INVESTIGATING THE ROLE OF *N*-DEACETYLASE/*N*-SULFOTRANSFERASE 2 IN HEPARIN BIOSYNTHESIS

Ryan Matthew Bullis

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Pharmaceutical Sciences (Chemical Biology and Medicinal Chemistry)

Chapel Hill
2013

Approved by:

Michael Jarstfer, Ph.D.

Jian Liu, Ph.D.

Robert Linhardt, Ph.D.

Andrew Lee, Ph.D.

Qisheng Zhang, Ph.D.

ABSTRACT

RYAN BULLIS: Investigating the Role of *N*-deacetylase/*N*-sulfotransferase 2 in Heparin Biosynthesis
(Under the direction of Jian Liu, Ph.D.)

Heparan sulfate (HS) is a highly sulfated polysaccharide, produced ubiquitously in the human body, which plays a key role in signaling and regulatory events. Heparin (HP), a structural analog of HS produced in mast cells, maintains a higher density of sulfation. HP is widely regarded for its anticoagulant properties. However, its mode of production, by extraction from animal tissues, has been recently compromised. This presents an opportunity to develop a new approach to synthesize HP and improve its safety and efficacy. The overall goal of our research is to develop an enzymatic-based approach to the design of HP-like compounds. A key component of this goal is to understand the biosynthetic pathway of HP.

The biosynthesis of HS has long been thought of as a stepwise process, consisting of: initiation of polysaccharide synthesis on a core protein, elongation, and modification. Recent research has shown that this process is a more dynamic event involving cooperation between the glycosyltransferase enzymes, epimerase, and sulfotransferases. Because the process is nontemplate driven, interactions among enzymes are key components. *N*-deacetylase/*N*-sulfotransferase (NDST) is the initial sulfotransferase to modify the polysaccharide. Most of the subsequent reactions rely on the GlcNS residue for substrate binding. NDST-2 is responsible for synthesizing HP *in vivo*, while NDST-1 is involved with HS synthesis. We aim to characterize

the substrate specificity and modification patterns of NDST-2 to gain a thorough understanding of HP biosynthesis.

We present a study of NDST-2 modification using a structurally-defined oligosaccharide library. We identified a pentasaccharide as the smallest oligosaccharide modified by NDST-2. We determined that NDST-2 does not have a directional mode of action, and is unaffected by pre-existing *N*-sulfation on the oligosaccharide. We demonstrated by one pot reaction that NDST-2 can cooperate with C_5-epimerase, and 2-*O*-sulfotransferase to form an IdoUA2S-GlcNS domain *in vitro*. The domain was formed in both short and long oligosaccharides already carrying GlcNS and IdoUA2S. This was a key discovery in uncovering the biosynthetic pathway of HP, as this repeating domain comprises up to 90% of the overall structure of HP. The result uncovers a new mechanism for control of the biosynthesis of HP.

To my sweet baby boy, Noah M. Bullis, whose smile and laughter always provides escape when times are tough and constantly reminds me of the bigger picture in life.

To my fiancé, Maria C. Cruz, whose love and encouragement has lit the path through my graduate career and made the completion of this dissertation possible.

ACKNOWLEDGEMENTS

First I would like to extend my sincerest thanks to my advisor, Dr. Jian Liu. His expertise in the field of glycobiology is unmatched and I cannot express how much I have learned from him. He has guided me through my toughest of times in the lab, when both of us wanted to throw in the towel. Now, coming out the other end of the gauntlet, he has truly helped to transform me in so many ways. I came to UNC as an inexperienced student and with his extended support I will leave UNC as a seasoned scientist ready to make a difference in the world. He tells me often how he has never seen such a sudden transformation in any student he has advised, but I have also never heard of any advisor as patient, forgiving, and understanding as Dr. Liu.

Next, I would like to acknowledge Dr. Linhardt for his assistance in supporting my graduate education through a joint supplemental grant between our laboratories. I would also like to extend my appreciation for the incredible hospitalities upon my visits to RPI. I have learned a great deal of the business side of drug development through my involvement in the BRB startup meetings. In addition, I was able to attend the glycosaminoglycan summer conference at RPI which truly broadened the scope of my knowledge and understanding in the field of glycobiology. I would also like to send my wholehearted gratitude to my other committee members, Dr. Mike Jarstfer, Dr. Andrew Lee, and Dr. Qisheng Zhang, for their time

spent providing intelligent suggestions and guidance through the development of my research project.

Finally, I would like to extend my unfeigned thanks to the members of my lab, both former and current. I would like to thank Dr. Juzheng Sheng for his mentorship with the design of the oligosaccharide substrate library and his guidance in the development of the NDST-2 substrate specificity study. I also would like to show my great appreciation and admiration for our lab 'Momma' Yongmei Xu. Yongmei always has helpful advice and suggestions and is truly a wonderful scientist and person. I would like to acknowledge the help that she has given to my project by providing starter oligosaccharides, which were only minimally modified for use in the second section of my project. I would also like to thank the former members of my lab for their support, Dr. Renpeng Liu, Dr. Liz Chappell, Dr. Kai Li, Dr. Xianxuan Zhou, and Justin Roberts, as well as current labmates, Truong Pham, Kasemsiri Chandarajoti, Tim O'Leary, Po-Hung Hsieh, Dr. Wen Zhou, and Susan Woody. Finally, I would like to send a special thank you to the "Liu ladies", Dr. Heather Bethea, Dr. Courtney Law, and Dr. Sherket Peterson. These three former labmates were always there for me when I needed them most to discuss science or life in general. Without the help of everyone involved, this dissertation and the scientific knowledge I have gained would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ΔUA	$\Delta^{4,5}$-unsaturated uronic acid
ΔUA2S	$\Delta^{4,5}$-unsaturated 2-*O*-sulfated uronic acid
2OST	2-*O*-sulfotransferase
3OST	3-*O*-sulfotransferase
6OST	6-*O*-sulfotransferase
AnMan	2,5-anhydromannitol
AT	Antithrombin
C_5-Epi	C_5-epimerase
CHO	Chinese hamster ovary
CS	Chondroitin sulfate
DEAE	Diethyl aminoethyl
DMSO	Dimethyl sulfoxide
DS	Dermatan sulfate
EPS	Epimerization site
ESI-MS	Electrospray ionization mass spectrometry
EXT	Exostosin
EXTL	Exostosin like
FBS	Fetal bovine serum
FGF(R)	Fibroblast growth factor (receptor)

GAG	Glycosaminoglycan
Gal	Galactose
GalT-I/II	Galactosyltransferase I/II
GlcNAc	*N*-acetyl glucosamine
GlcNAc6S	*N*-acetylated, 6-*O*-sulfated glucosamine
GlcNAcT-I/II	Glucosaminyltransferase I/II
$GlcNH_2$	Unsubstituted glucosamine
GlcNS	*N*-sulfated glucosamine
GlcNTFA	*N*-trifluoracetylated glucosamine
GlcUA	Glucuronic acid
GlcUAT-I/II	Glucuronyltransferase I/II
GlmU	Glucosamine-1-phosphate acetyltransferase/*N*-acetylglucosamine-1-phosphate uridyl transferase
GPI	Glycosylphosphatidylinositol
HI-FBS	Heat inactivated fetal bovine serum
HP	Heparin
HPPG	Heparin proteoglycan
HPLC	High performance liquid chromatography
HS	Heparan sulfate
HSPG	Heparan sulfate proteoglycan
HSV-I	Herpes simplex virus-1
IdoUA	Iduronic acid

IdoUA2S	2-*O*-sulfated iduronic acid
LMW HP	Low molecular weight heparin
MOI	Multiplicity of infection
MRRS	Mode of reaction recognition site
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MWCO	Molecular weight cut-off
NDST	*N*-deacetylase/*N*-sulfotransferase
NST	*N*-sulfotransferase
PAMN-HPLC	Polyamine high performance liquid chromatography
PAP	3'-phosphoadenosine 5'-phosphate
PAPS	3'-phosphoadenosine 5'-phosphosulfate
PFU	Plaque forming unit
PMSF	Phenylmethanesulfonylfluoride
RPIP-HPLC	Reverse phase ion-pairing high performance liquid chromatography
SFM	Serum free media
TFA	Trifluoroacetic acid
UDP	Uridine diphosphate
VEGF	Vascular endothelial growth factor
Xyl	Xylose
XylT-I/II	Xylosyltransferase I/II

Chapter I

Introduction

Heparan Sulfate Proteoglycans

Heparan sulfate (HS) is a highly sulfated, linear polysaccharide that resides in nearly every tissue at the cell surface and within the extracellular matrix. It is classified as a glycosaminoglycan (GAG) (1). GAGs are negatively charged heteropolysaccharides consisting of repeating disaccharide units of uronic acid and an amino sugar usually ranging between 10-100kDa in molecular weight (1). The major categories of GAGs are keratan sulfate, chondroitin sulfate (2) and dermatan sulfate (DS), heparin (HP) and HS, and hyaluronic acid (1). The distinguishing features among each category of GAG are attributed to differences in monosaccharide identities that comprise the repeating disaccharide units and glycosidic linkages between each monosaccharide (3). For example, HP and HS contain the amino sugar glucosamine, while CS and DS contain galactosamine. They can further be divided into two classifications: sulfated and non-sulfated. Hyaluronic acid is the only GAG to be non-sulfated (3). All other GAGs are highly sulfated at various positions and deprotonated at physiological pH giving rise to a myriad of signaling events and biological functions (1).

GAGs are normally covalently linked as side chains to a core protein and exist as a proteoglycan (PG). Only HA exists as a free polysaccharide not linked to any core protein (3). In the case of HS there are several associated core proteins that give rise to the HSPG, such as serglycin, syndecans, glypicans, perlecans and agrin (3). Perlecans and agrin are found in the extracellular matrix, while syndecans and glypicans remain bound to the cell surface. Glypicans are linked to the cell surface through a glycosylphosphatidylinositol motif, while syndecans normally possess a transmembrane domain (4-8). Serglycin is the only known intracellular core protein, and is the only PG to carry HP as its GAG side chains. This HPPG is found in the secretory granule of hematopoietic cell types, such as connective tissue mast cells (9). The core protein is mostly responsible for the distribution of HS among various tissue types, while the HS GAG is responsible for protein signaling leading to an array of physiological events (3).

Chemical Structure of Heparan Sulfate

Heparan sulfate naturally exists as a heterogeneous molecule containing various modifications throughout. The placement of these modifications throughout the polysaccharide backbone is maintained by a wide array of biosynthetic enzymes present in the lumen of the golgi apparatus (10). In addition, there are multiple isoforms for many of these various enzymes that maintain unique substrate binding capacity and unique modification patterns. On the whole, these modifications represent unique binding sites for various biologically relevant proteins, implicating the role of HS across much of human biology (11).

The basic building block of HS structure is a repeating disaccharide backbone of D-glucuronic acid (GlcUA) covalently linked to D-*N*-acetyl glucosamine (GlcNAc) through β-1,4 and α-1,4 alternating glycosidic linkages (Figure 1) (12). The glucosamine residue may undergo several modifications on its way to becoming fully functional HS. It can be sulfated at the 3-OH and 6-OH positions, and the *N*-position can either be acetylated, sulfated, or, to a lesser extent, be presented as a free amine. Only 1-7% of the glucosamine residues remain as a free amine (13). The GlcUA residue can become L-iduronic acid (IdoUA) by interconversion at the C_5-position. While GlcUA is only found in the $^4C_1$ chair conformation, IdoUA can reside in either the $^1C_4$ chair or $^2S_0$ skew boat, thus allowing for much more conformational flexibility (14, 15). In addition, GlcUA and IdoUA can be sulfated at the 2-OH position. These modifications do not run to completion and are governed by the presence or absence of pre-existing modifications. The ability of these molecules to possess such large combinations of modifications gives rise to heterogeneous macromolecules that are capable of binding and signaling many various proteins and controlling a multitude of biological events.

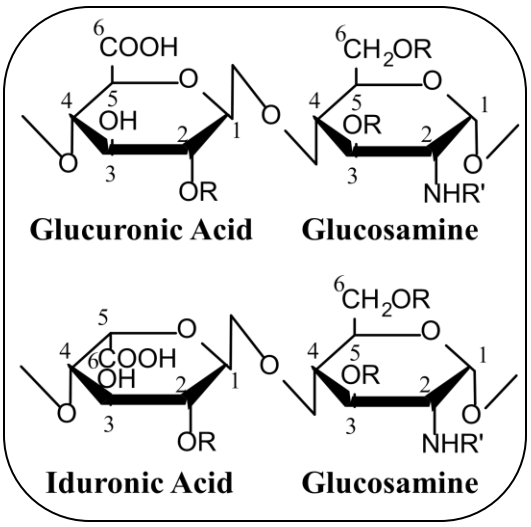

Figure 1. Disaccharide repeating units of HS. The uronic acid monosaccharide units can exist as either GlcUA or IdoUA. Sulfation (R=-SO_3) at the 2-*O* position of IdoUA is common, sulfation at the 2-*O* position of GlcUA is less common. Sulfation at the 6-*O* position of glucosamine is common. Sulfation at the 3-*O* position of glucosamine is rare. Both *N*-acetyl (R'=-Ac, GlcNAc) and *N*-sulfo (R'=-SO_3, GlcNS) glucosamine are common. *N*-unsubstituted glucosamine (R'=H, $GlcNH_2$) is less common.

HS may be further characterized by the domain structures that define its overall topography. HS is known to possess stretches of highly sulfated domains which are flanked by domain structures carrying very low sulfation density (Figure 2). The highly sulfated domains carry out much of the protein signaling due to their negative charge density (16). These domains are often referred to as NS domains due to the presence of repeating GlcNS residues. Many of the HS biosynthetic enzymes rely on the presence of GlcNS residues in order to carry out their reactions, thus adding sulfation density to these domains. The domains carrying low sulfation density are known as NAc domains due to the presence of repeating GlcNAc residues (17). These two domain structures are separated by segments of repeating disaccharides carrying both the GlcNAc and GlcNS residues (18)

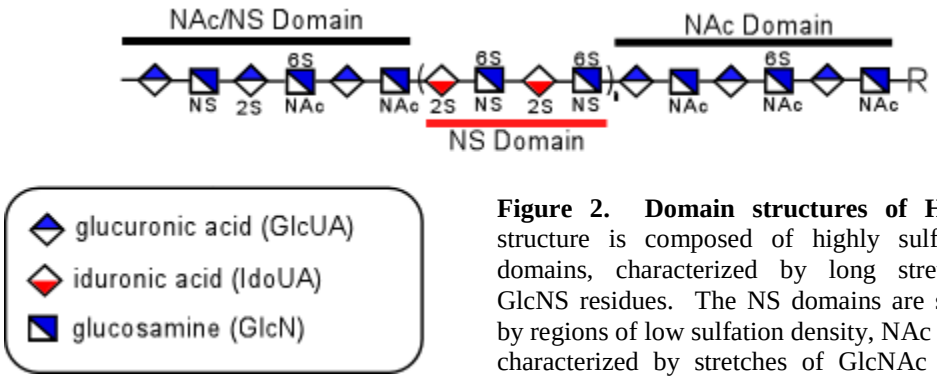

Figure 2. Domain structures of HS. HS structure is composed of highly sulfated NS domains, characterized by long stretches of GlcNS residues. The NS domains are separated by regions of low sulfation density, NAc domains, characterized by stretches of GlcNAc residues. There are also portions of the structure with alternating GlcNAc/GlcNS residues known as NAc/NS domains.

HP is a structural analogue of HS. It has been widely exploited therapeutically for its anticoagulant properties for over half a century (19). It maintains the same repeating disaccharide structure (UA-GlcN) but is different from HS mainly due to its high density of sulfation. On average, HP maintains 2.6 sulfo groups per disaccharide unit, while HS disaccharides carry 0.6 sulfo groups (3). HP is the most highly sulfated GAG that is currently known (20). Between 70-90% of its overall structure consists of the repeating disaccharide IdoUA2S-GlcNS6S, while only 30-60% of HS *N*-sulfated glucosamine residues (21). The remainder of HP structure is made up of other combinations of modification and the unique AT binding pentasaccharide (GlcNAc6S-GlcUA-GlcNS3S6S-IdoUA2S-GlcNS6S) that confers its anticoagulant activity (Figure 3) (22). In addition to degree of sulfation, fully elongated HP polysaccharide chains are generally longer (M_r = 60,000-100,000 Da) than HS (M_r = 22,000-45,000) (23, 24).

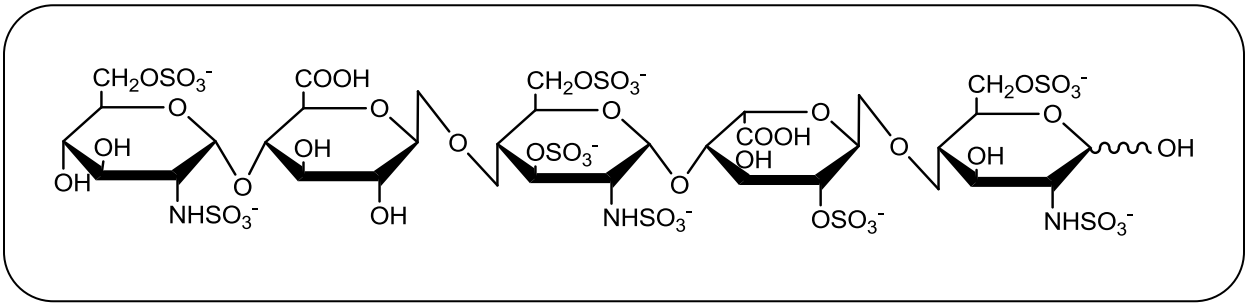

Figure 3. Chemical structure of AT-binding pentasaccharide.

Structural Analysis of Heparan Sulfate

The heterogeneous nature of HS polysaccharides has always presented a formidable challenge for researchers in attempting to decipher the specific monosaccharide sequence that constitutes each molecule. The sheer length, 100-400 monosaccharide units, in addition to the extensive array of potential modifications present, makes purification of a structurally defined HS nearly impossible (7). This is a critical barrier to the understanding the structure/function relationship between HS and their protein signaling partners. As a result, glycobiologists have determined that the best method for elucidating structures of polysaccharides is to use various methods of degradation into much shorter oligosaccharide or disaccharide fragments. Recent research has also developed dependable chemoenzymatic methods for designing structurally defined oligosaccharides which can also be utilized for probing structure/function relationships of resultant HS fragments (25). The oligosaccharides can then be analyzed by a variety of laboratory methods such as High Performance Liquid Chromatography (HPLC), and capillary electrophoresis. Electrospray ionization mass spectrometry (ESI-MS), and tandem mass spectrometry (MS/MS) are also useful methods for determining oligosaccharide chain length and modifications (26). Using these techniques in a combinatorial approach has allowed researchers to study HS structure, function, and synthesis like never before.

Depolymerization of HS polysaccharides into disaccharides for compositional analysis has been reported using both enzymatic and chemical approaches. Following depolymerization, these disaccharides are then able to be analyzed by anion-exchange chromatography, capillary electrophoresis, or reverse-phase ion pairing (RPIP) HPLC

analysis (27, 28). These techniques will reveal the relative composition of the HS polysaccharides. Using these methods, there have been 22 various disaccharides reported from naturally occurring HS and HP (7).

Enzymatic degradation of HS polysaccharides is achieved using the family of heparin lyase enzymes originally derived from *Flavobacterium heparium*. There are 3 known isoforms of heparin lyase, each expressing unique substrate modification sites based on the identities of the surrounding monosaccharide residues (Figure 4) (7). Heparin lyase I has shown the capability to cleave the glycosidic bond between GlcNS-IdoUA2S. Heparin lyase III has less selectivity as it cleaves the bond between GlcNAc/NS-GlcUA. Finally heparin lyase II has the least selectivity as it will cleave the glycosidic linkage between GlcNAc/NS-GlcUA/IdoUA (29). The disaccharide products of these enzymatic reactions can then be readily analyzed by RPIP-HPLC against appropriate standards to determine the modifications within (30, 31). This is made possible by the creation of a $\Delta^{4,5}$-unsaturated uronic acid residue during the reaction, which maintains a UV absorbance at 232nm (7). However, the creation of this uronic acid species results in an inability to determine the original identity of the uronic acid in the original polysaccharide sample. Using this enzymatic method has aided researchers in elucidating the identities and relative abundance of individual modified disaccharides at the polysaccharide or oligosaccharide levels. These discoveries have led to a more thorough understanding of biologically active sites within HS polysaccharides (32-34).

