>
<th>(k_{\text{cat}}/K_m) (s(^{-1}) mM(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>0.44</td>
<td>32</td>
<td>73</td>
<td>1.4</td>
<td>47</td>
<td>34</td>
</tr>
<tr>
<td>E271F</td>
<td>0.18 ± 0.01</td>
<td>26 ± 1</td>
<td>1.4 × 10(^2)</td>
<td>0.71 ± 0.12</td>
<td>28 ± 1</td>
<td>39</td>
</tr>
<tr>
<td>R313Y</td>
<td>0.62 ± 0.04</td>
<td>19 ± 1</td>
<td>30</td>
<td>0.67 ± 0.05</td>
<td>19 ± 1</td>
<td>28</td>
</tr>
<tr>
<td>E271F/R313Y</td>
<td>0.34 ± 0.02</td>
<td>23 ± 1</td>
<td>69</td>
<td>0.54 ± 0.04</td>
<td>17 ± 1</td>
<td>32</td>
</tr>
</tbody>
</table>

Acceptor substrate specificities of wild-type PmSTl and E271F/R313Y mutant. Fluorescent glycans with different glycosidic linkages and various monosaccharide units, including Galβ1-3Glcβ, Galp-4GlcNACP, Galp-3GalNAca, and Galp-3GlcNACP structures, were used to investigate the acceptor substrate specificities of the wild-type PmSTl and E271F/R313Y mutant. As shown in Figure 2, the E271F/R313Y mutant exhibited similar or slightly higher activity than the wild-type PmSTl towards different acceptors. Therefore, the acceptor promiscuity of PmSTl was not changed significantly by E271F and R313Y mutations.

Thermal stabilities of wild-type PmSTl and E271F/R313Y mutant. Incubating the wild-type PmSTl and E271F/R313Y mutant at 37 °C for up to 2 hours did not decrease their activities significantly (Figure 3). Therefore, both enzymes are considered quite stable and mutation does not affect the thermal stability of the PmSTl.
Example 2. A Sialyltransferase Mutant with Decreased Donor Hydrolysis and Reduced Sialidase Activities for Directly Sialylating Lewis’-

Methods

[0099] Site-directed mutagenesis, expression and purification of PmST1 mutants. Site-directed mutagenesis was performed using the QuikChange multi-site-directed mutagenesis kit from Stratagene according to the manufacturer’s protocol. The primers used were 5’ AATCTTTATGACGATGGCTCA GATGAATATGTTGATTTAGAAAAAG 3’ for M144D; 5’ AATCTTTATGACGATGGCTCACATGAATATGTTGATTTAGAAAAAG 3’ for M144H; 5’ ATCACGCTGTATTTAGATCCTCATTCCTTACCGGCATTAAATCAG 3’ for A35D; and 5’ ATCACGCTGTATTTAGATCCTCATTCCTTACCGGCATTAAATCAG 3’ for A35H. The expression and purification of the mutants were performed as previously described for the WT PmST1.

[0100] Kinetics of the donor hydrolysis activity of PmST1 and mutants by capillary electrophoresis analysis. The reactions were carried out in duplicate in a total volume of 10 µL at 37 °C for 15 min (WT), 40 min (D141A), 20 min (H31 IA), or 15 min (M144D and M144H) in Tris-HCl buffer (200 mM, pH 8.5) containing CMP-Neu5Ac (1, 2, 5, 10, 20 and 40 mM) and an enzyme (WT, 4 µg mL⁻¹; D141A, 1500 µg mL⁻¹; H31 IA, 40 µg mL⁻¹; M144D, 39 µg mL⁻¹; M144H, 5 µg mL⁻¹). The reactions were stopped by adding 10 µL of pre-chilled ethanol. The mixtures were incubated on ice for 30 min and centrifuged at 13,000 rpm for 5 min. The supernatants were diluted with borate buffer (25 mM, pH 9.5) and aliquots of 5 µL each were injected to a Beckman Coulter P/ACE™ MDQ Capillary Electrophoresis system equipped with a capillary (60 cm × 75 µm i.d.) and monitored at 254 nm. The apparent kinetic parameters were obtained by fitting the experimental data (the average values of duplicate assay results) into the Michaelis-Menten equation using Grafit 5.0.

[0101] Kinetics of the a2-3-sialyltransferase activity of PmST1 mutants by HPLC analysis. With LacpMU as the acceptor substrate, the reactions were performed in duplicate at 37 °C for 10 min (M144D) or 4 min (M144H) in a reaction mixture (10 µL) containing Tris-HCl (100 mM, pH 8.5), an enzyme (5 µg mL⁻¹), and different concentrations (0.2, 0.5, 1.0, 2.0, and 5.0 mM) of LacpMU with a fixed concentration (1 mM) of CMP-Neu5Ac or different concentrations (0.2,
0.5, 1.0, 2.0, 5.0, and 10.0 mM) of CMP-Neu5Ac with a fixed concentration (1 mM) of 
LacpMU. With LeβMU as the acceptor substrate, the reactions were carried out in duplicate at 
37 °C for 9 min (M144D) or 10 min (M144H) in a reaction mixture (10 μL) containing CAPSO 
(100 mM, pH 9.5), an enzyme (M144D, 39 μg ml⁻¹ or M144H, 5 μg ml⁻¹), and various 
concentrations of LeβMU (1.0, 5.0, 10.0, 15.0, 25.0, and 35.0 mM) with a fixed concentration 
(1 mM) of CMP-Neu5Ac or various concentrations (0.2, 0.5, 1.0, 2.0, 5.0, 10.0, 20.0, and 40.0 
mM) of CMP-Neu5Ac with a fixed concentration (1 mM) of LepMU. Reactions were stopped 
by adding 10 μL of pre-chilled ethanol. The mixtures were incubated on ice for 30 min and 
centrifuged at 13,000 rpm for 5 min. The supernatants were diluted with 25% acetonitrile and 
kept on ice until aliquots of 8 μL were injected and analyzed by the Shimadzu LC-6AD system 
equipped with a membrane on-line degasser, a temperature control unit, and a fluorescence 
detector (Shimadzu RF-10AXL). A reverse-phase Premier C18 column (250 × 4.6 mm i.d., 5 
μm particle size, Shimadzu) protected with a C18 guard column cartridge was used. The mobile 
phase was 25% acetonitrile. The fluorophore (MU)-labeled compounds were detected by 
excitation at 325 nm and emission at 372 nm. The apparent kinetic parameters were obtained by 
fitting the experimental data (the average values of duplicate assay results) into the Michaelis-
Menten equation using Grafit 5.0.

