Investigation of *Micromonospora viridifaciens* Mutant Sialidases in Transglycosylation and Novel Enzymatic Activity

by

Lydia Lan-Yun Cheng

B.Sc. (Hons.), Simon Fraser University, 2006

Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Doctor of Philosophy

in the
Department of Molecular Biology and Biochemistry
Faculty of Science

SIMON FRASER UNIVERSITY

Spring 2014

Approval

Name: **Lydia Lan-Yun Cheng**

Degree: **Doctor of Philosophy (Science)**

Title of Thesis: ***Investigations of Micromonospora viridifaciens Mutant Sialidases in Transglycosylations and Novel Enzymatic Activity***

Examining Committee:

Chair: Dr. Lisa Craig, Associate Professor

Dr. Andrew Bennet
Senior Supervisor, Department of Chemistry
Professor

Dr. Dipankar Sen
Supervisor, Department of MBB
Professor

Dr. Mark Paetzel
Supervisor, Department of MBB
Professor

Dr. Thor Borgford
Internal Examiner
Adjunct Professor
SFU/Department of Chemistry and MBB

By written/e-mail consultation (Halifax, NS)

Dr. Stephen L. Bearne
External Examiner
Professor, Department of Chemistry and MBB
Dalhousie University

Date Defended/Approved: December 05, 2013

Partial Copyright Licence

Abstract

Sialic acid is a monosaccharide with a nine-carbon atom backbone and there are more than fifty known derivatives of sialic acid depending on the substitution pattern, and these compounds make up the sialic acid family. The most widespread sialic acid family members are *N*-acetylneuraminic acid, *N*-glycolylneuraminic acid and Kdn. Sialic acid family members usually reside at the termini of sugar branch of glycoproteins and glycolipids in animals. As a result, the location of these sugars contributes to their roles in cellular and molecular recognition events. The present thesis focused on the biological mask function of sialic acid in the clearance and regulation of serum glycoconjugate lifetimes. Attaching sialic acid directly to terminal sugar residues can block the underlining galactose residues from recognition by the Ashwell receptors and tailor the glycoproteins clearance rate. Therefore, if sialic acid can be directly attached to recombinant glycoprotein drugs this will potentially decrease glycoprotein clearance rates; thereby allowing the glycoprotein drug's dosage to be reduced.

The present thesis focused on studying the ability of mutant *Micromonospora viridifaciens* sialidases to transfer sialic acid onto other sugar acceptors. The wild-type sialidase favors hydrolysis rather than transfer of sialic acid. We show that the Y370G mutant sialidase is flexible in transferring sialic acid onto several different sugar acceptors. To try and improve the yield for the transglycosylation reaction of the Y370G mutant sialidase, hydrophobic loops, which were incorporated on the enzyme to try and increase the local concentration of acceptor. The resultant loop mutants abolished the enzyme's activity. Therefore, we targeted mutagenesis of the loop region to create a library of Y370G-loop mutants. Unfortunately, screening of this mutant library did not yield any promising next generation Y370G mutants. We also screened another mutant library based on two tyr370 nucleophile mutants, histidine (Y370H) and methionine (Y370M), with the goal of finding novel mechanisms of action. However, screening of this library produced several revertant mutants, but no promising candidates with either histidine or methionine as the enzymatic nucleophile.

Keywords: sialic acid; sialidase; transglycosylations; Y370G; library

Dedication

Dedicated to my family, Mom, Dad, Brother, Husband, and Daughter

Acknowledgements

First of all I would like to thank my senior supervisor, Dr. Andrew Bennet for giving me the opportunity to pursue a PhD degree in his laboratory. He is always supportive and encouraging when experiments are not going smoothly; extraordinary excited when results are positive; cracking jokes when silly mistakes were made; and last, but not least, using lots of red ink to correct my writin. I really learned a lot throughout my graduate school year so I want to thank again to Dr. Andrew Bennet for giving me the interesting projects and for motivating my passion towards research.

I also want to thank Dr. Jacqueline Watson for guidance when I first started in the Bennet laboratory as an undergraduate and in the beginning of the graduate school year. Her perfect organization in both sample storage and lab book records allowed me to continue many of her projects. I want to thank my committee members Dr. Dipankar Sen and Dr. Mark Paetzel for the useful comments and discussions throughout the yearly committee meetings. I also appreciate the time they spent to write my reference letters and read my thesis. I want to thank both the external examiner Dr. Stephen L. Bearne and internal examiner Dr. Thor Borgford for willing to spend time to read my thesis and provide knowledgeable questions.

Of course I want to thank the Bennet lab members both past and present members that overlap with my graduate school year. We shared a lot of memories together, the success, failure, achievements and stress. Especially thanks to the Bennet lab girls, Dr. Shidmoossavee, Dr. Chakladar, Ms. Cerda, Ms. Kobra and Ms. Shamsi. We had a lot of cake, tea or coffee parties together. You all were more than friends to me. I felt very blessed to have a lot of wonderful people around me in my PhD years. Thanks to all.

Table of Contents

List of Figures

List of Tables

List of Figures S

List of Schemes

1. Introduction

1.1. Glycosidases

Glycosidases, which are also called glycosyl hydrolases (E.C. 3.2.1.X), are enzymes that catalyze the hydrolysis of carbohydrate *O*-, *N*- or *S*-glycosidic bonds (Figure 1.1). Glycosidases are categorized into ~132 families based on their primary amino acid sequences. These different families of glycosidases are compiled on the CAZy (CArbohydrate-Active EnZymes) web site (www.cazy.org). Glycosidases may be grouped into exo or endo acting enzymes, a classification that depends on the position of hydrolysis on the glycan chain. If the enzyme-catalyzed hydrolysis takes place at the non-reducing terminal end of the glycan chain it is called an exo-acting enzyme; whereas, if the hydrolysis occurs in the middle of a glycan chain it is called an endo-glycosidase. Glycosyl hydrolases can also be classified as either retaining or inverting enzymes depending on the stereochemical outcome of the hydrolysis reaction. If the stereocentre of the substrate and product are the same, the enzymes are referred to as retaining enzymes; if the configuration of the stereocentres are different, the enzyme are inverting glycosidases.

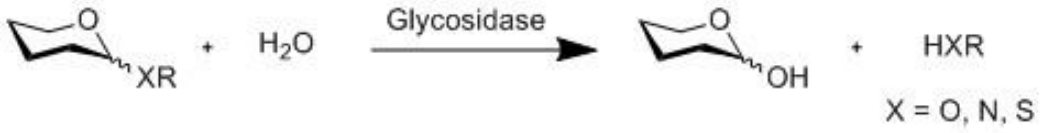


Figure 1.1 *General scheme of glycosidase hydrolysis of glycoside to glycohemiacetal and aglycon.*
The wavy bond at the anomeric position represents either alpha or beta configurations. The carbohydrate covalently linked to R group refers to chemical species such as proteins, peptides, lipids and saccharides.

1.1.1. Glycosylhydrolases

1.1.1.1. Inverting Glycosidases

Inverting glycosidases hydrolyze glycosidic bonds with a net inversion of configuration at the anomeric centre. The generally accepted mechanism is a one step single-displacement reaction in which an oxacarbenium ion-like transition state is involved. In general, two catalytic residues, which are glutamic and/or aspartic acids having carboxylic side chains, act as a Brönsted acid and a Brönsted base to promote the hydrolysis reaction. The two catalytic residues located within active site are typically 6-11 Å apart (67). Shown in Figure 1.2 is the proposed mechanism for hydrolysis by an inverting β-glycosidase.

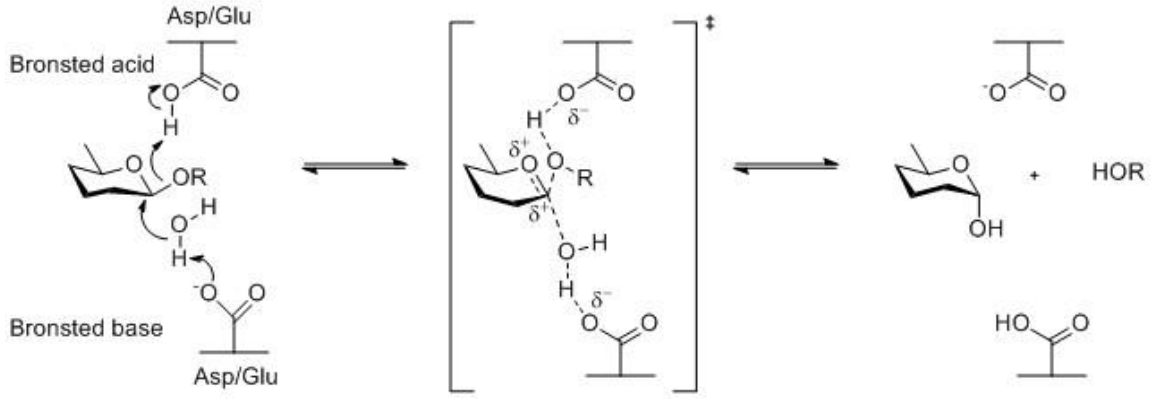


Figure 1.2 *Proposed Inverting mechanism of β-glycosidase. Two enzymatic residues involved in catalysis are typically carboxylate group that act as acid and base respectively. The one step hydrolysis is proposed through an oxicarbonium-like transition state with aglycon and incoming nucleophile bond at the anomeric position breaking and forming at the same time. The resultant anomeric position is inverted.*

1.1.1.2. Retaining Glycosidases

Retaining glycosidases are enzymes that hydrolyze glycosides with net retention of configuration at the anomeric centre. The most common mechanism, which was proposed by Koshland in 1953, involves a double displacement mechanism that occurs via a glycosyl-enzyme intermediate (55). The two enzymatic residues that are involved in catalysis are usually carboxylate groups, glutamic acid and/or aspartic acid that are

located ~5.5 Å apart (Figure 1.3). In the first step, one carboxylate group acts as a nucleophile attacking the anomeric centre while the other carboxylate acid group assists departure of the aglycon by acting as a Brönsted acid to yield a glycosyl-enzyme intermediate by way of an oxocarbenium ion-like transition state. The subsequent nucleophilic attack at the anomeric centre occurs with the aid of carboxylate side chain acting as a Brönsted base to deprotonate a water. The glycosyl product is produced through an oxocarbenium ion-like transition state.

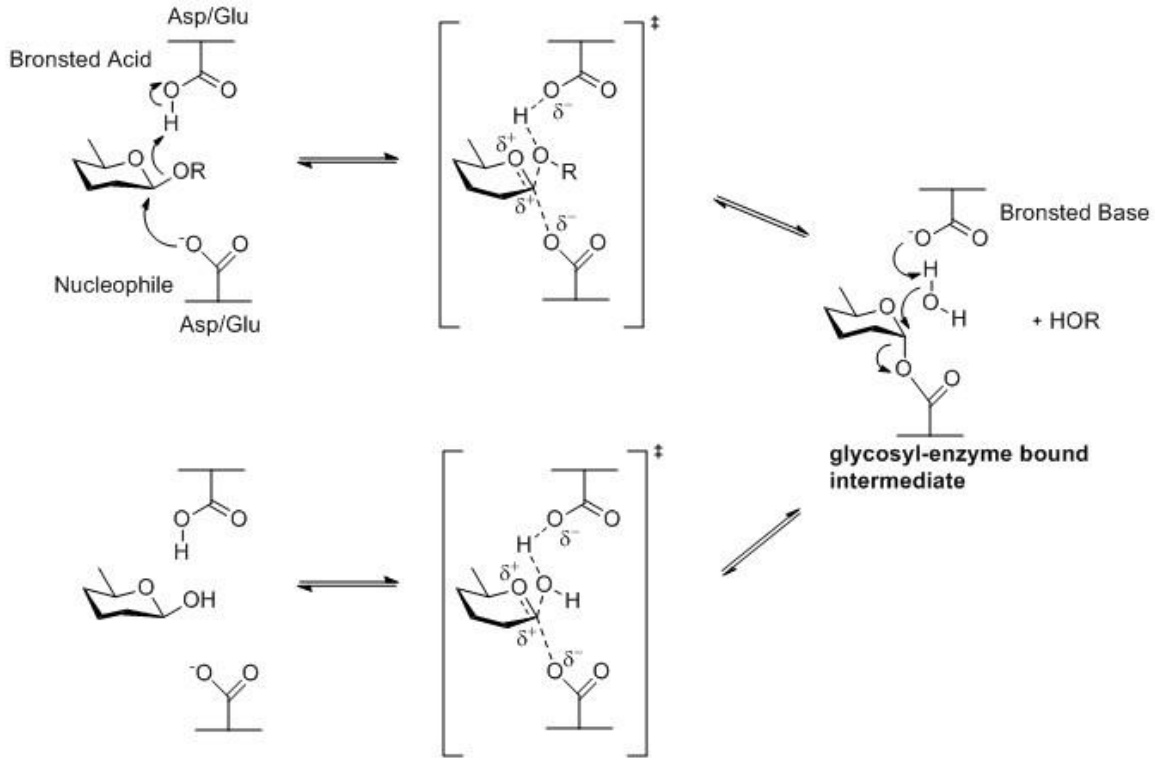


Figure 1.3 *Retaining mechanism of β-glycosidases.*
Two catalytic residues are commonly found to be carboxylate side chain which act as a nucleophile and the other as an acid/base. The catalysis is proposed to be two-step with each step inverting anomeric configuration, for a net retention of stereochemistry.

1.1.2. Glycosyltransferases

Glycosyltransferases (E.C. 2.4) are enzymes that catalyze the transfer of activated sugars to an acceptor molecule with the formation of a glycosidic bond (Figure 1.4). Glycosyltransferases are grouped into different families based on their amino acid sequence similarity, substrate specificities or stereochemical outcome of the coupling reaction (59, 94). Currently, there are 94 glycosyltransferase families that are defined based on their amino acid sequences and these are listed online on the CAZy webpage. The glycosyl donor substrates are generally sugar mono- or di-phosphate groups linked to either nucleotide, lipid, protein, sugar, antibiotic, or other small molecules (59, 94). The glycosyltransferases transfer glycosyl groups to acceptors with the stereochemistry at the anomeric centre being either retained or inverted in the glycosidic product.

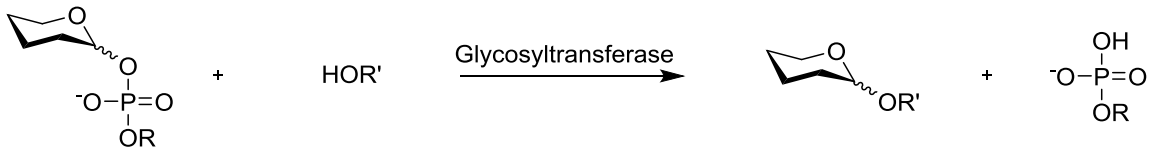


Figure 1.4 *Glycosyltransferase reactions.*
The enzymes catalyze the transfer of monosaccharide moieties from an activated nucleotide sugar that is either mono-, di- or tri- phosphate linked, to a glycosyl acceptor molecule.

1.1.2.1. Inverting Glycosyltransferases

The glycosidic bond formation and stereochemistry of the glycan as carried out by inverting glycosyltransferases are shown in Figure 1.5 below. Inverting glycosyltransferases are believed to catalyze the transfer reaction by a similar mechanism to that used by inverting glycosyl hydrolases. That is, a carboxylate group in the active site acts as a Brönsted base to assist the nucleophilic attack of an acceptor on the anomeric carbon. The reaction proceeds via a single step, direct displacement S_N2-like reaction with an oxocarbenium ion-like transition state to give an inversion in configuration.

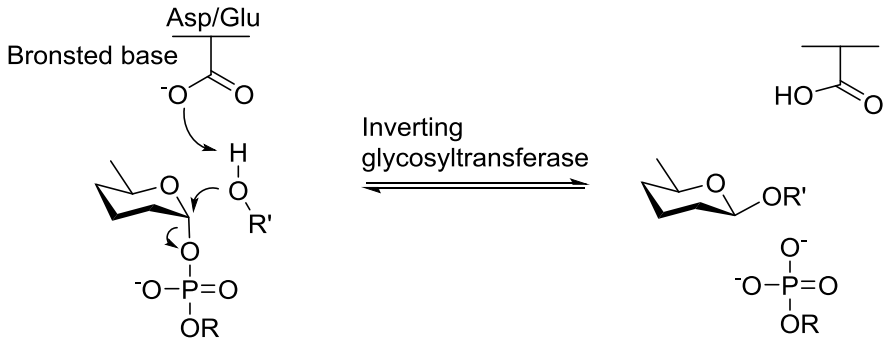


Figure 1.5 *Inverting mechanism of glycosyltransferase.*
The mechanism of action is proposed to follow a single displacement step similar to inverting glycosidases; therefore, with the net inversion configuration at the anomeric position.

1.1.2.2. Retaining Glycosyltransferases

The glycosidic bond formation and stereochemistry of the glycan resulting from catalysis by retaining glycosyltransferases are shown in Figure 1.6 below. The mechanism of retaining glycosyltransferases is less well understood in comparison to those for glycosyl hydrolases because assays are more complex and many of these enzymes are membrane associated making them difficult to purify in large quantity (37). The transition state for retaining glycosyltransferases are thought to have oxocarbenium ion character as well and the most recent research article suggests the mechanism of nucleophilic attack is from the same face as the leaving group departure called the S_Ni-like mechanism (60). The mechanism involves a front-face nucleophilic attack with possibly the formation of a hydrogen bond between the leaving group and the incoming nucleophile to assist the dissociation of the leaving group (60).

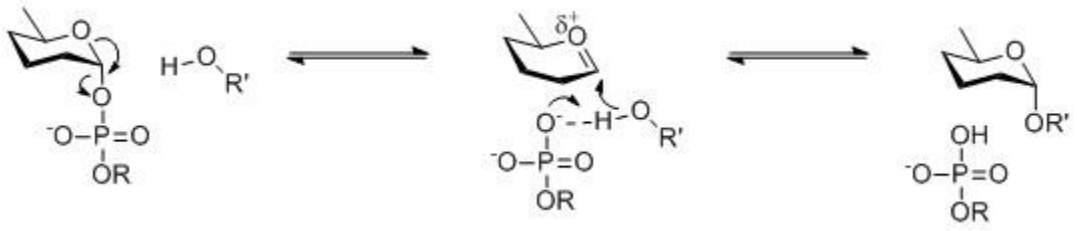


Figure 1.6 *The reaction of retaining glycosyltransferases.*
The proposed mechanism of action is an internal return S_{Ni}-like mechanism instead of two-step displacement that retaining glycosidase proposed.

1.2. Protein Glycosylation

Glycosylation is the enzymatic process of attaching carbohydrates to proteins, lipids, or other organic molecules. In eukaryotes, protein glycosylation is a post-translational modification that takes place in the endoplasmic reticulum (ER) and in the Golgi body. There are five main types of glycosylation and these are classified by the type of glycosidic linkage. These consist of *N*-, *O*- and *C*-linked glycosylation, glycosylphosphatidylinositol (GPI) anchored and phosphoglycosylation. *N*-Linked refers to glycan attached to the amino group of asparagines, *O*-linked are glycans bound to the hydroxyl group of serine or threonine, *C*-linked involve mannose covalently linked to the indole ring of a tryptophan residue. The glycan core of GPI-anchors connects to the phospholipid protein. The two types of glycosylation *N*- and *O*-linked will be discussed in more detail.

1.2.1. Protein N-Glycosylation

N-Linked glycosylation occurs when the connection between an oligosaccharide chain involves a glycosidic linkage to the side chain nitrogen atom of an asparagine amino acid residue in a protein. In eukaryotes, *N*-glycosylation is necessary for maturation of eukaryotic proteins and the steps involved in trafficking *N*-glycosylated proteins are best understood in yeast. Although glycan structures that decorate the protein surface are diverse. the initial glycosylation event for each protein involves a single distinctive glycan and this core glycan is the precursor for further modifications (Figure 1.7).

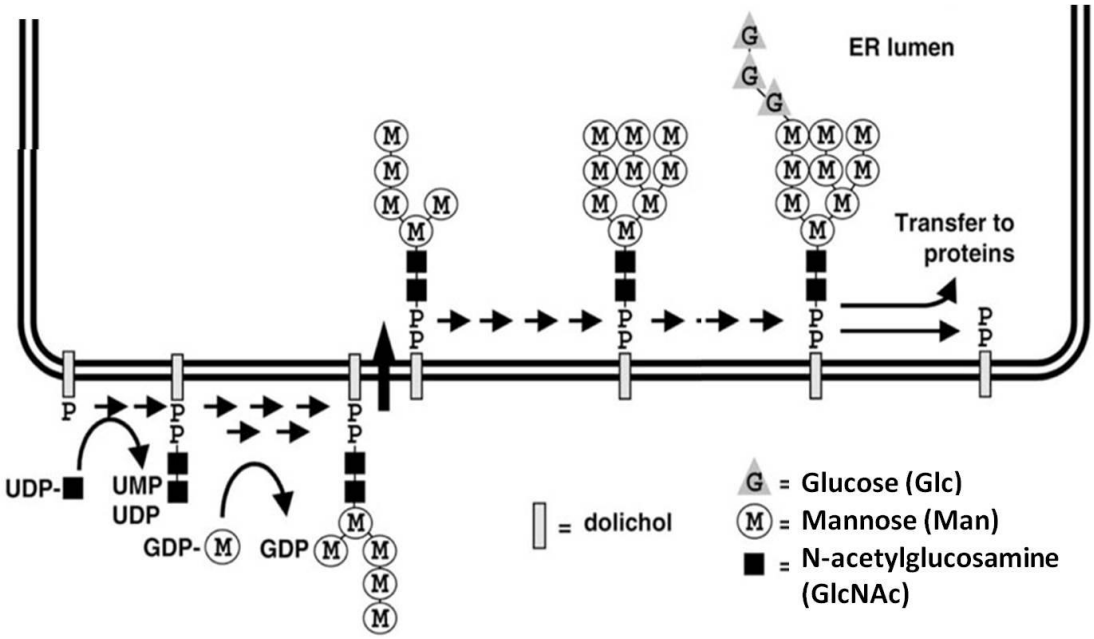


Figure 1.7 *ER Protein N-glycosylation in eukaryotes.*
The glycans for N-linked protein glycosylation is biosynthesized on dolichol phosphate with stepwise addition of carbohydrate by specific glycosyltransferase. The $Man_5(GlcNAc)_2$ glycan chain is made on the cytosolic face, which is then flipped to the luminal face of the ER membrane. The addition of carbohydrates on the glycans continued and the $Glc_3Man_9(GlcNAc)_2$ precursor is transferred to nascent proteins.

The building up of glycans starts from phosphorylation of dolichol, a lipid compound synthesized by the same metabolic route as cholesterol, to form dolichol phosphate. From the dolicholpyrophosphate base, the sequential addition of carbohydrate moieties then takes place on the cytosolic side of the endoplasmic reticulum (ER) using specific carbohydrate donors and glycosyltransferases enzymes to catalyze the formation of dolichol-$P_2GlcNAc_2Man_5$ (11). The dolichol-linked glycan then flips to the lumen side of the ER compartment with the assistance of the protein RAFT1 for further modification of the glycan. The final core glycan $GlcNAc_2Man_9Glc_3$ is transferred to an asparagine residue of the target protein that has the sequence Asn-x-Ser/Thr, where x is any amino acid (68). The final glycan composition varies because glycan protein maturation involves removal and/or addition of sugar moieties in both the ER and the Golgi body compartments. The resultant glycosylation patterns are diverse because of the different carbohydrates added, the sites of attachment, and the final

length of the glycan. Thus, glycosylation is a process that is called a complex post-translational modification.

The maturation processes include proper protein folding, disulfide bond formation, posttranslational folding, and assembly. The ER has a quality control system that involves the molecular chaperones Calnexin/Calreticulin that assist and monitor the folded status of the glycoprotein (90)(36). The action of glycosyl hydrolase and glycosyltransferase enzymes on these glycans modifies them to become high-mannose, hybrid, or complex oligosaccharide structures in the Golgi (Figure 1.8). The high mannose structures serve as a substrate to allow a large diversity of *N*-glycan subtypes, indicated by arrows of the figure below, to be made.

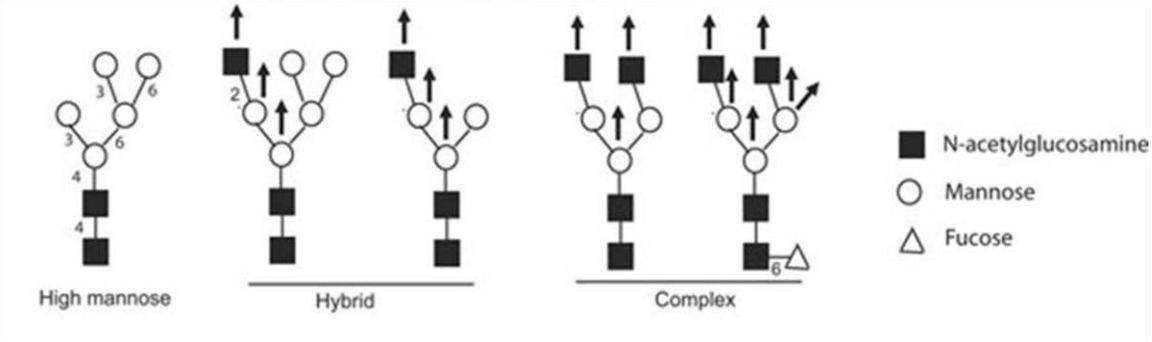


Figure 1.8 *Major types of N-glycans in vertebrates.*
Three N-glycans cores, high mannose, hybrid and complex can be made. The numbers beside the sugar blocks refer to the linkage from carbon-1 of sugar to the next carbon numbers positions of the other sugar. The arrows in the hybrid and complex glycan types mean the possible extension of sugar chains beyond the core structures.

Glycans also play other important roles besides the critical contribution of protein folding, maturation, and quality control. Studies show that *N*-glycans can increase protein stability, solubility and increase resistance to digestive proteases (123)(58). The position of glycans located on the outer surface of glycoproteins and glycolipids also has a function in binding and signaling purposes (42).

1.2.2. Protein O-Glycosylation

O-Linked glycosylation refers to glycans attached to the oxygen atom of an amino acid residue in a protein (95). The *O*-linked glycan can be classified as either a mucin or a non-mucin type (113). Mucin glycans are sugar chains in which an *N*-acetylgalactosamine (GalNAc) residue β-linked to the –OH side chain of the serine or threonine residue, that does not have a consensus amino acid sequence. Non-mucin glycans are polysaccharide chains containing to fucose, glucose, *N*-acetylglucosamine, galactose, xylose, and mannose (113).

Mucin type *O*-linked glycosylation usually occurs in the Golgi body in eukaryotes and is catalyzed by an *N*-acetylgalactosaminyltransferase (E.C. 2.4.1.41). The GalNAc residue that is attached to a serine/threonine is further elongated and modified by various transferases to yield eight different core *O*-linked glycans (Table 1.1) (45).

Table 1.1 *The mucin O-linked glycosylation core regions. The most common O-GalNAc glycan is core 1, Galβ1-3GalNAc-, and it is found in many glycoproteins and mucins.*

Core Type	Core Sequence
1	Galβ1-3GalNAc-
2	GlcNAcβ1 6 GalNAc- 3 Galβ1
3	GlcNAcβ1-3GalNAc-
4	GlcNAcβ1 6 GalNAc- 3 GlcNAcβ1
5	GalNAcα1-3GalNAc-
6	GlcNAcβ1-6GalNAc-
7	GalNAcα1-6GalNAc-
8	Galα1-3GalNac-

Starting with these core structures, the glycans can be further modified by the addition of galactose, GlcNAc, and sialic acid residues, which increases the structural diversity. Such *O*-linked glycosylation structures were found to cluster over a short stretch of peptide and therefore, they can play vital roles in multivalent recognition by acting as a structural motif for binding (45). Common examples include the different *O*-linked carbohydrate antigens presented on red blood cell surfaces and these are the significant determinate for our ABO blood system and RhD positive or negative status (80)(70). Also, sialyl Le^X, a tetrasacchride carbohydrate attached to *O*-glycans, functions as a ligand for E-selectin and this interaction triggers the inflammatory response (78). In general, *O*-linked sugar chains function in cell-cell interactions, protease resistance, antibody recognition, and as other possible functional determinants (45).

1.2.3. Glycosylation Patterns in Different Organisms

Unlike the genetic code, which has a universal function among organisms, glycosylation patterns that are modified through the secretory pathway of different taxa of living organisms vary (Figure 1.9) (4). One of the major reasons for scientists to want to understand diverse sugar structures is related to the biotechnology industry with the goal of producing glycoproteins to treat human diseases (4). Glycan structures on therapeutic proteins can influence pharmacokinetics, bioactivity, secretion, *in vivo* clearance, solubility, recognition and antigenicity (96, 116).

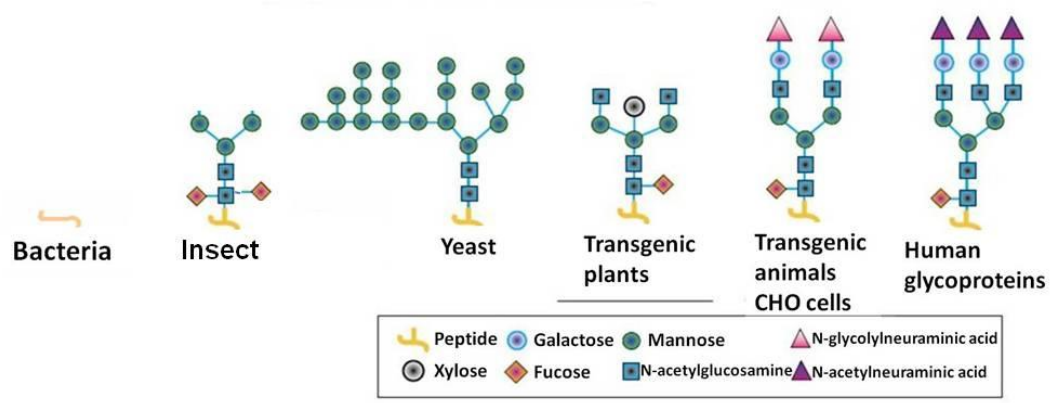


Figure 1.9 *N-linked glycosylation patterns in different cell types. The N-linked glycosylation process occurs in eukaryotes and archaea but rarely in bacteria. Proteins expressed in different systems have different glycosylation patterns. Each of the expression system has different glycosylation forms as well. The pattern shown in this figure is the general pattern.*

Currently, mammalian cells are used widely for the production of therapeutic glycoproteins. The common cell lines for production of glycoproteins that are approved for use in humans are Chinese hamster ovary (CHO) and baby hamster kidney (BHK) cells (78, 102). The major reason for using mammalian cell lines for glycoprotein production is their *in vivo* capability of producing similar glycosylation patterns to those of humans (54). However, disadvantages of using mammalian cell lines for this task include potential human virus contamination, heterogeneous mixture of different glycan structures and cost (102).

The use of other expression systems for producing glycoprotein therapeutics are constantly being investigated because of the potential for higher yields and lower cost. Even though *E. coli* was the first organism used for gene expression of recombinant proteins, introduced between 1972 – 1973, the lack of a glycosylation machinery makes glycoprotein production in this organism impossible (63). Other alternatives include yeast, insect, and plant cell lines that share similar *N*-glycosylation features to those of mammals (48). Nevertheless, the sugar chain structures diverge in different cell lines during processing in the Golgi body (38). Two general approaches are being investigated for producing “humanized” glycoproteins. One focus is on the *in vivo* installation or modification of the glycosylation machinery, while a separate approach is

focused on *in vitro* glycan addition or modification (47)' (65). In the current thesis, chapter 2 focuses on the *in vitro* modification approach and involves mutation of a sialidase enzyme so that it transfers sialic acid onto mannose sugars, with the goal of ultimately transferring this sugar onto glycoproteins produced in yeast.

1.3. Sialic Acids

Sialic acid is also named *N*-acetylneuraminic acid and is a monosaccharide with a nine-carbon backbone (113). The term "neuraminic acid" was termed by German scientist Ernst Klenk in 1941 following his discovery of a sugar cleavage product from brain lipids (53). The name "sialic acid" (from the Greek for *sialos* means saliva) was first introduced in 1952 for a sugar identified in his study of salivary mucins by Swedish biochemist Gunnar Blix (10).

