ABSTRACT

Title of Dissertation:	COMBINED CHEMICAL AND ENZYMATIC APPROACHES TO PROTEIN GLYCOSYLATION
	Sunaina Kiran Prabhu, Doctor of Philosophy, 2021
Dissertation directed by:	Professor, Lai-Xi Wang, Department of Chemistry and Biochemistry

Glycosylation is a key post-translational modification of proteins and influences the structure and biological functions of proteins. Glycoproteins are significant in treating a variety of diseases and make up a large fraction of biotherapeutics. The carbohydrate structures on the proteins regulate biological activity and pharmacokinetic properties, thereby dictating the efficacy and cost of glycoprotein drugs. However, glycoproteins expressed in biological systems are heterogeneous in nature and impose a challenge to structure-function studies as well as design of potent therapeutics. Thus, developing tools to modulate the glycan structures on proteins is highly significant. In my thesis, we have explored biological and chemoenzymatic methods to generate homogeneously glycosylated therapeutic proteins. First, we designed a glycosylation machinery in *Escherichia coli* (*E. coli*) using an N-glycosyl transferase enzyme to transfer a sugar handle onto a model protein. The protein was

then elaborated with a homogeneous glycoform using *in vitro* chemoenzymatic transglycosylation. Using this methodology, we produced a fully glycosylated human interferon alpha-2b that was biologically active and displayed significantly enhanced proteolytic stability. Next, we focused on expanding the toolbox of enzymes available to perform the chemoenzymatic glycan remodeling of proteins. Specifically, we compared the substrate specificities of the human α-L-fucosidase (FucA1) and two bacterial α-L-fucosidases (AlfC and BfFuc) with a panel of structurally well-defined core-fucosylated substrates. FucA1 was the only α-L-fucosidase to display hydrolytic activity towards full-length core-fucosylated glycopeptides and glycoproteins. Moreover, FucA1 showed low but apparent activity to remove core fucose from intact monoclonal antibodies. This finding reveals an opportunity to employ FucA1 to remove core fucose from therapeutic antibodies to improve their antibody-dependent cellular cytotoxicity. Finally, we explored modulation of core fucosylation of monoclonal antibodies through metabolic glycoengineering. We designed L-fucose analogs to potentially incorporate functionalized fucose into IgG-Fc glycan. We showed incorporation of a few fucose derivatives into antibodies and identified a concentration-dependent effect of some of the previously known analogs. While some of the novel compounds did not show effect, the study supplements the existing tools available for metabolic modulation of antibodies. In summary, these studies present feasible new approaches to produce therapeutic eukaryotic glycoproteins with desired, homogeneous glycosylation.

COMBINED CHEMICAL AND ENZYMATIC APPROACHES TO PROTEIN GLYCOSYLATION

by

Sunaina Kiran Prabhu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:
Professor Lai-Xi Wang, Chair
Professor Douglas Julin
Assistant Professor Myles Poulin
Assistant Professor Jinwoo Lee
Professor Xiaoping Zhu

Acknowledgements

My PhD career has been a challenging but fulfilling and memorable experience. I will cherish my time spent here at the University of Maryland and in the Wang lab for a lifetime. I hope the training and knowledge I have gained during this time will help me achieve my professional goals and motivate me to keep growing intellectually. Going through this journey would not have been possible without the support of people who inspired, motivated, advised and pushed me to do my best.

Firstly, I would like to thank Dr. Wang for giving me the opportunity to be a part of the Wang lab. I have learnt and grown scientifically under your guidance. You pushed me to work on my projects with a result-oriented and focused approach. This helped me to try harder during failures, which I had a fair share of during my PhD research.

Next, I would like to thank Dr. Qiang Yang for mentoring me and teaching me how to approach research problems methodically and creatively. Your positivity reinstated my belief in the *E. coli* project and helped me stay motivated through all the failures. And thank you for bringing good luck to the project! I would also like to thank Dr. Chao Li for your constant guidance and intellectual inputs through all my projects. It was always a pleasure to discuss scientific problems with you and learn from your experience.

I would like to extend my sincere thanks to all the current and past committee members. Your encouragement and support through the years during my various interactions with you have motivated me to perform better. Your questions and

constructive feedback have made me think scientifically and critically about my projects.

I would also like to thank all the current and past members of Wang lab. Working with all of you has been a pleasure and I have learnt from each one of you. I have enjoyed spending time with you during our outings.

Finally, I would like to express my gratitude to my parents, family members and friends for their constant support and for keeping me sane through the process.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADCC	Antibody dependent cellular cytotoxicity
AFP	α-fetoprotein
Asn	Asparagine
6-Az-Fuc	6-Azido fucose
BME	β-mercaptoethanol
CCK-8	Cell Counting Kit-8
CDC	Complement dependent cytotoxicity
CHO	Chinese hamster ovary
CT-F	Complex-type core fucosylated
DHB	Dihydroxybenzoic acid
DIC	N,N′-Diisopropylcarbodiimide
DMA	N, N-dimethylaniline
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNase	Deoxyribonuclease
Dol-P	Dolichol phosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
Endo	Endoglycosidase
ENGase	Endo-β-N-acetylglucosaminidase
EPO	Erythropoietin
EPPS	4-(2-Hydroxyethyl)-1-piperazinepropanesulfonic acid
ER	Endoplasmic reticulum
ESI	Electrospray ionization
FA	Formic acid
Fab	Antigen binding fragment
FBS	Fetal bovine serum
Fc	Crystallizable fragment
FcγRIIIa	Fc gamma IIIa receptor

FDA	Food and Drug Administration
2F-Fuc	2-deoxy-2-fluoro-fucose
6F-Fuc	6-fluoro fucose
2,2-diF-Fuc	2-deoxy-2,2-difluoro-fucose
2,6-diF-Fuc	2-deoxy-2,6-difluoro-fucose
6,6-diF-Fuc	6,6-difluoro fucose
6,6,6-triF-Fuc	6,6,6-trifluoro fucose
FKP	GDP-fucose pyrophosphorylase
FPLC	Fast protein liquid chromatography
Fuc	Fucose
FUT8	Fucosyl transferase 8
Gal	Galactose
GalNAc	N-acetylgalactosamine
GDP	Guanosine diphosphate
GFP	Green fluorescent protein
Glc	Glucose
GlcN	Glucosamine
GlcNAc/GN	N-acetylglucosamine
Gln	Glutamine
GM-CSF	Granulocyte macrophage colony stimulating factor
GMD	GDP-mannose 4,6-dehydratase
GnTIII	N-acetylglucosaminyltransferase III
gor	Glutathione reductase
GPI	Glycosyl phosphatidylinositol
GST	Glutathione-S-Transferase
HEK	Human embryonic kidney
hIFNα-2b/IFNα	Human interferon alpha-2b
HM	High mannose
HPLC	High pressure liquid chromatography
IBs	Inclusion bodies
IC50	Half maximal inhibitory concentration
IdeS	Immunoglobulin G-degrading enzyme of *Streptococcus pyogenes*
IFNγ	Interferon gamma

IgG	Immunoglobulin G
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IVIG	Intravenous IgG
KD	Knock-down
KO	Knock-out
LB	Luria Bertani
LC	Liquid chromatography
mAbs	Monoclonal antibodies
MALDI-TOF	Matrix-assisted laser desorption/ionization-Time of flight
Man	Mannose
MS	Mass spectrometry
NaCl	Sodium chloride
NaOAc	Sodium acetate
NGT	*N*-glycosyltransferase
NK	Natural killer
OD	Optical density
OGT	*O*-GlcNAc transferase
OST	Oligosaccharyltransferase
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PEI	Polyethylenimine
PMSF	Phenylmethylsulfonyl fluoride
PNGase F	Peptide-N-Glycosidase F
*p*NP-Fuc	*p*-Nitrophenyl α-L-fucoside
Pro	Proline
RNase B	Ribonuclease B
RP	Reverse phase
RPM	Revolutions per minute
RT	Room temperature
RTX	Rituximab
SCT	Sialylated complex type
SDS	Sodium dodecyl sulfate

Ser	Serine
SPPS	Solid-phase peptide synthesis
SPR	Surface plasmon resonance
TB	Terrific Broth
TFA	Trifluoro acetic acid
Thr	Threonine
Tris-HCl	Tris (hydroxymethyl) aminomethane hydrochloride
trxB	Thioredoxin reductase
UDP	Uridine-diphosphate
UGGT	UDP-glucose: glycoprotein glucosyltransferase
WT	Wild type

Chapter 1: Introduction

1.1 Protein glycosylation and the structural diversity

Glycosylation is the covalent attachment of carbohydrates to proteins and lipids resulting in glycoproteins and glycolipids, respectively. Specifically, protein glycosylation is one of the most ubiquitous and critical post-translational modifications of proteins that governs the fundamental properties of proteins[1-3]. Protein glycosylation is largely classified as N- and O- linked glycosylation depending on the amino acid residue that is modified with sugars. N-linked glycosylation (N-glycosylation) is the covalent attachment of sugars/glycans to the amido group of asparagine side chain at the consensus sequon N-X-S/T, where X cannot be proline. The N-glycans in eukaryotes contain an N-acetylglucosamine (GlcNAc)-β1 linkage to the Asn in proteins. In case of O-linked glycosylation (O-glycosylation), mono- or oligosaccharides are linked to the hydroxyl group of serine or threonine through N-acetylgalactosamine (GalNAc)-α1 linkage. However, there can be other types of O-glycosylation such as O-GlcNAcylation which involves GlcNAc monosaccharides attached through β-linkage to hydroxyl groups of serine or threonine[4]. Unlike N-glycosylation, O-linked glycosylation is not known to require a consensus sequon. Collectively, N- and O- glycosylation are abundant in proteins expressed on cell surface as well as in soluble secreted forms[2,5,6]. Figure 1.1 shows the various types of glycosylated proteins and lipids displayed on mammalian cell surface. Along with the N- and O- glycosylation described above, the mammalian cell surface is also adorned

with proteoglycans, glycosyl phosphatidylinositol (GPI)-anchored proteins and glycosphingolipids. Proteoglycans contain glycosaminoglycans like chondroitin-, dermatan- and heparan- sulfate linked to serine residues through a xylose. (GPI)-anchored proteins include protein chains anchored into the plasma membrane through C-terminal glycolipid modification. Glycosphingolipids refer to sphingolipids that are anchored in the plasma membrane and contain glycans linked through a glucose or galactose[2,6]. The various classes of glycosylation have distinct functional roles to play within the biological systems. Moreover, the glycan arrangement on the cell surface acts as an indicator of the physiological state of cells[7,8]. Anomalous glycosylation patterns are characteristic of certain diseases including cancers.

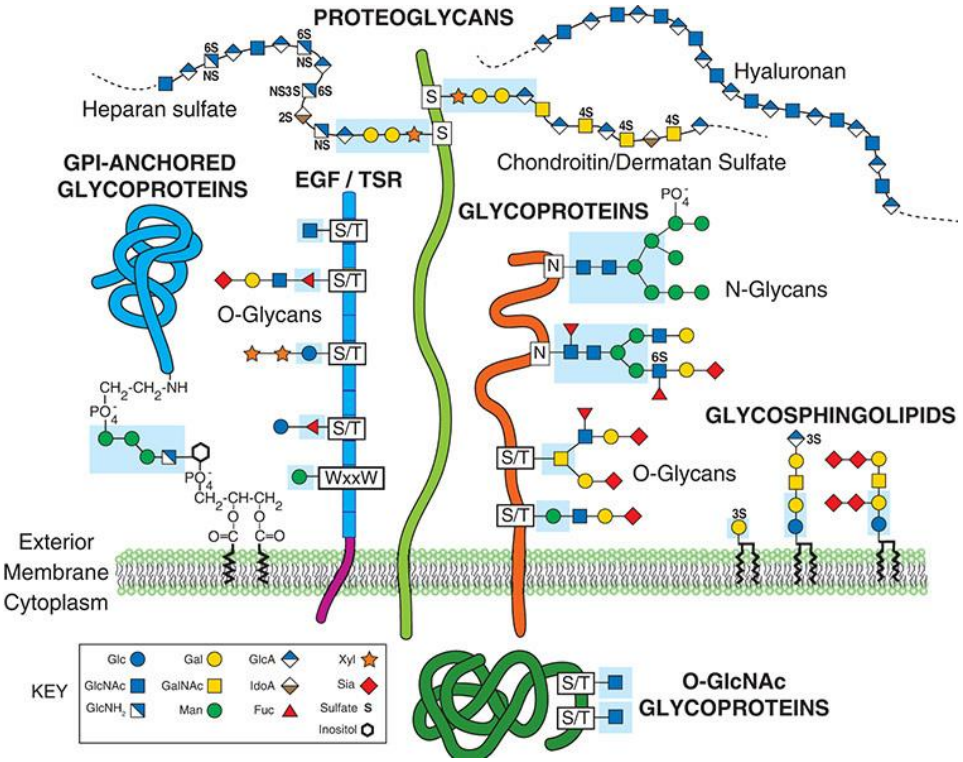


Figure 1.1. Types of glycosylation in mammals. This figure is taken from a book chapter by Varki et. al.[6]

In my research, I have focused on tools to modulate N-glycosylation of proteins for therapeutic application and functional analysis. The following introductory sections highlight the significance and need for development of tools to control N-glycosylation.

1.2 Biological implications of protein glycosylation

N-glycosylation influences the fundamental properties of glycoproteins including protein folding, solubility, stability, and biological activity[9,10]. During biosynthesis of N-glycoproteins, the endoplasmic reticulum (ER)-based quality control system, which consists of UDP-glucose:glycoprotein glucosyltransferase (UGGT), chaperones – calnexin and calreticulin and glucosidase II, monitors the folding of glycoproteins[10-12]. Misfolded proteins are recognized by UGGT potentially based on exposed hydrophobic surfaces and are subjected to cycles of glucosylation, folding and deglucosylation using the quality control proteins. Eventually, correctly folded proteins are transported to Golgi apparatus for maturation while misfolded proteins are degraded.[10-12] Moreover, glycosylation masks the underlying proteolytic cleavage sites on proteins and improves the stability of glycoproteins[13]. Also, the exposure of hydrophilic glycans on the surface of folded glycoproteins improves the solubility of glycoproteins in aqueous environments[13].

Apart from regulating the structural properties of proteins, glycosylation contributes strongly to protein functions within biological systems. Glycoproteins are involved in critical biological processes like cell adhesion, receptor signaling, intracellular transport, hormone secretion, development, and immune responses[9,13-19]. For example, core fucosylation influences the phosphorylation of epidermal growth factor receptor

(EGFR) and downstream signaling required for proliferation of epithelial cells[20]; mannose-6-phosphate glycan modification is essential for trafficking of newly synthesized lysosomal hydrolases into lysosomes, the malfunction of which leads to toxic build-up of substrates and causes metabolic diseases[10,21,22].

In the context of biotherapeutics, glycosylation has profound impacts on the biological activity, serum half-life, and efficacy of protein drugs[18]. For example, the glycan structures on monoclonal antibodies (mAbs) influence their binding to immune cell receptors and subsequent effector functions[18]. It is to be noted that a majority of biotherapeutics in the clinics are glycoproteins including mAbs, erythropoietin (EPO), enzymes and cytokines which are used to treat serious diseases[23,24]. For instance, glycosylation increases the circulatory half-life of EPO and thereby the efficacy of the glycosylated drug[25]. Similarly, glycosylation has been used to improve the physicochemical stability of therapeutic enzymes and glycoproteins like α-glucosidase, galactosidase and alpha-1-antitrypsin[26].

Glycans are also involved in interactions with glycan binding proteins/lectins and carbohydrate binding antibodies[27]. The glycan-lectin binding plays crucial roles in immunity and host-pathogen interactions[28,29]. For example, hemagglutinin expressed on influenza virus binds to sialic acid containing glycans on the surface of respiratory epithelial cells and facilitates entry of the virus into host cells to cause infection[30,31]. Conversely, receptors on immune cells recognize and bind to pathogenic cell surface carbohydrates that are different from host glycans, such as lipopolysaccharides present on bacterial cell surfaces, and prevent infections[7,32].

It is well documented that the glycosylation patterns of cells in diseased states such as cancers and autoimmune disorders is altered[7,33-35]. The aberrant glycosylation has been implicated in aiding disease progression and phenotype. For instance, increased N-glycan branching, cell surface fucosylation and sialylation are shown to participate in downregulating immune responses and promoting tumor growth and cancer metastasis[33,34]. The altered glycosylation patterns in cancers have led to identification of biomarkers that aid in diagnosis and act as therapeutic targets. Due to the reasons cited above, glycans present on pathogenic bacteria and viruses as well as the unique glycan epitopes on cancerous tissues are good targets for carbohydrate-based vaccines[36-38].

An understanding of the biological role of glycans has been historically developed through the study of glycosylation-related defects and by using mouse models to study the impact of induced mutations[13,14]. While these initial studies drastically increased the glycan knowledge in the field, it is still desirable to have access to pure, homogeneous glycans and glycan-defined proteins to perform functional studies. Glycoproteins obtained from serum and proteins expressed in common expression systems like the mammalian cells carry heterogeneous mixtures of glycoforms. Isolation of refined homogeneous glycans from such mixtures is difficult and results in low yields and purity[39]. Hence, efficient tools to modulate the heterogeneity and structural complexity of glycoproteins are highly valuable. Developing such tools requires an understanding of the origin of glycan diversity in biological systems and can be gained from an insight into the biosynthetic process of N-glycosylation.

1.3 *Biosynthesis of N-glycosylation in mammals*

N-glycosylation in mammals is a multi-step, complex process that takes place in two intracellular organelles, the ER and Golgi apparatus[40,41], as shown in Figure 1.2. The glycosylation process involves a set of enzymes including glycosyltransferases and glycosidases that catalyze transfer and removal of sugars respectively. The process begins on the cytoplasmic side of ER membrane, where the first sugar, GlcNAc is transferred to the membrane-anchored lipid carrier, dolichol phosphate (Dol-P), using uridine-diphosphate (UDP)-GlcNAc donor. Further, monosaccharide units are accordingly added stepwise from respective activated nucleotide sugar donors. Once $Man_5GlcNAc_2$-P-P-Dol chain is formed, it is flipped to the luminal side of the ER membrane where further monosaccharides are added. Ultimately, the 14-monosaccharide precursor $Glc_3Man_9GlcNAc_2$ is transferred at once from the dolichol pyrophosphate to a newly synthesized polypeptide chain at the consensus sequon by the oligosaccharyltransferase (OST) enzyme. Then, the terminal glucose residues are trimmed by glucosidases to form the mono-glucosylated glycoform, $Glc_1Man_9GlcNAc_2$, which aids in protein folding. As described earlier, this process is regulated by chaperone lectins – the membrane-bound calnexin and the soluble form, calreticulin along with the ER associated degradation pathway[10]. Correctly folded proteins are processed further to form $Man_8GlcNAc_2$-glycoproteins which are transferred to Golgi bodies for terminal glycosylation. The terminal glycosylation involves further trimming and elongation reactions. The final glycan structures on the glycoproteins are a result of the extent of processing carried out by the various enzymes distributed in the *cis*-, *medial*- and *trans*- Golgi compartments[40,41].

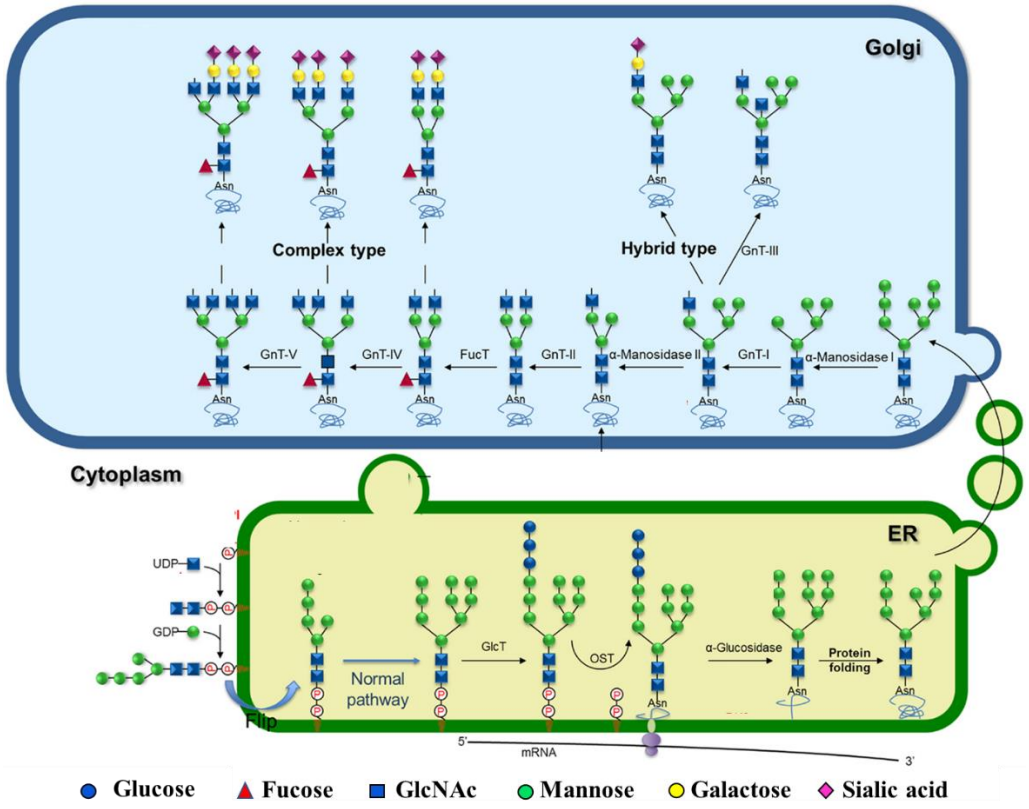


Figure 1.2. Biosynthesis of N-glycoproteins in mammals. This figure is adapted from a publication by Vasconcelos-dos-Santos et.al[42].

1.3.1 Heterogeneity of N-glycans

The process of glycosylation is highly variable and can lead to macro- and micro-heterogeneity in proteins[43]. Macro-heterogeneity arises from variability in occupancy of N-glycosylation sites and micro-heterogeneity is caused by the glycan structural diversity resulting from differential glycosylation during the biosynthetic process[43]. The non-template driven processing of glycans in the Golgi apparatus leads to formation of three major types of N-glycans namely, high-mannose, complex and hybrid type N-glycans (Figure 1.3). All N-glycoproteins share the same core structure of $Man_3GlcNAc_2$-β1-Asn; the difference lies in how the core structure is embellished

with other monosaccharides. The complex-type glycans represent the more processed glycoforms that may be adorned with monosaccharides like GlcNAc, galactose (Gal) and N-acetylneuraminic acid/sialic acid in the terminal positions, and fucose (Fuc) in the core position. The branching of sugars further determines the antennarity of the glycan structures with bi-, tri- and tetra- antennary glycoforms detected in natural and recombinant glycoproteins. In general, the extent of terminal glycosylation depends on enzyme expression, availability of nucleotide sugar donors, site of glycosylation, and overall physiological state of cells[41].

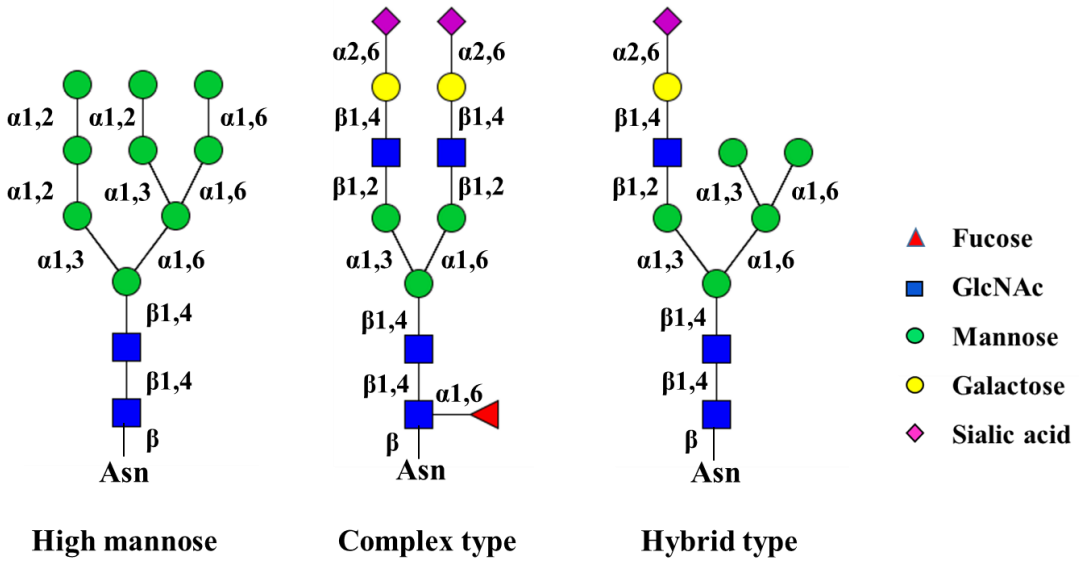


Figure 1.3. Major types of N-glycan structures in mammals. The three classes namely high mannose, complex type, and hybrid type N-glycans (from left to right) show the linkages between the monosaccharide units.

Understanding the impact of the specific structures of N-glycans is vital for designing potent therapeutics. For example, in the case of therapeutic mAbs, the presence of core

fucose has been demonstrated to reduce the binding of the IgG antibodies to the FcγIIIa receptor on natural killer (NK) cells by 50-fold, thereby resulting in lower antibody dependent cellular cytotoxicity (ADCC)[44]. Similarly, high mannose N-glycans such as $Man_5GlcNAc_2$ to $Man_9GlcNAc_2$ (termed Man5-Man9) lead to shorter circulatory half-lives and faster clearance of the glycoprotein from the body[45]. On the other hand, the presence of terminal galactose improves the complement dependent cytotoxicity (CDC) of mAbs[46]. And when the terminal galactose is capped with sialic acid residues, mAbs exhibit anti-inflammatory properties[17]. This is exemplified by the application of heavily sialylated intravenous IgG (IVIG) in the treatment of autoimmune diseases[47]. These distinct implications of antibody glycoforms on the biological effects and efficacy of therapeutic antibodies dictates the need to control the glycosylation profiles of therapeutic glycoproteins. Overall, glycosylation presents a vast functional and structural diversity that has huge implications in the functioning of an organism. It is important to examine the structure-function relationship of glycosylation and to devise methods for production of potent glycosylated biotherapeutics and to avoid undesired interactions within the body.

In the following section, the major approaches used to generate homogeneous glycoproteins are discussed.

1.4 Methods to control heterogeneity of N-glycoproteins

Numerous innovative methods have been explored to control the N-glycosylation profiles of glycoproteins for therapeutic applications and for functional studies. Some examples include gene editing of cell lines, use of small molecule inhibitors of

glycosylation enzymes, chemoenzymatic glycoengineering, and total chemical synthesis to generate homogeneous glycoproteins[39,48]. Each of these methods have been thoroughly explored and advanced through continuous efforts to generate important glycosylated proteins. The broad approaches used for controlling N-glycan heterogeneity are discussed below under three sub-categories.

1.4.1 Biosynthetic pathway modulation

One approach to controlling glycan heterogeneity through biosynthetic pathway modulation is to knock-out (KO), knock-down (KD) or overexpress specific glycosylation enzymes using cell line engineering and/or use of chemical inhibitors. Genetic engineering of cells has been applied to different expression systems including mammals, bacteria, yeast, and plants. Genetic engineering of mammalian cells, mainly Chinese hamster ovary (CHO) cells, has been primarily used to control the glycosylation of therapeutic mAbs[49]. Examples include KO of FUT8 and overexpression of N-acetylglucosaminyltransferase III (GnTIII) enzymes for reducing core fucosylation to improve the ADCC and efficacy of mAbs[50,51]. Yet another method is to target the enzymes involved in the de novo biosynthetic pathway of GDP-fucose, which is the donor of core fucose. For example, a double-KD of fucosyl transferase 8 (FUT8) and GDP-mannose 4,6-dehydratase (GMD) shows completely non-fucosylated antibody expression[52]. Non-fucosylated mAbs generated using these methods have been FDA-approved and more are under clinical trials[53,54]. In another example, overexpression of α2,6-sialyltransferase and β1,4-galactosyltransferase enzymes were used to increase the sialylation content of mAbs to improve their anti-inflammatory

properties[55]. Glycoengineering of yeast cells has been studied by multiple groups using both *Pichia pastoris* and *Saccharomyces cerevisiae* strains[56,57]. The native glycosylation of yeasts results in oligomannose glycan structures which are undesirable from a therapeutic perspective. Glycoengineering attempts have focused on replacing hyper-mannosylation with native human type complex glycans for potential expression of therapeutic glycoproteins. Several successful examples exist with improvement work required to ensure stability of the glycoengineered strains[48]. Similarly, efforts have been focused towards establishing glycosylation pathways in bacterial systems through functional transfer of glycosylation pathways or panel of enzyme into *E. coli* cells to make eukaryotic N-glycosylated proteins[58-60]. Despite appreciable efforts and success, the efficiency of glycosylation and applicability to a wide range of proteins remains to be improved.

The other approach involves use of small molecule inhibitors to intercept the glycosylation pathway to generate specific glycoforms[61]. For example, kifunensine is a strong inhibitor of mannosidase I and produces only high mannose glycans; similarly, swainsonine inhibits mannosidase II to generate hybrid type glycans[61]. While these are potent inhibitors, their use is limited to generating those respective glycoforms. Another class of inhibitors are the monosaccharide analogs that enter the metabolic pathways of respective sugars and inhibit the production of precursor nucleotide sugars. For example, 2-deoxy-2-fluoro-fucose (2F-Fuc) has been identified as a potent inhibitor of core fucosylation[62]. 2F-Fuc enters the salvage pathway of fucose to produce GDP-2F-Fuc and feedback inhibits the production of GDP-Fuc, thus controlling core

fucosylation of recombinant and cellular proteins. Similarly, 5-thio-GlcNAc has been used as a metabolic inhibitor of O-GlcNAcylation[63].

1.4.2 Total chemical synthesis

Total chemical synthesis can be used to make glycopeptides containing structurally defined glycoforms. Production of larger glycoproteins has been far more challenging. Efforts by multiple groups has led to efficient synthetic methods like the native chemical ligation and expressed protein ligation that have enabled improved yields of large polypeptide ligations to form complex proteins[64-66]. Such efforts have resulted in the synthesis of hormones and other biologically relevant proteins like granulocyte macrophage colony stimulating factor and EPO[64,67]. Apart from the protein fragment ligations, another key aspect in glycoprotein synthesis is the assembly of oligosaccharide structures to be attached to the peptide or protein chains. Since multiples linkages are possible between the participating monosaccharide units in a glycoform, several steps of protection and deprotection are required to achieve regio- and stereo- selectivity of the glycosidic linkages[66]. Thus, while the total synthesis of structurally defined glycoproteins can be very rewarding, the path to achieving it involves challenging, tedious procedures that require optimization specific to proteins.

1.4.3 Chemoenzymatic methods

Chemoenzymatic methods, as the name suggests, make use of both chemical and enzymatic tools for the synthesis of homogeneous glycopeptides and glycoproteins. Enzymes dictate the specificity of linkages between the participating sugars and ensure

accurate regioselectivity and configuration at the anomeric position. *In vitro* chemoenzymatic methods utilize various enzymes like glycosyltransferases, glycosidases and glycosynthases for glycopeptide and glycoprotein remodeling[39,68]. Glycosyltransferases, that are responsible for oligosaccharide production in biological systems, transfer single monosaccharides at a time in contrast to endoglycosidases that can transfer full-length glycans. Though glycosyltransferases have been helpful in the synthesis of several significant glycoconjugates, their current application is limited in scope for several reasons. They have fixed substrate specificities and can act on only specific types of N-glycans. A large set of enzymes is required to be able to perform glycan modification of the entire range of glycoforms that can be present on naturally produced glycoproteins. Also, most of the natural glycosyltransferases are membrane-bound proteins and recombinant expression of such enzymes in soluble form can be a challenging task. Additionally, they utilize nucleotide sugar donors for the transfer reactions which prove to be very expensive for large scale chemoenzymatic synthesis[39,68]. The discovery and characterization of efficient glycosidases and endo-β-N-acetylglucosaminidases (ENGases) that can perform transglycosylation has revolutionized *in vitro* chemoenzymatic synthesis of glycoproteins and oligosaccharides[39,68,69]. A summary of these enzymes is discussed in the following sections.

1.4.3.1 Glycoside hydrolases

Glycoside hydrolases or glycosidases are enzymes that hydrolyze glycosidic bonds. The catalytic mechanism of these enzymes has been studied and utilized to generate

mutants for improved efficiencies. Glycosidases perform catalysis of their respective substrates by either retaining or inverting mechanisms. Most of these enzymes have two key residues (a nucleophile and a general acid/base) in their catalytic sites that aid in catalysis[39,68]. Figure 1.4A shows the catalytic mechanism of retaining β-glycosidases - a double displacement mechanism. The first step is initiated by the general acid-base by protonating the glycosidic oxygen and activating the glycosidic bond to be hydrolyzed. This promotes nucleophilic attack by the second key residue in the catalytic site to form an enzyme-substrate complex. The enzyme-substrate covalent linkage is then broken by an activated water molecule by the action of the general acid-base. If there is an appropriate acceptor substrate in place of water, the outcome is transglycosylation instead of hydrolysis. The endoglycosidases (ENGases) in some cases may use a third mechanism called the substrate-assisted mechanism or neighboring group participation (Figure 1.4B). ENGases are enzymes that cleave between the two GlcNAc residues in the core structure of N-glycans. Here, the 2-acetamido group of the second GlcNAc acts as a nucleophile to form an oxazolinium ion intermediate. This intermediate formation appears to be promoted by a key residue in the catalytic site in ENGases from the GH18 and GH85 families, which does not directly participate in the catalysis. The intermediate formation is followed by activation of a water molecule by the general acid/base, as in the case of retaining glycosidases, leading to hydrolysis. Again, when the water molecules are replaced by acceptor substrates, transglycosylation occurs instead of hydrolysis[39,68,69]. As compared to glycosylation by glycosyltransferases, glycosidases are more promiscuous in their substrate specificities. Also, the enzymes are easily accessible. But there are limitations

in terms of the transglycosylation efficiency - the product can undergo hydrolysis through the formation of oxazolinium ion intermediate. To circumvent the issues associated with product hydrolysis, mutants of glycosidases were designed following outstanding research efforts by several groups including the Wang group.

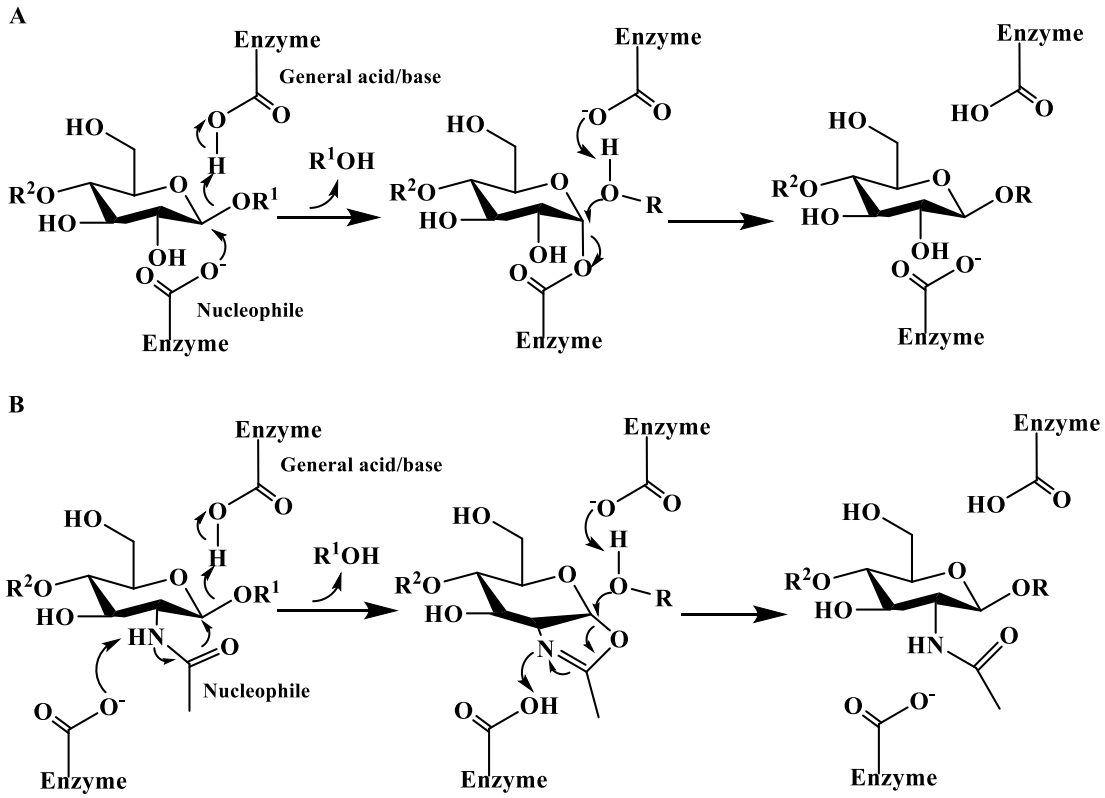


Figure 1.4. Catalytic mechanism of A) Retaining β-glycosidase and B) ENGase. R = H results in hydrolysis, R = acceptor results in transglycosylation

1.4.3.2 Glycosynthases

Glycosynthases are novel enzymes that are generated by the mutation of the nucleophilic residue in the respective glycosidases[70,71]. For instance, in case of the mutant generated from a β-retaining glycosidase, use of an activated donor like an α-

glycosyl fluoride mimics the intermediate in the retaining mechanism. The general acid-base assists in glycosidic bond formation to generate glycosyl product in the presence of an acceptor (Figure 1.5A). Due to a mutated nucleophile, usually an alanine mutation as shown in the figure, there is no product hydrolysis[39]. Several glycosynthases have been successfully designed using this concept[72,73]. On the other hand, glycoligases are designed by mutating the general acid-base in the catalytic site. Wang and coworkers have used this concept to design a fucoligase AlfC E274A that can perform direct core fucosylation of glycopeptides and glycoproteins including therapeutic antibodies[74].