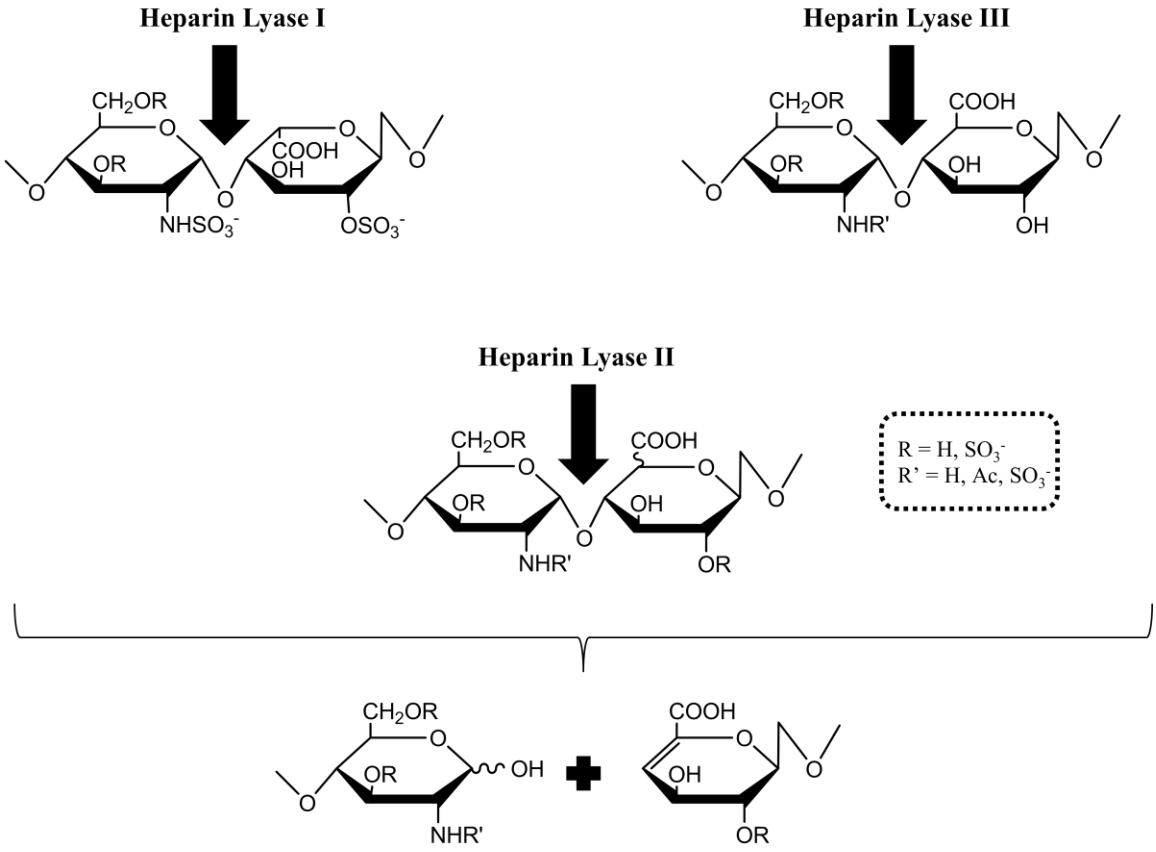

Figure 4. Substrate specificity among various heparin lyase isoforms. Heparin lyase I cleaves the disaccharide GlcNS-IdoUA2S. Heparin lyase III cleaves the disaccharide GlcNS/NAc-GlcUA. Heparin lyase II is capable of cleaving both substrates. Adapted from (7).

Another method for depolymerization of polysaccharides is nitrous acid degradation. Nitrous acid reacts with GlcNS or $GlcNH_2$ residues to form 2,5-anhydromannose residues, which are then reduced by sodium borohydride to form anhydromannitol (Figure 5) (35). At pH 1.5, nitrous acid will react with GlcNS residues. While at a higher ph (4.5-5.5) nitrous acid will exhibit selectivity for the unsubstituted free amine on the $GlcNH_2$ residue. In the case of GlcNAc residues, the polysaccharide must undergo deacetylation by treatment with hydrazine before going through the high pH degradation pathway (35). The combination of treatment with nitrous acid at both

low and elevated pH will lead to nearly complete digestion of the polysaccharide or oligosaccharide to be analyzed. The resultant disaccharide products can then be analyzed by RPIP-HPLC by coelution with appropriate standards (36). The key advantage of this method over lyase degradation lies in its ability to determine the configuration of the uronic acid residue in from the parent molecule (7). However, the main drawback to this method is the need for radioisotope or fluorescent labeling of the compound prior to digestion in order for HPLC detection (35, 37).

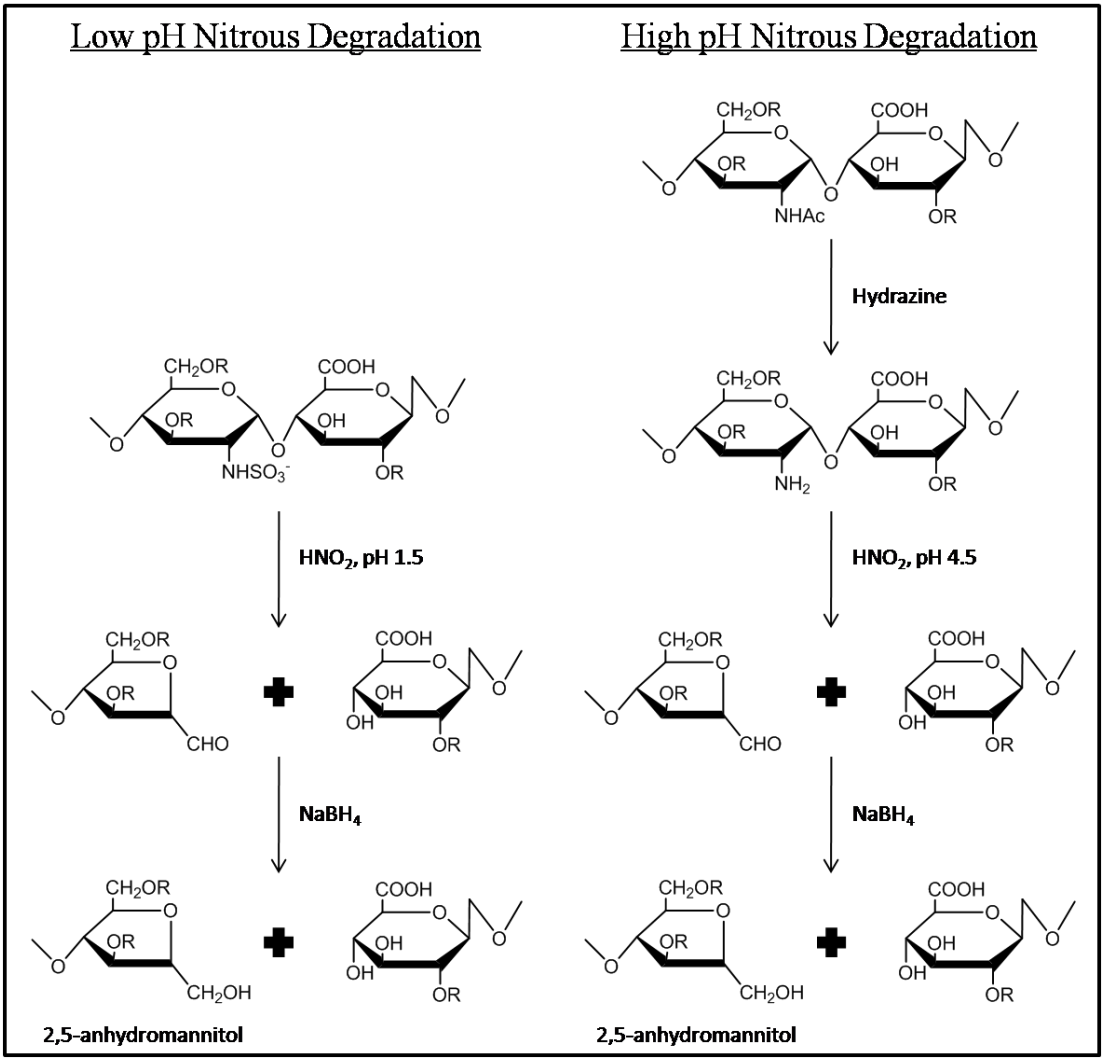

Figure 5. Nitrous acid degradation of heparan sulfate. At pH 1.5 nitrous acid cleaves the bond between *N*-sulfo glucosamine and uronic acid. When the *N*-position is acetylated hydrazine is used to convert to unsubstituted glucosamine, which is amenable to cleavage at pH 4.5. The uronic acid maintains its conformation throughout. Adapted from (7).

On the whole, a combinatorial approach of chemical and enzymatic depolymerization techniques is a useful tool for researchers to identify structural motifs within HS polysaccharides and oligosaccharides that result in biological function. This knowledge may then be exploited for use in therapeutics that may one day be used to treat a wide range of health problems.

Biosynthesis of Heparan Sulfate

Heparan sulfate biosynthesis can be divided into three separate phases before a fully functional HS proteoglycan is rendered. These phases are 1) Initiation, 2) Polymerization, and 3) Modification. The biosynthetic process mostly takes place in the Golgi apparatus, where the biosynthetic enzymes line the lumen. The initiation phase consists of attachment of a tetrasaccharide linkage region to the core protein. Next, the polymerization phase consists of repeated attachment of monosaccharides to form the polysaccharide backbone structure. Finally, the modification phase consists of an array of various deacetylation, sulfation, and epimerization provided by many various enzymes to produce biologically active HS (38).

Initiation

The initiation phase of proteoglycan formation is identical across each GAG classification, from chondroitin sulfate and dermatan sulfate to heparan sulfate and heparan. Each GAG chain is attached to the core protein through a common linkage region of –GlcUAβ3Galβ3Galβ3Xylβ3-L-[ser] (Figure 6). The transition step from initiation to elongation will determine the identity of the GAG to be synthesized (39). Initiation begins with the translation of a core protein and the transfer of xylose from

UDP-xylose to specific serine residues on the protein by xylosyltransferase (XT). Xylosyltransferase is a golgi associated type II transmembrane domain protein that performs the initial xylosylation in the early cisternae of the Golgi apparatus before moving into the lumen for subsequent modification (40, 41).The GAG attachment sites on the core protein contain the consensus sequence Ser-Gly/Ala-X-Gly (X stands for any amino acid) (42). There are two vertebrate β4-xylosysltransferases, XylT-I and XylT-II. These enzymes have 55% homology and both recognize the same consensus sequence for GAG attachment (40, 43). Elimination of XylT in CHO cells resulted in an inability for these cells to produce HS or CS, thus showing their necessity in GAG biosynthetic initiation (44).

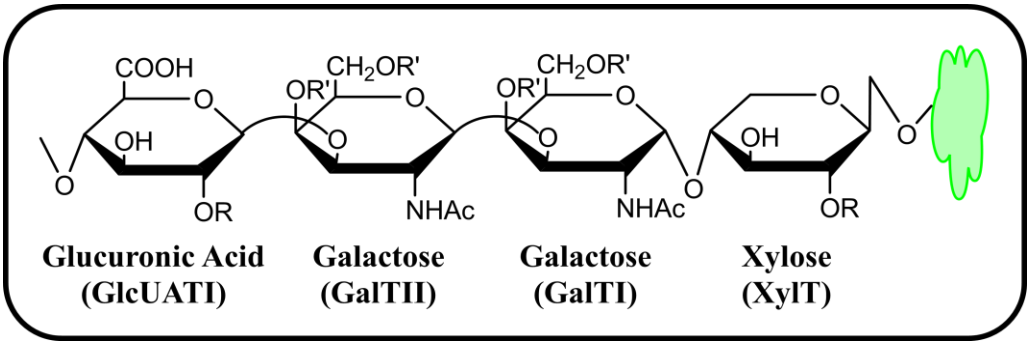

Figure 6. Chemical structure of HS linkage region and corresponding biosynthetic enzymes. The linkage region synthesis begins with the addition of xylose to the core protein (green) by the enzyme xylosyl transferase. Next two galactose units are added by galactosyltransferase type I and II respectively. Finally, the linkage region is completed with the addition of glucuronic acid by the enzyme glucuronic acid transferase type I.

The next step in the initiation pathway of GAG biosynthesis relies on the addition of two D-galactose residues from their UDP-galactose donors by the enzymes galactosyltransferase I and II (GalT-I & GalT-II). GalT-I is a member of the β4 galactosyltransferase family while GalT-II is a member of the β3 galactosyltransferase family (45, 46). Both of these galactosyltransferase enzymes are single isoforms

localized in the medial Golgi (47, 48). CHO cell mutants lacking GalT-I have been shown to be deficient in HS and CS (49). While small interfering RNA has been used to impede GalT-I, also resulting in a deficiency of HS and CS (47). These results show the requirement of both of these enzymes in the biosynthesis of GAGs.

Finally, the linkage region is completed by the addition of GlcUA by glucuronyltransferase 1 (GlcAT-I). GlcAT-I is a member of a family of β-1,3-glucuronyl transferases that exhibit action of glycoproteins and glycolipids (50, 51). CHO cells lacking GlcAT-I have been shown to be deficient in HS and CS, thus exhibiting its necessity for GAG chain synthesis (52).

Formation of the linkage region can be selectively regulated through a series of sulfations and phosphorylations. For example, the Gal residues in CS linkage regions may be selectively sulfated at the 4-*O* and 6-*O* positions, while they are not in HS (53). These sulfations may serve to either enhance formation of CS or block the synthesis of HS (39). Studies have shown that the presence of these sulfations on Galβ1-3Gal enhances the activity of GlcAT-I to complete the formation of the linkage region (53-55). Phosphorylation at the 2-*O* position of the xylose residue has also been shown to inhibit the action of GalT-I. This has been postulated as a rate limiting step in the formation of GAG chains (56).

Polymerization

Once the linkage region is in place, the repeating disaccharide structure of alternating GlcUA and GlcNAc is ready to be assembled. The polymerization of HS is controlled by the exostosin (EXT) gene family of glycosyltransferases (Figure 7). This

family consists of 5 currently characterized glycosyltransferase enzymes: EXTL1, EXTL2, EXTL3, EXT1, and EXT2 (57-59). As mentioned previously, the initial transfer of α-GlcNAc to the non-reducing end of the tetrasaccharide linkage region is the divergent point between synthesis of HS and CS. When CS is to be made, a β-*N*-acetylgalactosamine residue is added to the tetrasaccharide linkage region by a CS GalNAcT enzyme (42). The transfer of the initial α-GlcNAc monosaccharide to the linkage region relies on a transferase with GlcNAcT-I activity, wheras polymerization relies on GlcNAcT-II activity. Exostosin-like 2 (EXTL2) and EXTL3 have both demonstrated GlcNAcT-I activity (57, 58). While EXTL2 only possesses GlcNAcT-I activity, EXTL3 also has demonstrated GlcNAcT-II activity and most likely plays a role in HS polymerization as well. EXTL1 has only GlcNAcT-II activity and is most likely involved in polymerization of the growing polysaccharide chain (57). Each of the EXTL isozymes has the ability to recognize the amino acid sequence on the core protein that is proximal to the GAG attachment site. For the initial α-GlcNAc transfer, they prefer attachment sites that are flanked by acidic and hydrophobic amino acids with repeating serine-glycine units (28,42,60-63). This means that the EXTL enzyme is also able to interact with the core protein through a separate domain, in addition to binding the tetrasaccharide linkage region, in order to selectively polymerize a HS polysaccharide chain.

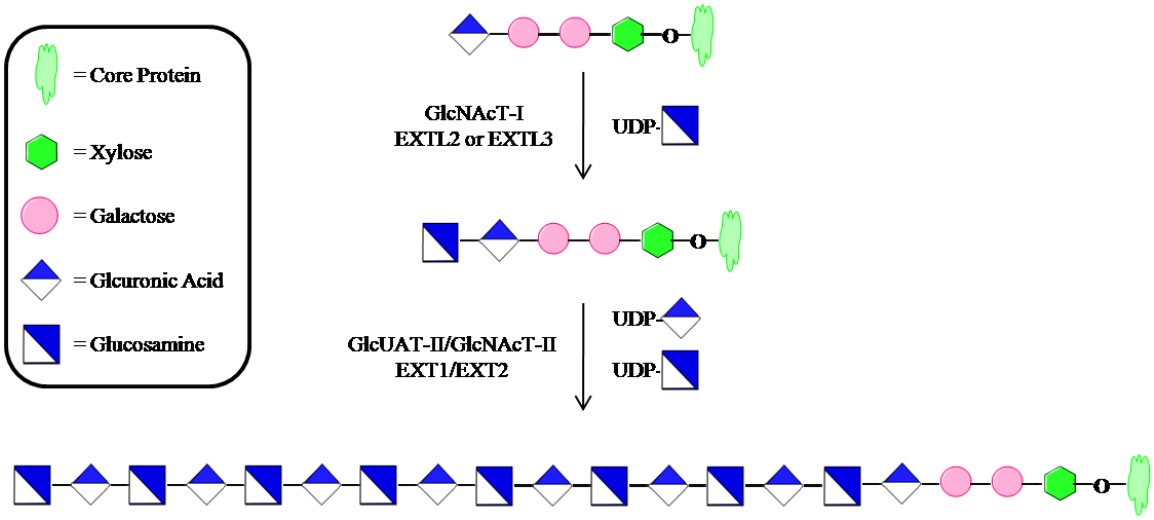

Figure 7. Elongation of the HS polysaccharide. Polymerization begins with the initial addition of GlcNAc to the tetrasaccharide linkage region, mediated by EXTL2 or EXTL3, both maintaining GlcNAcT-I activity. Next, an alternating disaccharide structure of GlcUA and GlcNAc is added by the enzyme complex of EXT1/EXT2, which maintains both GlcUAT-II and GlcNAcT-II activity.

After the initial α-GlcNAc is added to the tetrasaccharide linkage region, polymerization of the repeating disaccharide backbone of GlcUA and GlcNAc is mediated by a HS-polymerase enzymatic complex of EXT1 and EXT2, with most likely some influence by EXTL1 and EXTL3 (59). The activities of these glycosyltransferases were demonstrated in mice models. EXT1 knockout mice exhibited developmental defects and absence of HS, while EXT2 knockout mice exhibited the same developmental defects (10, 64). EXT1 and EXT2 were also expressed separately in yeast, which does not naturally produce HS. Both enzymes displayed GlcNAc transferase and GlcUA transferase activities, but at a lower level than when they were expressed together (65). This suggests that the biologically relevant form of these enzymes is a complex of EXT1 and EXT2 which are responsible for HS chain polymerization. In fact, it has been reported that these enzymes form a hetero-oligomeric complex *in vivo* that can be localized to the golgi apparatus (66).

The biosynthesis of heparan sulfate and heparin is most often presented as a linear process of initiation, polymerization, and modification of the polysaccharide backbone. However, this model may only be telling a one dimensional story of a much more highly dynamic and cooperative nature of biosynthetic machinery used to create fully functional heparan sulfate and heparin. The initial evidence pointing to this cooperative mode of heparan sulfate biosynthesis was reported by Lidholt, Kjellen, and Lindahl in 1989. They found that elongating polysaccharide chains in mouse mastocytoma microsomal fractions were synthesized to be larger when incubated in the presence of the sulfo donor 3'-phosphoadenosine 5'-phosphosulfate (PAPS) than those incubated without the donor. The sulfated polysaccharide chains were nearly tenfold larger by average molecular weight than their non-sulfated counterparts (67). In addition, they found that GlcUA transferase ability of EXT1 greatly favored a previously *N*-sulfated substrate versus a substrate with *N*-acetylated or *N*-unsubstituted glucosamine residues (68). Taken together, the evidence has existed for over two decades that *N*-sulfation and polysaccharide elongation have significant influence over one another and most likely occur simultaneously. Given that *N*-deacetylase/*N*-sulfotransferase (NDST) is responsible for the deacetylation and sulfation of the GlcNAc residues, it has been postulated that it may form an enzyme complex with the HS polymerase enzymes during elongation (69). There are also numerous other examples of coordination between HS biosynthetic modification enzymes that will be discussed further below.

Modification

As the polysaccharide backbone is built, a series of modification reactions take place. These are performed by 4 types of sulfotransferase enzymes and one epimerase

(Figure 8). The first modification to take place is a deacetylation reaction at the *N*-position of GlcNAc residues. Subsequently, most of these glucosamine residues are transferred a sulfo group from the sulfo donor PAPS. Both reactions are catalyzed by the bi-functional enzyme *N*-deacetylase/*N*-sulfotransferase. Next, the enzyme C_5-epimerase (C_5-epi) converts some of the GlcUA residues into IdoUA by interconversion at the C_5 position of the hexuronic ring. This reaction is followed by sulfation at the 2-*O*-position of some of the uronic acid residues by 2-*O*-sulfotransferase (2OST). Finally, the glucosamine residues can be sulfated at the 6-*O* and 3-*O* positions by 6-*O*-sulfotransferase (6OST) and 3-*O*-sulfotransferase (3OST) respectively (70). Each enzyme possesses unique substrate modification capabilities across each of their isoforms leading to final polysaccharide structures with unique properties.

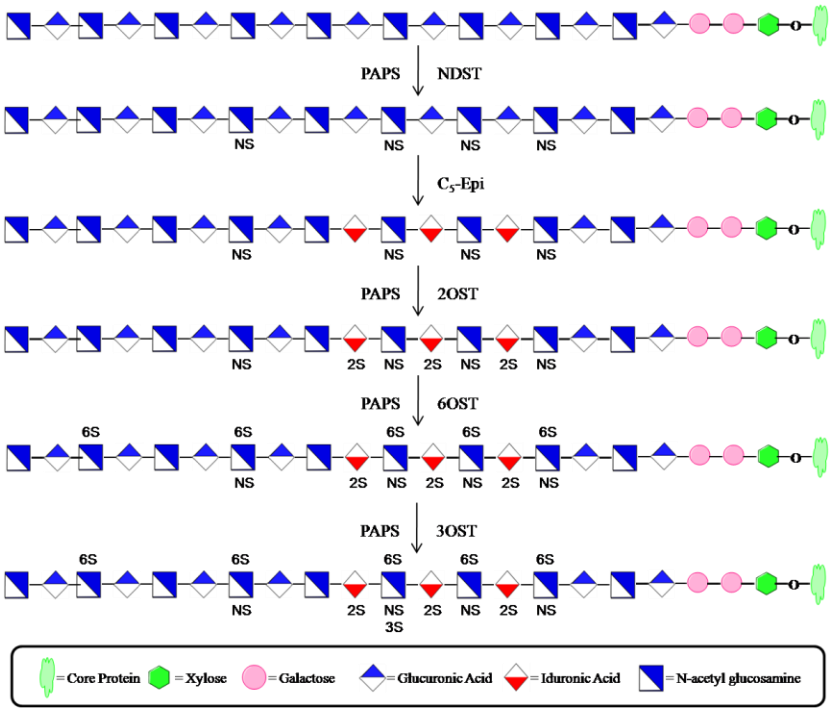

Figure 8. Enzymatic modifications of HS. The first modification to take place is the deacetylation and sulfation at the *N*-position of GlcNAc, facilitated by NDST. Next, C_5-Epi converts some of the GlcUA units to IdoUA. Following epimerization 2OST will add sulfo groups to the 2-*O*-position of uronic acid. Finally 6-*O* and 3-*O*-sulfation is added to glucosamine by the enzymes 6OST and 3OST respectively. The reactions do not run to completion giving rise to a heterogeneous HS polysaccharide.