The structure of one of the most abundant sialic acids, *N*-acetylneuraminic acid or Neu5Ac (5-acetamido-3,5-dideoxy-D-*glycero*-D-*galact*-onon-2-ulosonic acid) is shown below (Figure 1.10). The stereochemistry of α or β at the C2 chiral center is assigned based on the configurational relationship between the hemiacetal ring closure anomeric centre C2 to the highest-numbered chiral centre C7. When an oxygen atom at these two chiral centres are on the same side, *cis*-orientation, the α-configuration is assigned; whereas, when the oxygen atoms are at the opposite side, *trans*-orientation, the β-configuration is appointed. In nature, sialic acid is found only in the α-configuration when it is a component of a glycoconjugate, and in the β-configuration in the high-energy donor form, CMP-sialic acid.

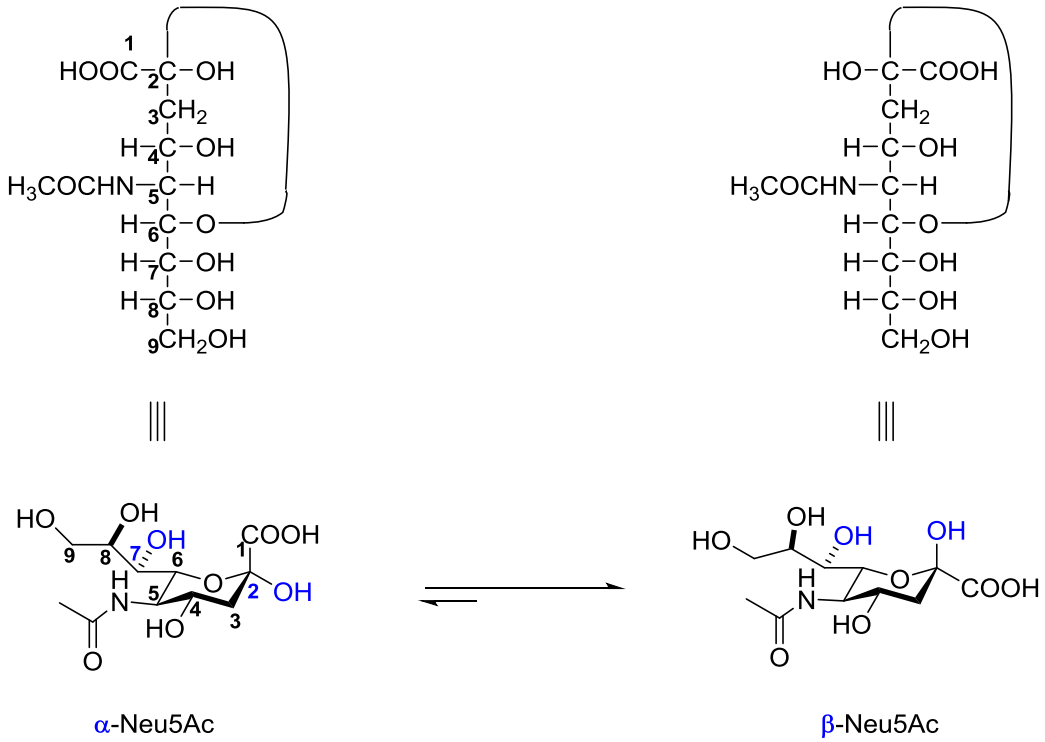


Figure 1.10 *Cyclized forms of N-acetylneuraminic acid.*
The hydroxyl group shown in blue at C2 and C7 determine α- or β-configuration when they are in cis- or trans- position respectively. The α-anomer is commonly found when Neu5Ac is bound to glycans; however, in solution more than 90 % of Neu5Ac mutarotate to β-anomer.

1.3.1. Structure Diversity

Sialic acids are α-keto acids in which various modifications generate a diverse family of more than 50 derivatives (5). Neuraminic acid, Neu is the unsubstituted form of sialic acid and modifications at the C-5 position include *N*-acetyl, *N*-glycolyl and substitution of the amino to a hydroxyl group; and these widespread forms are called, *N*-acetylneuraminic acid (Neu5Ac), *N*-glycolylneuraminic acid (Neu5Gc), and keto-deoxyneuraminic acid (Kdn), respectively (Figure 1.11) (10)(46). Sialic acids are widely distributed. For example, *N*-acetylneuraminic acids are found in vertebrates, echinoderms, insects, protozoa, fungi and bacteria; *N*-glycolylneuraminic acids are found in vertebrates, protozoa and fungi; and Kdns are found in vertebrates and bacteria(5). Additional diversity of sialic acid family members occurs by the combinations of these

three structures with modifications of the hydroxyl groups at C4, C7, C8 and C9 by acetate, lactate, sulfate, or phosphate esters, or by methyl ethers (5).

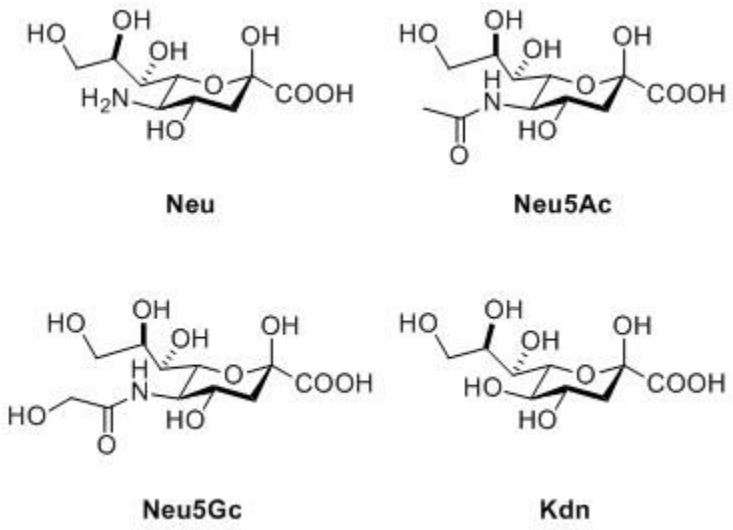


Figure 1.11 *Structure of sialic acids.*
Neuraminic acid (Neu), N-acetylneuraminic acid (Neu5Ac), N-glycolylneuraminic acid (Neu5Gc) and 2-keto-3-deoxyl-D-glycero-D-galacto-nonulosonic acid (Kdn) are different derivative of sialic acid that only different at C-5 position. The original definition of sialic acid raised from the structure of Neu and its derivatives and the most common forms of sialic acids are Neu5Ac, Neu5Gc and Kdn.

Sialic acids are commonly found α-linked at the outermost end of *N*-glycans, *O*-glycans and glycosphingolipids. The specific linkage and the underlying carbohydrate also increase the diversity of the glycans. The typical linkages found to the C2 carbon of sialic acid are to galactosyl C3 or C6 hydroxyl groups, the C6 primary hydroxyl of an *N*-acetylgalactosamine residue and C8 of another sialic acid residue (105). Other less common glycosidic bonds occur to C3 of *N*-acetylgalactosamine, C6 of *N*-acetylglucosamine and glucose, C4 of galactose and *N*-acetylgalactosamine, and C9 linked to another sialic acid (39, 57).

1.3.2. Biosynthetic Pathway

In the 1960s, the general biosynthetic pathway for Neu5Ac synthesis and glycosylation in both bacteria and vertebrates was elucidated (82). The steps for producing Neu5Ac in both bacteria and vertebrates are compared and pathways are shown in Figure 1.12 (82).

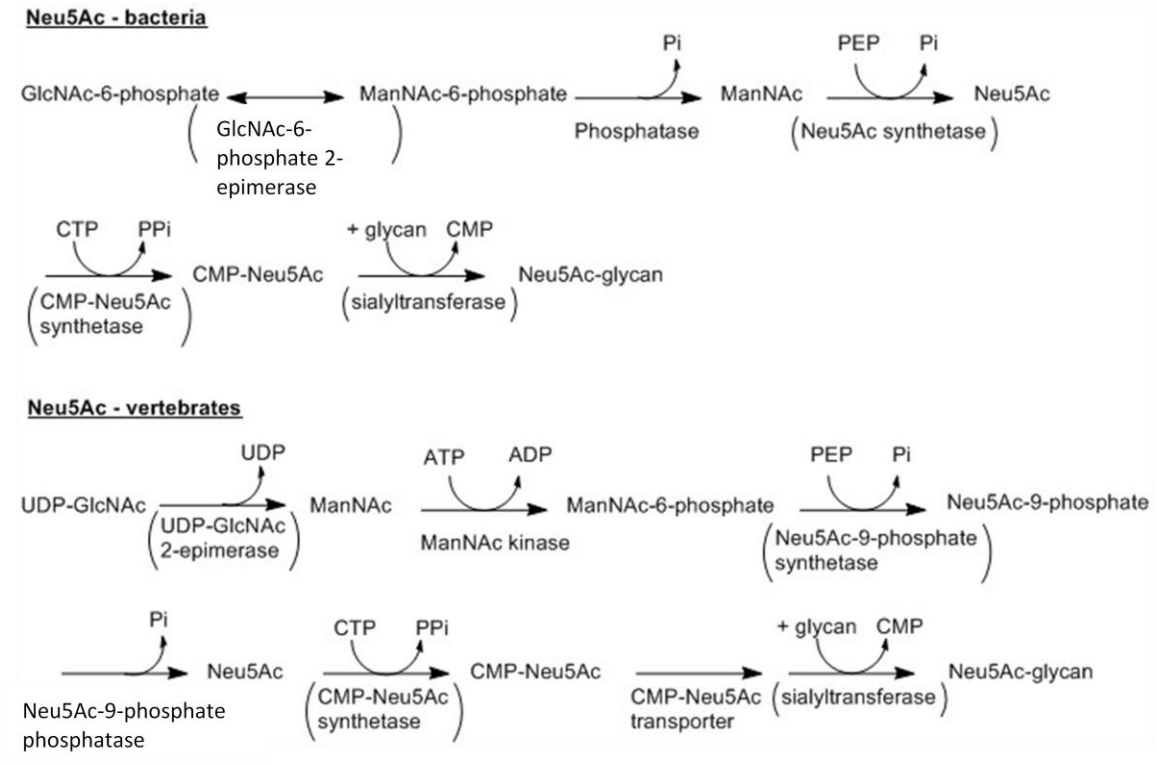


Figure 1.12 *Biosynthesis of Neu5Ac in bacteria and vertebrates. The biosynthesis of Neu5Ac in bacteria and vertebrates are compared. The steps in biosynthesizing are mostly conserved with minor difference in both systems. The similar steps in both systems are noted in brackets.*

The epimerase catalyzed conversion of *N*-acetylglucosamine (GlcNAc) to *N*-acetylmannosamine (ManNAc), which is unique to this pathway; therefore, biosynthesis of Neu5Ac is drawn generally starting from this step (5). In vertebrates, UDP-GlcNAc 2-epimerase converts UDP-GlcNAc to ManNAc and UDP. Subsequently, ManNAc kinase uses ATP to phosphorylate ManNAc to give ManNAc-6-phosphate (43). In bacteria,

GlcNAc-6-phosphate is epimerized by GlcNAc-6-phosphate 2-epimerase to give ManNAc-6-phosphate which is dephosphorylated by a phosphatase (26). In vertebrates, ManNAc-6-phosphate and phosphoenolpyruvate (PEP) are condensed to give Neu5Ac-9-phosphate by the enzyme Neu5Ac-9-phosphate synthetase, following which the phosphate is removed by a phosphatase to yield Neu5Ac (73). In bacteria, the condensation of ManNAc with PEP to form Neu5Ac is catalyzed by Neu5Ac synthetase (108). In both vertebrates and bacteria, Neu5Ac is activated to give CMP-Neu5Ac with the CMP group being provided by cytidine 5'-triphosphate (CTP), a reaction that is catalyzed by CMP-Neu5Ac synthetase (124)(33). In vertebrates CMP-Neu5Ac is synthesized in the nucleus and then transported to the Golgi body by a CMP-sialic acid transporter (32). The CMP-Neu5Ac donor is used to add Neu5Ac onto acceptors by the sialyltransferases in both bacteria and vertebrates. The linkage and acceptor substrate depend of the specificity of the various sialyltransferases (104).

1.3.3. Biological Interaction Proteins and Biological Functions

Proteins that recognize and interact with carbohydrates are named lectins. Lectins can be categorized into three groups, vertebrate pathogen lectins, vertebrate endogenous lectins, and other lectins. The first two groups will be discussed in more detail below. A subset of lectins that bind sialic acids are called Siglecs, sialic acid-binding immunoglobulin-type lectins. Indeed, the first lectins that were shown to bind to sialic acid were discovered in viruses where agglutination of red blood cells was detected upon addition of live virus at low temperatures (53). Studies showed the negatively charged carboxylate group at C1 of sialic acid is critical for recognition. Furthermore, the specific linkage of sialic acid to the glycoconjugate and the context of the underlying sugar also plays a role in recognition.

1.3.3.1. Vertebrate Pathogen Lectins

Pathogens like viruses, bacteria, and protozoa express lectins on their cell or particle surface that interact with sialic acid and these interaction give rise to various physiological effects. The receptors found on the surface of viruses that bind to sialic acid on red blood cells are called hemagglutinins. The most well-known viral lectins being influenza A hemagglutinins that binds to host cell surface sialic acids to initiate the infection process. It is known that a reservoir of influenza viruses is found in species of

water fowl; however, influenza viruses have adapted to infect humans and this causes the seasonal flu epidemics (122). Also, some bacteria interact with host cell sialic acids and the bacterial lectins can be expressed on the bacterial surface or secreted as a soluble lectin. Bacterial surface lectins are called adhesins and the soluble lectins are usually toxins. A typical soluble sialic acid binding lectin is cholera toxin that is produced by *Vibrio cholera* (35). This lectin recognizes the sialylated glycolipid, ganglioside GM1 and this binding event leads to overactivation of the signalling pathway that produces cyclic AMP in the gastrointestinal epithelial cells, which ultimately causing severe diarrhea and electrolyte imbalance. An often studied protozoa is *Plasmodium falciparum,* the malaria parasite, that contains EBA-175 antigen and BAEBL proteins in order to infect erythrocytes, a process that involves sialic acid binding lectins (77).

1.3.3.2. Vertebrate Endogenous Lectins

Vertebrates also express their own or intrinsic (endogenous) lectins that recognize sialic acid. Many of the endogenous lectins are found as part of the innate immune system. Common intrinsic lectins include Siglecs, factor H, selectins, neural adhesion molecule L1, calreticulin and laminin. Siglecs contain the largest number of family members compared to other sialic acid binding lectins (113). Examples of Siglecs that are found in vertebrates are sialoadhensin and Siglec-1, which are expressed in macrophages, CD22 and Siglec-2 in B-cells, and Siglecs-1, 3, 5, 7, and 10, which have shown to be involved in natural killer cells and granulocytes (25). Factor H is an inhibitory regulatory molecule of the alternative complement pathway of the innate immune system. The alternative pathway is one of the three complement pathways that enhance phagocytosis by making an antigen for an immune response in order to kill pathogens. Factor H binds to sialic acid presented on the cell surface glycoconjugates (glycosaminoglycans; GAGs) that are usually present on host cells but not on the pathogen surface; ensuring that the alternate complement pathway recognizes and attacks pathogens but not host cells (64). Selectins are a group of cell adhesion molecules (CAMs) that are transmembrane glycoproteins and possess similar structure motifs and calcium requirements for carbohydrate binding. There are three types of selectins, E-selectin found in the endothelial cells, L-selectins that are expressed on leukocytes, and P-selectin that are found on the surface of activated endothelial cells to activate platelets. All three subsets of selectins share high sequence homology and are

responsible for leucocyte trafficking to the plasma membrane (E-selectin), the tips of microfolds on leukocytes (L-selectin), or the secretory granules (P-selectin) (62).

1.3.3.3. Sialic Acid as Biological Masks

The terminal location of sialic acids residues contributes to its function of masking the penultimate oligosaccharide residues. The classic example involves the mammalian regulation of glycoproteins where terminal sialic acids prevent recognition of galactose residues by the Ashwell receptors in the liver, an interaction that mediates glycoprotein clearance (87). Therefore, sialic acids play a vital role in maintaining the serum half-lives of glycoproteins. Another instance includes the attachment of sialic acid to protect glycoconjugate sugars from carbohydrate binding receptors like galectins that are associated with inflammation, apoptosis, and immune cell responses (25).

1.4. Sialidases

Sialic acid hydrolases, which are also called neuraminidases or sialidases (EC 3.2.1.18), cleave glycosidic linkages between sialic acid and another sugar residue. Sialidases are a large family of enzymes, found in a range of organisms that include, viruses, microorganisms, and vertebrates, but not plants (99). Also, both viral and bacterial sialidase mechanisms of catalysis have been studied extensively and are covered in detail in the following subtopics. Also, the biochemical properties and functions of sialidases will be discussed.

1.4.1. Sialidase Sequence Families and Subtypes

The major categories of sialidase are viral, bacterial, and mammalian. For influenza type A virus, there are 9 different neuraminidase core sequences, assigned as N1 to N9. Each neuraminidase can be associated with a different hemagglutinin and these combinations give the various influenza A viral subtypes (e.g., H1N1, H5N1). According to hydrophobic cluster analysis, the viral sialidases belong to glycosyl hydrolase family 34 (GH34) (19). For bacterial sialidases, reviewed in 2011 there are 20 types of reported sialidases and at least 8 of these have had their crystal structures solved (51). The bacterial sialidases are grouped into GH 33 and 83 families. For

mammalian sialidases (GH33), the *NEUs* genes encode for sialidases that can be classified into four types: Neu 1 is found in lysosomes, Neu 2 a cytosolic enzyme, Neu 3 is a membrane localized enzyme, and Neu 4 is another lysosomal protein.

1.4.2. Enzyme Structure

In 1983, Varghese, Laver and Colman published the first sialidase x-ray crystal structure, which was of an influenza type A sialidase, and this is shown in Figure 1.13 (109). The tetrameric enzyme has each monomer made from six topological identical beta- sheets arranged in a propeller formation. The deep pocket observed from the top view of each monomer was identified as the catalytic site based on studies where sialidase crystals were soaked with sialic acid (109).

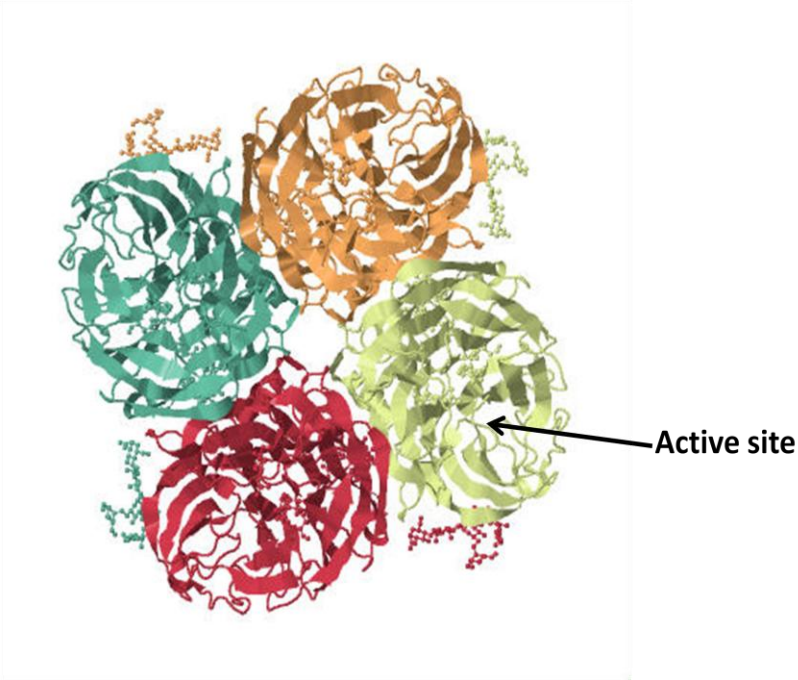
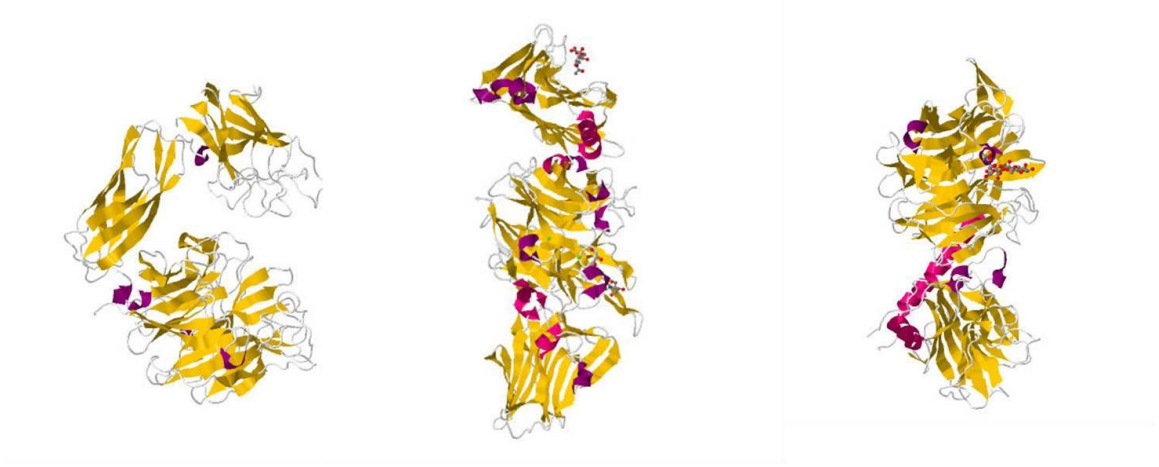


Figure 1.13 *Crystal structure of influenza virus A PDB code 2BAT. The influenza sialidases are usually found as homo-tetramers (each monomer is shown in different color above). The N-terminal stalk region of each monomer is used to join the subunits and anchor the protein to the virus. The active site for each of the monomer is located in the center region of β-propeller fold (pointed with an arrow).*

The bacterial sialidases have less than 30% sequence homology with the influenza viral enzymes but the topology and several motifs and residues are conserved (51). Bacterial sialidases consist of a catalytic domain with some enzymes having additional domains so that molecular weights of these enzyme range from 40 to 150 kDa (51). Non-pathogenic bacterial sialidases commonly have four to five copies of an aspartate box (Ser/Thr-X-Asp-X-Gly-X-Thr-Trp/Phe), but the function of this consensus sequence is unknown (12). Furthermore, the RIP motif (Arg-Ile/Leu-Pro) was also identified and this is located upstream of the aspartate box. The function of this Arg residue within the active site is to stabilize the negative charge of the carboxylate group on the sialioside (75). Selective bacterial sialidase structures are shown in Figure 1.14.



a) *Micromonospora viridifaciens* sialidase (1EUT)

b) *Vibrio cholerae* sialidase with α-2, 6 sialyl-lactose (PDB: 1E0P)

c) *Trypanosoma cruzi* trans-sialidase in complex with α-2, 6 sialyl-lactose (PDB: 1S0I)

Figure 1.14 *Selective bacterial sialidases with color highlights the different secondary structures.*
The β-propeller folds are colored in yellow and α-helices are in purple.

1.4.3. Active Site Residues and Their Catalytic Roles

The active site residues of viral and bacterial sialidases responsible for catalysis are conserved. Based on sequence analyses and x-ray crystallography data, seven conserved active site residues are identified (Figure 1.15). These key residues include

the three arginines which assist in stabilizing the negative charge of the carboxylate group of the substrate (115). The catalytic residues proposed to be involved in the catalytic mechanism are an aspartic acid, a glutamic acid and a tyrosine (61). The seventh conserved residue is another glutamate which forms a salt bridge with one of the conserved arginines in order to stabilize its position in the arginines triad to form hydrogen bonding with carboxyl group of sialosides (115).

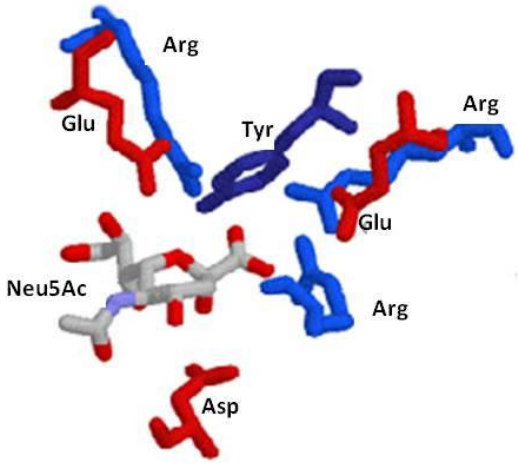


Figure 1.15 *Conserved active site residues of sialidases.*
Seven conserved sialidase active site residues were shown with the product Neu5Ac bound in the active site. The structure used is the active site of the influenza A sialidase (PDB code 2BAT).

All sialidases characterized to date are retaining glycosidases that give α-sialic acid as the first-formed product. Kinetic and structural studies have aided in determination of the catalytic roles for these conserved residues. The active site tyrosine residue acts as a nucleophile and the active site aspartic acid and glutamic acid function as Brönsted acids/ bases to assist in nucleophilic attack and nucleofuge departure. Studies on the *Trypanosoma cruzi trans*-sialidase demonstrated that Try342 is a nucleophile as the trapped 3-fluorosialoyl-enzyme intermediate has been characterized using X-ray crystallography (27). This structure also suggests that Asp59 is positioned within hydrogen bonding distance so that it acts as a general acid catalyst for the formation of the glycosyl-enzyme intermediate and as a general base catalyst to assist the transfer reaction of the acceptor (27). Our group has shown that mutation of

the conserved Tyr370 nucleophile in *M. viridifaciens* sialidase to an alanine, an aspartate or a glycine residue changes the mechanism from retaining to inverting (120).

1.4.4. Substrate Specificity

Unsurprisingly, sialidases from different species exhibit different linkage preferences and glycoconjugate types. The common sialoside linkages, α(2-3) and α(2-6) to galactose residues were assayed as substrates towards six sialidases, both bacterial and viral (21), which were from *Vibrio cholerae, Arthrobacter ureafaciens, Clostridium perfringens*, Newcastle disease virus, fowl plaque virus and influenza A N2. Most of these sialidases hydrolyze α(2–3) glycosidic linkages faster than α(2–6) linkages, except *A. ureafaciens* sialidase, which prefers α(2-6) linkages (21). Other enzymes that show a preference for α(2-6)-linked sialic acid are *M. viridifaciens* sialidases (120) while *Corynebacterium diphtheria* sialidase shows a preference for α(2,-3)-linked sialic acid (52).

1.4.5. Biological Properties and Functions

Influenza viral sialidases are found on the surface of the viral particles and they allow newly formed virus particles to be released from the host cell (Figure 1.16). The surface of the influenza virus also contains hemagglutinin proteins that recognize and bind to sialic acid residues on the host cell surface, an interaction that allows the virus to enter the host cell. The infected cell replicates and produces virion particles that bud from the host cell membrane. The sialidase then cleaves sialic acid residues from the virion's surface and the glycan structures of the infected cell and this allows the infection to spread.

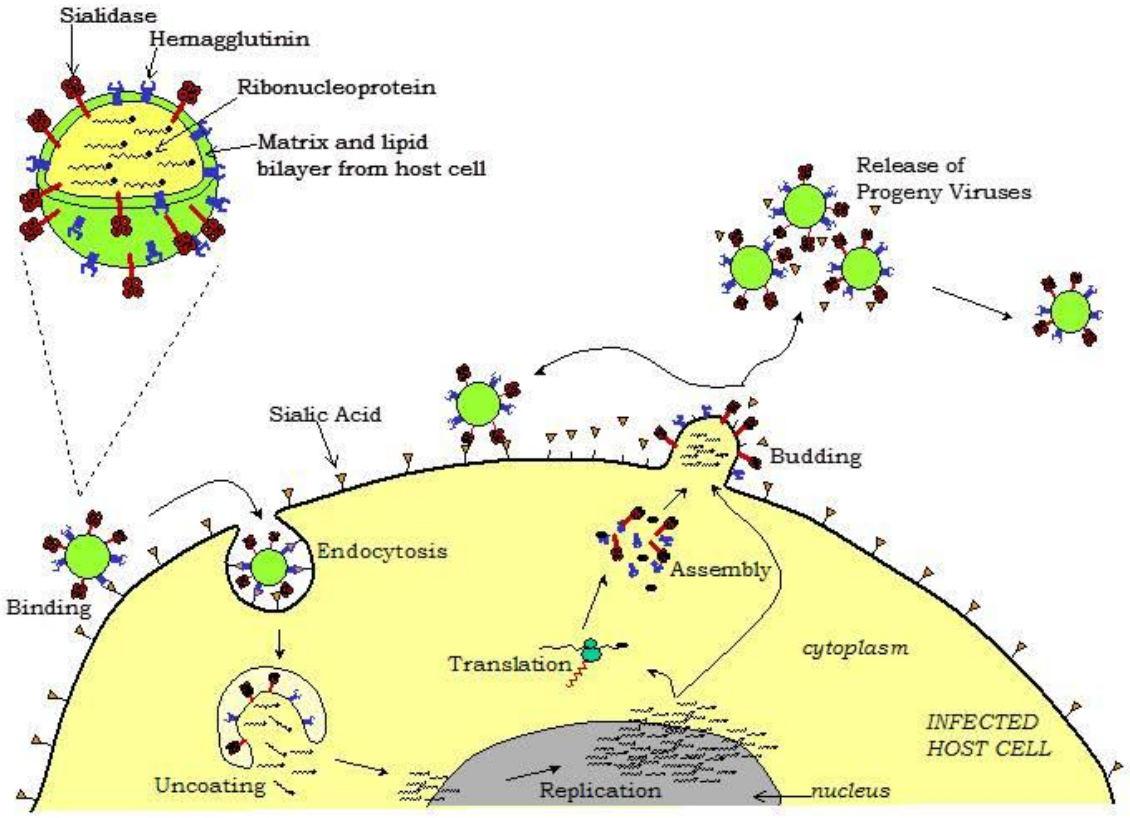


Figure 1.16 *Influenza Virus Infection Cycles.*
The virus enter host cell through hemagglutinin binding to cell surface sialic acid. The viral RNA uses host machinery system to replicate and translate the viral genetic materials. The assembled viral particles exit the cell by budding. The virions use its surface sialidase to hydrolyze sialic acid from binding to hemagglutinin and therefore, released newly made virus and ready to infect another host cell.

Bacterial sialidases can be secreted or cell bound. The cell bound proteins can play roles in exterior interactions that assist in entry into and infection of host cells. Also, the bacterial sialidase can be pre-stored inside the cell before being released for infection and/or food scavenging purposes (23). The secreted forms of sialidase can vary in size depending on the type of substrate available and whether the hydrolyzed sialic acid is used as a carbon source for bacteria growth (83). Examples of pathogenic bacteria that express a sialidase include *Bacteroides fragilis* and *Pseudomonas aeruginosa* with the function of their enzymes being to cleave the sialic acid residue from ganglioside-GM1 to make asialo-GM1 to which type 4 pilli bind preferentially (20).

1.4.6. Trans-sialylation

The transfer of sialic acid onto glycoconjugates can be achieved using either a chemical and or an enzymatic approach. In general, chemical syntheses involve numerous steps, many of which are protection or deprotection reactions, as well as purification of the products so that only a single regio- and stereo isomer is produced (19). The enzymatic approach is more facile in producing sialosides with a specific stereochemistry (19).

Bacterial sialidases can catalyze the hydrolysis of the terminal sialic acid that is coupled with a α(2, 3), α(2, 6), or α(2, 8) linkage to another sugar (51). Since all enzymes catalyze both the forward and the reverse reactions, using hydrolase enzymes to produce transferase products is also possible. However, due to the high concentration of water under physiological conditions the forward hydrolysis rate general dominates and the transglycosylation products made by the exo-α-sialidases are substrates for hydrolysis (19).

In spite of product hydrolysis many advantages are apparent for the use of sialidases to form transglycosylation products rather than using sialyltransferases. Most sialidases are soluble and secreted into the culture media, therefore, purification is less problematic compared to membrane bound sialyltransferases. In general, sialidases possess broader substrate specificities then do sialyltransferases. Sialidases are stable to the addition of various organic co-solvents (2)' (89). Furthermore, the sialyl-donors used by sialidases are often more stable and cost effective compared to CMP-Neu5Ac, the obligatory donor for sialyltransferases (2)' (114).

The first report for the use of a sialidase for the synthesis of a sialyloligosaccharides was performed with an immobilized *V. cholerae* enzyme (103), with *p*NP-α-Neu5Ac being the sialic acid donor to give both α(2,3) and α(2,6) sialosides with a combined yield of less than 15% after more than 45 h of reaction (103). Other sialidases from *C. perfringens*, *A. ureafaciens* and Newcastle disease virus have been used to perform transglycosylation reaction. Overall yields of transglycosylation were less than 10% and bacterial sialidases mainly produced the α(2,6) linkage product while viral sialidases gave α(2,3) regioisomers (51).