Similarly, the inherent transglycosylation property of ENGases has been utilized to design glycosynthases of these enzymes. As mentioned previously, ENGases function by a substrate-assisted mechanism and hydrolyze the β-1,4-glycosidic linkage between two GlcNAc residues with the 2-acetamido group in the second GlcNAc acting as the nucleophile. Thus, generating glycosynthase mutants of ENGases involves mutation of residues that promote the oxazolinium ion intermediate formation[39,75], as shown in Figure 1.5B. Endoglycosynthases can perform transglycosylation in the presence of activated sugar oxazoline donors that mimic the transition-state intermediate to transfer the hydrolyzed glycans to appropriate acceptors.

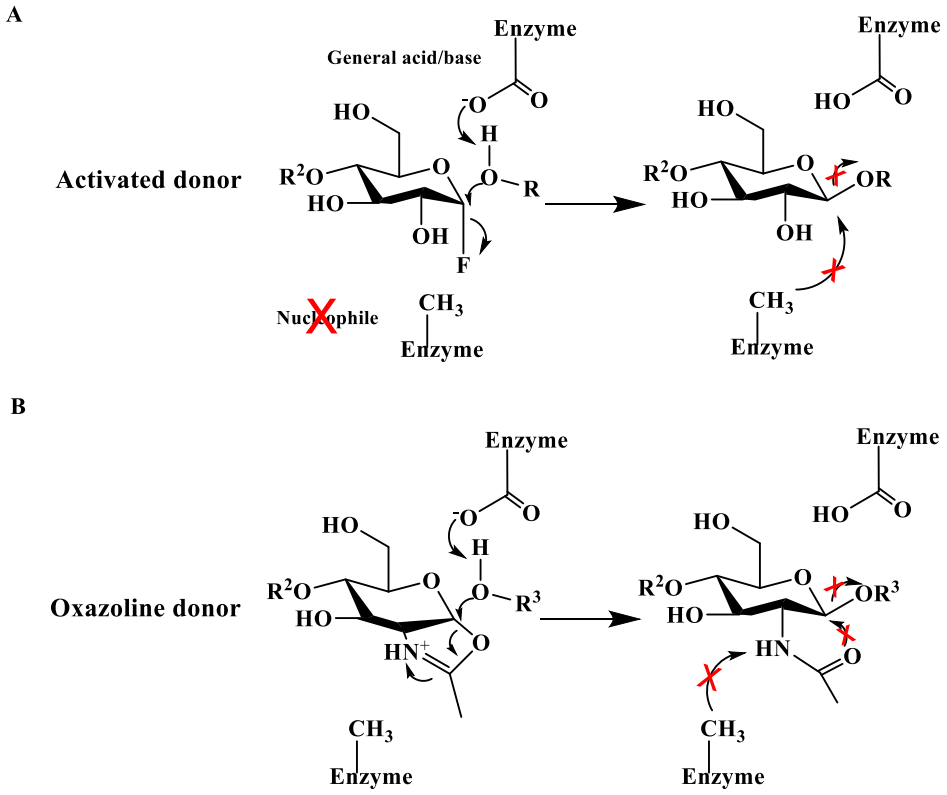


Figure 1.5. Catalytic mechanism of A) Retaining β-glycosynthase and B) ENGase-based glycosynthase. Under typical transglycosylation conditions, R^3OH= GlcNAc-acceptor

The chemical synthesis of various activated glycan oxazolines and the discovery of endoglycosidases with distinct substrate specificities has enabled efficient chemoenzymatic synthesis of glycoconjugates[76,77]. The ENGase-based glycosynthases have enabled convergent glycan remodeling approaches using wild-type ENGases for deglycosylation of the substrate to be remodeled, followed by transglycosylation to transfer desired full-length glycans *en bloc*, as shown in Figure 1.6[75,78,79]. Highly efficient glycosynthases have been designed that can transfer high mannose, hybrid, and complex type glycans to oligosaccharides, glycopeptides, and glycoprotein

substrates using corresponding sugar oxazolines. Some of the initial ENGase-based glycosynthases to be designed were from bacterial and fungal sources namely Endo-M from *Mucor hiemalis*[80,81], Endo-A from *Arthrobacter protophormiae*[82] and Endo-D from *Streptococcus pneumoniae*[83]. Each of these enzymes have distinct substrate specificities with defined acceptor substrates. For example, the Endo-A mutants can transfer high mannose glycans and Endo-M mutants can transfer high mannose as well as complex type glycans to GlcNAc-acceptors, but not when the acceptors are core-fucosylated[84]. To generate a library of enzymes that can modify a range of substrates with different glycoforms, the glycosynthase strategy has been applied to multiple other ENGases including Endo-S[85], Endo-S2[86], Endo-F3[87] and Endo-CC[84]. This repertoire of enzymes allows chemoenzymatic synthesis and remodeling of glycopeptides and glycoproteins with various glycoforms including bi- and tri- antennary glycans. However, continuous efforts are required to identify new enzymes with unique substrate specificities and higher efficiencies to fill the gaps in current methods.

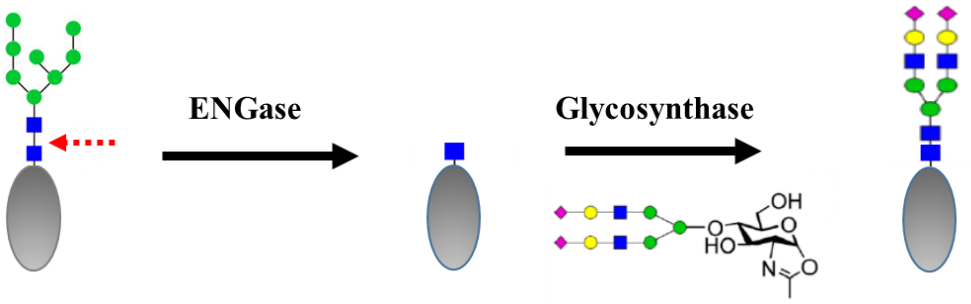


Figure 1.6. Chemoenzymatic remodeling of glycoproteins using ENGase-based enzymes. A heterogeneous glycoprotein can be first deglycosylated using an ENGase WT followed by transglycosylation using a glycosynthase to transfer desired glycoform from an oxazoline donor.

Chapter 2: Production of homogeneously glycosylated human interferon alpha by *in vivo* ApNGT-catalyzed glucosylation in *E. coli* and *in vitro* chemoenzymatic transglycosylation

The work reported in this chapter has been published in Bioorganic and Medicinal Chemistry journal [114]. My contributions to the study include planning and execution of experiments and analysis of results. I worked in collaboration with Dr. Qiang Yang, Wang lab, UMD for experiment planning and troubleshooting.

2.1 Introduction

The aim of this study was to use a convergent approach to generate homogeneous glycoproteins. As highlighted in Section 1.4, the key approaches used to control glycan heterogeneity are biosynthetic pathway perturbation and chemical and enzymatic tools for defined glycan synthesis. We explored a combined methodology to produce desired glycoproteins in a biological system with a sugar handle which can be further elaborated *in vitro* with full-length glycans using the ENGase-based glycosynthase technology. For this purpose, we decided to introduce a glycosylation machinery into *E. coli* for producing eukaryotic proteins with a monosaccharide tag for further modification. Several studies have been performed in the past to design a robust glycosylation system in bacterial cells, as will be discussed in this section. Though eukaryotic cells are attractive hosts for expressing glycoproteins, there are some

significant advantages to establishing a glycosylation pathway in bacterial expression systems. To name a few, the low cost of production, short culture duration, and easy handling and genetic manipulation of *E. coli* make it a model organism for protein expression[88,89]. Indeed, most therapeutic non-glycosylated proteins are expressed using *E. coli* and yeast expression systems. However, *E. coli* expression of heterologous glycoproteins requires an endogenous protein glycosylation machinery. Toward this end, several research groups have attempted to design a glycosylation system in *E. coli*. In collaboration with Aebi group, we have demonstrated a combined method that involves the functional transfer of a glycoengineered *Campylobacter jejuni* glycosylation machinery into *E. coli* to produce *N*-glycosylated eukaryotic proteins[58]. While this design allows *N*-glycosylation of the protein, the resulting glycoprotein contains an Asn-linked $(GalNAc)_5GlcNAc$-glycan, which is not a eukaryotic *N*-glycan. Further *in vitro* trimming of the outer GalNAc moieties by a bacterial α-*N*-acetylgalactosaminidase is required to generate the GlcNAc-protein, which then serves as the key intermediate for an *endo*-β-*N*-acetylglucosaminidase (ENGase)-catalyzed glycosylation to afford a eukaryotic glycoprotein. Moreover, it has been observed that expression of a heterologous eukaryotic glycoprotein such as the CH2 domain of immunoglobulin G (IgG) and a single-chain antibody F8 results in a low glycosylation efficiency (5–10% yield), which represents a bottleneck of the method[58]. In another study, DeLisa and co-workers have reported a bottom-up glycoengineering method for producing eukaryotic *N*-glycoproteins in *E. coli*[60]. In this approach, the genes responsible for constructing the *N*-glycan core ($Man_3GlcNAc_2$) are engineered into *E. coli* to build up the *N*-glycan glycolipid precursor. Then, PglB, the key bacterial

oligosaccharyltransferase from *C. jejuni*, is co-expressed to produce a recombinant glycosylated protein carrying a $Man_3GlcNAc_2$ *N*-glycan. However, the glycosylation yield is very low (*ca.* 1%), probably due to the low efficiency of the PglB-catalyzed transfer of a eukaryotic *N*-glycan from the corresponding glycolipid precursor[60]. Further study to improve the glycosylation efficiency of this method is required.

The discovery of a cytoplasmic *N*-glycosyltransferase in *Actinobacillus pleuropneumoniae* (ApNGT) that can transfer glucose (Glc) and galactose from a sugar donor to peptides and proteins at NX(S/T) *N*-glycosylation site represents an alternate avenue to protein glycosylation[90,91]. Several groups have demonstrated that ApNGT can transfer glucose from UDP-Glc to polypeptides and proteins carrying the consensus NX(S/T) *N*-glycosylation sequence (where X is any natural amino acid except proline). We have previously reported that *in vitro* transfer of glucose to peptides by ApNGT, coupled with subsequent sugar chain elongation provides an efficient way to synthesize glycopeptides carrying a full-length *N*-glycan[92]. Peng George Wang and co-workers have designed an engineered ApNGT (Q469A) that displayed a relaxed substrate specificity and higher efficiency of *in vitro* glucosylation with peptides and a bacterial protein containing multiple *N*-glycosylation sites[93]. In a later study, they have employed ApNGT Q469A, which can accept UDP-glucosamine (GlcN) as a donor substrate, to generate GlcN-peptides. Acetylation of the GlcN-peptides by an acetyltransferase resulted in GlcNAc-peptides, which were then extended with complex-type *N*-glycans by EndoM-N175Q to produce glycopeptides[94]. More recently, Jewett, MrKsich, and co-workers have explored the peptide acceptor sequence preference of a panel of NGT variants, which provides an exciting opportunity for

sequential glycosylation of a protein to make glycoproteins with multiple glycosylation sites[95,96]. For *N*-glucosylation in *E. coli,* Aebi and co-workers have reported that co-expression of ApNGT in *E. coli* is able to transfer glucose to bacterial adhesins[97]. In addition, *N*-glucosylation of heterologous proteins such as human EPO was detected in *E. coli* when ApNGT is co-expressed, although the human EPO forms inclusion bodies in *E. coli* and the extent of glucosylation is not determined[97]. More recently, Aebi and co-workers have shown that co-expression of ApNGT and three other glycosyltransferases in *E. coli* cytoplasm is able to generate a polysialylated protein[59]. While these studies have examined the application of ApNGT for *in vivo* glucosylation and explored further glycan modifications, attachment of natural eukaryotic *N*-glycans to these *E. coli*-expressed Glc-containing proteins has not been attempted so far. A closely related effort of transferring eukaryotic *N*-glycan to *O*-linked GlcNAc in proteins has been studied by Wang and co-workers[98]. *O*-GlcNAc transferase (OGT) was used to generate *O*-GlcNAc containing bovine protein in *E. coli* and a glycosynthase EndoM-N175Q was tested for *in vitro* transfer of sialylated complex type glycan, resulting in a 30% yield. In the present study, we report a case study of *in vivo* glucosylation of a eukaryotic protein by ApNGT in *E. coli* coupled with *in vitro* chemoenzymatic sugar chain elongation to produce a heterologous glycoprotein carrying a sialylated *N*-glycan. Human interferon α-2b (hIFNα-2b, referred to as IFNα hereafter) was used as a model eukaryotic protein and the overall experimental design is shown in Figure 2.1. Upon successful *in vivo* glucosylation with ApNGT in *E. coli*, EndoCC-N180H glycosynthase was found to efficiently transfer a sialylated *N*-glycan to IFNα-Glc. The purified sialylated IFNα displayed the expected anti-proliferative

activity in a cell proliferation test and a significantly enhanced stability towards protease degradation.

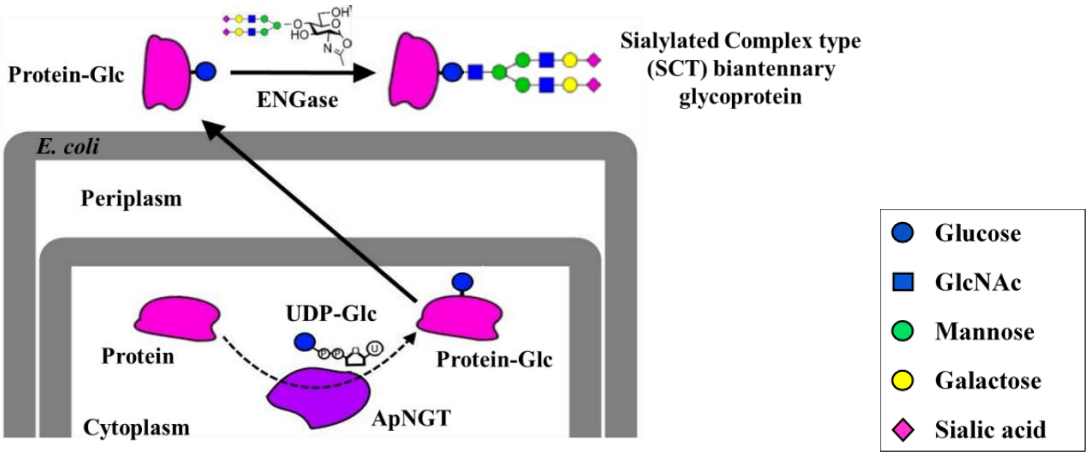


Figure 2.1. Schematic representation of *in vivo* ApNGT-catalyzed glucosylation and *in vitro* chemoenzymatic transglycosylation of eukaryotic proteins in *E. coli*. The N-glycosylated protein contains an innermost glucose instead of GlcNAc present in natural N-glycans.

2.2 Results and Discussion

2.2.1 Introduction of N-glycosylation sites into IFNα

IFNα is a 19.2 kDa α-helical protein consisting of 165 amino acids with a single *O*-glycosylation site at Thr-106. It is a cytokine that possesses anti-viral and anti-proliferative properties and is commonly used to treat Hepatitis C and several forms of cancers[99]. *N*-glycosylation of a modified IFNα variant with four engineered *N*-glycan sites that was expressed in a mammalian system shows up to a 25-fold increase in its serum half-life[100]. Based on this study, we chose the Pro-4 (P4) and/or Gln-158 (Q158) as the position to introduce an *N*-glycosylation site to test ApNGT-catalyzed *in vivo*

glucosylation in *E. coli*. P4 lies in the flexible region close to the protein *N*-terminus and Q158 is at the end of an α-helix close to the C-terminus. Each of the chosen sites were modified to attain an ANAT sequence at the $X_{-1}NX_{+1}X_{+2}$ positions to achieve high efficiency of glucosylation by ApNGT, as suggested by the previous report[97]. The structure of IFNα and the glycoengineered sequences are shown in Figure 2.2.

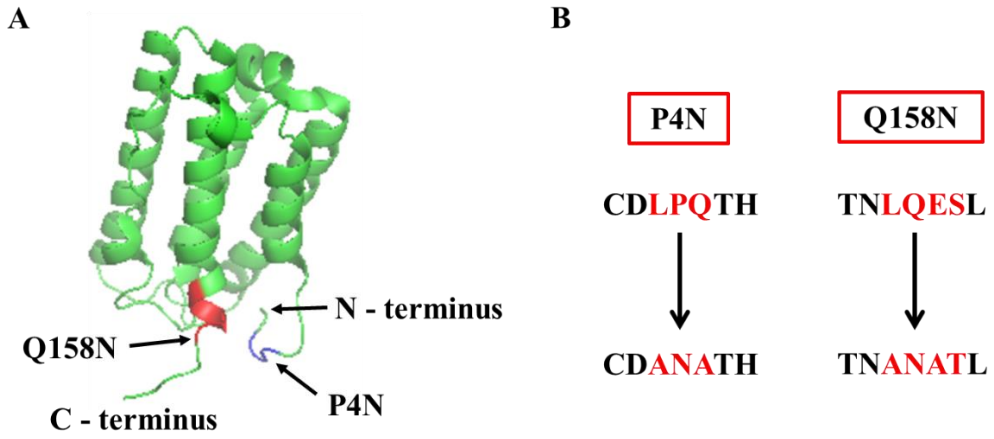


Figure 2.2. The structure of IFNα and the introduced *N*-glycosylation site. A) The structure of human IFNα (PDB code 2hym). The two positions chosen for site-directed mutagenesis are highlighted. Mutagenesis at P4 is shown in blue and at Q158 in red. B) The engineered *N*-glycosylation sites. The modified amino acids in both the mutants are highlighted in red

2.2.2 *In vivo* glucosylation of GST-IFNα in *E. coli*

In the initial study, IFNα wild-type (WT) and mutant proteins were expressed as fusion proteins with a Glutathione-S-Transferase (GST) tag at the *N*-terminus to enable soluble expression in *E. coli* cytoplasm[101]. The fusion protein design allows for release of IFNα by thrombin cleavage at an inserted thrombin protease site. The mutants, GST-

IFNα-P4N and GST-IFNα-Q158N were each co-expressed with ApNGT for *in vivo* glucosylation in *E. coli* with a protein yield of 50–70 mg/L. Mass spectrometric analyses of the purified fusion proteins showed completely glucosylated GST-IFNα-P4N (ESI-MS: Calculated, M = 45638 Da; found, M = 45639 Da after deconvolution) (Figure 2.3A) and about 80% glucosylated GST-IFNα-Q158N (ESI-MS: Calculated, M = 45620 Da; found, M = 45621 after deconvolution) (Figure 2.3B). SDS-PAGE profiles of the purified fusion proteins are shown in Figure 2.3C and 2.3D. Contrary to the efficient glucosylation of GST-IFNα fusion proteins, protease release of the GST tag posed some unexpected challenges. We found that the two mutants and the wild-type GST-IFNα, expressed in *E. coli* BL21(DE3) strain, could not be efficiently cleaved by the protease thrombin. The reaction, when carried out with an excess of enzyme showed a maximum of 50% cleavage with overnight incubation (Fig.2.3E) and increased only up to 70% when prolonged to incubation for two days. Intensive optimization of the reaction conditions did not show any further improvement in the GST tag removal. A possible explanation of this observation is improper folding of the fusion protein making the cleavage site hard to access by thrombin. In common *E. coli* strains, the cytoplasm hosts a reductive environment which may prevent the formation of disulfide bonds in proteins, thus causing misfolding of heterologous proteins. To test such a possibility, we expressed GST-IFNα in an Origami 2 (DE3) strain, which carries mutations in thioredoxin reductase (trxB) and glutathione reductase (gor) genes to provide an oxidative environment in the cytoplasm. Encouragingly, GST-IFNα expressed in Origami showed complete cleavage of GST tag by thrombin with overnight incubation under previously used reaction conditions (Figure 2.3F). This

observation prompted us to perform refolding of the soluble GST-IFNα expressed in *E. coli* BL21(DE3)21 and to examine thrombin cleavage of the refolded protein. Under the previously described reaction conditions, *ca.* 70% release of the GST tag was observed with overnight incubation indicating a moderate improvement in the protease cleavage. To test *in vivo* glucosylation in Origami 2 (DE3) strain, GST-IFNα-P4N and GST-IFNα-Q158N were co-expressed with ApNGT. Surprisingly, neither of the proteins showed any glucose transfer (data not shown). The expression levels of the enzyme and the fusion protein were observed to be lower in the Origami strain as compared to the BL21 cells. However, the expression of ApNGT was comparable or slightly more than that of the fusion protein thus potentially eliminating the lack of enzyme to be the reason behind lack of glucose transfer. To probe the lack of glucosylation of the proteins further, the two fusion proteins were also co-expressed with ApNGT Q469A, as the mutant enzyme has shown a broader substrate specificity and better glucosylation efficiency in a previous report[93]. Unfortunately, neither of the fusion proteins showed any glucosylation. While this result was unexpected, it provides an important implication in the mechanism of glucosylation by ApNGT in *E. coli*. It has been suggested that ApNGT requires substrate proteins in a not fully folded state for *in vivo* glucosylation to occur[97]. The lack of *in vivo* glucosylation in Origami 2 (DE3) strain compared to the highly efficient glucose transfer in BL21(DE3) supports this hypothesis and suggests co-translational glucosylation by ApNGT of the nascent recombinant protein (IFNα).

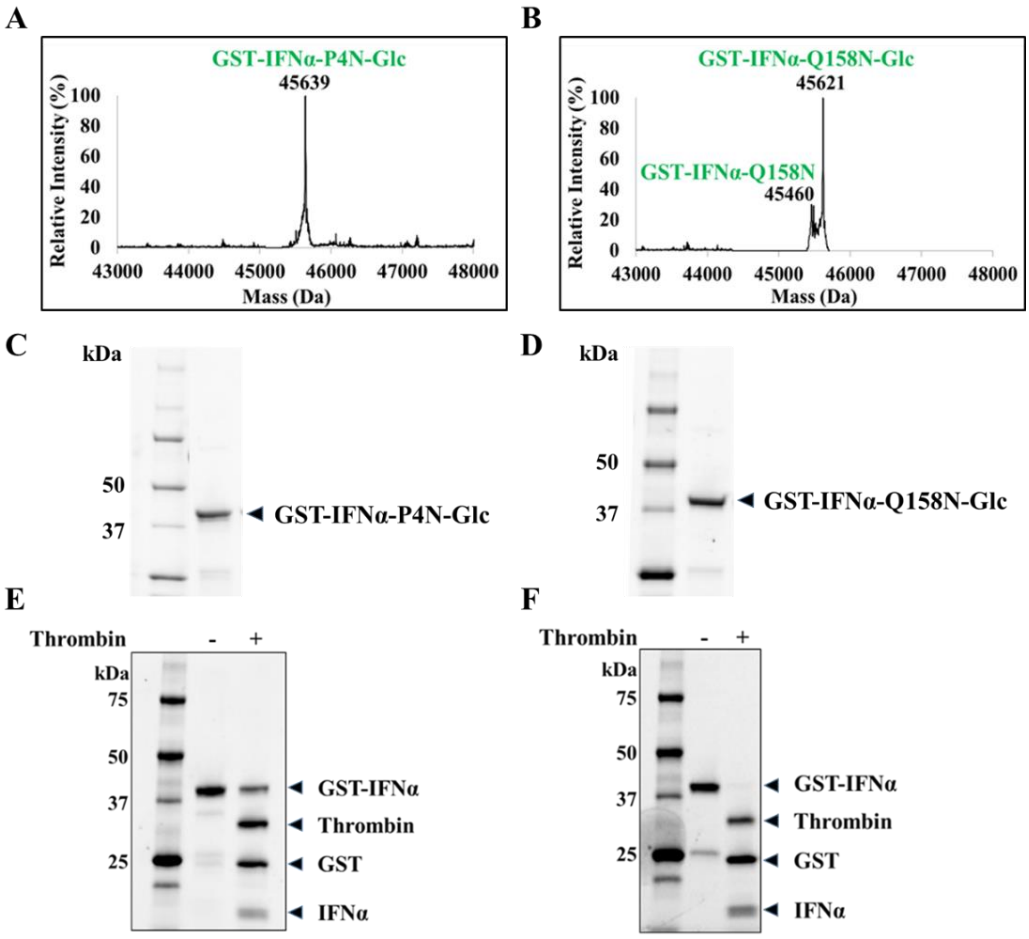


Figure 2.3. Co-expression of GST-IFNα mutants with ApNGT in *E. coli*. A) Deconvoluted LC-MS profiles of GST-IFNα-P4N. Calculated mass with two disulfide bonds = 45476 Da, calculated mass post glucosylation = 45638 Da. $(M+H)^+$ peak observed at 45639 Da. B) Deconvoluted LC-MS profiles of GST-IFNα-Q158N. Calculated mass with two disulfide bonds = 45458 Da, calculated mass post glucosylation = 45620 Da. $(M+H)^+$ peaks observed at 45460 and 45621 Da. SDS-PAGE showing C) GST-IFNα-P4N and D) GST-IFNα-Q158N, co-expressed with ApNGT in *E. coli* BL21(DE3) cytoplasm. SDS-PAGE showing overnight thrombin cleavage of GST-IFNα-WT expressed in E) *E. coli* BL21(DE3) F) *E. coli* Origami 2 (DE3).

In parallel to the examination of GST fusion protein, we also tested *in vivo* glucosylation of IFNα without any fusion tag. This work was focused on one of the mutants, Q158N. IFNα-Q158N was cloned into pET28a vector under T7 promoter for

expression in *E. coli* BL21(DE3). We purified IFNα from inclusion bodies (IBs) using established refolding protocols[102-104]. A final yield of *ca.* 120 mg/L was achieved. SDS-PAGE profile of the purified refolded protein is shown in Fig. 2.4A. Interestingly, co-expression of ApNGT and IFNα-Q158N showed a relatively low efficiency of glucosylation in IBs with a maximum glucose transfer of 35% under optimized culture conditions (Figure 2.4C). The low glucose transfer efficiency could be due to insufficient exposure of IFNα expressed in IBs to the cytoplasmic ApNGT. Finally, we attempted to express soluble glucosylated 6xHis-IFNα-Q158N in *E. coli* by optimizing the conditions. First, we tried to induce protein production at a relatively low temperature (16–25 ◦C), which resulted in soluble expression of 6xHis-IFNα-Q158N with a yield of *ca.* 20 mg/L. SDS-PAGE profile of the purified protein is shown in Fig. 2.4B. Second, to reach optimum glucose transfer, we executed stepwise induction of ApNGT and IFNα-Q158N. ApNGT was induced by arabinose 2 h prior to the induction of IFNα by IPTG. In this way, a maximum glucosylation of 55% was achieved at 16 ◦C (Figure 2.4D). In comparison, co-induction of ApNGT and IFNα-Q158N at 16 ◦C resulted in 40% glucosylation. Comparing these results to the 80% glucosylation of GST-IFNα- Q158N (containing the same acceptor sequence), we speculate that the local conformation of the glycosylation site and the protein folding kinetics likely affect the efficiency of glucose transfer by ApNGT.

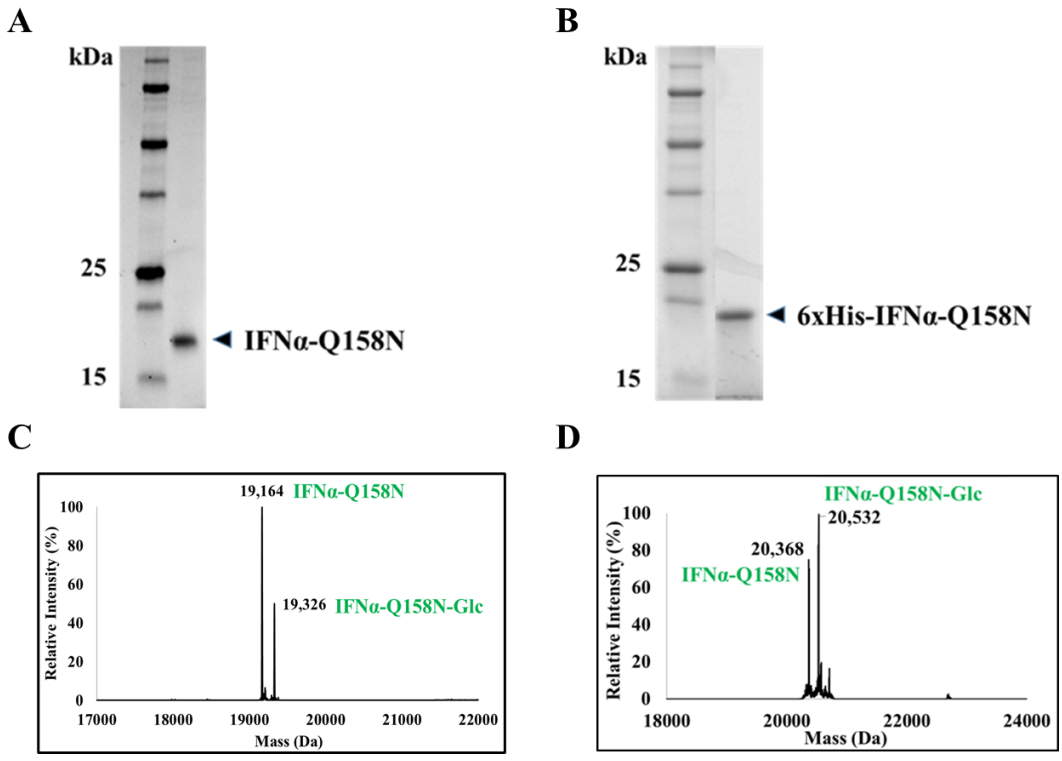


Figure 2.4. Alternate strategies of IFNα-Q158N expression in *E. coli*. SDS-PAGE showing A) IFNα-Q158N refolded and purified from *E. coli* BL21(DE3) IBs and B) 6xHis-IFNα-Q158N purified from *E. coli* BL21(DE3) cytoplasm. C) Deconvoluted LC-MS profile of IFNα-Q158N co-expressed with ApNGT in *E. coli* IBs. Calculated mass with two disulfide bonds = 19165 Da, calculated mass post-glucosylation = 19327 Da. D) Deconvoluted LC-MS profile of 6xHis-IFNα-Q158N co-expressed with ApNGT in *E. coli*. Calculated mass of 6xHis-IFNα-Q158N with two disulfide bonds = 20370 Da. Calculated mass post- glucosylation = 20532 Da. Additional higher mass peak observed in the profile is an unidentified non-glucose adduct.

2.2.3 *In vitro* chemoenzymatic transglycosylation of glucosylated IFNα- Q158N

Several glycosynthases, which are endoglycosidase mutants that demonstrate reduced product hydrolysis activity but can use sugar oxazolines as an activated donor substrate for transglycosylation, can transfer *N*-glycan *en bloc* to GlcNAc- or glucose-linked peptides and proteins to form homogeneous glycoproteins[39]. Such *in vitro* glycoengineering of glycoproteins relies strongly on the substrate specificities of the

parent endoglycosidases. Glycosynthases that can transfer different glycoforms to peptide and protein substrates have been developed. Previously we have shown that EndoA or EndoM mutants could transfer *N*-glycans to glucose-Asn-linked polypeptides that are generated by glucosylation with ApNGT[92]. However, very few studies have reported the enzymatic *N*-glycan transfer to intact glucose-proteins. To examine the *in vitro* enzymatic *N*-glycan transfer, we chose the glucosylated 6xHis-IFNα-Q158N recombinant protein as the acceptor substrate, which contains *ca.* 55% glucosylation (Figure 2.4D). An initial attempt to use EndoM-N175A and EndoM-N175Q mutants to transfer a sialylated *N*-glycan to the glucosylated IFNα failed to provide glycosylation product (data not shown). This result is consistent with previous reports that EndoM mutants (N175Q or N175A) can efficiently transfer complex type *N*-glycans from the corresponding glycan oxazolines to GlcNAc- or Glc-polypeptides, but the EndoM mutants are much less efficient to act on folded and intact GlcNAc-proteins[105,106]. Finally, an efficient *in vitro N*-glycan transfer was achieved by using the N180H mutant of EndoCC, an endoglycosidase from *Coprinopsis cinerea* that has been previously reported to be able to efficiently transfer a complex type *N*-glycan to deglycosylated ribonuclease B (GlcNAc-RNase B)[84]. Hence, we tested EndoCC-N180H for the transfer of SCT-glycan to the partially glucosylated 6xHis-IFNα-Q158N. Gratifyingly, the enzyme exhibited high efficiency in transferring SCT-glycan to the glucosylated IFNα substrate, with a 90% yield (Figure 2.5A). Thus, an *E. coli* co-expression of IFNα and ApNGT coupled with *in vitro* sugar chain extension with the EndoCC-N180H provides a feasible approach to produce glycosylated IFNα, which may be applicable to other therapeutic proteins. After the successful transglycosylation,

we attempted to purify the transglycosylation product with a SiaRich™ Pan-Specific Sialic Acid Affinity Column (Lectenz® Bio). The lectin column is expected to bind to α-2,3/6/8 linked sialic acid conjugates which are then eluted out using high salt concentration. The 6xHis-IFNα-Q158N mixture was applied to the lectin column, then the column was eluted with a gradient of salt solution and two protein fractions were obtained. However, to our surprise, the 6xHis-IFNα-Q158N-SCT was eluted first at a lower salt concentration, followed by the non-glycosylated 6xHis-IFNα-Q158N protein. The results suggest that the sialylated IFNα has only weak binding affinity with the lectin. To verify that the lectin binds to sialylated glycoprotein efficiently, we tested the lectin column with fetuin as a positive control, which contains multiple sialylated *N*- and *O*-glycans. As predicted, fetuin bound tightly to the affinity column and was eluted out with only high concentration of salt. The unexpected behavior of the sialylated and non-glycosylated IFNα suggests that the non-glycosylated IFNα exerts a relatively strong non-specific protein–protein interaction with the lectin while the large sialylated glycan attached might cancel out the non-specific protein–protein interaction, although at the same time the lectin-glycan interaction also plays a role in the affinity. The ESI-MS as a homogeneously glycosylated protein with an expected mass spectrometry profile (Calculated, M = 22534 Da; found, M = 22536 Da; deconvoluted data) (Figure 2.5B).