Glucosaminyl N-Deacetylase/N-Sulfotransferase

The first enzyme to modify the growing polysaccharide chain is *N*-deacetylase/*N*-sulfotransferase (NDST). It is the most essential of the modifications to the polysaccharide backbone as nearly every other HS biosynthetic enzyme, 6-*O*-sulfotransferase notwithstanding, relies on *N*-sulfation for substrate recognition and modification, including the HS polymerases as stated previously (71). This notion arises from the fact that in both HS and HP, the regions with large sulfation density always have long stretches of GlcNS residues. As mentioned previously, it is a bifunctional enzyme with two separate domains responsible for carrying out both deacetylation at the *N*-position of GlcNAc and addition of a sulfo group to the *N*-position, thus creating GlcNS (Figure 9) (72). There are 4 known isoforms of NDST (1-4) that possess the same overall structure and 65-80% nucleotide sequence identity (73). Sequence alignment and molecular modeling have revealed that these enzymes are type II transmembrane domain proteins with a 12-13 amino acid cytoplasmic tail, an approximately 20 amino acid transmembrane domain, a short domain that varies among each isoform, and a globular domain that houses the catalytic sites. They also contain a PAPS binding domain that is highly conserved across many various sulfotransferase enzymes (73). Other analyses have revealed that the sulfotransferase domains of the NDST enzymes are located near the carboxyl end of the protein with a deacetylase domain located more towards the amino terminus (74-77).

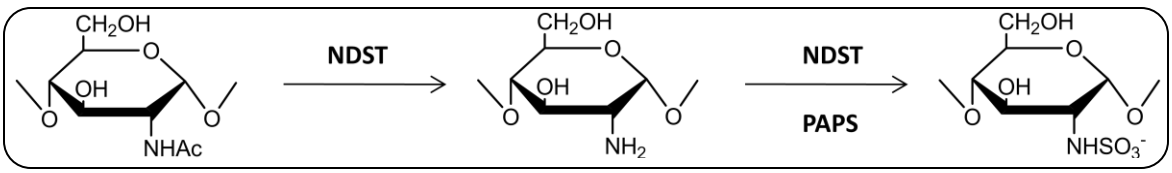

Figure 9. NDST reaction. NDST is a bifunctional enzyme that catalyzes the deacetylation of the *N*-position of glucosamine followed by a transfer of a sulfo group from the donor PAPS.

Although none of the x-ray crystal structures of the full length proteins have been solved, Kakuta et al has published an x-ray crystal structure of the sulfotransferase domain of NDST-1 (NST) in complex with PAP (Figure 10) (76). They found that NST consisted of a single α/β fold with a highly conserved five strand β-sheet that is characteristic of the PAPS binding site across most sulfotransferases. The PAP molecule is also further stabilized by a three strand β-sheet that runs anti-parallel and contains a disulfide bond. The PAPS binding site is sandwiched between the structural regions α1, α6, β1, β7, the 5'-phosphate binding loop between β1 and α1, and a random coil between β8 and α13. There is also a substrate binding cleft between α6 and a random coil that runs between β2 and α2. In addition, lysine 614 is able to form a hydrogen bond with the 5'-phosphate of PAPS (76, 78). The random coil that runs along the bottom of the substrate binding cleft is reported to differ in charge and composition among each isoform and has been postulated to be a major source of substrate preference among the NDSTs (73).

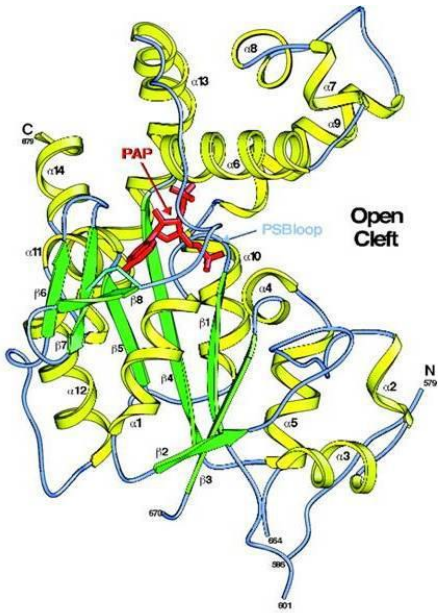

Figure 10. Crystal structure of NST in complex with PAP. Helices are colored in yellow, β-strands in green, random coil in blue and PAP in red. Structure taken from (76).

Each of the four isoforms has unique capabilities to show preference towards certain substrates and modify them in specific reproducible patterns. Based on overall tissue expression patterns the NDST isoforms can be divided into two groups: NDST1/NDST-2 and NDST-3/NDST-4 (73, 79, 80). The mRNAs for NDST-1 and NDST-2 are expressed in every tissue in the body and are therefore considered to be the most essential in the HS and HP biosynthesis pathways. It is important to note that these mRNAs are translationally regulated (81). NDST-1 is the protein that becomes translated in most cells, while NDST-2 is highly translated in connective tissue type mast cells, whereas NDST-1 is barely detected (79). HP naturally resides in mast cells, thus implicating the role of NDST-2 in HP biosynthesis. NDST-3 and NDST-4 have only been found in adult brain and fetal tissues and seem to have a very narrow functionality (73, 79, 80). The biological significance of each NDST isoform has been studied in mouse models and is discussed below.

NDST-1 is widely believed to be the most influential isoform in the biosynthetic pathway of HS. NDST-1studies in mice have shown that this enzyme is absolutely essential for synthesizing HS *in vivo*. Mice lacking NDST-1 were arrested in development where they underwent neonatal respiratory distress and death. This was marked by a reduction in lamellar bodies and microvilli versus the control group, in addition to higher glycogen content. The lungs of the NDST-$1^{-/-}$ mice also had a lower concentration of phospholipids and disaturated phosphatidylcholine (82). This evidence shows that NDST-1 is absolutely essential for early lung development and sustaining life. Additional knockout experiments have revealed the role of NDST-1 in many various biological processes such as: inhibition of tumor angiogenesis, neutrophil trafficking,

clearance of lipoproteins, and development of lobuloalveolar in the mammary gland (83-86).

In addition to studies on the biological function of NDST-1, Sheng et al performed multiple experiments using structurally defined oligosaccharide substrates to prove the much anticipated role of NDST-1 in HS biosynthesis (26). NDST-1 was seen to have no preference for its initial sulfation of the substrate, leading to multiple products for a single substrate. This initial observation shows that NDST-1 is a key player in producing the structural heterogeneity that defines HS. They also were able to demonstrate the role that this isoform plays in forming the domain structures prevalent in the overall structure of HS. It was reported that the enzyme moves in a processive fashion while sulfating the substrate consecutively in a direction from the nonreducing end towards the reducing end of the polysaccharide, stopping when the enzyme reaches the residue that is four units from the reducing end. This data provides a platform for the formation of the NS domain within HS (26). They also found that in the presence of a pre-existing *N*-sulfated substrate NDST-1 would allow for a gap of at least 5 sugar residues before initiating the subsequent *N*-sulfation. This data provides evidence that NDST-1 also has the substrate recognition capability to identify a pre-existing GlcNS causing the enzyme to form a NAc domain (26). With all of these recently discovered data, one can infer that NDST-1 is perhaps the most influential of the NDST isoforms when it comes to HS biosynthesis. In fact, studies have shown that the levels of HS polymerase expression (EXT 1 and EXT2) can have a significant impact on NDST-1 expression and affect the overall structure of HS that is produced. Overexpression of EXT2 in HEK293 cells led to greater NDST-1 expression and larger extent of *N*-

sulfation, while overexpression of EXT1 resulted in decreased expression. It has been postulated that NDST-1 may compete with EXT1 for binding to EXT2 (69). This data provides yet another potential example of HS biosynthetic enzyme cooperation that could be an essential element of HS biosynthesis.

If NDST-1 is essential to HS biosynthesis, but causes the formation of heterogeneous products with an appreciable number of NAc domains, there has to be another isoform that is responsible for the nearly 90% of NS domains that are prevalent in the structure of HP. That isoform is believed to be NDST-2. HP resides solely in the intracellular vesicles of connective tissue-type mast cells (87). In these specific cell types, NDST-2 is widely expressed, while NDST-1 is barely detectable (79). So it can be reasonably assumed that this isoform plays a key role in the biosynthesis of HP. The function of NDST-2 *in vivo* was studied using gene knockout experiments in mice. Researchers found that NDST-$2^{-/-}$ mice were unable to produce fully sulfated HP. As a result, the morphology of the mast cells was altered as they contained lower levels of histamine and mast cell proteases. Otherwise, these mice were viable and fertile. This study also led to the understanding that perhaps the key physiological role of HP is to sequester positively charged proteases within mast cells (87, 88). NDST-2 also possesses a greater ratio of deacetylase/sulfotransferase activity, also implicating its role in creating the highly sulfated domains of HP (73). There is also some evidence in microsomes suggesting that NDST-2 prefers to act on substrate locations containing pre-existing GlcNS residues (89, 90). This could help explain the much longer NS domains in HP when compared to HS.

Significantly less is known about the isoforms NDST-3 and NDST-4. These isoforms have very limited tissue distribution patterns and it is not clear as to the extent of their roles in HS biosynthesis. A murine knockout of NDST-3 revealed its tissue specific role in HS biosynthesis. NDST-$3^{-/-}$ mice were able to develop normally and were fertile, exhibiting only minor hematological and behavioral phenotypes. However, there was a detectable change in the sulfation level of HS located in the brain regions (80). NDST-3 has been shown to possess a significantly larger deacetylase/sulfotransferase activity than the other isoforms. NDST-4 displays the opposite activity with much more sulfotransferase activity (73).

In addition to the interaction between the HS polymerase enzymes, there also appears to be cooperation between NDSTs and 2-OST/C_5-epi which will be discussed below.

Glucuronosyl C_5-Epimerase

The structural heterogeneity of HS is not only determined by the array of sulfation on the polysaccharide backbone, but also by configuration of the uronic acid residues in the repeating disaccharide. Initially, a GlcUA monosaccharide will be added to the elongating chain by the HS polymerase enzymes, EXT1/EXT2. The pyranose ring of GlcUA can only exist in the chair conformation, thus limiting the substrate flexibility in binding to proteins (14). This potential limitation is erased by the enzyme C_5-epimerase. This enzyme has the ability to catalyze the reaction from D-glucuronic acid to L-iduronic acid by interconversion at the C_5-position (91). The reaction proceeds through a putative carbanion intermediate where a proton is removed from the C_5-position and added back

to the ring resulting in potential ring interconversion (Figure 11) (92, 93). IdoUA has the ability to take a chair or skew-boat conformation as it resides in the HS polysaccharide (14). Only one known isoform of C_5-Epi exists in the human genome. It is predicted to be a type II transmembrane domain protein with a 17 amino acid transmembrane domain and an 11 amino acid cytoplasmic tail. Overall it maintains an approximate molecular weight of 70kDa consisting of 618 amino acids (94). A murine gene knockout study revealed that a lack of this enzyme proved to be neo-natal lethal with mice displaying developmental defects of the lungs and kidneys (70). Thus, the conversion of the uronic acid residue into a more flexible conformation is absolutely essential for binding of HS to a wide array of proteins and for maintaining life. The substrate binding site for C_5-Epi requires a GlcNS residue to be located on the reducing end adjacent to the uronic acid residue to be modified. However, if the adjacent GlcNS residue or uronic acid residue is *O*-sulfated, the enzyme will not react (95-97). This means that epimerization relies on *N*-sulfation in order to proceed, but occurs before *O*-sulfation around the site of uronic acid epimerization (98). In addition, IdoUA residues are always located at sites of *N*-sulfation (99).

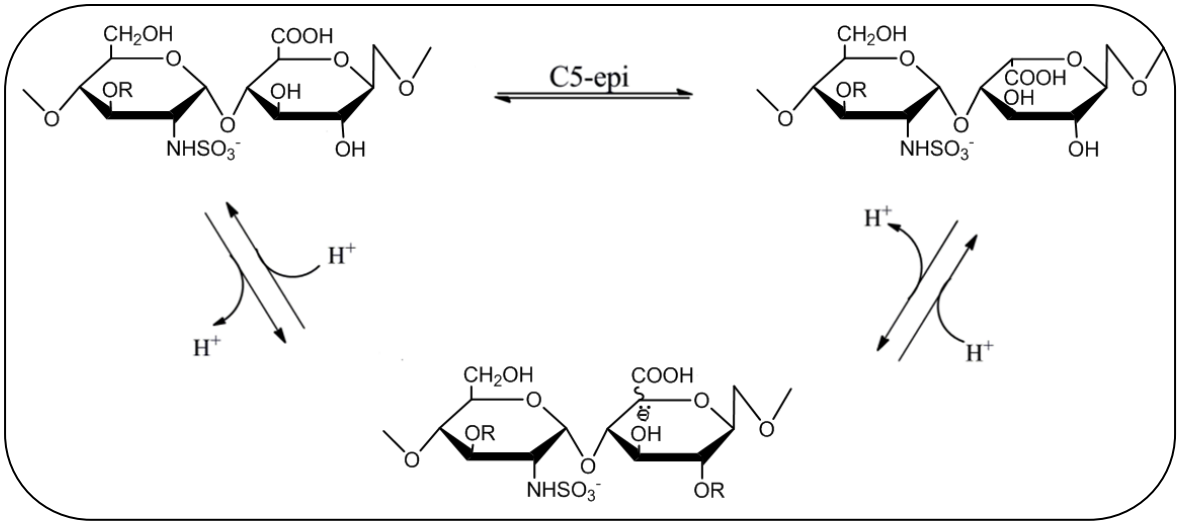

Figure 11. C_5-Epimerase mechanism. C_5-Epi abstracts a proton from carbon 5 of GlcUA forming a carbanion intermediate. The readdition of a proton at the carbon 5 position results in a conformational change to IdoUA.

The most prevalent problem for researchers attempting to characterize C_5-Epi has always been the "reversibility" of the enzyme, or its capability to convert back to GlcUA (92). Recent research has uncovered a mode of irreversibility for C_5-Epi that relies solely on a code of *N*-sulfation. Irreversibility of C_5-Epi depends on the sulfation state of the glucosamine residue that sits three residues to the non-reducing end of the GlcUA to be modified (100). This glucosamine at this site is termed the mode of reaction recognition site (MRRS), while the GlcUA is termed the epimerization product site (EPS) (Figure 12). When the MRRS is *N*-sulfated or *N*-unsubstituted, the EPS will remain reversible and can interconvert between GlcUA and IdoUA. However, when the MRRS is *N*-acetylated the EPS will become locked into an iduronic acid conformation (100). This data proves the hypothesis that the state of the glucosamine residues, either *N*-acetylated, *N*-sulfated or *N*-unsubstituted, are critically important for subsequent epimerization reactions to occur.

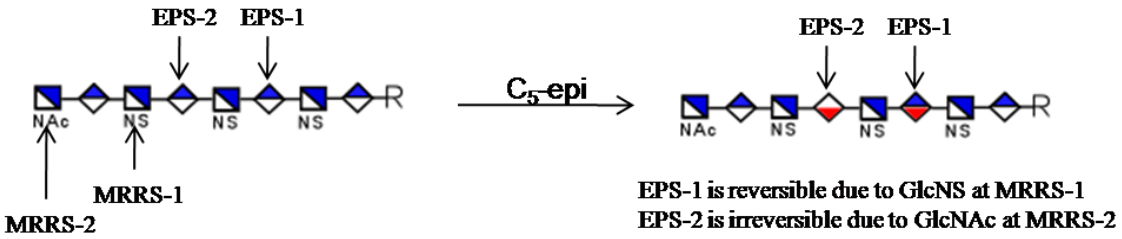

Figure 12. C_5-Epi mode of irreversibility. C_5-Epi is able to act in an irreversible mode when a GlcNAc residue is present at the location three monosaccharide units toward the nonreducing end (MRRS-2), locking the uronic acid residue into IdoUA conformation (EPS-2). If a GlcNS residue is located three monosaccharide units toward the nonreducing end (MRRS-1) of the uronic acid residue, the site will remain reversible and maintain a 50% chance of being either GlcUA or IdoUA (EPS-1).

Uronosyl 2-O-Sulfotransferase

Uronosyl-2-*O*-sulfotransferase is closely related to C_5-Epi because it has the ability to catalyze the transfer of a sulfo group to the 2-*O* position of either IdoUA or GlcUA (Figure 13) (101). It is the only sulfotransferase to modify the uronic acid residue in the HS biosynthetic pathway. The relationship of C_5-Epi and 2-OST has been demonstrated in mutant CHO cells defective for 2-OST. C_5-Epi was localized to the endoplasmic reticulum without 2-OST expression, but upon reintroduction of 2-OST, these proteins seemed to form a complex and move to the golgi apparatus. In addition, there was no epimerase activity in the mutant cells until the introduction of 2-OST, suggesting that these enzymes depend on one another for stability, localization, and function (48). 2-OST has a preference for IdoUA and is also associated with *N*-sulfation as it only modifies uronic acid residues with *N*-sulfation present at the adjacent GlcNS residue linked at the nonreducing end. Like C_5-Epi, there is only one isoform of this sulfotransferase in the human genome. It has a molecular weight of 43 kDa, consisting of 356 amino acids (102). Studies have confirmed that this gene is essential for development and sustaining life. 2-OST $^{-/-}$ mice displayed flaws in development of the

eyes and kidneys resulting in death during the neonatal period (103). Researchers have also demonstrated that 2-OST is critical for cell and axon migration in *C. elegans* (104). 2-OST has also shown the capability to decrease NDST activity, as CHO cells containing a 2-OST deficient mutation were shown to have more GlcNS residues than wild-type (105). Taken together, all of the data so far points to regulation and interaction between various HS biosynthetic enzymes. The HS polymerase enzymes, NDSTs, C_5-Epi, and 2-OST all have influence on one another and most likely these reactions are all occurring simultaneously as the polysaccharide is being constructed.

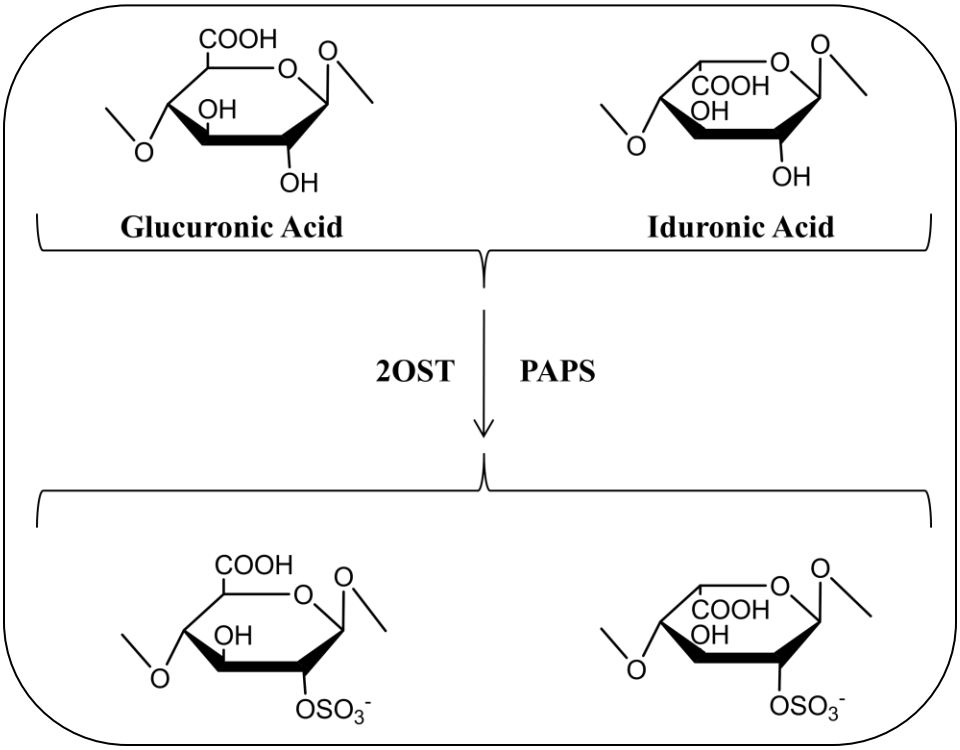

Figure 13. 2OST Reaction. 2OST catalyzes the transfer of a sulfo group from the donor molecule PAPS to the 2-*O*-position of the uronic acid monosaccharide unit. It has the ability to transfer to both GlcUA and IdoUA, but has shown a preference for IdoUA.

Glucosaminyl 6-O-Sulfotransferase

Beyond the 2-OST modification is where the true HS protein binding assembly takes place. *O*-sulfation at the glucosamine unit has long been understood as a pre-requisite for protein binding and downstream signaling events to occur (106). Seeing as how there is only a single isoform of 2-OST, it remains unable to produce a distinct binding code for proteins to read and interact. The key players in creating distinct HS sequences are 3-*O*-sulfotransferase and 6-*O*-sulfotransferase. Between these two enzymes there are ten isoforms, each with its own substrate binding capabilities and modification profiles (11). These are the steps that truly define the specific patterns that make up the heterogeneous HS polysaccharides, resulting in such a wide array of biological events.