[0102] Kinetics of the a2-3-sialidase activity of PmSTI mutants. The reactions were 
performed in duplicate in a total volume of 10 μL at 37 °C for 60 min (M144D) or 15 min 
(M144H) in MES buffer (100 mM, pH 5.5) containing Neu5Aca2-3LacpMU (0.4, 1, 2, 4, 10, 
20, 40 and 60 mM) and an enzyme (M144H, 1.36 mg ml⁻¹ or M144D, 1.05 mg ml⁻¹). Sample 
treatment after the reaction and analysis were carried out by HPLC similar to that described 
above for the a2-3-sialyltransferase assays.

[0103] The a2-3-sialidase activity assays of PmSTI and mutants. The reactions were 
carried out in duplicate in a total volume of 10 μL at 37 °C for 20 hr in MES buffer (100 mM, 
pH 5.5) containing Neu5Aca2-3LeβMU (1 mM) and an enzyme (4 mg ml⁻¹). Aliquots of 1 μL 
were withdrawn at 1 hr, 6 hr and 20 hr, and analyzed by HPLC as described above for the oc2-3-
sialyltransferase assays.
[0104] **Accession Codes.** The structure of PmSTl M144D mutant in complex with CMP-3F(a)-Neu5Ac was deposited with a PDB ID code 3S44.

[0105] **Materials and compound characterization.** Chemicals were purchased and used without further purification. $^1$H NMR (600 MHz) and $^{13}$C NMR (150 MHz) spectra were recorded on a Varian VNMRS 600 MHz spectrometer or $^1$H NMR (800 MHz) and $^{13}$C NMR (200 MHz) on a Bruker 800 MHz spectrometer. High resolution electrospray ionization (ESI) mass spectra were obtained at the Mass Spectrometry Facility in the University of California, Davis. Silica gel 60 Å was used for flash column chromatography. Thin-layer chromatography (TLC) was performed on silica gel plates using anisaldehyde sugar stain or 5% sulfuric acid in ethanol stain for detection. Gel filtration chromatography was performed with a column (100 cm x 2.5 cm) packed with BioGel P-2 Fine resins. Pasteurella multocida sialic acid aldolase, 1 N. meningitidis CMP-sialic acid synthetase (NmCSS),2 and wild-type PmSTl were expressed in E. coli and purified as described previously.

[0106] **Crystallization and structure determination.** PmSTl M144D mutant in Tris-HCl buffer (20 mM, pH 7.5) was concentrated to 13 mg mL$^{-1}$, and CMP-3F(αXαl)Neu5Ac was added to a final concentration of 2 mM. Binary CMP-3F(αXαl)Nεu5Ac crystals were grown by hanging drop with 3 μL of the sample mixed with an equal volume of reservoir buffer [24% polyethylene glycol] 3350, 100 mM HEPES (pH 7.5), 50 mM NaCl, and 0.4% Triton X-100]. Then, the binary crystals were soaking with 10 mM of CMP-3F(βXβl)-Neu5Ac and 10 mM of Le$^3$pProN$_3$ in buffer containing 26% poly (ethylene glycol) 3350, HEPES (100 mM, pH 7.5), NaCl (100 mM), and 0.4% Triton X-100 for overnight. All crystals were transferred to Paratone-N and frozen in a stream of nitrogen to 100 K for data collection. Diffraction data were collected at the Stanford Synchrotron Radiation Lightsource to 1.45 Å resolution. Data were processed with XDS and scaled with XSACE (Table 3). The structure was solved by Molecular Replacement using the program PHASER. Only the ligand-free open conformation structure (PDB ID: 2EX0) was successful in structure determination. The model was displayed and adjusted with COOT and refined with REFMAC. Final data processing and refinement statistics are shown in Table 3.

**Table 3: X-Ray data collection and refinement statistics for PmSTl M144D.**

<table>
<thead>
<tr>
<th>unit cell dimensions a,b,c (Å), β</th>
</tr>
</thead>
<tbody>
<tr>
<td>52.44, 61.57, 62.58, β = 114.15°</td>
</tr>
</tbody>
</table>

33
The NMR spectra were processed and analyzed using the software, NMRPipe.

[0107] **NMR analysis of WT PmSTl and M144D mutant.** Enzymes were expressed in E. coli BL21 (DE3) using M9 media containing $^{15}$NH$_4$Cl (1.0 g L$^{-1}$), Na$_2$HP0$_4$ 7H$_2$O (12.66 g L$^{-1}$), K$_2$HPO$_4$ (3.0 g L$^{-1}$), NaCl (0.5 g L$^{-1}$), MgSO$_4$ (0.2 g L$^{-1}$), CaCl$_2$ (50 µM), and glucose (0.3%). Expressions were induced by adding 0.5 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubating at 37 °C for 4 hr. The purifications were performed as previously described for the WT PmSTl. The purified enzymes were dialyzed with a phosphate buffer (10 mM, pH 7.0). NMR samples of $^{15}$N-labeled WT and M144D PmSTl (~ 0.7 mM) were prepared in 90%/10% of H$_2$O/D$_2$O containing 10 mM of phosphate (pH 7.0) in the presence or the absence of saturating CMP. $^{15}$N-H HSQC NMR experiments were performed at 37 °C on Bruker Avance III 800 spectrometer with an Ultrashield Bruker magnet equipped with a four-channel interface, triple-resonance probe, and cryo-probe with Z-axis pulsed field gradients. The number of complex points and acquisition times were: 256, 180 ms ($^1$H (F)); and, 512, 64 ms ($^1$H (F2)). The NMR spectra were processed and analyzed using the software, NMRPipe.
Example 3. Preparation of Oligosaccharides