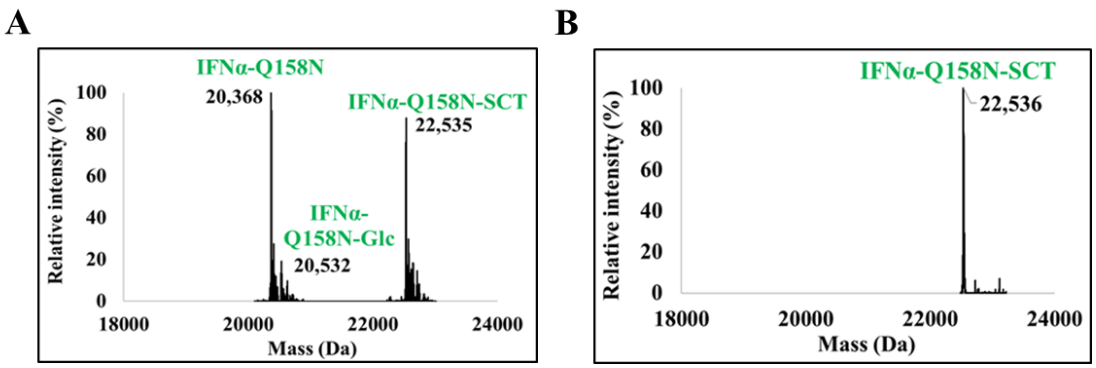


Figure 2.5. LC-MS analysis of *in vitro* transglycosylation of glucosylated IFNα-Q158N and enrichment of the transglycosylation product. Deconvoluted LC-MS profile of A) *In vitro* chemoenzymatic transglycosylation of IFNα-Q158N by EndoCC-N180H. Calculated mass of IFNα-Q158N-SCT, M = 22534 Da. 90% transglycosylation was observed. It should be noted that the protein is only 55% glucosylated resulting in a total yield of 50% sialylated protein. B) The IFNα-Q158N-SCT enriched using Lectenz affinity chromatography. The IFNα-Q158N referred to in this figure is the His-tagged protein

2.2.4. Biological activity and proteolytic stability of glycoengineered IFNα

Glycosylation of proteins is known to partially interfere with receptor binding and reduce the *in vitro* biological activity of proteins[107]. IFNα is a therapeutic protein that exhibits anti-viral and anti-proliferative properties and is used in the treatment of cancers and hepatitis C[99]. To test the effect of sialylated glycosylation on the biological activity of IFNα-Q158N, we conducted an *in vitro* anti-proliferative assay using Burkitt's lymphoma Daudi cells[100,108]. The growth of Daudi cells as a function of IFNα concentration was studied in microtiter plates using a colorimetric method. The anti-proliferative activities of the glucosylated and sialylated 6xHis-IFNα-Q158N were examined alongside a reference standard. Results from duplicate experiments showed the activity of 55% glucosylated 6xHis-IFNα-Q158N to be comparable with that of the

reference standard. Attaching a sialylated glycan to the protein showed a 3-fold decrease in its biological activity (Figure 2.6A, Table 2.1). This result is consistent with a previous report of IFNα-2b displaying a 10-fold reduction in biological activity upon glycosylation at four engineered sites[100]. The loss in activity was however, compensated by a 25-fold increase in serum half-life of the glycosylated IFNα-2b in mice[100]. Another study reports a 4-fold reduction in *in vitro* activity of a recombinant human EPO when attached with two additional *N*-glycan chains[109]. But the increase in sialic acid content shows improved serum half-life and an overall 4-fold increase in the *in vivo* potency of the modified glycoprotein[109]. In addition to the effect of sialylation on serum half-life of proteins, glycosylation also confers enhanced resistance to protease cleavage, mainly through steric hindrance[110]. Glycosylation of IFNγ at Asn-25 has been reported to drastically improve its proteolytic stability towards several proteases in *in vitro* assays[110]. The glycosylated IFNγ retains significant anti-viral activity post protease treatment, unlike the non-glycosylated and N25Q variants of IFNγ. Similarly, *O*-glycosylation of an engineered insulin at Thr-27 of B-chain doubled its proteolytic stability towards *in vitro* chymotrypsin treatment[111]. To examine the effect of glycosylation on the proteolytic stability of IFNα-Q158N, we studied the half-life of IFNα towards trypsin digestion which has the same cleavage specificity as plasmin found in blood. The decay of the proteins[112] was monitored using LC-MS with the reference standard used as an internal standard for quantitation[111]. Indeed, the sialylated IFNα (6xHis-IFNα-Q158N-SCT) showed at least a 7-fold higher resistance to tryptic digestion in comparison to the glucosylated protein (Figure 2.6B, Table 2.2). This increased stability is advantageous to IFNα as cytokines are known to be

susceptible to proteolytic degradation[113]. In summary, these results are in good agreement with the previous reports and indicate that while sialylated glycosylation resulted in some decrease in *in vitro* anti-proliferative activity, it significantly enhanced the proteolytic stability which could lead to increased overall biological activity *in vivo.*

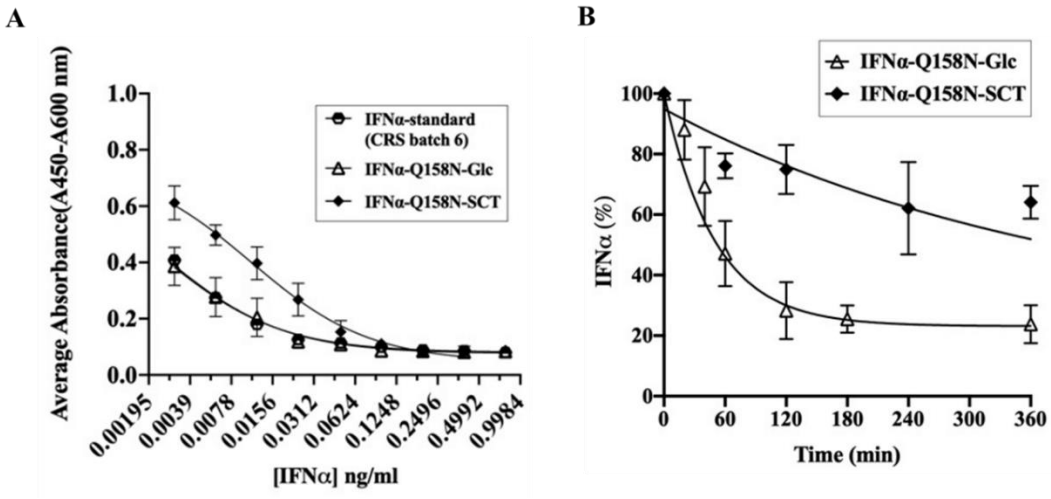


Figure 2.6. Effects of sialylated glycosylation on the anti-proliferative activity and proteolytic stability. A) *In vitro* anti-proliferative activity of glucosylated and sialylated IFNα-Q158N in comparison with IFNα reference standard. Data from duplicate experiments were fitted to obtain the dose response curves and half maximal inhibitory concentrations (IC50) of the IFNα proteins. B) *In vitro* trypsin digestion of glucosylated and sialylated IFNα-Q158N. Protein decay data from duplicate experiments were fitted using a first order decay equation to obtain the half-life of the proteins towards trypsin digestion.

Table 2.1. Summary of anti-proliferative dose-response curves of IFNα

Parameter	IFNα-standard (CRS batch 6)	IFNα-Q158N-Glc	IFNα-Q158N-SCT
IC50 (ng/ml)	0.0025	0.0026	0.0109
R squared	0.9932	0.9800	0.9835

Table 2.2. Summary of first order decay of IFNα upon trypsin digestion

Parameter	IFNα-Q158N-Glc	IFNα-Q158N-SCT
Half-life (min)	34.45	280.90
Rate constant, k (min^{-1})	0.0201	0.0025
R squared	0.9986	0.9048

2.3 Conclusion

An *E. coli*-based method to generate a homogeneously glycosylated eukaryotic protein is described. We used ApNGT co-expression to produce glucosylated human IFNα in *E. coli* which was then efficiently glycosylated *in vitro* using EndoCC-N180H to generate homogeneously sialylated IFNα. Biological activity of the purified product was confirmed alongside an IFNα reference standard using an anti-proliferative assay. This study demonstrates the feasibility of utilizing the ApNGT-based glycosylation system in combination with enzymatic sugar chain elongation to generate fully glycosylated heterologous eukaryotic proteins.

2.4 Experimental procedures

2.4.1. Gene constructs and plasmids

The coding region of wild type IFNα (Uniprot: P01563, IFNA2_HUMAN, Δ1-23) was modified for a P4N mutation as shown in Figure 2.2B. Codon-optimized synthetic gene of IFNα-P4N with an *N*-terminal GST tag, cloned into pGEX-4T-1 between *Bam*HI and *Not*I sites was procured from GenScript. pGEX-GST-IFNα-WT and pGEX-GST-IFNα-Q158N constructs were generated from pGEX-GST-IFNα-P4N using Q5® Site-Directed Mutagenesis Kit (New England BioLabs) and respective primers. IFNα-Q158N gene was amplified from pGEX-GST-IFNα-Q158N using PfuUltra II Fusion High-fidelity DNA Polymerase (Agilent) and cloned into pET28a(+) vector (Novagen) between *Nco*I and *Bam*HI sites. *N*-terminal 6xHis-tag was introduced into pET28a-IFNα-Q158N using the Q5 mutagenesis kit. Wild type ApNGT sequence (Uniprot:A3N2T3, NGT_ACTP2) was cloned into pET45b vector using *Kpn*I/*Xho*I sites, into pET28a vector using *Nde*I/*Xho*I sites and into pBAD33.1 vector using *Nde*I/*Sal*I sites. pBAD33.1 was a gift from Christian Raetz (Addgene plasmid # 36267; http://n2t.net/addgene:36267; RRID: Addgene_36267)[115].

2.4.2. Expression and purification of GST- IFNα in *E. coli*

Overnight grown pre-cultures of *E. coli* BL21(DE3) or Origami 2 (DE3) cells co-transformed with plasmids pGEX-GST-IFNα (respective mutants or WT) and pET28a-ApNGT were diluted 50 times into Terrific Broth (TB) medium supplemented with 100 μg/ml carbenicillin and 50 μg/ml kanamycin and cultured at 37 ◦C, 250 RPM. The broth was induced with 0.5 mM IPTG when OD600 reached 0.6–0.8, and continued to culture

overnight at 22 ◦C. Cells were harvested by centrifugation at 4000 RPM for 15 min at 4 ◦C. The cell pellet was resuspended in PBS buffer, pH-7.4. Cells were lysed by sonication (12 cycles of 30 s burst/30 s cooling). The cell lysate was centrifuged at 20,000 RPM for 12 min at 4 ◦C. The supernatant containing soluble proteins was filtered and loaded onto a Reduced Glutathione resin column (Pierce™ Glutathione Agarose, Cat. # 16100) and allowed to bind overnight at 4 ◦C. The flowthrough was collected, and the column was washed with 10 column volumes of PBS buffer. Bound proteins were eluted using 33 mM glutathione in 50 mM Tris-HCl, 200 mM NaCl at pH – 8.0. The eluted proteins were dialyzed against PBS buffer, pH-7.4 at 4 ◦C.

2.4.3. Removal of GST tag from GST-IFNα-WT by thrombin

5 µg of GST-IFNα-WT was treated with 2.5 U of Thrombin (Human Plasma, High Activity from Millipore Inc., Cat. # 605195) in PBS buffer, pH – 7.4. The reaction mixture (20 µl) was incubated at 22 ◦C overnight. Samples were analyzed by SDS-PAGE.

2.4.4. Expression and purification of His-tagged IFNα-Q158N in *E. coli*

Overnight grown pre-cultures of *E. coli* BL21(DE3) cells co-transformed with plasmids pET28a-IFNα-Q158N and pBAD33.1- ApNGT were diluted 50 times into TB medium supplemented with 50 µg/ml kanamycin and 25 µg/ml chloramphenicol and cultured at 37 ◦C, 250 RPM. ApNGT expression was induced with 0.4% L-arabinose at OD_{600} ~ 0.2 two hours prior to IFNα induction with 0.5 mM IPTG at OD_{600} ~ 0.6–0.8 and cultured overnight at 16–25 ◦C after another addition of 0.4% L-

arabinose. Protein purification was done as mentioned in Section 4.2. The supernatant containing His-tagged soluble proteins was filtered and purified using HisTrap™ columns (Cytiva, Cat. # 17524701), following manufacturer's protocol. The eluted proteins were dialyzed against 50 mM Tris-HCl, 2.5% sucrose, pH-8.5 at 4 ◦C.

2.4.5. Refolding of IFNα-Q158N expressed in *E. coli* inclusion bodies

Protein expression was done as outlined in Section 4.2 using constructs pET28a-IFNα-Q158N and pET45b-ApNGT. At the end of culture, cells were harvested by centrifugation at 4000 RPM for 15 min at 4 ◦C. The cell pellet was resuspended in 50 mM Tris-HCl, pH-8.5 containing 2 mM Ethylenediaminetetraacetic acid (EDTA) and 1 mM phenylmethylsulfonyl fluoride (PMSF). Cells were lysed by sonication (15 cycles of 30 s burst/30 s cooling). The cell lysate was centrifuged at 10,000*g* for 20 min at 4 ◦C. The cell pellet containing IBs was washed in 50 mM Tris-HCl, pH-8.0 containing 1% Triton X-100 by stirring for 1 h at room temperature. Washed IBs were centrifuged at 10,000*g* for 10 min at 4 ◦C. The IBs were washed again in 50 mM Tris-HCl, pH-8.0 containing 5 M urea, 20 mM DTT by stirring at room temperature for 1 h. This was followed by centrifugation at 15,000*g* for 30 min at 4 ◦C. The washed IBs were then solubilized in 6 M guanidine hydrochloride, 50 mM Tris-HCl, pH-8.0 containing 100 mM β-mercaptoethanol (BME) at room temperature for 1 h. The suspension was centrifuged at 10,000*g* for 30 min at 4 ◦C. Refolding of solubilized IBs was done by dialysis of the supernatant against refolding buffer containing 50 mM Tris-HCl, pH-8.0, 4 M guanidine hydrochloride, 0.2 mM EDTA, 1 mM reduced and 0.1 mM oxidized glutathiones at 4 ◦C. Refolding buffers with decreasing amount of

guanidine hydrochloride in steps of 2 M were used for subsequent dialyses. The final buffer contained 50 mM Tris-HCl, 2.5% sucrose, 0.2 mM EDTA at pH-8.5.

2.4.6. SDS-PAGE sample preparation and analysis

Approximately 1 μg of purified protein diluted in water was mixed with 5 μl of 2X SDS loading dye containing BME to make a final volume of 10 μl. The samples were denatured at 95 ◦C for 5–10 min and cooled to 4 ◦C before loading. 6 μl of protein ladder (Precision Plus Protein Standards, Bio-Rad, Cat. # 1610363) was loaded for reference. Precast polyacrylamide stain-free gels were used for electrophoresis (Mini-PROTEAN™ TGX Stain-Free™ Protein Gels, Bio-Rad, Cat. # 4568086, 4568106). The gels were run in 1X Tris-Glycine-SDS buffer at 220 V for 35 min. Gel imaging was done using Gel Doc™ EZ Imager (Bio- Rad).

2.4.7. Mass spectrometric analysis of IFNα proteins

IFNα samples were prepared in 0.1% formic acid (FA) and injected into a QExactivePlus OrbiTrap mass spectrometer in line with Ultimate 3000 HPLC system (ThermoFisher) for liquid chromatography-electrospray ionization mass spectrometry (LC-MS) analysis. Protein chromatography was performed on a C4 column (Xbridge™ BEH300, 3.5 μm, 2.1 × 50 mm, 300 Å, Waters) at a flow rate of 0.4 ml/min and a column temperature of 23 ◦C. Water containing 0.1% FA and acetonitrile containing 0.1% FA were used as solvents A and B respectively. Column equilibration was done with 5% B for 2 min, followed by sample injection and a 6 min linear gradient of 5–

95% B for protein separation. Column washing was done with 95% B for 2 min, followed by column equilibration with 5% B for 2 min. MS scan range was set at 400 to 3000 *m*/*z* with a scan speed of 0.9 Hz and a resolution of 140,000. MagTran software (Amgen) was used to deconvolute the *m*/*z* profiles.

2.4.8. Expression and purification of EndoCC-N180H in *E. coli*

Expression of EndoCC-N180H was done using a protocol adapted from previously reported literature[84]. Briefly, *E. coli* BL21(DE3) cells transformed with plasmid pET41b-EndoCC-N180H were cultured at 37 ◦C, 250 RPM overnight. The pre-culture was diluted 20 times into 1L Luria Bertani (LB) broth supplemented with 50 μg/ml kanamycin and cultured overnight at 30 ◦C, 250 RPM. Cells were harvested by centrifugation at 4000 RPM for 15 min at 4 ◦C. The cell pellet was resuspended in 20 ml bacterial cell lysis buffer (GoldBio ®) supplemented with 45 μg of lysozyme and 40 U of DNase and incubated at 37 ◦C for 1 h. The cell lysate was centrifuged at 20,000 RPM for 30 min at 4 ◦C. The supernatant containing soluble 6xHis-tagged enzyme was filtered and purified using a HisTrap^{TM} column. The eluted protein was dialyzed against PBS, pH-7.4 at 4 ◦C and stored at -80 ◦C.

2.4.9. Transglycosylation of glucosylated 6xHis-IFNα-Q158N by EndoCC-N180H

A mixture of the glucosylated 6xHis-IFNα-Q158N (400 μg, 0.8 mg/ ml, about 55% is glucosylated), SCT-oxazoline[78] (500 μg), and the N180H mutant of EndoCC (5 μg, 0.01 μg/μl) was incubated in a Tris-HCl buffer (20 mM, pH 7.5) at 30 ◦C. The

enzymatic transglycosylation was monitored by LC-MS analysis. After 2 h, an additional portion of SCT-oxazoline was added to push the reaction to completion. Then the sialylated IFNα-Q158N product was isolated by lectin affinity chromatography.

2.4.10. Enrichment of sialylated 6xHis-IFNα-Q158N

SiaRich™ Pan-Specific Sialic Acid Affinity Column (Lectenz® Bio) was used for enrichment of 6xHis-IFNα-Q158N-SCT using the recommended protocol, on an AKTA FPLC System (Cytiva). The column was equilibrated with at least 5 column volumes of binding buffer (10 mM EPPS, 10 mM NaCl, pH-7.5). The transglycosylation reaction mixture of 6xHis-IFNα-Q158N was dialyzed against the binding buffer to get rid of excessive SCT-oxazoline and salts. The dialyzed solution was diluted to 1 ml with binding buffer before loading on to the column at 0.2–0.5 ml/ min. The column was then washed with 10 column volumes of binding buffer at 1 ml/min. Elution was performed with 5 ml of 10 mM EPPS, 500 mM NaCl buffer, pH 7.5. 0.5 ml fractions were collected, and samples were analyzed using LC-MS. Column regeneration (10 mM EPPS, 1 M NaCl buffer, pH-7.5) and storage conditions were followed as recommended in the product datasheet provided by the manufacturer.

2.4.11. *In vitro* anti-proliferative assay

Daudi cells (ATCC® CCL-213™) were maintained in suspension in RPMI-1640 medium (ATCC) containing 10% fetal bovine serum (FBS), 100 U/ml penicillin and

100 μg/ml streptomycin in T-75 flasks (Sarstedt). Cells were sub-cultured every 2–3 days, when the cell density reached 1–2 × 10^6 cells/ml, with a seeding density of 2–4 × 10^5 cells/ ml. The anti-proliferative assay was conducted following previously reported protocols[100,108]. Interferon alfa-2b CRS batch 6 (EDQM Catalogue code: I0320301, Millipore Sigma) was used as a reference standard for comparison of the biological activity of engineered IFNα. Serial dilutions of IFNα (reference standard and test samples) were made in a 96-well plate to attain a final concentration range of 0.779–0.003 ng/ml (10 μl volume in each well). Each of the concentration was tested in duplicates. To each well, 200 μl of Daudi cells at a density of 5 × 10^4 cells/ml of culture medium was added. Wells with only 200 μl cells (no IFNα) and only 200 μl culture medium (no IFNα or cells) were run as controls. The well plate was incubated at 37 ◦C, 5% CO_2 for 96 h. 10 μl of Cell Counting Kit-8 (Sigma) was added to each well and the plate was incubated at 37 ◦C, 5% CO_2 for 3–4 h. The absorbance of formazan released by viable cells by reducing WST-8 [2-(2-methoxy-4-nitrophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H-tetrazolium, monosodium salt] dye was measured at 450 nm using a spectrophotometer. Dose-response curves of [IFNα] and absorbance (A450nm – A600nm) were plotted in each case using GraphPad Prism to obtain IC50 values.

2.4.12. *In vitro* trypsin digestion of 6xHis-IFNα-Q158N

10 μg of IFNα-Q158N-Glc or IFNα-Q158N-SCT was mixed with trypsin at a final enzyme concentration of 0.1 μg/ml in 50 mM Tris-HCl, pH-8.0 and incubated at 37 ◦C. 1 μg sample was drawn from the reaction mixture at 0, 20, 40, 60, 120, 180, 360 min

for IFNα-Q158N-Glc and 0, 60, 120, 240, 360, 840, 1680 min for IFNα-Q158N-SCT. Drawn samples were added to 0.1% formic acid containing 0.1 μg Interferon alfa-2b CRS batch 6 for LC-MS analysis. The decrease in relative intensity of IFNα-Q158N ($100 * I_t/I_0$) with respect to the internal standard was monitored over time, where I_t is the relative intensity of IFNα-Q158N at time t and I_0 is the relative intensity of IFNα-Q158N at t = 0. Half-lives of the proteins towards trypsin digestion were obtained by fitting the data to first-order decay as reported in the past[111,112] using GraphPad Prism. The first-order decay equation used for the fitting was $I_t = (I_0 - \text{Plateau})*\exp(-k*t) + \text{Plateau}$, where k is the rate constant and plateau represents the relative intensity of IFNα-Q158N at infinite time.

Chapter 3: Expression and substrate specificity analysis of α-L-fucosidases for modulating core fucosylation of proteins

The study reported in this chapter has been published in Bioorganic and Medicinal Chemistry journal[157]. My contributions to the study include planning and execution of experiments and analysis of results. Synthesis of core fucosylated substrates was done in collaboration with Dr. Chao Li, Dr. Guanghui Zong and Roushu Zhang, Wang lab, UMD.

3.1 Introduction

Core fucosylation, the attachment of an α-1,6- fucose moiety to the innermost *N*-acetyl glucosamine (GlcNAc) in *N*-glycans in mammalian systems, is one of the most frequent modifications on *N*-glycoproteins that can modulate the glycan conformations and regulates biological processes including cell adhesion, signal transduction, and development[116-118]. Further, the role of core fucose in tumor growth and progression has been demonstrated in several studies[34,119-121]. For example, gene knockout of FUT8, the α-1,6-fucosyltransferase that is responsible for core fucosylation in mammalian systems, in liver cancer cell lines and xenograft mouse models has shown reduced signaling through growth factor receptors and subsequent tumor suppression[122,123]. Additionally, core-fucosylated α-fetoprotein (AFP-L3) is upregulated in liver cancers and is an FDA approved biomarker in hepatocellular carcinoma[124]. In addition to

regulating biological functions, core fucosylation also presents structural implications. Presence of α-1,6 core fucose in glycoproteins impacts the conformations of *N*-glycans and thus modulates their affinity for respective lectins[117,125]. In the case of monoclonal antibodies, core fucosylation of IgG-Fc *N*-glycans at Asn-297 interferes with the favorable carbohydrate-carbohydrate interactions between the *N*-glycans from antibody Fc domain and Fc receptors, thus decreasing the affinity of antibody for FcγIIIa receptor[126]. Compelling data have shown that removal of core fucosylation in IgG-Fc *N*-glycans can significantly enhance FcγIIIa receptor binding (up to 50- fold), resulting in remarkably increased ADCC and enhanced therapeutic efficacy[18,44,127]. Given the structural effects and biological importance of core fucosylation, it is crucial to develop tools that can modulate the fucose content of glycoproteins and antibodies for structure–function relationship studies as well as for therapeutic applications. Many α-L-fucosidases have been reported but only a few have been confirmed to be able to act on the α-1,6-glycosidic linkage to the innermost GlcNAc residue (the core fucose)[74,128-135]. We have previously shown that a nonspecific α-L-fucosidase from bovine kidney could remove core fucose from Endo-S treated IgG antibody, but it has only low activity and requires a long incubation time to achieve complete defucosylation[131]. Later, we have found that the α-L-fucosidases from *Bacteroides fragilis* (BfFuc) and *Lactobacillus casei* (AlfC) could much more efficiently remove the core fucose from Fc-deglycosylated antibodies than the bovine kidney enzyme, although none of the enzymes show detectable defucosylation activity on intact IgG antibodies[74,132]. Wong and co-workers have independently reported that the two α-L-fucosidases (BfFuc and AlfC) could not act on intact glycoproteins including antibodies

either but hydrolyze the core fucose when the external *N*-glycans were removed[133]. On the other hand, FucA1, the human α-L-fucosidase, is responsible for the defucosylation of fucosylated substrates in lysosomes and its activity of defucosylation has been associated with tumor suppression[136-139]. Furthermore, an *in vitro* study with FucA1 has demonstrated that FucA1 could remove core fucose from free complex-type *N*-glycans[140]. However, the substrate specificity and substrate structure–activity relationships of FucA1 and the above-mentioned bacterial α-L-fucosidases still remain to be further characterized. One challenge in probing the substrate specificity of these α-L-fucosidases is the availability of structurally well-defined core-fucosylated *N*-glycans, glycopeptides, and glycoproteins due to the difficulties in isolation and synthesis of core-fucosylated oligosaccharides and glycoconjugates[140-142]. We describe in this paper the chemoenzymatic synthesis of an array of well-defined core-fucosylated oligosaccharides, glycopeptides, and glycoproteins including antibodies (Figure 3.1). These fucosylated compounds were used for a comparative study on the substrate specificity and activities of the human α-L-fucosidase (FucA1) and two bacterial α-L-fucosidases (AlfC and BfFuc). Our results showed that the three α-L-fucosidases had different substrate specificity and activities. The two bacterial enzymes were found to hydrolyze the core fucose only when the core fucose was exposed by endoglycosidase treatment to remove the external sugar residues of the *N*-glycans. In contrast, the human α-L-fucosidase, FucA1, was the only enzyme capable of removing core fucose from some intact full-length glycopeptides and glycoproteins. FucA1 also demonstrated a low but detectable activity to remove core fucose from intact antibodies. Furthermore, the nature of *N*-glycosylation on antibodies appeared to

influence the defucosylation activity of FucA1, which demonstrated higher hydrolytic activity on the high mannose glycoform than the complex-type glycoform of an antibody.

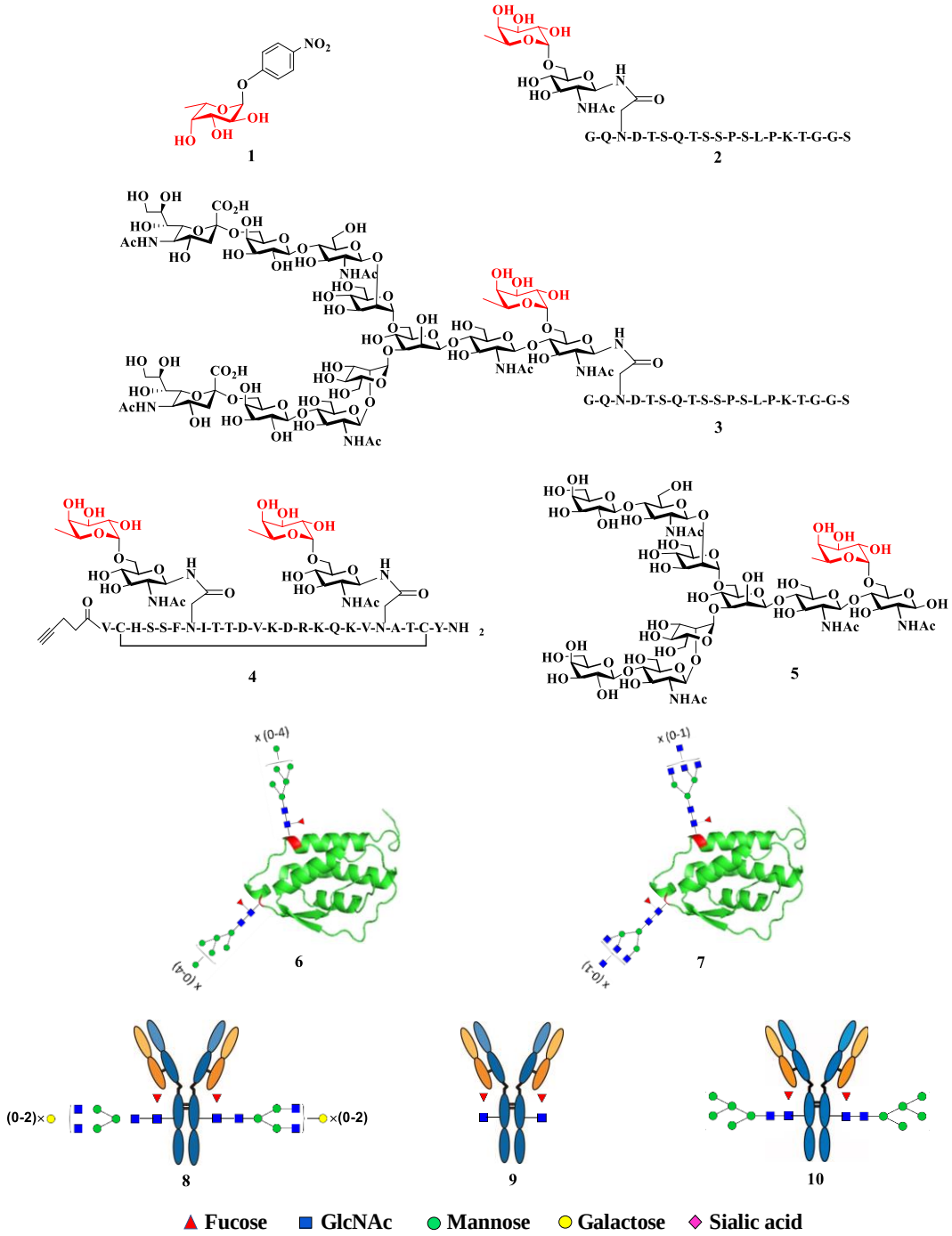


Figure 3.1. Substrates used for testing hydrolytic activity of α-fucosidases. *p*-nitrophenyl α-L-fucoside (**1**); CD52-GN-F: a 19-mer glycopeptide containing CD52 antigen, sortase A signal sequence and truncated *N*-glycan (**2**); CD52-SCT-F: CD52 glycopeptide containing core fucosylated biantennary sialylated complex type glycan (**3**); V1V2-GN2-F2: a 24-mer cyclic glycopeptide derived from the V1V2 region of HIV-1 gp120 glycoprotein (**4**); CT-F: core fucosylated complex-type free *N*-glycan (**5**); GM-CSF-HM-F: granulocyte macrophage colony stimulating factor containing fucosylated high mannose glycan (**6**); GM-CSF-WT: GM-CSF wild type containing a mixture of complex-type *N*-glycans (**7**); Rituximab: IgG1 antibody (**8**); RTX-GN-F: Deglycosylated Rituximab containing Fucα1,6GlcNAc (**9**); RTX-M5-F: Rituximab containing core fucosylated mannose-5 glycan (**10**)

3.2. Results and discussion

3.2.1. Synthesis of the fucosylated glycopeptides (**2**, **3** and **4**) and a fucosylated biantennary complex-type N-glycan (**5**)

A total of ten fucosylated substrates (**1**–**10**) (Figure 3.1) were applied to study the substrate specificity of the three enzymes. The synthesis of fucosylated glycopeptides (**2** and **3**) started with the preparation of the GlcNAc-peptide, CD52-GN (**11**), a 19-mer containing the CD52 antigen sequence and the sortase A signal sequence (LPKTGGS), followed by enzymatic core fucosylation and transglycosylation (Scheme 3.1). The inclusion of a sortase A signal sequence was for the purpose of further enzymatic site-specific conjugation for future applications[143],[144]. Solid-phase peptide synthesis (SPPS) on an automated peptide synthesizer following our previously published procedure[145] gave CD52-GN (**11**) in 70% yield after HPLC purification. Core fucosylation was achieved by using the fucoligase AlfC E274A that we have previously reported, which was able to use simple synthetic α-fucosyl fluoride as the donor substrate to attach an α1,6-linked fucose to a GlcNAc moiety in peptide[74]. The resulting core-fucosylated glycopeptide CD52-GN-F (**2**) was purified using preparative HPLC in 62% yield and characterized by LC-MS (Figure 3.2A,B). Transfer of a biantennary complex-type *N*-glycan to **2** was achieved by using an EndoF3-D165A mutant, which was specific for glycosylation of fucosylated GlcNAc acceptor substrate[87], giving the full-length glycopeptide (**3**) (Scheme 3.1). The glycopeptide (**3**) was purified by preparative HPLC in 42% yield and characterized by LC-MS (Figure 3.3A,B).

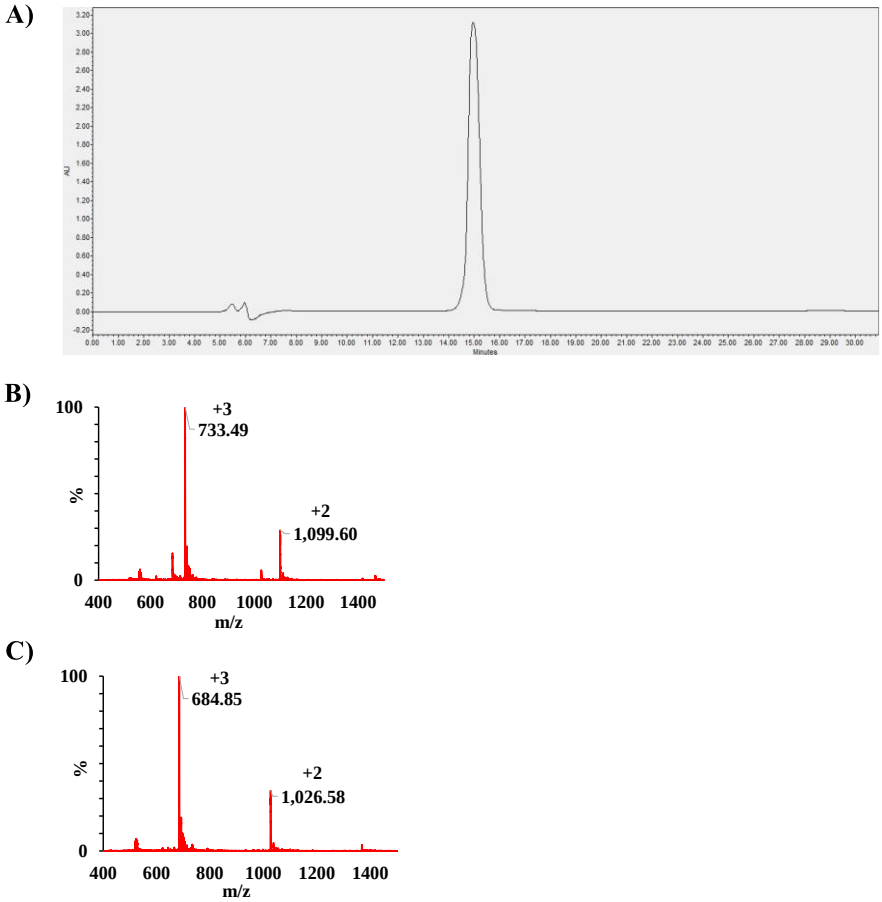
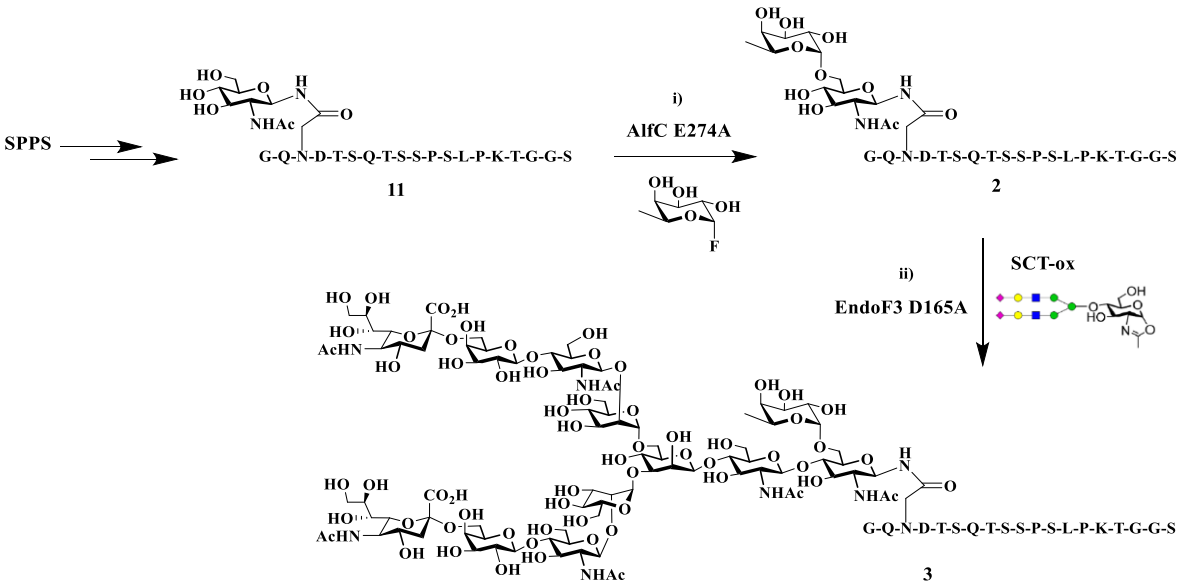


Figure 3.2. Synthesis and hydrolysis of CD52-GN-F (**2**). A) Preparative-HPLC on a RP C18 column. A linear gradient of 5 to 35% acetonitrile containing 0.1% TFA over 30 minutes at a flow rate of 12 ml/min. t_R = 14.959 minutes. B) ESI-MS of CD52-GN-F: calculated, *M* = 2197.0 Da; found (*m/z*), 733.49 $[M + 3H]^{3+}$, 1099.60 $[M+2H]^{2+}$. C) ESI-MS of the hydrolysis product, CD52-GN (**11**): calculated, *M* = 2051.4 Da; found (*m/z*): 684.85 $[M+3H]^{3+}$, 1026.58 $[M+2H]^{2+}$.