6-*O*-sulfotransferase catalyzes the addition of a sulfo group from the sulfo donor PAPS to the 6-*O*-position of glucosamine (Figure 14). There are three isoforms of 6OST-(1-3), and each isoform is predicted to maintain type II transmembrane topology consisting of 401, 506, and 470 amino acids respectively. They display between 50-57% homology across each isoform (106). Regulation of 6-*O*-sulfation has been implicated as a major factor in HS binding to and signaling FGF2 and FGF1, resulting in cell differentiation (107). Sulfation at the 6-O-position has been observed in 2 of the characteristic HS domain structures, the NS and NS/NAc domains. Considering that the NS domains normally have 2-*O*-sulfation at the IdoUA residue, while in the NS/NAc alternating domain there is no 2-*O*-sulfation, the 6-*O*-sulfation is occurring in very different contexts. Studies using all three isoforms have shown that 6-*O*-sulfation can occur at both the GlcNS and GlcNAc locations (108, 109). It was also noted that in the

context of a GlcUA or IdoUA at the nonreducing end of the GlcNS residue, IdoUA was the favored substrate. Moreover 6-*O*-sulfation occurred more favorably at sites that were linked to a 2-*O*-sulfate uronic acid residue (109). The critical differences among these isozymes are their target substrate preferences.

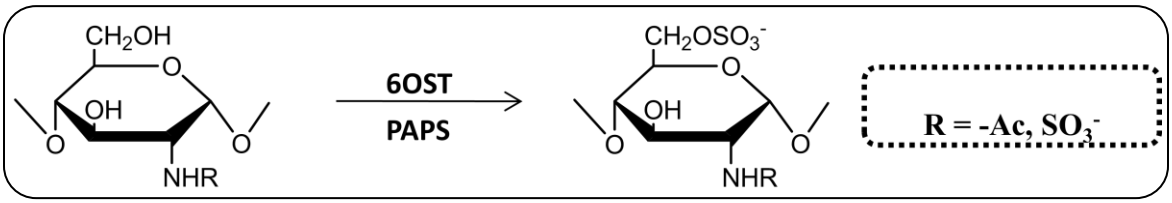

Figure 14. 6OST reaction. 6OST catalyzes the transfer of a sulfo group from the donor molecule PAPS to the 6-*O*-position of glucosamine. The reaction may proceed independent of the sulfation state of the 3-*O* and *N*-positions.

Each isoform differs in its preference of uronic acid at the nonreducing end of the glucosamine residue to be modified. 6-OST-1 seems to prefer IdoUA-GlcNS, but has also been shown to modify GlcNAc. 6-OST-2 can modify both IdoUA-GlcNS and GlcUA-GlcNS, but is very dependent on substrate concentration. 6-OST-3 has the ability to act on both substrates with no dependency on concentration (106, 110). These substrate specificities seem to imply that 6-OST-1 performs most of the modification in the NS domain, where IdoUA is most prevalent, while 6-OST-2 and 6-OST-3 most likely modify the glucosamine residues within the NS/NAc domain of HS. The expression of each isoform also seems to be tissue specific and may imply that each isoform is responsible for creating structures that will be used in signaling proteins within each specific physiological location. Northern blot analysis on murine tissue has shown that the expression of 6-OST-1 is greatest in the liver, where there is a large abundance of the disaccharide 2SIdoUA-GlcNS6S. 6-OST-2 is mainly expressed in the brain, corresponding to the elevated level of GlcUA-GlcNS6S disaccharide (106).

Glucosaminyl 3-O-sulfotransferase

Glucosaminyl 3-*O*-sulfotransferase is arguably the most crucial modification for protein binding and signaling. 3-OST transfers a sulfo group from the donor PAPS to the 3-*O*-position of glucosamine (Figure 15). Unlike the other HS biosynthetic enzymes, 3-OST-1 is not a transmembrane protein. Instead, it resides inside the lumen of the golgi apparatus and lacks a cytoplasmic domain (111). 3-OST-1 does, however, maintain approximately 50% homology to the NDSTs, mostly in the C-terminal sulfotransferase domain (112). 3-*O*-sulfation is present in binding sequences of antithrombin and the viral gD-envelope protein of HSV-1. It is a relatively rare modification, accounting for approximately 0.5% of HS sulfation, and its presence creates finite structures for HS ligand binding (30, 36). The ability of 3-OST to provide such an array of unique structures that various proteins can recognize and bind most likely comes from its full arsenal of seven distinct isoforms in the human genome, 3-OST-1, 3-OST-2, 3-OST-3A, 3-OST-3B, 3-OST-4, 3-OST-5, and 3-OST-6 (111, 113-115). Each isoform has the ability to transfer a sulfo group to the 3-*O*-position of glucosamine, but the modification depends on the conformation and sulfation state of the uronic acid linked to the nonreducing end (116a). Studies in zebrafish have revealed that some of the isoforms are expressed ubiquitously, while others are localized to very specific tissues (117). This data provides the notion that certain isoforms provide certain tissue specific modifications to finite HS polysaccharide structures that are needed for specific protein signaling within that location.

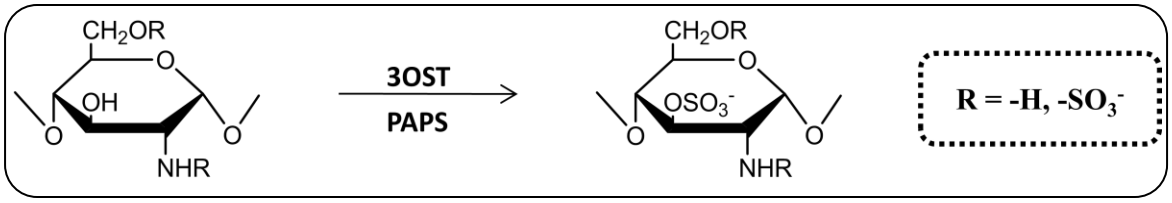

Figure 15. 3OST reaction. 3OST catalyzes the transfer of a sulfo group from the donor molecule PAPS to the 3-*O*-position of glucosamine. The glucosamine must either be sulfated or unsubstituted at the *N*-position for the reaction to proceed.

Perhaps the most extensively studied protein binding site within HS structures is the antithrombin-binding pentasaccharide. This structure contains a critical 3-*O*-sulfation that is required for binding to antithrombin, creating a conformational change in the protein that enhances its affinity for thrombin resulting in inhibition of blood coagulation (118). In addition, 3-*O*-sulfation has been implicated in the binding of herpes simplex virus type 1 to the cell membrane through interaction with the viral glycoprotein gD, resulting in the initiation of viral entry (32). Interestingly, a very similar modification has also been implicated in the binding of cyclophilin B, a cyclosporine A-binding protein, which signals for migration and integrin-mediated adhesion of peripheral blood T lymphocytes (119). As one can see, the 3-*O*-sulfo group is very important for protein signaling and based on the variations provided by the different isoforms, researchers have only scratched the surface in terms of proteins that bind these sites in HS polysaccharides.

Human 3-OST-1 consists of 307 amino acid residues and has a molecular weight of approximately 36kDa. It has a 93% homology to mouse 3-OST-1. 3-OST-1 is believe to be a key player and provide the final modification in the development of the AT-binding site (111). This isoform is capable of recognizing and sulfating the disaccharides GlcUA-GlcNS and GlcUA-GlcNS6S, as well as IdoUA-GlcNS and IdoUA-GlcNS6S, provided there is no 2-*O*-sulfation present (111, 120). It has been postulated that the

presence of 2-*O*-sulfation at the nonreducing end of the GlcNS residue to be modified will inhibit catalysis by 3-OST-1. Mice knockout studies have shown that it is not the only enzyme capable of creating the AT-binding domains and that there may be some level of cooperation between the various isoforms (121). They also showed lethality of this knockout in a specific genetic background, in addition to intrauterine growth retardation.

A crystal structure of 3-OST-1 has been solved at 2.5Å resolution bound to PAP (116a). This structure displayed striking similarities to the structure of NST-1. It displays the common α/β fold that is consistent with structures of most other sulfotransferase enzymes. Based on site-directed mutagenesis studies it was determined that the residues Arg-67, Lys-68, Arg-72, Glu-90, His-92, Asp-95, Lys-123, and Arg-276 are all necessary for the enzyme to maintain activity. When compared to the structure of NST-1, it was determined that Arg-67, Arg-72, His-92, and Asp-95 are conserved among each 3-OST isoform. These residues are not conserved in the structure of NST-1, thus may play a role in substrate recognition of 3OST-1 (116a).

Recently, two additional crystal structures have been solved which shed light on the substrate binding specificities of 3-OST-1 and 3-OST-3 (116b). The first structure is a ternary complex between 3-OST-1 and a heptasaccharide substrate. The second structure is a ternary complex between 3-OST-3 and a tetrasaccharide substrate. Both substrates maintain a common trisaccharide motif of GlcNS6S-IdoUA2S-GlcNS6S. It is evident that the substrates bind 3-OST-1 and 3-OST-3 in a different orientation through a change in the configuration of the IdoUA2S residue in the trisaccharide motif. This monosaccharide unit maintains a $^{2}S_{0}$ skew boat conformation when bound to 3-OST-3,

while adopting the $^{1}C_{4}$ conformation when binding to 3-OST-1. This conformational difference results in a change in the positioning of the reducing end of the substrate causing 3-OST-3 interaction with the substrate to be mediated by a metal ion, while 3-OST-1 maintains a more direct interaction. In addition to differences in substrate binding, the substrate specificity between the two isoforms is clearly mediated by the 2-*O*-sulfation at the uronic acid residue to the nonreducing end of the 3-*O*-sulfation site. For 3-OST-1, the GlcUA cannot be 2-*O*-sulfated due to the placement of this residue ~2.8Å from the carboxylate on the reducing end IdoUA2S, resulting in charge repulsion due to the close proximity of these residues after binding the enzyme. But after binding to 3-OST-3 the charge repulsion is less because the negatively charged groups are now 3.2Å away, and there is an additional hydrogen bonding interaction with Lys259. Thus, the specificities between 3-OST-1 and 3-OST-3 are clearly due to a difference in substrate binding.

3-OST-2 maintains a substrate preference of GlcUA2S-GlcNS and IdoUA2S-GlcNS (31, 122). Studies have shown that 3-OST-2 function is activated in the pineal glands of rats only during the daytime (123, 124). The physiological significance of this observation remains to be fully understood. 3-OST-3A and 3-OST-3B catalyze the addition of a 3-*O*-sulfation to the disaccharides containing the sequences IdoUA2S-GlcNS and IdoUA2S-GlcNH$_2$. The 3-OST-3 isoforms are responsible for the 3-*O*-sulfation that is seen in the gD-binding sequence that binds herpes simplex virus type 1, allowing for viral entry into the cell (32). 3-OST-3B has been found to be an important player in the notch signaling pathway in drosophila. Interfering RNA was used to disrupt 3-OST-3B translation, resulting in neurogenic phenotypes with a reduction in levels of

Notch protein (125). 3-OST-2 and 3-OST-3A/B are believed to not be involved in the formation of the AT-binding domain. Cell extracts that were transfected with these enzymes showed 300-fold lower activity towards antithrombin versus cell extracts transfected with 3-OST-1 (31). 3-OST-4 and 3-OST-6 have also demonstrated the capability of creating a binding site to facilitate HSV-1 fusion in CHO cells (115, 126). 3-OST-5 has demonstrated the unique capability of sulfating both substrates that will lead to AT-binding sites, as well as, those that will lead to HSV-1 entry (114).

Researchers have attempted to decipher the mechanisms by which the isoforms, in particular how 3-OST-1, 3-OST-3, and 3-OST-5, recognize their various preferred substrates to carry out their modification. The substrate recognition mechanism is a result of amino acid residues at two sites: the catalytic site and the gate (127). Residues at the catalytic site can recognize monosaccharide units adjacent to the acceptor glucosamine, but these residues are all conserved among the 3-OST isoforms. Thus, it was determined that the bulk of substrate specificity was most likely coming from the gate residues of each isoform. These amino acids interact with the monosaccharide unit that is three sugar units to the nonreducing end of the acceptor monosaccharide (127). Point-mutations at the gate amino acids were introduced into both 3-OST-1 and 3-OST-3 isoforms. 3-OST-1 mutations consisted of E88G and H271G, essentially decreasing the size of these gate amino acids. The resultant mutant displayed more 3-OST-5 like activity by modifying polysaccharides to have reduced activity towards antithrombin and greater interaction with the gD viral envelope protein (127). In contrast, they installed point mutations in the gate amino acids of 3-OST-5 in an attempt to make it have more 3-OST-1 like properties. These mutations essentially replaced small amino acids with

larger ones, consisting of S120E and A306H. These mutants displayed the ability to produce polysaccharides with more AT-binding ability and less interaction with the gD viral protein (127). These site-directed mutagenesis studies did not completely shift the activity one way or the other, thus there seems to be more unknown factors playing a role in the substrate recognition mechanisms of the 3-OST isoforms. These postulated factors could be charge/charge interactions or conformational flexibility resulting from the incorporation of the IdoUA residue (127).

BIOSYNTHESIS OF HEPARIN

Considering that most of the biosynthetic enzymes which define heparan sulfate structure are also involved in the biosynthesis of the structural analog heparin, a key question to ask is what sets heparin apart? Given that heparin is only produced in mast cells and that heparan sulfate is produced nearly everywhere else in the body, there must be something special about the mast cell.

The structure of HS is governed by the presence of three basic domain regions. The NAc domains consist of mostly GlcUA-GlcNAc repeating disaccharide units. The NAc/NS domains consist of random intermittent GlcNS residues mostly linked to GlcUA. Finally, the NS domains consist of long stretches of GlcNS residues (128). These stretches, however, are not fully sulfated at all positions and seem to vary depending on tissue type and location (129, 130). The structure of HP is essentially one long mostly homogeneous NS domain, as it contains upwards of 90% IdoUA2S-GlcNS6S disaccharide units (21). Based on these observations, it is clear that there must

be some difference in the regulation of biosynthesis for these two structurally analogous GAGs.

To begin to understand the difference between HS and HP biosynthesis, one should look at the expression of HS biosynthetic enzymes in the mast cells versus expression in epithelial cells. Considering the linkage region to the core protein serglycin is identical to those in HS, CS, and DS, one would expect that the enzymes involved in initiation are most likely the same. Also given that 2-*O*-sulfation and C_5-epimerization are present in HP structure and that there is only a single isoform of 2-OST and C_5-epimerase in the human genome, this step must also be consistent with HS biosynthesis (94, 102). As far as the other *O*-sulfation reactions, there appears to be some variance. Researchers showed that 6-OST-$1^{-/-}$ mice were able to maintain a normal level of trisulfated HP disaccharide unit when compared with wild type mice. This result reveals that HP 6-*O*-sulfation is most likely maintained by 6-OST-2 and 6-OST-3 (131). Based on the much understood knowledge of 3-*O*-sulfation as a requirement for HS binding to antithrombin, and the fact that HP obviously possesses these same AT binding sites, one would expect 3-OST-1-like activity in the mast cells. In fact, it has been discovered that in an immortalized mouse mast cell line, C57.1, there is an overexpression of 3-OST-1 (111). The presence of other 3-OST isoforms in mast cells is not well understood.

The truly appreciable difference in enzyme expression between mast cells and epithelial cells comes from the NDSTs. It is known that the mRNA transcripts for both NDST-1 and NDST-2 are observed ubiquitously among all tissue types (81). It has also been demonstrated that NDST-$1^{-/-}$ mutation is neonatal lethal & show a deficiency of HS formation, indicating this isoform is critical for HS biosynthesis (82). In addition, Sheng

et al was able to demonstrate the role that this isoform plays in forming the characteristic domain structures that define HS structure (26). In comparison, NDST-2 has shown some ability to compensate for a deficiency of NDST-1 in CHO cells, although not to the full extent we would expect for fully functional HS (132, 133). NDST-$2^{-/-}$ mice only exhibited abnormal mast cell morphology due to a reduced amount of histamine and proteases in their secretory granules, with no other obvious developmental abnormalities (134). Meanwhile HS extracted from various tissues of the NDST-$2^{-/-}$ mice is unchanged (129). Researchers have also demonstrated that NDST-2 is the major isoform in mouse mastocytoma, while the NDST-1 transcript could hardly be detected (79). In addition, pre-existing sulfation has shown to increase the activity of NDST-2, perhaps providing a processive mechanism by which long strands of GlcNS residues are produced in HP synthesis (89, 90). Taking all of this data together, NDST-1 is responsible for HS synthesis and NDST-2 is responsible for HP synthesis.

If the NDST enzymes are responsible for differentiating between HS and HP biosynthesis, but both mRNAs are expressed ubiquitously and abundantly, there must be some type of control on which enzyme will act on the polysaccharide substrate. More specifically, what determines why NDST-2 is the active isoform in mast cells, while NDST-1 is active in all other cells? There is some evidence regarding transcriptional, translational, and post-translational control of the NDST isoforms. The transcription factor GA-binding protein has shown capability of binding NDST-2 to a GGAA motif in the 5'-UTR and increasing its expression in skin mast cells of mice. Researchers found that a mutation in the *mi* allele caused disruption of localization of the GA-binding protein through a faulty *mi* transcription factor. This mutation resulted in deficient HP

synthesis in the mast cells of the *mi*/*mi* mice (135). There also appears to be a translational control mechanism, which could help to explain the presence of high levels of NDST-2 mRNA, but lack of enzyme influence on the modification of HS (81). Each of the NDST isoforms maintains unusually long sequences in the 5'-UTR region, that have predicted secondary structures, and contain multiple AUG start codons that result in short open reading frames. These mRNA features present a problem for the initiation of cap-dependent translation. *In vivo* observations have shown that each isoform can be differentially expressed between various cell types and culture conditions. The 5'-UTR of NDST-2, NDST-3, and NDST-4 have been found to possess internal ribosome entry sites, which could explain a method of cap-independent initiation of their translation and expression (81). Take together, this data reveals that there are many levels of enzyme regulation that occur between cell types that control which isoform is to be expressed, even after transcription has taken place. In fact, a database screening has revealed that many of the HS biosynthetic enzymes maintain a very similar 5'-UTR to that of the NDSTs and most likely are under the same regulatory processes (81). Translational control has also been demonstrated for HS biosynthetic enzymes in *Drosophila* (136). This may explain why HS that resides among the same tissues maintains a very similar modification pattern, while across different tissues will vary greatly, and results in HP production in the case of mast cells. In addition to transcriptional and translational control, there also appears to be post-translational control of the NDST isoforms as researchers have found that glycosylation of the NDST-2 protein is essential for cellular activity (137). This same effect of post-translational control *via* glycosylation has been seen for CS biosynthetic enzymes as well (138).

GAGosome model

Based on the close association and in some cases physical interaction of many of the various HS biosynthetic enzymes, Esko and Selleck proposed a macromolecular enzyme physical complex that is formed during the biosynthesis of HS and HP, termed GAGosome (10). This complex would depend on physical interactions between the biosynthetic enzymes, as well as concentrations that are determined based on regulation of enzyme expression through a variety of means. Researchers have shown these interactions to be prevalent throughout the synthesis of HS polysaccharides. During the formation of the tetrasaccharide linkage region, there is a physical interaction between XylT and GalT-I (139). Examples of interaction during the polymerization phase include EXT1 and EXT2 as they appear to form a complex that localizes to the Golgi upon co-expression (65, 66). In addition, NDST-1 has been linked to interaction with EXT2 as the expression pattern of the HS polymerase enzymes effects the expression of NDST-1. As mentioned earlier, it is believed that NDST-1 competes for binding with EXT1 to EXT2 thus providing another example of physical interaction (69). Finally, 2-OST and C_5-epimerase have also been shown to be completely dependent upon one another and also form a physical complex (48). In addition, it also appears that the reversibility of C_5-Epi is completely dependent on the sulfation state of the glucosamine residue that is 3 monosaccharide units to the nonreducing end, thus making the case for elongation, *N*-deacetylation/*N*-sulfation, epimerization, and 2-*O*-sulfation all occurring simultaneously (100). It has also been speculated that the lack of a transmembrane domain in the 3-OST isoforms indicate that they may also form complexes with other HS biosynthetic enzymes *in vivo* (111). With all of these physical interactions and co-factor like dependencies

among the various HS biosynthetic enzymes, it seems the GAGosome model is a viable alternative to the stepwise layout of HS synthesis proposed in the past. This model would also help to define how the substrate is modified in a dynamic fashion, moving quickly from enzyme to enzyme, all while the polysaccharide chain is being elongated.

Researchers have also fairly recently added more speculation to this model in regards to HS versus HP biosynthesis. Upon their analysis of HS structure derived from mouse liver that was either NDST-$1^{-/-}$ or NDST-$2^{-/-}$, they postulated that NDST-1 is the preferred NDST isoform that is incorporated into the GAGosome (Figure 16) (140). They speculate that because NDST-2 is abundantly expressed in the mast cells, while the mRNA transcript for NDST-1 is nearly non-existant, it becomes incorporated into the complex and HP production occurs (79).

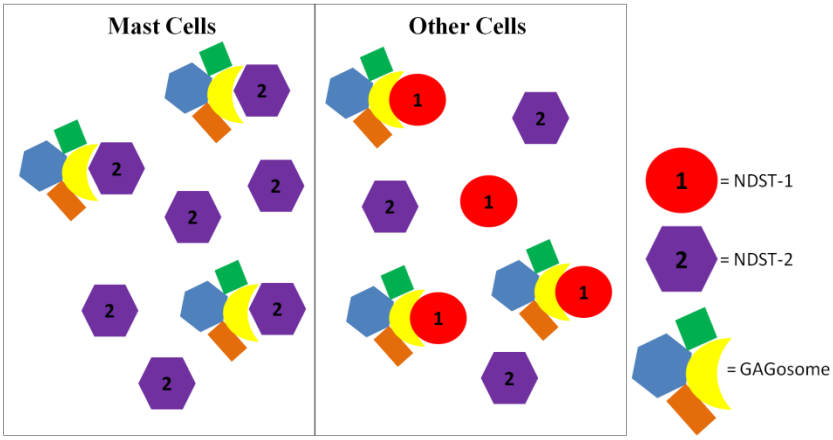

Figure 16. GAGosome biosynthetic model. The GAGosome model assumes that the physical interactions observed between various HS biosynthetic enzymes may be part of a larger macromolecular physical complex. This complex would allow for regulation among the enzymes, in addition to quick processing of fully functional HS and HP biomolecules. It is also postulated that NDST-1 is the preferred isozyme for the complex and is present in most cells, which synthesize HS. While in mast cells NDST-2 becomes incorporated due to its abundant expression and a deficiency of NDST-1, making way for HP synthesis. Figure adapted from (137).