[0108] One-pot three-enzyme synthesis of SLe\(^{\beta}\)ProN\(_3\) with different sialic acid forms.
Le\(^{\beta}\)ProN\(_3\) (20-25 mg),\(^9\) a sialic acid precursor (mannose, ManNAc, ManNGc or their
derivatives, 1.5 equiv.), sodium pyruvate (5 equiv.), and CTP (1.5 equiv.) were dissolved in Tris-
HCl buffer (10 mL, 100 mM, pH 7.5-8.5) containing MgCl\(_2\) (20 mM) and appropriate amounts
of Pm aldolase (0.5 mg), NmCSS (0.3-0.5 mg), and PmSTI mutant M144D (0.5-0.9 mg). The
reactions were carried out by incubating the reaction mixture in an incubator shaker at 37 °C for
4-6 h. The product formation was monitored by TLC developed with EtOAc:MeOH:H\(_2\)O:HOAc
= 4:2:1:0.2 (by volume) and stained with p-anisaldehyde sugar stain. When an optimal yield was
achieved, the reaction was stopped by adding the same volume (10 mL) of cold EtOH and
incubation at 4 °C for 30 min. The mixture was then centrifuged and the precipitates were
removed. The supernatant was concentrated, passed through a BioGel P-2 gel filtration column,
and eluted with water to obtain partially purified product. A silica gel column was then used to

[0109] NMR chemical shifts and HRMS data of SLe\(^{\beta}\)ProN\(_3\) containing different sialic acid
forms synthesized by the one-pot three-enzyme system.

[0110] Neu5Aca2-3Le\(^{\beta}\)ProN\(_3\) (la). 33 mg, yield 93%. \(^1\)H NMR (600 MHz, D\(_2\)O): \(\delta \) 5.09
(d, 1H, \(J = 4.2 \) Hz), 4.50 (d, 1H, \(J = 7.8 \) Hz), 4.07 (dd, 1H, \(J = 10.4 \) and 3.2 Hz), 4.01-3.82 (m,
11H), 3.74 (d, 1H, \(J = 4.2 \) Hz), 3.66-3.59 (m, 9H), 3.56-3.50 (m, 4H), 3.36-3.30 (m, 2H), 2.72
(dd, 1H, \(J = 12.6 \) and 4.8 Hz), 2.01 (s, 3H), 2.00 (s, 3H), 1.87 (m, 2H), 1.75 (t, 1H, \(J = 12.3 \) Hz),
1.12 (d, 3H, \(J = 6.6 \)). \(^13\)C NMR (150 MHz, D\(_2\)O): \(\delta \) 175.20, 174.41, 174.05, 101.79, 101.15,
99.82, 98.76, 75.81, 75.42, 75.07, 74.98, 73.51, 73.07, 72.07, 72.03, 69.42, 69.35, 68.47, 68.28,
67.87, 67.47, 67.36, 66.84, 62.76, 61.64, 59.81, 55.98, 51.86, 47.93, 39.95, 28.27, 22.39, 22.20,
15.43. HRMS (ESI) m/z calcd for C\(_{39}\)H\(_{57}\)N\(_3\)O\(_{23}\)Na (M+Na) 926.3319, found 926.3342.

[0111] Neu5Gc2-3Le\(^{\beta}\)ProN\(_3\) (lb). 28 mg, yield 87%. \(^1\)H NMR (600 MHz, D\(_2\)O): \(\delta \) 5.13
(d, 1H, \(J = 4.2 \) Hz), 4.56-4.54 (m, 2H), 4.15 (s, 2H), 4.07 (dd, 1H, \(J = 10.4 \) and 3.2 Hz), 4.01-
3.82 (m, 12H), 3.78-3.60 (m, 8H), 3.56-3.54 (m, 3H), 3.52 (dd, 1H, \(J = 10.4 \) and 7.8 Hz), 3.36-
3.30 (m, 2H), 2.78 (dd, 1H, \(J = 12.6 \) and 4.8 Hz), 2.06 (s, 3H), 1.87 (m, 2H), 1.75 (t, 1H, \(J = 12.3 \)
Hz), 1.19 (d, 3H, \(J = 6.6 \)). \(^13\)C NMR (150 MHz, D\(_2\)O): \(\delta \) 175.73, 174.17, 173.84, 101.55, 100.91,
99.62, 98.52, 75.59, 75.20, 74.83, 74.75, 73.30, 72.57, 71.84, 69.20, 69.12, 69.10, 67.98, 67.65, 67.23, 67.13, 66.67, 66.61, 62.50, 60.91, 59.79, 59.60, 55.75, 51.33, 47.71, 39.78, 28.04, 22.16, 15.20. HRMS (ESI) m/z calcd for C_{14}H_{35}N_{2}O_{24}Na (M+Na) 942.3291, found 942.3292.

[0112] Kdna2-3Le \[^1\] \( \beta \)ProN\(_3\) (le). 27 mg, yield 85%. \(^1\)H NMR (600 MHz, D\(_2\)O): \( \delta \) 5.04 (d, 1H, J = 4.2 Hz), 4.47-4.45 (m, 2H), 3.85 (dd, 1H, J = 9.6 and 2.4 Hz), 3.81-3.62 (m, 10H), 3.56 (d, 1H, J = 4.0 Hz), 3.49-3.29 (m, 12H), 3.18-3.14 (m, 3H), 2.65 (dd, 1H, J = 12.6 and 4.8 Hz), 1.98 (s, 3H), 1.87 (m, 2H), 1.69 (t, 1H, J = 12.3 Hz), 1.10 (d, 3H, J = 6.6). \(^{13}\)C NMR (150 MHz, D\(_2\)O): \( \delta \) 174.36, 174.17, 101.72, 101.09, 99.76, 98.74, 75.70, 75.35, 75.02, 74.93, 74.04, 73.43, 72.26, 71.99, 70.32, 69.84, 69.34, 69.26, 67.79, 67.36, 67.28, 66.79, 62.73, 61.60, 59.73, 55.91, 47.84, 39.51, 28.21, 22.31, 15.37. HRMS (ESI) m/z calcd for C\(_y\)H\(_{54}\)N\(_0\) Na (M+Na) 885.3077, found 885.3103.