1.5. Research Objectives

The focus of work described in the present thesis was to optimize transglycosylation using the sialidase from *M. viridifaciens*. The various advantages of using sialidases to chemoenzymatic synthesis of sialylated molecules were discussed above. By comparison product hydrolysis using wild type sialidase for synthesis gives low transglycosylation yields. This thesis focused on the use of mutant sialidases to try and achieve regiospecific coupling with higher transglycosylation yields.

The first project demonstrated how substitution of the tyrosine nucleophile to a glycine in the *M. viridifaciens* sialidase (Y370G) solved the issue of transglycosylation product degradation, while the enzyme displayed a low specificity for the acceptor sugar. Subsequent projects focused on engineering two hydrophobic loops into the Y370G mutant sialidase to attempt to incorporate a sugar binding site as a means to improve transglycosylation. The third project focused on screening a random mutagenetic library of Y370M and Y370H mutant sialidase in order to discover new sialidase mechanisms.

1.6. Contributions of the author to the work presented in this thesis

Description of candidate's contribution to the manuscript (Chapter 2)

CHAPTER 2. Neuraminidase Substrate Promiscuity Permits a Mutant *Micromonospora viridifaciens* Enzyme to Synthesize Artificial Carbohydrates.

This chapter is an exact copy of a paper submitted to Biochemistry.

For this contribution, I performed all the Y370G mutant sialidase coupling reactions with various acceptor sugars and monitored the reactions on HPLC-PAD. I purified all of the transglycosylation products that were made by using Y370G and 2,6-sialyltransferase. I prepared all samples for NMR spectral characterization. Dr. Fahimeh Shidmoossavee synthesized 2FPh-β-NeuAc and analyzed both the ^{1}H- and ^{13}C- NMR spectra of the purified coupled sialosides. Dr. Andrew Bennet supervised the project and was mostly involved in writing and editing the paper.

Description of candidate's contribution to the thesis (except Chapter 2)

In the rest of the thesis, all experimental work was performed by the candidate except:

- The genetic insertion of the hydrophobic loops into *M. viridifaciens* Y370G mutant sialidase gene was performed by Dr. Jacqueline Watson.

- Prepared of Y370M and Y370H plasmids libraries was done by Dr. Jacqueline Watson.

- The following substrates, FMU-man, Neu5Ac α-2,6 Gal β-FMU, Gal β-FMU and Gal-ELF, were synthesized by Dr. Jefferson Chan.

- Dr. Fahimeh Shidmoossavee made 2FPh-Neu5Ac and Sia-gal-X.

All experiments described in this thesis were part of the thesis work.

2. Neuraminidase Substrate Promiscuity Permits a Mutant *Micromonospora viridifaciens* Enzyme to Synthesize Artificial Carbohydrates

† Funded by the Natural Sciences and Engineering Research Council of Canada.

*Lydia L. Cheng, Fahimeh S. Shidmoossavee and Andrew J. Bennet**

Departments of Chemistry and Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, V5A 1S6, Canada.

e-mail bennet@sfu.ca

2.1. Abstract

Mutation of the nucleophilic amino acid residue tyrosine to the small nonpolar residue glycine (Y370g) in the active site of *Micromonospora viridifaciens* neuraminidase (*Mv*NA) produces an efficient catalyst for the transfer of *N*-acetylneuraminic acid from an artificial substrate (i.e., phenyl *N*-acetyl-β-D-neuraminide) to a sugar acceptor (e.g., D-lactose, D-glucose, D-mannose, D-raffinose, D-allose or D-fructose) to give *N*-acetyl-α-neuraminide coupled carbohydrate products. In addition, this mutant enzyme (*Mv*NA Y370g) catalyzes the transfer of a sugar residue from the artificial substrate 2-fluorophenyl *N*-acetyl-β-D-neuraminide to methyl glycopyranosides acceptors. Interestingly, when trans-glycosylation reactions are conducted in aqueous solutions containing 30% (v/v) acetonitrile, the α-anomeric acceptors of methyl glucopyranoside and galactopyranoside generate higher product yields than do their corresponding β-anomers. Specifically, a 64 hr reaction with 2-fluorophenyl *N*-acetyl-β-D-neuraminide as the limiting reagent and the acceptors methyl α-D-galactopyranoside, methyl α-D-glucopyranoside or methyl α-D-mannopyranoside gives trans-glycosylation product yields of 36.4%, 30.7% or 35.7%, respectively. Using methyl α-D-galactopyranoside as the acceptor, trans-glycosylation catalyzed by both *Mv*NA Y370g and a naturally occurring 2,6-sialyltransferase yields an identical product, which we identified as methyl *N*-acetyl-α-D-neuraminyl-(2→6)-α-D-galactopyranoside. The *Mv*NA Y370g-catalyzed coupling of *N*-acetylneuraminic acid to these three methyl α-D-glycopyranoside acceptors is favored by factors of 24–30 fold over the competing hydrolysis reaction. These coupling efficiencies likely arise from non-selective interactions between the acceptor glycopyranoside and *Mv*NA Y370g, which preferentially places a carbohydrate hydroxyl group rather than water in close proximity to the active site where this functionality intercepts the nascent neuraminyl oxacarbenium ion that is formed during cleavage of the glycosidic bond in the aryl *N*-acetyl-β-D-neuraminide donor. The ability to transfer *N*-acetylneuraminic acid from a stable and readily accessible donor to acceptor carbohydrates that are not substrates for sialyltransferases is one step on the path for pseudo-human glycoprotein production in non-mammalian cell lines.

2.2. Introductions

Sialic Sialic acids are a family of structurally diverse nine-carbon keto-sugars typically found as terminal carbohydrate residues of glycoconjugates, a position that impacts key biological recognition events (6)(69)(86)(100). In mammalian systems, sialic acid—also known as *N*-acetylneuraminic acid—is often an important carbohydrate end-cap for modulators of cellular responses in a range of physiological processes, including differentiation, proliferation and apoptosis (85)(71)(112). Many of these processes are mediated by asialoglycoprotein receptors (ASGPRs) that bind glycoconjugates with carbohydrate structures that terminate in galactose or *N*-acetylgalactosamine residues; if, however, these carbohydrate moieties are capped with sialic acid, then the underlying galactose or *N*-acetylgalactosamine-containing carbohydrate epitopes are bound less tightly to the ASGPRs (8)(31). Interestingly, the rate of glycoprotein clearance from the circulation—a process that is mediated by the ASGPRs of liver cells—depends on the number of appended sialic acid residues and the number of exposed galactose and/or *N*-acetylgalactosamine moieties (72)(9)(56).

Neuraminidases (sialidases; E.C. 3.2.1.18) are a specific type of glycosyl hydrolase (GH) that acts on sialic acid, removing it from the end-cap position of a glycoconjugate. Given sialic acid's important role in mammalian biological recognition events, it is not surprising that, in addition to neuraminidases from the host cell, bacterial and viral neuraminidases can affect molecular recognition events that impact the organism. During *Streptococcus pneumoniae* infection, for example, a bacterial neuraminidase catalyzes the removal of sialic acid residues from the host organism's blood platelets. This causes the previously protected sugar structures on the surface of platelets to be recognized by Ashwell receptors (a type of ASGRP) in the liver (8)(31) and thereby initiates clearance of these platelets from the circulation, an outcome which lessens coagulation during sepsis (40). This natural system of molecular recognition has implications for the design of pharmaceuticals. In particular, if a recombinant glycoprotein such as a glycoprotein hormone or therapeutic agent, lacks terminal sialic acid residues on the carbohydrate chains, then that recombinant molecule will be cleared more rapidly from the circulation (31).

The synthesis of glycoconjugate-containing therapeutic enzymes and protein hormones with long serum half-lives remains challenging despite recent advances in recombinant DNA technologies. One approach to lengthening the circulatory half-live of recombinant glycoproteins is to introduce additional asparagine *N*-glycosylation sites within the protein, effectively increasing the sialic acid content of the therapeutic and thus lengthening its circulatory half-life (34). Alternatively, glycoproteins expressed in systems that synthesize glycosylation structures that are different than those of humans (e.g., yeast and insect cell lines) (41) can be more closely matched to human type *N*-glycosylation patterns by subsequently attaching *N*-acetylneuraminic acid to the termini of the oligosaccharide chains; ideally, this is done using enzymatic methods (44). This approach has been successful in cases where the acceptor glycoconjugate contains a terminal sugar residue that is recognized by a sialyltransferase (84) or a *trans*-sialidase(66).

In another approach, neuraminidases, which normally remove terminal *N*-acetylneuraminic acid residues from glycoproteins, can be used to generate sialoglycoproteins by catalyzing the reaction in reverse. For example, wild-type neuraminidases are capable of transferring a *N*-acetylneuraminic acid residue onto an acceptor sugar (88). The usefulness of this approach is limited, however, because the trans-glycosylation product is also a substrate for the normal hydrolytic reaction and thus the enzyme degrades the sialylated products as they are produced (3).

Using site-directed mutagenesis of the *M. viridifaciens* neuraminidase (*Mv*NA) gene, we produced a mutant neuraminidase that works in the synthetic direction to give modest yields of sialic acid linked to an acceptor molecule. Specifically, we changed the active site nucleophilic tyrosine residue of *Mv*NA to a glycine residue (Y370g) to give a mutant enzyme (*Mv*NA Y370g) that transfers sialic acid from the non-natural substrate phenyl *N*-acetyl-β-D-neuraminides (Ph-βNeu5Ac) to an acceptor to give α-2,6-sialyl-lactose in yields of up to 13% (118). In contrast to wild-type neuraminidases, this mutant enzyme is an extremely poor catalyst for the hydrolysis of natural *N*-acetyl-α-D-neuraminides (117), which contain unactivated carbohydrate aglycones, and thus the concentration of the coupled product remains constant (118).

As a first step toward increasing the yield of coupled products using mutant neuraminidases, we aimed to understand the mechanism of Y370g mutant-catalyzed reactions and then apply that knowledge toward the selection of optimal conditions for trans-glycosylation (91). Since both academic and industry researchers express proteins using non-mammalian cell systems and recombinant DNA techniques (24)—an approach that produces non-sialylated glycoproteins—we decided to test the ability of our Y370g mutant sialidase to transfer *N*-acetylneuraminyl units onto a variety of carbohydrates. This paper presents a study of the range of acceptor sugars that undergo *Mv*NA Y370g mutant-catalyzed coupling with *N*-acetylneuraminic acid and the identity of the reaction products.

2.3. Materials and Methods

2.3.1. Chemicals and reagents

Raffinose, ribose and *N*-acetylneuraminic acid were purchased from British Drug Houses Ltd, Eastman Kodak Co and Rose Scientific, respectively. Dithiothreitol was purchased from Bioshop and cytidine 5′-triphosphate disodium salt (CTP) was purchased from 3B Scientific Corporation. Pyridine and acetic acid were purchased from Anachemia, and sodium acetate and magnesium chloride were purchased from Caledon Laboratories Ltd. All other chemicals, reagents and enzymes were purchased from Sigma-Aldrich. Ph-β-Neu5Ac and 2FPh-βNeu5Ac were synthesized according to literature procedures (91)(30).

2.3.2. Enzymes

*Mv*NA Y370g mutant was expressed and purified as previously reported (117). *Escherichia coli N*-acetylneuraminic acid (Neu5Ac) aldolase was purchased from Codexis. *Neisseria meningitides* CMP-Neu5Ac synthase (50)(30) and *photobacterium sp*. JT-ISH-224 α-2,6-sialyltransferase (49) were expressed as reported.

2.3.3. Trans-glycosylation reactions

We added *N*-acetylneuraminic acid donor Ph-βNeu5Ac (40 mM), various acceptor sugars (200 mM) and *Mv*NA Y370g mutant (5×10^{-6} M) to a buffered solution (pH 5.25, 100 mM acetate buffer) containing acetonitrile (30% v/v) and incubated at 37 °C overnight; negative control reactions (the reaction mix without enzyme) were performed in parallel. We examined whether any coupled product was formed by taking an aliquot (5 µL) from the coupling reaction mixture, heating it at 85 °C for 15 min to inactivate the enzyme, diluting the sample 1/60 in H_2O (v/v) and analyzing the resultant solution by HPLC-PAD (Dionex ICS-3000 HPLC using amperometric detection). Specifically, we injected the sample (10 µL) at a flow rate of 0.5 mL/min with isocratic elution (0.15 M NaOH and 0.015 M NaOAc) onto a Carbopac PA20 column (3×150 mm). We identified the trans-glycosylation product by comparing the chromatogram with that of the negative control. After new peaks were identified, we treated an aliquot from the coupling reaction with wild type *M. viridifaciens* neuraminidase and reanalyzed to confirm that the new peak disappeared upon neuraminidase treatment, and thus we concluded that it was *N*-acetylneuraminide.

In a second set of trans-glycosylation reactions we used Y370g mutant neuraminidase to transfer Neu5Ac from 2Fph-βNeu5Ac (sialyl donor) to four methyl glycoside acceptor sugars: α- and β-D-galactopyranoside, and α- and β-D-glucopyranoside. In a typical experiment, we added neuraminyl donor (5.0 mg), acceptor sugar (25.0 mg) and *Mv*NA Y370g mutant (0.37 mg) to acetate buffer (pH 5.25, 100 mM) containing acetonitrile (30% v/v) in a total volume of 250 µL and incubated at 37 °C for 16 h. We terminated the reactions by adding ethanol to give a final concentration of 70% (v/v) and then precipitated the enzyme by storing the samples at –20 °C for 20 min. The supernatant was separated from the precipitate by centrifugation (13000 rpm for 10 min at 4 °C), evaporated to dryness and re-suspended in D_2O for ^1H NMR spectral analysis.

To identify the regiochemistry of coupling, we increased the concentrations of both acceptor and donor to achieve Y370g mutant enzyme (0.37 mg), 2FPh-NeuAc (25.0 mg) and acceptor (55.0 mg) in pH 5.25 acetate buffer (100 mM, containing acetonitrile 30% v/v) in a total volume of 250 µL and then incubated the sample at 37 °C

for 64 h. After precipitation of the *Mv*NA Y370g mutant (as described above), we analyzed the sample by NMR spectroscopy to determine the coupling to hydrolysis product ratio. Subsequently, we degraded the hydrolysis product (*N*-acetylneuraminic acid) by adding *N*-acetylneuraminic acid aldolase and pyruvate decarboxylase at pH 7.03 (100 mM phosphate buffer) to the sample and incubating at room temperature overnight (16 hours). At this time, ^{1}H NMR spectral analysis showed the absence of *N*-acetylneuraminic acid. We then removed the protein by ethanol precipitation and centrifugation (as detailed above). Following evaporation of the supernatant, we re-suspended the sample in milliQ water and loaded it onto an anion exchange column (DOWEX 1×2–200 anion exchange resin–acetate form). The column was washed thoroughly with water and subsequently *N*-acetylneuraminide products were eluted using a pyridinium acetate buffer (20 mM, pH 5.01). Fractions were collected in 1.5 ml Eppendorf tubes and evaporated to dryness. Fractions containing pure *N*-acetylneuraminide product were combined and the structure and chemical formula of each product were confirmed by NMR spectroscopy and high resolution mass spectrometry. All NMR peak assignments are based on ^{1}H−^{1}H COSY and ^{1}H−^{13}C HMQC experiments; coupling constants are reported in hertz (Hz). The ESI-MS analyses were performed by the Mass Spectrometry and Proteomics Facility at the University of Notre Dame.

2.3.4. 2,6-Sialyltransferase-catalyzed synthesis

We used a published protocol for the enzyme-catalyzed synthesis of 2,6-*N*-acetylneuraminides with minor modifications (16). Specifically, we added fresh supernatant (200 µL) from an expression of CMP-Neu5Ac synthase (50) to a reaction medium containing *N*-acetylneuraminic acid (20 mg), CTP (100 mg), $MgCl_2$ (200 mM) and DTT (4 mM) in Tris buffer (0.8 mL; 500 mM; pH 8.01), incubated for 1.5 hr at 37 °C and centrifuged the sample (5000 rpm) for 5 min at rt. We transferred the supernatant to a clean Eppendorf tube, added methyl α-D-galactopyranoside (150 mg) and supernatant (200 µL) from an expression of 2,6-sialyltransferase (16)(49) and then incubated the sample overnight at 37 °C. The progress of the enzymatic reaction was monitored by ^{1}H NMR spectroscopy. Purification and analysis of the trans-glycosylation product was performed as described above.

2.4. Results

2.4.1. Trans-glycosylation reactions

We investigated the ability of the *Mv*NA Y370g mutant enzyme to use different sugars as acceptors of the *N*-acetylneuraminyl group from the donor Ph-βNeuAc; the HPLC-PAD chromatograms from coupling reactions containing D-lactose, D-glucose, D-mannose, D-raffinose, D-allose and D-fructose as acceptors are shown in Figure S 2.1 (Supporting Information). Several acceptor sugars gave rise to a newly observed peak in the HPLC-PAD chromatogram that disappeared if the reaction mixture was subsequently treated with wild-type *Mv*NA. In the presence of Ph-βNeuAc, the mutant enzyme catalyzed trans-glycosylation with acceptors D-lactose, D-glucose, D-mannose, D-allose, and D-fructose to form single dominant *N*-acetylneuraminyl product peaks (Figure S2.1), whereas, the acceptor D-raffinose gave two peaks that are consistent with two newly formed *N*-acetylneuraminide-containing products. In contrast, the carbohydrate acceptors L-xylose, D-ribose, D-lyxose, and D-isoascorbic acid did not produce any observable product peak in the HPLC-PAD chromatogram.

All of the productive trans-glycosylation acceptor carbohydrates we tested are present in solution as a mixture of pyranose and furanose anomers and as such would require the synthesis of an impracticable number of standard compounds for product identification to proceed via HPLC-PAD chromatography. Instead, we chose to identify the products by NMR spectroscopy, a method with which we could characterize the type of glycosidic linkage formed between sialic acid and commercially available methyl glycopyranosides. Figure 2.1a and 2.1b show the ^{1}H NMR spectra we acquired following an overnight incubation of *Mv*NA Y370g mutant (0.51 mg) in a coupling reaction mixture of 2FPh-βNeu5Ac (5.0 mg) and methyl α-D-galactopyranoside (~15 mg) or methyl β-D-galactopyranoside (~15 mg), respectively. The corresponding NMR spectra for the coupling reactions to methyl α- and β-D-glucopyranosides (76) are shown in Figure S2.2 (supporting information).

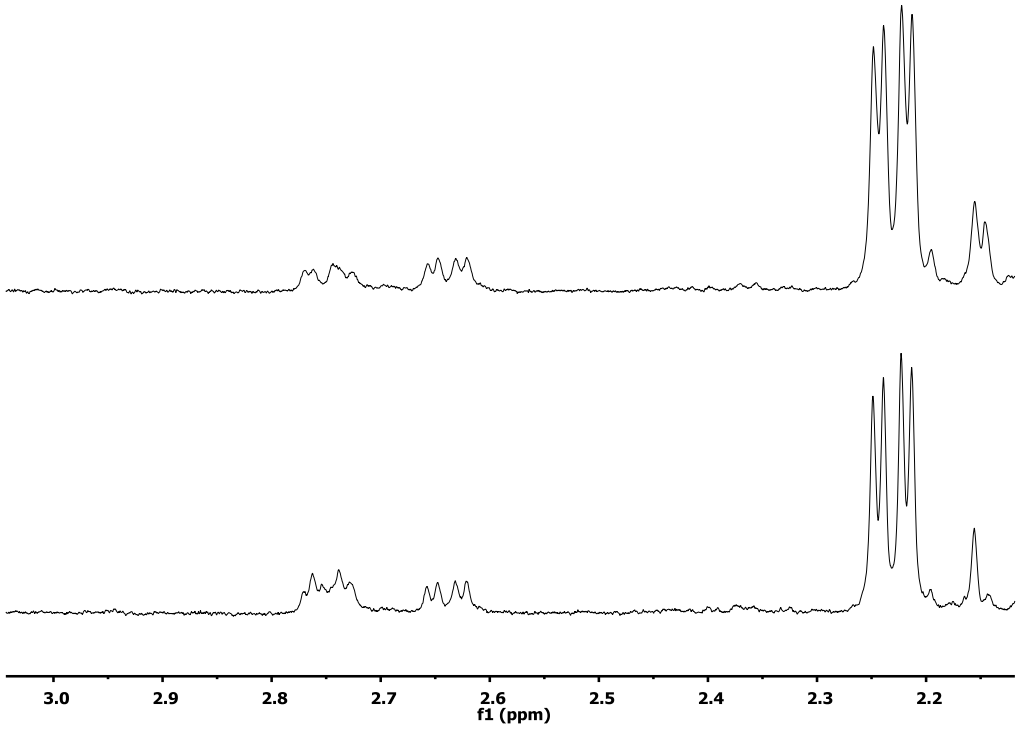


Figure 2.1 *The part of the ^{1}H NMR spectra showing the resonances for the various H-3_{eq} protons of the coupling reaction mixture for the* Mv*NA Y370g mutant-catalyzed coupling of 2FPh-βNeu5Ac to an acceptor sugar.*
(a) methyl β-D-galactopyranoside acceptor, and (b) methyl α-D-galactopyranoside acceptor.

Both the galactoside and the glucoside α-anomers showed a greater propensity for coupling than did their corresponding β-anomers. We calculated the yield of coupled and hydrolysis products by integrating the respective peaks in the ^{1}H NMR spectra (Table 2.1).

Table 2.1 *Comparison of trans-Glycosylation and Hydrolytic Activities of the Micromonospora viridifaciens Y370g Mutant Sialidase with Various Carbohydrate Acceptors.*

	Coupling Experiment 1[a]			Coupling Experiment 2[b]		
Acceptor	**Coupling**	**Hydrolysis**	**Starting Material**	**Coupling**	**Hydrolysis**	**Starting Material**
Me α-galactoside	13.4%	78.7%	8.9%	36.4%	42.4%	21.2%
Me β-galactoside	6.0%	86.2%	7.8%	–	–	–
Me α-glucoside	24.8%	75.2%	ND[c]	30.7%	43.9%	25.4%
Me β-glucoside	3.8%	96.2%	ND[c]	–	–	–
Me α-mannoside	–	–	–	35.7%	42.0%	22.3%

[a] Reaction conditions: Y370g mutant enzyme (0.37 mg), 2FPh-NeuAc (5.0 mg) and acceptor (25.0 mg) in pH 5.25 acetate buffer (100 mM containing acetonitrile 30% v/v); incubated for 16 hrs at 37 °C. [b] Reaction conditions: Y370g mutant enzyme (0.37 mg), 2FPh-NeuAc (25.0 mg) and acceptor (55.0 mg) in pH 5.25 acetate buffer (100 mM containing acetonitrile 30% v/v); incubated for 64 hrs at 37 °C. [c] ND = not detected.

2.4.2. Trans-glycosylation product isolation and characterization

We used anion exchange chromatography to purify the products of mutant enzyme-catalyzed trans-glycosylation of a *N*-acetylneuraminyl group and the acceptors methyl α-D-galactopyranoside, methyl α-D-glucopyranoside, and methyl α-D-mannopyranoside, as well as that of wild-type 2,6-sialyltransferase with CMP-Neu5Ac and methyl α-D-galactopyranoside. Based on the known specificity of 2,6-

sialyltransferase, we assigned methyl *N*-acetyl-α-D-neuraminyl-(2→6)-α-D-galactopyranoside as the transglycosylation product (88). The ^{1}H and ^{13}C NMR spectra for three products formed in the *Mv*NA Y370g mutant-catalyzed coupling reactions are shown in Supporting Information (Figures S2.3–S2.8). All resonance assignments were based on $^{1}H-^{1}H$ COSY, $^{1}H-^{13}C$ HMQC NMR experiments, and all high resolution m/z ratios (in negative ion mode) for the three purified methyl *N*-acetyl-α-D-neuraminyl-(2→6)-α-D-glycopyranosides were in the m/z range of 484.1639–484.1687, which brackets that expected ($M-H^{+}$; m/z = 484.1666) for the molecular formula $C_{18}H_{31}NO_{14}$ (Supporting Information).

Based on a series of low intensity signals (Figure 2.2), the ^{13}C NMR spectrum for the mutant enzyme mediated coupling of a *N*-acetylneuraminyl glycosidic bond to methyl α-D-galactopyranoside is consistent with the formation of more than one regioisomer in this coupling reaction. However, the major product is an α-2→6 linked sugar as shown by a comparison of the ^{13}C NMR spectra for the products formed by the mutant neuraminidase and the 2,6-sialyltransferase (Figure 2.2). In addition, we acquired $^{1}H-^{13}C$ two dimensional HMBC correlation spectra, which show two and three bond ^{1}H to ^{13}C correlations for all isolated methyl glycoside coupling products. In the case of the mannoside coupled product, we observed the presence of cross peaks that are consistent with the major linkage being *N*-acetyl-α-D-neuraminyl-(2→6)-α-D-mannopyranoside. However, for the coupling product when methyl α-D-glucopyranoside was the acceptor we could not unambiguously determine the regiochemistry but we suggest that it is likely the 2→6-isomer, as compound that has been synthesized (76), although no ^{13}C NMR spectra was reported that we could have used to compare with our material. Also, we were unable to identify the regiochemistry of minor stereoisomers because of low NMR signal intensities and the complexity of the ^{1}H NMR spectra.

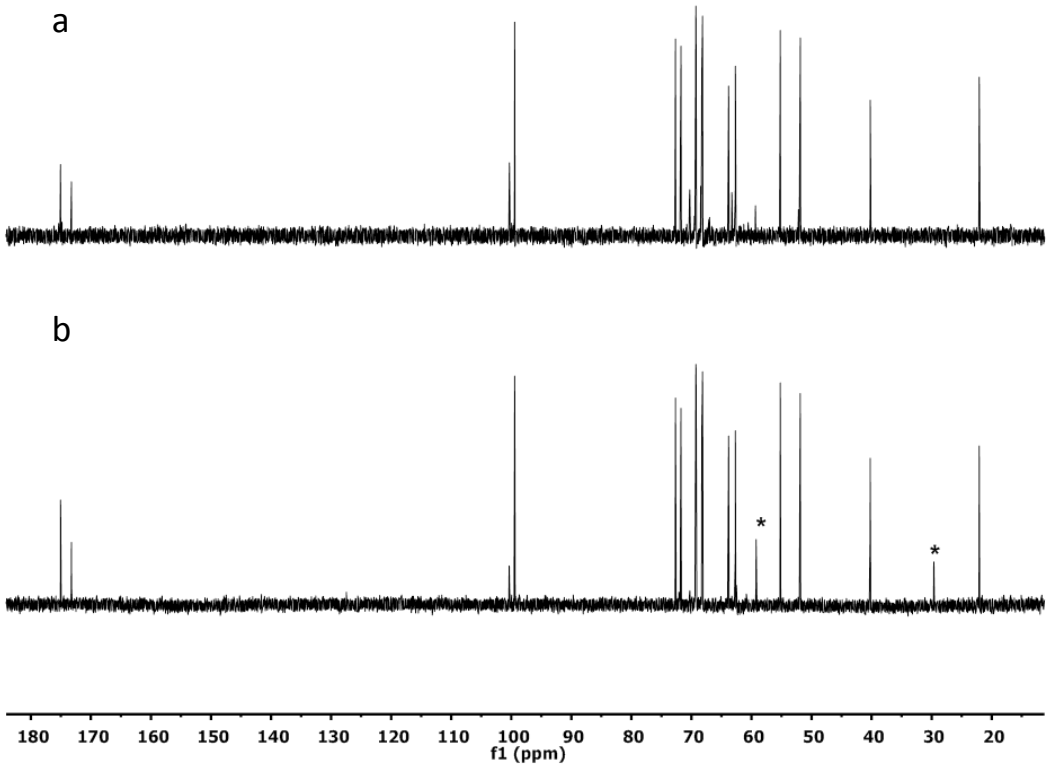


Figure 2.2 *^{13}C NMR spectra for the coupling reactions of methyl α-D-galactopyranoside to give methyl* N*-acetyl-α-D-neuraminyl-(2→6)-α-D-galactopyranoside.*

(a) reaction catalyzed by MvNA Y370g, (b) reaction catalyzed by photoacterium sp. JT-ISH-224 α-2,6-sialyltransferase using CMP-sialic acid as donor; the two peaks marked by asterisks are impurities.

2.5. Discussion

The approach of adding *N*-acetylneuraminic acid residues to carbohydrates using a neuraminidase that normally functions as a catabolic enzyme but that can work efficiently in a synthetic direction is especially advantageous because catabolic enzymes usually exhibit greater substrate promiscuity than anabolic enzymes. That is, sialyltransferases (*N*-acetylneuraminyl transferases; E.C. 2.4.99.X) generally display high selectivities for the neuraminide acceptor (which becomes the aglycone in the product), although a few examples of promiscuity with regard to the acceptor structure have been reported (107)(106). In contrast, wild-type *Mv*NA efficiently catalyzes the removal of Neu5Ac residues irrespective of both the identity of and position of coupling

to the aglycone (117)(1). The promiscuity of wild-type *Mv*NA towards the aglycone structure is preserved in the *Mv*NA Y370g mutant, as this genetically modified enzyme catalyzes trans-glycosylation of *N*-acetylneuraminic acid with a variety of aglycone acceptors, including lactose, glucose, raffinose (β-D-fructofuranosyl α-D-galactopyranosyl-(1→6)-α-D-glucopyranoside), allose and fructose, to give neuraminides products. Moreover, addition of the wild-type *Mv*NA to the mutant enzyme coupling mixtures removed these trans-glycosylation products, presumably by hydrolysis. Based on this observation, we conclude that the Y370g-catalyzed trans-glycosylation reactions form novel *N*-acetylneuraminides.

We identified the products of *Mv*NA Y370g-catalyzed coupling of the *N*-acetylneuraminyl group to various carbohydrate acceptors, focusing on methyl glycosides of sugars that yielded a single observable product peak in the HPLC-PAD chromatogram (Figure S2.1). We used commercial samples of methyl α- and β-D-glucopyranosides, methyl α- and β-D-galactopyranosides and methyl α-D-mannopyranoside in the coupling reactions and employed NMR spectroscopy to characterize the reaction products. The regiochemistry of coupling to the methyl galactoside acceptors was confirmed by comparing the results to those from reactions conducted with a wild-type sialyltransferase. Previously we noted that *Mv*NA Y370g preferentially forms 2,6-sialyl lactose during the coupling reaction with Ph-βNeu5Ac and lactose (118) despite the fact that the catalytic efficiency of the wild-type *Mv*NA (k_{cat}/K_m) is higher for 2,3-sialyl lactose as substrate (117). In order to minimize coupling reaction times, we employed the more reactive 2FPh-βNeu5Ac rather than Ph-βNeu5Ac as the *N*-acetylneuraminyl group donor. Specifically, 2FPh-βNeu5Ac has a 25-fold higher rate of turnover in a *Mv*NA Y370g-catalyzed reaction than the parent compound ($k_{cat(2FPh\text{-}\beta Neu5Ac)} \sim 25 \times k_{cat(Ph\text{-}\beta Neu5Ac)}$ (118)).

Our results suggest that for both methyl glucopyranosides and galactopyranosides it is the α-anomer that is the better acceptor, as it produces correspondingly higher coupling yields (Table 2.1, coupling experiment 1). This result is surprising, as the most common aglycone found in *N*-acetylneuraminides is a β-galactoside linkage that often is present in sialylated glycoconjugates as a lactose or a *N*-acetyllactosamine constituent (110)(111). The non-specific binding of methyl α-

glycopyranoside acceptors to the active site of *Mv*NA Y370g during catalysis results in a selection factor for coupling over hydrolysis of 24–30, we calculate this factor based on an acceptor concentration of ~1.13 M (55 mg in 250 μL), ~39 M H_2O (30% aqueous acetonitrile v/v) and the variation in the ratio of hydrolysis to coupling, with the extremes being 1.43 and 1.16 for reactions with methyl α-D-glucopyranoside and methyl α-D-galactopyranoside, respectively (Table 2.1; coupling experiment 2).