Scheme 3.1: Chemoenzymatic synthesis of core fucosylated glycopeptides



Reagents and conditions: (i) AlfC E274A (0.0014 mol equiv of **11**), α-fucosyl fluoride (6.36 mol equiv of **11**), PBS (pH-7.4), 37^{o}C, 2 h. (ii) Endo F3 D165A (0.0009 mol equiv of **2**)**,** SCT-ox (2 mol equiv of **2**), PBS (pH-7.4), 30^{o}C, 0.5 h.

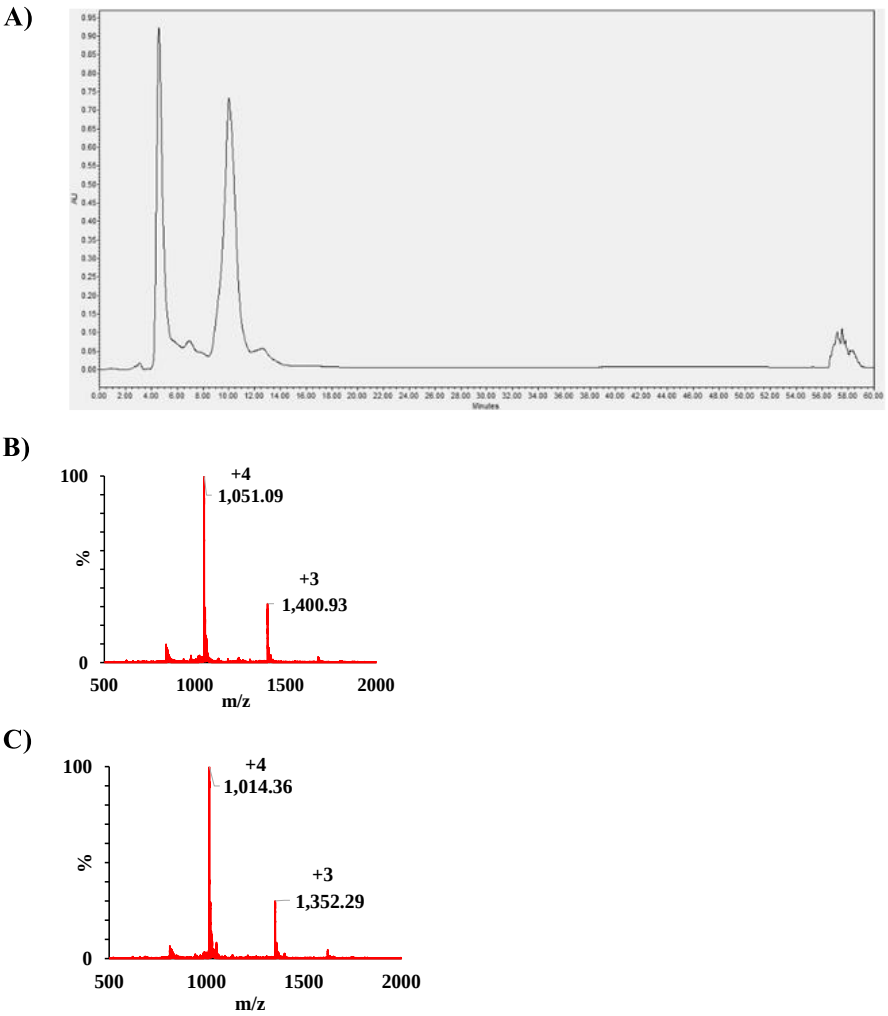
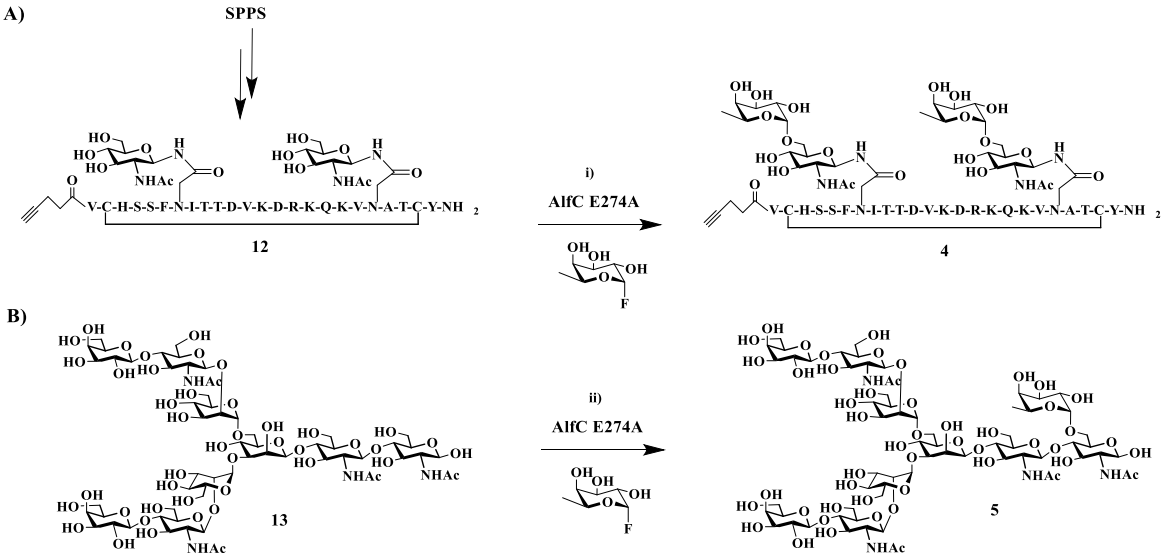


Figure 3.3. Synthesis and hydrolysis of CD52-SCT-F (**3**). A) Preparative-HPLC on a RP C18 column. A linear gradient of 5 to 25% acetonitrile containing 0.1% TFA over 40 minutes at a flow rate of 4 ml/min. t_R = 10.009 minutes. B) ESI-MS of CD52-SCT-F: calculated, *M* = 4199.1 Da; found (*m/z*), 1051.09 $[M + 4H]^{4+}$, 1400.93 $[M + 3H]^{3+}$. C) ESI-MS of CD52-SCT: calculated, *M* = 4053.1 Da; found (*m/z*), 1014.36 $[M + 4H]^{4+}$, 1352.29 $[M + 3H]^{3+}$.

To further assess the hydrolysis of core fucose from a glycopeptide containing more than one *N*-glycosylation site, we also synthesized fucosylated glycopeptide **4**, a 24-mer cyclic glycopeptide derived from the V1V2 region of envelope glycoprotein gp120 of HIV-1 strain ZM10938. First, the precursor GlcNAc-peptide, V1V2-GN2 (**12**), was synthesized using SPPS by modifying our previously described procedure[146]. The

purified glycopeptide (62% yield) was fucosylated by AlfC E274A fucoligase to provide fucosylated glycopeptide **4**, which was purified by preparative HPLC in 70% yield and its identity was confirmed by LC-MS analysis (Figure 3.4A,B). We also synthesized a fucosylated biantennary complex-type *N*-glycan (**5**) using the fucoligase AlfC E274A for direct core fucosylation (Scheme 3.2). But in this case, we found that the full-length *N*-glycan was a poor substrate, and a large amount of the enzyme was required to drive the reaction. The fucosylated product (**5**) was purified by preparative HPLC in 27% yield and its identity was confirmed by MALDI-TOF-MS analysis (Figure 3.5A).

Scheme 3.2: Chemoenzymatic synthesis of core fucosylated glycopeptide 4 and *N*-glycan 5



Reagents and conditions: (i) AlfC E274A (0.008 mol equiv of **12**), α-fucosyl fluoride (10 mol equiv of **12**), PBS (pH-7.4), 37°C. (ii) AlfC E274A (2 mg/ml), α-fucosyl fluoride (10 mol equiv of **13**), PBS (pH-7.4), 37°C, 3 h.

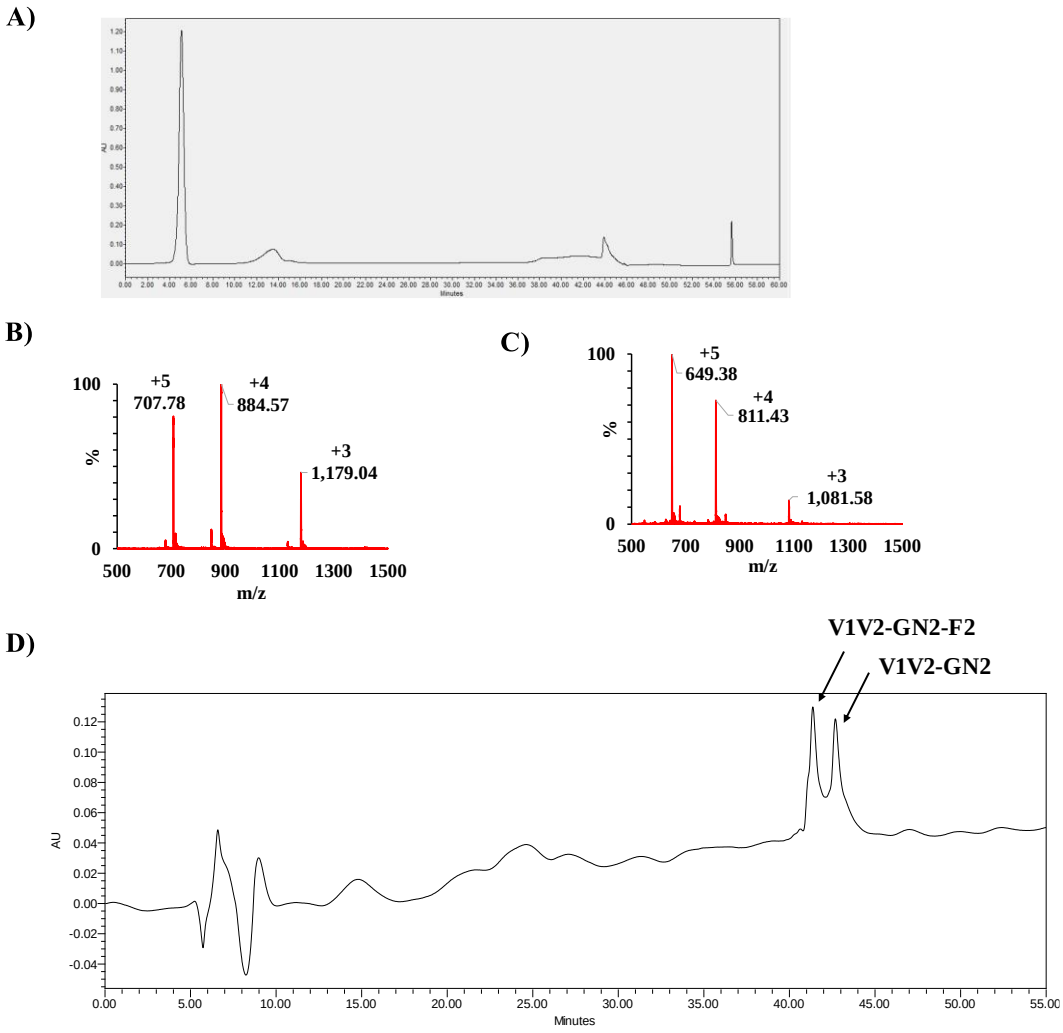
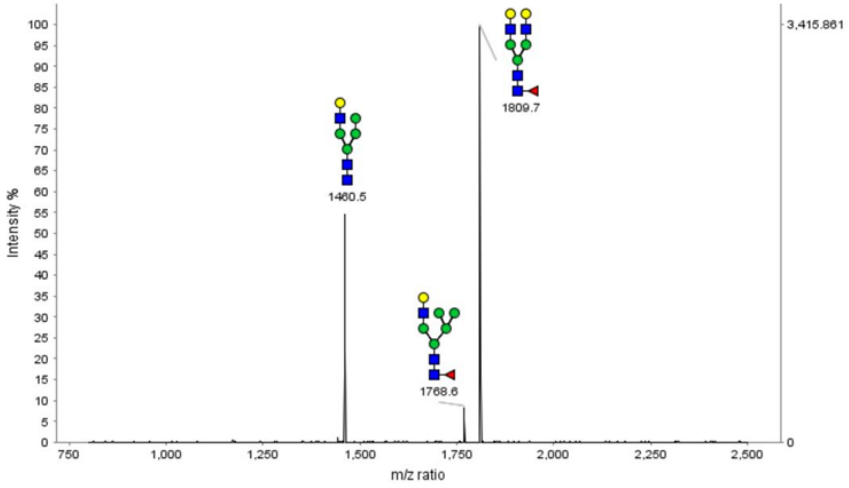


Figure 3.4. Synthesis and hydrolysis of V1V2-GN2-F2 (**4**). A) Preparative RP-HPLC of V1V2-GN2-F2, **4.** t_R = 10.117 minutes (Initial 5 min of the run was not captured on profile due to a technical error) B) ESI-MS spectrum of V1V2-GN2-F2: calculated, *M* = 3533.7 Da; found (*m/z)*, 707.78 $[M + 5H]^{5+}$, 884.57 $[M + 4H]^{4+}$, 1179.04 $[M + 3H]^{3+}$. C) ESI-MS spectrum of V1V2-GN2: calculated, *M* = 3241.7 Da; found (*m/z*), 649.38 $[M + 5H]^{5+}$, 811.43 $[M + 4H]^{4+}$, 1081.58 $[M+3H]^{3+}$. D) Analytical RP-HPLC of a 1:1 mixture of **4** and **12**. A linear gradient of 5 to 40% acetonitrile containing 0.1% TFA over 50 minutes at a flow rate of 0.5 ml/min. t_R: V1V2-GN2-F2 = 41.367 minutes, V1V2-GN2 = 42.686 minutes.

A)



B)

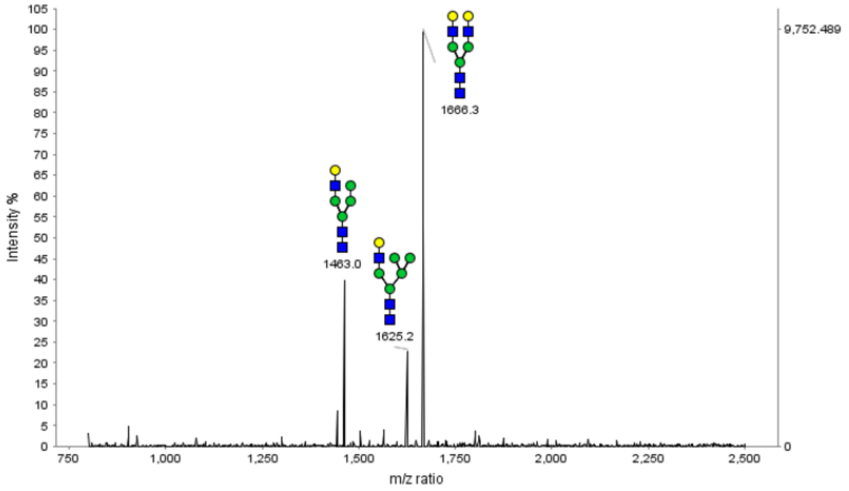


Figure 3.5. MALDI-TOF profile of (A) CT-F glycan (**5**). Calculated, M = 1787 Da; found (*m/z*), 1809.7 $[M + Na]^+$ and (B) the hydrolysis product, CT (**13**). Calculated, M = 1641 Da; found (*m/z*), 1666.3 $[M + Na]^+$

3.2.2. Preparation of the core-fucosylated glycoproteins (**6** and **7**) and antibodies (**9** and **10**)

To study the hydrolysis of core fucose in the context of intact glycoproteins, two types of glycoforms of the granulocyte macrophage colony stimulating factor (GM-CSF)

were produced. GM-CSF-HM-F (**6**) was expressed in HEK293T FUT8+ cells in the presence of kifunensine, an α1,2-mannosidase inhibitor[147]. The His-tagged protein was purified using immobilized metal affinity chromatography (IMAC). A smear of protein bands on SDS-PAGE indicated a heterogeneously glycosylated protein (Figure 3.6A). Glycan characterization was done by releasing *N*-glycan from the protein using PNGase F, followed by glycan enrichment and MALDI-TOF-MS analysis. A mixture of partially fucosylated high mannose (Man5-9) glycoforms was seen in the released *N*-glycan (Figure 3.6B). On the other hand, the complex-type glycoform, the His-tagged GM-CSF-WT (**7**), was expressed using the above method in the absence of kifunensine. Purification and characterization of the glycoprotein were done as explained above (Figure 3.7A). The released *N*-glycan showed a mixture of glycoforms with tri- and tetra-antennary complex-type *N*-glycans as the predominant forms (Figure 3.7B). For testing the activity of the α-L-fucosidases on intact antibodies, we chose Rituximab (**8**), a therapeutic monoclonal antibody as a model and prepared two variants of core-fucosylated glycoforms (**9,10**). First, Rituximab was treated with endoglycosidase EndoS2 to hydrolyze the Fc glycan, giving the truncated fucosylated glycoform, Fucα1,6GlcNAc- RTX (RTX-GN-F, **9**). Then, a high-mannose (Man5) *N*-glycan was transferred to **9** using the EndoS2-D184M mutant[86] to afford the Man5 glycoform (**10**) (Scheme 3.3). The identities of the antibody glycoforms (**8**–**10**) were confirmed by LC-MS analysis (Figure 3.8A, 3.9A, 3.10A).

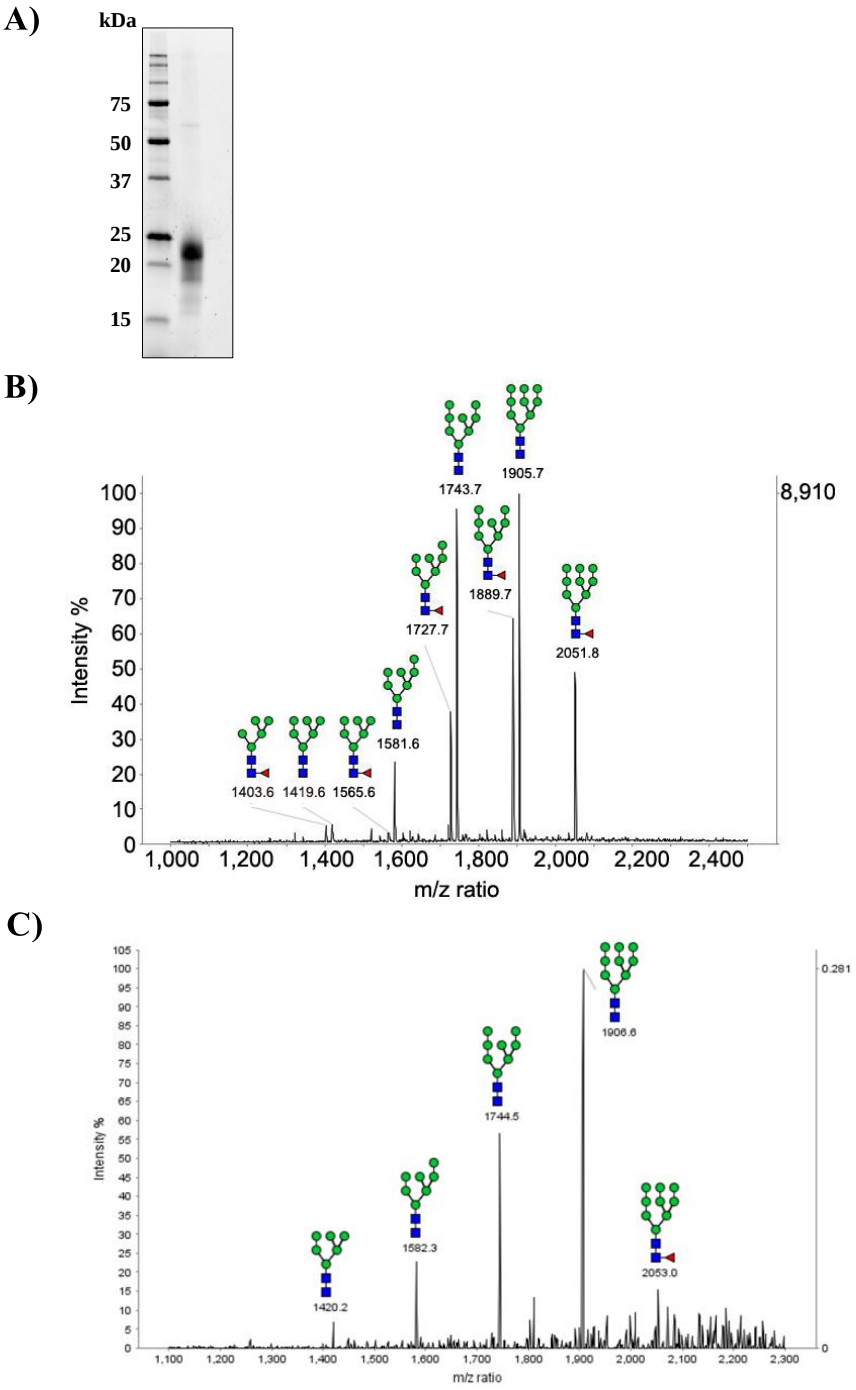


Figure 3.6. Expression and characterization of GM-CSF-HM-F (**6**). (A) SDS-PAGE of the purified protein showing a smear of bands (B) MALDI-TOF MS profile of *N*-glycan released from **6 (C)** MALDI-TOF MS profile of *N*-glycan released from the hydrolysis product, GM-CSF-HM

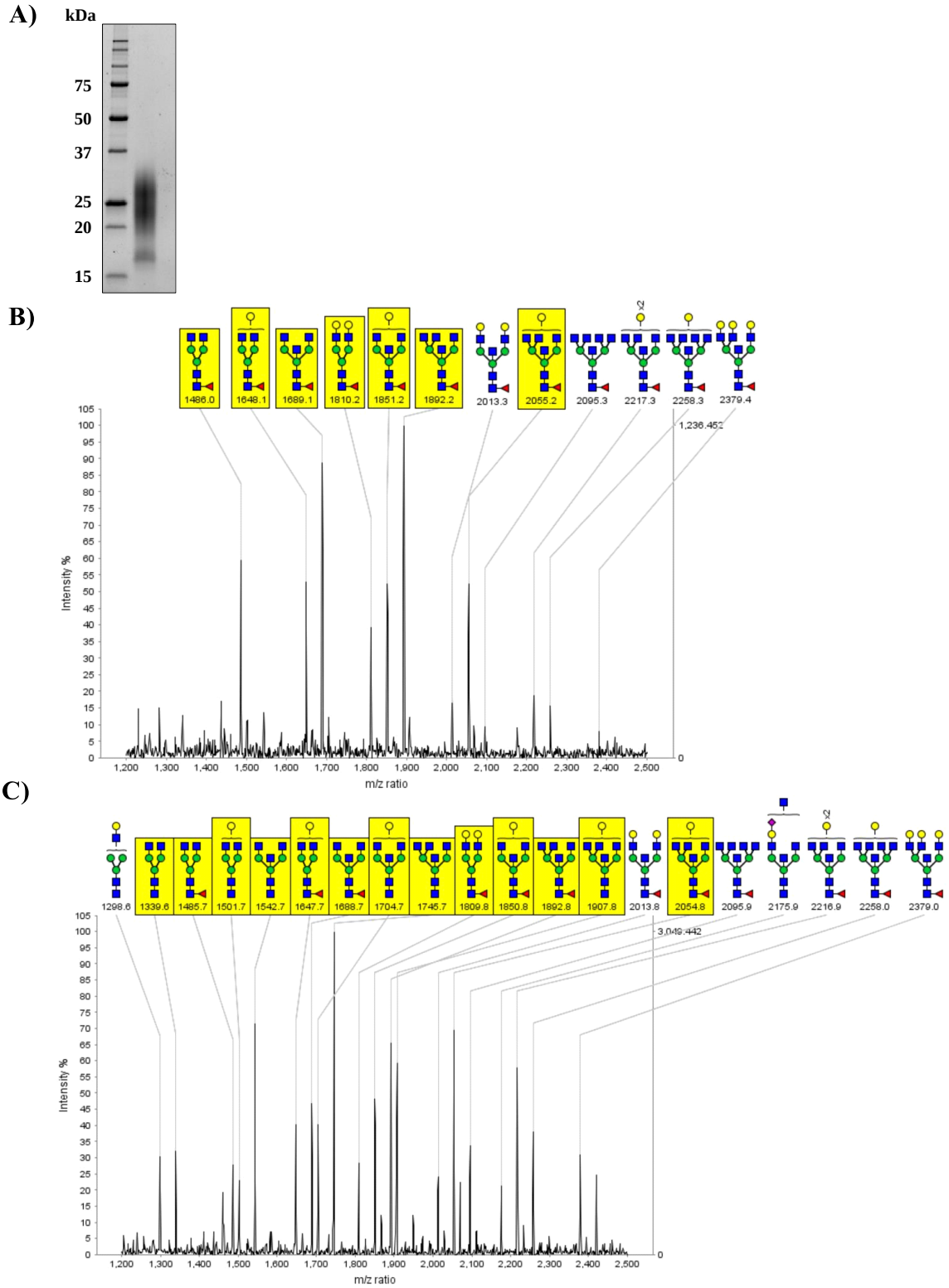
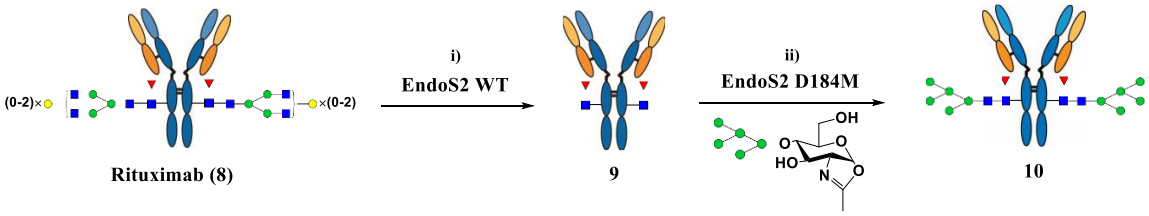


Figure 3.7. Expression and characterization of GM-CSF-WT (**7**). (A) SDS-PAGE of the purified protein showing a smear of bands. (B) MALDI-TOF MS profile of free *N*-glycan released from **7**. (C) MALDI-TOF MS profile of free *N*-glycan released from hydrolysis product of **7.** Highlighted glycoforms were monitored to calculate the hydrolysis yield.

Scheme 3.3: Chemoenzymatic synthesis of defined core-fucosylated glycoforms of antibody (9, 10)



Reagents and conditions: (i) Commercial Rituximab (4 mg/ml), immobilized EndoS2 WT (0.1 mg/ml), PBS (pH-7.4), RT, 3 h; (iii) **9** (10 mg/ml), EndoS2 D184M (0.2 mg/ml), Man-5 oxazoline (80 mol equiv of **9**), PBS (pH-7.4), 30°C

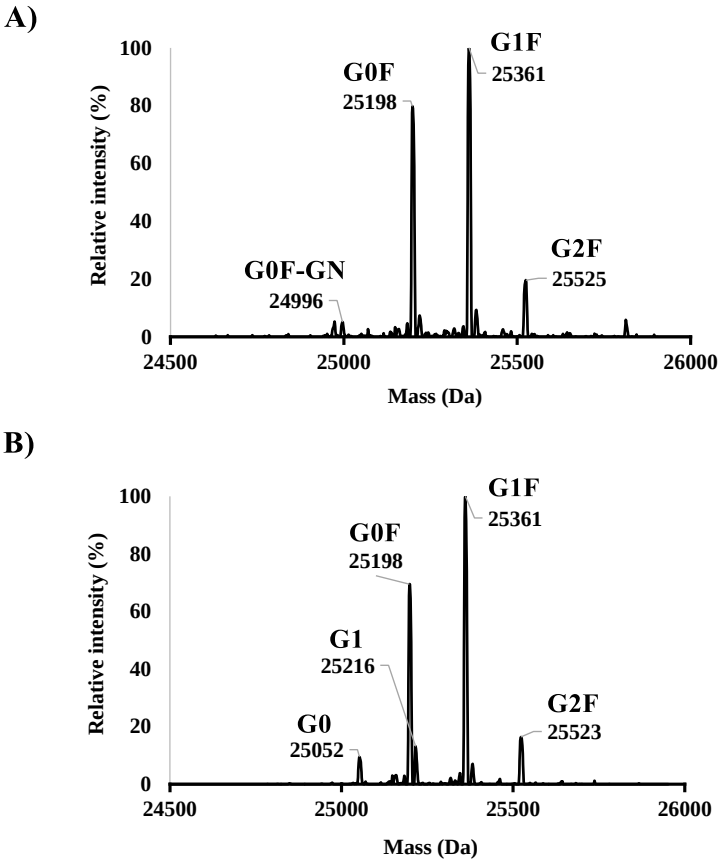


Figure 3.8. Characterization and hydrolysis of Rituximab **(8)**. Deconvoluted ESI-MS profiles showing Fc monomer of IdeS-treated antibody. (A) Rituximab, **8** and (B) **8** treated with FucA1 for seven days. G0F, G1F, G2F refer to core-fucosylated complex type glycan containing 0, 1 and 2 terminal galactose residues, respectively. G0F-GN refers to G0F complex type glycan without a terminal GlcNAc on one mannose arm. G0 and G1 refer to the non-fucosylated complex type glycan containing 0 and 1 terminal galactose residues, respectively.

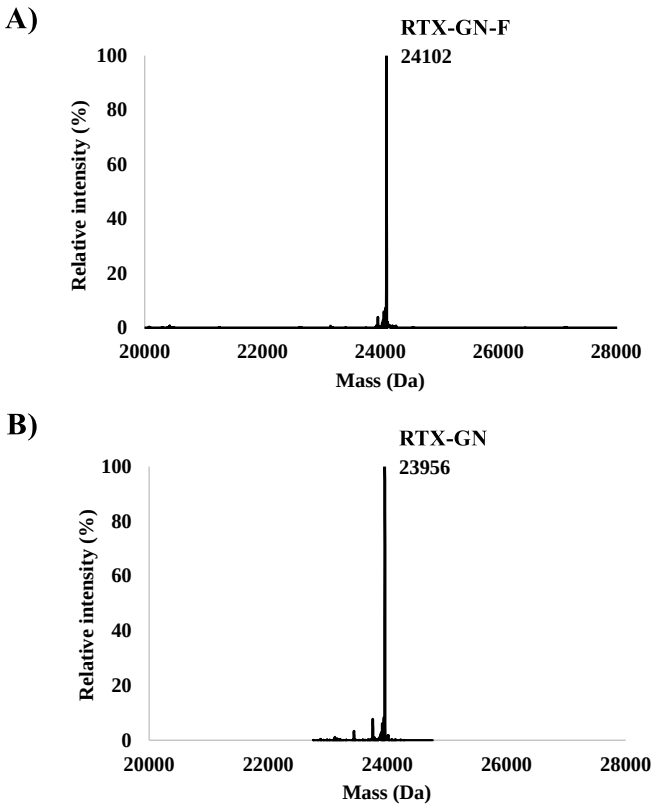


Figure 3.9. Chemoenzymatic synthesis and hydrolysis of RTX-GN-F **(9).** Deconvoluted ESI-MS profiles showing Fc monomer of IdeS-treated antibody. (A) RTX-GN-F, **9** and (B) the hydrolysis product, RTX-GN

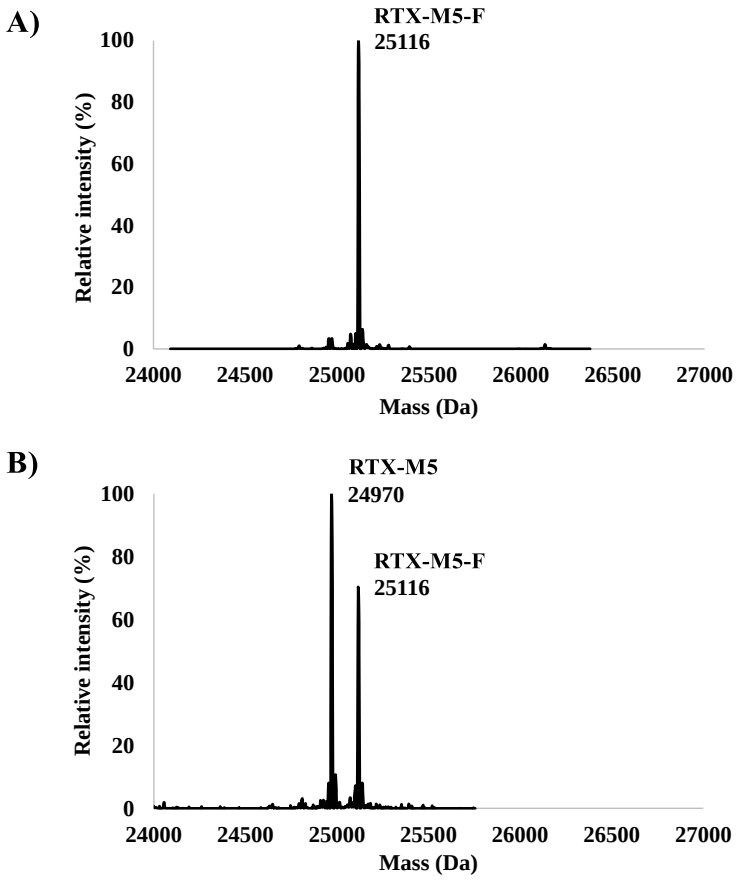


Figure 3.10. Chemoenzymatic synthesis and hydrolysis of RTX-M5-F **(10).** Deconvoluted ESI-MS profiles showing Fc monomer of IdeS-treated antibody. (A) RTX-M5-F, **10** and (B) **10** treated with FucA1 for seven days

3.2.3. Hydrolysis of *p*-nitrophenyl α-L-fucoside by α-L-fucosidases

Recombinant expression and purification of α-L-fucosidases was performed by following previously reported procedures. The human α-L-fucosidase, FucA1, was expressed as a fusion protein with an *N*-terminal green fluorescent protein (GFP) tag in HEK293T cells[140]. AlfC and BfFuc were expressed in *E. coli* BL21(DE3) strain[132]. The three His-tagged enzymes were purified by IMAC and analyzed using SDS-PAGE (Figure 3.11). Previous studies have characterized that the optimal pH of AlfC[128] and BfFuc[133] is around pH 7, while FucA1 has an acidic condition (pH 4.5) for the optimal

enzymatic hydrolysis[139]. To compare the substrate specificity and hydrolytic efficiency of the enzymes, we performed reactions at the respective optimum pH of each enzyme. Reactions with AlfC and BfFuc were done in PBS at pH 7.4. Meanwhile, FucA1 reactions were set up in sodium acetate buffer at pH 4.5.