[0113] Neu5AcN\(_3\)a2-3Le"\(^p\)ProN\(_3\) (Id). 33 mg, yield 89%. \(^1\)H NMR (800 MHz, D\(_2\)O): \( \delta \) 5.06 (d, 1H, J = 4.0 Hz), 4.47-4.45 (m, 2H), 4.01 (s, 2H), 3.96-3.76 (m, 11H), 3.72-3.59 (m, 10H), 3.54-3.45 (m, 3H), 3.37 (dd, 1H, J = 10.4 and 7.8 Hz), 3.34-3.27 (m, 2H), 2.71 (dd, 1H, J = 12.6 and 4.8 Hz), 1.98 (s, 3H), 1.81 (m, 2H), 1.74 (t, 1H, J = 12.3 Hz), 1.10 (d, 3H, J = 6.6). \(^{13}\)C NMR (200 MHz, D\(_2\)O): \( \delta \) 174.43, 174.09, 171.36, 101.75, 101.17, 99.82, 98.81, 75.79, 75.40, 75.08, 74.99, 73.47, 72.74, 72.10, 72.06, 69.43, 69.33, 68.35, 68.19, 67.86, 67.44, 67.35, 66.86, 62.72, 61.67, 59.79, 55.98, 52.06, 51.92, 47.91, 39.97, 28.28, 22.38, 15.44. HRMS (ESI) m/z calcd for C\(_{34}H_{86}N_{0}\) Na (M+Na) 967.3356, found 967.3396.

[0114] Kdn1a2-3Le \(^p\)\(^{\beta}\)ProN\(_3\) (le). 27 mg, yield 84%. \(^1\)H NMR (600 MHz, D\(_2\)O): \( \delta \) 5.11 (d, 1H, J = 4.0 Hz), 4.54 (d, 1H, J = 8.0 Hz), 4.52 (d, 1H, J = 8.0 Hz), 4.07 (dd, 1H, J = 9.6 and 3.2 Hz), 4.03-3.83 (m, 11H), 3.78 (d, 1H, J = 3.2 Hz), 3.72-3.66 (m, 6H), 3.60-3.49 (m, 5H), 3.38-3.35 (m, 3H), 2.75 (dd, 1H, J = 12.6 and 4.8 Hz), 2.04 (s, 3H), 1.83 (m, 2H), 1.78 (t, 1H, J = 12.3 Hz), 1.17 (d, 3H, J = 6.6). \(^{13}\)C NMR (150 MHz, D\(_2\)O): \( \delta \) 174.15, 173.64, 101.49, 100.88, 99.60, 98.52, 75.53, 74.25, 74.78, 74.73, 73.25, 72.77, 71.93, 71.80, 69.38, 69.16, 69.08, 68.30, 67.61, 67.10, 66.59, 62.50, 62.46, 61.38, 59.57, 55.73, 47.67, 39.50, 28.02, 22.13, 15.17. HRMS (ESI) m/z calcd for C\(_{12}H_{33}N_{0}\) Na (M+Na) 910.3141, found 910.3137.

[0115] 9-N\(_3\)-Neu5Acoc2-3Le \(^p\)ProN\(_3\) (If). 28 mg, yield 91%. \(^1\)H NMR (800 MHz, D\(_2\)O): \( \delta \) 5.11 (d, 1H, J = 4.0 Hz), 4.56 (d, 1H, J = 8.0 Hz), 4.51 (d, 1H, J = 8.0 Hz), 4.03-4.02 (m, 2H), 3.74-3.62 (m, 2H), 3.60-3.49 (m, 5H), 3.38-3.35 (m, 3H), 2.75 (dd, 1H, J = 12.6 and 4.8 Hz), 2.04 (s, 3H), 1.83 (m, 2H), 1.78 (t, 1H, J = 12.3 Hz), 1.17 (d, 3H, J = 6.6). \(^{13}\)C NMR (150 MHz, D\(_2\)O): \( \delta \) 174.15, 173.64, 101.49, 100.88, 99.60, 98.52, 75.53, 74.25, 74.78, 74.73, 73.25, 72.77, 71.93, 71.80, 69.38, 69.16, 69.08, 68.30, 67.61, 67.10, 66.59, 62.50, 62.46, 61.38, 59.57, 55.73, 47.67, 39.50, 28.02, 22.13, 15.17. HRMS (ESI) m/z calcd for C\(_{12}H_{33}N_{0}\) Na (M+Na) 910.3141, found 910.3137.
3.98-3.85 (m, 9H), 3.79 (d, 1H, J = 3.2 Hz), 3.71-3.69 (m, 8H), 3.61-3.49 (m, 6H), 3.40-3.36 (m, 3H), 2.77 (dd, 1H, J = 12.6 and 4.8 Hz), 2.05 (s, 3H), 1.83 (m, 2H), 1.79 (t, 1H, J = 12.3 Hz), 1.17 (d, 3H, J = 6.6). 13C NMR (200 MHz, D2O): δ 174.90, 174.16, 173.74, 101.51, 100.88, 99.58, 98.50, 75.63, 75.21, 74.81, 74.73, 73.24, 72.64, 72.42, 71.81, 70.39, 69.19, 69.09, 68.70, 68.20, 67.61, 67.15, 66.59, 61.38, 59.60, 55.72, 53.01, 51.59, 47.68, 39.77, 28.02, 22.14, 15.18. HRMS (ESI) m/z calcd for C34H58NaO22Na (M+Na) 951.3407, found 910.3407.