BIOLOGICAL FUNCTIONS OF HEPARAN SULFATE

The vast array of heparan sulfate structures that are made possible through the various HS biosynthetic enzymes and their multiple isoforms has lead researchers to the understanding that HS plays critical roles in signaling and regulation of many various physiological processes. Several of these interactions have been thoroughly characterized. The first interaction and perhaps best characterized is the binding of HS and HP to antithrombin and regulation of the blood coagulation cascade. HS has also been demonstrated to be a key regulator of cell growth, proliferation, and differentiation through interactions with various fibroblast growth factors. HS also has been shown to play key roles in pathophysiological processes. It provides a mechanism of fusion with the herpes simplex virus type I gD viral envelope protein, allowing for viral entry and proliferation within host cells. Each one of these characterized processes depends on unique structures within the HS polysaccharide in order to take place. The search for new interactions is ongoing and researchers have only begun to scratch the surface of protein interactions and regulatory processes for which HS and HP are responsible.

Anticoagulation

Blood flow throughout the human body is the result of a finely tuned balance of clotting factors and anticoagulant agents. Coagulation is a process by which a cascade of serine protease pro-enzymes cleaves one another leading to their activation and results in the formation of a platelet-lined fibrin clot (Figure 17) (141). On the other hand, human physiology has the capability to counteract the clotting mechanism through a series of anticoagulation factors that are able to bind the serine proteases rendering them unable to

form a blood clot at times when it is not needed. The coagulation cascade can be initiated upon injury, or extrinsically, where exposure to tissue factor or collagen begins the serine protease cascade. The cascade can also become initiated intrinsically, when the vasculature becomes damaged, initiating the activation of tissue factor XII. Factor IIa, commonly referred to as thrombin, performs the final cleavage of fibrinogen into fibrin resulting in the formation of a blood clot (142). There are also many anticoagulant factors that are present in higher concentrations *in vivo* to keep the blood flowing normally when there is no vascular injury. These proteins are commonly referred to as serpins, or serine protease inhibitors. They include antithrombin, heparin cofactor II, protein C inhibitor, and tissue factor pathway inhibitor. Each of these proteins has a common structural feature by which they perform their inhibitory activities, the reactive center loop. This loop maintains a protease substrate site that the enzyme is able to recognize and cleave, resulting in a covalent joining of the serpin to the protease and permanent inhibition (143-145). In addition to the structural homogeneity among the serpins, they also maintain the ability to undergo dramatic structural rearrangement to both the serpin and protease. Finally, serpins are also greatly affected by the binding of glycosaminoglycan molecules, such as HP and HS, which serve to enhance their affinity towards their target proteases (146).

Perhaps the most well known and extensively studied of the serpins is antithrombin (AT). AT is widely accepted as the most important anticoagulant serpin due to its high affinity for thrombin, as well as its ability to inhibit factor Xa (147). While it does act on other various protease tissue factors, these are the two most important and influential interactions in the cascade. The importance of AT in regulation

of blood coagulation was studied in mice where ATIII$^{-/-}$ mice displayed embryonic lethality. This was shown to be due to an abundance of fibrin deposition in the liver and myocardium. These results demonstrate the critical role that AT plays in the regulation of blood flow (148). HP has the ability to bind to antithrombin, resulting in activation that dramatically increases its affinity for factors IIa and Xa. In fact, the binding affinity for thrombin and factor Xa is increased 9,000 and 17,000 fold respectively upon binding HP (149). It is widely accepted that a unique pentasaccharide sequence within the structures of HP and HS is responsible for the binding to AT. This structure is GlcNAc-GlcUA-GlcNS3S6S-IdoUA-GlcNS6S (Figure 3) (150). The most critical modifications in the AT-binding pentasaccharide are the presence of the 3-*O*-sulfation on the central glucosamine moiety, in addition to 6-*O*-sulfation on the nonreducing end glucosamine. When 3-O-sulfation is absent, the pentasaccharide displays a decrease in activity towards binding AT of approximately 20,000 fold (151). In the absence of the 6-*O*-sulfation, all AT-binding activity is lost (152). It has also been demonstrated that HP lacking the pentasaccharide structure can bind to AT, however at much higher concentrations (153).

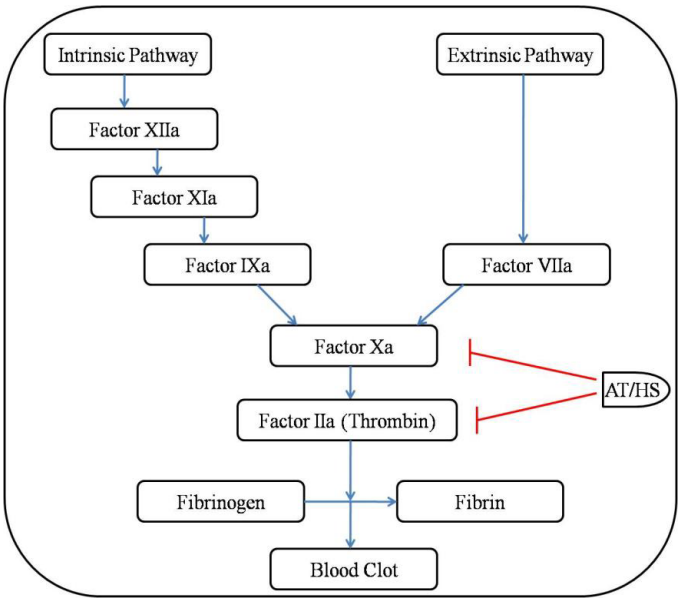

Figure 17. Blood coagulation cascade. The blood coagulation cascade consists of inactive serine protease enzymes. The pathway can be triggered by an intrinsic or extrinsic signal, by which the proteases are cleaved to form activated enzymes which cleave the next enzyme and so forth. The two pathways converge at factor Xa, which is responsible for activating thrombin. Thrombin directly cleaves fibrinogen into fibrin resulting in a blood clot. HS containing an AT binding sequence has the ability to inhibit both Factor Xa and thrombin, thus inhibiting the formation of a blood clot.

The action of HP binding to AT activates the serpin, upon which it is released into the plasma. HP binds into the active site of AT through hydrogen bonding interactions between the sulfate groups and carobxylate groups on the HP molecule with an arrangement of basic residues across multiple helices in the protein that are perfectly spaced to accommodate the pentasaccharide (Figure 18) (154). After binding to AT, conformational changes are induced both locally around the binding site, as well as, throughout the enzyme which increases its affinity for thrombin. The most important of these conformational changes causes the reactive center loop, containing the protease cleavage site, to be extended outward (146). This extension results in greater exposure to thrombin, upon which thrombin will cleave the reactive center loop and form an irreversible covalent linkage between the two proteins. This linkage results in a conformational change in thrombin and permanent inactivation of the thrombin protease capabilities (143). However, in order for the inactivation to take place, there must also be another interaction between the HP molecule bound to AT and a specific exosite on the thrombin protein. Therefore, for inactivation of thrombin and complete anticoagulant activity, the HP molecule must consist of the AT-binding pentasaccharide in addition to a minimum of 12 non-specific sugar residues linked at the nonreducing end that is able to form a ternary complex between AT, HP, and Thrombin. This minimum linkage is termed the C-domain (155).

AT also has the ability to interact with factor Xa through a very similar process. Factor Xa also has the capability to cleave the reactive center loop and cause conformational change within the protease. However, factor Xa has the ability to interact with AT through a direct protein-protein interaction at a specific exosite and does not

need the HP molecule to aid in the interaction (156). Therefore, many low-molecular weight HPs that are currently used therapeutically are only able to exhibit anti-Xa activity due to their inability to form the ternary complex with thrombin. However, since naturally occurring HP is much longer it is able to utilize the full anticoagulant properties of activated AT.

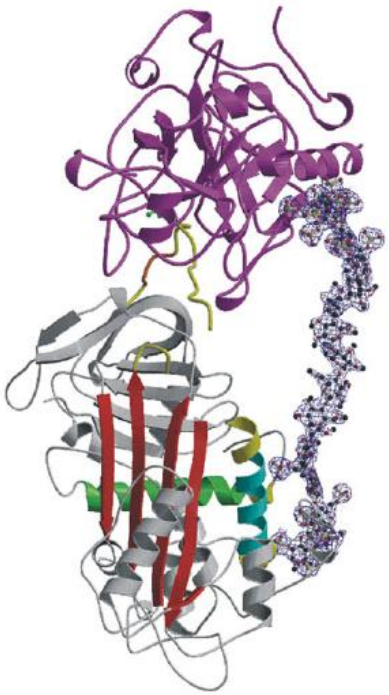

Figure 18. Ternary complex of Antithrombin, Thrombin, and a Heparin mimetic. Antithrombin is represented as a ribbon diagram with helices colored green and cyan, the central β-sheet red, and random coils grey. Thrombin is colored magenta. The heparin mimetic is represented as a ball and stick diagram. Binding of the AT-binding pentasaccharide region causes the reactive center loop (yellow) to become exposed, thereby cleaved and covalently bound to thrombin. In addition, the HP mimetic bridges the two structures by binding to an exosite on thrombin, creating a stable ternary complex. Structure taken from (137).

HP/HS-activated AT also has the ability to inhibit other proteases in the blood coagulation cascade. Researchers have demonstrated the tissue factor IXa also binds to activated thrombin in much the same manner as factor Xa (157). It binds to the same exosite on AT and utilizes the reactive center loop. However, factor IXa was found to also possess a binding site for HP, thus displaying properties of thrombin binding as well (158). Taken with the understanding of how thrombin interacts with AT through a HP bridge, it can be postulated that the large size of the HP molecule indeed plays a large

role in full anticoagulant effect versus an AT-binding pentasaccharide alone. HP has also been shown to cause inhibition of factor XIa through both AT-mediated and protein C-mediated methods, in addition to direct HP binding (159, 160).

HP also has been shown to play somewhat less significant roles in the activation of other serpins which act against proteases in the blood coagulation cascade. Heparin cofactor II is activated by HP and is also able to form a complex with thrombin leading to inactivation (161). However, a key difference lies in its ability to not rely on a HP bridge for interaction with the exosite on the thrombin protein, instead employing its *N*-terminal tail to perform the required interaction with an exosite on thrombin for its complete abolishment of anticoagulant properties (162). Although HP cofactor II does selectively inhibit thrombin, it does so at an order of magnitude lower rate than does AT (163). HP has also been shown to bind to protein C inhibitor, resulting in the inhibition of thrombin, factor Xa, and factor XIa (164, 165). Finally, HP has the ability to bind tissue factor pathway inhibitor, the most important inhibitor of the extrinsic pathway of the coagulation cascade, resulting in inhibition of factors VIIa and Xa (166). It is clear that HP plays many roles in the activation of serpins that play crucial roles in the formation of blood clots.

Cell Proliferation and Differentiation

Fibroblast growth factor (FGF) consists of a family of 22 structurally related growth factors that are able to bind to and activate the receptor tyrosine kinases FGFR-1, FGFR-2, FGFR-3, and FGFR-4 (167-170). The binding of FGF to the FGFR will transmit a signal to the cell that will mediate a specific response. The FGF family is very

important throughout the life of an organism. They play critical roles in cellular differentiation, migration and proliferation during embryonic growth. The FGF family also is important for regulating the nervous system, repair of wounds, and tumor angiogenesis (171). HP and HS have been shown to aid in the binding of the FGF to the FGFR, thus playing a critical role in regulation of each of the functions mentioned previously (172-174). Crystal structures have been solved which reveal the mechanism of receptor activation and the role that HP/HS play in regulating this process. Upon binding of FGF and HP to the FGFR induces dimerization, resulting in *trans*-autophosphorylation and stimulation of protein tyrosine kinase activity. HP is able to stabilize the complex by situating into a positively charged crevice of the D2 domains of the FGFR-1 dimer when FGF-2 is bound (Figure 19) (175).

Due to the nature of HS/HP interactions with the FGF and the biological relevance of FGF regulation, much attention has been paid to the structures of HP/HS that bind each of the growth factors thus regulating their downstream signaling events. Much of the work to date has focused on FGF-1 and FGF-2 due to extensive knowledge

Figure 19. Crystal structure of ternary FGF-2, HP, and FGFR1 complex. Ribbon diagram showing FGF-2 in yellow, D2 and D3 of the ligand binding portion in green and blue, and heparin in red. Taken from (175).

of their interaction with HP/HS. Researchers have determined that the minimal sequence required to bind FGF-2 is a pentasaccharide with the structure: UA-GlcNS-UA-GlcNS-IdoUA2S- (Figure 20) (176). The minimum structure for FGF-1 binding is quite different, mainly because it contains 6-*O*-sulfation. It consists of between 5-7 monosaccharide units and contains the critical trisulfated disaccharide unit IdoUA2S-GlcNS6S (177). It has been postulated that the presence of the 6-*O*-sulfation is a distinguishing feature between binding of the two isoforms and may represent a bias towards binding FGF-1 versus FGF-2 (178). It has been hypothesized that binding selectivity arises from an array of factors such as the length of the polysaccharide, level of IdoUA/GlcUA, in addition to the arrangement of the sulfo groups (16). Thus, research is ongoing in regards to defining true methods of selectivity between the FGF isoforms.

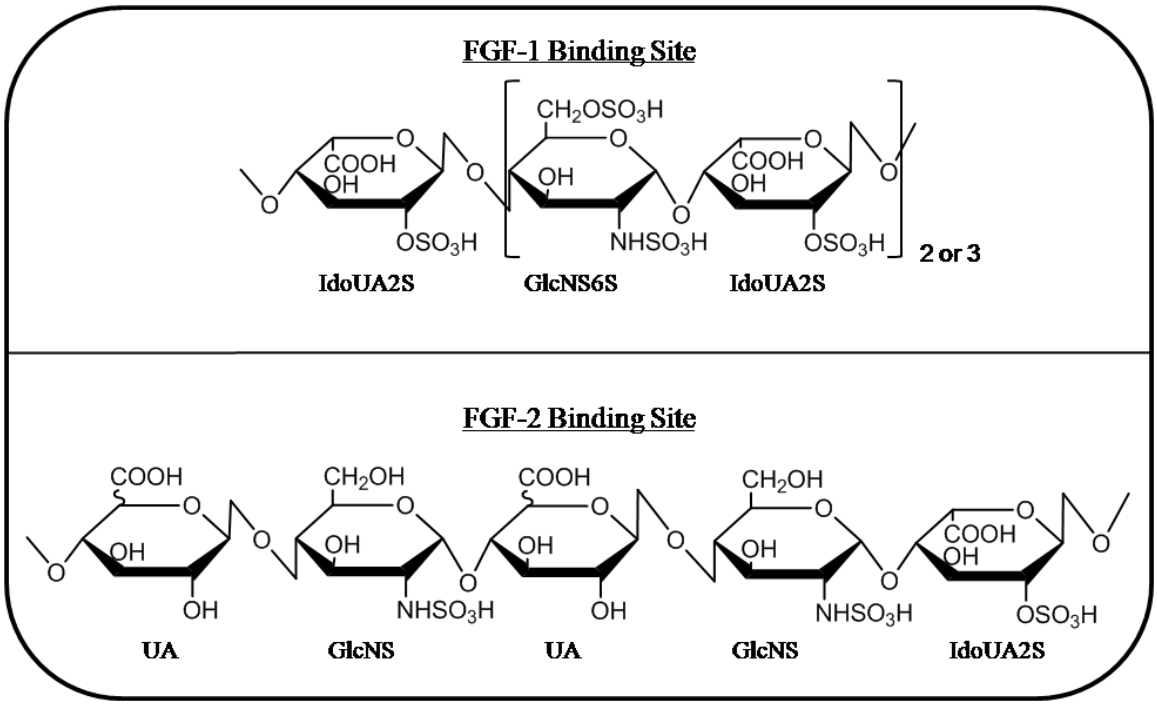

Figure 20. Chemical structure of heparan sulfate FGF-1 and FGF-2 binding sites.

Viral Infection

As stated previously, HS has been shown to serve as a docking site for herpes simplex virus type I through an interaction with the viral envelope protein gD. Upon binding by gD to HS and cell surface entry receptor, the viral particle fuses with the cell membrane, eventually gaining access to the cell, resulting in infection (179). Research into the exact structure required for binding the gD envelope protein has revealed that a rare 3-*O*-sulfation plays a key role (32, 114, 115). As mentioned previously, AT-binding in the blood coagulation cascade requires a 3-*O*-sulfation in the minimal pentasaccharide structure (150). However, the context of the monosaccharide structures adjacent to the target glucosamine residue determines both the substrate specificity of the 3-OST isoform to perform the modification, as well as, the function of the applied 3-*O*-sulfation. For gD

binding, the 3-*O*-sulfation must be linked to a IdoUA2S residue at the nonreducing end of the target glucosamine and the minimal length is an octasaccharide (Figure 21) (180). The main isoform responsible for carrying out this modification is 3-OST-3, however, 3-OST-5 also has the capability to produce this structure.

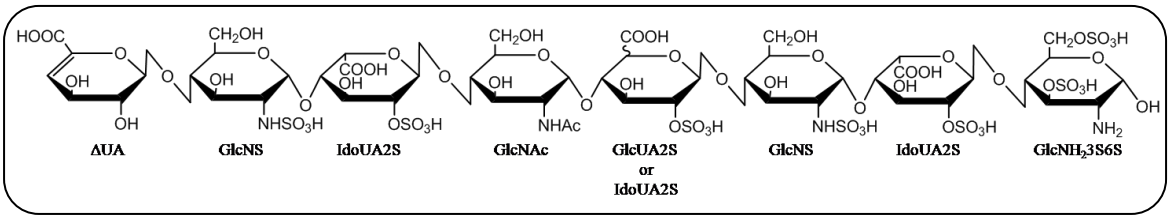

Figure 21. Chemical structure of gD-binding octasaccharide required for viral entry.

Inflammation

Inflammation is the body's physiological response to malicious events, resulting from bacterial infection, environmental exposure, and certain cancers (181-183). Initiation of the inflammatory response is a result of neutrophil recruitment to the affected site and migration into the tissue, whereby multiple events occur ranging from rearrangement of the vasculature to signaling for other various cells to respond (184). Neutrophils are the most abundant type of white blood cell and are the first to arrive on the scene when harmful stimuli are presented. The process by which neutrophils are recruited to the site of injury is known as chemotaxis. Chemotaxis occurs when the neutrophils are rolling along the blood vessel wall along a gradient of chemokines and cytokines, such as Interleukin-8, C5a, fMLP, and Leukotriene B4, which are released from endothelial cells at the site of injury. Once reaching the affected site, the neutrophils are able to transverse the cell membrane into the affected site and release its own cytokines, resulting in the recruitment of other various immune system cell types.

Neutrophils are able to function not only as a recruiter, but as a phagocyte that can digest foreign particles. In addition, they are capable of releasing anti-bacterial proteins through the process of degranulation, as well as the release of neutrophil extracellular traps that maintain anti-microbial properties (184).

HS plays a critical role in the recruitment of neutrophils to the affected site, as well as crossing from the blood vessel into the endothelial cell (Figure 22). The recruitment is initiated upon release of chemokines and cytokines from the affected cell, which are presented on the HS polysaccharide chains and attract the neutrophils along a gradient to the site of injury (185). In addition to the presentation of attractants, HS has been shown to interact with L-selectin presented on the outside of the neutrophil membrane. This interaction slows the free-flowing neutrophils from the blood, causing them to roll along the surface of the blood vessel up the attractant gradient, toward the site of injury (186). Finally, HS also stimulates the integrins, which are leukocyte adhesion molecules, to bring the leukocyte to a stop through adhesion at the epithelial membrane and movement out of the blood vessel (187, 188).

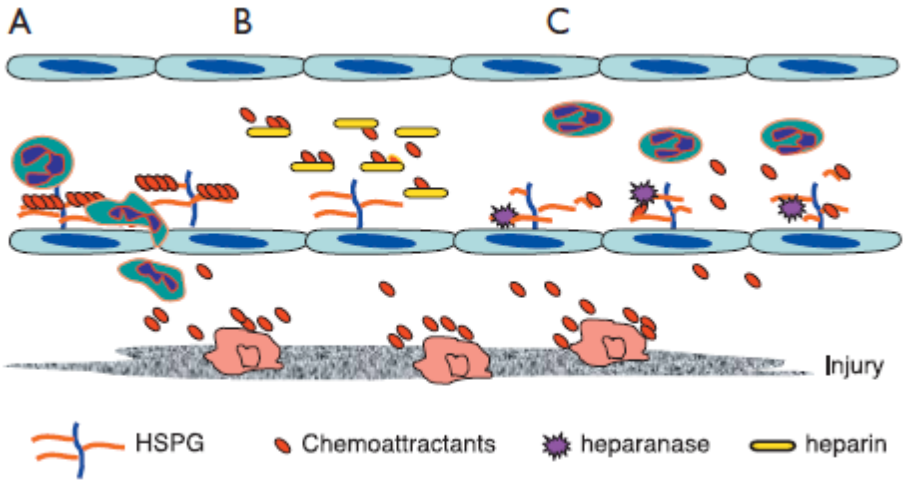

Figure 22. Roles of HS and HP in mediating the inflammatory response. (A) HS has the ability to sequester chemokines and cytokines, thereby recruiting the neutrophil to the site of injury and influencing transmigration out of the blood vessel. (B) When HP is administered it can result in an anti-inflammatory effect by competing for chemoattractant binding. (C) Upregulation of heparanase also results in an anti-inflammatory effect by shearing the HS to which the chemoattractants are bound. Taken from (184).