Previously, we reported that hydrolysis of *N*-acetyl-α-D-neuraminyl lactoses by wild-type *Mv*NA favors hydrolytic cleavage of the 2→3 regioisomer over the 2→6 isomer (117). Furthermore, we concluded that *Mv*NA-catalyzed hydrolysis reactions proceed via a transition state for glycosidic bond cleavage in which nucleophilic attack by the enzymatic tyrosine residue is not synchronous with cleavage of the glycosidic C–O bond (15, 17)(17), (15) (Figure 2.2a), that the transition state for deglycosylation of the wild-type *Mv*NA tyrosinyl bound intermediate involves little or no nucleophilic attack by water (Figure 2.2b) (16). Similarly, we reasoned that the mutant enzyme *Mv*NA Y370g catalyzes hydrolysis of aryl *N*-acetyl-β-D-neuraminides via a transition state that involves little or no nucleophilic attack by water (Figure 2.2c) (91).

If, as is likely, the *Mv*NA Y370g catalyzed coupling reactions proceed via a transition state that is similar to that of wild-type deglycosylation (Figure 2.2b–c), then the indiscriminate nature of these trans-glycosylation reactions must arise from a solely non-selective binding interaction between the acceptor sugar and the enzymatic active site. Therefore, increasing coupling efficiency and selectivity in these trans-glycosylation reactions only requires increasing the carbohydrate binding efficiency close to the catalytic center; and this increase could be accomplished by the mutation of proximal active site amino acid residues.

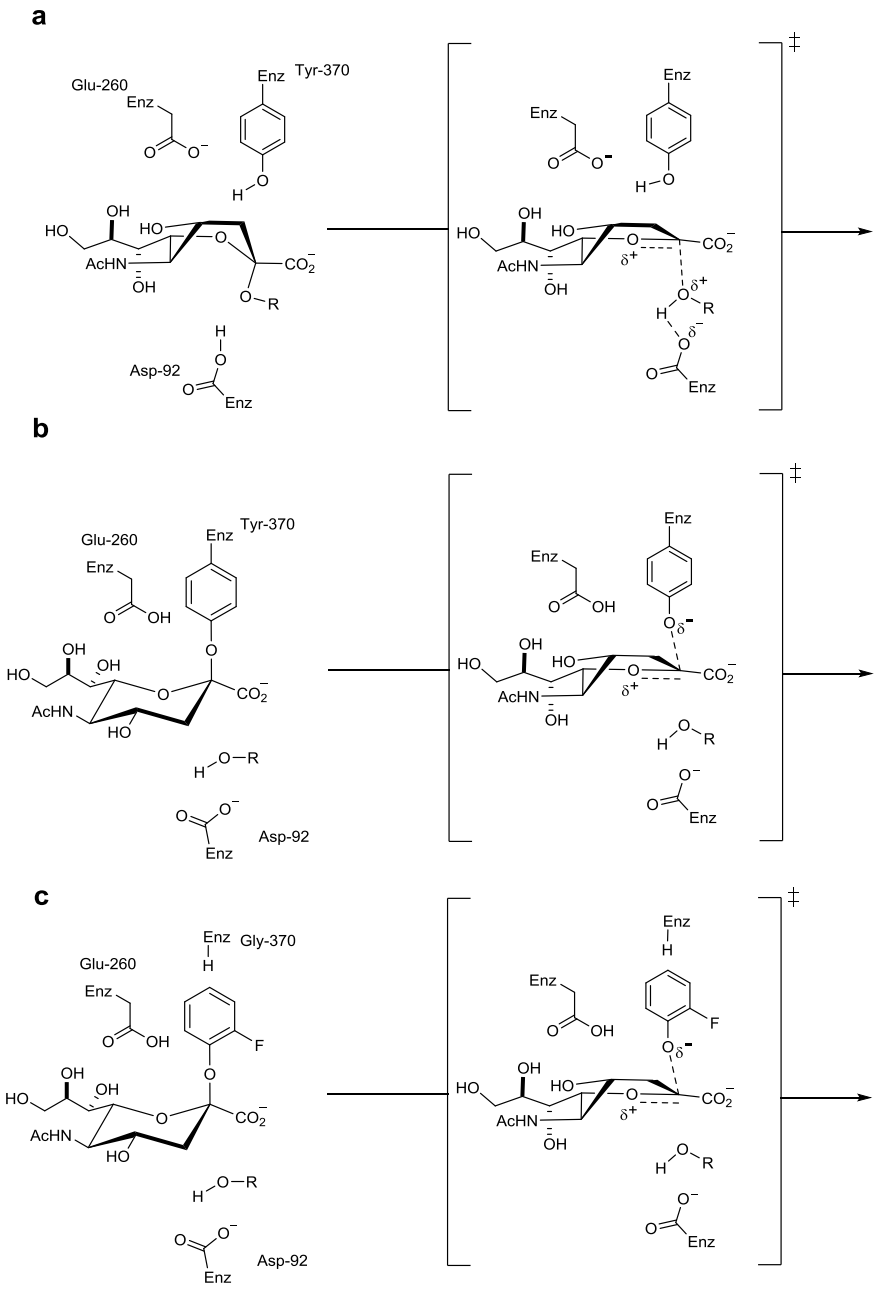


Figure 2.3 *Proposed transition states.*
(a) glycosylation of wild-type MvNA in which cleavage of the anomeric C–O bond is occurring with little or no nucleophilic assistance from the enzymatic tyrosine residue, (b) deglycosylation of wild-type MvNA in which cleavage of the tyrosinyl-bound intermediate occurs with little or no nucleophilic assistance, and (c) trans-glycosylation for MvNA Y370g catalysis in which cleavage of the anomeric C–O bond to the fluorophenyl leaving group is occurring with little or no nucleophilic assistance from non-selectively bound sugars (R = carbohydrate) or solvent waters (R = H).

2.6. Conclusion

A nucleophile mutant of the neuraminidase from *Micromonospora viridifaciens* catalyzes *N*-acetylneuraminyl transfer reactions for which it displays a high level of promiscuity toward carbohydrate acceptors and an accompanying low selectivity towards water. The formation of products that originate from the attack of a carbohydrate primary hydroxyl group on the nascent *N*-acetylneuraminyl cation are preferred to those initiated by the sterically more encumbered pyranosyl ring secondary ring alcohol functionalities.

To increase coupling efficiency we suggest that mutation of amino acid residues that are close to the catalytic site is required in order to incorporate a binding site for the carbohydrate acceptor.

2.7. Acknowledgements

Glycerol stocks of *Neisseria meningitides* CMP-Neu5Ac synthase and *photoacterium sp*. JT-ISH-224 α-2,6 sialyltransferase were kindly provided by Dr. Warren Wakarchuk.

2.8. Supporting Information

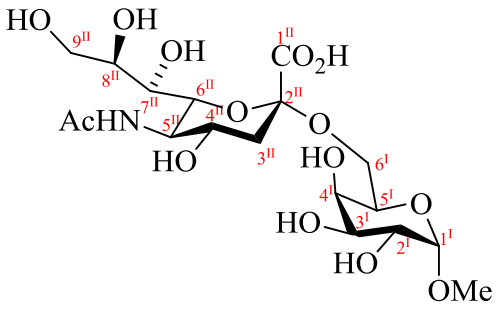
HPLC-PAD chromatograms of *Mv*NA Y370g mutant-catalyzed trans-glycosylation reactions (Figure S2.1). ^{1}H NMR spectra of *Mv*NA Y370g mutant-catalyzed reaction mixtures for methyl α- and β-D-glucopyranoside couplings (Figure S2.2). NMR spectroscopic characterization data for, and ^{1}H and ^{13}C spectra of, three methyl *N*-acetyl-α-D-neuraminyl-(2→6)-α-D-glycopyranosides (Figures S2.3–S2.8). This material is available free of charge via the Internet at http://pubs.acs.org.

2.8.1. General Methods

^{1}H- and ^{13}C-NMR spectra were acquired on a Bruker 600 MHz spectrometer equipped with a QNP cryoprobe, chemical shifts are listed in parts per million, and coupling constants (*J*) are given in hertz.

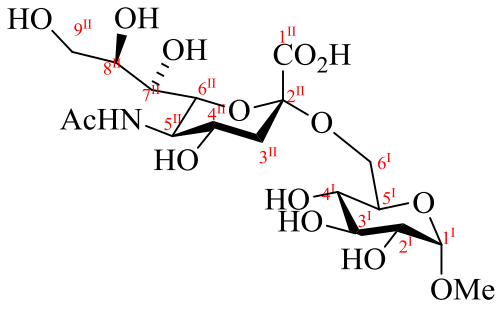
2.8.2. Methyl 5-N-acetyl-α-neuraminyl-(2→6)-α-D-galactopyranoside:

^{1}H NMR (600 MHz, D_2O) δ: 1.76 (t, 1 H, $J_{3a,4}$ = 12.1, H-3$^{II}_{ax}$), 2.05 (s, 3 H, NHCOC*H*$_3$), 2.74 (dd, 1 H, $J_{3e,3a}$ = 12.5, $J_{3e,4}$ = 4.6, H-3$^{II}_{eq}$), 3.43 (s, 3 H, OCH_3), 3.59 (dd, 1 H, $J_{7,8}$ = 8.9, $J_{7,6}$ = 1.6, H-7^{II}), 3.63–3.71 (m, 2 H, H-6$_a^{I}$, H-9$_a^{II}$), 3.71–3.81 (m, 2 H, H-4^{II}, H-6^{II}), 3.80–3.92 (m, 5 H, H-2^{I}, H-5^{I}, H-5^{II}, H-8^{II}, H-9$_b^{II}$), 3.95 (t, 1 H, $J_{6b,6a}$ = $J_{6b,5}$ = 10.1, H-6$_b^{I}$), 3.97–4.05 (m, 2 H, H-3^{I}, H-4^{I}), 4.74–4.89 (m, 1 H, H-1^{I} overlapped with solvent peak, measured using COSY and HSQC); ^{13}C NMR (150 MHz, D_2O) δ: 21.56 (NHCO*C*H$_3$), 39.35 (C-3^{II}), 51.30 (C-5^{II}), 54.67 (O*C*H$_3$), 62.27 (C-9^{II}), 63.28 (C-6^{I}), 67.39 (C-4^{II}), 67.60 (C-2^{I}), 67.78 (C-7^{II}), 68.57–68.99 (3 C, C-3^{I},C-4^{I}, C-5^{I}), 70.93 (C-8^{II}), 72.23 (C-6^{II}), 98.90 (C-1^{I}), 99.29 (C-2^{II}), 171.99 & 174.50 (2 C, C-1^{II} and NHAc).HRMS-FAB (*m*/*z*): [M–H^{+}] calcd for $C_{18}H_{30}NO_{14}$, 484.1666; found, 484.1687.



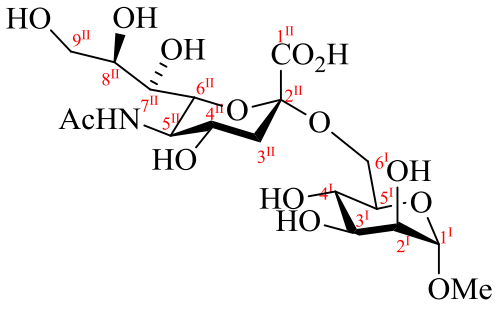
2.8.3. Methyl 5-N-acetyl-α-neuraminyl-(2→6)-α-D-glucopyranoside:

^{1}H NMR (600 MHz, D_2O) δ: 1.73 (t, 1 H, $J_{3a,4}$ = 11.1, H-3$^{II}_{ax}$), 2.05 (s, 3 H, NHCOC*H*$_3$), 2.76 (dd, 1 H, $J_{3e,3a}$ = 11.9, $J_{3e,4}$ = 3.8, H-3$^{II}_{eq}$), 3.43 (s, 3 H, OCH_3), 3.49 (t, 1 H, $J_{4,3}$ = $J_{4,5}$ = 9.5, H-4^{I}), 3.59 (dd, 1 H, $J_{2,3}$ = 9.8, $J_{2,1}$ = 3.5, H-2^{I}), 3.56–3.82 (m, 7 H, H-3^{I}, H-5^{I}, H-6$_a^{I}$, H-4^{II}, H-7^{II}, H-6^{II}, H-9$_a^{II}$), 3.82–3.95 (m, 3 H, H-6$_b^{I}$, H-5^{II}, H-9$_b^{II}$), 3.99 (dd, 1 H, $J_{8,9a}$ = 11.9, $J_{8,9b}$ = 3.8, H-8^{II}), 4.69–4.94 (m, 1 H, H-1^{I} overlapped with solvent residual peak and measured by COSY and HSQC); ^{13}C NMR (150 MHz, D_2O) δ: 21.54 (NHCO*C*H$_3$), 39.61 (C-3^{II}), 51.39 (C-5^{II}), 54.64 (O*C*H$_3$), 62.13 (C-9^{II}), 62.32 (C-6^{I}), 67.77–67.79 (2 C, C-4^{II}, C-6^{II}), 68.98 (C-4^{I}), 69.71 (C-8^{II}), 70.64 (C-2^{I}), 71.28–72.51 (3 C, C-3^{I}, C-5^{I}, C-7^{II}), 98.75 (C-1^{II}), 99.65 (C-2^{II}), 172.93–174.56 (2 C, C-1^{II} and NHAc). HRMS-FAB (*m*/*z*): [M – H$^+$] calcd for $C_{18}H_{30}NO_{14}$, 484.1672; found, 484.1649.



2.8.4. Methyl 5-N-acetyl-α-neuraminyl-(2→6)-α-D-mannopyranoside:

^{1}H NMR (600 MHz, D_2O) δ: 1.77 (t, 1 H, $J_{3a,4}$ = 11.9, H-3$^{II}_{ax}$), 2.06 (s, 3 H, NHCOC*H*$_3$), 2.76 (dd, 1 H, $J_{3e,3a}$ = 12.3, $J_{3e,4}$ = 4.1, H-3$^{II}_{eq}$), 3.42 (s, 3 H, OCH_3), 3.61 (dd, 1 H, $J_{7,8}$ = 8.8, $J_{7,6}$ = 1.2, H-7^{II}), 3.67 (dd, 1 H, $J_{9a,9b}$ = 11.8, $J_{9a,8}$ = 5.8, H-9$_a^{II}$), 3.69–3.74 (m, 3 H, H-4^{II}, H-4^{I}, H-5^{I}), 3.74–3.81 (m, 2 H, H-6$_a^{I}$, H-3^{I}), 3.84–3.98 (m, 5 H, H-2^{I}, H-9$_b^{II}$, H-8^{II}, H-5^{II}, H-6^{II}), 4.01 (dd, 1 H, $J_{6a,6b}$ = 10.7, $J_{6b,5}$ = 2.7, H-6$_b^{I}$), 4.77 (m, 1 H, H-1^{I} overlapped with solvent peak, measured using COSY and HSQC); ^{13}C NMR (150 MHz, D_2O) δ: 21.55 (NHCO*C*H$_3$), 39.42 (C-3^{II}), 51.36 (C-5^{II}), 54.32 (O*C*H$_3$), 62.19 (C-9^{II}), 62.60 (C-6^{I}), 66.07–67.61(2 C, C-4^{II}, C-4^{I}), 67.78 (C-7^{II}), 69.35 (C-2^{I}), 69.97–70.57 (2 C, C-5^{I}, C-3^{I}), 71.06–72.10 (2 C, C-6^{II}, C-8^{II}), 99.39 (C-2^{II}), 100.38 (C-1^{I}), , 172.50–174.54 (2 C, C-1^{II} and NHAc).HRMS-FAB (*m*/*z*): [M–H$^+$] calcd for $C_{18}H_{30}NO_{14}$, 484.1666; found, 484.1639.



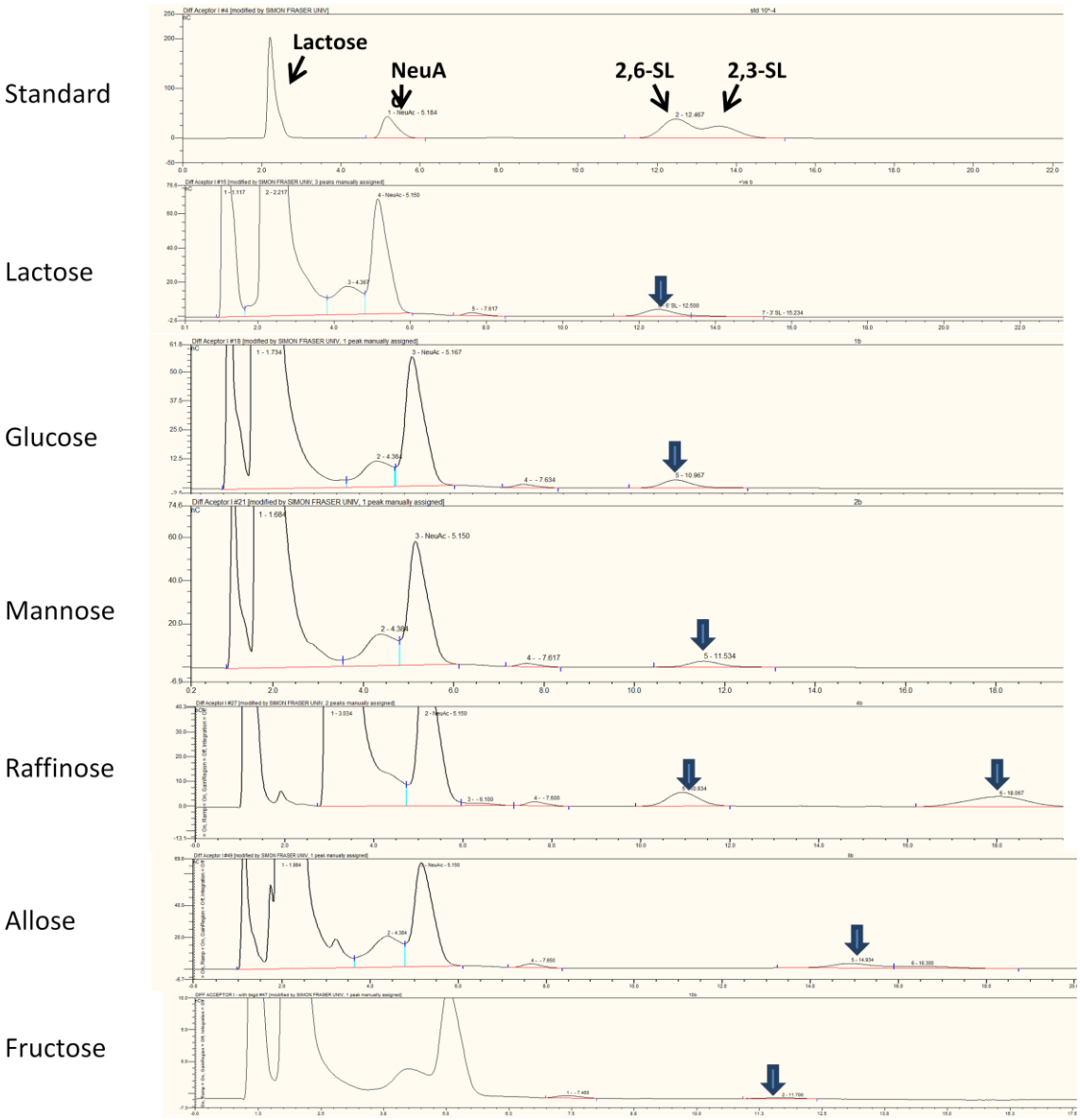


Figure S 2.1 *HPLC-PAD chromatograms of MvNA Y370g mutant-catalyzed trans-glycosylation reactions.*

The peaks that correspond to a new signal that disappeared on treatment with wild-type MvNA are highlighted by the blue arrows. The two neuraminide standards are 2,6-sialyl lactose (2,6-SL) and 2,3-sialyl lactose (2,3-SL).

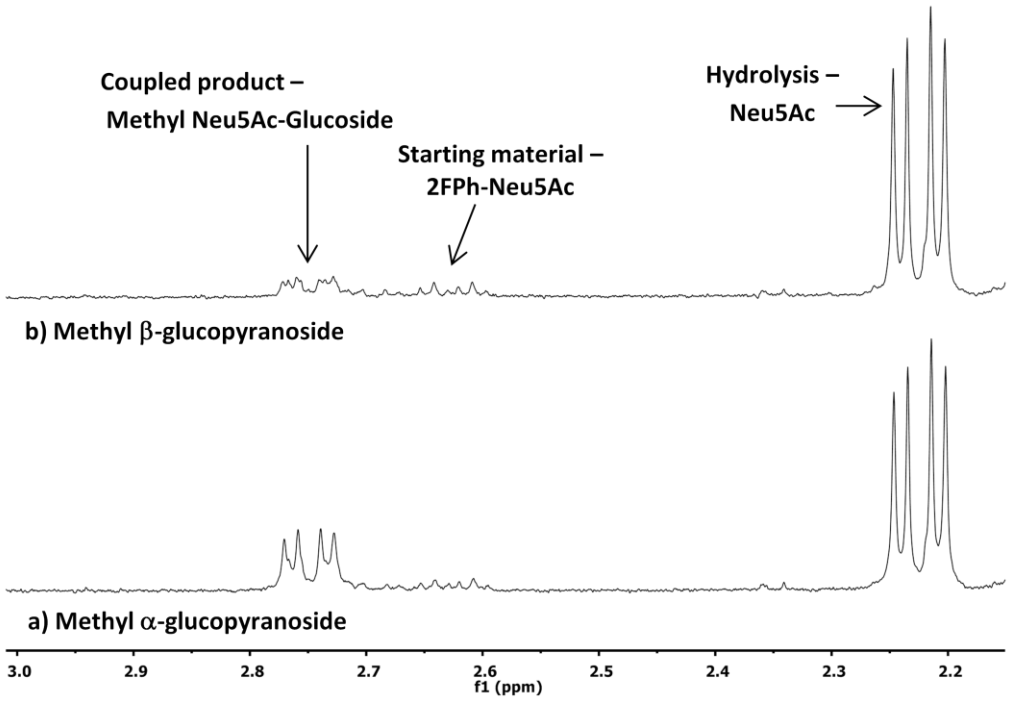


Figure S 2.2 *Slices of the ^{1}H NMR spectra showing the resonances for the various H-3_{eq} protons of the coupling reaction mixture for the* Mv*NA Y370g mutant-catalyzed coupling of 2FPh-βNeu5Ac to an acceptor sugar. (a) methyl α-D-glucopyranoside acceptor and (b) methyl β-D-glucopyranoside acceptor.*

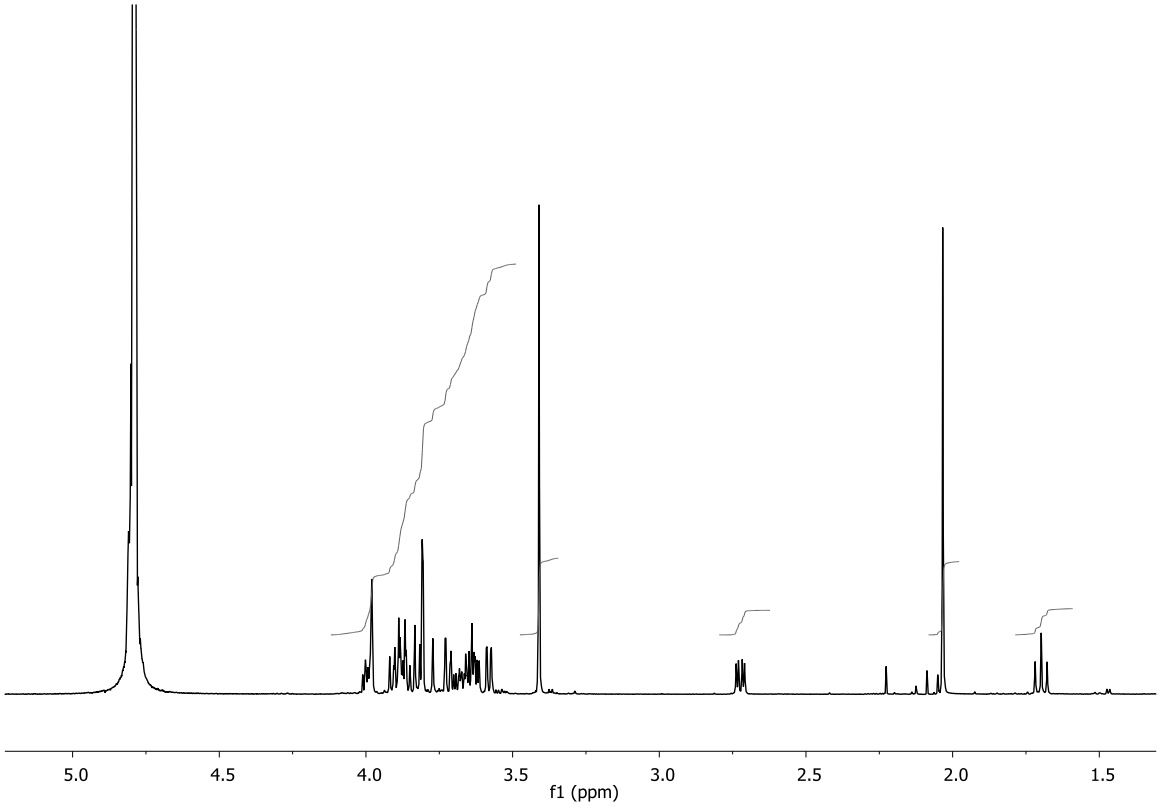


Figure S 2.3 *^{1}H-NMR spectrum of methyl 5-*N*-acetyl-α-neuraminyl-(2→6)-α-D-galactopyranoside made by* Mv*NA Y370g mutant-catalyzed coupling of 2FPh-βNeu5Acto methyl α-D-galactopyranoside.*

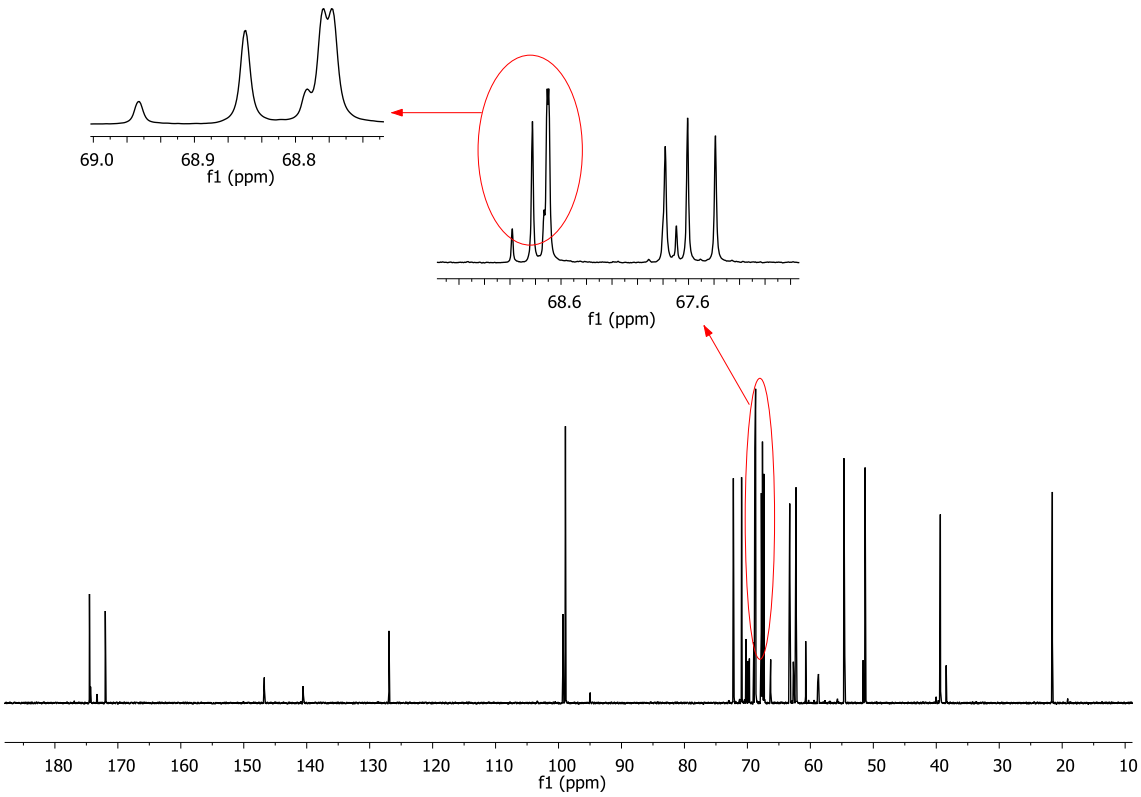


Figure S 2.4 *^{13}C-NMR spectrum of methyl* 5-N-*acetyl-α-neuraminyl-(2→6)-α-D-galactopyranosidemade by* MvNA *Y370g mutant-catalyzed coupling of 2FPh-βNeu5Acto methyl α-D-galactopyranoside.*

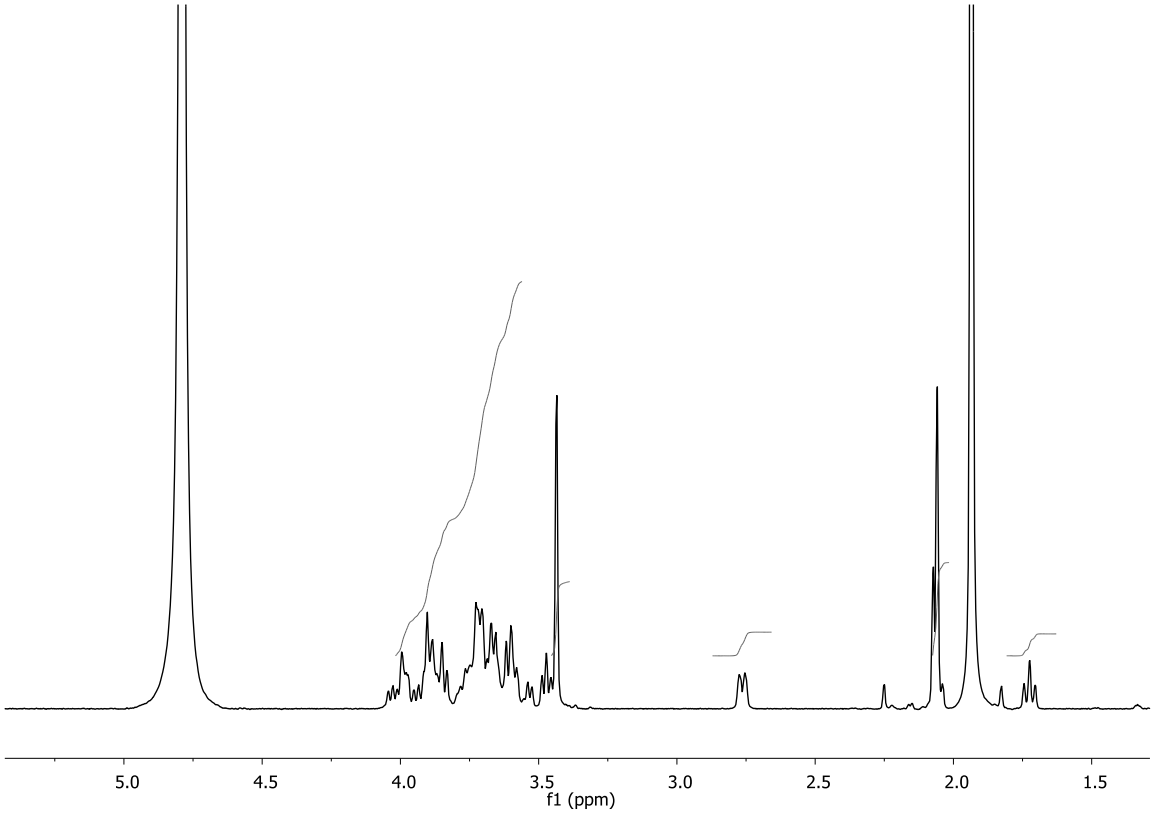


Figure S 2.5 *^{1}H-NMR spectrum of methyl 5-*N*-acetyl-α-neuraminyl-(2→6)-α-D-glucopyranosidemade by* Mv*NA Y370g mutant-catalyzed coupling of 2FPh-β*Neu5Ac *to methyl α-D-glucopyranoside.*