[0116] 9-Ac-Neu5Ac2-3Le*ProN3 (lg). 20 mg, yield 62%. 1H NMR (600 MHz, D2O): δ 5.12 (d, 1H, J = 4.0 Hz), 4.55-4.53 (m, 2H), 4.44 (dd, 1H, J = 11.4 and 1.8 Hz), 4.20 (dd, 1H, J = 11.4 and 6.6 Hz), 4.14-3.86 (m, 11H), 3.79 (d, 1H, J = 3.0 Hz), 3.74-3.65 (m, 8H), 3.58-3.56 (m, 2H), 3.54 (dd, 1H, J = 9.6 and 7.8 Hz), 3.41-3.36 (m, 2H), 2.78 (dd, 1H, J = 12.6 and 4.8 Hz), 2.16 (s, 3H), 2.05 (s, 6H), 1.85 (m, 2H), 1.81 (t, 1H, J = 12.6 Hz), 1.18 (d, 3H, J = 6.6). 13C NMR (150 MHz, D2O): δ 174.18, 174.41, 174.00, 173.96, 101.80, 101.14, 99.80, 98.78, 75.88, 75.53, 75.09, 74.80, 73.54, 72.89, 72.08, 69.75, 69.57, 69.45, 69.37, 68.46, 68.37, 67.89, 67.37, 66.86, 65.98, 61.66, 59.84, 55.95, 51.85, 47.60, 40.05, 28.29, 22.41, 22.23, 20.44, 15.45. HRMS (ESI) m/z calcd for C36H59N5O24Na (M+Na) 968.3448, found 968.3427.

[0117] 9-OAc-Neu5Gca2-3Le*βProN3 (lh). 21 mg, yield 64%. 1H NMR (600 MHz, D2O): δ 5.12 (d, 1H, J = 4.0 Hz), 4.55 (dd, 1H, J = 7.8 and 4.2 Hz), 4.45 (m, 1H), 4.22-4.19 (m, 1H), 4.14 (s, 2H), 4.13-3.79 (m, 13H), 3.73-3.66 (m, 8H), 3.61-3.59 (m, 2H), 3.55 (t, 1H, J = 7.8 Hz), 3.41-3.37 (m, 2H), 2.80 (dd, 1H, J = 12.6 and 4.8 Hz), 2.16 (s, 3H), 2.06 (s, 6H), 1.88-1.85 (m, 2H), 1.81 (t, 1H, J = 12.6 Hz), 1.19 (d, 3H, J = 6.6). 13C NMR (150 MHz, D2O): δ 175.93, 174.54, 174.41, 174.01, 101.80, 101.15, 99.81, 99.77, 75.86, 75.53, 75.09, 75.00, 73.54, 72.62, 72.09, 69.75, 69.63, 69.45, 69.37, 68.30, 68.23, 67.90, 67.36, 66.86, 65.93, 61.65, 61.17, 59.83, 55.99, 51.54, 47.95, 40.10, 28.29, 22.41, 20.44, 15.45. HRMS (ESI) m/z calcd for C36H59N5O24Na (M+Na) 984.3397, found 984.3397.

Results and discussion

[0118] Donor hydrolysis by PmSTI causes low yield sialylation of Leα. In order to understand why PmSTI-catalyzed sialylation of Leα resulted in low yields, time course studies were carried out using a fluorescently labeled Leα acceptor (4-methylumbelliferyl β-Leα or Leα-BMU) in a high performance liquid chromatography (HPLC) assay. As shown in Figure 4, PmSTI-catalyzed...
sialylation of Le^xβMU (1 mM) using one equivalent of donor CMP-Neu5Ac reached a low yield (1.1-1.3%) plateau quickly within 2 min. Every additional dose of donor substrate CMP-Neu5Ac (shown by arrows in Figure 4) increased the product formation which always reached a plateau quickly. Monitoring the CMP-Neu5Ac consumption (% consumption numbers are shown in parentheses in Figure 4) in the reaction mixture by capillary electrophoresis studies confirmed a quick consumption of CMP-Neu5Ac. These indicated that donor (CMP-Neu5Ac) hydrolysis activity of PmSTl, where water molecules compete with the poor Le^x acceptor for the consumption of sugar nucleotide (CMP-Neu5Ac) donor of the sialyltransferase (Figure 5), contributed significantly to the low yield of PmSTl-catalyzed sialylation. In fact, donor hydrolysis has been observed in other glycosyltransferase-catalyzed reactions that lead to lower synthetic yields. The donor hydrolysis were observed frequently in co-crystallization of glycosyltransferases with a corresponding sugar nucleotide donor where its sugar component was usually cleaved off and only the hydrolyzed nucleotide was observed in the substrate binding pocket of the enzyme. Therefore, inert donor derivatives of glycosyltransferases have been commonly applied in the x-ray crystal structure studies of glycosyltransferases. Two recent papers discussed the donor hydrolysis activities of human blood group A and B glycosyltransferases (GTA and GTB) which are Mn^{2+}-dependent and the UDP-Gal hydrolysis activity of GTB is increased in the presence of an acceptor substrate analog. Nevertheless, the effect of donor hydrolysis of glycosyltransferases on glycosylation processes has not been investigated in detail. In addition, no strategy has been reported for improving the yields of glycosyltransferase-catalyzed reactions by decreasing donor hydrolysis activity.

[0119] Asp141 and His311 influence PmSTl donor hydrolysis activity. As shown in Table 4, D141A mutation decreased the efficiency of CMP-Neu5Ac hydrolysis activity of PmSTl by 1,000-fold mainly due to the decrease in the turnover number. H311A mutation also decreased the CMP-Neu5Ac hydrolysis activity by 16-fold, mainly contributed by a decreased turnover number without affecting the binding affinity significantly.
Table 4: Apparent kinetics of the CMP-Neu5Ac hydrolysis activity of WT PmSTI and mutants.

<table>
<thead>
<tr>
<th></th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_m$ (s$^{-1}$ mM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>1.5 ± 0.2</td>
<td>27 ± 1</td>
<td>18</td>
</tr>
<tr>
<td>$^a$D141A</td>
<td>1.4 ± 0.2</td>
<td>(2.5 ± 0.1) × 10$^{-2}$</td>
<td>1.8 × 10$^{-2}$</td>
</tr>
<tr>
<td>$^b$H311A</td>
<td>1.8 ± 0.2</td>
<td>2.1 ± 0.1</td>
<td>1.1</td>
</tr>
<tr>
<td>M144D</td>
<td>7.3 ± 0.5</td>
<td>6.5 ± 0.1</td>
<td>0.89</td>
</tr>
<tr>
<td>M144H</td>
<td>13 ± 3</td>
<td>7.1 ± 6</td>
<td>5.5</td>
</tr>
</tbody>
</table>

*PmSTI D141A and H311A mutants were generated previously. See, Ni, et al. (2006) Biochemistry 45, 2139-2148.