Each of these implications of HS for the inflammatory response has been previously well described. *In vitro* studies have revealed that Heparan sulfate can interact with several chemokines and cytokines *in vitro* such as, microphage inflammatory protein-1α, RANTES, interleukin-2, interleukin-8, and interleukin-10 (189-192). It has been speculated that the ability of HS to express such a variety of modification sequences may give it the ability to discriminate between chemoattractants in order to achieve the desired inflammatory response (193). Upon binding chemokines and cytokines, HS is believed to trigger integrin activation, leading to fusion of the neutrophil with the epithelial membrane. Studies with adult T leukemia cells have shown that HS is required for binding the chemokine macrophage inflammatory protein-1β, which then triggers integrin activation and adhesion to endothelial cells (194). Finally, interactions with selectins have also been characterized. Interestingly, the HS sequence that binds L-

selectin was found to contain an abundance of unsubstituted glucosamine residues, perhaps indicating that this rare modification is present in the binding motif for selectins (195). The binding of E-selectin was inhibited by the presence of HP and CS, indicating that it normally binds with HS under normal conditions (196). The Esko group was able to provide *in vivo* evidence of the effect of HS on neutrophil trafficking. They used a mouse model that was selectively deficient in NDST-1 in both neutrophils as well as endothelial cells, thus lacking functional HS in these cells. Their results showed that neutrophil trafficking was deficient in these mice due to a lack of interaction between L-selectin and the HS that was deficient on the endothelial membrane (86). Taken together, these results clearly demonstrate that HS plays a critical role in the inflammatory response.

Heparin has shown much clinical potential in controlling inflammation in a variety of disorders such as ulcerative colitis, bronchial asthma, and injuries to the skin (197-199). It is believed that HP competes with the HS on the endothelial membrane for binding to L-selectin, thus decreasing neutrophil trafficking to the response site (186). The expression of heparanase, an endo-β-glucuronidase, has also been shown to be upregulated in a variety of inflammatory disorders such as inflammatory bowel disease, Crohn's disease, ulcerative colitis, and rheumatoid arthritis (200-202). Heparanase has the ability to degrade both HP and HS (203, 204). One would be made to believe that this would decrease the inflammatory response due to the reduction of HS on the cell surface. However, it has been postulated that upregulation of heparanase causes increased HS production and perhaps oversulfation of the HS polysaccharides (205). This scenario would result in increased cytokine presentation to neutrophils and increased

inflammatory response. In addition, heparanase could also degrade HP so that inflammatory proteases within mast cells are no longer sequestered, but become released more rapidly at the site of infection (206). Therefore, inhibition of heparanase, in addition to use of HP may both aid to treat patients with chronic inflammatory conditions.

Tumor Progression

Heparan sulfate has been shown to regulate various aspects of cancer such as tumor growth kinetics, invasiveness, and potential for metastasis (207, 208). HS can either aid or inhibit cancer growth based on the modification sequence on the polysaccharide chain, or whether it is located at the cell surface or unbound in the extracellular matrix (209). Cancer cells have the unique ability to alter both the sequence of the HS polysaccharide, as well as, the level of expression and distribution of core proteins responsible for harboring the HS chains (210). For instance, glypican-3 has been shown to be a negative regulator of breast cancer growth, but is mutated in Simpson-Golabi-Behmel syndrome (211). Researchers believe that the origin of this syndrome is a point mutation in the gene Xq26, which encodes for glypican-3 (211). Other glypicans have also been downregulated in many various tumor types, as well as maintaining an essential part in regulating cellular signals that are essential in growth and differentiation, such as insulin-like growth factor 2 and bone morphogenic protein (212). Also syndecan-1 downregulation has been identified as a key player in multiple myeloma and uterine carcinoma and is almost non-existant once the tumor becomes advanced (213). This data points to a common link for tumorigenesis through mutation resulting in a deficiency of expression of various core proteins and the lack of HS thereof.

As noted previously, HS can bind various FGF isoforms resulting in dimerization of the FGFR and stabilization of the growth factor/receptor complex in order for tyrosine kinase signaling to occur (175). The binding of various FGF isoforms, in addition to cytokines and other various structural proteins, depends on a specific binding motif located in the HS polysaccharide sequence. Cancer cells have displayed the unique ability to increase the density of these sequences within their HS structures, as well as to increase the overall abundance of these HS polysaccharides presented at the cell surface (214). This results in an overall increase in FGF signaling leading to the advancement of tumor cell proliferation. HS also plays a key role in the development of angiogenesis, as HSPG attached to the endothelial cell surface also modulate binding of FGF and vascular endothelium growth factor (215). While binding at the tumor cell surface and the endothelial cell surface is occurring, HS in the extracellular matrix attempts to inhibit the cancer growth by binding and sequestering growth factors.

HS also has the ability to modulate tumor metastasis through many of the same interactions that are involved in inflammation. Tumor cells often carry the lectin, P-selectin, at their cell surface. This lectin has the ability to bind to platelets at the surface of the endothelial cell and facilitate adhesion. The adhesion of tumor cells to endothelial cells also occurs through the ability of HS to activate integrins on the cell surface. Once adhesion takes place, the cancer cell is able to cross into the blood stream and metastasize to other parts of the body. Studies have shown the use of HP to be beneficial in inhibition of metastasis, by competing with the binding of P-selectin and by competition with the HS present on the cancer cell surface attempting to bind integrins on the epithelial cell (208).

Heparanase has the ability to cleave the HS chains from the core protein into oligosaccharide fragments ranging from 12-20 monosaccharide units (204, 216, 217). This action will reduce the potential of HS for sequestering growth factors at the cell membrane (218). The upregulation of heparanase has been seen in many various cancer types such as esophageal carcinomas, head and neck cancer, hepatocellular carcinoma, pancreatic tumors, as well as, other various human cancer cell lines (219). Pharmacological use of heparin mimetics and low molecular weight HP has shown benefits in the inhibition of heparanase activity resulting in anti-tumor effects (214).

Statement of Problem

Heparan sulfate (HS) is a highly sulfated polysaccharide that is ubiquitously expressed across all tissues. Heparin (HP) is a structural analog of HS that carries even more sulfation and is only produced within mast cells. The array of modifications, sulfation and epimerization, is a nontemplate driven process that is a direct result of substrate specificities and cooperation among the various enzymatic isoforms. The biosynthesis of HS/HP has always been presented as a step-wise process consisting of three main phases: 1) initiation of the polysaccharide chain attachment to the core protein; 2) elongation of the chain; and 3) modification with sulfation and epimerization. However, multiple experiments have shown that interactions between certain enzymes along the biosynthetic pathway are absolutely essential for production of biologically functional HS. These observations have begun to shift the biosynthetic paradigm from a static, stepwise process, to a more dynamic and cooperative cycle.

The divergent point of HS and HP structure is the presence of long stretches of a trisulfated disaccharide unit that accounts for up to 90% of HP. A noted difference in the isoforms responsible for the synthesis of each polysaccharide is the presence of NDST-2 in mast cells, while NDST-1 is absent. The main goal of this dissertation was to probe NDST-2 substrate specificities and modification profile using a structurally defined oligosaccharide library. We then hope to combine our understanding of NDST-2 modification patterns with existing knowledge of enzymatic capabilities of C_5-Epi and 2OST in an attempt to investigate the biosynthetic pathway of HP. We believe the biosynthesis of HP occurs in a cooperative cycle with elongation and modification occurring simultaneously.

Chapter II

Materials and Methods

Culturing Insect Cells

A 1ml cryovial containing >1.0 x 10^7 Sf9 cells in 80% Grace's Insect Medium, supplemented with 10% heat inactivated fetal bovine serum (HI-FBS), and 10% dimethyl sulfoxide (DMSO), was quickly thawed in a 37°C water bath. This cell suspension was then transferred to a 25-cm^2 T-flask containing 10ml pre-warmed Grace's Insect Media, supplemented with 10% FBS and 50 U/ml penicillin-streptomycin (complete media). The cells were then incubated at 27°C in a non-humidified, non CO_2, incubator until the monolayer reached approximately 90% confluency, or approximately 4.0 x 10^6 cells. At this point the cells were passaged using the sloughing method. The sloughing method entails the use of a sterile glass pipette to draw up approximately 5ml of culture media at a time within the T-flask and streaming the media in a side-to-side motion across the monolayer, resulting in the release of healthy cells into the media suspension. Once the cells were completely released into the suspension, cell viability was examined using the tryptan blue cell exclusion method. Between 250µL-1ml of cell suspension was mixed with an equal volume of 0.4% tryptan blue solution. The suspension dilution was then loaded onto the hemacytometer at a concentration of 10-50 cells/mm^2 and counted under

an inverted microscope with a 10X ocular and 10X objective (Figure 23). The cells that were stained blue were counted as dead, all others were viable. Cell viability was maintained at >95% for healthy log-phase cell growth. Once cell viability was determined, the cells were seeded between 2 x 10^4-5 x 10^4 cells/cm^2 into a 75-cm^2 T-flask containing 20 ml pre-warmed complete media. The cells were again passaged, using the sloughing method at 90% confluency, or approximately 1.2 x 10^7 cells, and analyzed for viability using the tryptan blue exclusion method. For optimal cell health, the cells were maintained in this state for approximately 3 months before initiation from a new frozen stock.

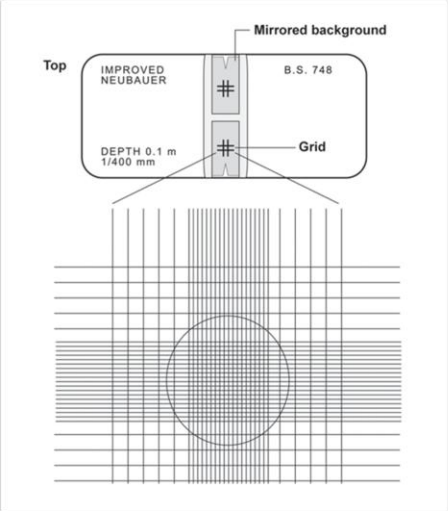

Figure 23. Counting cells using a hemacytometer

10µL of cell suspension is added to the hemacytometer. The hemacytometer is then placed under a 10X objective. Cells are counted in the large, central gridded square, 1mm^2 area. Cell count is the multiplied by 10^4 to gather cells/ml. This procedure is repeated and the average cell count is used

Once the cells reached optimal log-phase growth, maintaining >95% viability, frozen stocks were created. To create a frozen stock cells were sloughed and counted from 75cm^2 T-flasks to obtain >1 x 10^7 cells per vial to be frozen. The cell suspension was placed in 50 ml conical centrifuge tubes and centrifuged at 200 x g for 5 minutes to form a pellet. The spent media was removed from the tube and the pellet was resuspended in freezing medium, containing 80% Grace's Insect Medium, 10% HI-FBS, and 10% DMSO, to a concentration of >1 x 10^7 cells/ml. The freezing suspension was

then aliquoted into individual 1 ml cyrovials for freezing and storage. Once dispensed, the cryovials were stored in the -20°C freezer for 30 minutes, then transferred to the -80°C freezer for overnight storage. The following day, the frozen cells were transferred into the vapor phase of the liquid nitrogen tank for long-term storage.

Generating the Recombinant Baculovirus Expression Vector

Cloning into pFastBac-Mel-HT

The catalytic domain of human NDST-2, A66-G883, was determined previously (75). The catalytic domain of the NDST-2 cDNA sequence was amplified by polymerase chain reaction (PCR) using a pair of primers:
5'-ATATGCGAATTCAATGGCACGGCCTCCAGTTCCACCT and
5'-ATATGCGGCCGCATCAGCCCAGACTGGAATGCTGCAG (the cleavage sites of EcoRI and NotI are underlined). The PCR reaction was conducted by mixing the following components: 10X Pfu buffer, 5µL; 10mM dNTP mix (New England BioLabs), 1µL; 50-100ng NDST-2 cDNA; forward and reverse primers (25µM), 1.5µL each; Pfu Turbo DNA polymerase (Stratagene) 1µL; water to a final volume of 50µL. The PCR cycle utilized the following parameters: 1) Initial Denaturing: 95°C, 5min; 2) Denaturing: 95°C, 1min; 3) Annealing: 62°C, 1min; 4) Elongation: 72°C, 2.5min; Repeat cycle steps 2-4, 40X. Once the PCR reaction was finished, a 10µL aliquot was examined by agarose gel electrophoresis using a 1% agarose gel.

The baculovirus expression vector was constructed by inserting the honeybee melittin secretion signal sequence and six histidine residues at the N-terminal end of the pFastBac expression vector (Life Technologies, Inc.) as described by Liu, et al (122, 220). The resultant baculovirus expression vector was named pFastBac-Mel-HT. The

PCR-amplified NDST-2 catalytic domain product (1µg) and pFastBac-Mel-HT expression vector (3µg) were then subjected to a consecutive restriction enzymatic digest by EcoRI and NotI (New England BioLabs). The digested PCR products and vector were purified using the QIAquick Gel Extraction Kit (Qiagen) and 10µL of each sample were analyzed by agarose gel electrophoresis using a 1% agarose gel. The relative concentration of each sample was visualized in order to achieve an insert to vector ratio of 3:1 for ligation into the transfer vector. Ligation of the digested PCR product into the digested pFastBac-Mel-HT expression vector was achieved using the Rapid DNA Ligation Kit (Roche).

Transformation and Analysis of Recombinant Expression Vector

The ligation product was chemically transformed into DH5 alpha cells and grown overnight on Luria-agar (LB-agar) plates supplemented with 50µg/ml carbenicillin and 7µg/ml gentamycin at 37°C. The next day, a single colony was isolated and used to inoculate 3ml LB media containing the same concentration of antibiotics used previously and was shaken at 250rpm overnight at 37°C. The following day, the amplified expression vector was isolated using the QIAprep Spin Miniprep Kit (Qiagen). The purified expression vector containing NDST-2 was then sent for sequencing at the UNC-CH Genome Analysis Facility. The vector was sequenced using both forward and reverse primers for NDST-2 listed previously. The vector was also analyzed for correct positioning of NDST-2 using the sequencing primer:

5'-TATTCCGGATTATTCATACC.

Generating the Recombinant Bacmid

Transforming DH10Bac™E. coli

The NDST-2 recombinant baculovirus was prepared from the NDST-2 pFastBac-Mel-HT baculovirus expression vector using the Bac-to-Bac Baculovirus Expression System (Life Technologies). Generation of the recombinant baculovirus began with chemical transformation of the pFastBac-Mel-HT expression vector containing NDST-2 into DH10Bac cells (Life Technologies). DH10Bac cells contain the bacmid, or baculovirus DNA, with a mini-*att*Tn7 transposon target site and a helper plasmid. The pFastBac-Mel-HT expression vector contains a mini-Tn7 element that is able to transpose with the mini-*att*Tn7 site on the bacmid DNA in the presence of the transposition proteins provided by the helper plasmid, thus recombining the gene of interest into the baculovirus DNA. Once the DH10Bac cells were transformed with the expression vector containing NDST-2, they were grown overnight on LB-agar plates containing 50µg/ml kanamycin, 7µg/ml gentamycin, 10µg/ml tetracycline, 100µg/ml X-Gal and 40µg/ml isopropyl-β-D-thiogalactopyranoside.

Isolating Recombinant Bacmid DNA

Colonies containing recombinant bacmid DNA were identified by blue/white screening due to the disruption of expression of the *lacZα* peptide. This results in colonies harboring the recombinant bacmid DNA being white among a background of blue colonies containing the unmodified bacmid. Once a white colony was selected, it was used to inoculate 2ml of LB medium containing 50µg/ml kanamycin, 7µg/ml gentamycin, and 10µg/ml tetracycline, which was then grown overnight at 37°C with

constant shaking at 250rpm. After incubation overnight, the cells were pelleted by centrifugation at 9,000 x g for 15 minutes, and the recombinant bacmids were isolated using the QIAprep Spin Miniprep Kit (Qiagen).

Analyzing Recombinant Bacmid DNA by PCR

PCR was once again employed to analyze the recombinant bacmid isolate. This method was chosen over restriction enzyme digestion due to the large size of the recombinant bacmid DNA. This PCR method utilized specialized primers (pUC/M13) that hybridized to sites flanking the mini-*att*Tn7site within the *lacZα*-complementation region that determined the presence of the NDST-2 insert. The forward and reverse primers were: 5'-CCCAGTCACGACGTTGTAAAACG and 5'- AGCGGATAACAATTTCACACAGG. The PCR reaction was composed of 100ng recombinant bacmid DNA, 5µL 10X PCR buffer, 1µL 10mM dNTP mix, 1.5µL 50mM $MgCl_2$, 1.25µL both 10mM forward and reverse primers, 38.5µL water, and 0.5µL *Taq* polymerase (5 units/µL). The reaction was run at the following parameters: 1) Initial Denaturing: 93°C, 3min; 2) Denaturing: 94°C, 45sec; 3) Annealing: 55°C, 45sec; 4) Elongating: 72°C, 5min; Repeat steps 2-4 35X; 5) Final elongation: 72°C, 7min. After, completion of PCR, 10µL of product was used for 1% agarose gel analysis and the presence of NDST-2 was confirmed.

<u>Producing Recombinant Baculovirus</u>

Transfecting Sf9 insect cells

Sf9 cells growing in log phase (1.5-2.5 x 10^6 cells/ml) and >95% viability were used for transfection. Plating media was prepared using 1.5ml Supplemented Grace's Insect Medium containing 10% FBS (without antibiotics) and 8.5ml Grace's Insect

Medium, unsupplemented (without FBS or antibiotics). Cells were then seeded in a 6-well plate at 8 x 10^5 Sf9 cells per well. They were allowed to adhere to the plate for 15 minutes, whereby the medium was removed and replaced with 2.5ml plating medium. The transfection mixture was then prepared using CellFectin II (Invitrogen). 8µl of CellFectin II was diluted into 100µl Grace's Medium, unsupplemented. Next, 1µl of the recombinant bacmid generated in the previous section was diluted into 100µl Grace's Medium, unsupplemented and gently mixed. The two solutions were then combined to a total volume of 209µl and incubated for 30 minutes at room temperature. After incubation, the mixture was then added, dropwise, to each plate well and incubated at 27°C for 5 hours. After 5 hours, the transfection media was replaced with complete media and incubated at 27°C for 72 hours.

Isolating P1 Viral Stock

The budded virus was released into the medium 72 hours post-transfection. Obvious signs of infection were noted: cessation of growth, granular appearance, detachment from the monolayer, and cell lysis. At this point the medium was collected from each well and transferred to 15 ml conical tubes. These tubes were spun at 500 x g for 5 minutes to remove the cellular debris and clarify the baculoviral stock. The supernatant was then transferred to a fresh tube, covered with aluminum foil to protect from the light, and stored at 4°C for near term usage. An aliquot of P1 stock also placed in the -80°C freezer for long term storage.

Amplifying Baculoviral Stock

The initial P1 viral stock is a low titer stock, not suitable for optimal protein expression. The next step was to amplify this viral stock to generate a higher viral titer,

which could then be utilized for protein expression. The titer of a baculovirus stock is measured in plaque forming units (pfu), and was determined using the BaculoTiter Assay Kit (Invitrogen) according to the manufacturer's instructions. Amplification of our P1 stock allowed us to increase our viral concentration from 5×10^6 pfu/ml to 5×10^7 pfu/ml in the P2 stock, or by one order of magnitude. Typically, amplification was performed twice to reach a viral titer of approximately 5×10^8 pfu/ml in the P3 stock. Amplification beyond P3 was not ideal as deleterious mutations could arise, resulting in a reduction of target protein activity. Amplification was performed using healthy cells in log-phase growth, with a low passage number (5-20) and >95% viability. The cells were seeded in a 6-well plate at 2×10^6 cells/well. The cells were allowed to attach at room temperature for 1 hour. After 1 hour, the appropriate amount of P1 viral stock was added to each well. The amount of stock added is dependent upon the titer of the viral stock, number of cells to be infected, and the desired multiplicity of infection (MOI). The MOI is defined as the number of virus particles/cell (pfu/cell). The appropriate amount of P1 viral stock to use for amplification was calculated using the following formula: Inoculum required (ml) = ((MOI) (number of cells))/titer of viral stock. We found that the optimal MOI for our P1 amplification was 0.1 pfu/cell. Once the P1 stock was added to each well, the cells were incubated for 72 hours at 27°C. After 72 hours, the medium was collected and spun at 500 x g. The supernatant was then collected, protected from light and stored at 4°C for immediate use and -80°C for long term storage.

Expressing NDST-2

Initial NDST-2 expression was done in a 6-well plate format using the previously generated P3 viral stock. 6×10^5 cells were seeded in each well and allowed to attach for

30 minutes at room temperature. After attachment to the bottom of the well, the media was removed and replaced with 300µl fresh complete medium. Next, the appropriate amount of viral stock was added to each well. In the initial experiment, we used a range of MOIs to infect our cells: 1, 2, 5, 10, and 20. Based on our enzymatic activity assay described below, we determined our optimal MOI for this expression method to be 2. We also tested a range of infection times: 24hr, 48hr, 72hr, and 96hr. We determined our optimal infection time to be 72 hours. After infection, the enzyme was secreted into the media, the medium was then collected and centrifuged at 500 x g for 5 minutes to remove cellular debris. This clarified crude media was then used to carry out our initial sulfotransferase activity measurements. Large scale expression, analysis, and purification methods are discussed below.