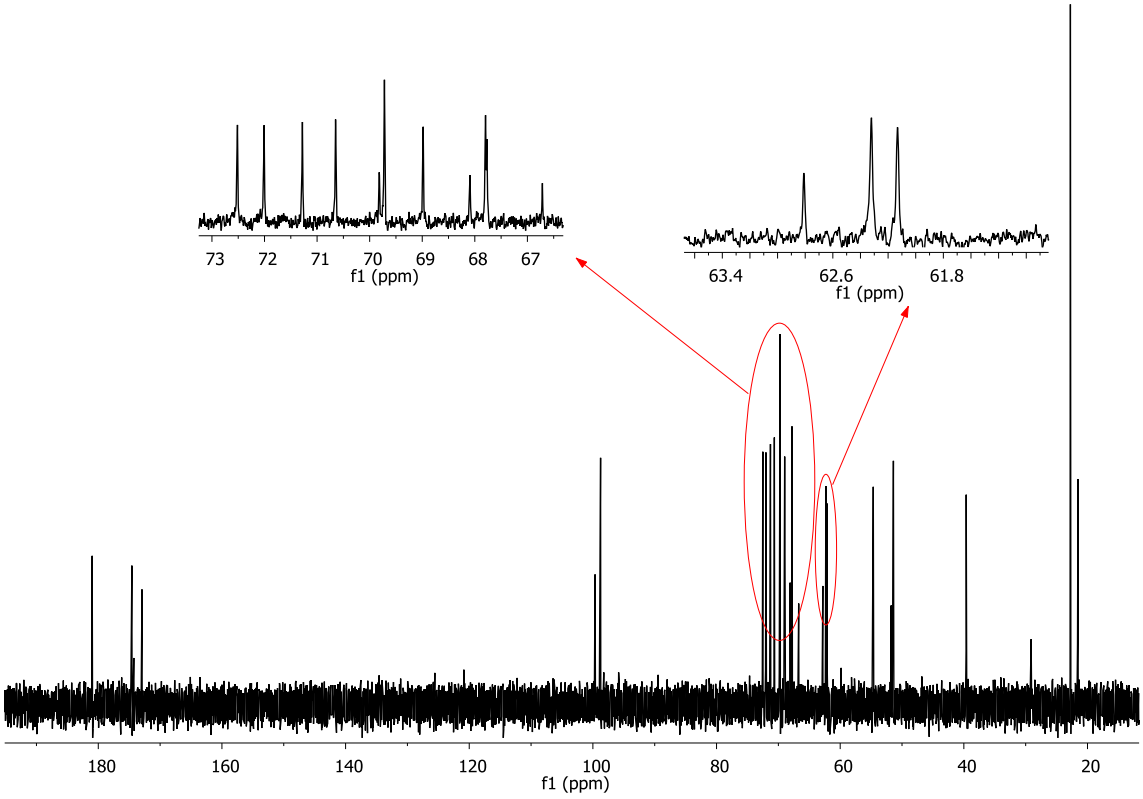


Figure S 2.6 *^{13}C-NMR spectrum of methyl* 5-N-*acetyl-α-neuraminyl-(2→6)-α-D-glucopyranosidemade by* MvNA *Y370g mutant-catalyzed coupling of 2FPh-βNeu5Ac to methyl α-D-glucopyranoside.*

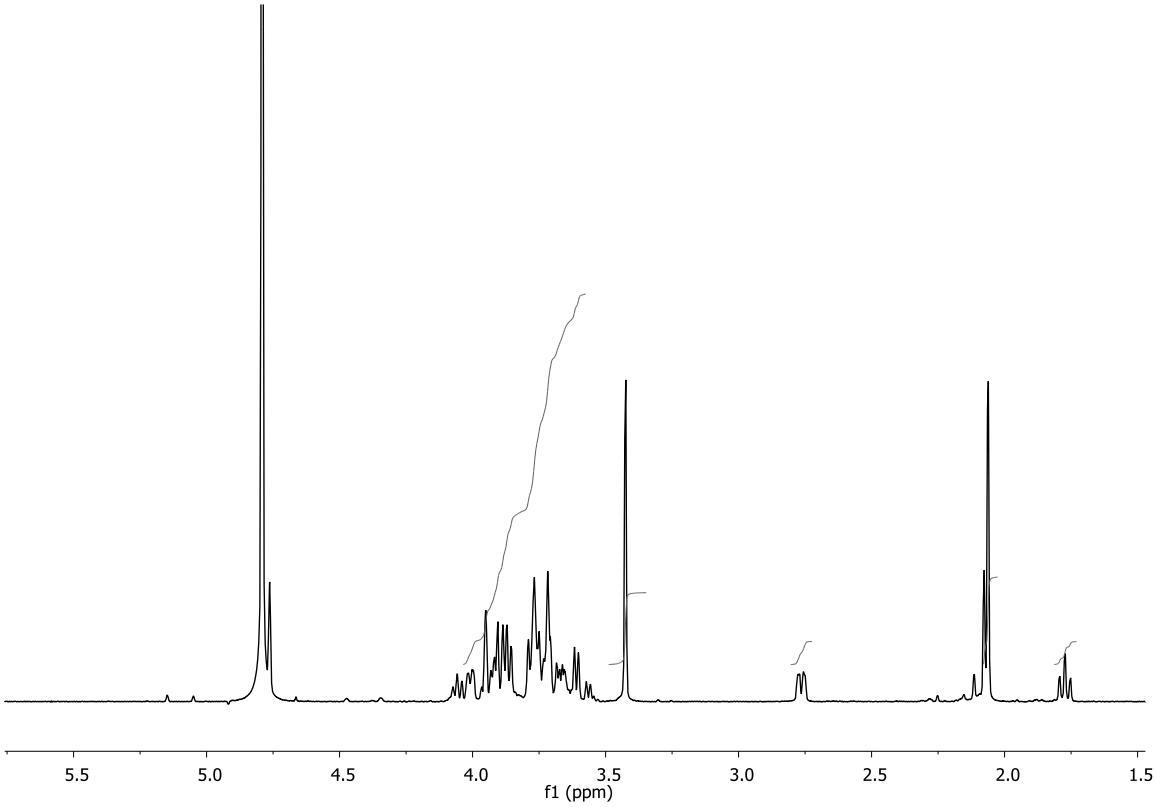


Figure S 2.7 *^{1}H-NMR spectrum of methyl 5-*N*-acetyl-α-neuraminyl-(2→6)-α-D-mannopyranosidemade by* Mv*NA Y370g mutant-catalyzed coupling of 2FPh-βNeu5Ac to methyl α-D-mannopyranoside.*

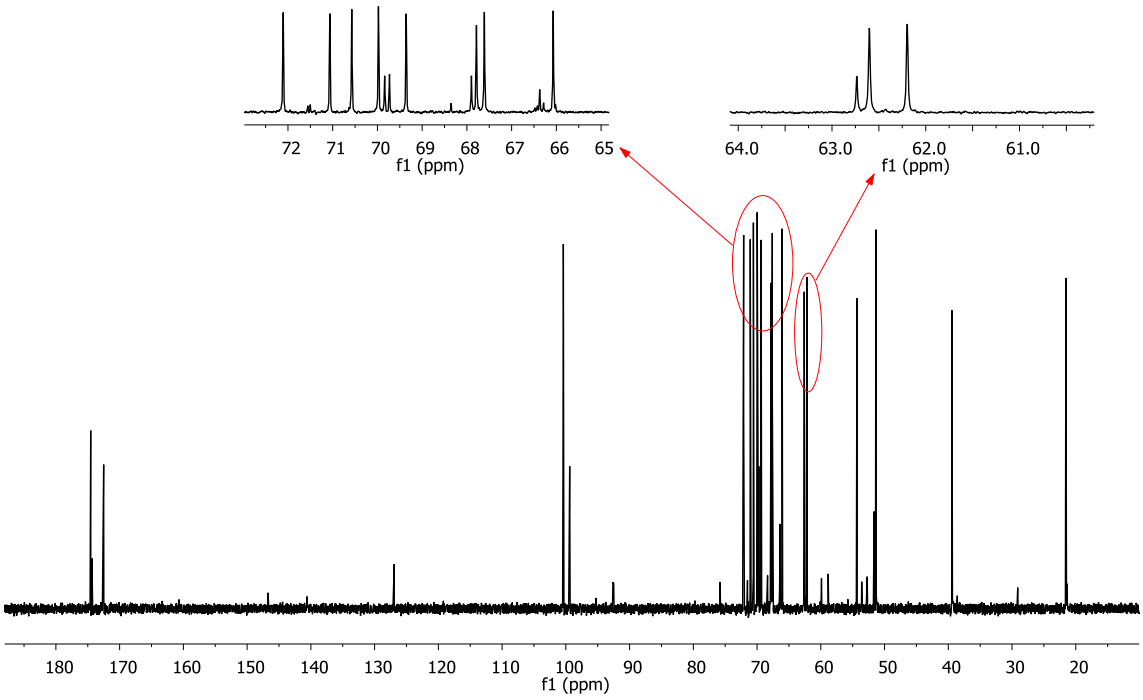


Figure S 2.8 *^{13}C-NMR spectrum of methyl 5-*N*-acetyl-α-neuraminyl-(2→6)-α-D-mannopyranosidemade by* MvNA *Y370g mutant-catalyzed coupling of 2FPh-βNeu5Ac to methyl α-D-mannopyranoside.*

3. Library Screening and Analyzing the Y370G Mutant Sialidase with Hydrophobic Loops

3.1. Introduction

Sialic acids are a family of nine carbon monosaccharides that are found throughout nature in species that range from mammals, plants, fungi, yeasts to bacteria (113). These monosaccharide units are regularly positioned on the termini of glycoconjugates and as a result they have vital roles in numerous biological interactions (22). Given the importance of terminal sialic acid residues we planned to alter the serum half life of glycoconjugates in the circulation as a possibly useful therapeutic approach (7). In other words, if sialic acid residues can be transferred onto glycoconjugate drugs, which have been produced recombinantly, then the life time of the resulting therapeutic agent in the circulation will be increased and thus the dosage of the agent can be reduced.

The aim of this project was to create and to screen a library of Y370G mutant sialidases that incorporated a proximal sugar binding site so that transglycosylation was improved. Of note, the Y370G mutant has been shown to possess a low activity against natural substrates (121). Furthermore, the results given in the previous chapter demonstrated the flexibility of the Y370G mutant sialidase to accommodate different acceptor sugars when catalyzing a transglycosylation reaction. The transfer of sialic acid from 2-fluorophenyl β- D-sialoside to methyl α- D-galactopyranoside, methyl α-D-glucopyranoside, and methyl α-D-mannopyranoside are 46.2%, 41.2%, and 46.0% respectively. The critical coupling conditions include using 30% acetonitrile and just over a 2-fold excess of the acceptor sugar relative to the donor sugar. The use of organic solvent and excess acceptor sugar helped increase the yield of coupled disaccharide. In this chapter the results of installing a sugar binding site into the Y370G mutant sialidase are reported.

The inspiration for this project came from the *Trypanosoma cruzi trans*-sialidase, (13) an enzyme that very efficiently catalyzes sialyl-transferase reactions. One of the structural features that promotes transferase activity is a hydrophobic binding pocket which binds lactose, that is positioned above the catalytic site (14). This hydrophobic pocket contains both a tryptophan (W312) and tyrosine (Y119) residues that are ideally positioned to bind lactose (14). Thus, we proposed to install a similar hydrophobic pocket onto our MvNA Y370G mutant enzyme.

The comparison of the MvNA (1EUR) and TcTS (1MSO) active site domains is shown in Figure 3.1.

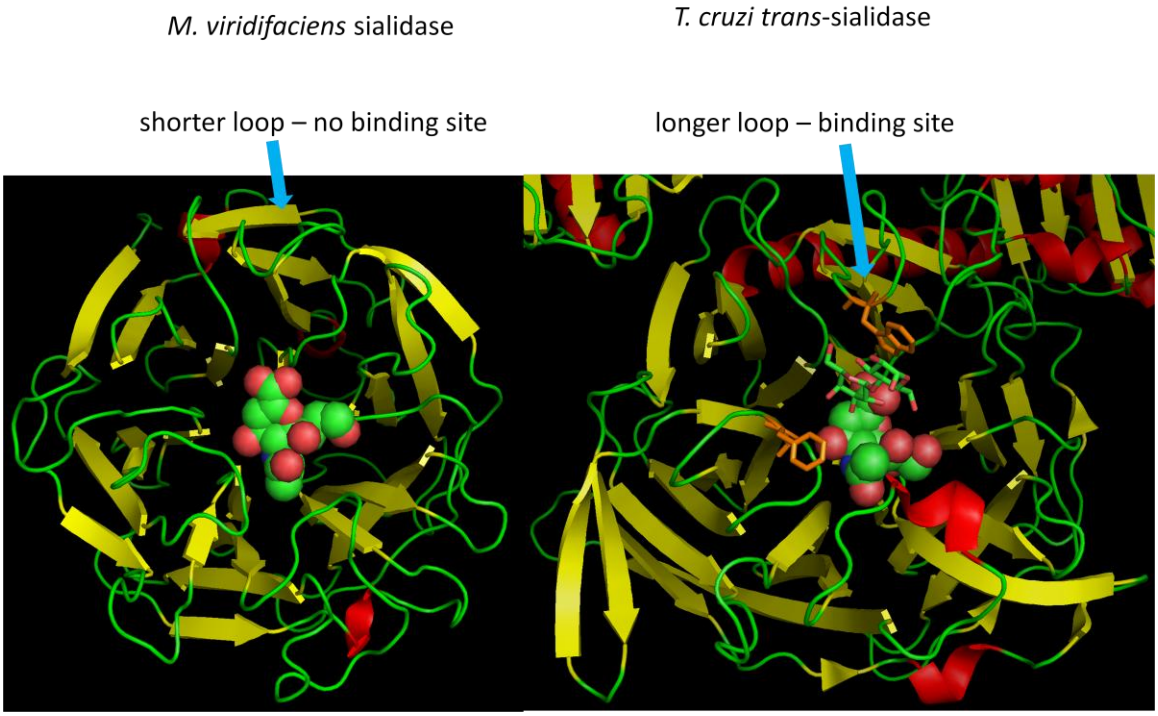


Figure 3.1 *Crystal structure comparison of M. viridifaciens sialidase and T. cruzi trans-sialidase.*
a) The crystal structure of M. viridifaciens sialidase with DANA inhibitor bound at the active site show in space filling structure and nucleophile Y370 show in ball and stick (PDB code: 1EUR). b) The crystal structure of T. cruzi trans-sialidase with lactose bound at the active site in space filling structure and nucleophile Y342 and substrate binding residues Y119 and W312 show in stick representation (PDB code: 1MSO).

The tryptophan containing loop in *Tc*TS consists of six amino acids (F308, K309, G310, R311, W312, and L313) whereas the *Mv*NA enzyme contains only two amino acids (T340 and S341) at this location (Figure 3.1). Therefore, the two amino acids in *Mv*NA were swapped for the six amino acids in the *Tc*TS loop to try and install a carbohydrate binding motif. Also, a second mutation series included replacing an additional residue with a conservative change in which glutamine 339 (*Mv*NA) was replaced by an asparagine (*Tc*Ts N307). These two variations in the Trp containing loops were named as TS6 (for the mutant *Mv*NA in which TS340-341 were replaced with six amino acids) and QTS7 (where QTS 339-341 *Mv*NA were replaced with seven amino acids). The Tyr containing loop in *Tc*TS (Y119) is the same length as that found in *Mv*NA. Therefore, a simple substitution was made to install a Tyr residue into the sequence (G154Y) in order to form the second important binding interaction. In addition, the *Tc*TS Y119 has an adjacent Trp that may be important, so a second *Mv*NA Tyr loop mutant was also made (G154Y/F155W – orange loops Figure 3.1). The hydrophobic loops inserted or substituted on Y370G mutants are shown in Figure 3.2.

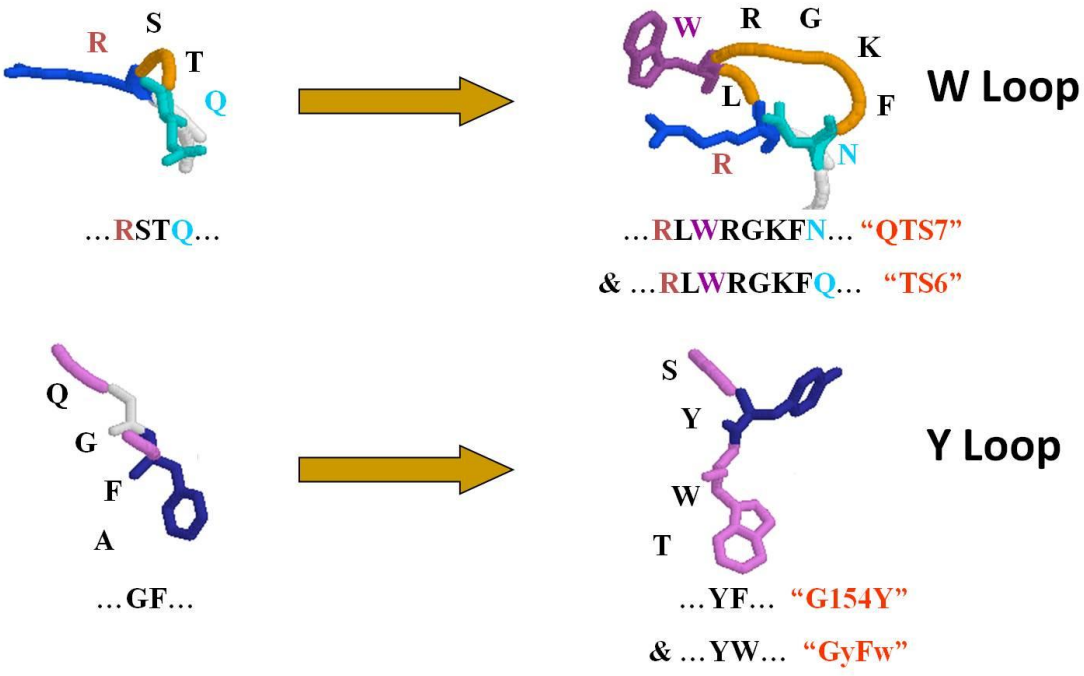


Figure 3.2 *Summary of hydrophobic loop mutations made on MvNA Yg mutant sialidase.*

The research detailed in this chapter involved using these four different tryptophan and tyrosine loops Yg mutant as templates for the generation of mutant proteins libraries followed by screening of these libraries for transglycosylation activity (TS6-Yg-Gy, TS6-Yg-GyFw, QTS7-Yg-Gy, and QTS7-Yg-GyFw) followed by characterization of selected candidate proteins.

3.2. Materials and Methods

3.2.1. Chemicals and reagents.

Bactotryptone and urea were purchased from Bioshop Inc. Yeast extract was purchased from EMD chemicals. Glycerol was purchased from VWR. Potassium phosphate, dibasic and monobasic were purchased from Caledon Laboratories Ltd., imidazole was purchased from AK Scientific, Inc. and IPTG was purchased from Invitrogen. All other chemicals, reagents and enzymes were purchased from Sigma-Aldrich Canada, Ltd.

3.2.2. Substrate synthesis.

2-Fluorophenyl β-D-*N*-acetylneuraminide was made by Fahimeh Shidmoossavee according to the literature procedure (92), (29, 93)(20, 65). The following substrates were made by Dr. Jefferson Chan, FMU-man; Neu5Ac α-2,6 Gal β-FMU; Gal β-FMU(18); and Gal-ELF (28, 74). NMR spectra were acquired on either a Bruker 600 or 400 MHz spectrometer. Chemical shifts are reported in parts per million downfield from the signal for TMS. The residual signal from deuterated chloroform and that from external TMS-salt (D_2O) were used as ^{1}H NMR references; for ^{13}C NMR spectra, natural abundance signals from $CDCl_3$ and external TMS-salt (D_2O) were used as references. Coupling constants (J) are given in Hertz.

3.2.3. Enzymes.

Micromonospora viridifaciens sialidase was expressed and purified as previously reported (120). Vent DNA polymerase and T4 DNA ligase were purchased from New England Biolabs. All restriction endonucleases were purchased from Thermo Scientific.

3.2.4. W & Y Loop Library Construction.

The construction and expression of the mutant protein library followed a previously published method with minor modifications mentioned in the following (121). A mixture of the four pJWβ plasmids containing the TS6-Yg-Gy, TS6-Yg-GyFw, QTS7-Yg-Gy and QTS7-Yg-GyFw mutant sialidase genes were used as the template DNA to create the library. Site-directed mutations were made to the following four amino acids FKGR that are a component of the "W loop" – N/QFKGRWL. Four primers were designed to produce a library of mutants using the strand-overlap extension method (Table 3.2).

Table 3.1 *DNA oligonucleotide sequences.*
N refers to any four nucleotides, 25 % of A, T, C, or G. K refers to 50 % of either T or G nucleotide and M refers to 50 % of either A or C nucleotide. The red G and C nucleotides are where the substitution site.

Primer	Nucleotide Sequence
Q/N4N-F'	5'-GCCGCCAGTMANNNKNNKNNKNNKTGGCTGCGCAGTCAGGGC-3'
Q/N4N-R'	5'-GCGCAGCCAMNNMNNMNNMNNNTKACTGGCGGCGTTGGAG-3'
Eco*'	5'-CCGGAATTCACTGCGAATCCGTACCTCCGC-3'
ggHinR'	5'-CCCAAGCTTCAGCCAGGCGAGGTTG-3'
C-I-25-A5-F'	5'-CCTGACTGCGCA**G**GCCAATTCGAAGGACG-3'
C-I-25-A5-R'	5'-CGAATTGGC**C**TGCGCAGTCAGGGCACC-3'
Cut-OmpT'	5'-GGCCGAATTCGGGGCACCCGTCCC-3'

Mutagenesis was ensured by making the mutations in each strand in separate PCR experiments. The 5'-end (front) of the gene was amplified using the Eco*' primer with the reverse mutagenic primer Q/N4N-R' to give a 1.04 kb product. The 3'-end (back) of the gene was amplified using the forward mutagenic primer Q/N4N-F' with the ggHinR' reverse primer to produce a 0.20 kb fragment. These two PCR fragments were purified and joined by using an overlapping sequence at the site of mutation to prime each other in an extension reaction of 1 cycle. The whole 1.24 kb fragment was then amplified with Eco*' and ggHinR' primers. All PCR reactions used vent DNA polymerase with 1x thermo polymerase buffer in 10% DMSO and 0.25 mM dNTPs. The front and back fragment portion of PCR reactions were performed at a denaturing temperature of 95 °C for 1 min (1 cycle) followed by denaturing at 95 °C for 0.5 min, annealing at 55 °C for 1 min, elongation at 72 °C for 1 min (28 cycles), the final elongation at 72 °C was allowed to react for 2 min (1 cycle), and the resulting samples were stored at 4 °C prior to the next step. Both the overlapping extension PCR and the full fragment amplification were done at 50 °C (annealing temperature) with the rest of the PCR conditions being the same as those detailed above. The full 1.24 kb gene and pJWβ plasmid were digested with *Eco* RI and *Hind* III. The digested insert gene and plasmid vector were separated on 1% agarose gel and were then sliced and purified. The purified restriction enzyme-digested insert was ligated into purified digested vector with 2:1 ratio using T4 DNA ligase. The ligation reaction was incubated at 16 °C overnight followed by the addition of 1 µL fresh ligase and incubation was continued for another 3 hours at room temperature. The ligation mixture was then transformed into *E. coli* BL21-Gold(DE3) competent cells (Cedar Lane Laboratories). Twenty clones were picked to perform restriction and sequence diversity check. The four residues were completely randomized using saturation mutagenesis which would create 4.2×10^6 of possible mutant clones (8 × 32 × 32 × 32 × 32) and encode 1.6×10^5 different proteins (4 × 20 × 20 × 20 × 20).

3.2.5. Library Expression.

Using the ligation mixture, 30–40 ng of DNA was used to transform *E. coli* BL21-Gold (DE3) competent cells. Transformations were carried out according to the

instructions of the manufacturer. Approximately 5000–6000 colonies were observed and dispensed into 31 U-bottomed 96 well-plates (Falcon) using colony picker (Genetix QPix 2). Into each well was added 200 µL of Luria broth containing 30 µg/ml of kanamycin and then each plate was covered with an airpore (Qiagen) membrane. Cells were grown at 37 °C and 200 rpm for 16 h. An aliquot 150 µL of each culture was transferred to a fresh plate containing 20 µL of 80 % sterile glycerol for long-term storage at –80 °C. To the remaining 50 µL of culture, 200 µL of TB/Kan (12 g of bactotryptone, 24 g of yeast extract, 4 mL of glycerol, and 100 mL of K phosphate in 1 L containing 30 µg/mL kanamycin) was added and returned to incubator for 6–7 h. Plates were placed on ice for 5 min before 1 mM final concentration of IPTG was added. Following induction the plates were kept for 16 hrs at room temperature (180 rpm). After placing the resultant plates on ice for 10 min the plates were centrifuged (3000 rpm 20 min at 4 °C). The supernatants were then used to screen for sialyltransferase activity. A number of individual crude supernatants were analyzed by SDS-PAGE to compare the variation in expression levels between individual mutants in the library (Figure 3.3).

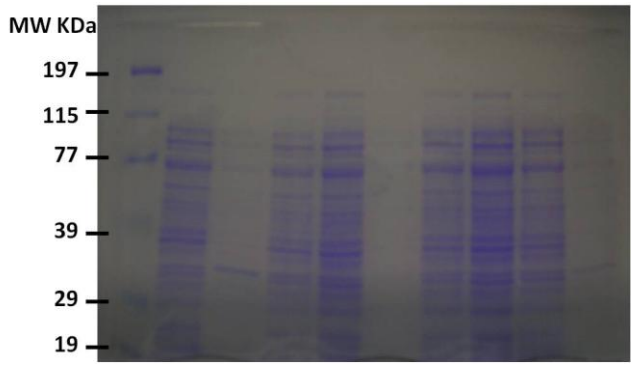
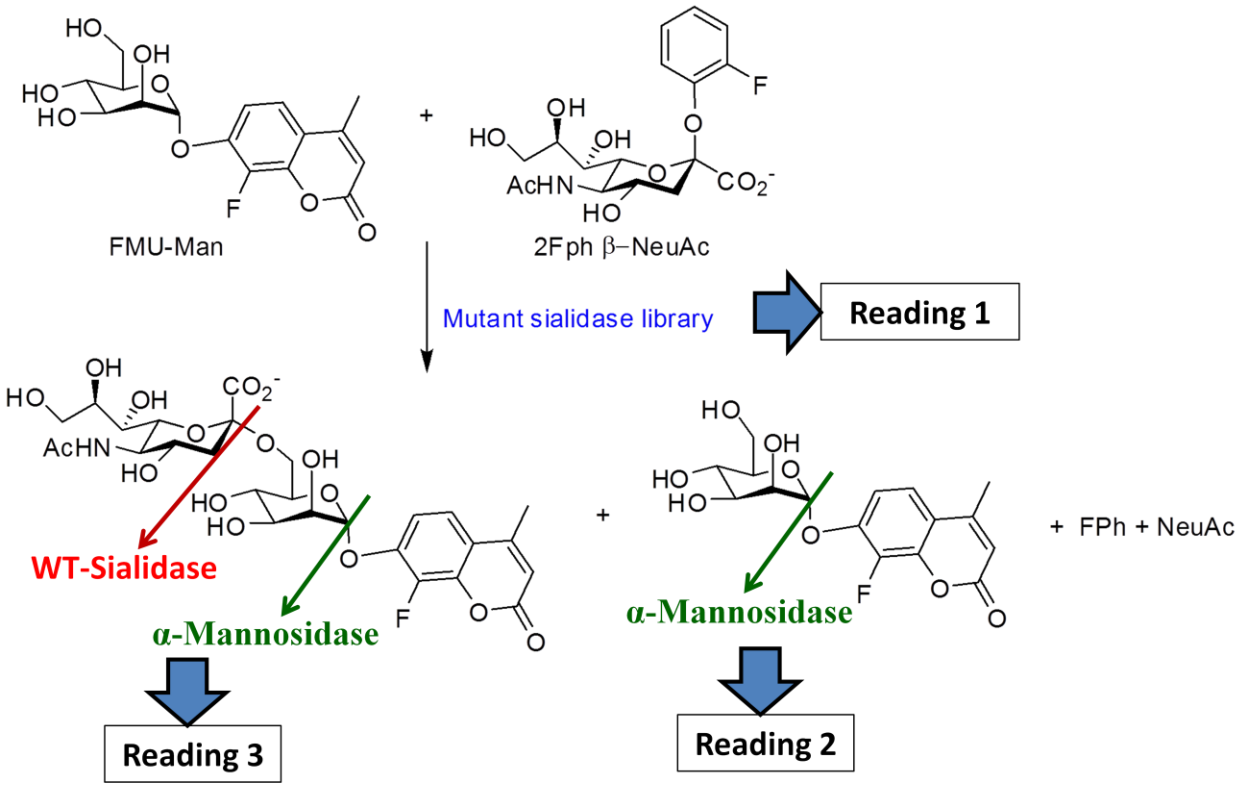


Figure 3.3 *SDS-PAGE gel of selective library mutant clones supernatant protein expression level.*

3.2.6. Library Screening Protocol 1.

Supernatants containing expressed mutant proteins were transferred to black plates (nunc) in order to screen for sialyltransferase activity. Transglycosylation reactions were set up both in pH 5.25 100 mM acetate buffer and pH 7.01 100 mM HEPES buffer. Supernatants (50 µL) were incubated with 1 mM 2FPh-NeuAc sialyl donor, 0.5 mM FMU-Man acceptor, and 20% (v/v) acetonitrile in a total reaction volume

of 100 µL. After the reactions were sealed with tape to minimise evaporation they were incubated at room temperature for 3 days. The rationale behind library screening protocol 1 is shown in Scheme 3.1. That is, the plates were mounted into a Cary Eclipse fluorescence plate reader and the background fluorescent was measured (excitation wavelength of 355 nm and emission wavelength of 485 nm). Then a second fluorescent reading was taken after α-mannosidase (from Jack Bean) has been added to hydrolyze unreacted FMU-Man. Finally, a third reading was acquired after wild-type sialidase had hydrolyzed all sialylated products. For wells in which a significant difference between readings 3 and 2 were observed, the bacteria were grown (using the master plate stocks) for DNA sequencing in order to identify consensus sequences.



Scheme 3.1 *The fluorescent reading of each step in library screening protocol 1. Mutant clones were selected against high fluorescent signals of the difference between reading 3 and reading 2 compared to the background culture reading.*

3.2.7. Library Screening Protocol 2.

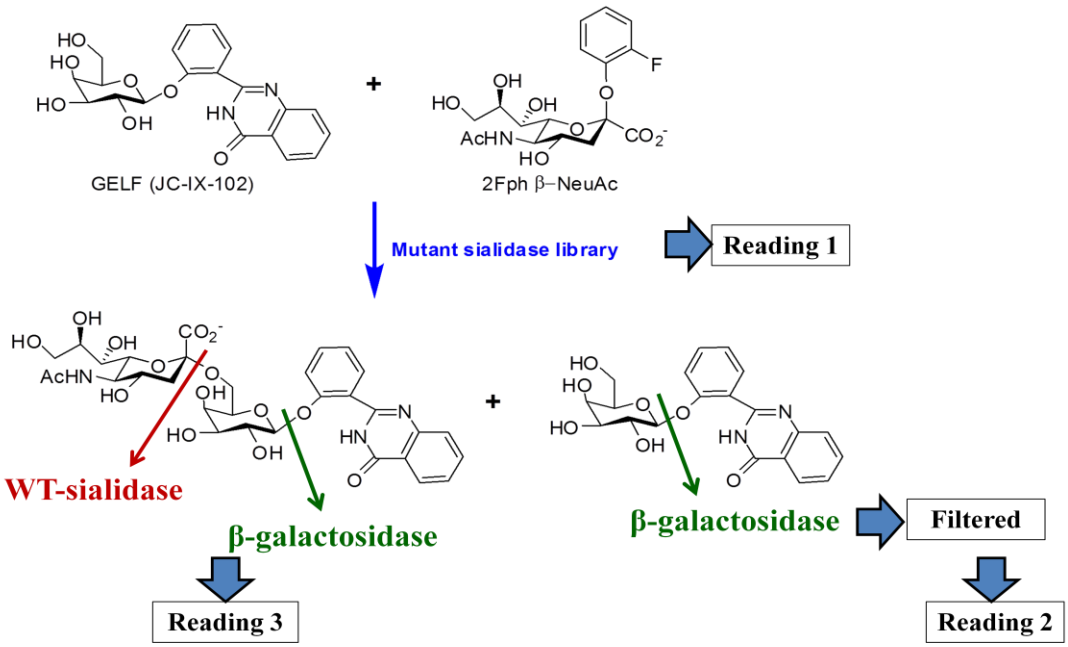
The expression and screening protocol was slighted modified for the second screen by expressing protein for 41 h at room temperature instead of 16 h, combining 6 wells from each plate into a single well in order to reduce substrate usage and performing the screen at pH 7.10 100 mM Tris-HCl buffer.