[0120] PmSTI mutants with decreased CMP-Neu5Ac hydrolysis activity. As shown in Table 4, both M144D and M144H mutations decreased the efficiency of donor hydrolysis. M144D mutation decreased the efficiency of donor hydrolysis by 20-fold due to a 4.9-fold increase of the $K_m$ value and a 4.2-fold decrease of the $k_{cat}$ value. M144H mutation caused a less significant 3.3-fold decrease in the efficiency of donor hydrolysis due to a significant 8.7-fold increase in the $K_m$ value which is offset by a 2.6-fold increase in the $k_{cat}$ value.

[0121] a2-3-Sialyltransferase activities of PmSTI mutants. As shown in Table 5, when a good sialyltransferase acceptor 4-methylumbelliferyl $\beta$-lactoside (LacPMU) was used, the M144D mutation decreased the cc2-3-sialyltransferase activity by 18-fold due to a 9-fold increase of $k_{cat}$ value and a 2-fold decrease of $k_{cat}/K_m$ value. When a poor sialyltransferase acceptor Le$^\beta$MU was used, the M144D mutation did not change the efficiency of the a2-3-sialyltransferase activity of PmSTI significantly. In comparison, M144H mutation only decreased the rx2-3-sialyltransferase activity weakly (1.3-fold) when Lac$^\beta$MU was used as an acceptor and increased the efficiency of ct2-3-sialyltransferase activity by 2.6-fold when Le$^\beta$MU was used as an acceptor.
Table 5: Apparent kinetics of the a2-3-sialyltransferase activity of WT PmSTl and mutants.

<table>
<thead>
<tr>
<th></th>
<th>$K^*$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K^*$ (s$^{-1}$ mM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WT</td>
<td>M144D</td>
<td>M144H</td>
</tr>
<tr>
<td>LacβMU</td>
<td>1.4</td>
<td>22±1</td>
<td>2.1±1</td>
</tr>
<tr>
<td>LacβMU</td>
<td>1.4±0.01</td>
<td>2.1±0.1</td>
<td>0.4±0.01</td>
</tr>
<tr>
<td>CMP-Neu5Ac</td>
<td>0.39±0.03</td>
<td>2.1±0.1</td>
<td>0.4±0.05</td>
</tr>
</tbody>
</table>

aData are from Yu, H., et al. (2005) J. Am. Chem. Soc. 127, 1761 8-1 76 19. bWith LacPμU, cWith LeβΜU.

[0122] PmSTl M144D mutant has a decreased a2-3-sialidase activity. M144D and M144H mutations also decreased the a2-3-sialidase activity of PmSTl by 5588- and 594-fold respectively when Neu5Aca2-3LacPμU was used as the sialidase substrate (Table 6). While the PmSTl M144D mutant showed no sialidase activity when Neu5Aca2-3LeβMU was used as the substrate, PmSTl M144H has increased sialidase activity compared to the WT PmSTl using the SLeβ substrate. For example, the PmSTl M144H mutant cleaved 10.0%, 24.5%, and 34.0% of Neu5Ac from Neu5Aca2-3LeβMU in 1 h, 6 h, and 20 h, respectively. In comparison, WT PmSTl removed 2.0%, 7.0%, and 7.5% of Neu5Ac from Neu5Aca2-3LeβMU under the same reaction conditions. The decreased a2-3-sialidase activity by M144D mutation allows the potential application of the PmSTl M144D mutant in sialylation of glycoconjugates containing terminal galactoside or Leβ where the decreased a2-3-sialidase activity has the most advantages as these reactions are challenging for prompt monitoring.
Table 6: Apparent kinetics of the a2-3-sialidase activity of WT PmSTl, M144D, and M144H mutants using Neu5Aca2-3LacBMU as the sialidase substrate.

<table>
<thead>
<tr>
<th></th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_m$ (s$^{-1}$ mM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>$^a$24</td>
<td>$^a2.3 \times 10^7$</td>
<td>$^a9.5$</td>
</tr>
<tr>
<td>M144D</td>
<td>20 ± 2</td>
<td>(3.5 ± 0.1) x $10^{-2}$</td>
<td>1.7 x $10^{-3}$</td>
</tr>
<tr>
<td>M144H</td>
<td>1.7 ± 0.3</td>
<td>(2.7 ± 0.2) x $10^{-2}$</td>
<td>1.6 x $10^{-2}$</td>
</tr>
</tbody>
</table>


[0123] PmSTl M144D mutant has a similar expression level as the WT PmSTl. The PmSTl M144D mutation did not change the enzyme expression level in E. coli. About 98 mg of C-His$_6$-tagged PmSTl M144D protein can be routinely purified from one liter of E. coli cell culture using Ni$^{2+}$-affinity column (Figure 6). This expression level is very similar to that (100 mg) of the WT PmSTl and allows the application of the mutant in preparative and large-scale synthesis of SLe$^x$ antigens.

[0124] X-Ray crystal structure of PmSTl M144D mutant. The structure of the PmSTl M144D mutant with CMP-3F(ax$_{a,b}$)-Neu5Ac was determined to 1.45A resolution with $i\theta$$_f$$_{ave}$ and $R_{free}$ values of 18.7% and 21.5% respectively Table 3. Figure 7 shows the structural comparison between WT PmSTl and M144D mutant with bound CMP donor. Figure 7A shows the overall structure of WT PmSTl with CMP bound (white tubes), aligned with the C-terminal domain of the M144D mutant (grey tubes) also with CMP bound (space filled atoms). Figure 7B shows the stereo view of the superposition near the active site. WT PmSTl is shown as white tubes with bound CMP-3F(ax$_{a,b}$)-Neu5Ac (sticks with white carbon bonds) and lactose acceptor (sticks with dark grey carbon bonds). The M144D mutant is shown as grey tubes with CMP bound (sticks with light grey carbon bonds). Figure 7C shows the active site of the ternary crystal structure of PmSTl (PDB ID: 2HZ) with bound CMP-3F(ax$_{a,b}$)-Neu5Ac and lactose. The mutation site M144 is underlined.