Identification of NDST-2 Activity

NDST-2 Activity Assay

The sulfotransferase activity assay was used to determine the enzymatic activity of NDST-2 produced by the baculovirus expression system. During the initial expression stages, 10µl of crude media containing the secreted enzyme was assayed for activity. After full purification, approximately 14ng of NDST-2 was measured for full characterization of sulfo transfer capabilities. The experiment was carried out by incubating either the crude media or purified enzyme with 1µg heparosan K5 polysaccharide, 1µl of 1mM PAPS, and 1-5 x 10^5 cpm of [^{35}S]PAPS in 100µl of buffer containing 50mM MES, pH 7.0, 10mM $MnCl_2$, 5mM $MgCl_2$, and 1% Triton X-100. The reaction was incubated at 37°C for 30 minutes and quenched by the addition of UPAS

buffer containing 50mM NaOAc pH 5.0, 250 mM NaCl, 3M urea, 1mM EDTA, and 0.01% Triton X-100.

Polysaccharide Purification by DEAE-Sephacel and Radioisotope Quantification

The reaction mixture from the NDST-2 sulfotransferase activity assay was loaded onto a 200µl diethylaminoethyl-sephacel (Sigma-Aldrich) column for purification of the [^{35}S]-polysaccharide. The column was washed 4 times with 1ml UPAS buffer, 3 times with 1ml 250mM NaCl and 0.001% Triton X-100. The purified [^{35}S]-polysaccharide was eluted with 1M NaCl and 0.001% Triton X-100. The quantity of [^{35}S] present in the purified polysaccharide sample was determined by liquid scintillation counting.

Disaccharide Analysis of Heparan Sulfate Polysaccharides

Heparin Lyase Degradation

Heparosan K5 polysaccharides modified with NDST-2 were digested with heparin lyases to disaccharides for further analysis and confirmation of the presence of *N*-sulfation on the glucosamine monosaccharide unit. This was achieved by digesting the modified polysaccharides containing approximately 10,000 cpm of [^{35}S] with 10µl of each heparin lyase I, II, and III. The reaction was carried out in 200µl of buffer consisting of 50mM Na_2HPO_4, pH 7.0 and was incubated overnight at 37°C. The following day, the reaction was terminated by boiling at 100°C for 2 minutes and centrifuged at 13,000rpm in a benchtop centrifuge for 2 minutes. 100µl of the sample was then diluted in 130mM tetrabutlyammonium dihydrogenphosphate (TBA) and 100µl reverse phase ion pairing (RPIP) buffer A consisting of 38mM $NH_4H_2PO_4$, 2mM H_3PO_4, and 1mM TBA. The dilution was then injected onto an RPIP C_{18}-HPLC column (0.46 x 25cm) (Vydac) which was run at a flow rate of 0.5ml/min and UV was detected at

232nm. The column was equilibrated with 14% buffer B which consisted of 38mM $NH_4H_2PO_4$, 2mM H_3PO_4, 1mM TBA and 50% CH_3CN. The disaccharides were then eluted using a step-wise buffer B gradient. The gradient increased from 14% to 30% at 45 minutes, then to 39% at 60 minutes, and finally 100% at 85 minutes. The disaccharide composition of the NDST-2 modified sample was then determined by comparison to the elution times of disaccharide standards (Seikagaku Corporation 400576 Unsaturated Heparan/Heparin disaccharide mixture H mix). The mix of disaccharide standards was diluted into 130mM TBA and 180µl RPIP buffer A and analyzed using the same method and conditions described previously.

Large Scale Expression of NDST-2

Culturing Sf9 Insect Cells in Serum Free Media

Sf9 cells adapted for growth in Sf-900III serum free media (SFM) (Invitrogen) were obtained to be utilized in large scale baculovirus expression. These cells are originally derived from *Spodoptera frugiperda* cell line IPLB-Sf-21-AE and adapted in Sf-900 III SFM (SKU 12658) (2, 221). They are suitable for baculovirus expression of recombinant proteins on a large scale, grown in suspension cultures (222, 223). Sf-900 III is a complete 1X serum-free medium containing L-glutamine, Pluronic F68, and a reduced level of hydrolysate. Initially, the cells were rapidly thawed from frozen in a 37°C water bath before being transferred into a 125ml shaker flask containing 28.5ml of pre-warmed Sf-900 III SFM. The culture was then incubated at 27°C on an orbital shaker platform rotating at 140rpm. The cells were passaged when they reached >2 x 10^6 viable cells/ml. The cells were then seeded at a density of 3-5 x 10^5 cells/ml in a 125, 250, or 500ml shaker flask containing 30, 60, or 120ml of pre-warmed Sf-900 III SFM,

depending on the desired scale of protein expression. Cells were then incubated under the same conditions as above and passaged repeatedly once they reach the threshold mentioned previously. Cells were also available for cryopreservation as long as the viability was >90%. Flasks were chosen for freezing, and the cells were counted. They were then centrifuged at 100 x g for 10 minutes, or until a pellet had formed. Then they were resuspended in the appropriate amount of cryopreservation medium (containing 46% fresh Sf-900 III SFM, 46% conditioned Sf-900 III SFM, and 8% DMSO) to give a final cell density of >1 x 10^7 cells/ml. 1.5ml of suspension was then aliquoted into each cryovial, and freezing was achieved as mentioned in the previous cell culture section.

Expression of NDST-2 in Large Scale Shaker Flasks

Large scale cultures were infected at an MOI 2 once they reach >2 x 10^6 cells/ml. These cultures were incubated in the manner mentioned in the previous section for 72 hours. After 72 hours, the cultures were harvested, transferred to 500ml, conical centrifuge tubes and centrifuged at 500 x g for 5 minutes to get rid of cellular debris and clarify the supernatant. After collecting the supernatant, glycerol was added to a final concentration of 10%, and the crude media containing NDST-2 was stored at -80°C until purification.

Purification of NDST-2 Coupled to the FPLC System

Toyopearl Heparin AF HC-650M Chromatography

Initial purification of the crude supernatant from NDST-2 expression in Sf-900 III SFM utilized a Toyopearl Heparin AF HC-650M (Tosoh Bioscience, LLC #0020031) column coupled with the FPLC system. Harvested and clarified media was buffered with MOPS to a final concentration of 20mM and the pH was adjusted to 7.0.

Phenylmethanesulfonylfluoride (PMSF) was added to a final concentration of 1mM. The crude medium was then chilled on ice for 1 hour. After 1 hour, the solution was centrifuged at 9,000 rpm and filtered through a 0.45µm membrane to remove any cellular debris. The solution was then diluted 1:1 with water and loaded on a 10ml Toyopearl Heparin AF HC-650M at a flow rate of 5ml/min. The column was then connected to the FPLC system and washed for 10 minutes at 3ml/min with buffer A, consisting of 10mM MOPS, pH 7.0, 2% glycerol, 0.3% CHAPS, and 50mM NaCl. After washing the gradient was applied over 67 minutes for 0-75% buffer B, consisting of 10mM MOPS, pH 7.0, 2% glycerol, 0.3% CHAPS, and 1M NaCl. The column was then washed with 75% buffer B for 7 minutes and 100% buffer B for 13 minutes. The UV peak detected at 280nm was collected, analyzed by 1% SDS-PAGE gel analysis, NDST-2 activity assay, and pooled for further purification.

Nickel Sepharose 6 Fast Flow™Affinity Chromatography for His_6-Tagged NDST-2

Since our protein contains a 6xHis affinity tag applied during construction of the baculovirus expression vector, we were able to utilize the Nickel Sepharose 6 Fast Flow (GE Healthcare Bio-Sciences AB #17-5318-01) affinity chromatography column for our final purification step. The pooled fractions from the previous purification step were loaded onto the 2ml Nickel Sepharose Fast Flow 6 column at 1ml/min. The column was connected to the FPLC system and the method was initiated with a wash at 1ml/min of buffer A, consisting of 20mM TRIS, 500mM NaCl, 30mM Imidazole, at pH 7.6. Following the wash, the gradient was initiated at 1ml/min for 40 minutes of 0-100% buffer B, consisting of 20mM TRIS, 500mM NaCl, 300mM Imidazole, at pH 7.6. Finally, the column was washed with 100% buffer B for 5 minutes. The collected

fractions were eluted in 5% glycerol. Each fraction was then analyzed by NDST-2 activity assay and 1% SDS-PAGE gel analysis. The fractions containing NDST-2 activity, and appear as a single band on SDS-PAGE were pooled and stored at -80°C.

Chemoenzymatic Synthesis of Structurally Defined Oligosaccharides

KfiA and pmHS$_2$ Guided Oligosaccharide Backbone Elongation

Methods have recently been developed which enabled the synthesis of structurally defined oligosaccharides. These oligosaccharides are important tools for the characterization of HS and HP biosynthetic enzymes. The method takes advantage of bacterial glycosyltransferases that are responsible for the construction of heparosan, a bacterial capsular polysaccharide. *N*-acetyl glucosaminyl transferase of *Escherichia coli* K5, KfiA, is capable of transferring a GlcNAc residue, from UDP-GlcNAc, to the non-reducing end of GlcUA through an α-1,4 linkage (224). Heparosan synthase-2, pmHS2, from *Pastuerella multocida* is capable of transferring a GlcUA residue, from UDP-GlcUA, to the non-reducing end of glucosamine through a β-1,4 linkage (225). In addition, the oligosaccharides that are built maintain a p-nitrophenol (pnp) tag at the reducing end which allows for constant monitoring due to its UV absorbance at 310nm.

The synthesis began with 6mg pnp-GlcUA (Sigma-Aldrich). This starting material was mixed with 18μmol of UDP-GlcNAc (Sigma-Aldrich) and 2mg KfiA in a solution consisting of 25mM TRIS-HCL (pH 7.2) and10mM $MnCl_2$in a total volume of 40ml. The reaction was incubated overnight at room temperature. The following day, 20μl of the reaction mixture was analyzed for completion by polyamine-HPLC (PAMN), which is described below. Once the reaction was deemed to be complete, the disaccharide was purified by 100ml C_{18} column RPIP-HPLC. The reaction mixture was

centrifuged at 6,000 rpm for 10minutes and the pH was adjusted to 1.5. The mixture was then loaded onto the column at a flow rate of 1ml/min. The column was then washed for 10 minutes with buffer A, consisting of 0.1% trifluoroacetic acid (TFA) in water. A gradient was then applied at a flow rate of 2ml/min for 40 minutes from 0-100% buffer B, consisting of 0.1% TFA in acetonitrile. The peak corresponding to UV absorbance at 310nm was collected and pooled. The disaccharide was then analyzed by electrospray ionization-mass spectrometry (ESI-MS), which is discussed below. Following ESI-MS analysis, the fraction pool was dried using a centrivap and resuspended in 1ml water.

The disaccharide was then elongated to a trisaccharide by addition of a GlcUA monosaccharide unit to the growing oligosaccharide backbone. This was achieved by mixing the 1ml solution containing the disaccharide from the previous step with 27μmol UDP-GlcUA and 2mg pmHS2 in a solution consisting of 25mMTRIS-HCl, pH 7.2 and 10mM $MnCl_2$, to a total volume of 40ml. After overnight incubation at room temperature, the same procedure was followed for analysis and purification as described in the previous step.

For additional elongation reactions the conditions were essentially identical to the conditions above, with the reaction volumes adjusted accordingly depending on the amount of substrates.

GlcNTFA Deacetylation and Subsequent Sulfation with NST-1

A selective approach for building *N*-sulfated oligosaccharides was recently developed by the Liu group (226). This method utilized an unnatural monosaccharide donor residue, UDP-GlcNTFA. It was determined that this monosaccharide unit served as an acceptor molecule for KfiA, and could be readily incorporated into the growing

oligosaccharide backbone. Moreover, pmHS2 was also able to incorporate the next GlcUA residue with no problems. The GlcNTFA residue could then be selectively deprotected under mild basic conditions, yielding $GlcNH_2$, which, could then be sulfated by a truncated version of NDST-1 expressing only the *N*-sulfotransferase domain, otherwise known as NST.

UDP-GlcNTFA was synthesized by a chemoenzymatic approach, which consisted of preparing GlcNTFA 1-phosphate and coupling it to UDP. The synthesis began with 11mg of $GlcNH_2$ 1-phosphate (Sigma-Aldrich), which was dissolved in 200µl of anhydrous methanol and mixed with 60µl of $(C_2H_5)_3N$ and 130µl of *S*-ethyl trifluorothioacetate (Sigma-Aldrich). The reaction was incubated for 24 hours at room temperature resulting in GlcNTFA 1-phosphate. This was converted to UDP-GlcNTFA by using glucosamine-1-phosphate acetyltransferase/*N*-acetylglucosamine-1-phosphate uridyl transferase (GlmU) in a buffer containing 46mM TRIS-HCl pH7.0, 5mM $MgCl_2$, 200µM dithiothreitol, 2.5mM UTP, and 0.012 units/µl inorganic phosphatase (Sigma-Aldrich). The proteins were removed from the reaction mixture by using centrifugal filters (10,000 MWCO; Millipore) and dialyzed against water using 1,000 MWCO membranes for 4 hours. The product was confirmed by ESI-MS and concentration was determined by PAMN-HPLC by comparison with appropriate standards.

Selective de-*N*-trifluoroacetylation was achieved by resuspending the purified oligosaccharides in a 200µl solution consisting of a 2:2:1 ratio of CH_3OH, water, and $(C_2H_5)_3N$. The reactions were incubated overnight at room temperature. The following day, the oligosaccharides were dried and reconstituted in 1ml of water for subsequent elongation.

A typical *N*-sulfation reaction was accomplished by incubating 1mg de-*N*-trifluoracetylated oligosaccharide with 400µM PAPS and 2.5mg NST in a buffered solution consisting of 50mM MES pH 7.0, 1% and Triton X-100 to a final volume of 20ml. The reaction was incubated overnight at room temperature. The reaction was purified by 1ml Q-sepharose column and the fractions corresponding with the absorbance peak at 310nm were collected, pooled, and dialyzed against water using 1,000 MWCO dialysis membranes. After dialysis, the samples were dried overnight. The following day they were resuspended in 1ml of water and the structures were confirmed by ESI-MS

C_5-Epi and 2OST Modification

Epimerase and 2OST reactions were run simultaneously in a one pot reaction typically using 1mg oligosaccharide substrate. The oligosaccharide was incubated with 2.5mg C_5-Epi, 2.5mg 2OST, and 400µM PAPS in a buffered solution consisting of 50mM MES, pH 7.0, and 1mM $CaCl_2$ in a total volume of 10ml. The reaction was incubated at 37°C overnight. The reaction was purified by Q sepharose-HPLC as described in a later section and confirmation of 2-*O*-sulfation was achieved by ESI-MS.

NDST-2 Modification of Structurally Defined Oligosaccharides

Modification reactions using structurally defined oligosaccharide substrates typically utilized 300µg of oligosaccharide for each reaction. The substrate was incubated with 6.5µg of purified NDST-2 and 400µM PAPS in a buffered solution containing 50mM MES, pH 7.0, 1% Triton X-100, and 10mM $MnCl_2$ to a final volume of 20ml. These reactions were incubated for 72 hours at 37°C. At this point, they were purified by Q sepharose-HPLC, using the same conditions described below. The UV peaks corresponding with absorbance at 310nm were pooled separately and dialyzed

against water using 1,000 MWCO dialysis membranes. Following dialysis, these samples were further desalted by BioGel P-2 size-exclusion column (0.75 x 200cm) (BioRad), which was equilibrated with 0.1M ammonium bicarbonate and run at a flow rate of 4ml/hour. The UV absorbance at 310nm was identified among the fractions collected, and they were used directly for ESI-MS analysis.

Analysis of NDST-2 Modified Structurally Defined Oligosaccharides

Q Sepharose Fast Flow™-HPLC Purification and Analysis

The 1ml Q sepharose Fast Flow (GE Healthcare Bio-Sciences) was equilibrated with buffer A (20mM NaOAc, pH 5.0). The column was connected to the HPLC and the method was initiated for a 60 minute gradient of 0-100% buffer B (20mM NaOAc, pH 5.0, 1M NaCl) with the UV detector set at 310nm.

PAMN-HPLC Analysis

PAMN-HPLC analysis was performed using a polyamine II column (250 x 4.6mm) (YMC Group) which was equilibrated with water. The 10µg oligosaccharide substrate sample/reaction product was diluted with 200µl water and injected into the column. The column was then run with a gradient of 0-100% buffer B (1M KH_2PO_4) at a flow rate of 0.5ml/min for 40 minutes with the UV detector set at 310nm. After 40 minutes, the column was washed for 10 minutes with 100% buffer B.

DEAE-HPLC Analysis

DEAE-HPLC Analysis was performed using a TSKGelDNA-NPR column (Tosoh Bioscience, LLC) which was equilibrated with buffer A, consisting of 20mM NaOAc, pH 5.0. The 10µg oligosaccharide sample/reaction product was diluted with 200µl buffer A and injected into the column. The column was then run with a gradient of 0-100% buffer

B (20mM NaOAc, pH 5.0, 1M NaCl) at a flow rate of 0.4ml/min for 60 minutes with the UV detector set at 310nm. After 60 minutes, the column was washed for 20 minutes with 100% buffer B. For radiolabeled samples, the outlet line was connected to a liquid scintillation counter and radioactive intensity was recorded.

Mass Spectrometry Analysis of NDST-2 Modified Oligosaccharides

Mass spectrometry analyses were performed on a Thermo LCQ-Deca. The oligosaccharides were dissolved in 50% methanol. A syringe pump (Harvard Apparatus) was used to introduce the sample via direct infusion (35µl/min). Experiments were performed in negative ionization mode with a spray voltage of 5KV and a capillary temperature of 275°C. The automatic gain control was set to 1×10^7. Data were acquired and processed using Xcalibur 1.3 software.

Tandem Mass Spectrometry Analysis of NDST-2 Modified Oligosaccharides

Tandem mass spectrometry analysis was achieved by selecting a precursor ion using an isolation width of 3Da and the activation energy was 40% normalized collision energy. The automatic gain control was set to 2×10^7. The product ions in MS/MS data were labeled according to Domon/Costello nomenclature (227). All other methods, hardware, and software were consistent with the previous section.

Identifying the Cooperative Roles of NDST-2, C_5-Epi, and 2OST using Structurally Defined Ideal Substrates in a One Pot Reaction

Structurally defined ideal oligosaccharide substrates were utilized in order to investigate the cooperative role of NDST-2, C_5-Epi, and 2OST, when mixed together in a one-pot reaction. A typical reaction was performed by incubating 10µg oligosaccharide substrate with 163ng of purified NDST-2, 8µg of purified 2OST, 8µg of purified C_5-Epi,

and 800µM PAPS in a buffered solution consisting of 50mM MES, pH 7.0, 1mM $CaCl_2$, 1% Triton X-100, and 10mM $MnCl_2$, to a final volume of 200µl. The reactions were incubated overnight at 37°C and analyzed by PAMN-HPLC the following day.

These reactions were also carried out using 300µg oligosaccharide substrate, 4.9µg NDST-2, 240µg 2OST, and 240µg C_5-Epi in a total volume of 6ml. All other reaction conditions were the same as listed for the small scale reaction. The larger scale reactions were purified by Q sepharose-HPLC, and the peaks corresponding to UV absorbance at 310nm were collected and pooled. These peaks were then dialyzed against water using 1,000 MWCO dialysis membranes for 6 hours. They were then dried overnight and further desalted the following day by size-exclusion chromatography on a Bio-Gel P2 column as described previously. The peaks were located from the P2 fractions by UV absorbance at 310nm and were directly used to characterize the structure by ESI-MS.

Low pH Nitrous Acid Degradation Disaccharide Analysis to Identify the Structure of ΔUA Residue Product from One Pot Reaction

For identification of the iduronic acid residue resulting from the one-pot reaction of NDST-2, 2OST, and C_5-Epi, a low-pH nitrous acid degradation and disaccharide analysis method was used. The reaction was carried out in the same manner as the small scale reaction (10µg) described above, with the addition of 50nmol [^{35}S] PAPS (4×10^6 cpm) to the reaction mixture. After overnight incubation at 37°C, the reaction mixture was purified by size exclusion chromatography on a BioGel P10 column (BioRad), and radiolabeled oligosaccharides were identified by liquid scintillation counting and pooled.

Nitrous degradation was performed by incubating the ^{35}S-labeled oligosaccharides with 20µl water and 40µl 1M HNO_2, which was prepared from a fresh mixture (v/v 1:1) of 0.5M H_2SO_4 and 0.5M $Ba(NO_2)_2$. The reaction was then incubated on ice for 30 minutes, followed by a quenching with a mixture of 20µl 1M Na_2CO_3 and 1M $NaHCO_3$ (v/v 1:1). 20µl 0.5M $NaBH_4$ was then added and incubated at 50°C for 30 minutes. Finally 20µl 10M acetic acid was added to stop reduction. The disaccharide mixture was then purified by BioGel P2 column using the method described previously. The mixture was then analyzed using the C_{18} RPIP-HPLC procedure described previously. The ^{35}S peak eluted at approximately 45 minutes corresponds to the eluted position of the disaccharide IdoUA2S-AnMan, thus demonstrating that the one-pot reaction did indeed provide a 2-*O*-sulfated IdoUA residue at the expected position.

Chapter III

Determination of the substrate specificity of *N*-deacetylase/*N*-sulfotransferase isoform 2

Heparan sulfate *N*-deacetylation/*N*-sulfation is the most critical modification within the HS biosynthetic pathway for yielding biologically active HSPGs. Nearly every downstream modification is carried out at the *N*-sulfated glucosamine monosaccharide unit, or at the uronic acid that is adjacent to the *N*-sulfated glucosamine (10). This long understood fact of HS biosynthetic enzyme substrate recognition gives rise to domains within the polysaccharide structure that carry a large density of negative charge, in addition to greater conformational flexibility. These two characteristics of the NS domain form the basis of HS interactions and regulations within a multitude of physiological processes (18).