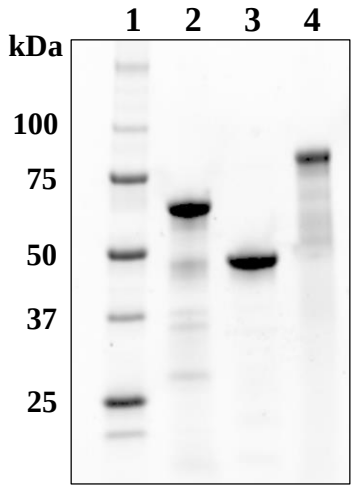


Figure 3.11. Expression and purification of recombinant α-fucosidases. The purified α-fucosidases were run on a reducing SDS-PAGE. Lane 1 – Protein ladder containing protein standards of specified sizes, Lane 2 – 10xHis-AlfC. Calculated molecular mass = 69 kDa. Lane 3 – 6xHis-BfFuc. Calculated molecular mass = 39 kDa. Lane 4 – 6xHis-GFP-FucA1. Calculated molecular mass = 85 kDa

The purified enzymes were first tested for activity with *p*NP-Fuc (**1**) at 37 ◦C. Reaction progress was monitored by measuring the UV absorbance (at 410 nm) of the *p*NP product formed. BfFuc and FucA1 showed efficient hydrolysis of **1** while AlfC was required at a 10-fold higher concentration to achieve comparable hydrolysis (Figure 3.12).

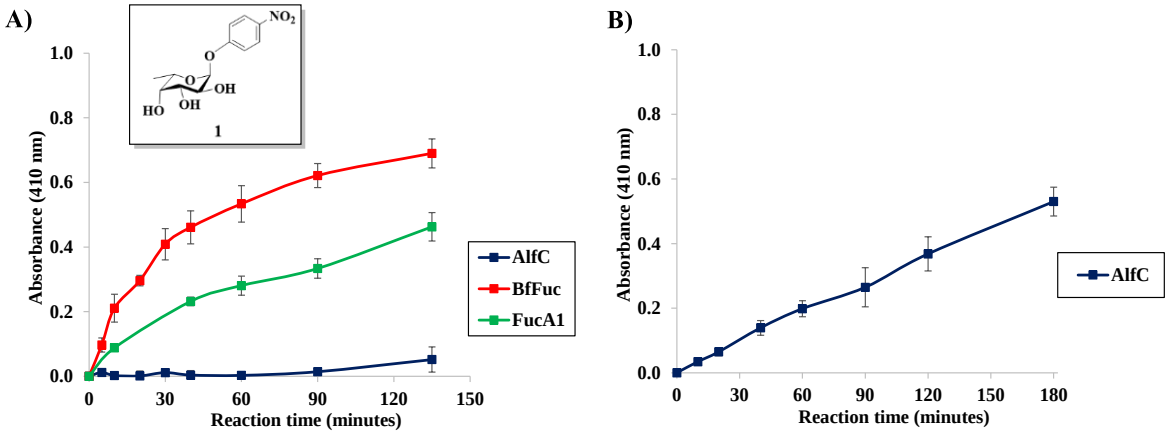


Figure 3.12. Hydrolysis of fucosylated saccharide by α-fucosidases. Time course profile of hydrolysis of (A) 5 mM *p*-nitrophenyl L-fucoside (**1**) by 0.5 µg/ml enzyme (n = 3). (B) 5 mM *p*-nitrophenyl L-fucoside (**1**) by 5 µg/ml AlfC (n = 3).

3.2.4. Hydrolysis of Fucα1,6GlcNAc-peptides

Next, we evaluated the hydrolysis of glycopeptides containing core-fucosylated GlcNAc. We used two glycopeptide substrates - the 19-mer CD52-GN-F (**2**) and the 24-mer V1V2-GN2-F2 (**4**) carrying one and two core fucose residues, respectively. The yields of defucosylation were estimated based on relative abundance of the cleaved product using mass spectrometric analysis. Additionally, a mixture of **4** and its hydrolysis product, V1V2-GN2 (**12**) were separated and quantified using RP-HPLC (Figure 3.4D) and the results were confirmed using mass spectrometry. In case of glycopeptide **2,** AlfC and BfFuc hydrolyzed the substrate efficiently, while FucA1 showed much lower hydrolytic activity than the two bacterial α-L-fucosidases (Figure 3.13A, 3.2C). Similar trends were observed with glycopeptide **4** (Figure 3.13B, 3.4C). Thus, the two bacterial α-L-fucosidases (AlfC and BfFuc) are much more active on the truncated Fucα1,6GlcNAc-peptides than the human α-L-fucosidase (FucA1).

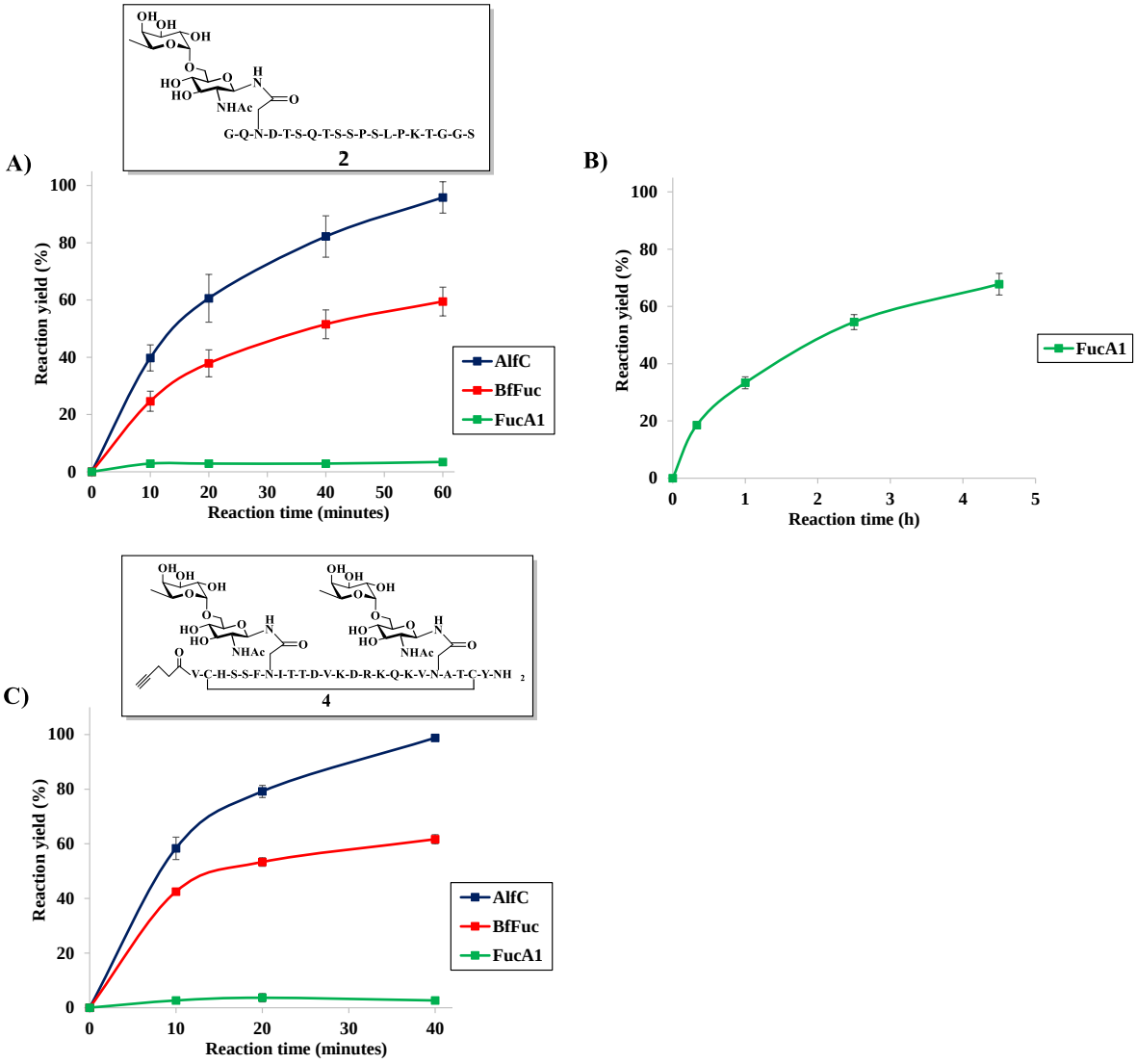


Figure 3.13. Hydrolysis of Fucα1,6GlcNAc-peptides by α-fucosidases. Time course profile of hydrolysis of (A) CD52-GN-F (**2**) at 0.01 mg/ml enzyme concentration (n = 3). (B) CD52-GN-F (**2**) by FucA1 at 1 mg/ml enzyme concentration (n = 3). (C) V1V2-GN2-F2 (**4**) at 0.01 mg/ml enzyme concentration (n = 3).

3.2.5. Hydrolysis of full-length fucosylated complex-type N-glycan and intact glycopeptide by α-L-fucosidases

Next, we turned our attention to test the substrate specificity of the three α-L-fucosidases on full-length fucosylated *N*-glycans and glycopeptides. Recent studies have demonstrated *in vitro* activity of FucA1 with synthetic aryl α-fucosides[139] and

with core-fucosylated biantennary sialylated complex-type free *N*-glycan[140]. We first assessed the activity of FucA1 and the other two α-L-fucosidases with core-fucosylated biantennary complex-type *N*-glycan (**5**). As expected, FucA1 efficiently defucosylated *N*-glycan (**5**). However, AlfC and BfFuc did not show any fucosidase activity on the full-length core-fucosylated *N*-glycan (**5**) (Figure 3.14A, 3.5B). Then, we tested the activity of these enzymes on the core-fucosylated full-length *N*-glycopeptide (**3**). The glycopeptide (**3**) was incubated with the α-L-fucosidases at 37 ◦C. The reactions were monitored by LC-MS analysis. We found that FucA1 showed considerable hydrolysis of the glycopeptide at 1 mg/ml enzyme concentration and the reaction could be pushed to completion by increasing the amount of enzyme and/or prolonging the incubation time (Figure 3.14B, 3.3C). In contrast, like the case of the full-length *N*-glycan (**5**), AlfC and BfFuc did not show hydrolytic activity on the intact full-length fucosylated glycopeptide (**3**) either (Figure 3.14B). Taken together, these results reveal that AlfC and BfFuc are much more active to defucosylate the truncated Fucα1,6GlcNAc-peptide substrates than FucA1, but the human enzyme (FucA1) is the only one that showed apparent activity to defucosylate intact full-length core-fucosylated *N*-glycopeptides.

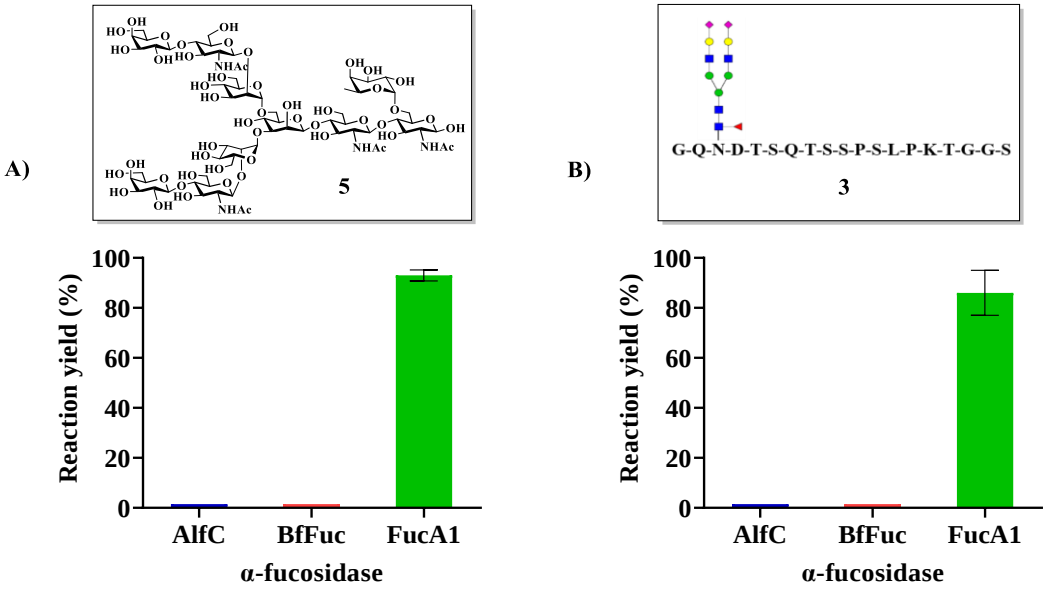


Figure 3.14. Hydrolysis of full-length fucosylated complex-type *N*-glycan and intact glycopeptide by α-fucosidases. (A) Hydrolysis of complex-type free *N*-glycan **(5)** by 1 mg/ml enzyme in 7 h (n = 3) (B) Hydrolysis of intact glycopeptide CD52SS-SCT-F **(3)** by 3 mg/ml enzyme in 21 h (n = 2).

3.2.6. Hydrolysis of fucosylated intact glycoproteins by the human α-L-fucosidase FucA1

The successful defucosylation of intact glycopeptide by FucA1 encouraged us to examine its activity with biologically relevant glycoproteins. Human granulocyte macrophage colony stimulating factor, GM-CSF is a cytokine involved in immune functions[148] and contains two conserved *N*-glycosylation sites. We expressed two types of glycoforms of GM-CSF using HEK293T cells – a) a fucosylated high mannose variant and b) the wild-type protein containing mainly complex-type *N*-glycans. First, we tested the hydrolysis of GM-CSF-HM-F (**6**) by the α-L-fucosidases. The extent of defucosylation was estimated by release of the total *N*-glycans and measurement of the released *N*-glycans using MALDI- TOF-MS analysis. We have previously found that

the estimate by MALDI-TOF-MS analysis of the fucosylated and non-fucosylated high-mannose *N*-glycans is consistent with the more accurate HPLC quantification[149]. Gratifyingly, an overnight incubation of GM-CSF-HM-F with FucA1 showed a high amount of defucosylation (estimated as 70% of released total *N*-glycan) (Figure 3.15A, 3.6C). Prolonging the incubation to two overnights under the reaction conditions showed 85% of released total *N*-glycan to be defucosylated (data not shown). Next, we assessed the defucosylation of GM-CSF-WT. The WT protein contains a mixture of bi- to tetra-antennary complex-type *N*-glycans, with fucosylated tetra- and tri-antennary glycans terminating in GlcNAc as the two most abundant glycoforms, respectively. MALDI-TOF-MS analysis of *N*-glycan released from FucA1-treated GM-CSF-WT showed defucosylation but the reaction progress was slow. We observed about 50% of total *N*-glycan released from the glycoprotein to be hydrolyzed in seven days (Figure 3.15B, 3.7C), while the protein control incubated without the enzyme did not show any loss of fucose. Upon comparing defucosylation of the two variants of GM-CSF by FucA1, hydrolysis of complex-type glycoprotein appears less efficient than that of the high-mannose type.

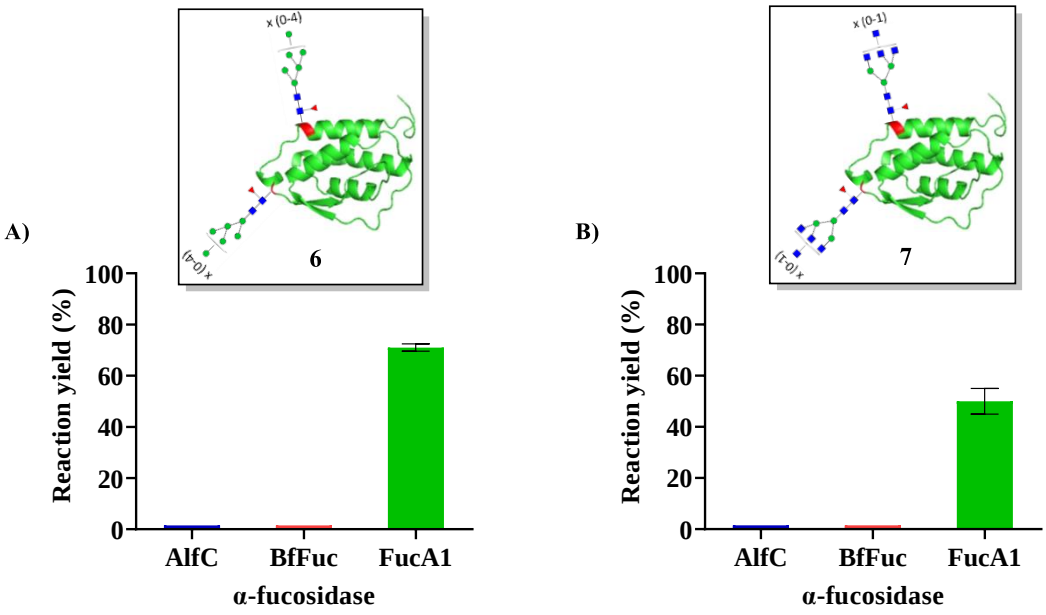


Figure 3.15. Hydrolysis of fucosylated intact glycoproteins by α-fucosidases. (A) Hydrolysis of GM-CSF-HM-F (**6**) by 1.5 mg/ml enzyme in 21 h (n = 3) (B) Hydrolysis of GM-CSF-WT (**7**) by 0.6 mg/ml enzyme in 7 days (n = 2)

3.2.7. Defucosylation of intact IgG antibody by the three α-L-fucosidases

Next, we evaluated the removal of core fucose from intact antibodies. The conserved glycosylation site at N297 in antibody Fc is buried within the antibody dimer making the site difficult to access for modifications. Hence, discovery of an enzyme that can efficiently modulate core fucose of intact antibodies is very valuable. To study the hydrolytic capability of FucA1 with antibodies, we used commercial monoclonal antibody, Rituximab (**8**). About 10% hydrolysis of the antibody was observed in seven days at an equimolar ratio of antibody to enzyme (Figure 3.16A, 3.8B). The result was reproducible in multiple independent reactions. Although the defucosylation yield was low, to our knowledge, this was the first example of direct enzymatic defucosylation of intact antibodies by an α-L-fucosidase. In contrast, AlfC and BfFuc did not show any

hydrolytic activity on the intact antibody. Wong and co-workers as well as our group have previously reported that AlfC and BfFuc could efficiently remove the core fucose from the endoglycosidase-pretreated; truncated Fucα1,6GlcNAc-antibody intermediate for glycan remodeling to produce non-fucosylated antibody glycoforms[74,86,132,133]. We thus compared the hydrolysis of RTX-GN-F (**9**) by the three α-L-fucosidases. In agreement with previous reports, AlfC and BfFuc were found to be efficient in removing core fucose from the deglycosylated antibody (Figure 3.16B, 3.9B). Interestingly, FucA1 showed only low hydrolytic activity on the Fucα1,6GlcNAc-antibody (**9**) under a similar condition and a large amount of enzyme was required to give hydrolytic activity comparable to the AlfC and BfFuc (Figure 3.16B). We also observed that at a high antibody substrate concentration (>20 mg/ ml), the AlfC showed higher apparent initial rate for the defucosylation of the Fucα1,6GlcNAc-antibody (**9**) than BfFuc, confirming our previous result[132]. Interestingly, when the substrate concentration is low (<5 mg/ ml), the BfFuc showed a higher apparent rate of hydrolysis of **9** (data not shown), consistent with the observation from Wong and co-workers[133]. These results suggest that AlfC may have lower substrate affinity (higher K_M) for the antibody but higher turnover rate than the BfFuc enzyme, which should be clarified in future kinetic studies. As an initial study to evaluate the effect of different *N*-glycans on the defucosylation of antibody by FucA1, we prepared a Man5 glycoform of Rituximab, RTX-M5-F (**10**) and tested it with FucA1. We found that FucA1 had much higher activity on the fucosylated Man5 glycoform (**10**) than the parent Rituximab (**8**), giving ca. 55% yield of defucosylation in a total of seven-day incubation (Figure 3.16C, 3.10B), while incubation of Rituximab (**8**) under similar condition gave

ca. 10% defucosylation (Figure 3.16A). This result suggests that FucA1 can distinguish between different core-fucosylated *N*-glycans for defucosylation. In contrast, AlfC and BfFuc did not show hydrolytic activity on the core-fucosylated high-mannose antibody glycoform (**10**) (Figure 3.16C). The finding that FucA1 possesses low but apparent defucosylation activity on intact monoclonal antibodies is significant, as this discovery raises an exciting opportunity to enzymatically remove the core fucose directly from recombinant monoclonal antibodies to enhance their antibody-dependent cellular cytotoxicity and the overall therapeutic efficacy[18,44,127]. The defucosylation activity of FucA1 on intact antibodies may be further improved through directed evolution[150,151].

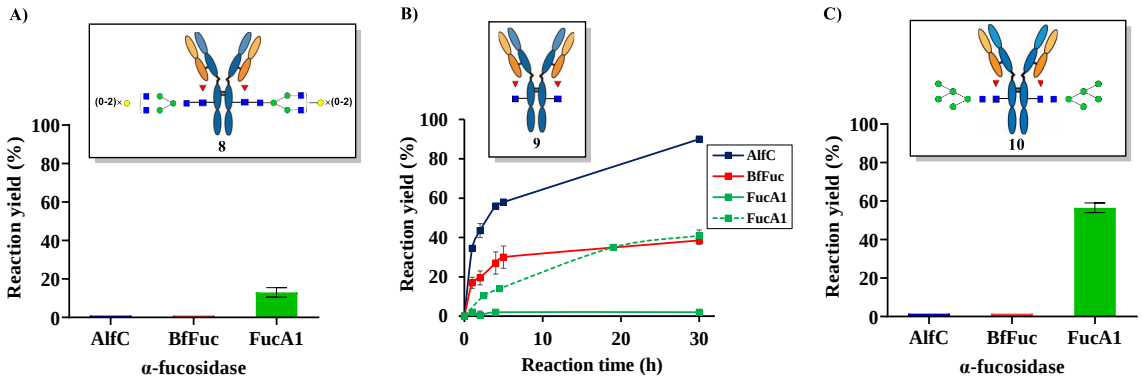


Figure 3.16. Hydrolysis of fucosylated intact antibody by α-fucosidases. (A) Hydrolysis of Rituximab (**8**) by 2.8 mg/ml enzyme in 7 days (n = 4). (B) Time-course profile of hydrolysis of RTX-GN-F (**9**) by 0.15 mg/ml enzyme (n = 2). Green dotted line shows activity of FucA1 at 1 mg/ml enzyme concentration (n = 2). (C) Hydrolysis of RTX-M5-F (**10**) by 1.6 mg/ml enzyme in 7 days (n = 3).

3.2.8. Surface plasmon resonance (SPR) analysis on the binding of FcγIIIa receptor with the antibodies treated by FucA1

To characterize the structural and functional integrity of the antibodies treated with FucA1, we assessed the receptor binding affinities of the partially defucosylated glycoforms obtained by FucA1 treatment. It is well established that the core fucose on the Fc *N*-glycans impacts the local conformations of Fc domain[152] and affects the antibody binding to FcγIIIa receptor (FcγRIIIa)[44,126]. Defucosylation has been shown to significantly enhance the antibody's FcγRIIIa affinity (up to 50-fold) for the typical complex type Fc glycoform[153] but only moderately (2–3 fold) enhance the affinity for FcγRIIIa in the case of high-mannose type glycoform[154,155]. Thus, the affinity of the partially defucosylated antibody mixtures for the FcγRIIIa might serve as a quick estimate on the structural and functional integrity of the antibodies after treatment with the fucosidase, as Fc denaturation might occur during the long incubation to potentially contribute to the observed defucosylation. First, we used the core-fucosylated rituximab and the partially defucosylated antibody generated by FucA1 treatment to measure the binding to recombinant FcγRIIIA V158 receptor using surface plasmon resonance (SPR) analysis. Our data showed a marginal enhancement (1.3-fold) in the binding affinity of the FucA1-hydrolyzed antibody (Figure 3.17A, 3.17B). The KD value changed from 71 nM of the intact rituximab to 56 nM of the partially (5–10%) defucosylation of rituximab. The data suggest that the Fc domain maintained its active conformations without denaturation during the long fucosidase incubation and a small fraction of the antibody was defucosylated, as reflected by the slight increase in the FcγRIIIa affinity. Next, we analyzed the binding affinities of the Man5 glycoforms of

rituximab. We found that the mixture, which contains ca. 55% defucosylated glycoforms after FucA1 treatment, showed about a 1.5- fold enhancement in binding affinity to the FcγRIIIa (Figure 3.17C, 3.17D). The K_D value changed from 103 nM of the fucosylated Man5 glycoform to 70 nM of the partially defucosylated Man5 glycoform. To provide an accurate side-by-side comparison of the impact of defucosylation of the Man5 glycoforms, we produced a non-fucosylated rituximab Man5 standard through glycoengineering. Binding analysis of the rituximab Man5 glycoform with FcγRIIIa V158 showed a K_D of 43 nM (Figure 3.17E), which translates to a 2.4-fold enhancement in binding affinity over the core-fucosylated glycovariant. The K_D value of rituximab Man5 glycoform agrees with the previous reported values of 32 nM for the Man5 glycoform and 27 nM for the Man8-9 glycoforms of IgG1[156]. Indeed, in comparison to the core-fucosylated complex-type Fc glycoforms of antibodies, the non-fucosylated oligomannose glycoform appears to possess only a 2–3-fold enhancement in affinity for FcγRIIIa, as demonstrated in several previous studies[154,155]. The moderate impact of core-fucosylation on FcγRIIIa affinity was in sharp contrast to the dramatic increase (up to 50-fold) in binding affinity upon defucosylation of core-fucosylated complex-type antibody Fc glycoforms. Our results are consistent with previous studies and the 1.5-fold enhancement in FcγRIIIa affinity of the partially defucosylated Man5 glycoforms confirmed that about 50% defucosylation occurred in the sample and that the Fc domain maintained an active conformation without denaturation during the FucA1 treatment.

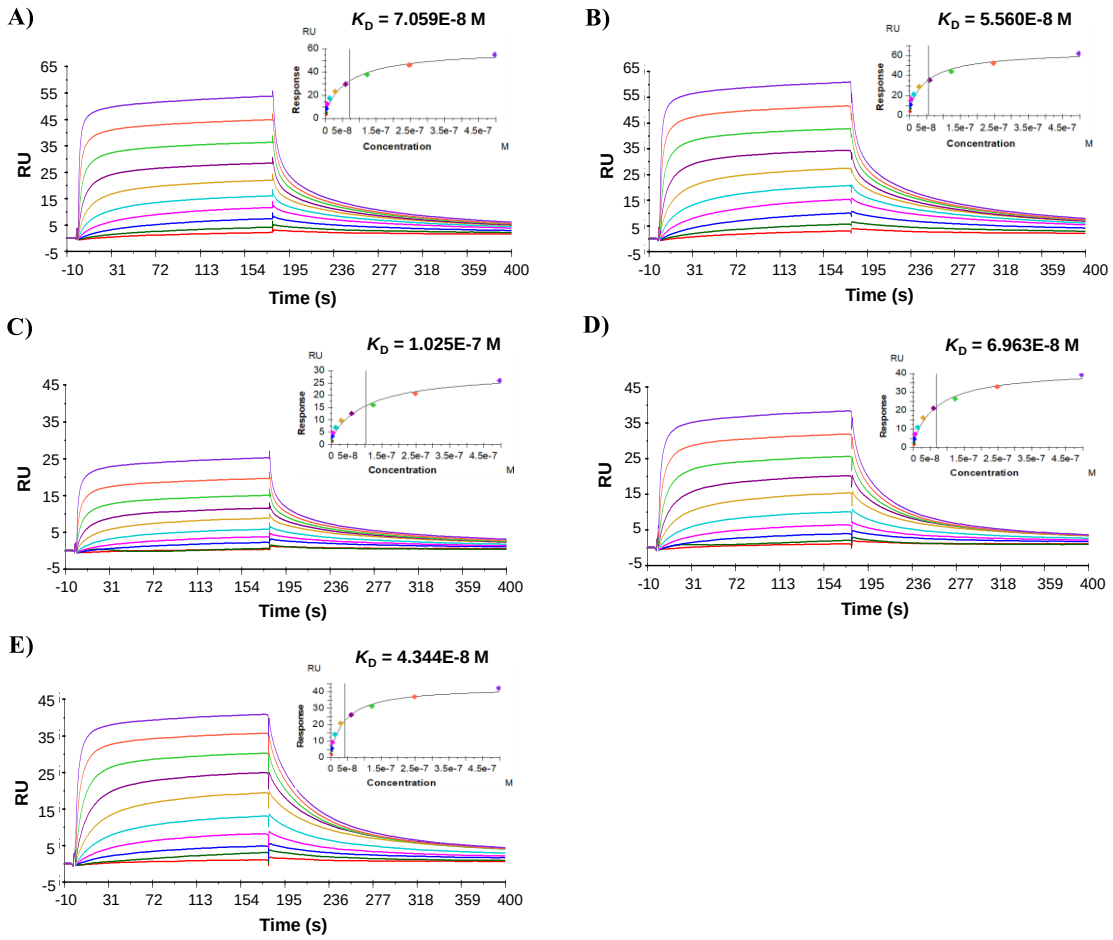


Figure 3.17. SPR sensorgrams showing binding profiles of antibody glycoforms to FcγRIIIA receptor. (A) Intact Rituximab (B) Rituximab defucosylated by FucA1 (10% yield) (C) RTX-M5-F (D) RTX-M5-F defucosylated by FucA1 (55% yield) SPR sensorgrams showing binding profiles of antibody glycoforms to FcγRIIIA receptor. (E) RTX-M5 standard

3.3 Conclusion

A comparative study of the substrate specificity and relative activity of the bacterial α-L-fucosidases from *Bacteroides fragilis* (BfFuc) and *Lactobacillus casei* (AlfC) and the human α-L-fucosidase (FucA1) is described. This study was enabled by the synthesis of an array of structurally well-defined core-fucosylated *N*-glycopeptides and glycoproteins including a few antibody glycoforms. The experimental data reveal that the two bacterial α-L-fucosidases hydrolyze only truncated Fucα1,6GlcNAc-peptide

substrates and they are not active to remove the core fucose when it is present in intact full-length *N*-glycans or *N*-glycoproteins. In contrast, the human α-L-fucosidase (FucA1) shows low activity on truncated Fucα1,6GlcNAc-peptide substrates but FucA1 can remove the core fucose from intact full-length *N*-glycopeptides and glycoproteins. In addition, FucA1 appears to be the only α-L-fucosidase that shows low but apparent activity to remove core fucose from intact IgG antibodies. The ability of FucA1 to defucosylate intact antibodies points to a promising way to directly remove the core fucose from intact therapeutic antibodies to improve their antibody-dependent cellular cytotoxicity for treatment of human diseases.

3.4 Experimental

3.4.1. Materials

p-Nitrophenyl α-L-fucoside was purchased from Sigma-Aldrich. Monoclonal antibody Rituximab (Genentech Inc., South San Francisco, CA) was purchased from Premium Health Services Inc. (Columbia, MD). Sialoglycan complex-type oxazoline (SCT-oxazoline) was synthesized using previously established procedure[86]. Man5-oxazoline was synthesized using previously described method[146]. AlfC E274A, EndoF3 D165A, EndoS2 WT and EndoS2-D184M were overexpressed and purified using previously reported methods[74,86,87].

3.4.2. Methods

Analytical HPLC was performed with a Waters 626 HPLC system using a YMC column (5 μm, 4.6 × 250 mm). Preparative HPLC was performed with a Waters 600 HPLC system equipped with a SymmetryPrep™ C18 column (7 μm, 19 × 300 mm) using a flow rate of 12 ml/ min or an XBridge™ Prep Shield RP18 column (5 μm, 10 × 250 mm) using a flow rate of 4 ml/min. Water containing 0.1% TFA and acetonitrile containing 0.1% TFA were used as Solvents A and B respectively for both analytical and preparative chromatography purposes. LC-MS analysis of saccharides and glycopeptides was performed using a Waters Alliance e2695 HPLC system connected to an SQ Detector 2. The C18 columns used for analysis include a long Thermo Scientific Hypersil column (3 μm, 4.6 × 250 mm) employing a flow rate of 1 ml/min and a short XBridge™ column (3.5 μm, 2.1 × 50 mm) using a flow rate of 0.4 ml/min. Mass spectrometric analysis of CT-F glycan was performed using Bruker Autoflex III MALDI-TOF (Bruker Daltonics) in reflectron positive mode using 2,5-dihydroxybenzoic acid (DHB) matrix. The matrix was prepared by dissolving 100 mg DHB and 20 μl of N, *N*-dimethylaniline (DMA) in 50% aqueous acetonitrile solution. LC-MS analysis of antibodies was carried out using Thermo Scientific Exactive™ Plus Orbitrap mass spectrometer on a Waters XBridge™ BEH300 C4 column (3.5 μm, 2.1 × 50 mm). The method involved a 9-minute linear gradient of 5 to 90% acetonitrile containing 0.1% formic acid at a flow rate of 0.4 ml/min. Deconvolution of raw data was done using MagTran (Amgen).

3.4.3. Expression and purification of AlfC, BfFuc and FucA1 α-L-fucosidases

α-L-fucosidases from *Lactobacillus casei* (AlfC) and *Bacteroides fragilis* (BfFuc) were overexpressed in *E. coli* BL21(DE3) competent cells and purified as previously reported[132]. DNA construct for expression of human α-L-fucosidase (FucA1), pGEn2-FucA1 containing the catalytic domain of the enzyme was purchased from Glycozyme, CCRC, UGA through DNASU plasmid repository. Overexpression of FucA1 was done in HEK293T cells following the reported procedures31. Briefly, a day prior to transfection, HEK293T cells were seeded at a density of 7×10^5/ ml in serum-free FreeStyle™ F17 Expression Medium (Thermo Fisher Scientific) and cultured at 37 ◦C, 8% CO2, 150 rpm. The following day, transient transfection was performed with 1 μg DNA (pGEn2-FucA1 expression vector) and 2 μg polyethylenimine (PEI) per million cells. The expression culture was incubated at 37 ◦C, 8% CO2, 150 rpm for 3 days. At harvest, the broth was spun down at 1000 rpm for 10 min at 4 ◦C and the supernatant was purified using a HisTrap™ column (GE Healthcare), following manufacturer's protocol. The eluted protein was buffer exchanged to PBS, pH-7.4 and characterized using SDS-PAGE. Long-term storage of enzyme aliquots was done at -80 ◦C.

3.4.4. Synthesis of core-fucosylated substrates

3.4.4.1. Synthesis of CD52-GN-F (2)

Core-fucosylated CD52 antigen (**2**) was synthesized from CD52-GN containing a signal peptide sequence (**11**) following previously outlined procedure[145]. **11** (10 mg, 5 μmol) and α-fucosyl fluoride (1.7 mg, 10 μmol) were mixed with AlfC E274A (0.5

mg/ml) in 1 ml PBS, pH-7.4 and incubated at 37 ◦C. Reaction progress was monitored by LC-MS. Complete transfer was achieved with further addition of α-fucosyl fluoride (3.7 mg, 21.8 μmol). The fucosylated product was purified on a C18 column using RP-HPLC using a linear gradient of 5 to 35% B over 30 min at a flow rate of 12 ml/min. tR = 14.96 min. ESI-MS: calcd. for **2,** *M* = 2197.0 Da; found (*m*/*z*), 733.49 [M+3H]3+, 1099.60 [M+2H]2+.

*3.4.4.2. Synthesis of CD52-SCT-F (**3**)*

Core-fucosylated biantennary sialylated CD52 (**3**) was generated from **2** as reported before[87]. EndoF3-D165A (0.5 mg/ml) was added to a mixture of **2** (2 mg, 0.91 μmol) and SCT-oxazoline (4 mg, 2 μmol) in 120 μl PBS, pH-7.4. The reaction mixture was incubated at 30 ◦C for 0.5 h. The final product was purified using RP-HPLC on a C18 column using a linear gradient of 5 to 25% B over 40 min at a flow rate of 4 ml/min. tR = 10.01 min. ESI-MS: calcd. for **3**, *M* = 4199.1 Da; found (*m*/*z*), 1051.09 [M+4H]4+, 1400.93 [M+3H]3+

*3.4.4.3. Synthesis of V1V2-GN2-F2 (**4**)*

First, V1V2-GN2 (**12**) was synthesized using SPPS by modifying our previously described procedure[146]. To synthesize the core fucosylated compound **4**, **12** (1.43 mg, 0.44 μmol) and α-fucosyl fluoride (0.29 mg, 1.76 μmol) were incubated with AlfC E274A (0.5 mg/ml) in 0.5 ml PBS, pH-7.4 at 37 ◦C. The reaction was monitored by LC-MS. The reaction was pushed to completion with further addition of α-fucosyl fluoride (0.45 mg, 2.64 μmol) while maintaining the pH of the mixture. After 1 h of

incubation, the fucosylated peptide was purified using RP-HPLC on a C18 column using stepwise gradient of solvent B starting with 5 to 20% over 10 min followed by 20 to 40% over 30 min at a flow rate of 4 ml/min. tR = 10.12 min. ESI-MS: calcd. for **4**, *M* = 3533.7 Da; found (*m*/*z*), 707.78 [M + 5H]5+, 884.57 [M+4H]4+, 1179.04 [M+3H]3+.