[0125] The structure resides in the open conformation similar to the wild-type structure with no substrate (rmsd of 0.50 A for 385 equivalent a-carbons). However, the M144D structure contains well-ordered electron density in the active site that clearly defines the CMP nucleotide. The sialic acid moiety is disordered, likely due to dynamics and/or multiple conformations in the
open state of the enzyme. In the M144D structure, the CMP moiety does not bind as deeply into
the pocket of the active site as the WT PmST1. The base and ribose are situated about 1.5 and
2.0 Å respectively, farther out of the active site compared to the WT PmST1. In the wild-type
structure, Glu338 forms bidentate hydrogen bond interactions with both the 2’ and 3’ OH of the
CMP ribose. In the M144D structure, an ordered water molecule mediates the interaction
between the ribose and Glu338. The more shallow binding of the donor nucleotide in the M144D
structure does not pull down the β-strand and the ensuing loop that contains Trp270. In
comparison, in the wild-type enzyme, donor-nucleotide binding pulls down a β-strand causing
Trp270 to pop out of the C-terminal domain, where it helps define the acceptor binding site in
the sialyltransferase reaction.

**[0126]** PmST1 M144D mutant is more efficient than M144H mutant in sialylating Le\(^\alpha\).
Overall, the M144D mutation decreased the undesired CMP-Neu5Ac hydrolysis activity
significantly (20-fold) without appreciably changing the efficiency of the 2-3-sialyltransferase
activity when Le\(^\alpha\) was used as an acceptor. As a result, M144D showed an overall improved
activity in sialylation of Le\(^\alpha\) for the formation of sialyl Le\(^\alpha\) (SLe\(^\alpha\)) structures. In comparison,
M144H mutant which has a 3.3-fold decreased CMP-Neu5Ac hydrolysis activity and 2.6-fold
increased 2-3-sialyltransferase activity using Le\(^\alpha\) as an acceptor was less effective for directly
sialylating Le\(^\alpha\).

**[0127]** Synthesis of SLe\(^\alpha\) containing diverse sialic acid forms using PmST1 M144D mutant.

The application of the PmST1 M144D mutant obtained by protein structure-based rational
design in the synthesis of SLe\(^\alpha\) containing diverse naturally occurring and non-natural sialic acid forms
was demonstrated using an efficient one-pot three-enzyme chemoenzymatic synthetic system
(Figure 9). The system contained PmST1 M144D mutant, an *Neisseria meningitidis* CMP-sialic
acid synthetase (NmCSS), and a *Pasteurella multocida* sialic acid aldolase. N-Acetylmannosamine (ManNAc), mannose, and their derivatives were used for in situ synthesis of
CMP-sialic acids and derivatives. Le\(^\alpha\) trisaccharide used as the sialyltransferase acceptor was
synthesized using a one-pot two-enzyme system containing a bifunctional L-fucokinase/GDP-
fucose pyrophosphorylase (FKP) cloned from *Bacteroides fragilis* and a recombinant
*Helicobacter pylori* al-3-fucosyltransferase as shown previously. As shown in Figure 9, SLe\(^\alpha\)
tetrasaccharides containing natural sialic acid forms including N-acetylneuraminic acid
(Neu5Ac), N-glycolylneuraminic acid (Neu5Gc), 2-keto-3-deoxy-D-glycero-D-galacto-
nonulosonic acid (Kdn), as well as 9-O-acetylated Neu5Ac and Neu5Gc were obtained in
excellent (85-93%) to good yields (62-64%). The relatively lower yields for the synthesis of
SLε̈ containing the 9-O-acetyl sialic acid forms were due to the de-O-acetylation process
leading to the formation of non-(9-acetylated SLε̈ oligosaccharides. In addition, SLε̈ containing
non-natural sialic acid forms including those with an N-azidoacetyl group or an azido group at C-
5 or a C-9 azido group were also successfully obtained in excellent yields (84-91%).

[0128] Although the foregoing invention has been described in some detail by way of
illustration and example for purposes of clarity of understanding, one of skill in the art will
appreciate that certain changes and modifications may be practiced within the scope of the
appended claims. In addition, each reference provided herein is incorporated by reference in its
entirety to the same extent as if each reference was individually incorporated by reference.
Where a conflict exists between the instant application and a reference provided herein, the
instant application shall dominate.
WHAT IS CLAIMED IS:

1. An isolated glycosyltransferase, wherein the amino acid corresponding to position 120 of SEQ ID NO:1 is any amino acid other than M, the amino acid corresponding to position 247 of SEQ ID NO:1 is any amino acid other than E, or the amino acid corresponding to position 289 of SEQ ID NO:1 is any amino acid other than R, wherein the glycosyltransferase has decreased α2-3 sialidase or donor substrate hydrolysis activity compared to a control glycosyltransferase, wherein the amino acid of the glycosyltransferase corresponding to position 120 of SEQ ID NO:1 is M, the amino acid corresponding to position 247 of SEQ ID NO:1 is E, and the amino acid corresponding to position 289 of SEQ ID NO:1 is R, and wherein the glycosyltransferase is a member of the glycosyltransferase family 80 (GT80).

2. The isolated glycosyltransferase of claim 1, wherein the isolated glycosyltransferase has decreased α2-3 sialidase activity, and the amino acid of the glycosyltransferase corresponding to position 247 of SEQ ID NO:1 is any amino acid other than E, or the amino acid of the glycosyltransferase corresponding to position 289 of SEQ ID NO:1 is any amino acid other than R.

3. The isolated glycosyltransferase of any of claims 1-2, wherein the ratio of α2-3 sialidase activity for the control glycosyltransferase to the α2-3 sialidase activity of the isolated glycosyltransferase is at least 5:1.

4. The isolated glycosyltransferase of claim 3, wherein the ratio is at least 10:1.

5. The isolated glycosyltransferase of claim 3, wherein the ratio is at least 100:1.
6. The isolated glycosyltransferase of claim 3, wherein the ratio is at least 1000:1.

7. The isolated glycosyltransferase of any of claims 1-6, wherein the isolated glycosyltransferase comprises:
   the amino acid corresponding to position 247 of SEQ ID NO:1 is any amino acid other than E, and
   the amino acid corresponding to position 289 of SEQ ID NO:1 is any amino acid other than R.