3.2.8. Trial Library Screening Protocol 3.

Mock reactions with 0%, 5%, 20%, and 100% coupling product were set up to assess the validity and sensitivity of the screening methodology. Different proportion of Sia-gal-FMU and gal-FMU were set up in 100 µL reactions. β-Galatosidase was first added to cleave the gal-FMU and the pH of the reactions were adjusted from 7.10 to 4.07 by 100 mM acetic acid for 20 min to precipitate free FMU. The precipitated FMU was then filtered through 0.45 µm filter by centrifuge at 6576 ×g 20 min. The supernatant collected on a black nunc plate was adjusted back to pH 6.99 by 150 mM phosphate buffer followed by addition of WT-sialidase and β-galactosidase. Each step was monitored by using a fluorescence plate reader to assess the validity of the method.

3.2.9. Library Screening Protocol 4.

A different fluorescent acceptor, Gal-ELF was used in this screening protocol (Scheme 3.2) in order to try and remove hydrolyzed “ELF” from the unreacted Gal-ELF acceptor. The expression and transglycosylation conditions were the same as screening 2 protocol, except 1 mM of gal-ELF acceptor and 13% of DMSO organic solvent were used. After the addition and incubation of β-galactosidase to the coupling wells, the reactions from the plates were centrifuged and filtered through 0.25 mm size filter. The filtrates were collected in a black screening plate and WT-sialidase and β-galactosidase were added to hydrolyze any sia-gal-ELF compound that was formed in the coupling reactions.



Scheme 3.2 *The fluorescent reading of each step in library screening protocol 4.*

3.2.10. Expression and Purification.

Select positive hit mutants from library screenings were expressed and purified (1 L-scale). Specifically, a fresh colony was used to inoculate 2 mL LB/Kan media for 1.75 hrs 37 °C and 180 rpm. Cultures were then transferred to 70 mL TB/Kan media to grow for 3 hrs 37 °C and 180 rpm and this 70 mL culture was then used to seed 1 L TB/Kan media for 2 hrs 37 °C and 180 rpm. The resulting culture was first placed on ice for 15 min (to cool to room temperature) and then IPTG (1 mM final concentration) was added and the culture was kept at room temperature for 41 h (180 rpm). The cells were harvested by placing the culture media on ice for 15 min followed by centrifugation at 9963 ×g for 12 min (4 °C). The supernatant was first filtered using 2 layers of Whatman#1 filter and then mixed with 5 mL Ni-NTA resins (Qiagen) at 10 °C and 100 rpm. The binding resin was collected by centrifuging at 5978 ×g 10 min of 4 °C. A slurry of the binding resin was transferred to a column and washed with binding buffer (5 mM imidazole, 1 × phosphate buffer, 100 mM NaCl pH 7.4) followed by washing buffer (10 mM imidazole, 1 × phosphate buffer, 100 mM NaCl pH 7.4) and then eluting buffer (100

mM imidazole, 1 × phosphate buffer, 100 mM NaCl pH 7.4). Fractions were collected from the column and analyzed using 10% SDS-PAGE.

3.2.11. Inclusion Body Purification.

To purify inclusion bodies, a published protocol was followed with some modifications (81). Specifically, the cell pellet from 1 L of media was suspended in binding buffer (Ni-binding buffer with 6 M urea, 1 × PMSF, and 1 mM DTT). After lysozyme 1% (v/v) was added the mixture was incubated for 30 min at room temperature on a rocker. Six cycles of sonication with 20 s on and 40 s off were applied to break up the cells. The cell debris and supernatant were separated by centrifuging at 39845 ×g for 1 h (4 °C). The cell pellet was collected and resuspended in 40 mL of 0.1 M Tris pH 8.05 buffer with 10 µL DNAase I added. The resulting sample was then centrifuged at 39845 ×g 10 min (4 °C) and the resuspension and centrifugation cycle was repeated twice more. The isolate inclusion bodies were solubilised by addition of 10 mL 8 M urea and 10 mL Triton buffer (20 mM Tris-HCl pH 8.05, 0.2 M NaCl, 10% glycerol, 0.1% Triton-X100). This mixture was loaded slowly onto a 5 mL Ni-NTA column and the column was washed and eluted with 10 mM and 300 mM imidazole in Triton buffer. Fractions were collected and analyzed using 10% SDS-PAGE gel. Fractions containing target protein size were pooled and dialyzed three times with each time against 4 L buffer (20 mM Tris pH 8, 200 mM NaCl, 10 % glycerol and 0.1 % Triton X-100) overnight at 4 °C.

3.2.12. HPLC Analysis of Transglycosylation Reactions.

The positive hits from library screening protocol 2 were prepared and analyzed on HPLC with a gradient pump system from Agilent Technologies using a Phenomenex C18 column 150 x 60 mm (3 µm particle size). The two mobile phase solutions were water containing 0.1% formic acid (solvent A) and acetonitrile (solvent B). Isocratic mobile phases (12% v/v solvent B and 88% v/v solvent A) were delivered at 1.0 mL/min for two minutes, which was followed by running a gradient mobile phase (12 to 30% v/v solvent B from two to twelve minutes). Samples were prepared by removing 5 µL of supernatant transglycosylation reaction from the screening plate and 20 µL of cold ethanol was added and stored at –20 °C for 20 min to precipitate the protein. Samples

were then centrifuged at 26563 ×g for 10 min at 4 °C and the supernatant was evaporated to dryness (speed vac). The dried samples were re-suspended with 50 µL of dH_2O and analyzed on HPLC-UV_{290nm}. The retention time for sia-gal-FMU was 9.85 min and man-FMU was 11.29 min.

3.2.13. NMR Analysis of Transglycosylation Reactions.

Transglycosylation reactions were stopped by the addition of ethanol to give a final concentration of 70% (v/v) followed by storage at –20 °C for 20 min in order to precipitate the Y370G mutant enzyme. The supernatant and precipitate were separated by centrifugation (34532 ×g for 10 min at 4 °C) and the supernatant was evaporated to dryness and re-suspended in D_2O for ^1H NMR spectral analysis.

3.2.14. Time Point Fluorescent Analysis of Transglycosylation Reactions.

The positive hits from screening protocol number 4 were de-convoluted into individual clones for re-screening for each clone transglycosylation activity. The expressed supernatant of each clone were set up in parallel using screening protocol number 4 conditions. At time points 1, 8, 16, and 48 hours, the coupling reactions were stopped by addition of β-galactosidase, centrifuge filtered following WT-sialidase treatment, and the fluorescence intensities were measured using a fluorometer. The fluorescent time point assays were repeated three times, with the last trial being performed in the presence of 0.01% BSA.

3.2.15. Switching pJWOSH to pJWSH and pJWFimi Constructs.

The expression yields for both I-1-A11 and I-25-A5 pJWOSH were low and more importantly the gene product did not show enzymatic activity. Therefore different constructs were tried to see if the expression and activity problems could be solved. The pJWOSH construct consisted of an OmpT secretion signal and when this section was omitted the resulting construct was named pJWSH. The pJWSH construct was produced by PCR to amplify only the sialidase gene without the N-terminal OmpT sequence. Primers of Cut-OmpT' and ggHin' were used under the same conditions as described above. The PCR product was sub-cloned into pET28a vector with *EcoRI* and

HindIII sticky ends and used to transform *E. coli* BL21 DE3 cells. Proteins from positive transformants were expressed, purified, and their activity was checked. The pJWFimi constructs were produced in a similar manner to pJWHS. Instead of sub-cloning the mutant sialidase gene into a pET28a vector, the gene was sub-cloned into a pJWFimi vector (pET28a consist Fimi sequence).

3.2.16. Transglycosylations with Various Acceptors.

The pJWFimi construct of I-1-A11 and I-25-A5 mutant genes was expressed and supernatant proteins were concentrated 5-fold before the transglycosylation reactions were set up. Three different acceptors, GELF, gal-FMU, and galactose (10 mM) were set up in parallel for both mutant enzymes. The reactions consisted of concentrated mutant enzymes (57% v/v), 2FPh-NeuAc (10 mM), BSA (0.01% v/v), DMSO (10% v/v for GELF and gal-FMU only) in Tris-HCl buffer (100 mM, pH 7.12). Reactions were incubated at 37 °C overnight and were stopped and prepared as described above before analysis by NMR spectroscopy. The reaction with a galactose acceptor had shown a positive result; therefore, various PNP-sugars were used as a potential acceptors. The PNP-sugar acceptors included PNP-α/β-galactopyranosides, PNP-α/β-glucopyranosides, PNP-α-mannopyranoside, and PNP-β-*N*-acetylgalactosaminopyranoside were all tested in parallel with both mutant enzymes. 2FPh-NeuAc sialyl donor (16 mM), a PNP-sugar acceptor (8 mM) and a mutant enzyme were incubated. That is, solutions containing pH 7.02 phosphate buffer (100 mM), acetonitrile (10% v/v), for 2 days at room temperature prior to NMR spectroscopic analysis as described above.

3.2.17. Incorporation of Lactose-binding and Linker Domains into the Active Site Sialidase Gene.

Plasmids consisting of full size sialidase ($pJWOSH_{full}$) and I-1-A11 active site mutant (from the pJWOSH plasmid referred to as pJWβ I-1-A11) were prepared from overnight cultures followed by miniprep (Qiagen). Restriction digestion using *MscI* and *HindIII* were performed on both plasmids using the manufacturer's recommended conditions (Fisher Scientific). The DNA fragments digested from $pJWOSH_{full}$ containing the linker and lactose-binding domain as 883 bp. This fragment was ligated into pJWβ I-

1-A11 and used to transform *E. coli* BL21 DE3 competent cells using the same conditions as detailed above. The mutant incorporated with these other domains is referred to as pJWOSH$_{full}$-I-1-A11. The positive transformants were confirmed by target gene digestion with *EcoRI* and *HindIII* and gene fragments analyzed on 1 % agarose gel and DNA sequencing before expression, purification and activity check. Eleven other mutant clones from screening protocol 4 were also incorporated into the full size construct in order to obtain a greater yield of protein and hopefully a more stable protein in order to characterize transglycosylation reactions.

3.3. Results

3.3.1. DNA sequence diversity.

Based on twenty randomly selected transformants, 18 of the 20 sequences were different; hence, 90% mutational randomization was observed. Three sequences contained a stop codon, which would lead to a truncated protein. Two thirds of the sequences contained a either Q or N amino acid next to the four targeted mutation sites and the remainder one third were either K or H. The exact sequences of the twenty mutants are listed below. The mutant sequences were aligned against one of the template sequence TS6-Yg-Gy. The four amino acids of the W-loop random mutation are highlighted in yellow and the sequences that were different from the template sequence are shaded in gray color (Figure 3.4).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                         10         20         30         40         50
TS6-Yg-Gy       PDAPAGSARA KVLLFSNAAS QFKGRWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-1             PDAPAGSARA KVLLFSNAAS Q.VIIWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-2             PDAPAGSARA KVLLFSNAAS NLFKSWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-3             LP.RPCRLRA GQGPALLQRR PKTILWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-4             PDAPAGSARA KVLLFSNAAS NRCN.WLRSQ GTIRMSCDDG QTWPVSKVFQ
```

```
I-5          PDAPAGSARA KVLLFSNAAS QPEN.WLRSQ GTIRMSCDDG QTWPVSKVFQ
I-6          PDAPAGSARA KVLLFSNAAS NPRDNWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-7          PDAPAGSARA KVLLFSNAAS QFKGRWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-8          PDAPAGSARA KVLLFSNAAS QIGPDWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-9          PDAPAGSARA KVLLFSNAAS KEWKLWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-10         PDAPAGSARA KVLLFSNAAS QVINNWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-11         PDAPAGSARA KVLLFSNAAS HYRHLWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-12         PDAPAGSARA KVLLFANAAS QDENQWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-13         LP.RPGRLRA GQGPALLQRR PLFVSWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-14         PDAPAGSARA KVLLFSNAAS KVTHKWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-15         PDAPAGSARA KVLLFSNAAS NEHLFWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-16         PDAPAGSARA KVLLFSNAAS HWDVTWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-17         PDAPAGSARA KVLLFSNAAS QFSTMWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-18         LP.RPCRLRA GQGPALLQRR QHRLPWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-19         PDAPAGSARA KVLLFSNAAS NFKGRWLRSQ GTIRMSCDDG QTWPVSKVFQ
I-20         PDAPAGSARA KVLLFSNAAS HMAVQWLRSQ GTIRMSCDDG QTWPVSKVFQ

                  ....|....| ....|....| ....|....| ....|....| ....|....|
                           60         70         80         90        100
TS6-Yg-Gy    PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-1          PGPMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-2          PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-3          PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-4          PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-5          PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-6          PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
```

```
I-7          PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-8          PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-9          PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-10         PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-11         PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-12         PGPMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-13         PGPMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-14         PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-15         PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-16         PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-17         PGSMSGSTLS ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-18         PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-19         PGPMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
I-20         PGSMSGSTLT ALPDGTYGLL YEPGTGIRYA NFNLAWLKLA AALEHHHHHH
```

Figure 3.4 *Amino acid sequence alignments of 20 random clones. Dots referred to the stop codon.*

3.3.2. Library Screening – 1.

A total of seven 96-well plates were screened using screening protocol 1. The background control reactions were performed in parallel and the difference between fluorescene reading 3 and reading 2 were 0.27 and -18.87 at pH 5.25 and pH 7.01, respectively. As a result, a fluorescence difference reading of greater than 50 was considered as a positive hit. The 40 positive hits were sent for sequencing. Unfortunately, 13 out of the 40 sequences turned out to contain a stop codon in either the targeted 4 amino acid mutagenesis positions or in other parts of the complete sequence (Table 3.3). Therefore, this screening protocol was not effective in determining mutant clones that had sialic acid transferring activity.

Table 3.2 *Translation of relevant codons from library screening. Selective of potential hits were sequenced from screening 672 individuals of the WY loop library for coupling activity onto FMU-Man. The * signs denote sequences that contain a stop codon in the middle of the sequence. The ** notation refers to poor quality sequence data.*

Clone from the original plates	Clone transfer to the master plate	N/Q 339	F340	K341	G342	R343
I-10 H1	I-a-A1	N	A	R	V	V
I-10 D3	I-a-A2*	Q	Stop	S	S	H
I-10 E5	I-a-A3	H	T	F	T	Q
I-10 H5	I-a-A4	N	P	C	K	K
I-10 G7	I-a-A5	N	F	K	G	R
I-17 C1	I-a-B1	H	V	G	I	D
I-17 H2	I-a-B2	H	E	L	V	Q
I-17 H3	I-a-B3	Q	N	F	M	H
I-17 H4	I-a-B4	Q	S	L	V	S
I-17 H12	I-a-B5	H	G	S	V	H
I-18 B1	I-a-C1	Q	Q	E	V	Y

I-18 H1	I-a-C2*	P	S	S	S	V
I-18 A4	I-a-C3	Q	C	C	S	G
I-18 H12	I-a-C4	N	H	R	S	K
I-19 A9	I-a-D1	Q	M	G	S	T
I-19 D10	I-a-D2	Q	V	S	M	L
I-29 E1	I-a-E1	Q	M	W	S	Q
I-29 F4	I-a-E2*	P	I	V	Y	M
I-29 E5	I-a-E3	Q	C	K	H	T
I-29 B11	I-a-E4*	K	Stop	T	G	Q
I-30 E1	I-a-F1*	L	F	G	T	R
I-30 F1	I-a-F2	N	Y	S	T	K
I-30 G1	I-a-F3	H	L	K	N	S
I-30 H1	I-a-F4	N	F	K	G	R
I-30 C2	I-a-F5*	H	R	Y	Y	F

I-30 D2	I-a-F6	Q	Y	N	I	F
I-30 E2	I-a-F7*	K	R	R	G	L
I-30 F2	I-a-F8	H	G	V	T	D
I-30 C4	I-a-F9**	--	--	--	--	--
I-30 F5	I-a-F10**	--	--	--	--	--
I-30 H6	I-a-F11**	--	--	--	--	--
I-30 C8	I-a-F12*	H	L	W	L	R
I-30 H9	I-a-G1	N	E	C	T	L
I-30 H11	I-a-G2*	H	E	L	H	K
I-30 E12	I-a-G3	N	E	S	F	C
I-31 H1	I-a-H1*	N	F	K	G	R
I-31 G2	I-a-H2*	H	C	V	L	Stop
I-31 H2	I-a-H3*	H	Stop	M	N	A
I-31 H7	I-a-H4	Q	W	L	W	M

I-31 F10	I-a-H5*	Q	P	Stop	A	S

3.3.3. Expression and Purification – 1.

The three mutant clones, I-a-F6, I-a-H4, and I-a-E3 that exhibited the highest measured fluorescence differences using protocol number 1 were expressed on a 1 L scale. Three other control sequences were expressed in parallel with these selected mutants. The negative control containing just media, and two positive controls, Yg-active sialidase, and the original WY loop mutant clones TS6-Yg-Gy, were also expressed. The expression level of three mutant clones, I-a-F6, I-a-H4, and I-a-E3 was barely detectable on SDS-PAGE.

3.3.4. Library Screening – 2.

Subsequently, 15 more library plates were screened using screening protocol number 2. From this screening batch, three wells showed more than twice the fluorescent signal intensity compared to that of the background reaction. Each well contained a combination of six mutant clones; therefore, these three wells were de-convoluted into eighteen individual mutant clones and the screening procedure was repeated. Out of eighteen single clones, only one stood out from the background signal; moreover, the screening supernatant from this clone was reanalyzed on the HPLC but no transglycosylation product was detected.

3.3.5. Trial Library Screening – 3.

In order to try and solve the problems that are associated with high background readings, after the addition of β-galactosidase the pH was lowered to try to precipitate the hydrolysis product FMU (pK_a = 6.4); however, this procedure was unsuccessful. That is, different percentages of mock coupling reactions were set up to test if a pH precipitation approach was feasible. When the pH of the reaction mixture was adjusted to 4.07 the fluorescent intensity of the mock reactions dropped around 8-fold before filtration (the actual fluorescent reading range from 391.5–522.1 at pH 7.12 depending

on the percent of mock coupling dropped to 31.4–66.0 at pH 4.07) and this value dropped another 15-fold after filtration (the fluorescent reading range were 20.4–40.4). Nevertheless, after filtration when the pH was adjusted back to 6.99 the fluorescent intensity increased 10-fold (fluorescent reading 331.6–434.6) without the addition of WT-sialidase and β-galactosidase. Even after the addition of these two enzymes the fluorescent intensity increased around 1-2 fold but some wells that had less mock coupling ended up with higher fluorescent readings than those that represented more mock coupling reactions. Therefore, the intrinsic fluctuations of such high fluorescent signal intensities meant that incorporation of a low pH precipitation step did not solve the problem. That is, lowering the pH was unsuccessful in completely precipitating FMU.

3.3.6. Library Screening – 4.

All 31 library plates were screened using this screening protocol. Twenty five wells stood out using this screening with signal intensities that ranged from 5 to 11 times higher than the background signal. These 25 wells were each de-convoluted into six individual clones and re-screened to elucidate which clone or clones gave rise to the positive hits. The de-convolution experiments were repeated twice using the same screening conditions but no positive hits were detected. In order to try and increase coupling the DMSO concentration was reduced from 13% to 8% and the 2FPh-NeuAc concentration was increased from 1 mM to 2 mM. This attempt resulted in 1 clone exhibiting a signal 20 x higher than the background even though at least 25 positive hits were expected based on the initial screen. Furthermore, these experiments were repeated with 0.1% BSA added to the coupling reactions. Under these conditions, 14 clones had fluorescence signals that were 10-20 times higher than the background. These 15 positive hits from different experiments were sent for DNA sequencing to check for any consensus pattern (Table 3.4). The DNA sequence suggested I-18-D8 clone had a stop codon in W-loop region and clones I-2-A5, I-20-B11 and I-25-A5 had 1 nucleotide deletion next to the C-terminal end of the W-loop region. The clone I-3-H4 had no growth when cultured.

Table 3.3 *Translation of relevant codons from sequencing potential hits from screening protocol 4 de-convolution experiments.*
*The notation of * means no growth; ** means the sequence data is inconclusive; *** means 1 base pair deleted after the W-loop mutations (frameshift results in no his tag); **** means a stop codon is present in the middle of the sequence.*

Clone from the screening plates	Clone on the master plate	Y-Loop	N/Q 339	F340	K341	G342	R343
Plate I - F4	I-3-H4*	--	--	--	--	--	--
Plate II - B2	I-18-F8	YW	H	V	R	K	E
Plate II - D12	I-22-D12	-- **	Q	V	E	L	Y
Plate I - 2E	I-1-F8	YF	K	Y	W	T	V
Plate I - 5A	I-1-A11	YF	H	P	T	H	L
Plate I - 5F	I-2-A5 ***	YW	Q	V	L	R	I
Plate I - 8E	I-17-A2	YF	K	K	K	E	Y
Plate II - 2B	I-18-D8 ****	YF	H	M	Y	Stop	V
Plate II - 3G	I-20-B9	YF	Q	C	G	Q	K
Plate II - 4G	I-20-B11 ***	YW	H	V	W	R	S
Plate II - 6D	I-18-G12	YW	K	I	V	R	P

Plate II - 6E	I-18-H12	YF	N	H	R	S	K
Plate II - 6F	I-19-G12	YF	Q	Q	C	T	N
Plate II - 7G	I-25-A7	YW	Q	L	R	G	S
Plate II - 11F	I-25-A5 ***	YW	N	V	L	R	I

These twelve positive clones were expressed on a larger scale and their activities were measured. Coupling conditions were set up as detailed in the experimental and monitoring of transferase activity was performed using fluorometer. However, several attempts were tried before any positive signal being observed. At time point 48 h with addition of BSA, four clones, I-1-A11, I-25-A5, I-1-F8, and I-25-A7 stood out with 10 ×, 6 ×, 2 ×, and 1.5 × higher than the background intensity, respectively.

3.3.7. Expression, Purification and Transglycosylation Activities of I-1-A11 and I-25-A5 Mutants.

The two clones in pJWOSH construct, I-1-A11 and I-25-A5 that stood out from the twelve “positive” hits were expressed, purified, and assessed for transferase activity by NMR spectroscopy. The expression level for I-1-A11 was about 0.2 mg/L after Ni-NTA column purification which is about 100× lower than that of the wild type. For I-25-A5 frameshift mutation was fixed by inserting an extra base C nucleotide after gly 343 using the PCR overlap extension method to frame shift the protein back to incorporate a His-tag. The corrected reading frame of the I-25-A5 mutant clone was then expressed and purified by Ni-NTA column. The expression level for I-25-A5 was also low. Both mutant enzymes were concentrated before the transglycosylation reactions were attempted. However, based on the acquired NMR spectra, only starting materials were observed after the coupling reactions had been allowed to proceed for three days at

room temperature. The experiment was repeated but still no coupled product was obtained.

Furthermore, the gene products of I-1-A11 and I-25-A5 pJWOSH construct were attempted to be expressed and purified. The expression sample of the supernatants for both mutants and the associated cell pellets were analyzed on SDS-PAGE gel. The gel data suggests that most of the target protein ended up in the cell pellet. Thus, purification of inclusion bodies was tried and protein with the correct size was isolated. Yet, NMR analysis of coupling reactions using these purified mutant proteins still showed no apparent trans-glycosylation.

In order to try and solve problems associated with low enzyme activities several different constructs and coupling conditions were tried for mutant I-1-A11 and I-25-A5. That is, constructs without the OmpT signal (pJWSH) and that with a different secretion signal, Fimi (pJWFSH) were produced and the proteins expression levels were similar to those of the pJWOSH construct. However, the NMR spectra of coupling reactions showed that no product was formed. Also, galactose was used as the acceptor and the NMR spectra showed a trace of sialyl-galactose was formed. Therefore, PNP-β-galactose was tried as the acceptor for further kinetic analysis of the glycosylation reaction using UV spectrometry. Seven other PNP-α/β-sugars (galactose, glucose, mannose, and N-acetylgalactosamine) were also tried as acceptors, with the hope that appreciable amounts of coupling would occur and thus allow for further characterization studies on the mutant enzymes would be feasible. However, only very low signals at a chemical shift of around 2.7 ppm appeared (expected ^{1}H signal of H3 for an α-sialoside).

3.3.8. Expression, Purification, and Transglycosylation of Mutant Candidates in Full Length Protein.

The various constructs of the active site mutant sialidases showed low expression yields and despite being “positive” hits had barely detectable transglycosylation or hydrolysis activity; therefore, the lectin and linker domain of the full length enzyme were incorporated in an attempt to produce more stable proteins with higher expression levels. That is, the full construct of MvNA has been shown to give

high expression yields (120), pure mutant proteins being obtained in ranges from 15–89 mg/L (Table 3.5). The coupling reactions were set up using lactose as acceptor and the reactions were monitored by ^{1}H NMR spectroscopy. Based on the NMR spectra, most of the reactions showed different rates of 2FPh-β-Neu5Ac hydrolysis but no coupling (Figure 3.5). The reaction catalyzed by the I-17-A2 mutant show coupling product but hydrolysis was still the dominant reaction. The reaction with I-25-A7 mutant showed a very small amount of coupling product and no hydrolysis. Although, this result could not be replicated after two other attempts.

Table 3.4 *Summary analysis of full length mutant clones from the WY loop library screening.*
Notation "--" refers to neither protein, hydrolysis, nor coupling detected.

pJWOSH$_{full}$	**Protein Quantification (mg/L)**	**W-loop**	**Y-loop**	**% of hydrolysis**	**% of coupling**
I-1-A11	24.4	HPTHL	YF	29.1	--
I-25-A5	63.8	NVLRI	YW	13.8	--
I-1-F8	27.5	KYWTV	YF	12.3	--
I-18-G12	48.0	KIVRP	YW	2.0	--
I-18-H12	15.0	NHRSK	YF	32.0	--
I-19-G12	68.5	QQCTN	YF	5.7	--
I-25-A7	61.3	QLRGS	YW	--	<0.0

I-17-A2	65.4	KKKEY	YF	84.3	4.7
I-2-A5	40.0	QVLRI	YW	6.54	--
I-20-B11	26.2	HVWRS	YW	2.0	--
I-20-B9	88.5	QCGQK	YF	76.7	--

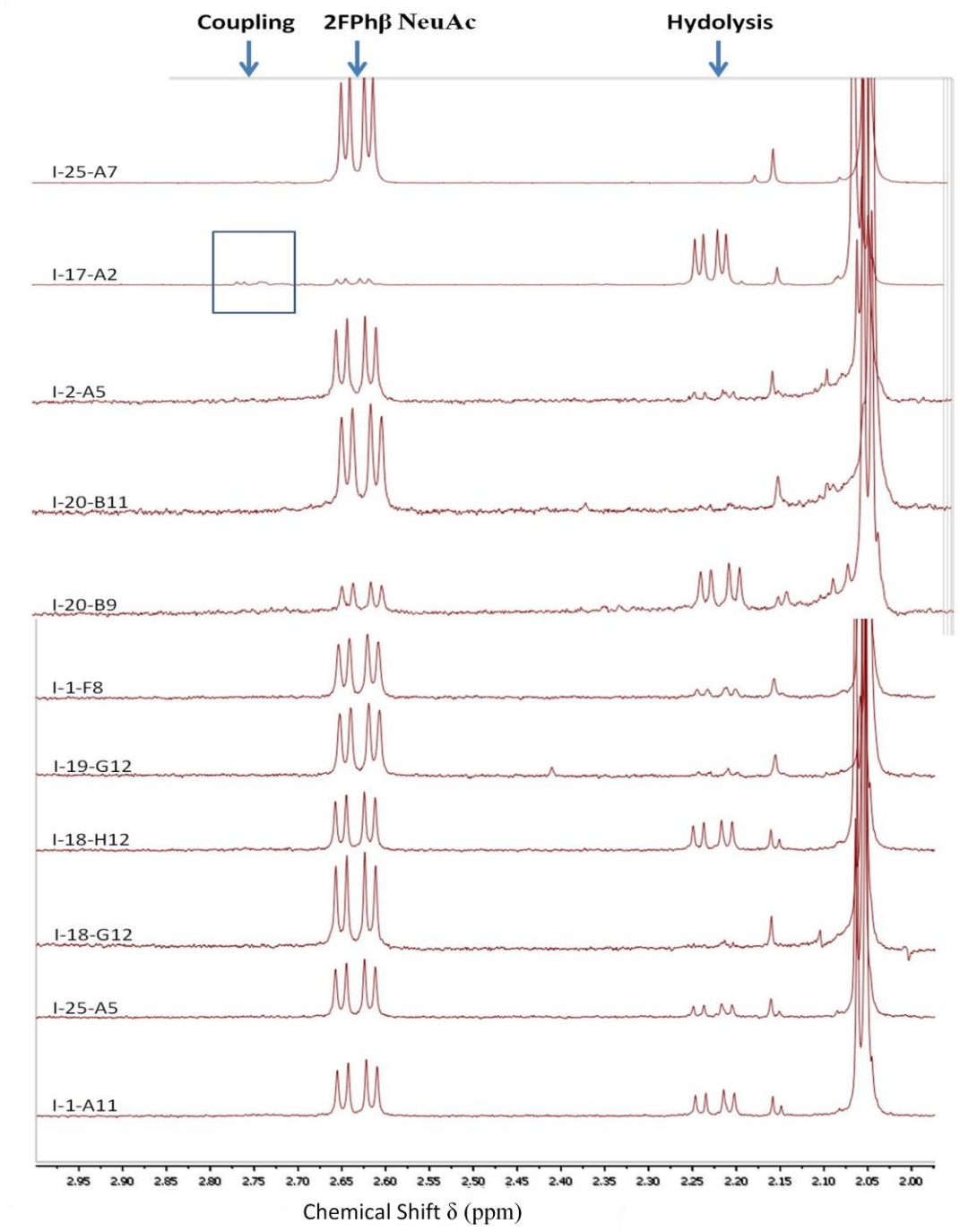


Figure 3.5 *^{1}H NMR spectra of transglycosylation reactions of full length mutant clones. The square box on I-17-A2 mutant spectrum shows the signal of transglycosylation product.*

3.4. Discussion

After examining the sequence data for the positive hits identified using screening protocol 1, for which around 30% of these hits coded for truncated mutant enzymes where transglycoslyation activity is unlikely, it was decided that the screening method needed to be adjusted. That is, reading 2 is taken after the addition of α-mannosidase, which cleaves unreacted (8-fluoro-4-methylumbelliferyl α-D-mannoside; FMU-Man) and, generally gives rise to a high background fluorescence intensity reading. As a result, positive hits are based on the difference between two high fluorescence readings (reading #3 - reading #2) and thus this protocol is intrinsically error prone. A similar screening protocol in which six clones were combined into each well (protocol 2) so as to conserved synthetic acceptor resulted in high background signals, which continues to be an issue for the detection of true trans-glycosylation activities. That is, the de-convolution analysis of 18 individual clones from three hits (six clones per well) would have been expected to give, at least three hits, yet only one hit was observed. An observation that suggests, once again, that hits from the six clones in a single well combination methodology result from artifacts. Therefore, a third screening protocol was appraised in order to try and filter out the high background signal problem.

That is, in this alternative approach it was decided to try any remove free fluorophore following addition of glycosidase enzyme. In this example, as the pK_a of FMU is 6.4 (97) and that protonated coumarins are much less soluble in aqueous media than are the corresponding anionic phenolate form it was reasoned that lowering the pH of the screening solution (to ~4.1) should result in precipitation of free FMU, which could be removed by filtration and thus lower the background fluorescent reading. However, even after filtration of the acidic media the fluorescent background (reading #2) was still high when the pH was adjusted back up to values near 7.0. An outcome that suggested FMU either did not precipitate efficiently or that the resultant particle size was smaller than the filter pores (0.45 µm). In summary, screening protocol #3 did not solve the intrinsic problems associated with high background fluorescent intensity readings.

In an extension to the idea that an insoluble fluorophore, which should precipitate when released following addition of a glycosidase to the screening reaction mixture and lead to an efficient method for lowering the background fluorescent reading, should give

lower background reading a new fluorescent leaving group "ELF" was chosen. Thus, 2-(4-oxo-3,4-dihydroquinazolin-2-yl)phenyl β-D-galactopyranoside (abbreviated as GELF) was used in screening protocol number 4.

Using ELF as the fluorescent leaving group allowed removal of unreacted substrate before detection of the trans-glycosylation product. However, after performing an initial screen (six clones in each well), the necessary deconvolution experiments once again did not give the expected number of hits for transferase activity. A contributing cause for the production of inactive enzyme is likely the obligatory presence of an organic co-solvent (DMSO), which is necessary for the solubilisation of GELF, hindering enzyme function and stability. To try and compensate BSA was added to stabilize the mutant enzymes, however, only 14 clones were detected to be “hits”, and one of these sequences contained a stop codon in the W loop region. These clones contained mainly basic, hydrophobic, and aromatic amino acids in the variable region (Table 3.4). Three clones had a one base pair deletion after the targeted W loop mutation, which gave rise to a Gly instead of a Trp amino acid following the four targeted residues in the W loop region. Clearly, the presence of a Gly amino acid might be significant for trans-glycosylation activity. Thus, it was decided to maintain the Gly residue and frame shift the remaining short peptide sequence back to that of the wild-type by insertion of a base pair, which also re-inserts the C-terminal his-tag.

3.4.1. Expression, Purification, and Trans-glycosylation of I-1-A11 and I-25-A5 Active Site Mutants.

All three constructs (pJWOSH, pJWSH, pJWFSH) gave similar expression yields of protein, all of which displayed barely detectable trans-glycosylation activities. A consistent problem when performing screening assays using fluorescent leaving groups is the need to add either acetonitrile or DMSO to the coupling reactions, which can affect enzymatic reactions by destabilizing protein folds. This explanation is consistent with the observation that trans-glycosylation products are more commonly observed when a natural sugar is used as acceptor (no organic solvent is needed to solubilize the acceptor). In combination, this data suggests that the mutant enzymes are prone to denaturation. Therefore, other natural domains of the *M. viridifaciens* sialidase enzyme

(galactose binding domain and hinge domain) were incorporated with the hope that the mutant enzymes would be stabilized.

3.4.2. Expression, Purification, and Trans-glycosylation of Mutant Candidates in the Full Length of Sialidase.

As expected, the full length sialidase constructs gave expression yields that were more than 100-times greater than those of the active site constructs. Even though all selected library mutants were expressed in the full length sialidase construct the expression yield of these mutant protiens varied by nearly 6-fold. The lowest expression level was ~15 mg/L I-18-H12 that contained NHRSKWLR in the W loop and YF in the Y loop. The W loop region in this particular mutant sialidase contains three basic amino acids that are most likely positively charged under physiological conditions. Whereas, the highest expression level was ~90 mg/L I-20-B9 that contained QCGQKWLR residues in W loop region and YF in the Y loop region, of which only one of the variable residues is likely charged.

The pJWOSH$_{full}$ construct of eleven mutant clones showed differences in their hydrolytic and transferase activities. The percent of hydrolysis per mg of mutant proteins were calculated, clones I-18-H12, I-17-A2, and I-1-A11 were the three most active catalysts for hydrolysis and the variable W-loop regions for these three clones contained, NHRSK, KKKEY, and HPTHL amino acid sequences, respectively. Out of the eleven clones only I-17-A2 and I-25-A7 showed traces of transferring activity and their variable W-loop amino acid sequences are KKKEY and QLRGS, respectively. Both of these clones contain basic amino acids of K (Lys) and R (Arg) in the W loop region, although given the low *trans*-glycosylation activities no definite conclusions can be made as to what residues are needed in either the W or Y loops to promote an anabolic (coupling) reaction over the catabolic (hydrolysis) reaction.

3.5. Conclusions

The sequence of positive hits mutant clones from WY loop library screen consisted of similar numbers of YF and YW mutants in the Y loop region, which suggests that the Y-loop region only has a minor perturbation to the enzymatic activity.

Mutant full-length sialidases (active site domain, lactose binding domain, and hinge domain) have a greatly increased expression level compared to the active site mutant enzymes. Even though this library screen did not yield mutants that favor transferase activity, the W and Y loop region mutants possess greatly reduced hydrolysis activities.

4. Screening of Y370M & Y370H Random Mutant Library

4.1. Introduction

Enzymes are natural, often highly specific, catalysts that accelerate the rate of various chemical reactions. To understand how such large biological molecules catalyze specific reactions is a continuing area of interest. The goal of this project is to study the catalytic mechanism of *Micromonospora viridifaciens* Y370 mutant sialidases.

The substitution of the active site nucleophile Y370 to either a methionine or a histidine (Y370M and Y370H) residue results in the incorporation of intrinsically nucleophilic residues (79). Studies on the rates of reaction of various nucleophiles were measured with methyl iodide in methanol to mimic the hydrophobic condition in the enzyme active site (79). The nucleophilic strength of three nucleophiles, dimethyl sulfide (methionine mimic), imidazole (histidine mimic) and phenol (tyrosine mimic) were compared. Methionine and histidine nucleophiles mimic showed greater strength in acting as nucleophile rather than tyrosine mimic. However, experiments on the Y370M and Y370H mutant sialidases revealed the enzymatic activity had been greatly reduced (119). This contradiction between expectation and reality raises the question of whether it is possible to evolve an enzyme to use methionine or histidine as a nucleophile during catalysis. Thus, random mutations were made on the Y370H and Y370M sialidase mutant genes by error-prone PCR to generate a library of mutant proteins and to screen for improved enzyme activity, and then to study and analyze any de novo catalytic mechanisms (98).

A former post doctoral fellow, Dr. Jacqueline Watson made the plasmid library based on the active site domain for mutants Y370H and Y370M MvNA (pJWβ construct) by error-prone PCR (Gene Morph II kit from stratagene). The active site domain of enzyme encoded 400 amino acids (1.2 kb). Residues R10 – F396 were susceptible to

random mutations. The plasmid library Ymh10 encoded enzymes contain between 0-5 amino acids substitutions per protein and therefore, this plasmid library was used in this project.

4.2. Material and Methods

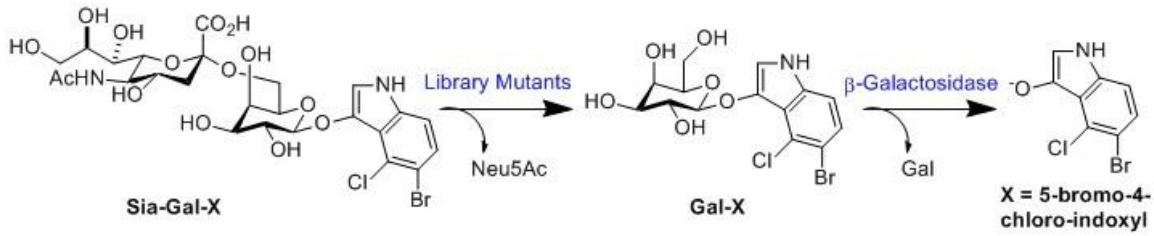
4.2.1. Trial Library Transformation, Expression and Screening Protocol - 1.

A small scale library was generated in order to validate the expression and screening conditions. One aliquot (~100 µL) of *Escherichia coli* BL21 (DE3) Gold ultracompetent cells (Agilent Technology) were transformed with 30-40 ng of the DNA ligation mixture. Transformations were carried out according to the instructions of the manufacturer. The mutant transformants generated were pick into a single 96 well plate in parallel with both negative control (pure media) and positive control (wild type active site domain sialidase). Terrific broth with kanamycin antibiotic (12 g of bactotryptone, 24 g of yeast extract, 4 mL of glycerol, and 100 mL of K phosphate in 1 L containing 30 µg/mL kanamycin) media was used to grow the *E. coli* bacteria. The expression conditions were the same as our previously published protocols (121). The expressed supernatant was used to screen for activity. For each of the well, 150 µL of expressed supernatant, 100 mM MOPS pH 7.01 buffer, 10 µM Sia-2,6 Gal-FMU substrate, and 0.1% β-galactosidase were prepared in final volume of 170 µL. The hydrolysis activity of mutants were monitored using a Cary eclipse fluorimeter (excitation wavelength 360 nm/emission wavelength 465 nm, excitation slit 5 nm/emission slit 5 nm, and 600v PMT voltage) with 1 min reading cycle for 10 min.

4.2.2. Library Transformation, Expression and Screening Protocol – 2.

A different library screening method was used to solve the fluorescent media issue (Scheme 4.1). Substrate Sia-2,6 Gal-X (X = 5-bromo-4-chloro-3-hydroxyindole) was synthesized by Dr. Fahimeh Shidmoossavee. The transformation protocol used was the same as that for trial #1 except the transformants were plated on M9 minimum kanamycin agar plate (for 500 mL of media consist, 3.4 g Na_2HPO_4, 1.5 g KH_2PO_4, 0.3 g

NaCl, 0.5 g NH_4Cl, 7.5 g agar, and 30 µg/mL kanamycin) with 10 µL of 1 M IPTG, 100 µL of 0.36g/mL glucose, and 200 µL of 20 mg/mL Sia-2,6 Gal-X plated on top of agar plates (~200 mL agar medium per plate). A total of ten 240 × 240 × 20 mm petri traies with 200 mL agar medium per plate were incubated at 37 °C for 7 days.



Scheme 4.1 *Screening analysis of Ymh mutant sialidase library.*

4.2.3. DNA Sequencing of Individuals from the Library.

Freshly re-streaked colonies were used to seed 5 mL of LB/Kan media (Luria Broth media, 10 g bacto-trptone, 5 g yeast extract, 10 g NaCl in 1 L of water) in 14 mL culture tubes. After an overnight incubation, plasmid DNA was purified (Qiaspin Ket, Qiagen). DNA sequencing was performed on 13 blue colonies using both the T7 and T7term universal primers to sequence front and back portion of the gene, respectively.

4.2.4. Mutant Proteins Expression and Purifications.

All the mutants clones were expressed and purified using the published protocol (121).

4.2.5. Kinetic Analysis.

The sialidase activity of YmhII-11 and YmhII-2 were compared. The supernatant of the clones were incubated with PNP-α-NeuAc and PNPS-α-NeuAc (0.1 mM) at pH 7.01 (100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; HEPES buffer) and pH 5.25 (100 mM acetate buffer) in the presence of 0.01% BSA. The hydrolysis activity

of both clones were monitored using a UV-Vis Cary 3E spectrometer (405 nm) at room temperature.

A pH-rate profile experiment was performed with YmhII-11 using PNP-α-NeuAc as the substrate. A series of buffers with different pH values ranging from pH 4.06 to 9.29 were prepared as 1 M stock solutions containing 1 M sodium chloride. Acetate buffers were used for pH values of 4.60, 5.05, and 5.60 and the HEPES buffers were used for pHs of 7.03 and 7.80 and a glycine buffer were used for pH 8.75 experiments. At each pH, five to seven different substrate concentrations were measured in 100 mM buffer, 0.01 % BSA with 16 µg of YmhII-11 enzyme. The hydrolysis reactions were monitored using a Cary 3E UV-vis spectrophotometer at room temperature and at a wavelength of 400 nm. All Michaelis-Menten parameters were calculated using a standard nonlinear least-squares fit of the rate versus substrate concentration data (GraphPad Prism).

4.3. Results

4.3.1. Trial Screening Protocol -1.

Unexpectedly, one third of the wells in the screening plate were detected to have similar activities to the wild type enzyme. This experiment was repeated to make sure that no contamination had occurred during preparation of the plate, again the experiment gave greater than 40% hits included the negative control. DNA sequencing of ten selective hits showed that one of the selected clones would give a truncated protein (i.e., the sequence contained a stop codon). The stepwise addition of substances into the wells of the fluorescent plate was performed in order to discover the source of the high background intensity, which turned out to be caused by the growth media used in the experiments.

4.3.2. Library Screening Protocol – 2.

As a result of the high background readings, an alternative screening method was developed using agar plates with Sia-2,6 Gal-X (Scheme 4.1) as substrate. A total of ~10,000 colonies were produced and the sialidase activity was determined by the

presence of a blue coloured colony. The blue pigment originates from the hydrolyzed "X" component (5-bromo-4-chloro-indoxyl) undergoes dimerization upon oxidization to form the blue insoluble pigment 5,5'-dibromo-4,4'-dichloro-indigo. The blue signal indicates the presence of a functional sialidase that hydrolyzes Sia-2,6 Gal-X to produce Gal-X, which is then hydrolyzed by the β-galactosidase present within the *E. coli* host. These library plates were monitored sequentially day by day for seven days and a total of 13 blue colonies emerged.

The amino acid sequence alignments for these 13 clones and the wild type sialidase are shown in Figure 4.1. The numbering of each mutant involved the YmhII abbreviation that refers to the tyrosine nucleophile substitution to either a methionine or a histidine residue. The numerical value refers to the order of blue colony appearance through the sequential seven days of monitoring. Wh means whole or complete amino acid sequence was obtained, while T7t means only the sequence using the T7 terminator primer gave good sequences but the sequence using T7 promoter primer was inconclusive. The amino acids that are different to the wild type are underlined and the nucleophile at amino acid position 370 is bolded.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                         10         20         30         40         50
WT-Act-Sia      MTANPYLRRL PRRRAVSFLL APALAAATVA GASPAQAIAG APVPPGGEPL
YmhII-1 Wh      FTANPYLRRL PRRRAVSFLL APALAAATVA GASPAQAIAG APVPPGGEPL
YmhII-2 Wh      FTANPYLRRL PRRRAVSFLL APALAAATVA GASPAQAIAG APVPPGGEPL
YmhII-3 Wh      FTANPYLRRL PRRRAVSFLL APALAAATVA GASPAQAIAG APVPPGGEPL
YmhII-4 Wh      FTANPYLRRL PRRRAVSFLL APALAAATVA GASPAQAIAG APVPPGGEPL
YmhII-5 T7t     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
YmhII-6 Wh      FTANPYLRRL PRRRAVSFLL APALAAATVA GASPAQAIAG APVPPGGEPL
YmhII-8 Wh      FTANPYLRRL PRRRAVSFLL APALAADTVA GASPAQAIAG APVPPGGEPL
YmhII-9 Wh      FTANPYLRRL PRRRAVSFLL APALAAATVA GASPAQAIAG APATPGGEPL
YmhII-11 T7t    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