*3.4.4.4. Synthesis of the fucosylated biantennary complex-type N-glycan (**5**)*

Sialylated biantennary complex-type glycan (SCT) was extracted from SGP present in egg yolk powder (Henningsen Foods) and treated with neuraminidase to generate asialo complex-type *N*-glycan, using previously reported methods[78]. Purified CT-glycan (**13**) was fucosylated by incubating with AlfC E274A (2 mg/ml) and α-fucosyl fluoride (10 mol equivalent of **13**) in PBS, pH-7.4 at 37 ◦C for 1 h. The final product was purified by preparative HPLC on a C18 column using a linear gradient of 0-5% acetonitrile containing 0.1% formic acid. MALDI-TOF-MS: calcd. for **5**, *M* = 1787 Da; found (*m*/z), 1809.7 [M + Na]+ .

*3.4.4.5. Expression and purification of GM-CSF (**6**,**7**)*

GM-CSF-HM-F (**6**) was expressed using HEK 293 T FUT8+ cells following a previously reported procedure[158]. Briefly, cells were seeded at a density of 7×10^5/ml in serum-free FreeStyle™ F17 Expression Medium (Thermo Fisher Scientific) and cultured at 37 ◦C, 8% CO2, 150 rpm. The overnight grown culture was transiently transfected with 1 μg DNA (pcDNA3.1-GM-CSF plasmid) and 3 μg polyethylenimine (PEI) per million cells. 4 μM of kifunensine was supplemented post-transfection to

restrict protein glycosylation to high mannose glycoform. The expression culture was incubated at 37 ◦C, 8% CO2, 150 rpm for 3 days. At harvest, the broth was spun down at 1000 rpm for 10 min at 4 ◦C and the supernatant was purified using a HisTrap™ column (GE Healthcare), following manufacturer's protocol. The eluted protein was buffer exchanged to PBS, pH-7.4 and characterized using SDS-PAGE and mass spectrometric analysis. Expression and purification of GM-CSF-WT (**7**) was performed as above using HEK293T cells without supplementation of kifunensine.

3.4.4.6. Chemoenzymatic remodeling of Rituximab

Commercial Rituximab (**8**, 4 mg/ml) was deglycosylated by incubating with immobilized EndoS2 WT (0.1 mg/ml) in PBS, pH-7.4 at room temperature for 3 h and purified on a Protein A column to obtain RTX-GN-F, **9** as reported previously[132]. The samples were analyzed by MALDI-TOF-MS. **9** (10 mg/ml) was incubated with EndoS2 D184M (0.2 mg/ml) and Man5-oxazoline (80 mol equivalent of **9**) in PBS, pH-7.4 at 30 ◦C to generate RTX-M5-F (**10)**. Reaction progress was monitored using LC-MS and the final product was purified using Protein A chromatography. ESI-MS analysis of the Fc monomer released by IdeS treatment of the RTX-M5-F showed a species of 25116 Da (after deconvolution) which agrees with the calculated molecular mass of the Fc monomer carrying a Man5-F glycan. To prepare the RTX-M5 standard, **9** (10 mg/ml) was first treated with AlfC (0.1 mg/ml) in PBS, pH- 7.4 at 37 ◦C to produce RTX-GN and purified using Protein A chromatography. Purified RTX-GN was then incubated with EndoS2 D184M (0.2 mg/ml) and Man5-oxazoline (80 mol equivalent) in PBS, pH-7.4 at 30 ◦C. Purification and analysis were done as described above. ESI-MS

analysis of the Fc monomer released by IdeS treatment of the RTX-M5 showed a species of 24971 Da (after deconvolution).

3.4.5. Testing hydrolytic activity of α-L-fucosidases with fucosylated saccharide

The hydrolytic activity of the α-L-fucosidases with simple saccharides was tested by incubating **1** (5 mM) and the respective enzyme (0.0005 mg/ml) in 100 µl PBS, pH-7.4 or sodium acetate, pH 4.5 at 37 ◦C. Samples were collected through the course of the reaction and quenched with 0.1 M sodium hydroxide solution. Reaction products were monitored by measuring the absorbance at 410 nm. The above reaction was also conducted with 10-fold higher concentration of AlfC (0.005 mg/ml) while maintaining other reaction conditions constant.

3.4.6. Testing hydrolytic activity of α-L-fucosidases with Fucα1,6GlcNAc-peptides

To test the hydrolysis of core α-1,6 fucose linked to GlcNAc peptides, **2** (5 mg/ml) was incubated with each of the enzymes (0.01 mg/ml) in 10 µl PBS, pH-7.4 or NaOAc, pH-4.5 at 37 ◦C. Intermediate reaction samples were collected and analyzed by LC-MS. ESI-MS: calcd. for hydrolysis product, M = 2051.4 Da; found (m/z), 684.85 [M + 3H]3+ , 1026.58 [M + 2H]2+ . The activity of FucA1 was tested at ten-fold higher concentration (1 mg/ml) by incubating with **2** (5 mg/ml) under similar reaction conditions. Similarly, **4** (4 mg/ml) was mixed with respective enzymes (0.01 mg/ml) in 10 µl PBS, pH-7.4 or NaOAc, pH-4.5 at 37 ◦C. Samples were collected at 10, 20, 40

and 60 min and analyzed using LC-MS. ESI-MS: calcd. for hydrolysis product, *M* = 3241.7 Da; found (*m*/*z*), 649.38 [M + 5H]5+, 811.43 [M + 4H]4+, 1081.58 [M + 3H]3+.

3.4.7. Testing hydrolytic activity of α-L-fucosidases with complex-type N-glycan and intact glycopeptide

To test the hydrolysis of complex-type glycoform by the α-L-fucosidases, **5** (5 mg/ml) was incubated with respective enzymes (1 mg/ml) in 8 µl PBS, pH-7.4 at 37 ◦C. Reaction samples were analyzed by MALDI-TOF-MS. MALDI-TOF-MS: calcd. for hydrolysis product **5**, *M* = 1641 Da; found (*m*/*z*), 1666.3 [M + Na]+. Hydrolysis of intact glycopeptide was tested by mixing **3** (3 mg/ml) and FucA1 (3 mg/ml) in 10 µl of 100 mM NaOAc, pH-4.5 at 37 ◦C for 21 h. Reactions with AlfC (2.9 mg/ml) and BfFuc (1.3 mg/ml) were performed in 8 µl PBS, pH-7.4 at 37 ◦C. Sample analysis was performed by LC-MS. ESI-MS: calcd. for hydrolysis product, *M* = 4053.1 Da; found (*m*/*z*), 1014.36 [M + 4H]4+, 1352.29 [M + 3H]3+.

3.4.8. Testing hydrolytic activity of α-L-fucosidases with glycoproteins

The hydrolytic activity of the α-L-fucosidases with intact glycoproteins was tested by incubating **6** (1.5 mg/ml) and FucA1 (1.5 mg/ml) in 70 mM NaOAc, pH-4.5 in 10 µl total volume at 37 ◦C for 45 h. Reactions with AlfC and BfFuc were performed in 10 µl PBS, pH-7.4 at 37 ◦C. Sample analysis was done by *N*-glycan release and MALDI-TOF-MS. Hydrolysis of complex-type intact glycoprotein was tested by incubating **7**

(0.55 mg/ml) and the respective enzymes (0.6 mg/ml) in 180 µl of 40 mM NaOAc, pH-4.5 or PBS, pH-7.4 at 37 ◦C for 7 days.

3.4.9. Testing hydrolytic activity of α-L-fucosidases with antibodies

The hydrolytic activity of the α-L-fucosidases with intact antibody was tested by incubating **8** (2.8 mg/ml) and the respective enzymes (2.8 mg/ml) in 18 µl of 70 mM NaOAc, pH-4.5 or PBS, pH-7.4 at 37 ◦C. Samples were treated with IdeS protease (0.02 mg/ml) in PBS at 37 ◦C for 15 min to generate Fc monomers and analyzed by LC-MS. To test the hydrolysis of RTX-GN-F, **9** (22 mg/ml) was incubated with respective enzymes (0.15 mg/ml) in 10 µl PBS, pH-7.4 or NaOAc, pH-4.5 at 37 ◦C. Samples were collected through the course of the reaction and analyzed by LC-MS post-IdeS treatment. The hydrolytic activity of the α-L-fucosidases with antibody containing fucosylated Man5 glycan was tested by incubating **10** (1.6 mg/ml) and the respective enzymes (1.6 mg/ml) in 18 µl of 100 mM NaOAc, pH-4.5 or PBS, pH-7.4 at 37 ◦C. Samples were analyzed by LC-MS post-IdeS treatment.

3.4.10. N-glycan release and MALDI-TOF-MS analysis

N-glycans were released from glycoproteins using Peptide: *N*-glycosidase F (PNGase F) under denaturing condition following the protocol recommended by New England Biolabs (NEB). The cleaved glycans were purified using HyperSep™ Hypercarb™ SPE cartridges (Thermo Scientific). The eluted samples were lyophilized and analyzed

by MALDI-TOF-MS in reflectron positive mode using 2,5-dihydroxybenzoic acid (DHB) matrix. Glycan assignment was done using GlycoWorkbench.

3.4.11. SPR binding analysis

The experiment was carried out by capturing each antibody glycoform onto the protein A chip and flowing serial dilutions of FcγRIIIA V158 as the analyte. After each cycle, the surface was regenerated by injecting a glycine HCl buffer (10 mM, pH 2.0). The antibodies were captured at 200 RU. The receptor in 2-fold serial dilutions (from 500 nM to 0.976 nM) was injected at 30 µl/min for 180 s, followed by a 300 s dissociation. The experimental data were fit to a 1:1 Langmuir binding model using the BIA Evaluation software (GE Healthcare) to obtain the steady state kinetic data.

Chapter 4: Modulating core fucosylation of antibodies through metabolic glycoengineering using L-fucose analogs

4.1 Introduction

Metabolic glycoengineering provides a tool to manipulate the biosynthetic glycosylation pathway of cells using exogenously supplied unnatural monosaccharides[159-161]. This approach has long been studied to identify L-fucose analogs that can metabolically inhibit protein and cell surface fucosylation, some of which have been investigated in the context of cancer cell growth[62,122]. Compelling examples include per-acetylated 6,6,6-trifluorofucose (**6**), 2-fluorofucose (**1**) and 6-alkynylfucose (**8**) which show efficient inhibition of core fucosylation and minimal incorporation of the deacetylated sugar analogs into the target proteins[62,162-165]. An alternate effort has been towards identifying fucose probes that can be competently incorporated into cell surface glycans for imaging or for labeling glycoproteins, including recombinant antibodies. The progress in this direction has been far more challenging due to cell line specific effects, cytotoxicity, or low sensitivity of fucose probes. For example, 6-azido fucose (6-Az-Fuc, **7**) has long been described as an efficient probe for cell surface labeling using azide-alkyne click chemistry but has not found wide applications due to stated cytotoxicity of the azido-sugar in cells[166-168]. Similarly, a recent study used 7-alkynylfucose (**9**) to label cellular glycans and detected the incorporated sugar using azide-biotin click and streptavidin blotting[169]. While the efficiency of the analog incorporation into cellular glycoproteins was found to be high,

a follow-up study shows limited labeling of secreted proteins, which in turn varied across expression cell lines[170]. Thus, use of this fucose probe to tag recombinant proteins in cells requires further investigation.

There have been some promising reports of efficient protein labeling with fucose probes. Senter and coworkers screened a library of compounds with modifications at different positions in fucose[171]. Among these, fucose analogs carrying modifications at the C-6 position of fucose including 6-thio, 6-chloro and 6-fluoro fucose etc., showed 70 – 90% incorporation into recombinant antibodies[171]. Although, among the fucose analogs that were efficiently incorporated, only 6-thiofucose (**10**) functionalizes the target protein and can be used for further applications. In the study, the authors generated a site-specifically labeled antibody-drug conjugate by ligating a drug to the thio-modified antibody using maleimide reaction[171]. In another study, Hossler et.al. used D-arabinose (**17**) and L-galactose (**16**) independently to completely inhibit antibody core fucosylation and demonstrated concomitant incorporation of the monosaccharides, attached to the innermost GlcNAc[172]. Moreover, D-arabinose incorporation in the antibody N-glycan core showed an improvement in FcγIIIa receptor binding by 2-fold and ADCC activity that was comparable to that of non-fucosylated antibody[172]. Further structural and/or functional analysis of the arabinose-modified antibodies using homogeneous glycoforms will be required to understand the direct impact of arabinose addition in the N-glycan core.

In general, for an L-fucose analog to be a good incorporator, the efficiency of analog uptake by cells and conversion to corresponding GDP-analog through the salvage pathway of fucose metabolism (Figure 4.1) are key determinants. Build-up of GDP-

fucose-analog in cells feedback inhibits formation of GDP-fucose and thus controls core fucosylation[62]. Furthermore, the utilization of these GDP-fucose analogs by FUT8 determines the efficiency of incorporation of the analog into the N-glycan core of glycoproteins. Availability of functionalized fucose analogs that can bio-orthogonally label purified proteins in quantitative yields are valuable for antibody-drug conjugate designs and for structure-function relationship studies. The objective of our study was to evaluate a panel of L-fucose analogs to identify efficient incorporators into the fucose metabolic pathway to possibly tag recombinant antibodies for further studies. We synthesized an array of L-fucose analogs with substitutions at the C-6 position of fucose as C-6 modifications are reported to be well tolerated in the salvage pathway[166]. Also, the corresponding GDP-analog precursors are reported to be accepted as substrates by FUT8[166,167]. We performed side-by-side evaluation of fluorinated, azido-modified, alkynylated and thio-modified L-fucose analogs to compare novel and known fucose probes for their effect on antibody fucosylation. Our data shows that the novel compounds synthesized with azide and alkene functional groups linked using carbon spacers to the methyl group of fucose do not have any effect on the core fucosylation of antibodies. The lack of effect possibly implies that the analogs are not processed by the enzymes of the salvage pathway to form enough corresponding GDP-analog precursors. Among the other compounds tested, peracetylated 6-fluoro-fucose (**4**), 6,6-difluoro-fucose (**5**) and 6-thio-fucose (**10**) show efficient incorporation into the N-glycan core of antibodies. While **5** has been previously reported to be incorporated into mAb, the highest incorporation was said to be 11%[162]. Our results show a concentration-dependent effect of **5** and *c.a.* 80% incorporation was observed at an

analog concentration of 2 mM. Interestingly, peracetylated 6-Az-fucose (**7**) did not show any detectable labeling of antibody up to 2 mM concentration in HEK293 cells. And contrary to the previous reports[166,167], the azido-sugar was not found to be cytotoxic up to a high concentration of 1 mM. Peracetylated 7-alkynyl-fucose (**9**) was found to be ineffective till 1 mM concentration. A partial incorporation was observed at 2 mM and addition of the analog at 5 mM concentration led to cytotoxicity. Similarly, L-galactose (**16**) and D-arabinose (**17**) were found to only partially label antibodies. Further studies are being undertaken to assess a potential cell-line dependent effect.

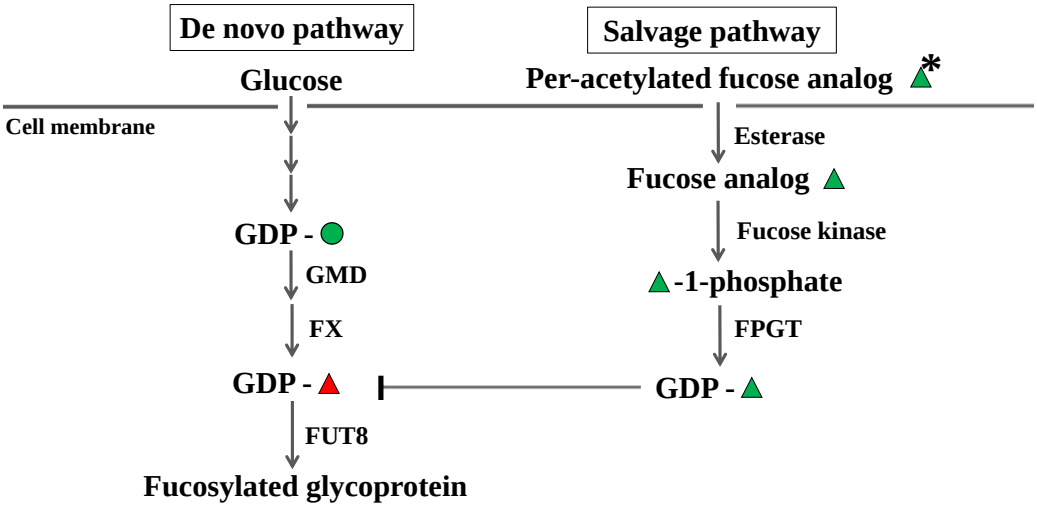


Figure 4.1. Fucose metabolic pathways in mammalian cells. GMD: GDP-mannose 4,6 dehydratase; FX: GDP-keto-6-deoxymannose 3,5-epimerase, 4-reductase; FPGT - Fucose-1-phosphate guanylyltransferase; FUT8 - Fucosyltransferase VIII

4.2 Results and Discussion

4.2.1 Panel of L-fucose analogs for modulation of antibody core fucosylation

A total of 15 L-fucose analogs **1-15**, were synthesized to study the effect on core fucosylation of monoclonal antibodies. Figure 4.2 shows the panel of compounds tested for antibody labeling. The rationale behind the design was to include small, bio-orthogonal groups at the C-6 position of L-fucose to potentially allow the fucose analogs to be accepted as substrates by FUT8 for antibody labeling. All the free hydroxyl groups in the monosaccharides were protected by peracetylation to increase their hydrophobicity for efficient diffusion across cell membranes. Synthesis of compounds was done by Dr. Yuanwei Dai and Dr. Xiao Zhang in the Wang lab. The synthetic details and characterization data will be discussed elsewhere. Compounds **16** and **17**, which are free sugars unlike the synthesized compounds, were purchased from Biosynth Carbosynth® and Millipore Sigma respectively.

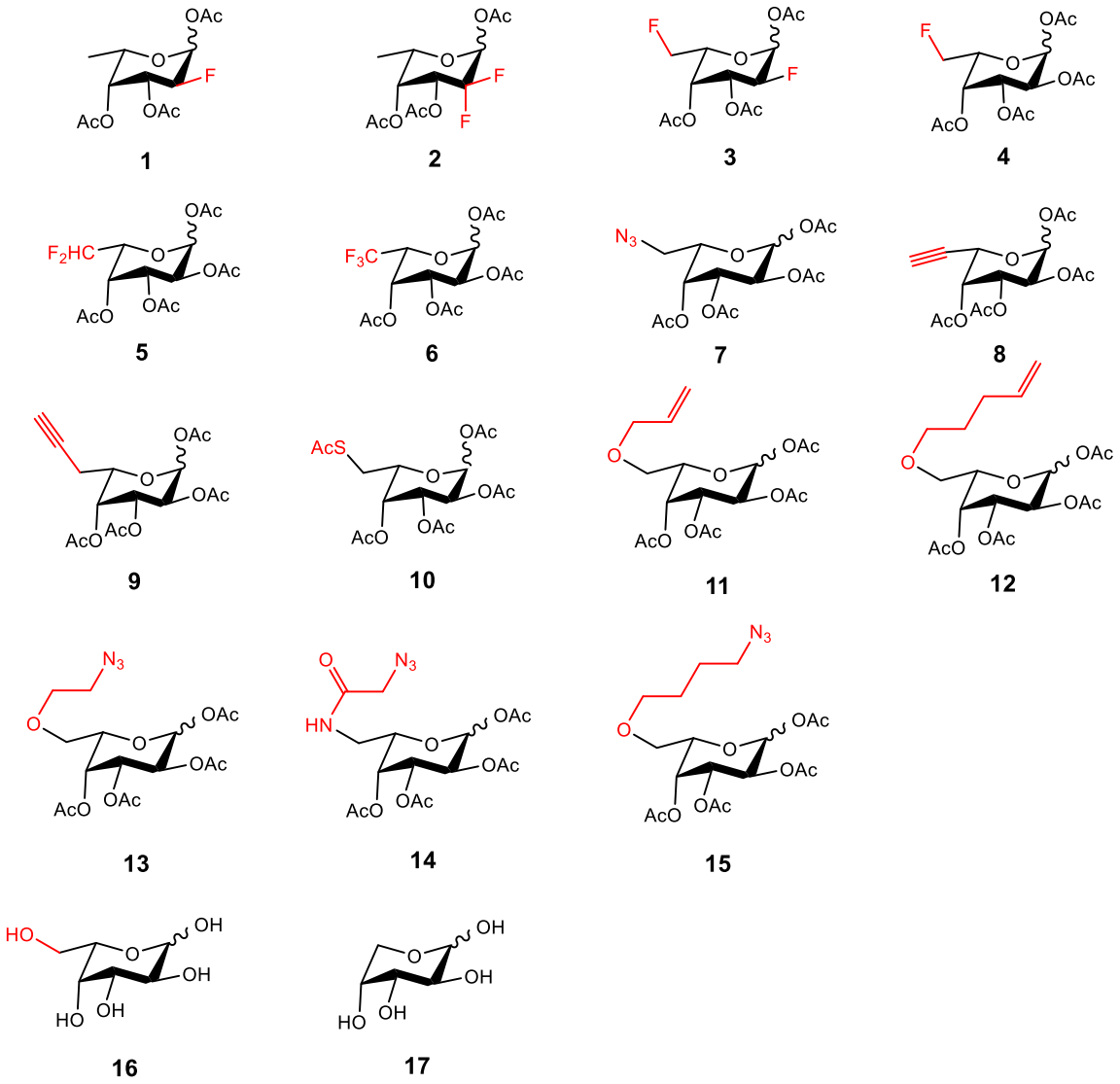


Figure 4.2. Panel of compounds tested for modulating core fucosylation of antibodies - 2F-Fuc (**1**), 2,2-diF-Fuc (**2**), 2,6-diF-Fuc (**3**), 6F-Fuc (**4**), 6,6-diF-Fuc (**5**), 6,6,6-triF-Fuc (**6**), 6-Az-Fuc (**7**), 6-Alkynyl-Fuc (**8**), 7-Alkynyl-Fuc (**9**), 6-Thio-Fuc (**10**), alkene-linked compounds (**11, 12**), azide-linked compounds (**13**, **14**, **15**), L-Galactose (**16**) and D-Arabinose (**17**)

4.2.2 Effect of fucose analogs on core fucosylation of antibodies

To evaluate the effect of the modified sugar analogs on the core fucosylation of recombinant antibodies, we chose Herceptin as a model antibody. Herceptin is a monoclonal antibody and contains two heavy chains and two light chains that are linked

together with disulfide bonds. The antibody is further classified into two fragments – the antigen binding fragment (Fab) and the crystallizable fragment (Fc) involved in generating effector functions. The Fc portion of the antibody contains an N-glycosylation site at N297 position on each of the heavy chains.

We expressed Herceptin in HEK293T and CHO-S cells through transient transfection and exposed the cells to the sugar analog to be tested for three to five days. We also expressed Herceptin in the presence of up to 1% (v/v) of dimethylsulfoxide (DMSO), the solvent used to dissolve the fucose analogs, as a control run. Purified antibodies were treated with IdeS to cleave below the antibody hinge region and generate IgG-Fc monomers to analyze the glycosylated Fc using LC-ESI-MS. To get an accurate estimate of potential analog incorporation into the Fc glycan, we performed N-glycan release using PNGase F. Purified N-glycans were then analyzed using MALDI-TOF-MS and relative quantities of glycoforms were calculated and compared. Figure 4.3 shows the deconvoluted LC-MS profile of antibody expressed in the presence of DMSO. The major glycoforms detected on the antibody Fc were G0F, G1F and G2F (core-fucosylated complex type glycan containing 0, 1 and 2 terminal galactose residues, respectively). Table 4.1 shows the summary of incorporation of compounds **1-17** into Herceptin at various concentrations.

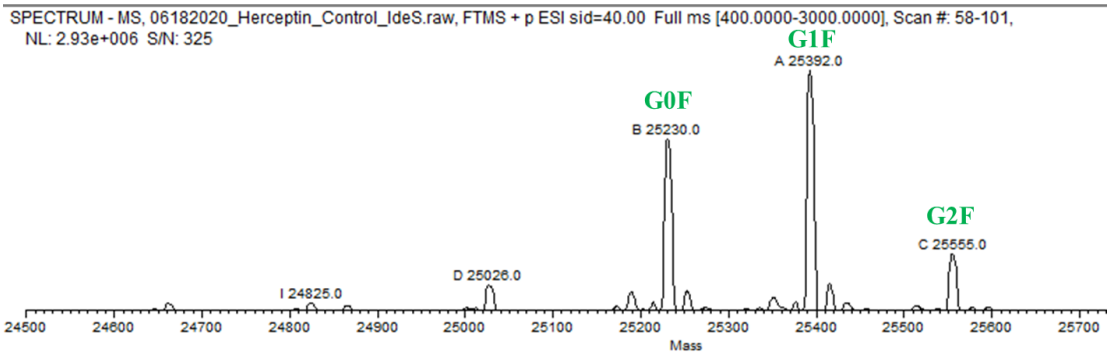


Figure 4.3. Deconvoluted LC-ESI-MS profile showing Fc monomer of IdeS-treated Herceptin expressed in the presence of DMSO.

Table 4.1. Efficiency of incorporation of L-fucose-based analogs into Herceptin at various concentrations

Compound #	Compound description	Incorporation (%)		
		at 0.1 mM	at 1 mM	at 2 mM
1	2F-Fuc	ND*	ND*	-
2	2,2-diF-Fuc	No effect	-	-
3	2,6-diF-Fuc	No effect	-	-
4	6F-Fuc	95	99	-
5	6,6-diF-Fuc	30	50	80
6	6,6,6-triF-Fuc	1	-	-
7	6-Az-Fuc	No effect	No effect	No effect
8	6-Alkynyl-Fuc	10	15	20
9	7-Alkynyl-Fuc	No effect	No effect	40
10	6-Thio-Fuc	-	99	-
11	Alkene-linker-Fuc 1	-	No effect	-
12	Alkene-linker-Fuc 2	-	No effect	Cytotoxic
13	Azide-linker-Fuc 1	-	No effect	-
14	Azide-linker-Fuc 2	-	No effect	No effect
15	Azide-linker-Fuc 3	-	No effect	-
16	L-Galactose	55 (10 mM)	55 (25 mM)	50 (50 mM)
17	D-Arabinose	30 (10 mM)	50 (25 mM)	45 (50 mM)

* ND - Not Detected by the mass spectrometry method used for detection. No further analysis was done to identify elemental composition
Hyphen '-' implies that the condition was not tested

No effect implies that the fucosylation as well as the overall glycosylation profiles were unaltered

4.2.2.1 Evaluation of fluorinated L-fucose analogs in HEK293T cells

Fluorinated fucose analogs have been reported to be good inhibitors of fucosyl transferases by previous studies, with a few exceptions. 2-deoxy-2-fluoro fucose (2F-Fuc) was shown to metabolically inhibit cell surface fucosylation by 90-95% in CHO-K1 cells[62]. Similarly, two independent studies showed complete[163] and 80% inhibition[165] of antibody fucosylation in CHO cells respectively. Our results show similar effects with 10-15% antibody core fucosylation at 0.1-1 mM of **1** in HEK293T cells (Figure 4.4). We did not notice incorporation of the analog in the antibody using MALDI-TOF-MS analysis of the released N-glycans (Table 4.1). However, it is important to note that the mass difference between the analog and L-fucose is only 2 Da which makes it hard to be detected. We expect the incorporation to be low as has been reported by other studies[62,163,165]. Also, in agreement with the previous report, we observed identical inhibition efficiency of the per-acetylated and non-protected 2F-Fuc sugars[163] (data not shown). However, cell permeability of sugar analogs may depend on the cell type used and on the sugar modifications.

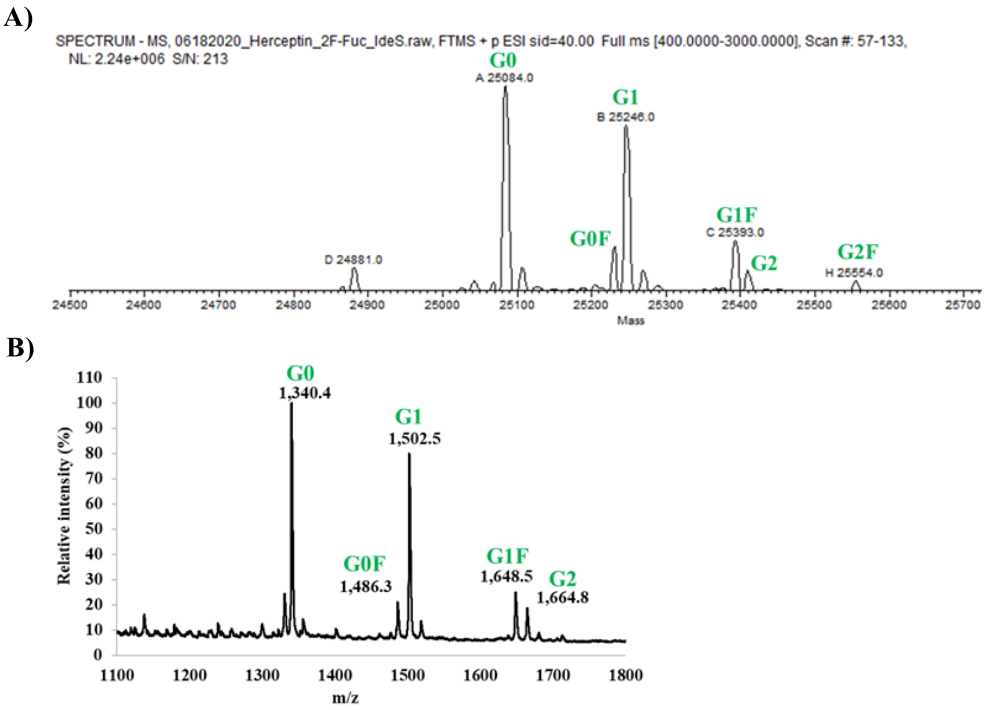


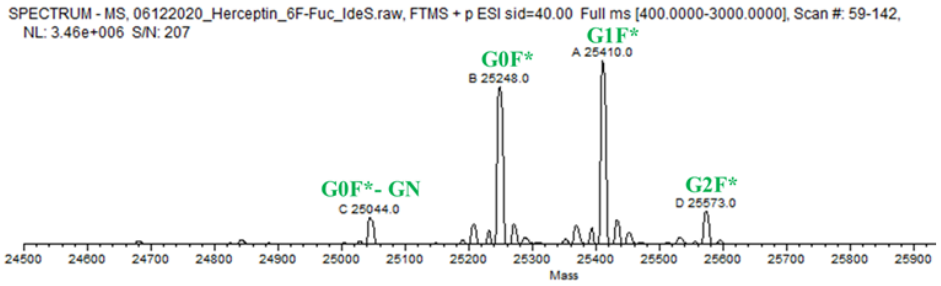
Figure 4.4. Herceptin expressed in the presence of 0.1 mM 2F-Fuc (**1**). A) Deconvoluted LC-ESI-MS profile showing Fc monomer (B) MALDI-TOF-MS profile of released N-glycan

Next, we tested the novel fluorinated L-fucose analogs 2-deoxy-2,2-difluoro fucose (2,2-diF-Fuc, **2**) and 2-deoxy-2,6-difluoro fucose (2,6-diF-Fuc, **3**) that were synthesized by the Wang group and previously tested for potential inhibition of cancer cell proliferation[173]. Interestingly, neither of these L-fucose analogs were found to have any effect on antibody fucosylation at the tested 0.1 mM concentration. A lack of effect may be indicative of low cell permeability or low efficiency of processing by the salvage pathway. Indeed, our previous study shows that **2** is a poor substrate of L-fucokinase/GDP-fucose pyrophosphorylase (FKP), the bacterial counterpart of the human enzyme that catalyzes the conversion of L-fucose to its corresponding GDP-fucose[173]. Additionally, GDP-2-deoxy-2,6-difluoro-L-fucose was shown to be a strong

inhibitor of FUT8[173]. Thus, it is unlikely that **3** would incorporate well into N-glycan core of target proteins.

Allen et.al. compared the 6-fluorinated L-fucose analogs namely, 6-fluoro fucose (6F-Fuc, **4**), 6,6-difluoro fucose (6,6-diF-Fuc, **5**) and 6,6,6-trifluoro fucose (6,6,6-triF-Fuc, **6**) for their inhibitory properties towards antibody core fucosylation[162]. Among the three analogs, **6** was found to be a highly potent inhibitor that showed minimal incorporation into the antibody. On the other hand, **4** and **5** showed core fucose inhibition with concomitant incorporation of the analog into the antibody, which is undesirable of a good inhibitor. To probe the incorporation efficiency of these fucose derivatives further, we performed a dose response analysis with **4** and **5** in the 0.1 – 2 mM concentration range. The efficiency of incorporation of **4** was found to be very high and a 0.1 mM addition shows almost complete replacement of core fucose with the fucose derivative (Table 4.1, Figure 4.5), which is consistent with the previous reports[162,171]. On the other hand, **5** requires higher concentrations for high incorporation. A dose response effect was observed with *c.a.* 80% incorporation of the analog at a concentration of 2 mM (Table 4.1, Figure 4.6). This data was encouraging as the highest incorporation reported in the previous study was 11%[162]. The incorporation of **6** was found to be minimal at 0.1 mM (Table 4.1, Figure 4.7), which agrees with the previous report[162].

A)



B)

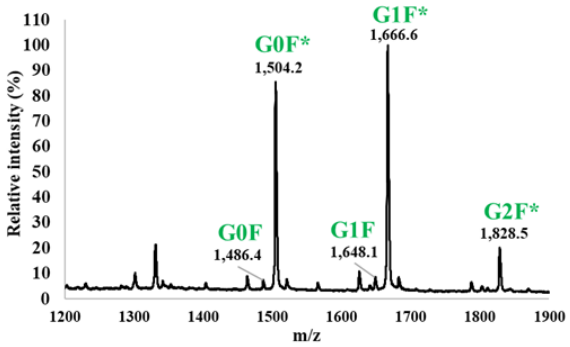


Figure 4.5. Herceptin expressed in the presence of 0.1 mM 6F-Fuc (**4**). A) Deconvoluted LC-ESI-MS profile showing Fc monomer (B) MALDI-TOF-MS profile of released N-glycan. F* indicates incorporation of fuocse analog in place of L-fucose.