8. The isolated glycosyltransferase of any of claims 1-7, wherein the isolated glycosyltransferase has decreased donor substrate hydrolysis activity, and wherein the amino acid corresponding to position 120 of SEQ ID NO:1 is any amino acid other than M.

9. The isolated glycosyltransferase of claim 8, wherein the ratio of donor substrate hydrolysis activity for the control α2-3 sialidase to the donor substrate hydrolysis activity of the isolated glycosyltransferase is at least 2:1.

10. The isolated glycosyltransferase of any of claims 1-9, wherein the amino acid corresponding to position 120 of SEQ ID NO:1 is any amino acid selected from the group consisting of V, I, L, F, W, P, S, T, A, G, C, Y, N, Q, D, E, K, R, and H.

11. The isolated glycosyltransferase of any of claims 1-10, wherein the amino acid corresponding to position 247 of SEQ ID NO:1 is any amino acid selected from the group consisting of V, I, L, M, F, W, P, S, T, A, G, C, Y, N, Q, D, K, R, and H.

12. The isolated glycosyltransferase of any of claims 1-11, wherein the amino acid corresponding to position 289 of SEQ ID NO:1 is any amino acid selected from the group consisting of V, I, L, M, F, W, P, S, T, A, G, C, Y, N, Q, D, E, K, and H.

13. The isolated glycosyltransferase of any of claims 1-12, wherein the amino acid corresponding to position 120 of SEQ ID NO:1 is D, E, H, K or R, the amino acid corresponding to position 247 of SEQ ID NO:1 is F, Y or W, or
the amino acid corresponding to position 289 of SEQ ID NO:1 is Y, F or W.

14. The isolated glycosyltransferase of any of claims 1-13, wherein the amino acid corresponding to position 120 of SEQ ID NO: 1 is D or H, the amino acid corresponding to position 247 of SEQ ID NO:1 is F, or the amino acid corresponding to position 289 of SEQ ID NO:1 is Y.

15. The isolated glycosyltransferase of any of claims 1-14, wherein the glycosyltransferase is an α2-3 sialyltransferase.

16. The isolated glycosyltransferase of claim 15, comprising a motif in the sialyltransferase domain comprising at least one member selected from the group consisting of sialyltransferase motif A (YDDGS) and sialyltransferase motif B (KGH).

17. The isolated glycosyltransferase of any of claims 1-16, wherein the control glycosyltransferase is SEQ ID NO:1.

18. The isolated glycosyltransferase of claim 17, wherein the glycosyltransferase comprises a polypeptide sequence having at least 80% sequence identity to SEQ ID NO: 1.

19. The isolated glycosyltransferase of claim 1, wherein the isolated glycosyltransferase comprises a polypeptide sequence selected from the group consisting of SEQ ID NO: 3 (M120D), SEQ ID NO: 5 (M120H), SEQ ID NO: 7 (E247F), SEQ ID NO: 9 (R289Y) and SEQ ID NO: 11 (E247F/R289Y).

20. A recombinant nucleic acid encoding an isolated glycosyltransferase of any of claims 1-19.

21. A cell comprising an recombinant nucleic acid of claim 20.

22. The cell of claim 21, wherein the cell is selected from the group consisting of bacteria, yeast, insect, mammalian and plant cells.
23. A method of preparing an oligosaccharide, the method comprising:
   a) forming a reaction mixture comprising an acceptor sugar, a donor substrate
      comprising a sugar moiety and a nucleotide, and the glycosyltransferase of any of
      claims 1-19, under conditions sufficient to transfer the sugar moiety from the
      donor substrate to the acceptor sugar, thereby forming the oligosaccharide.

24. The method of claim 23, wherein the glycosyltransferase comprises a
    polypeptide sequence selected from the group consisting of SEQ ID NO: 3 (M120D),
    SEQ ID NO: 5 (M120H), SEQ ID NO: 7 (E247F), SEQ ID NO: 9 (R289Y) and SEQ ID NO: 11
    (E247F/R289Y).

25. The method of claim 23, wherein the glycosyltransferase comprises a
    polypeptide sequence selected from the group consisting of SEQ ID NO: 3 (M120D) and
    SEQ ID NO: 5 (M120H).

26. The method of claim 23, wherein the isolated glycosyltransferase
    comprises a polypeptide sequence selected from the group consisting of SEQ ID NO: 7
    (E247F), SEQ ID NO: 9 (R289Y) and SEQ ID NO: 11 (E247F/R289Y).

27. The method of claim 23, wherein the donor substrate comprises a cytidine
    5'-monophosphate(CMP)-sialic acid.

28. The method of claim 27, wherein the CMP-sialic acid comprises cytidine
    5'-monophosphate N-acetylneuraminic acid (CMP-Neu5Ac) or a CMP-Neu5Ac analog.

29. The method of claim 28, further comprising:
   b) forming a reaction mixture comprising a CMP-sialic acid synthetase, cytidine
      triphosphate, and N-acetylneuraminic acid (Neu5Ac) or a Neu5Ac analog, under
      conditions suitable to form the CMP-Neu5Ac or CMP-Neu5Ac analog.

30. The method of claim 29, wherein steps a) and b) are performed in one pot.
31. The method of claim 29, further comprising:
   c) forming a reaction mixture comprising a sialic acid aldolase, pyruvic acid or
derivatives thereof, and N-acetylmannosamine or derivatives thereof, under
conditions suitable to form the Neu5Ac or Neu5Ac analog.

32. The method of claim 31, wherein steps a), b), and c) are performed in one
pot.

33. The method of any of claims 23-32, wherein the oligosaccharide is an α2-3-linked
sialyloligosaccharide.

34. The method of claim 23, wherein the oligosaccharide is a fucosylated
oligosaccharide.
FIG. 1

FIG. 2
FIG. 3

Relative activity (%)

WT
E271F/R313Y

FIG. 4

% Conversion

(86.5%)
(99.2%)
(97.8%)
(99.5%)
(90.1%)
(88.3%)

Time (min)
FIG. 7