```
YmhII-11 Wh  FTANPYLRRL PRRRAVSFLL APALAAATVA GASPAQAIAG APVPPGGEPL
YmhII-10 Wh  FTANPYLRRL PRRRAVSFLL APALAAATVA GASPAQAIAG APVPPGGEPL
YmhII-12 Wh  FTANPYLRRL PRRRAVSFLL APALAADTVA GASPAQAIAG APVPPGGEPL
YmhII-13 Wh  FTANPYLRRL PRRRAVSFLL APALAAATVA GASPAQAIAG APATPGGEPL

                ....|....| ....|....| ....|....| ....|....| ....|....|
                         60         70         80         90        100
WT-Act-Sia   YTEQDLAVNG REGFPNYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL
YmhII-1 Wh   YTEQDLAVNG REGFPNYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL
YmhII-2 Wh   YTEQDLAVNG REGFPNYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL
YmhII-3 Wh   YAEQDLAVNG REGFPNYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL
YmhII-4 Wh   YTEQDLAVNG REGFPNYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL
YmhII-5 T7t  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
YmhII-6 Wh   YTEQDLAVNG REGFPNYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL
YmhII-8 Wh   YTEQDLAVNG RESFPNYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL
YmhII-9 Wh   YTEQDLAGNG MEGFPYYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL
YmhII-11 T7t ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
YmhII-11 Wh  YTEQDLAVNG REGFPNYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL
YmhII-10 Wh  YTEQNLAVNG REGFPNYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL
YmhII-12 Wh  YTEQDLAVNG RESFPNYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL
YmhII-13 Wh  YTEQDLAGNG MEGFPYYRIP ALTVTPDGDL LASYDGRPTG IDAPGPNSIL

                ....|....| ....|....| ....|....| ....|....| ....|....|
                        110        120        130        140        150
WT-Act-Sia   QRRSTDGGRT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFNFHVYS
```

YmhII-1 Wh QRRSTDGGRT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFNFHVYS

YmhII-2 Wh QRRSTDGGRT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFNFHVYS

YmhII-3 Wh QRRSTDGGRT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFNFHVYS

YmhII-4 Wh QRRSTDGGRT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFNFHVYS

YmhII-5 T7t ~~~~~~~~GT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFNFHVYS

YmhII-6 Wh QRRSTDGGRT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFNFHVYS

YmhII-8 Wh QRRSTDGGWT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFNFHVYS

YmhII-9 Wh QRRSTDGGRT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFDFHVYS

YmhII-11 T7t ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

YmhII-11 Wh QRRSTDGGRT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFNFHVYS

YmhII-10 Wh QRRSSDGGRT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFNFHVYS

YmhII-12 Wh QRRSTDGGWT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFNFHVYS

YmhII-13 Wh QRRSTDGGRT WGEQQVVSAG QTTAPIKGFS DPSYLVDRET GTIFDFHVYS

....|....| ....|....| ....|....| ....|....| ....|....|

160 170 180 190 200

WT-Act-Sia QRQGFAGSRP GTDPADPNVL HANVATSTDG GLTWSHRTIT ADITPDPGWR

YmhII-1 Wh QRQGFAGSRP GTDPADPNVL HANVATSTDG GLTWSHRTIT ADIHPGSGLA

YmhII-2 Wh QRQGFAGSRP CTDPADPNVL HANVATSTDG GLTWSHRTIT ADITPDPGWR

YmhII-3 Wh QRQGFAGSRP GTDPADPNVL HANVATSTDG GLTWSHRTIT ADITPDPGWR

YmhII-4 Wh PRQGFAGSRP GTDPADPNVL HANVATSTDG GLTWSHRTIT ADITPDPGWR

YmhII-5 T7t QRQGFAGSRP GTDPADPNVL HANVATSTDG GLTWSHRTIT ADITPDPGWR

YmhII-6 Wh QRQGFAGSRP CTDPADPNVL HANVATSTDG GLTWSHRTIT ADITPDPGWR

YmhII-8 Wh QRQGFAGSQP GTDPADPNVL HANVATSTDG GLTWSHRTIT ADITPDPGWR

YmhII-9 Wh QRQGFAGSRP GTDPADPNVL HANAATSTDG GLTWSHRTIT ADFTPDPGWR

YmhII-11 T7t ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

```
YmhII-11 Wh   QRQGFAGSRP CTDPADPNVL HANVATSTDG GLTWSHRTIT ADITPDPGWR
YmhII-10 Wh   QRQGFAGSRP GTDPADPNVL HANVATSTDG GLTWSHRTIT ADITPDPGWR
YmhII-12 Wh   QRQGFAGSQP GTDPADPNVL HANVATSTDG GLTWSHRTIT ADITPDPGWR
YmhII-13 Wh   QRQGFAGSRP GTDPADPNVL HANAATSTDG GLTWSHRTIT ADFTPDPGWR


                  ....|....| ....|....| ....|....| ....|....| ....|....|
                          210        220        230        240        250
WT-Act-Sia    SRFAASGEGI QLRYGPHAGR LIQQYTIINA AGAFQAVSVY SDDHGRTWRA
YmhII-1 Wh    QPLAASGEGI QLRYGPHAGR LIQQYTIINA AGAFQAVSVY SDDHGRTWRA
YmhII-2 Wh    SRFAASGEGI QLRYGPHAGR LIQQYTIINA AGAFQAVSVY SDDHGRTWRA
YmhII-3 Wh    SRFAASGEGI QLRYGPHAGR LIQQYTIINA AGAFQAVSVY SDDHGRTWRA
YmhII-4 Wh    SRFAASGEGI QLRYGPHAGR LIQQYTIINA AGAFQAVSVY SDDHGRTWRA
YmhII-5 T7t   SRFAASGEGI QLRYGPHAGR LIQQYTIINA AGAFQAVSVY SDDHGRTWRA
YmhII-6 Wh    SRFAASGEGI QLRYGPHAGR LIQQYTIINA AGAFQAVSVY SDDHGRTWRA
YmhII-8 Wh    SRFAASGEGI QLRYGPHAGR LIQQYTIINA AGAFQAVSVY SDDHGRTWRA
YmhII-9 Wh    SRFAASGEGI QLRYGRHAGR LIQQ.TIINA AGAFQAVSVY GDDHGRTWRA
YmhII-11 T7t  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
YmhII-11 Wh   SRFAASGEGI QLRYGPHAGR LIQQYTIINA AGAFQAVSVY SDDHGRTWRA
YmhII-10 Wh   SRFAASGEGI QLRYGPHAGR LIQQYTIINA AGAFQAVSVY SDDHGRTWRA
YmhII-12 Wh   SRFAASGEGI QLRYGPHAGR LIQQYTIINA AGAFQAVSVY SDDHGRTWRA
YmhII-13 Wh   SRFAASGEGI QLRYGRHAGR LIQQ.TIINA AGAFQAVSVY GDDHGRTWRA


                  ....|....| ....|....| ....|....| ....|....| ....|....|
                          260        270        280        290        300
WT-Act-Sia    GEAVGVGMDE NKTVELSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT
```

```
YmhII-1 Wh    GEAVGVGMDE NKTVELSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT
YmhII-2 Wh    GEAVGVGMDE NKTVELSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT
YmhII-3 Wh    GEAVGVGMDE NKTVELSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT
YmhII-4 Wh    GEAVGVGMDE NKTVELSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT
YmhII-5 T7t   GEAVGVGMDE NKTVELSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT
YmhII-6 Wh    GEAVGVGMDE NKTVELSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT
YmhII-8 Wh    GEAVGVGMDE NKTVELSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT
YmhII-9 Wh    GEAVGVGMDE NKTVVLSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT
YmhII-11 T7t ~~~~~~~MDE NKTVELSDGR VLLNSRDSAR SGYRKVAVST NGGHSYGPVT
YmhII-11 Wh   GEAVGVGMDE NKTVELSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT
YmhII-10 Wh   GEAVGVGMDE NKTVELSDGR VLLNSRDSAR SGYRKVAVSI DGGHSYGPVT
YmhII-12 Wh   GEAVGVGMDE NKTVELSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT
YmhII-13 Wh   GEAVGVGMDE NKTVVLSDGR VLLNSRDSAR SGYRKVAVST DGGHSYGPVT

              ....|....| ....|....| ....|....| ....|....| ....|....|
                       310        320        330        340        350
WT-Act-Sia    IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS
YmhII-1 Wh    IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS
YmhII-2 Wh    IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS
YmhII-3 Wh    IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS
YmhII-4 Wh    IDSDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS
YmhII-5 T7t   IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS
YmhII-6 Wh    IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS
YmhII-8 Wh    IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS
YmhII-9 Wh    IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS
YmhII-11 T7t IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS
```

YmhII-11 Wh IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS

YmhII-10 Wh IDRDPPDPTN NASIIRAFPD APAGSARAKV LLFSNAPSQT SRSQGTIRMS

YmhII-12 Wh IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS

YmhII-13 Wh IDRDLPDPTN NASIIRAFPD APAGSARAKV LLFSNAASQT SRSQGTIRMS

....|....| ....|....| ....|....| ....|....| ....|....|

360 370 380 390 400

WT-Act-Sia CDDGQTWPVS KVFQPGSMS**Y** STLTALPDGT YGLLYEPGTG IRYANFNLAW

YmhII-1 Wh CDDGQTWPVS KVFQPGSMS**Y** STLTALPDGT YGLLYEPGTG IRYANFNLAW

YmhII-2 Wh CDDGQTWPVS KVFQPGSMS**Y** STLTALPDGT YGLLYEPGTG IRYANFNLAW

YmhII-3 Wh CDDGQTWPVS KVFQPGSMS**Y** STLTALPDGT YGLLYEPGTG IRYANFNLAW

YmhII-4 Wh CDDGQTWPVS KVFQPGSMS**Y** STLTALPDGT YGLLYEPGTG IRYANFNLAW

YmhII-5 T7t CDDGQTWPVS KVFQPGSMS**Y** STLTALPDGT YGLLYEPGTG IRYANFNLAW

YmhII-6 Wh CDDGQTWPVS KVFQPGSMS**Y** STLTALPDGT YGLLYEPGTG IRYANFNLAW

YmhII-8 Wh CDDGQTWPVS KVFQPGSMS**H** STLTALPDGT YGLLYEPGTG IRYANFNLAW

YmhII-9 Wh CDDGQTWPVS EVFQPGSMS**M** STLTALPYGT YGLLYEPGTG IRYANFNLAW

YmhII-11 T7t CDDGQTWPVS KVFQPGSMS**K** STLTALHDGT YGLLYEPGTG IRYADFNLAW

YmhII-11 Wh CDDGQTWPVS KVFQPGSMS**Y** STLTALPDGT YGLLYEPGTG IRYANFNLAW

YmhII-10 Wh CDDGQTWPVS KVFQPGSMS**Y** STLTALPDGT YGLLYEQGTG IRYANFNLAW

YmhII-12 Wh CDDGQTWPVS KVFQPGSMS**H** STLTALPDGT YGLLYEPGTG IRYANFNLAW

YmhII-13 Wh CDDGQTWPVS EVFQPGSMS**M** STLTALPYGT YGLLYEPGTG IRYANFNLAW