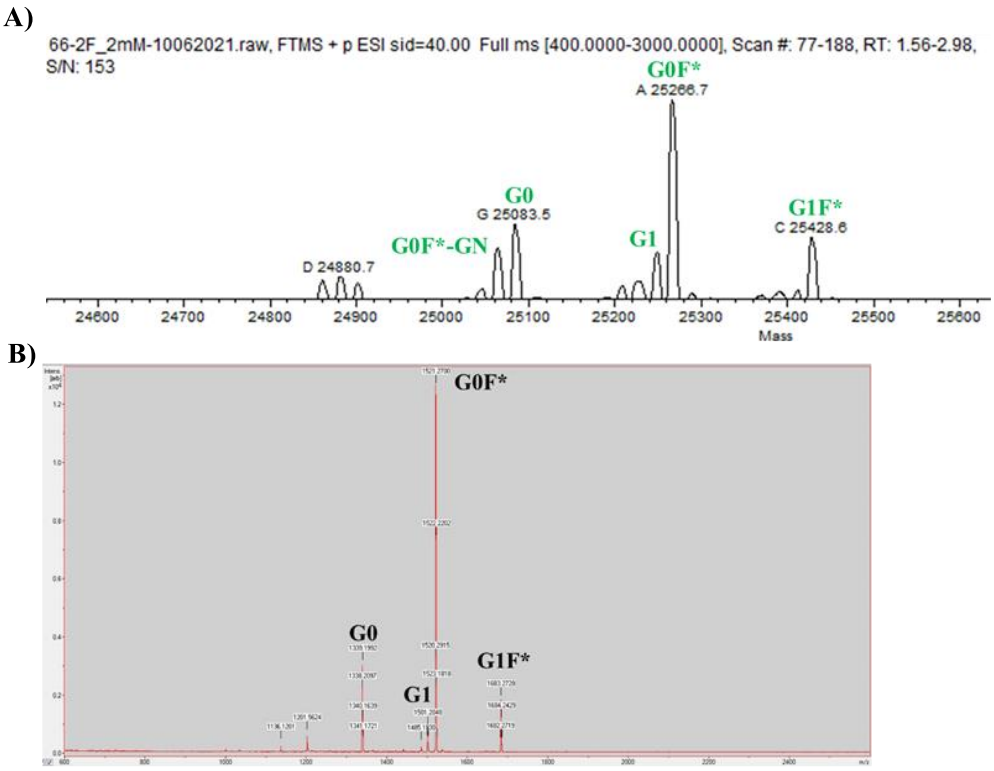


Figure 4.6. Herceptin expressed in the presence of 2 mM 6,6-diF-Fuc (**5**). A) Deconvoluted LC-ESI-MS profile showing Fc monomer (B) MALDI-TOF-MS profile of released N-glycan

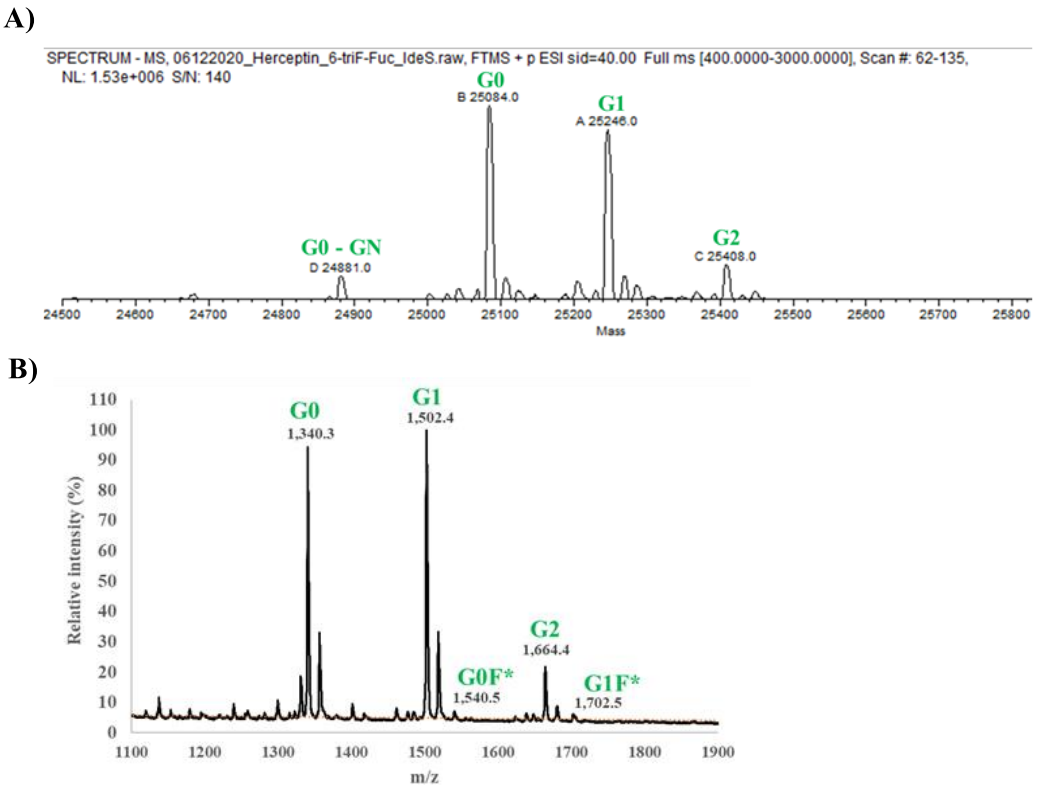


Figure 4.7. Herceptin expressed in the presence of 0.1 mM 6,6,6-triF-Fuc (**6**). A) Deconvoluted LC-ESI-MS profile showing Fc monomer (B) MALDI-TOF-MS profile of released N-glycan

4.2.2.2 Evaluation of azido-modified L-fucose analogs in HEK293T cells

Azido fucose was one of the first L-fucose analogs to be studied for labeling of cell surface glycans using click chemistry. One of the initial studies performed by Bertozzi and co-workers showed that only 6-Az-Fuc could label Jurkat cells at 0.1 mM concentration when tested alongside 2- and 4-azido fucose[166]. However, 6-Az-Fuc was reported to be toxic to the cells at the effective concentration. Also, Wong and co-workers reported cell surface labeling in Jurkat cells by 0.2 mM 6-Az-Fuc but observed cytotoxic effect on the cells[167,168]. Bertozzi and co-workers further investigated the efficiency of 6-Az-Fuc labeling by treating zebrafish embryos with 6-Az-Fuc-1-P and

GDP-6-Az-Fuc[174]. Incubation with GDP-6-Az-Fuc showed strong labeling unlike 6-Az-Fuc-1-P implying that the conversion of 6-Az-Fuc-1-P to GDP-6-Az-Fuc in the salvage pathway may be a limiting factor[174]. Further, a study conducted by Reinhold and co-workers showed that recombinant human FUT8 could use GDP-6-Az-Fuc as a substrate to label N-glycans in neuraminidase[175]. Taken together, site-specific labeling of proteins with azido sugar appears as an attractive strategy if the cytotoxicity of the analog can be reduced. Reduced cytotoxicity would allow higher concentrations to be tested to increase the potential intracellular GDP-6-Az-Fuc concentration and efficiency of labeling.

We sought to evaluate novel azide-modified L-fucose analogs (**13 – 15**) synthesized with varying lengths of spacers between azide and L-fucose at the C-6 position to potentially reduce cytotoxicity. We did not observe any effect of the analogs supplemented at 1 mM on the fucose content of the antibody. Moreover, increasing the amount of **14** to 2 mM also did not show any effect. Interestingly, no cytotoxicity was observed with any of the analogs at the tested concentrations (data not shown). We also performed a dose-response analysis of peracetylated 6-Az-Fuc (**7**) in the concentration range of 0.1 - 5 mM. Interestingly, our data shows that the fucose analog was not cytotoxic to cells till a high concentration of 1 mM (Figure 4.8). Gradual reductions in cell concentrations were recorded when the analog concentration was increased from 1 to 2 mM with high cytotoxicity at 5 mM. Yet, even at a high concentration of 2 mM, there was no effect of the analog on the antibody core fucosylation. A lack of effect by these analogs on the core fucosylation and overall glycosylation profile of the antibody

may indicate low cell permeability or low efficiency of processing by the enzymes involved in the salvage pathway.

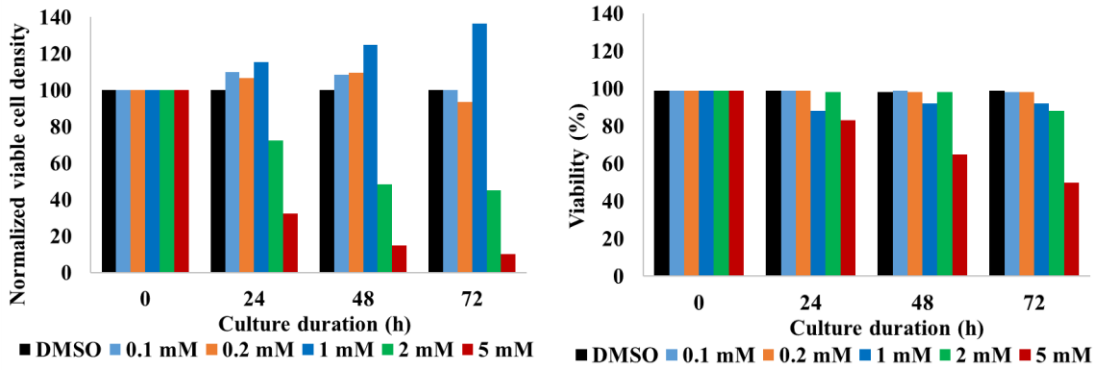


Figure 4.8. Dose-dependent effect of 6-Az-Fuc (**7**) on HEK293T cell growth. A) Viable cell density normalized with repsect to the control culture B) Cell viability

4.2.2.3 Evaluation of alkynylated L-fucose analogs in HEK293T cells

Another L-fucose analog that is known to metabolically inhibit cellular and protein fucosylation is 6-alkynylated L-fucose (**8**)[163-165]. A low amount of concomitant incorporation of the analog (3-5%) was reported by previous studies[163,165]. Our studies with **8** at a low concentration of 0.1 mM are consistent with the previous reports and the fucose derivative shows strong inhibition of antibody core fucosylation. However, the incorporation was seen to be slightly higher (10%) than the previous reports. To assess if the incorporation can be increased with higher amount of analog fed to the cells, we tested high concentrations of the analog. We supplemented the cultures with 1 and 2 mM of the peracetylated sugar and observed a rise in incorporation to 15 and 20% respectively (Table 4.1, Figure 4.9). This implies that the GDP-6-alkynyl fucose is not a strong inhibitor of FUT8 and instead can be used as a substrate. However, much

higher amounts of the analog may be required for good labeling which poses the risk of cytotoxicity to cells.

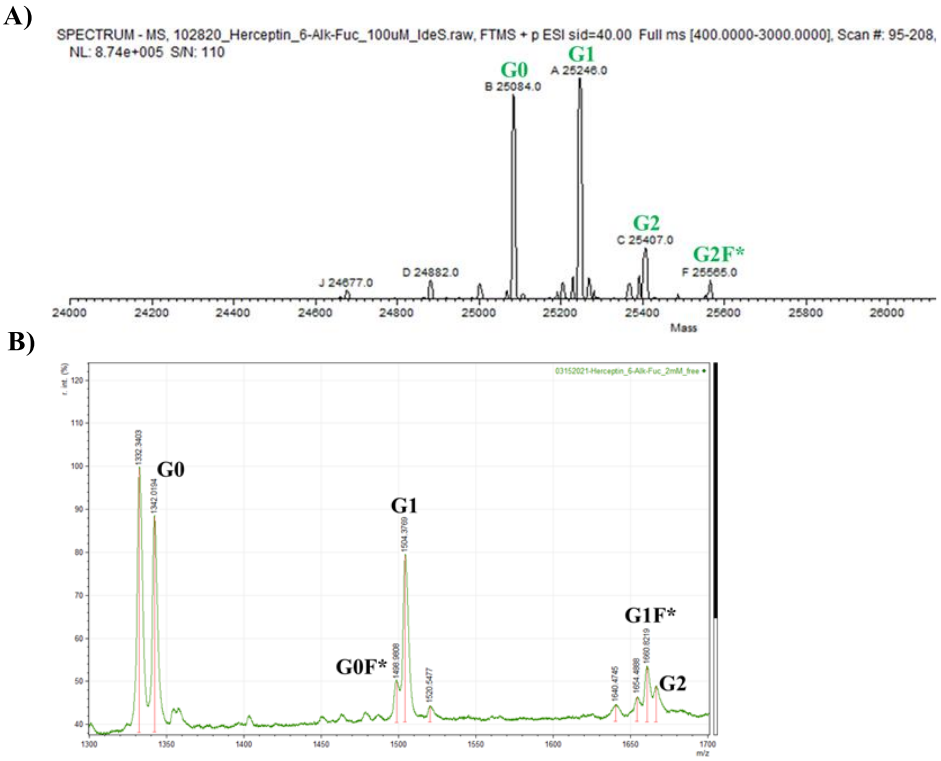


Figure 4.9. Herceptin expressed in the presence of 2 mM 6-Alk-Fuc **(8).** A) Deconvoluted LC-ESI-MS profile showing Fc monomer (B) MALDI-TOF-MS profile of released N-glycan

7-alkynyl fucose (**9**) was reported to be a sensitive incorporator in cell surface glycans[169]. And a more recent study reported the cell labeling efficiency to be highly variable depending on the cell line used[170]. The efficiency of labeling proteins secreted into medium was also shown to be very low[170]. We tested the analog **9** in a dose-dependent manner to probe its capability to label antibodies. Our observations show that there is no incorporation of the analog till 1 mM supplementation. However, increasing the concentration further to 2 mM shows partial labeling (40%) of antibody glycan (Table 4.1, Figure 4.10). It is important to note that most of the remaining

antibody N-glycan was core fucosylated implying that there is also a lack of strong inhibition by the analog. Moreover, trying to push the incorporation to completion with a 5 mM analog addition did not yield positive results due to cytotoxicity of the analog at such high concentrations (data not shown).

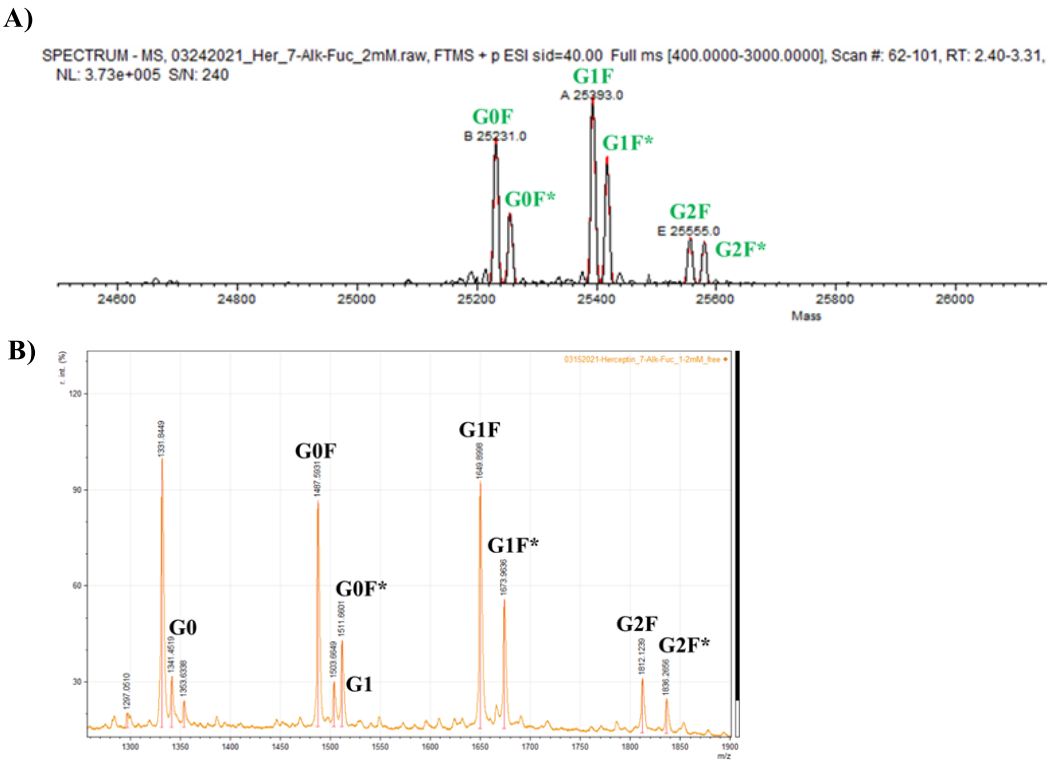


Figure 4.10. Herceptin expressed in the presence of 2 mM 7-Alk-Fuc **(9).** A) Deconvoluted LC-ESI-MS profile showing Fc monomer (B) MALDI-TOF-MS profile of released N-glycan

To assess the effect of size and complexity of the glycoprotein on the labeling outcome, we simultaneously tested incorporation of **9** into EPO at 0.5 mM in HEK293T cells. As with the antibody at concentrations less than 1 mM of the analog, there was no observation of incorporation or perturbation of core fucosylation in EPO. The bottleneck in robust labeling of proteins with 7-alkynyl fucose could possibly be the

formation of GDP-7-alkynyl fucose derivative in cells. Kizuka and coworkers measured the intracellular GDP-7-alkynyl fucose and GDP-Fuc concentration upon addition of 0.1 mM of **9** in HEK293 cells[169]. Their results show a high amount of GDP-Fuc and low levels of GDP-7-alkynyl fucose in the cells explaining the low efficiency of labeling. A similar estimation of intracellular nucleotide sugar concentrations at high feeding amounts may help in identifying the limiting step in the pathway.

*4.2.2.5 Evaluation of 6-thio-L-fucose (**10**) in HEK293T cells*

Peracetylated 6-thio-fucose (**10**) has been previously used to label antibodies expressed in CHO cells with a yield of 70%[171]. In our study, we evaluated the effect of supplementing 1 mM of **10** in HEK293T cells to label Herceptin. We observed high efficiency of labeling with almost complete incorporation of the thio-modified sugar into the antibody Fc (Table 4.1, Figure 4.11). This implies that the thio-modification at the C-6 position can be well tolerated by the salvage pathway and the fucose derivative is recognized as a substrate by FUT8. Moreover, the addition of **10** at 1 mM did not affect the cell growth or antibody production. Thus, the high efficiency of labeling and minimal cytotoxicity present 6-thio-fucose as a good incorporator in recombinant proteins, possibly in a cell-line independent manner.

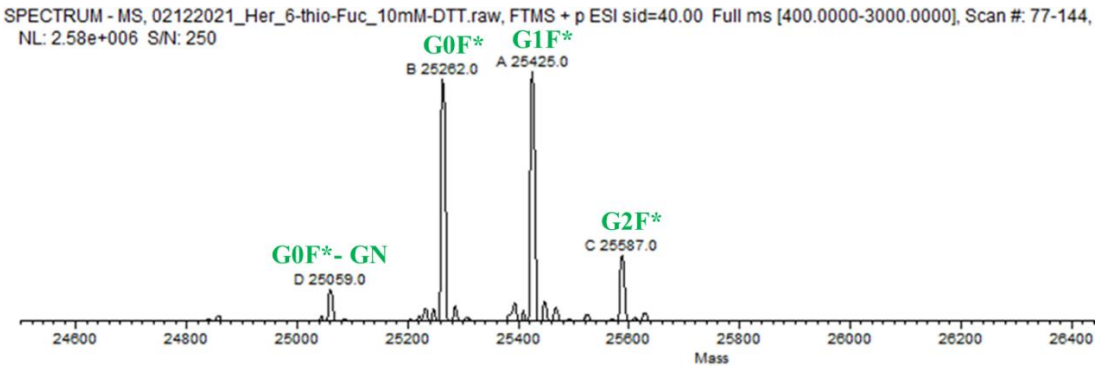


Figure 4.11. Deconvoluted LC-ESI-MS profile showing Fc monomer of Herceptin expressed in the presence of 1 mM 6-Thio-Fuc (**10**)

*4.2.2.6 Evaluation of compounds **11**, **12**, **16** and **17** in HEK293T cells*

As with the novel azide-modified analogs, we synthesized L-fucose analogs **11** and **12** with alkene group at the C-6 position linked by varying carbon chain lengths. When these compounds were tested for labeling of Herceptin at 1 mM concentration, no effect was observed on the core fucosylation of the antibody (Table 4.1). Further, compound **12** caused severe cytotoxicity to the cells at 2 mM concentration. Thus, the novel L-fucose analogs were found to be ineffective for protein labeling.

A recent study identified L-Galactose (**16**) and D-Arabinose (**17**) to be effective in labeling antibodies expressed in CHO cells[172]. Moreover, the incorporated D-Arabinose in place of L-fucose in the antibody N-glycan core did not seem to affect the overall properties of the antibody. Instead, a marginal improvement in FcγIIIa receptor binding and a significant improvement in ADCC were reported with arabinose incorporation[172].

Our aim was to test the two monosaccharides **16** and **17** for their incorporation efficiency and potentially use the labeled antibodies for functional assays. Our initial tests with **16** and **17** at 10 mM showed only partial incorporation of the monosaccharides into Herceptin (Table 4.1). Hence, we performed side-by-side analysis of both the sugars at 25- and 50-mM supplementation. Surprisingly, even at such high concentrations we achieved the same level of incorporation that was observed at 10 mM (Table 4.1, Figure 4.12) and there appeared to be a saturation in antibody labeling levels. These results were unlike the high efficiency of labeling reported by the previous study. However, upon careful examination of the results presented in the previous report, we noticed a similar plateau of incorporation at *c.a.* 50% in one of the cell lines expressing a bispecific antibody. Further studies to investigate a cell-line dependent effect of the monosaccharides are ongoing in CHO cells.

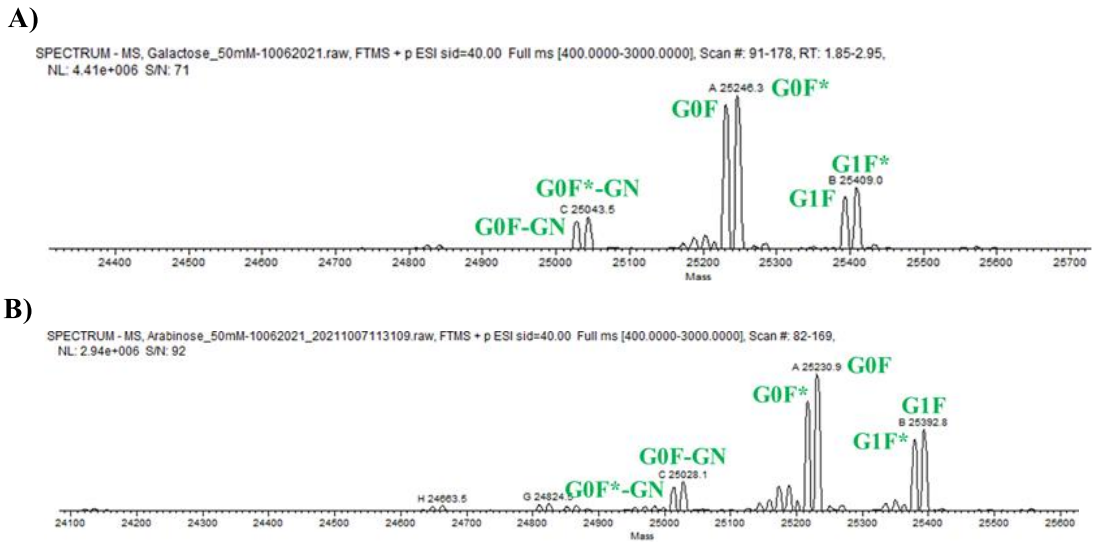


Figure 4.12. Deconvoluted LC-ESI-MS profile showing Fc monomer of Herceptin expressed in the presence of A) 50 mM L-Galactose **(16)** (B) 50 mM D-Arabinose **(17)**

4.2.2.7 Comparison of fucose analog effect in HEK293 and CHO cells

We chose HEK293T cells for our initial screening studies as transient transfection affords higher protein yields in HEK293T compared to CHO cell lines. However, since most previous studies have tested L-fucose analogs in CHO cells, we examined a few analogs side-by-side in CHO-S and HEK293T cells to validate our results listed in Table 4.1. We tested compounds **4**, **7**, **9**, **11** and **13** at the concentrations specified in Table 4.2. LC-ESI-MS profiles of the purified antibodies treated with IdeS show comparable results in both the cell lines with all the L-fucose analogs tested (Table 4.2). Among the sugar derivatives, only 6-fluoro fucose (**4**) showed an effect on core fucosylation of Herceptin in both cell lines. Taken together, the effect of various L-fucose analogs was seen to be similar in both the cell lines. This implies that the efficiency of fucose derivative incorporation reported in Table 4.1 is likely applicable to both the cell lines that are commonly used for recombinant glycoprotein expression.

Table 4.2. Comparison of L-fucose analog incorporation in Herceptin expressed in CHO-S and HEK293T cells

Compound #	**Compound**	**Test concentration (mM)**	**Incorporation (%)**	
			CHO-S	**HEK293T**
4	6F-Fuc	0.1	80	95
7	6-Az-Fuc	0.1	No effect	
9	7-Alkynyl-Fuc	1		
11	Alkene-linker-Fuc 1	1		
13	Azide-linker-Fuc 1	1		

4.3 Conclusion

A systematic side-by-side evaluation of L-fucose analogs to metabolically label therapeutic monoclonal antibodies is presented in this study. We synthesized a panel of compounds with bio-orthogonal groups at the C-6 position of L-fucose to test their ability to hijack the fucose salvage pathway and to be recognized as substrates by FUT8. Our results show that peracetylated L-fucose with 6-fluoro, 6,6-difluoro and 6-thio modifications can efficiently incorporate into the N-glycan core of antibodies. Peracetylated 7-alkynyl-fucose, L-galactose and D-arabinose showed 40-50% incorporation into Herceptin. On the other hand, 6-azido-fucose and novel analogs designed with carbon spacers between azide- or alkene- and L-fucose at C-6 failed to modulate the core fucosylation of antibodies. Moreover, the observations made with the fucose derivatives were found to be identical across HEK293 and CHO cells. The lack of effect by some of the fucose analogs on the fucosylation profile of the antibodies could be due to low cell permeability or low efficiency of processing by the enzymes involved in the salvage pathway. Further studies will have to be taken up for investigated the results.

4.4 Experimental

4.1 Material

CHO-S cells were purchased from Thermo Fisher Scientific. L-Galactose and D-Arabinose were purchased from Biosynth Carbosynth and Millipore Sigma respectively.

4.2 Herceptin gene constructs

The coding sequences of Herceptin heavy and light chains were retrieved from the DrugBank online public database (DrugBank Accession Number - DB00072). Signal peptides for the two chains were used based on optimized sequences reported by Haryadi et. al[176]. The heavy and light chains containing the N-terminal signal peptides MDWTWRFLFVVAAATGVQS and MDMRVPAQLLGLLLLWLSGARC respectively were inserted into separate pcDNA3.1 vectors using NheI/XhoI sites. The codon-optimized synthetic genes were purchased from GenScript.

4.3 Expression and purification of antibodies

Recombinant Herceptin was transiently expressed in HEK293T or CHO suspension cultures. The cells were seeded at a density of 12 x 10^5/ml in serum-free FreeStyle™ F17 Expression Medium (Thermo Fisher Scientific) and cultured at 37°C, 8% CO2, 150 rpm. After 20 – 24 h, the culture was transiently transfected with a total of 2 µg/mL DNA (Heavy chain : Light chain = 1:3) and 6 µg/mL polyethylenimine (PEI). The expression culture was incubated at 37°C, 8% CO2, 150 rpm for 6 days. The cultures were fed with L-glucose and GlutaMAX™ (Thermo Fisher Scientific) at regular intervals to maintain the cells in a healthy state. At harvest, the broth was spun down at 1000 rpm for 10 minutes at 4°C and the supernatant was purified using Protein A chromatography (GE Healthcare). The antibody product bound to the Protein A column was eluted out with 0.1 M Citrate buffer, pH – 3.3. The eluted protein was immediately neutralized with Tris-HCl buffer, pH – 8.5 and buffer exchanged to PBS, pH-7.4 using

Amicon spin columns. The purified antibody was characterized using non-reduced and reduced gel electrophoresis and mass spectrometric (MS) analysis.

4.4 Testing *in vivo* effect of fucose analogs on antibodies

L-fucose analogs were dissolved in dimethyl sulfoxide (DMSO) to prepare stock solutions for evaluating effects on antibody fucosylation. The fucose analog solutions were added to the cells 2-3 h post-transfection at desired concentrations. Cell growth, antibody production and glycosylation profiles were monitored to assess the effect of analog supplementation. Viable cell densities were measured manually using Trypan blue dye exclusion method. Antibody yields were calculated by measuring the absorbance of purified antibodies at 280 nm. Glycosylation profiles were studied using mass spectrometric methods.

4.5 SDS-PAGE sample preparation and analysis

Reduced sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed by dissolving about 1 µg of purified antibody in 2X SDS loading dye containing β-mercaptoethanol and denaturing the sample at 95^{o}C for 5 minutes followed by cooling to 4^{o}C. The denatured samples were loaded onto precast polyacrylamide stain-free gels (Mini-PROTEAN™ TGX Stain-Free™ Protein Gels, Bio-Rad, Cat. # 4568086, 4568106) along with a protein ladder (Precision Plus Protein Standards, Bio-Rad, Cat. # 1610363) for reference. Electrophoresis was performed in 1X Tris-Glycine-SDS buffer at 220V for 35 minutes. Gel imaging was done using Gel

Doc™ EZ Imager (Bio-Rad). For non-reduced SDS-PAGE, the disulfide bonds were maintained intact by preparing the samples in a 2X SDS loading dye without the reducing agent, β-mercaptoethanol.

4.6 Mass spectrometric analysis of antibodies

Intact antibody analysis was performed using liquid chromatography (LC)-MS analysis on a Thermo Scientific Exactive™ Plus Orbitrap mass spectrometer on a Waters XBridge™ BEH300 C4 column (3.5 µm, 2.1 x 50 mm) using a 9-min linear gradient of 5 - 90% acetonitrile containing 0.1% formic acid at a flow rate of 0.4 mL/min. To analyze the antibody Fc containing the N-glycosylation site, antibodies were treated with IdeS protease (0.02 mg/ml) in PBS at 37°C for 15 min to generate Fab dimer and Fc monomer fragments. Antibody fragment analysis was performed on a Waters XBridge™ BEH300 C8 column (3.5 µm, 2.1 x 50 mm) using a 6-min linear gradient of 25 - 35% acetonitrile containing 0.1% formic acid at a flow rate of 0.4 mL/min. MS raw data was deconvoluted using MagTran (Amgen). Further, to analyze the released N-glycans, MALDI-TOF-MS analysis was employed. Peptide: N-glycosidase F (PNGase F) was used to cleave N-glycans from the antibody Fc under denaturing conditions as recommended by NEB. The released N-glycans were purified using HyperSep™ Hypercarb™ SPE cartridges using recommended procedure (Thermo Scientific). The eluted samples were lyophilized and analyzed by MALDI-TOF-MS in linear positive mode using 2,5-dihydroxybenzoic acid (DHB) matrix. Glycan assignment was done using GlycoWorkbench.

Chapter 5: Conclusions and Future directions

N-glycosylation of proteins is critical in regulating the structure and function of proteins including therapeutic glycoproteins. Homogeneous glycosylation is important for the clinical outcome of therapeutic glycoproteins and for performing mechanistic studies. Numerous methodologies have been explored to generate glycoproteins with desired glycoforms. In my research, I have investigated biological as well as chemoenzymatic approaches to generate homogeneously glycosylated eukaryotic proteins. The research work is outlined in three main chapters in the dissertation.

In the first study, we designed an *E. coli*-based glycosylation system. A two-step approach was used to glucosylate IFNα by co-expressing an N-glycosyl transferase enzyme in *E. coli*. The glucosylated-IFNα was then used for a subsequent *in vitro* transglycosylation reaction catalyzed by an endoglycosynthase EndoCC-N180H. Using this method, we produced IFNα containing a homogeneous biantennary sialylated complex type glycan with glucose as the reducing end sugar. Further, we evaluated the anti-proliferative property of the glycosylated IFNα and compared it with the glucosylated counterpart and an IFNα reference standard. The biological activity of the glycosylated protein was found to be marginally lower. However, this was compensated by an increase in the proteolytic stability of the protein assessed via resistance to trypsin digestion. While we successfully demonstrated a proof-of-concept of the two-step method to produce biologically active eukaryotic glycoprotein *using E. coli*, there are some limitations and scope of improvement for future applications. Firstly, ApNGT has strict acceptor peptide/protein preferences which necessitates

mutagenesis of target proteins to introduce desired acceptor sequon into the protein[97]. This limits efficient glucosylation of natural protein substrates. Moreover, it has been observed in two independent studies[97], including ours, that the native ApNGT prefers partially or fully unfolded proteins to transfer glucose efficiently. This poses a challenge in achieving completely glucosylated proteins *in vivo*. In our study, we compared three different expression methodologies to optimize the ApNGT-catalyzed glucosylation of IFNα. While the glucose transfer to GST-tagged IFNα was complete, the glucosylation efficiency of IFNα expressed in inclusion bodies and in the soluble form with a His-tag was only partially complete. ApNGT Q469A mutant[93], which is shown to have a broader substrate specificity, was tested in some of these cases but did not improve the glucosylation efficiency significantly. Thus, to universally apply ApNGT-catalyzed glucosylation to eukaryotic proteins in *E. coli*, further assessment of ApNGT substrate specificity is warranted. Furthermore, characterization of NGTs from other species for glucosylation of complex protein substrates needs to be evaluated. A more challenging and necessary improvement is the donor substrate specificity of the NGTs. The well characterized ApNGT and other lesser studied NGTs[96] currently cannot accept UDP-GlcNAc as a donor substrate and thus, cannot transfer GlcNAc to acceptor substrates. Hence, the N-glycans transferred to the proteins, as shown in our study, are attached by a non-natural linkage to the Asn in proteins. In the context of therapeutics, this can pose immunogenicity-related issues. A structure-based mutagenesis of the native ApNGT and eventually other NGTs will be required to introduce GlcNAc transferring ability into these enzymes.

In the second study, we performed a comparative analysis of α-L-fucosidases from human (FucA1), bacteria - *Lactobacillus casei* (AlfC) and *Bacteroides fragilis* (BfFuc) for their ability to hydrolyze core fucose from substrates. We synthesized a panel of structurally well-defined core-fucosylated glycopeptide and glycoprotein substrates of varying complexities. Our side-by-side analysis with the three enzymes shows that all three enzymes have distinct substrate specificities and relative activities. AlfC and BfFuc were very efficient in hydrolyzing core fucose from peptide and protein substrates containing only Fuca-1,6GlcNAc. Both the bacterial enzymes did not display any hydrolytic activity towards full-length glycans in the context of glycopeptides/glycoproteins and even as free N-glycans. To our surprise and delight, FucA1 showed appreciable activity towards core fucose in intact glycopeptides and glycoproteins containing full-length complex type glycans. Further, FucA1 showed good activity towards monoclonal antibody containing core-fucosylated Man5 glycan and marginal but evident activity towards antibody containing core fucosylated complex type glycan. This is the first account of core fucosidase activity towards intact monoclonal antibodies and glycoproteins. This discovery paves way to developing an efficient enzyme that can be used to directly modulate the core fucose content in antibodies *in vitro* and potentially apply to therapeutic antibodies to improve their ADCC. The key limitation currently is the low efficiency of the enzyme. The results summarized above pertaining to glycoproteins and antibodies were generated by using large quantities of FucA1 and incubating reaction mixtures for long durations, up to several days in some cases. While the ability of FucA1 to hydrolyze intact glycoproteins and antibodies is exciting, the application of this enzyme for glycan

remodeling purposes is not feasible. Apart from the uneconomical set-up of the reaction mixtures, the long incubations of the substrate proteins under the harsh reaction conditions (pH - 4.5) can be detrimental to the integrity and activity of the proteins. Thus, to develop an efficient fucosidase that can hydrolyze intact proteins and antibodies in quantitative yields, enzyme engineering needs to be assessed. Use of directed evolution or structure-informed mutagenesis studies will be required to develop a FucA1 mutant with higher efficiencies. Such structure-based mutagenesis studies would benefit from the crystal structure of the FucA1 along with homology modeling with closely related fucosidases. Additionally, kinetic evaluation of the catalytic activity of FucA1 with different protein substrates and glycoforms to identify substrate affinity and catalytic turn over will be useful.

Finally, we attempted to modulate the core fucosylation of monoclonal antibodies by using L-fucose analogs. We synthesized an array of L-fucose derivatives containing various bio-orthogonal groups in the C-6 position of L-fucose including fluorine atoms, azide- , alkynyl- , alkene- and thio- modifications. This research was performed in collaboration with a colleague Dr. Yuanwei Dai in the Wang lab. The design and synthesis of compounds was performed by Dr. Dai. The main objective of this study was to identify fucose analogs that can be incorporated into the antibody N-glycan core to functionalize the antibody for further applications like cell surface imaging, antibody-drug conjugate design etc. We evaluated several novel compounds as well as some known fucose analogs for their incorporation efficiency. Our results show that it is tricky to identify novel compounds that can efficiently enter the fucose metabolic pathway and be accepted as a substrate by the FUT8 enzyme to be added to target

proteins. Our results also confirmed some of the previous observations reported in literature whereas in some cases, the reported analogs were found to potentially work in a cell-line dependent and context-specific manner. For example, the compound peracetylated 7-alkynyl fucose was reported to label cellular proteins efficiently but not the recombinant proteins[169]. Our results show that the compound did not have any effect on antibody fucosylation till a high concentration of 1 mM. But at 2 mM addition to the cell culture, about 40% incorporation into antibody was observed. Similarly, 6-Az-Fuc, which is known to cause cytotoxicity in Jurkat cells in the 0.1 – 0.2 mM range[166-168], did not show any cytotoxicity in HEK293T cells up to 1 mM. Yet, the compound did not show any effect on the core fucosylation of antibodies. In summary, three compounds 6F-Fuc, 6,6-diF-Fuc and 6-thio-Fuc show high yield of incorporation into monoclonal antibodies without causing cytotoxicity. Out of the three compounds, only 6-thio-fucose can be utilized to further modify the incorporated fucose derivative. Further assessments are required to identify other modifications that hijack the salvage pathway of fucose metabolism to efficiently label recombinant proteins. Identification of the bottleneck in the metabolic pathway for the compounds that show low efficiency will also be necessary. The deprotected sugar analog that enters the salvage pathway first gets converted to a sugar analog-1-phosphate that is then converted to GDP-analog[62]. The sugar analogs designed for labeling must be good substrates of both these enzymes to form enough GDP-analog in the cells to feedback inhibit formation of GDP-fucose. The next requirement is that the GDP-analog must be a good substrate of the FUT8 enzyme to be effectively incorporated into protein substrates.