```
                ....|....| ....|...
                        410
WT-Act-Sia      LKLAAALEHH HHHH....
YmhII-1 Wh      LKLAAALEHH HHHHStop
YmhII-2 Wh      LKLAAALEHH HHHHStop
YmhII-3 Wh      LKLAAALEHH HHHHStop
YmhII-4 Wh      LKLAAALEHH HHHHStop
YmhII-5 T7t     LKLAAALEHH HHHHStop
YmhII-6 Wh      LKLAAALEHH HHHHStop
YmhII-8 Wh      LKLAAALEHH HHHHStop
YmhII-9 Wh      LKLAAALEHH HHHHStop
YmhII-11 T7t    LKLAAALEHH HHHHStop
YmhII-11 Wh     LKLAAALEHH HHHHStop
YmhII-10 Wh     LKLAAALEHH HHHHStop
YmhII-12 Wh     LKLAAALEHH HHHHStop
YmhII-13 Wh     LKLAAALEHH HHHHStop
```

Figure 4.1 *Amino acid sequences alignments of 13 positive hits to that of the wild type sialidase (active site domain only).*
Wh refers to the whole amino acids sequence obtained from DNA sequencing from both directions using T7 promoter and T7 terminator primers. T7t refers to amino acids sequence data obtained from DNA sequencing using T7 terminator primer. YmhII-11 T7t and Wh are aligned and nucleophile is sequenced to be lysine and tyrosine respectively. Other amino acids difference in the alignment is underlined.

In the following mutants reversion of the nucleophile back to a tyrosine residue had occurred, YmhII-1, 2, 3, 4, 5, 6, and 10. The mutants YmhII-8 and 12 had a histidine at amino acid position 370 and had identical sequences, while mutants YmhII-9 and 13 had a methionine at position 370 and they were also identical. Disconcertingly, YmhII-9 and 13 had a stop codon at amino acid 225. The initial sequencing of YmhII-11

showed a lysine residue at amino acid position 370 but upon re-sequencing this turned out to be a tyrosine. Re-sequencing of YmhII-11 showed this had an identical sequence to that of YmhII-2.

4.3.3. Confirmations of Library Hits Mutants.

The sialidase activity of positive hits identified by library screening protocol 2 was confirmed using Sia-2,6-X as the substrate on agar plates and below is the summary table of this experiment (Table 4.1). After five days of incubation, YmhII-1, 2, 3, 4, 6, 7, 8, 9, 10, and 11 had blue colonies, while the Yh-WT control, YmhII-5, 12, and 13 showed no activity. Out of the ten mutant clones which gave blue colonies, we were interested in YmhII-8, 9, and 11 because the sequence data showed that the nucleophile was a His, Met, and Lys, respectively.

Table 4.1 *Summary of thirteen library mutant hits' sialidase activity with Yh-WT as control.*
Each clone was streaked on individual 100 × 15 mm agar plate to avoid cross plasmids transfer.

Incubation Time	Mutant Clones	Color of Colony
1 Day	YmhII-1, 3, 6, 8	Blue (remaining clones showed no colony on the plate)
2 Days	YmhII-2, 4, 7, 9, 10	Blue
	Yh-WT, YmhII-5, 11, 12, 13	White
5 Days	YmhII-11	Mixtures of blue and white
	Yh-WT, YmhII-5, 12, 13	White

4.3.4. Library Mutant Hits Expressions and Purifications.

Clones of Yh-WT, YmhII-8, 9, and 11 were expressed in 500 mL of culture media and purification of the resultant proteins was attempted. Unfortunately, YmhII-9 showed no growth upon overnight LB/Kan media incubation, while YmhII-8 growth media turned a pinkish colour and centrifugation of cells within the media was unsuccessful. Re-streaking was again performed for clones of YmhII-8 and YmhII-9's *E. coli* glycerol stock on plates with Sia-X; however, only white colonies were observed. Both Yh-WT and YmhII-11 were expressed and purified normally. Kinetic analysis of the purified YmhII-11 was followed and Yh-WT was used as control.

4.3.5. Kinetic Analysis.

The Michaelis-Menten plots were constructed for YmhII-11 at various pH values. For each pH, five to seven different PNP-α-NeuAc concentrations were measured and the K_{cat} and K_m were calculated and the data is summarized in Table 4.2 and shown in Figure 4.2 below. Based on the data obtained, the optimal pH for YmhII-11 mutant enzymes was around 5.05, which is similar to that for the wild type sialidase.

Table 4.2 *Summary of k_{cat}, K_m, and k_{cat}/K_m parameters obtained from Michaelis-Menten experiments at different pH. YmhII-11 mutant sialidase activity was measured using PNP-α-NeuAc substrate at different pH.*

pH	k_{cat} (s^{-1})	Error [a]	Km (M)	Error [a]	k_{cat}/K_m ($s^{-1}M^{-1}$)	Error [b]
pH 4.60	7.74E-02	8.08E-03	6.22E-05	2.36E-06	1.25E+03	1.38E+02
pH 5.05	1.35E-01	5.21E-03	2.87E-06	6.69E-07	4.70E+04	1.11E+04
pH 5.60	7.70E-02	7.36E-03	1.42E-05	5.55E-06	5.44E+03	2.20E+03
pH 7.03	5.26E-03	2.91E-04	2.95E-05	7.87E-06	1.78E+02	4.85E+01
pH 7.80	3.86E-03	4.19E-04	1.26E-04	4.41E-05	3.06E+01	1.12E+01
pH 8.75	2.36E-03	1.25E-03	3.43E-04	2.51E-04	6.89E+00	6.21E+00

[a] *Given error are the standard errors from the Michaelis-Menten fit of rates versus substrate concentrations.*
[b] *Errors are calculated according to reference* (101).

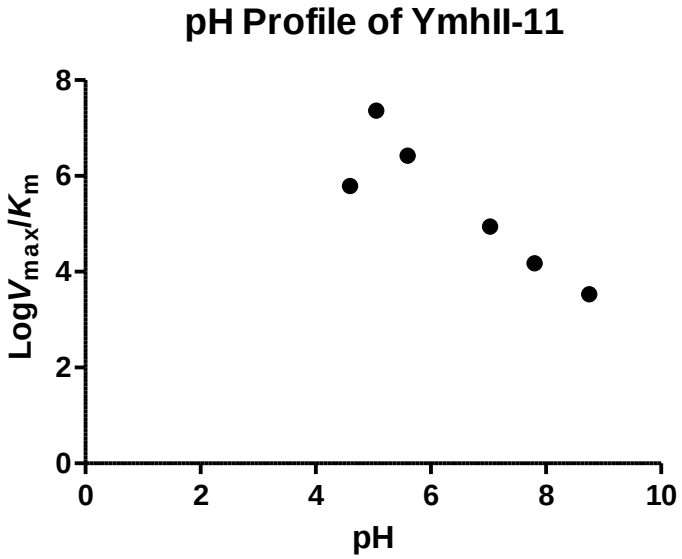


Figure 4.2 *The plot of pH vs LogV_{max}/K_m of YmhII-11 mutant enzyme. The error bar for each data point is not drawn since this clone was later determined to be wild type revertant.*

4.3.6. Activity Comparison.

Since the later sequence data of YmhII-11 showed identical DNA sequence to that of the YmhII-2, the sialidase activity of these clones were compared. The hydrolysis slope of YmhII-11 and YmhII-2 to two different substrates at two different pHs were measured and summarized in Table 4.3. By comparing the ratio of hydrolysis slopes at different pHs and substrates suggested these two mutants are different. However, the crude supernatant of YmhII-11 did not show hydrolysis activity toward PNP-α-NeuAc substrate which contradicts with purified YmhII-11 kinetic analysis. We did not characterize the YmhII-11 further because the DNA sequence suggested to be same sequence as YmhII-2 and the nucleophile is reverted to tyrosine.

Table 4.3 *Hydrolysis comparison of YmhII-2 and YmhII-11. Both YmhII-11 and YmhII-2 mutant clones' crude supernatants were used to test for the hydrolysis. The notation ND refers to not detected.*

pH	Clones	Substrates	Slope (Abs/min)
7.01	YmhII-2	PNP-α-NeuAc	0.00502
		PNPS-α-NeuAc	0.01581
	YmhII-11	PNP-α-NeuAc	ND
		PNPS-α-NeuAc	0.00046
5.25	YmhII-2	PNP-α-NeuAc	0.00105
		PNPS-α-NeuAc	0.02760
	YmhII-11	PNP-α-NeuAc	0.00015
		PNPS-α-NeuAc	ND

4.4. Discussion

Of the 13 blue colonies two pairs of clones turned out to have identical protein DNA sequences were YmhII-8, -12 and YmhII-9, -13. The identical plasmids from different clones suggested that the incubation time during the rescuing step of the transformation was too long. Usually the rescuing steps were performed between 30 to 60 min depends on DNA sensitivity. For this library the rescue step was performed for

60 min; therefore, if the incubation time can be reduced in the future, the clones with identical plasmids may be minimized.

In addition, DNA sequencing of YmhII-11, YmhII-6 and YmhII-7 was attempted multiple times and involved preparing plasmids and repeated sequencing before acceptable data for the T7term portion of YmhII-11 and the whole sequence of YmhII-6 were obtained, however no good sequencing data for YmhII-7 was forthcoming. This result can be explained by more than one plasmid being present in the same cell. The same competent cell could have either the vector without an insert, vector with an insert or multiple vectors with different inserts. In the former case, sequencing of a mixture of empty vector and vector with an insert would be difficult to obtain. In the latter case, a mixture of vectors with different inserts would provide noisy data and the outcome of sequencing could vary. For instance, for YmhII-11 the DNA sequence was first read to be a lysine-370 mutant, but a later sequence suggested that this was a tyrosine-370 revertant.

4.4.1. Library Mutant Hits Expression.

During the expression and purification of selected library mutant clones, different issues became apparent. Clone YmhII-9 showed no growth when a large scale expression was performed, yet the screening plate twice showed a blue colony. The expression culture of clone YmhII-8 showed a pale pink colour and we were unable to collect any cells after centrifugation. These contradictory data suggest that the glycerol stocks from the cultures that were seeded from the blue colonies identified in the screening protocol (number 2) were not clean. As the library plates were incubated for a week at 37 °C incubator, microbes or fungus could infect the library mutant clones. Indeed, many of the library plates started to have visible fungus spores after three days in the incubator; thus, contamination of the colonies was the likely cause of YmhII-8 mutant clone.

4.4.2. Kinetic Analysis.

The pH profile of YmhII-11 experiments were performed assuming the Y370 was mutated to a lysine; however the later DNA sequence of YmhII-11 suggested that the nucleophile at position 370 is a tyrosine residue. In fact, the entire nucleotides sequence

is identical to that of the YmhII-2. The activity comparisons of these two clones were performed to deduce whether YmhII-11 clone was contaminated with YmhII-2 or not. The ratio of hydrolysis slop from the two clones with different pH and substrates were divided and compared. The activity data suggested YmhII-11 and YmhII-2 were different enzymes. However, more experiments are required before a detailed kinetic analysis is attempted for the YmhII-11 mutant.

4.5. Conclusions

Unfortunately, this library screening project did not provide any active mutant that is either a Ym or a Yh nucleophile. For all the positive hits detected from the library screening, the nucleophile is converted back to the wild-type tyrosine. The screening process needs to be repeated and several experimental steps are suggested for the next screening.

To reduce the number of identical plasmid copies at the transformation step by decreasing the rescuing incubation time in the incubator from 1h to 0.5 h. Furthermore, reducing the cell density in library plating could also minimize transfer or diffusion of mutant sialidase and decrease the false positive screening. It would be ideal if the library screening plates were replicated so that positive hits could be traced back to another identical plate that is not incubated or screened for a long period of time.

5. Conclusions and Future Directions

5.1. Y370G Mutant Sialidase Transglycosylations Assessment

Micromonospora viridifaciens Y370G mutant sialidase is capable of transferring sialic acid from aryl β-sialosides onto several sugar acceptors to give α-sialosides. Furthermore, the transglycosylation yield is higher using α-glycoside acceptors compared to β-linked sugars, the anomeric configuration of natural substrates such as α-D-sialosyl-(2→6)-β-D-galactopyranosyl-D-glucopyranose (2,6-sialyllactose). The transglycosylation products using 2-fluorophenyl β-D-sialoside as donor and methyl α-D-galactopyranoside, methyl α-D-glucopyranoside, and methyl α-D-mannopyranoside as acceptors are isolated and the coupling yields are 46.2%, 41.2%, and 46.0%, respectively. The transglycosylation products made by Y370G and a 2,6-sialyltransferase were compared and shown to be identical, thus, the coupled product from Y370G mutant sialidase with Me-α-gal as the acceptor is methyl α-D-sialosyl-(2→6)-α-D-galactopyranoside.

The Y370G mutant sialidase were used to attempt a transglycosylation reaction with a potential glycoprotein therapeutic but the coupling was unsuccessful. The most probable explanation is the limited amount of acceptor concentration. The Y370G mutant sialidase has a broad substrate specificity and does not hydrolyze transglycosylation products, but the enzyme still favours hydrolysis activity.

5.2. W and Y Loops Mutant Sialidases Library Screening

The *T. cruzi* trans-sialidase contains a hydrophobic loop on top of the active site that binds lactose and excludes water to favour transferase rather than hydrolase activity. These Trp and Tyr containing loops (WY loops) were incorporated into *M.*

viridifaciens Y370G mutant sialidase by Dr. Watson. The WY loops on the Y370G had abolished the enzyme activity; therefore, four amino acids on the Trp containing loop were targeted for random mutagenesis in order to improve transferase activity. Even though this library screen did not provide mutants that favour transfer over hydrolysis, but the W and Y loop region mutants have reduced the hydrolysis rate greatly.

Nevertheless, from this W-loop mutant library we learned that the full sialidase construct is more stable and results in higher expression yields compared to the active site construct. Hence, future sialidase mutant libraries should use the full length sialidase construct. The present library focused on mutation of four amino acids in the W-loop, and the library can also be randomly mutagenized on the whole construct.

The present screening method using expressed mutant sialidases' supernatants and GELF substrates had its limitations. Since the protein expression level varies significantly, one can incorporate green fluorescent protein gene in the library construct to indicate expression levels. The GELF substrate had solubility problems hence addition of DMSO organic solvent was necessary, but it hinders enzyme activity measurement.

5.3. Y370M and Y370H Mutant Sialidases Library Screening

In methanol, methionine and histidine are expected to be a better nucleophiles than tyrosine based on the published nucleophilicity value (79), yet the Y370M and Y370H mutant *M. viridifaciens* sialidases eliminated enzymes activity. Random mutagenesis based on these two clones was performed and screening for an active sialidase, in which the purpose of elucidating novel sialidase mechanism is investigated.

The first attempt of library screening performed in this thesis is unsuccessful; all the positive hits turned out to be revertants to tyrosine at position 370. Nevertheless, improvements can be considered to improve screening methods in the future. One of the issues encountered in the present screening is the repetitive plasmid sequence in different clones and inconsistence of sialidase activity. The screening process with bacteria colonies on the agar plates preceded more than one week, so replicate plates were recommended to minimize the contaminations. Reduce cell density on screening

plates or pick each colony into individual wells can decrease the chance false positive hits. Using less time in the rescuing incubation at the transformation step might lessen the amount of identical plasmids copies and reduce the number of identical clone hits.

6. References

1. Aisaka K, Igarashi A, Uwajima T. 1991. *Purification, crystallization, and characterization of neuraminidase from micromonospora viridifaciens.* Agriculture Biological Chemistry. 55 : 997-1004

2. Ajisaka K, Fujimoto H, Isomura M. 1994. Carbohydrate Research. 259 : 103-15

3. Ajisaka K, Fujimoto H, Isomura M. 1994. *Regioselective transglycosylation in the synthesis of oligosaccharides: Comparison of b-galactosidases and sialidases of various origins.* Carbohydrate Research. 259 : 103-115

4. Al-Rubeai M. 2002. *Cell engineering, volume 3 : Glycosylation,* Hingham, MA, USA: Kluwer Academic Publishers

5. Angata T, Varki A. 2002. *Chemical diversity in the sialic acids and related alpha-keto acids: an envolutionary perspective.* Chemical Review. 102 : 439-469

6. Angata T, Varki A. 2002. *Chemical diversity in the sialic acis and related a-keto acids: an evolutionary perspective.* Chemical Review. 102 : 439-470

7. Ashwell G, Harford J. 1982. *Carbohydrate-specific receptors of the liver.* Annual Review of Biochemistry. 51 : 531-554

8. Ashwell G, Harford J. 1982. *Carbohydrate-specific receptors of the liver.* Annual Review of Biochemistry. 51 : 531-554

9. Basset C, Devauchelle V, Durand V, Jamin C, Pennec YL, et al. 1999. *Glycosylation of immunoglobulin A influences its receptor binding..* Scandinavian Journal of Immunology. 50 : 572-579

10. Blix FG, Gottschalk A, Klenk E. 1957. *Proposed nomenclature in the field of neuraminic and sialic acids.* Nature. 179 : 1088

11. Burda P, Aebi M. 1999. *The dolichol pathway of N-linked glycosylation.* Biochimica et Biophysica Acta. 1426 : 239-257

12. Buschiazzo A, Alzari PM. 2008. *Structural insight into sialic acid enzymology.* Current Opinion in Chemical Biology. 12 : 565-572

13. Buschiazzo A, Tavares GA, Campetella O, Spinelli S, Cremona ML, et al. 2000. *Structural basis of sialyltransferase activity in trpanosomal sialidases.* EMBO Journal. 19 : 16-24

14. Buschiazzo A, Amaya MF, Cremona ML, Frasch AC, and Alzari PM. 2002. *The crystal structure and mode of action of trans-sialidase, a key enzyme in trypanosoma cruzi pathogenesis.* Molecular Cell. 10 : 757-768

15. Chan J, Lewis AR, Indurugalla D, Schur M, Wakarchuk W, Bennet AJ. 2012. *Transition state analysis of vibrio cholerae sialidase-catalyzed hydrolyses of natural substrate analogues.* Journal of American Chemistry Society. 134 : 3748-3757

16. Chan J, Lu A, Bennet AJ. 2011. *Turnover is rate-limited by deglycosylation for micromonospora viridifaciens sialidase-catalyzed hydrolyses: Conformational implications for the michaelis complex.* Journal of American Chemistry Society. 133 : 1877-1884

17. Chan J, Watson JN, Lu A, Borgford TJ, Bennet AJ. 2012. *Bacterial and viral sialidases: Contribution of the conserved active site glutamate to catalysis.* Biochemistry. 50 : 433-441

18. Chan J, Lu A, Bennet AJ. 2011. *Turnover is rate-limited by deglycosylation for micromonospora viridifaciens sialidase-catalyzed hydrolyses: Conformational implications for the michaelis complex.* Journal of the American Chemical Society. 133 : 2989-97

19. Chen X, Varki A. 2010. *Advances in the biology and chemistry of sialic acids.* ACS Chemical Biology. 5 : 163-176

20. Comolli JC, Waite LL, Mostov KE, Engel JN. 1999. *Pili binding to asialo-GM1 on epithelial cells can mediate cytotoxicity or bacterial internalization by pseudomonas aeruginosa.* Infection and Immunology. 67 : 3207-3214

21. Corfield AP, Higa H, Paulson JC, Schauer R. 1983. *The specificity of viral and bacterial sialidases for α(2–3)- and α(2–6)-linked sialic acids in glycoproteins.* Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology. 744 : 121-126

22. Corfield T. 1992. *Bacterial sialidases - roles in pathogenicity and nutrition.* 2 : 509-521

23. Corfield T, Wagner SA, Clamp JR, Krians MS, Hoskins LC. 1992. *Mucin degradation in the human colon. production of sialidase, sialate 0-acetylesterase, W-acetylneuraminate lyase, arylesterase and glycosulfatase activities by strains of fecal bacteria.* Infection and Immunity. 60 : 3971-3977

24. Cregg JM, Cereghino JL, Shi JY, Higgins DR. 2000. *Recombinant protein expression in pichia pastoris.* Molecular Biotechnology. 16 : 23-52

25. Crocker PR, Varki A. 2001. *Siglecs, sialic acids and innate immunity.* Trends in Immunology. 22 : 337-342

26. Daines DA, Wright LF, Chaffin DO, Rubens CE, Silver RP. 2000. *NeuD plays a role in the synthesis of sialic acid in escherichia coli K1.* FEMS Microbiology Letters. 189 : 281-284

27. Damager I, Buchini S, Amaya MF, Buschiazzo A, Alzari P, et al. 2008. *Kinetic and mechanistic analysis of trypanosoma cruzi trans-sialidase reveals a classical ping-pong mechanism with Acid/Base catalysis.* Biochemistry. 47 : 3507-3512

28. Diwu Z, Klaubert DH, Haugland RP. 1999. *Spectral properties and biological applications of ELF enzyme substrates that yield fluorescent precipitates at the enzymatic activity sites.* Part of the SPIE Conference Advances in Fluorescence Sensing Technology IV. 3602 : 265-274

29. Dookhun V, and Bennet AJ. 2005. *Unexpected stability of aryl β-N-acetylneuraminides in neutral solution: Biological implications for sialyl transfer reactions.* Journal of the American Chemical Society. 127 : 7458-7465

30. Dookhun V, Bennet AJ. 2005. *Unexpected stability of aryl b-N-acetylneuraminides in neutral solution: Biological implications for sialyl transfer reactions..* Journal of American Chemistry Society. 127 : 7458-7465

31. Drickamer K. 1991. *Clearing up glycoptotein hormones.* Cell. 67 : 1029-1032

32. Eckhardt M, Mühlenhoff M, Bethe A, Gerardy-Schahn R. 1996. *Expression cloning of the golgi CMP-sialic acid transporter.* Proceedings of the National Academy of Sciences of the United States of America. 93 : 7572-7576

33. Edwards U, Frosch M. 1992. *Sequence and functional analysis of the cloned neisseria meningitidis CMP-NeuNAc synthetase.* FEMS Microbiology Letters. 75 : 161-166

34. Elliott S, et al. 2003. *Enhancement of therapeutic protein in vivo activities through glycoengineering.* Nature BioTechnology. 21 : 414-421

35. Friedman RL, Iglewski BH, Roerdink F, Alving CR. 1982. *Suppression of cytotoxicity of diphtheria toxin by monoclonal antibodies against phosphatidylinositol phosphate.* Biophysical Journal. 37 : 23-24

36. Gelebarta P, Opasb M, Michalaka M. 2005. *Calreticulin, a Ca2+ binding chaperone of the endoplasmic reticulum.* The International Journal of Biochemistry and Cell Biology. 37 : 260-266

37. Gloster TM, Vocadlo DJ. 2012. *Developing inhibitors of glycan processing enzymes as tools for enabling glycobiology.* Nature Chemical Biology. 8 : 683-694

38. Gouchee CF, Gramer J, Andersen DC, Bahr JB, Rasmussen JR. 1991. *The oligosaccharides of glycoproteins; bioprocess factors affecting oligosaccharide structure and their effect on glycoprotein properties.* Bio Technology. 9 : 1347-1355

39. Green ED, Adelt G, Baenziger JU, Wilson S, Van Halbeek H. 1988. *The asparagine-linked oligosaccharides on bovine fetuin. structural analysis of N-glycanase-released oligosaccharides by 500-megahertz 1H NMR spectroscopy.* Journal of Biological Chemistry. 263 : 18253-18268

40. Grewal PK, Uchiyama S, Ditto D, Varki N, Le DT, et al. 2008. *The ashwell receptor mitigates the lethal coagulopathy of sepsis.* Nature Medicine. 14 : 648-655

41. Hamilton SR, et al. 2003. *Production of complex human glycoproteins in yeast.* Science. 301 : 1244-1246

42. Helenius A, Aebi M. 2001. *Intracelluar function of N-linked glycans.* Science. 291 : 2364-2369

43. Hinderlich S, Weidemann W, Yardeni T, Horstkorte R, Huizing M. 1997. *UDP-GlcNAc 2-epimerase/ManNAc kinase (GNE): a master regulator of sialic acid synthesis.* Journal of Biological Chemistry. 272 : 24319-24324

44. Hollister JR, Jarvis DL. 2001. *Engineering lepidopteran insect cells for sialoglycoprotein production by genetic transformation with mammalian 1,4-galactosyltransferase and 2,6-sialyltransferase genes.* Glycobiology. 11 : 1-9

45. Hounsell EF, Davies MJ, Renouf DV. 1996. *O-linked protein glycosylation structure and function.* Glycoconjugate Journal. 13 : 19-26

46. Inoue S, Kitajima K, Inoue Y. 1996. *Identification of 2-keto-3-deoxy-D-glycero-D-galactonononic acid (KDN, deaminoneuraminic acid) residues in mammalian tissues and human lung carcinoma cells.* Journal of Biological Chemistry. 271 : 24341-24344

47. Jarvis DL, Kawar ZS, Hollister JR. 1998. *Engineering N-glycosylation pathways in the baculovirus insect cell system.* Current Opinion in Biotechnology. 9 : 528-533

48. Jenkin N, Parckh RB, James DC. 1996. *Getting the glycosylation right: implications for the biotechnology industry.* Nature BioTechnology. 14 : 975-981

49. Kakuta Y, Okino N, Kajiwara H, Ichikawa M, Takakura Y, et al. 2008. *Crystal structure of vibrionaceae photobacterium sp JT-ISH-224 a2,6-sialyltransferase in a ternary complex with donor product CMP and acceptor substrate lactose: Catalytic mechanism and substrate recognition.* Glycobiology. 18 : 66-73

50. Karwaski MF, Wakarchuk WW, Gilbert M. 2002. *High-level expression of recombinant neisseria CMP-sialic acid synthetase in escherichia coli.* Protein Expression and Purification. 25 : 237-240

51. Kim S, Oh DB, Kang HA, Kwon O. 2011. *Features and applications of bacterial sialidases.* Application of Microbial Biotechnology. 91 : 1-15

52. Kim S, Oh DB, Kwon O, Kang HA. 2010. *Identification and functional characteriztion of the NanH extracellular sialidase from corynebacterium diphtheriae.* Journal of Biochemistry. 147 : 523-533

53. Klenk E. 1941. *Neuraminsäure, das spaltprodukt eines neuen gehirnlipoids.* Zeitschrift für Physikalische Chemie (Journal of Physical Chemistry). 268 : 50-58

54. Kornfeld R, Kornfeld S. 1985. *Assembly of asparagine-linked oligosaccharides.* Annual Review of Biochemistry. 54 : 631-634

55. Koshland D. 1953. *Stereochemistry and the mechanism of enzymatic reactions.* Biological Review. 28 : 416-436

56. Kronman C, Velan B, Marcus D, Ordentlich A, Reuveny S, Shafferman A. 1995. *Involvement of oligomerization, N-glycosylation and sialylation in the clearance of cholinesterases from the circulation.* Biochemistry Journal. 311 : 959-967

57. Kubo H, Irie A, Inagaki F, Hoshi M. 1990. *Gangliosides from the eggs of the sea urchin, anthocidaris crassispina.* Journal of Biochemistry. 108 : 185-192

58. Kundra R, Kornfeld S. 1999. *Asparagine-linked oligosaccharides protect lamp-1 and lamp-2 from intracellular proteolysis.* Journal of Biological Chemistry. 274 : 31039-31046

59. Lairson LL, Henrissat B, Davies GJ, Withers SG. 2008. *Glycosyltransferases: structures, funcations, and mechanisms.* Annual Review of Biochemistry. 77 : 521-555

60. Lee SS, Hong SY, Errey JC, Izumi A, Davies GJ, Davies BG. 2011. *Mechanistic evidence for a front-side, S_Ni-type reaction in a retaining glycosyltransferase.* Nature chemical biology. 7 : 631-638

61. Lentz MR, Webster RG, Air GM. 1987. *Site-directed mutation of the active site of influenza neuraminidase and implications for the catalytic mechanism.* Biochemistry. 26 : 5351-5358

62. Ley K. 2003. *The role of selectins in inflammation and disease.* Trends in Molecular Medicine. 9 : 263-268

63. Lobban P, Kaiser A. 1973. *Enzymatic end-to end joining of DNA molecules.* Journal of Molecular Biology. 78 : 453-471

64. Maillet F, Kazatchkine MD. 1983. *Modulation of the formation of the human amplification C3 convertase of complement by polycations.* Immunology. 50 : 27-33

65. Maras M, Saelens X, Laroy W, Piens K, Claeyssens M, et al. 1997. *In vitro conversion of the carbohydrate moiety of fungal glycoproteins to mammalian-type oligosaccharides.* European Journal of Biochemistry. 249 : 701-707

66. Marcha I, Cerutti M, Mir AM, Juliant S, Devauchefle G, et al. 2001. *Expression of a membrane-bound form of trypanosoma cruzi trans-sialidase in baculovirus-infected insect cells: A potential tool for sialylation of glycoproteins produced in the baculovirus-insect cells system.* Glycobiology. 11 : 593-603

67. McCarter JD, Withers SG. 1994. *Mechanisms of enzymatic glycoside hydrolysis.* Current Opinion Structural Biology. 4 : 885-892

68. Mellquist JL, Kasturi L, Spitalnik SL, Shakin-Eshleman SH. 1998. *The amino acid following an asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency.* Biochemistry. 37 : 6833-6837

69. Miyagi T, Takahashi K, Hata K, Shiozaki K, Yamaguchi K. 2012. *Sialidase significance for cancer progression.* Glycoconjugate Journal. 29 : 567-577

70. Mohandas N, Narla A. 2005. *Blood group antigens in health and disease.* Current opinion in hematology. 12 : 135-140

71. Monti E, Bonten E, D'Azzo A, Bresciani R, Venerando B, et al. 2010. *Sialidases in vertebrates: A family of enzymes tailored for several cell functions..* Advances in Carbohydrate Chemistry and Biochemistry. 64 : 403-479

72. Mortensen B, Huseby NE. 1997. *Clearance of circulating g-glutamyltransferase by the asialoglycoprotein receptor. enzyme forms different sialic acid content are eliminated at different clearance rates and without apparent desialylation.* Clinica Chimica Acta. 258 : 47-58

73. Nakata D, Close BE, Colley KJ, Matsuda T, Kitajima K. 2000. *Molecular cloning and expression of the mouse N-acetylneuraminic acid 9-phosphate synthase which does not have deaminoneuraminic acid (KDN) 9-phosphate synthase activity.* Biochemical and Biophyscal Research Communications. 273 : 642-648

74. Naleway JJ, Fox CMJ, Robinhold D, Terpetschnig E, Olson NA, Haugland RP. 1994. *Synthesis and use of new fluorogenic precipitating substrates.* Science Direct, Tetrahedron Letters. 35 : 8569-8572

75. Newstead SL, Potter JA, Wilson JC, Xu G, Chien CH, et al. 2008. *The structure of clostridium perfringens NanI sialidase and its catalytic intermediates.* Journal of Biological Chemistry. 283 : 9080-9088

76. Okamoto K, Hasegawa T, Toyomaki Y, Yamakawa M, Okukado N. 1992. *3b-hydroxysialic acid glycosides .1. calcium-binding ability and chemical and enzymatic stabilities.* Chemical and Pharmaceutical Bulletin. 40 : 2728-2734

77. Orlandi PA, Klotz FW, Haynes JD. 1992. *A malaria invasion receptor, the 175-kilodalton erythrocyte binding antigen of plasmodium falciparum recognizes the terminal Neu5Ac(alpha 2-3)gal- sequences of glycophorin A.* Journal of Cell Biology. 116 : 901-909

78. Pan L, Yamauchi K, Sawai T, Nakadate T, Kojima Y, et al. 2011. *Inhibition of binding of E- and P-selectin to sialyl-lewis X molecule suppresses the inflammatory response in hypersensitivity pneumonitis in mice.* American Journal of Respiratory and Critical Care Medicine. 161 : 1689-1697

79. Pearson RG, Sobel H, Songstad J. 1968. *Nucleophilic reactivity constants toward methyl iodide and trans-[pt(py)$_2$Cl$_2$].* Journal of the American Chemical Society. 90 : 319-326

80. Reid ME, Mohandas N. 2004. *Red blood cell blood group antigens: Structure and function.* Seminars in hematology. 41 : 93-117

81. Rogl H, Kosemund K, Kuhlbrandt W, Collinson I. 1998. FEBS letters. 432 : 21-6

82. Rosenberg A, Schengrund CL, eds. 1976. *Biological Roles of Sialic Acids,* New York and London: Plenum Press.

83. Sakurada K, Ohta T, Hasegawa M. 1992. *Cloning, expression, and characterization of the micromonospora viridifaciens neuraminidase gene in streptomyces lividans.* Journal of Bacteriology. 174 : 6896-6903

84. Sato M, Sadamoto R, Niikura K, Monde K, Kondo H, Nishimura SI. 2004. *Site-specific introduction of sialic acid into insulin.* Angewandte Chemie International Edition. 43 : 1516-1520

85. Schauer R. 2009. *Sialic acids as regulators of molecular and cellular interactions.* Current Opinion in Structural Biology. 19 : 507-514

86. Schauer R. 2000. *Achievements and chanllenges of sialic acid research.* Glycoconjugate Journal. 17 : 485-499

87. Schauer R. 1985. *Sialic acids and their roles as biological masks.* Trends in Biochemical Sciences. 10 : 357-360

88. Schmidt D, Sauerbrei B, Thiem J. 2000. *Chemoenzymatic synthesis of sialyloligosccharides with sialidases employing transglycosidation method.* Journal of Organic Chemistry. 65 : 8518-8526

89. Schmidt D, Sauerrei B, Thiem J. 2000. 65 : 8518-26

90. Schrag JD, Bergeron JJ, Li Y, Borisova S, Hahn M, et al. 2001. *The structure of calnexin, an ER chaperone involved in quality control of protein folding.* Molecular cell. 8 : 633-644

91. Shidmoossavee FS, Cheng L, Watson JN, Bennet AJ. 2010. *Brønsted analysis of an enzyme-catalyzed pseudo-deglycosylation reaction: Mechanism of desialylation in sialidases..* Biochemistry. 49 : 6473-6484

92. Shidmoossavee FS, Cheng L, Watson JN, Bennet AJ. 2010. *Bronsted analysis of an enzyme-catalyzed pseudo-deglycosylation reaction: Mechanism of desialylation in sialidases.* Biochemistry. 49 : 6473-84

93. Shidmoossavee FS, Cheng L, Watson JN, Bennet AJ. 2010. *Bronsted analysis of an enzyme-catalyzed pseudo-deglycosylation reaction: Mechanism of desialylation in sialidases.* Biochemistry. 49 : 6473-84

94. Sinnott ML. 1990. *Catalytic mechansims of enzymatic glycosyl transfer.* Chemical Review. 90 : 1171-1202

95. Steen PV, Rudd PM, Dwek RA, Opdenakker G. 1998. *Concepts and principles of O-linked blycosylation.* Critical Reviews in Biochemistry and Molecular Biology. 33 : 151-208

96. Storring PL. 1992. *Assaying glycoprotein hormones - the influence of glycosylation on immunoreactivity.* TIBTECH. 10 : 427-432

97. Sun WC, Gee KR, Haugland RP. 1998. *Synthesis of novel fluorinated coumarins: excellence UV-light excitable fluorescent dyes.* Biooranic Medicine of Chemistry Letter. 8 : 3107-3110

98. Sylvestre J, Chautard H, Cedrone F, Delcourt M. 2006. *Directed evolution of biocatalysts.* Organic Process Research & Development. 10 : 562-571

99. Taylor G. 1996. *Sialidases: Structures, biological significance andtherapeutic potential.* Current Opinion in Structural Biology. 6 : 830-837

100. Taylor G. 1996. *Sialidases: structures, biological significance and therapeutic potential.* Current Opinion in Structural Biology. 6 : 830-837

101. Taylor JR. 1982. The study of uncertainties in physical measurements. In *An Introduction to Error Analysis,* 73. Mill Valey, California: University science books

102. Taylor ME, Drickamer K. 2011. *Introduction to glycobiology.* Oxford New York: Oxford University Press. 3rd ed.

103. Thiem J, Sauerbrei B. 1991. *Chemoenzymatic syntheses of sialyloligosaccharides with immobilized sialidase.* Angewandte Chemie International Edition. 30 : 1503-1505

104. Tsuji S. 1996. *Molecular cloning and functional analysis of sialyltransferases..* Journal of Biochemistry. 120 : 1-13

105. Tsuji S, Datta AK, Paulson JC. 1996. *Systematic nomenclature for sialyltransferases.* Glycobiology. 6 : v-vii

106. Van Pelt J, Dorland L, Duran M, Hokke CH, Kamerling JP, Vliegenthart JFG. 1990. *(1990) sialyl-a-2-6-mannosyl-b-1-4-N-acetylglucosamine, a novel compound occurring in urine of patients with b-mannosidosis, J. biol. chem. 265, 19685-19689.* Journal of Biological Chemistry. 265 : 19685-19689

107. VanDorst, J. A. L. M., Tikkanen JM, Krezdorn CH, Streiff MB, Berger EG, et al. 1996. *Exploring the substrate specificities of a-2,6- and a-2,3-sialyltransferases using synthetic acceptor analogues.* European Journal of Biochemistry. 242 : 674-681

108. Vann WF, Tavarez JJ, Crowley J, Vimr E, Silver RP. 1997. *Purification and characterization of the escherichia coli K1 neuB gene product N-acetylneuraminic acid synthetase.* Glycobiology. 7 : 697-701

109. Varghese JN, Laver WG, Colman PM. 1983. *Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 A resolution.* Nature. 303 : 35-40

110. Varki A. 2007. *Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins.* Nature. 446 : 1023-1029

111. Varki A. 1997. *Sialic acids as ligands in recognition phenomena.* FASEB Journal. 11 : 248-255

112. Varki A, Gagneux P. 2012. *Multifarious roles of sialic acids in immunity..* Annals of the New York Academy of Sciences. 1253 : 16-36

113. Varki A, Schauer R, eds. 2008. *Essentials of Glycobiology,* New York: Cold Spring Harbor Laboratory Press. . 2nd Edition ed.

114. Vetere A, Paoletti S. 1996. *Complete synthesis of 3'-sialyl-N-acetyllactosaminie by regioselective transglycosylation.* FEBS letters. 399 : 203-206

115. Vimr ER. 1994. *Microbial sialidases: Does bigger always mean better?* Trends in Microbiology. 2 : 271-277

116. Wasley LC, Timony G, Murtha P, Stoudernine J, Domer aJ, et al. 1991. *The importance of N- and O-linked oligosaccharides for the biosynthesis and in vitro and in vivo biologic activities of erythropoietin.* Blood. 77 : 2624-2632

117. Watson JN, Dookhun V, Borgford TJ, Bennet AJ. 2003. *Mutagenesis of the conserved active-site tyrosine changes a retaining sialidase into an inverting sialidase.* Biochemistry. 42 : 12682-90

118. Watson JN, Indurugalla D, Cheng L, Narine AA, Bennet AJ. 2006. *The hydrolase and transferase activity of an inverting mutant sialidase using non-natural b-sialoside substrates.* Biochemistry. 45 : 13264-13275

119. Watson JN, Newstead S, Narine AA, Taylor G, Bennet AJ. 2005. *Two nucleophilic mutants of the micromonospora viridifaciens sialidase operate with retention of configuration by two different mechanisms.* Chembiochem. 10 : 1999-2004

120. Watson JN, Dookhun V, Borgford TJ, Bennet AJ. 2003. *Mutagenesis of the conserved active-site tyrosine changes a retaining sialidase into an inverting sialidase.* Biochemistry. 42 : 12682-90

121. Watson JN, Indurugalla D, Cheng LL, Narine AA, Bennet AJ. 2006. *The hydrolase and transferase activity of an inverting mutant sialidase using non-natural beta-sialoside substrates.* Biochemistry. 45 : 13264-75

122. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. 1992. *Evolution and ecology of influenza A viruses.* Microbiology Review. 56 : 152-179

123. Wormald MR, Dwek RA. 1999. *Glycoproteins: glycan presentation and protein-fold stablility.* Structure. 7 : 155-160

124. Zapata G, Crowley JM, Vann WF. 1992. *Sequence and expression of the escherichia coli K1 neuC gene product.* Journal of Bacteriology. 174 : 35-319