In summary, the approaches outlined in this dissertation enable production of therapeutic glycoproteins carrying desired glycan structures. The *E. coli*-based glycosylation method offers a feasible strategy for eukaryotic as well as prokaryotic protein production with homogenous glycoforms for structure-function studies or for therapeutic application post some improvements. The human core fucosidase presents a promising candidate for therapeutic as well as research-related core fucose modification of glycoproteins and antibodies. Similarly, generating a functionalized antibody using metabolic engineering can be useful to design antibody-based reagents and therapeutic candidates.

Bibliography

1. O'Connor SE, Imperiali B. Modulation of protein structure and function by asparagine-linked glycosylation. Chemistry & Biology. 1996;3(10):803-812.

2. Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology. 2002;12(4):43R-56R.

3. Weerapana E, Imperiali B. Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. Glycobiology. 2006;16(6):91R-101R.

4. Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. Annu Rev Biochem. 2011;80:825-858.

5. Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol. 2012;13(7):448-462.

6. Varki A, Cummings RD, Esko JD, et al. Essentials of Glycobiology. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.; 2015.

7. Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. Nat Rev Nephrol. 2019;15(6):346-366.

8. Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. Cell. 2006;126(5):855-867.

9. Dwek RA. Glycobiology: Toward Understanding the Function of Sugars. Chem Rev. 1996;96(2):683-720.

10. Helenius A, Aebi M. Intracellular functions of N-linked glycans. Science. 2001;291(5512):2364-2369.

11. Xu C, Ng DT. Glycosylation-directed quality control of protein folding. Nat Rev Mol Cell Biol. 2015;16(12):742-752.

12. Caramelo JJ, Parodi AJ. Getting in and out from calnexin/calreticulin cycles. J Biol Chem. 2008;283(16):10221-10225.

13. Varki A. Biological roles of oligosaccharides: all of the theories are correct. Glycobiology. 1993;3(2):97-130.

14. Varki A. Biological roles of glycans. Glycobiology. 2017;27(1):3-49.

15. Dwek RA, Butters TD, Platt FM, Zitzmann N. Targeting glycosylation as a therapeutic approach. Nat Rev Drug Discov. 2002;1(1):65-75.

16. Haltiwanger RS, Lowe JB. Role of glycosylation in development. Annu Rev Biochem. 2004;73:491-537.

17. Nimmerjahn F, Ravetch JV. Fcgamma receptors as regulators of immune responses. Nat Rev Immunol. 2008;8(1):34-47.

18. Jefferis R. Glycosylation as a strategy to improve antibody-based therapeutics. Nat Rev Drug Discov. 2009;8(3):226-234.

19. Walsh G, Jefferis R. Post-translational modifications in the context of therapeutic proteins. Nat Biotechnol. 2006;24(10):1241-1252.

20. Taylor ES, Pol-Fachin L, Lins RD, Lower SK. Conformational stability of the epidermal growth factor (EGF) receptor as influenced by glycosylation, dimerization and EGF hormone binding. Proteins. 2017;85(4):561-570.
21. Coutinho MF, Prata MJ, Alves S. Mannose-6-phosphate pathway: a review on its role in lysosomal function and dysfunction. Mol Genet Metab. 2012;105(4):542-550.
22. Gary-Bobo M, Nirdé P, Jeanjean A, Morère A, Garcia M. Mannose 6-phosphate receptor targeting and its applications in human diseases. Curr Med Chem. 2007;14(28):2945-2953.
23. Seeberger PH, Cummings RD. Glycans in Biotechnology and the Pharmaceutical Industry. In: Varki A, Cummings RD, Esko JD, et al. Essentials of Glycobiology. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press;Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.; 2015.
24. Dalziel M, Crispin M, Scanlan CN, Zitzmann N, Dwek RA. Emerging principles for the therapeutic exploitation of glycosylation. Science. 2014;343(6166):1235681.
25. Darling RJ, Kuchibhotla U, Glaesner W, Micanovic R, Witcher DR, Beals JM. Glycosylation of erythropoietin affects receptor binding kinetics: role of electrostatic interactions. Biochemistry. 2002;41(49):14524-14531.
26. Solá RJ, Griebenow K. Effects of glycosylation on the stability of protein pharmaceuticals. J Pharm Sci. 2009;98(4):1223-1245.
27. Sharon N. Lectins: carbohydrate-specific reagents and biological recognition molecules. J Biol Chem. 2007;282(5):2753-2764.
28. Gagneux P, Varki A. Evolutionary considerations in relating oligosaccharide diversity to biological function. Glycobiology. 1999;9(8):747-755.
29. Mikulak J, Di Vito C, Zaghi E, Mavilio D. Host Immune Responses in HIV-1 Infection: The Emerging Pathogenic Role of Siglecs and Their Clinical Correlates. Front Immunol. 2017;8:314.
30. von Itzstein M, Wu WY, Kok GB, et al. Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature. 1993;363(6428):418-423.
31. Lew W, Chen X, Kim CU. Discovery and development of GS 4104 (oseltamivir): an orally active influenza neuraminidase inhibitor. Curr Med Chem. 2000;7(6):663-672.
32. Zamyatina A, Heine H. Lipopolysaccharide Recognition in the Crossroads of TLR4 and Caspase-4/11 Mediated Inflammatory Pathways. Front Immunol. 2020;11:585146.
33. Dube DH, Bertozzi CR. Glycans in cancer and inflammation--potential for therapeutics and diagnostics. Nat Rev Drug Discov. 2005;4(6):477-488.
34. Taniguchi N, Kizuka Y. Glycans and cancer: role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics. Adv Cancer Res. 2015;126:11-51.
35. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer. 2015;15(9):540-555.

36. Zhu H, Rollier CS, Pollard AJ. Recent Advances in Lipopolysaccharide-Based Glycoconjugate Vaccines. Expert Rev Vaccines. 2021.
37. Medina-Ramírez M, Sanders RW, Sattentau QJ. Stabilized HIV-1 envelope glycoprotein trimers for vaccine use. Curr Opin HIV AIDS. 2017;12(3):241-249.
38. Asín A, García-Martín F, Busto JH, Avenoza A, Peregrina JM, Corzana F. Structure-based Design of Anti-cancer Vaccines: The Significance of Antigen Presentation to Boost the Immune Response. Curr Med Chem. 2021.
39. Li C, Wang LX. Chemoenzymatic Methods for the Synthesis of Glycoproteins. Chem Rev. 2018;118(17):8359-8413.
40. Roth J. Protein N-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. Chem Rev. 2002;102(2):285-303.
41. Stanley P, Taniguchi N, Aebi M. N-Glycans. In: Varki A, Cummings RD, Esko JD, et al. Essentials of Glycobiology. Cold Spring Harbor (NY)2015.
42. Vasconcelos-Dos-Santos A, Oliveira IA, Lucena MC, et al. Biosynthetic Machinery Involved in Aberrant Glycosylation: Promising Targets for Developing of Drugs Against Cancer. Front Oncol. 2015;5:138.
43. Sola RJ, Griebenow K. Glycosylation of therapeutic proteins: an effective strategy to optimize efficacy. BioDrugs. 2010;24(1):9-21.
44. Shields RL, Lai J, Keck R, et al. Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. J Biol Chem. 2002;277(30):26733-26740.
45. Goetze AM, Liu YD, Zhang Z, et al. High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans. Glycobiology. 2011;21(7):949-959.
46. Hodoniczky J, Zheng YZ, James DC. Control of recombinant monoclonal antibody effector functions by Fc N-glycan remodeling in vitro. Biotechnol Prog. 2005;21(6):1644-1652.
47. Washburn N, Schwab I, Ortiz D, et al. Controlled tetra-Fc sialylation of IVIg results in a drug candidate with consistent enhanced anti-inflammatory activity. Proc Natl Acad Sci U S A. 2015;112(11):E1297-1306.
48. Wang LX, Lomino JV. Emerging technologies for making glycan-defined glycoproteins. ACS Chem Biol. 2012;7(1):110-122.
49. Donini R, Haslam SM, Kontoravdi C. Glycoengineering Chinese hamster ovary cells: a short history. Biochem Soc Trans. 2021;49(2):915-931.
50. Yamane-Ohnuki N, Kinoshita S, Inoue-Urakubo M, et al. Establishment of FUT8 knockout Chinese hamster ovary cells: an ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity. Biotechnol Bioeng. 2004;87(5):614-622.
51. Ferrara C, Brünker P, Suter T, Moser S, Püntener U, Umaña P. Modulation of therapeutic antibody effector functions by glycosylation engineering: influence of Golgi enzyme localization domain and co-expression of heterologous beta1, 4-N-acetylglucosaminyltransferase III and Golgi alpha-mannosidase II. Biotechnol Bioeng. 2006;93(5):851-861.

52. Imai-Nishiya H, Mori K, Inoue M, et al. Double knockdown of alpha1,6-fucosyltransferase (FUT8) and GDP-mannose 4,6-dehydratase (GMD) in antibody-producing cells: a new strategy for generating fully non-fucosylated therapeutic antibodies with enhanced ADCC. BMC Biotechnol. 2007;7:84.
53. Ureshino H, Kamachi K, Kimura S. Mogamulizumab for the Treatment of Adult T-cell Leukemia/Lymphoma. Clin Lymphoma Myeloma Leuk. 2019;19(6):326-331.
54. Sharman JP, Farber CM, Mahadevan D, et al. Ublituximab (TG-1101), a novel glycoengineered anti-CD20 antibody, in combination with ibrutinib is safe and highly active in patients with relapsed and/or refractory chronic lymphocytic leukaemia: results of a phase 2 trial. Br J Haematol. 2017;176(3):412-420.
55. Raymond C, Robotham A, Spearman M, Butler M, Kelly J, Durocher Y. Production of α2,6-sialylated IgG1 in CHO cells. MAbs. 2015;7(3):571-583.
56. Chiba Y, Jigami Y. Production of humanized glycoproteins in bacteria and yeasts. Curr Opin Chem Biol. 2007;11(6):670-676.
57. Jacobs PP, Geysens S, Vervecken W, Contreras R, Callewaert N. Engineering complex-type N-glycosylation in Pichia pastoris using GlycoSwitch technology. Nat Protoc. 2009;4(1):58-70.
58. Schwarz F, Huang W, Li C, et al. A combined method for producing homogeneous glycoproteins with eukaryotic N-glycosylation. Nat Chem Biol. 2010;6(4):264-266.
59. Keys TG, Wetter M, Hang I, et al. A biosynthetic route for polysialylating proteins in Escherichia coli. Metab Eng. 2017;44:293-301.
60. Valderrama-Rincon JD, Fisher AC, Merritt JH, et al. An engineered eukaryotic protein glycosylation pathway in Escherichia coli. Nat Chem Biol. 2012;8(5):434-436.
61. Elbein AD. Glycosidase inhibitors: inhibitors of N-linked oligosaccharide processing. Faseb j. 1991;5(15):3055-3063.
62. Rillahan CD, Antonopoulos A, Lefort CT, et al. Global metabolic inhibitors of sialyl- and fucosyltransferases remodel the glycome. Nat Chem Biol. 2012;8(7):661-668.
63. Gloster TM, Zandberg WF, Heinonen JE, Shen DL, Deng L, Vocadlo DJ. Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells. Nat Chem Biol. 2011;7(3):174-181.
64. Fernández-Tejada A, Brailsford J, Zhang Q, Shieh JH, Moore MA, Danishefsky SJ. Total synthesis of glycosylated proteins. Top Curr Chem. 2015;362:1-26.
65. Chaffey PK, Guan X, Li Y, Tan Z. Using Chemical Synthesis To Study and Apply Protein Glycosylation. Biochemistry. 2018;57(4):413-428.
66. Gamblin DP, Scanlan EM, Davis BG. Glycoprotein synthesis: an update. Chem Rev. 2009;109(1):131-163.
67. Wang P, Dong S, Shieh JH, et al. Erythropoietin derived by chemical synthesis. Science. 2013;342(6164):1357-1360.
68. Danby PM, Withers SG. Advances in Enzymatic Glycoside Synthesis. 2016.

69. Fairbanks AJ. The ENGases: versatile biocatalysts for the production of homogeneous N-linked glycopeptides and glycoproteins. Chem Soc Rev. 2017;46(16):5128-5146.
70. Boltje TJ, Buskas T, Boons GJ. Opportunities and challenges in synthetic oligosaccharide and glycoconjugate research. Nat Chem. 2009;1(8):611-622.
71. Cobucci-Ponzano B, Moracci M. Glycosynthases as tools for the production of glycan analogs of natural products. 2012.
72. Jahn M, Stoll D, Warren RA, et al. Expansion of the glycosynthase repertoire to produce defined manno-oligosaccharides. Chem Commun (Camb). 2003(12):1327-1329.
73. Jakeman DL, Sadeghi-Khomami A. A β-(1,2)-glycosynthase and an attempted selection method for the directed evolution of glycosynthases. Biochemistry. 2011;50(47):10359-10366.
74. Li C, Zhu S, Ma C, Wang LX. Designer alpha1,6-Fucosidase Mutants Enable Direct Core Fucosylation of Intact N-Glycopeptides and N-Glycoproteins. J Am Chem Soc. 2017;139(42):15074-15087.
75. Wang LX, Amin MN. Chemical and chemoenzymatic synthesis of glycoproteins for deciphering functions. Chem Biol. 2014;21(1):51-66.
76. Li B, Zeng Y, Hauser S, Song H, Wang LX. Highly efficient endoglycosidase-catalyzed synthesis of glycopeptides using oligosaccharide oxazolines as donor substrates. J Am Chem Soc. 2005;127(27):9692-9693.
77. Rising TW, Claridge TD, Davies N, Gamblin DP, Moir JW, Fairbanks AJ. Synthesis of N-glycan oxazolines: donors for endohexosaminidase catalysed glycosylation. Carbohydr Res. 2006;341(10):1574-1596.
78. Giddens JP, Wang LX. Chemoenzymatic Glyco-engineering of Monoclonal Antibodies. Methods Mol Biol. 2015;1321:375-387.
79. Wang LX. Chemoenzymatic synthesis of glycopeptides and glycoproteins through endoglycosidase-catalyzed transglycosylation. Carbohydr Res. 2008;343(10-11):1509-1522.
80. Umekawa M, Huang W, Li B, et al. Mutants of Mucor hiemalis endo-beta-N-acetylglucosaminidase show enhanced transglycosylation and glycosynthase-like activities. J Biol Chem. 2008;283(8):4469-4479.
81. Umekawa M, Li C, Higashiyama T, et al. Efficient glycosynthase mutant derived from Mucor hiemalis endo-beta-N-acetylglucosaminidase capable of transferring oligosaccharide from both sugar oxazoline and natural N-glycan. J Biol Chem. 2010;285(1):511-521.
82. Huang W, Li C, Li B, et al. Glycosynthases enable a highly efficient chemoenzymatic synthesis of N-glycoproteins carrying intact natural N-glycans. J Am Chem Soc. 2009;131(6):2214-2223.
83. Fan SQ, Huang W, Wang LX. Remarkable transglycosylation activity of glycosynthase mutants of endo-D, an endo-β-N-acetylglucosaminidase from Streptococcus pneumoniae. J Biol Chem. 2012;287(14):11272-11281.
84. Higuchi Y, Eshima Y, Huang Y, et al. Highly efficient transglycosylation of sialo-complex-type oligosaccharide using Coprinopsis cinerea endoglycosidase and sugar oxazoline. Biotechnol Lett. 2017;39(1):157-162.

85. Iwamoto M, Sekiguchi Y, Nakamura K, Kawaguchi Y, Honda T, Hasegawa J. Generation of efficient mutants of endoglycosidase from Streptococcus pyogenes and their application in a novel one-pot transglycosylation reaction for antibody modification. PLoS One. 2018;13(2):e0193534.
86. Li T, Tong X, Yang Q, Giddens JP, Wang LX. Glycosynthase Mutants of Endoglycosidase S2 Show Potent Transglycosylation Activity and Remarkably Relaxed Substrate Specificity for Antibody Glycosylation Remodeling. J Biol Chem. 2016;291(32):16508-16518.
87. Giddens JP, Lomino JV, Amin MN, Wang LX. Endo-F3 Glycosynthase Mutants Enable Chemoenzymatic Synthesis of Core-fucosylated Triantennary Complex Type Glycopeptides and Glycoproteins. J Biol Chem. 2016;291(17):9356-9370.
88. Rosano GL, Ceccarelli EA. Recombinant protein expression in Escherichia coli: advances and challenges. Front Microbiol. 2014;5:172.
89. Khow O, Suntrarachun S. Strategies for production of active eukaryotic proteins in bacterial expression system. Asian Pac J Trop Biomed. 2012;2(2):159-162.
90. Choi KJ, Grass S, Paek S, St Geme JW, Yeo HJ. The Actinobacillus pleuropneumoniae HMW1C-like glycosyltransferase mediates N-linked glycosylation of the Haemophilus influenzae HMW1 adhesin. PLoS One. 2010;5(12):e15888.
91. Schwarz F, Fan YY, Schubert M, Aebi M. Cytoplasmic N-glycosyltransferase of Actinobacillus pleuropneumoniae is an inverting enzyme and recognizes the NX(S/T) consensus sequence. J Biol Chem. 2011;286(40):35267-35274.
92. Lomino JV, Naegeli A, Orwenyo J, Amin MN, Aebi M, Wang LX. A two-step enzymatic glycosylation of polypeptides with complex N-glycans. Bioorg Med Chem. 2013;21(8):2262-2270.
93. Song Q, Wu Z, Fan Y, et al. Production of homogeneous glycoprotein with multisite modifications by an engineered N-glycosyltransferase mutant. J Biol Chem. 2017;292(21):8856-8863.
94. Xu Y, Wu Z, Zhang P, et al. A novel enzymatic method for synthesis of glycopeptides carrying natural eukaryotic N-glycans. Chem Commun (Camb). 2017;53(65):9075-9077.
95. Kightlinger W, Lin L, Rosztoczy M, et al. Design of glycosylation sites by rapid synthesis and analysis of glycosyltransferases. Nat Chem Biol. 2018;14(6):627-635.
96. Lin L, Kightlinger W, Prabhu SK, et al. Sequential Glycosylation of Proteins with Substrate-Specific *N*-glycosyltransferases. ACS Cent Sci. 2020;6(2):144-154.
97. Naegeli A, Neupert C, Fan YY, et al. Molecular analysis of an alternative N-glycosylation machinery by functional transfer from Actinobacillus pleuropneumoniae to Escherichia coli. J Biol Chem. 2014;289(4):2170-2179.
98. Wu Z, Jiang K, Zhu H, et al. Site-Directed Glycosylation of Peptide/Protein with Homogeneous O-Linked Eukaryotic N-Glycans. Bioconjug Chem. 2016;27(9):1972-1975.

99. Wang YS, Youngster S, Grace M, Bausch J, Bordens R, Wyss DF. Structural and biological characterization of pegylated recombinant interferon alpha-2b and its therapeutic implications. Adv Drug Deliv Rev. 2002;54(4):547-570.
100. Ceaglio N, Etcheverrigaray M, Conradt HS, Grammel N, Kratje R, Oggero M. Highly glycosylated human alpha interferon: An insight into a new therapeutic candidate. J Biotechnol. 2010;146(1-2):74-83.
101. Rabhi-Essafi I, Sadok A, Khalaf N, Fathallah DM. A strategy for high-level expression of soluble and functional human interferon alpha as a GST-fusion protein in *E. coli*. Protein Eng Des Sel. 2007;20(5):201-209.
102. Srivastava P, Bhattacharaya P, Pandey G, Mukherjee KJ. Overexpression and purification of recombinant human interferon alpha2b in *Escherichia coli*. Protein Expr Purif. 2005;41(2):313-322.
103. Ahmed N, Bashir H, Zafar AU, et al. Optimization of conditions for high-level expression and purification of human recombinant consensus interferon (rh-cIFN) and its characterization. Biotechnol Appl Biochem. 2015;62(5):699-708.
104. Valente CA, Monteiro GA, Cabral JM, Fevereiro M, Prazeres DM. Optimization of the primary recovery of human interferon alpha2b from *Escherichia coli* inclusion bodies. Protein Expr Purif. 2006;45(1):226-234.
105. Zou G, Ochiai H, Huang W, Yang Q, Li C, Wang LX. Chemoenzymatic synthesis and Fcγ receptor binding of homogeneous glycoforms of antibody Fc domain. Presence of a bisecting sugar moiety enhances the affinity of Fc to FcγIIIa receptor. J Am Chem Soc. 2011;133(46):18975-18991.
106. Yang Q, An Y, Zhu S, et al. Glycan Remodeling of Human Erythropoietin (EPO) Through Combined Mammalian Cell Engineering and Chemoenzymatic Transglycosylation. ACS Chem Biol. 2017;12(6):1665-1673.
107. Elliott S, Egrie J, Browne J, et al. Control of rHuEPO biological activity: the role of carbohydrate. Exp Hematol. 2004;32(12):1146-1155.
108. Silva MM, Lamarre B, Cerasoli E, et al. Physicochemical and biological assays for quality control of biopharmaceuticals: interferon alpha-2 case study. Biologicals. 2008;36(6):383-392.
109. Egrie JC, Dwyer E, Browne JK, Hitz A, Lykos MA. Darbepoetin alfa has a longer circulating half-life and greater in vivo potency than recombinant human erythropoietin. Exp Hematol. 2003;31(4):290-299.
110. Sareneva T, Pirhonen J, Cantell K, Julkunen I. N-glycosylation of human interferon-gamma: glycans at Asn-25 are critical for protease resistance. Biochem J. 1995;308 (Pt 1):9-14.
111. Guan X, Chaffey PK, Wei X, et al. Chemically Precise Glycoengineering Improves Human Insulin. ACS Chem Biol. 2018;13(1):73-81.
112. Ramon J, Saez V, Baez R, Aldana R, Hardy E. PEGylated interferon-alpha2b: a branched 40K polyethylene glycol derivative. Pharm Res. 2005;22(8):1374-1386.
113. Zhao W, Oskeritzian CA, Pozez AL, Schwartz LB. Cytokine production by skin-derived mast cells: endogenous proteases are responsible for degradation of cytokines. J Immunol. 2005;175(4):2635-2642.

114. Prabhu SK, Yang Q, Tong X, Wang LX. Exploring a combined Escherichia coli-based glycosylation and in vitro transglycosylation approach for expression of glycosylated interferon alpha. Bioorg Med Chem. 2021;33:116037.
115. Chung HS, Raetz CR. Interchangeable domains in the Kdo transferases of Escherichia coli and Haemophilus influenzae. Biochemistry. 2010;49(19):4126-4137.
116. Miyoshi E, Moriwaki K, Nakagawa T. Biological function of fucosylation in cancer biology. J Biochem. 2008;143(6):725-729.
117. André S, Kozár T, Kojima S, Unverzagt C, Gabius HJ. From structural to functional glycomics: core substitutions as molecular switches for shape and lectin affinity of N-glycans. Biol Chem. 2009;390(7):557-565.
118. Schneider M, Al-Shareffi E, Haltiwanger RS. Biological functions of fucose in mammals. Glycobiology. 2017;27(7):601-618.
119. Geng F, Shi BZ, Yuan YF, Wu XZ. The expression of core fucosylated E-cadherin in cancer cells and lung cancer patients: prognostic implications. Cell Res. 2004;14(5):423-433.
120. Takahashi M, Kuroki Y, Ohtsubo K, Taniguchi N. Core fucose and bisecting GlcNAc, the direct modifiers of the N-glycan core: their functions and target proteins. Carbohydr Res. 2009;344(12):1387-1390.
121. Keeley TS, Yang S, Lau E. The Diverse Contributions of Fucose Linkages in Cancer. Cancers (Basel). 2019;11(9).
122. Zhou Y, Fukuda T, Hang Q, et al. Inhibition of fucosylation by 2-fluorofucose suppresses human liver cancer HepG2 cell proliferation and migration as well as tumor formation. Sci Rep. 2017;7(1):11563.
123. Wang Y, Fukuda T, Isaji T, et al. Loss of α1,6-fucosyltransferase inhibits chemical-induced hepatocellular carcinoma and tumorigenesis by down-regulating several cell signaling pathways. FASEB J. 2015;29(8):3217-3227.
124. Sato Y, Nakata K, Kato Y, et al. Early recognition of hepatocellular carcinoma based on altered profiles of alpha-fetoprotein. N Engl J Med. 1993;328(25):1802-1806.
125. André S, Kozár T, Schuberth R, Unverzagt C, Kojima S, Gabius HJ. Substitutions in the N-glycan core as regulators of biorecognition: the case of core-fucose and bisecting GlcNAc moieties. Biochemistry. 2007;46(23):6984-6995.
126. Ferrara C, Grau S, Jäger C, et al. Unique carbohydrate-carbohydrate interactions are required for high affinity binding between FcgammaRIII and antibodies lacking core fucose. Proc Natl Acad Sci U S A. 2011;108(31):12669-12674.
127. Wang LX, Tong X, Li C, Giddens JP, Li T. Glycoengineering of Antibodies for Modulating Functions. Annu Rev Biochem. 2019;88:433-459.
128. Rodriguez-Diaz J, Monedero V, Yebra MJ. Utilization of natural fucosylated oligosaccharides by three novel alpha-L-fucosidases from a probiotic Lactobacillus casei strain. Appl Environ Microbiol. 2011;77(2):703-705.
129. Rodríguez-Díaz J, Carbajo RJ, Pineda-Lucena A, Monedero V, Yebra MJ. Synthesis of fucosyl-N-acetylglucosamine disaccharides by transfucosylation

using α-L-fucosidases from Lactobacillus casei. Appl Environ Microbiol. 2013;79(12):3847-3850.
130. Cobucci-Ponzano B, Moracci M. Glycosynthases as tools for the production of glycan analogs of natural products. Nat Prod Rep. 2012;29(6):697-709.
131. Huang W, Giddens J, Fan SQ, Toonstra C, Wang LX. Chemoenzymatic glycoengineering of intact IgG antibodies for gain of functions. J Am Chem Soc. 2012;134(29):12308-12318.
132. Li C, Li T, Wang LX. Chemoenzymatic Defucosylation of Therapeutic Antibodies for Enhanced Effector Functions Using Bacterial α-Fucosidases. Methods Mol Biol. 2018;1827:367-380.
133. Tsai TI, Li ST, Liu CP, et al. An Effective Bacterial Fucosidase for Glycoprotein Remodeling. ACS Chem Biol. 2017;12(1):63-72.
134. Nakamura S, Miyazaki T, Park EY. α-L-Fucosidase from Bombyx mori has broad substrate specificity and hydrolyzes core fucosylated N-glycans. Insect Biochem Mol Biol. 2020;124:103427.
135. Wan L, Zhu Y, Zhang W, Mu W. α-L-Fucosidases and their applications for the production of fucosylated human milk oligosaccharides. Appl Microbiol Biotechnol. 2020;104(13):5619-5631.
136. Panmontha W, Amarinthnukrowh P, Damrongphol P, Desudchit T, Suphapeetiporn K, Shotelersuk V. Novel mutations in the FUCA1 gene that cause fucosidosis. Genet Mol Res. 2016;15(3).
137. Ezawa I, Sawai Y, Kawase T, et al. Novel p53 target gene FUCA1 encodes a fucosidase and regulates growth and survival of cancer cells. Cancer Sci. 2016;107(6):734-745.
138. Jiang J, Kallemeijn WW, Wright DW, et al. In vitro and in vivo comparative and competitive activity-based protein profiling of GH29 α-l-fucosidases. Chem Sci. 2015;6(5):2782-false.
139. Liu SW, Chen CS, Chang SS, et al. Identification of essential residues of human alpha-L-fucosidase and tests of its mechanism. Biochemistry. 2009;48(1):110-120.
140. Li T, Huang M, Liu L, Wang S, Moremen KW, Boons GJ. Divergent Chemoenzymatic Synthesis of Asymmetrical-Core-Fucosylated and Core-Unmodified N-Glycans. Chemistry. 2016;22(52):18742-18746.
141. Li L, Liu Y, Ma C, et al. Efficient Chemoenzymatic Synthesis of an N-glycan Isomer Library. Chem Sci. 2015;6(10):5652-5661.
142. Calderon AD, Liu Y, Li X, et al. Substrate specificity of FUT8 and chemoenzymatic synthesis of core-fucosylated asymmetric N-glycans. Org Biomol Chem. 2016;14(17):4027-4031.
143. Wu Z, Guo X, Wang Q, Swarts BM, Guo Z. Sortase A-catalyzed transpeptidation of glycosylphosphatidylinositol derivatives for chemoenzymatic synthesis of GPI-anchored proteins. J Am Chem Soc. 2010;132(5):1567-1571.
144. Wu Z, Guo Z. Sortase-Mediated Transpeptidation for Site-Specific Modification of Peptides, Glycopeptides, and Proteins. J Carbohydr Chem. 2012;31(1):48-66.

145. Huang W, Zhang X, Ju T, Cummings RD, Wang LX. Expeditious chemoenzymatic synthesis of CD52 glycopeptide antigens. Org Biomol Chem. 2010;8(22):5224-5233.
146. Amin MN, McLellan JS, Huang W, et al. Synthetic glycopeptides reveal the glycan specificity of HIV-neutralizing antibodies. Nat Chem Biol. 2013;9(8):521-526.
147. Yang Q, Li C, Wei Y, Huang W, Wang LX. Expression, glycoform characterization, and antibody-binding of HIV-1 V3 glycopeptide domain fused with human IgG1-Fc. Bioconjug Chem. 2010;21(5):875-883.
148. Mellstedt H, Fagerberg J, Frödin JE, et al. Augmentation of the immune response with granulocyte-macrophage colony-stimulating factor and other hematopoietic growth factors. Curr Opin Hematol. 1999;6(3):169-175.
149. Yang Q, Wang LX. Mammalian α-1,6-Fucosyltransferase (FUT8) Is the Sole Enzyme Responsible for the N-Acetylglucosaminyltransferase I-independent Core Fucosylation of High-mannose N-Glycans. J Biol Chem. 2016;291(21):11064-11071.
150. Romero PA, Arnold FH. Exploring protein fitness landscapes by directed evolution. Nat Rev Mol Cell Biol. 2009;10(12):866-876.
151. Yang KK, Wu Z, Arnold FH. Machine-learning-guided directed evolution for protein engineering. Nat Methods. 2019;16(8):687-694.
152. Subedi GP, Barb AW. The Structural Role of Antibody N-Glycosylation in Receptor Interactions. Structure. 2015;23(9):1573-1583.
153. Li T, DiLillo DJ, Bournazos S, Giddens JP, Ravetch JV, Wang LX. Modulating IgG effector function by Fc glycan engineering. Proc Natl Acad Sci U S A. 2017;114(13):3485-3490.
154. Zhou Q, Shankara S, Roy A, et al. Development of a simple and rapid method for producing non-fucosylated oligomannose containing antibodies with increased effector function. Biotechnol Bioeng. 2008;99(3):652-665.
155. Kanda Y, Imai-Nishiya H, Kuni-Kamochi R, et al. Establishment of a GDP-mannose 4,6-dehydratase (GMD) knockout host cell line: a new strategy for generating completely non-fucosylated recombinant therapeutics. J Biotechnol. 2007;130(3):300-310.
156. Okbazghi SZ, More AS, White DR, et al. Production, Characterization, and Biological Evaluation of Well-Defined IgG1 Fc Glycoforms as a Model System for Biosimilarity Analysis. J Pharm Sci. 2016;105(2):559-574.
157. Prabhu SK, Li C, Zong G, Zhang R, Wang LX. Comparative studies on the substrate specificity and defucosylation activity of three α-l-fucosidases using synthetic fucosylated glycopeptides and glycoproteins as substrates. Bioorg Med Chem. 2021;42:116243.
158. Zhang R, Yang Q, Boruah BM, et al. Appropriate aglycone modification significantly expands the glycan substrate acceptability of α1,6-fucosyltransferase (FUT8). Biochem J. 2021;478(8):1571-1583.
159. Dube DH, Bertozzi CR. Metabolic oligosaccharide engineering as a tool for glycobiology. Curr Opin Chem Biol. 2003;7(5):616-625.

160. Campbell CT, Sampathkumar SG, Yarema KJ. Metabolic oligosaccharide engineering: perspectives, applications, and future directions. Mol Biosyst. 2007;3(3):187-194.
161. Almaraz RT, Mathew MP, Tan E, Yarema KJ. Metabolic oligosaccharide engineering: implications for selectin-mediated adhesion and leukocyte extravasation. Ann Biomed Eng. 2012;40(4):806-815.
162. Allen JG, Mujacic M, Frohn MJ, et al. Facile Modulation of Antibody Fucosylation with Small Molecule Fucostatin Inhibitors and Cocrystal Structure with GDP-Mannose 4,6-Dehydratase. ACS Chem Biol. 2016;11(10):2734-2743.
163. Okeley NM, Alley SC, Anderson ME, et al. Development of orally active inhibitors of protein and cellular fucosylation. Proc Natl Acad Sci U S A. 2013;110(14):5404-5409.
164. Kizuka Y, Nakano M, Yamaguchi Y, et al. An Alkynyl-Fucose Halts Hepatoma Cell Migration and Invasion by Inhibiting GDP-Fucose-Synthesizing Enzyme FX, TSTA3. Cell Chem Biol. 2017;24(12):1467-1478.e1465.
165. Zimmermann M, Ehret J, Kolmar H, Zimmer A. Impact of Acetylated and Non-Acetylated Fucose Analogues on IgG Glycosylation. Antibodies (Basel). 2019;8(1).
166. Rabuka D, Hubbard SC, Laughlin ST, Argade SP, Bertozzi CR. A chemical reporter strategy to probe glycoprotein fucosylation. J Am Chem Soc. 2006;128(37):12078-12079.
167. Sawa M, Hsu TL, Itoh T, et al. Glycoproteomic probes for fluorescent imaging of fucosylated glycans in vivo. Proc Natl Acad Sci U S A. 2006;103(33):12371-12376.
168. Hsu TL, Hanson SR, Kishikawa K, Wang SK, Sawa M, Wong CH. Alkynyl sugar analogs for the labeling and visualization of glycoconjugates in cells. Proc Natl Acad Sci U S A. 2007;104(8):2614-2619.
169. Kizuka Y, Funayama S, Shogomori H, et al. High-Sensitivity and Low-Toxicity Fucose Probe for Glycan Imaging and Biomarker Discovery. Cell Chem Biol. 2016;23(7):782-792.
170. Ma C, Takeuchi H, Hao H, et al. Differential Labeling of Glycoproteins with Alkynyl Fucose Analogs. Int J Mol Sci. 2020;21(17).
171. Okeley NM, Toki BE, Zhang X, et al. Metabolic engineering of monoclonal antibody carbohydrates for antibody-drug conjugation. Bioconjug Chem. 2013;24(10):1650-1655.
172. Hossler P, Chumsae C, Racicot C, et al. Arabinosylation of recombinant human immunoglobulin-based protein therapeutics. MAbs. 2017;9(4):715-734.
173. Dai Y, Hartke R, Li C, Yang Q, Liu JO, Wang LX. Synthetic Fluorinated l-Fucose Analogs Inhibit Proliferation of Cancer Cells and Primary Endothelial Cells. ACS Chem Biol. 2020;15(10):2662-2672.
174. Dehnert KW, Beahm BJ, Huynh TT, et al. Metabolic labeling of fucosylated glycans in developing zebrafish. ACS Chem Biol. 2011;6(6):547-552.

175. Wu ZL, Zhou H, Ethen CM, V NR. Core-6 fucose and the oligomerization of the 1918 pandemic influenza viral neuraminidase. Biochem Biophys Res Commun. 2016;473(2):524-529.
176. Haryadi R, Ho S, Kok YJ, et al. Optimization of heavy chain and light chain signal peptides for high level expression of therapeutic antibodies in CHO cells. PLoS One. 2015;10(2):e0116878.