Formation of the domain structures in HS is initiated by the NDST-1 enzyme. Sheng et al. demonstrated that NDST-1 will form a non-sulfated region, known as the NAc domain, when it encounters a pre-existing GlcNS residue. The GlcNS residue forces NDST-1 to leave a gap of at least 4 unmodified monosaccharide units before it will engage another glucosamine unit for modification (26). This data has provided the long sought evidence for the presence of unmodified regions within HS structure, also known as NAc domains.

The mRNA transcripts for both NDST-1 and NDST-2 are expressed abundantly and ubiquitously throughout all cell types that make HS (73, 79). However, it was demonstrated *in vivo* that a lack of NDST-1 led to respiratory distress and neonatal death in mice. They were shown to be deficient in HS production and their embryonic growth was retarded (82-86). These results are evidence that even though mRNA transcripts for NDST-2 are found in all cell types, the enzyme does not have the capability to compensate for the loss of NDST-1 during HS biosynthesis. There is evidence pointing to transcriptional, translational, and post-translational control of NDST-2 expression, which may help to explain why the enzyme does not compensate for NDST-1 *in vivo* (81, 135, 137). When NDST-2 expression was silenced in mice, the only abnormal phenotype observed was a deficiency and abnormal formation of connective tissue mast cells, including a lack of protease and histamine storage capability (88, 134). Interestingly, HP resides within these specialized cells and is responsible for the sequestration of histamines and proteases inside the mast cells, which are released upon arrival at the site of an injury. Under normal conditions NDST-2 expression is quite abundant in mast cells, while NDST-1 mRNA is barely detected (79). Thus, HP formation inside the mast cells must be directed by NDST-2, while HS formation in all other cells is under the direction of NDST-1.

These *in vivo* studies have given rise to the need for researchers to gain a better understanding of the substrate recognition mechanisms and modification capabilities of the family of NDST isoforms. In this chapter, we present a study of the substrate recognition and modification capability of NDST-2. This study was performed using structurally defined oligosaccharide substrates that were designed using a recently

reported chemoenzymatic synthesis approach. Our data was compared and contrasted with the recently reported study on the substrate modification pattern of NDST-1, in hopes of uncovering key mechanistic differences between the two closely related isozymes. Interestingly, we found that NDST-2 acts in a very different manner from its counterpart NDST-1. Unlike NDST-1, NDST-2 does not move in one uniform direction. Instead it prefers to modify the substrate towards either the reducing or non-reducing end of the oligosaccharide substrate in a seemingly random manner. In addition, NDST-2 activity is not inhibited, but rather enhanced by the presence of a pre-existing *N*-sulfation within the oligosaccharide substrate. This data is consistent with observations that were originally uncovered in microsomal fractions over two decades ago. Finally, we were able to demonstrate that a pentasaccharide is the smallest oligosaccharide to be modified by NDST-2. These results will help to explain why there are no NAc domains found in HP structure, as it is essentially one long NS domain consisting of nearly 90% of the trisulfated disaccharide unit IdoUA2S-GlcNS6S (21).

Expression of NDST-2, Column Selection and Purity

The *N*-deacetylase domain of NDST-2 was first reported by Duncan et al. and was comprised of A66-P604 at the *N*-terminus of the protein. Based on sequence alignment with NDST-1, the *N*-sulfotransferase domain was known to exist towards the *C*-terminus and the total size of NDST-2 was known to be comprised of 883 amino acid residues. Thus, NDST-2 expression was achieved by cloning the catalytic domain (A66-G883) from NDST-2 cDNA into the expression vector pFastBac-Mel-HT for expression using the Bac-to-Bac expression vector system (Invitrogen) (Figure 24). This expression vector contains an *N*-terminal honeybee mellitin secretion signal, followed by an *N*-terminal

6xHis purification tag. Next, the expression vector was transformed into DH10Bac cells, which contain a parent bacmid, or baculovirus DNA, as well as a helper plasmid to supply the proteins needed for recombination. After recombination took place between the expression vector and the parent bacmid, the recombinant bacmid containing the NDST-2 gene was isolated and used to transfect Sf9 insect cells. Approximately 72 hours post-transfection obvious signs of infection were observed, indicating that the budded virus has been released into the media. The media was then collected and amplified through several rounds of infection with Sf9 cells before it was suitable for protein expression. Once the virus was amplified to a suitable viral titer, it was used to infect Sf9 cells adapted for growth in serum free medium. These cells were grown in 125ml, 250ml, and 500ml shaker flasks at 27°C, rotating on an orbital shaker platform at 250rpm and were infected during log-phase growth. The infection time course most suitable for optimal enzymatic expression was determined to be 72 hours and the optimal viral concentration was MOI 2. After incubation, the media containing the secreted NDST-2 enzyme was collected and clarified by centrifugation for further purification.

Several attempts at purification of NDST2 were made but proved unsuccessful. Our initial attempts focused on direct purification of NDST-2 using a Ni-sepharose column based on the *N*-terminal histidine$_6$ tag that was introduced from the expression vector. However, it became quite clear that the many media components were interfering with the binding of the His$_6$ tag to the Ni-sepharose resin. Attempts were made to concentrate and dilute the crude medium, but all attempts proved to be unsuccessful.

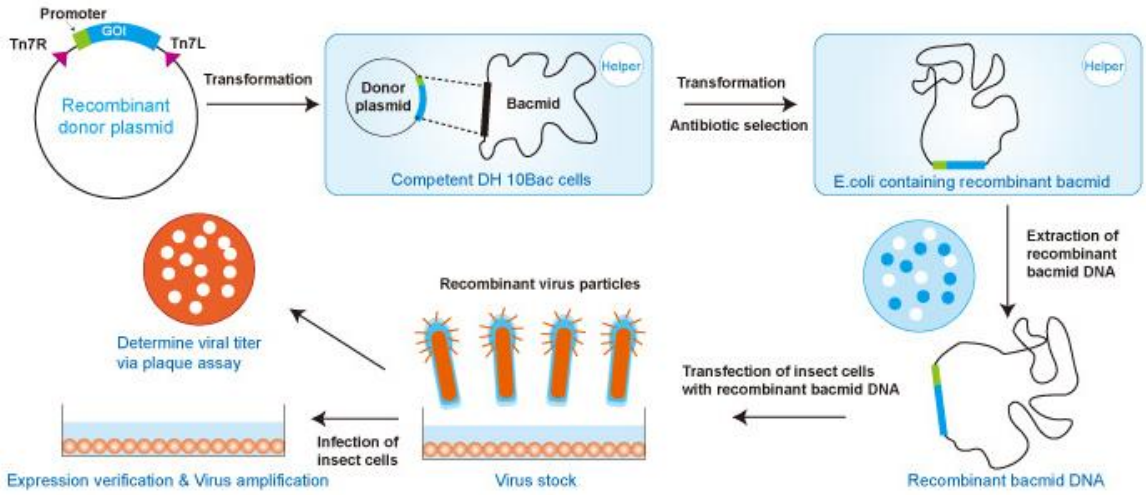

Figure 24. Baculovirus expression system. The process begins when the gene of interest is cloned into the donor plasmid. The donor plasmid is then transformed into competent DH10 Bac cell, where recombination occurs, with the aid of helper proteins, between the donor plasmid and the parent bacmid DNA contained within the cells. The bacmid containing the gene of interest is selected by blue/white screening, isolated, and used to transfect Sf9 insect cells. After infection, budded viral particles are released into the media. The media is harvested and used for reinfection of Sf9 cells, resulting in protein expression. Figure taken from GenScript.

Upon literature review, we discovered that the Lindahl group had purified NDST-2 in 1991 from mouse mastocytoma cells using Blue Sepharose (228). Blue sepharose consists of the dye cibacron blue F3-GA covalently linked to sepharose beads and has long been exploited for affinity chromatography applications (GE Healthcare Life Sciences, Hi-Trap Blue HP). The NDST-2 from our crude media also would not bind this column, as all activity was found in the flow through fraction. We then decided to try a host of other affinity purification columns. Our plan was to try two anion exchange affinity based columns, DEAE and Q-sepharose (GE Healthcare) at pH9 and pH8 respectively. These attempts also proved to be failures with all of the activity retained in the flow through fraction. Finally, we attempted a cation exchange column, Heparin-Sepharose 6 Fast Flow (GE Healthcare). This column consists of highly-negative charged heparin ligands covalently linked to a cross-linked sepharose matrix. In this

attempt we were finally able to see some ability to maintain enzymatic activity in the eluant fraction. However, the best recovery yield we were able to achieve using this column was approximately 10%. With the understanding that cation exchange affinity chromatography had given us the most promising results to this point, we again consulted the literature. After an extensive search, we discovered a group that had recently published a method of NDST-2 purification in which they utilized a Toyopearl AF Heparin 650M affinity column (229). After examining the data sheets for both Toyopearl AF Heparin 650M and Heparin Sepharose 6 Fast Flow, it became clear that the Toyopearl column was superior in stability, flow rate, and particle size, allowing for a greater density of heparin ligands/ml bed volume when compared with Heparin Sepharose 6 Fast Flow. Thus, on our initial attempt with we loaded 1.75L of crude media containing NDST-2 onto a 10mL Toyopearl AF Heparin 650M column coupled with the FPLC system and were able to achieve approximately 90% recovery yield in the eluant fractions (Figure 25) (Table 1). However upon further analysis by 1% SDS-PAGE and Bradford Assay, we realized that even though this column did clean up our sample significantly while retaining sufficient yield, there was still a large concentration of proteins other than NDST-2 present in the eluant fractions maintaining NDST-2 activity (Figure 27A:Lane 2). The presence of other various heparin-binding proteins greatly decreased the overall purity of our NDST-2 sample. At this point, we decided to focus our attention back to the Ni-sepharose column and the *N*-terminal His_6 tag on our NDST-2 protein. We loaded the eluant fractions obtained from the Toyopearl AF Heparin 650M purification step onto a 2ml Ni sepharose column coupled with the FPLC system (Figure 26). NDST-2 activity was assayed in each of the eluant fractions and was it was

observed that nearly all activity was present in the fractions that were eluted after the major 280nm UV peak observed between 2-8 minutes elution time. This initial observation gave us hope that our sample had reached optimal purity and was confirmed by 1% SDS-PAGE analysis. Our final NDST-2 purification analysis revealed that our protein was >90% pure with a final recovery yield of nearly 20% (Figure 27A:Lane 3) (Table 1). We speculate that the relatively poor yield provided by the Ni sepharose purification step is likely a result of peptidase cleavage of the melittin signal peptide that could interfere with the His_6 tag, thus disrupting binding to the column (230). Overall, we were able to express NDST-2, purify the enzyme to exceptional purity, and maintain a high level of enzymatic activity.

The *N*-sulfo heparosan recovered from the NDST-2 activity assay was subjected to heparin lyase degradation to determine the identity of the product formed from NDST-2 treatment. A single peak was observed with an elution time of 13 minutes, corresponding to the elution time of the disaccharide standard, ΔUA-GlcNS (Figure 27B). This data confirms that the purified NDST-2 enzyme is sulfating at the *N*-position of heparosan and possesses the expected enzymatic activity.

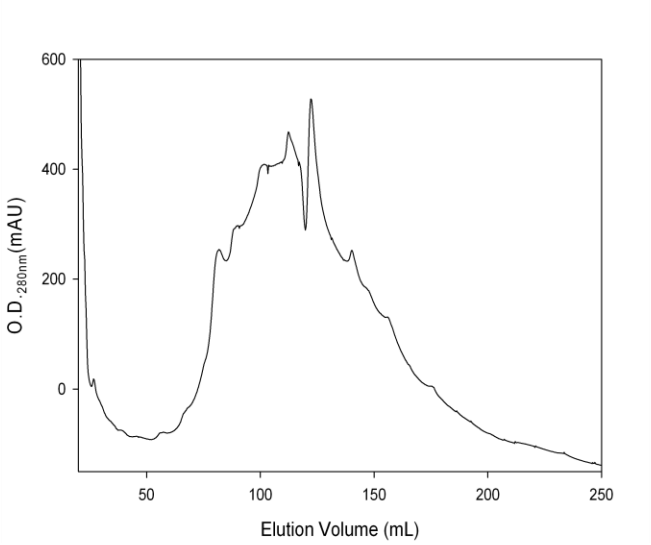

Figure 25. FPLC chromatogram of the purification of NDST-2 by AF Heparin 650M affinity chromatography. Approximately 1.75L of crude media containing secreted NDST-2 was loaded onto a 10ml Toyopearl AF Heparin 650M chromatography column. The UV absorbance at 280nm was plotted against elution volume. The fractions corresponding with elution volumes 60ml-180ml were shown to maintain enzymatic activity and were collected and pooled for further purification.

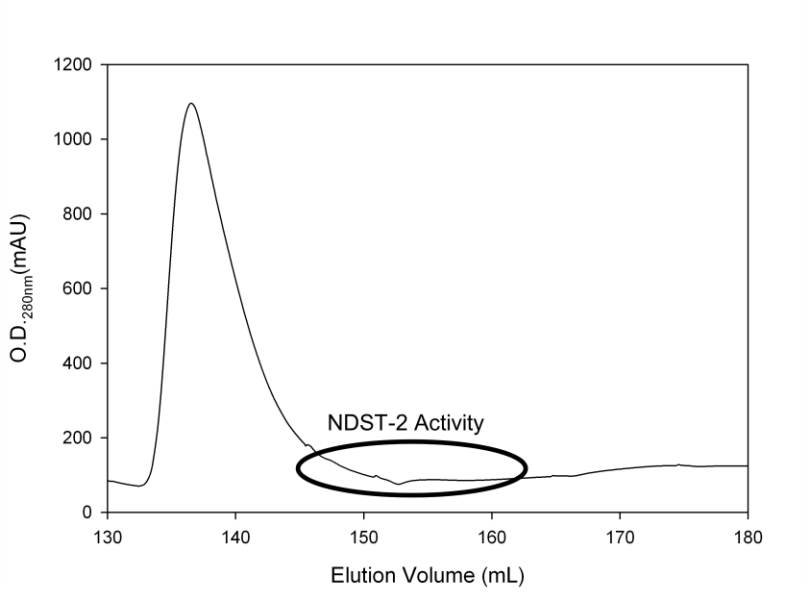

Figure 26. FPLC chromatogram of the purification of NDST-2 by Nickel Sepharose 6 Fast Flow chromatography. 120ml of AF Heparin 650M eluant containing secreted NDST-2 was loaded onto a 2ml nickel sepharose chromatography column. The UV absorbance at 280nm was plotted against elution volume. The fractions corresponding with elution volumes 146ml-162ml were shown to maintain enzymatic activity and were collected and pooled for purification analysis.

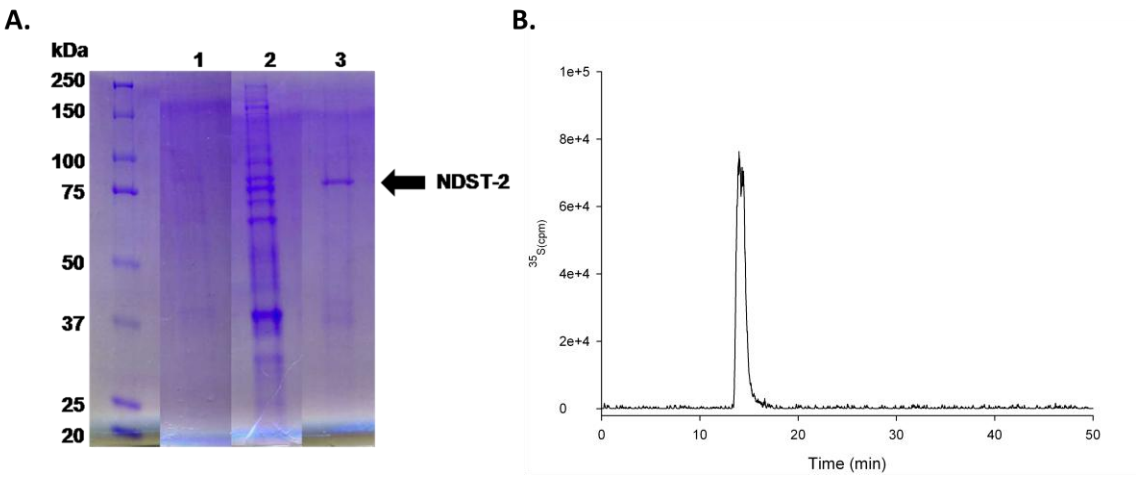

Figure 27. Analysis of purified NDST-2. (A) SDS-PAGE analysis. A small aliquot (10µl) of crude media containing NDST-2 (lane 1), NDST-2 purified by AF Heparin 650M (lane 2) and (270ng) NDST-2 purified by nickel sepharose (lane 3) were analyzed by SDS-PAGE using a 10% TRIS-HCl gel stained with Coomassie Blue. The purity of NDST-2 is greater than 90% following Ni sepharose purification. (B) Heparin lyase degradation. K5 polysaccharide modified by NDST-2 was subjected to enzymatic degradation by heparin lyases I, II, and III. The eluted peak corresponds to the position of the disaccharide standard ΔUA-GlcNS, thus confirming the presence of *N*-sulfation resulting from modification with NDST-2.

	Protein (mg)	Activity (U)	Specific Activity (U/ng)	Recovery %	Purification Fold
Crude Media	276.5	$3.1x10^9$	0.2	-	-
AF Heparin 650M	70.3	$2.8x10^9$	0.8	88.8	3.6
Nickel Sepharose 6 Fast Flow	0.5	$5.2x10^8$	21.0	18.6	26.9

Table 1. NDST-2 purification analysis. An initial volume of 1.75L yielded approximately 500µg active NDST-2 after Ni sepharose purification. 1 unit = 1pmol SO_4 transferred to polysaccharide substrate.

Synthesis of Structurally Defined Oligosaccharide Library

In order to characterize the mode of action of NDST-2, we designed a library of structurally defined oligosaccharides (Table 2). The synthesis of the oligosaccharide library began with the starting material *p*-nitrophenyl (pNP) β-glucuronide (Figure 28). The pnp-group allows for UV detection at 310nm, facilitating the monitoring of our oligosaccharides throughout synthesis and subsequent reactions. Elongation was performed by the alternating use of two bacterial glycosyltrasnferase enzymes, *N*-acetylglucosaminyltransferse from *Escherichia coli* K5 strain (KfiA) and heparosan synthase 2 (pmHS2), which was derived from *P. multocida*. Initially, KfiA was used to transfer a GlcNAc residue from the donor UDP-GlcNAc to the pnp-GlcUA starting material, creating a pnp-labeled disaccharide. Once this reaction was complete, pmHS2 was utilized to transfer a GlcUA monosaccharide unit from UDP-GlcUA to the disaccharide, resulting in a pnp-labeled trisaccharide. These steps were alternated to achieve the desired size of oligosaccharide substrate (231).

Two of the oligosaccharide substrates in the library carried an *N*-sulfation at the 6^{th} monosaccharide unit. This was achieved by using an unnatural form of glucosamine containing a trifluoroacetic acid group (GlcNTFA) that can be readily removed upon treatment with triethylamine, methanol, and water, creating a free amine at the *N*-position (Figure 28). Upon treatment with *N*-sulfotransferase and PAPS, this position can be readily sulfated, thus giving us the ability to selectively add *N*-sulfation at desired positions within the oligosaccharide.

The library consisted of 7 oligosaccharides that were essentially backbone structures and contained only alternating GlcUA and GlcNAc monosaccharide units to

the desired oligosaccharide length (Figures 29-35). It is worth noting that as a result of this library design there were 4 structures that maintained a GlcNAc residue on the nonreducing end of the structure (Tetra-1, Hexa-1, Octa-1, and Deca-1), while the other 3 structures maintained a GlcUA residue at the nonreducing end (Penta-1, Hepta-1, and Nona-1). In addition to the 7 unmodified backbone structures, there were two structures that maintained a GlcNS residue at the 6th monosaccharide position (Nona-2 and Deca-2) (Figures 36-37). Each of the substrates in the oligosaccharide substrate library were purified to greater than 90% purity and the structures were confirmed by ESI-MS before use in the NDST-2 reactions (Table 2).

	Structure of substrate	Purity by HPLC (%)	Calculated MW (Da)	Measured MW (Da)
Tetra-1	GlcNAc-GlcUA-GlcNAc-GlcUA-PNP	96%	897.7	897.6 ± 0.3
Penta-1	$(GlcUA\text{-}GlcNAc)_2$-GlcUA-PNP	94%	1073.9	1073.6 ± 0.1
Hexa-1	GlcNAc-$(GlcUA\text{-}GlcNAc)_2$-GlcUA-PNP	98%	1277.1	1277.3 ± 0.5
Hepta-1	$(GlcUA\text{-}GlcNAc)_3$-GlcUA-PNP	90%	1453.2	1453.7 ± 0.5
Octa-1	GlcNAc-$(GlcUA\text{-}GlcNAc)_3$-GlcUA-PNP	98%	1656.4	1656.3 ± 0.3
Nona-1	$(GlcUA\text{-}GlcNAc)_4$-GlcUA-PNP	99%	1832.5	1832.3 ± 0.5
Deca-1	GlcNAc-$(GlcUA\text{-}GlcNAc)_4$-GlcUA-PNP	97%	2035.7	2036.8 ± 0.1
Nona-2	GlcUA-GlcNAc-GlcUA-GlcNS-$(GlcUA\text{-}GlcNAc)_2$-GlcUA-PNP	97%	1870.5	1870.2 ± 0.2
Deca-2	GlcNAc-GlcUA-GlcNAc-GlcUA-GlcNS-$(GlcUA\text{-}GlcNAc)_2$-GlcUA-PNP	90%	2073.7	2073.5 ± 0.2

Table 2. Library of structurally defined oligosaccharide substrates.

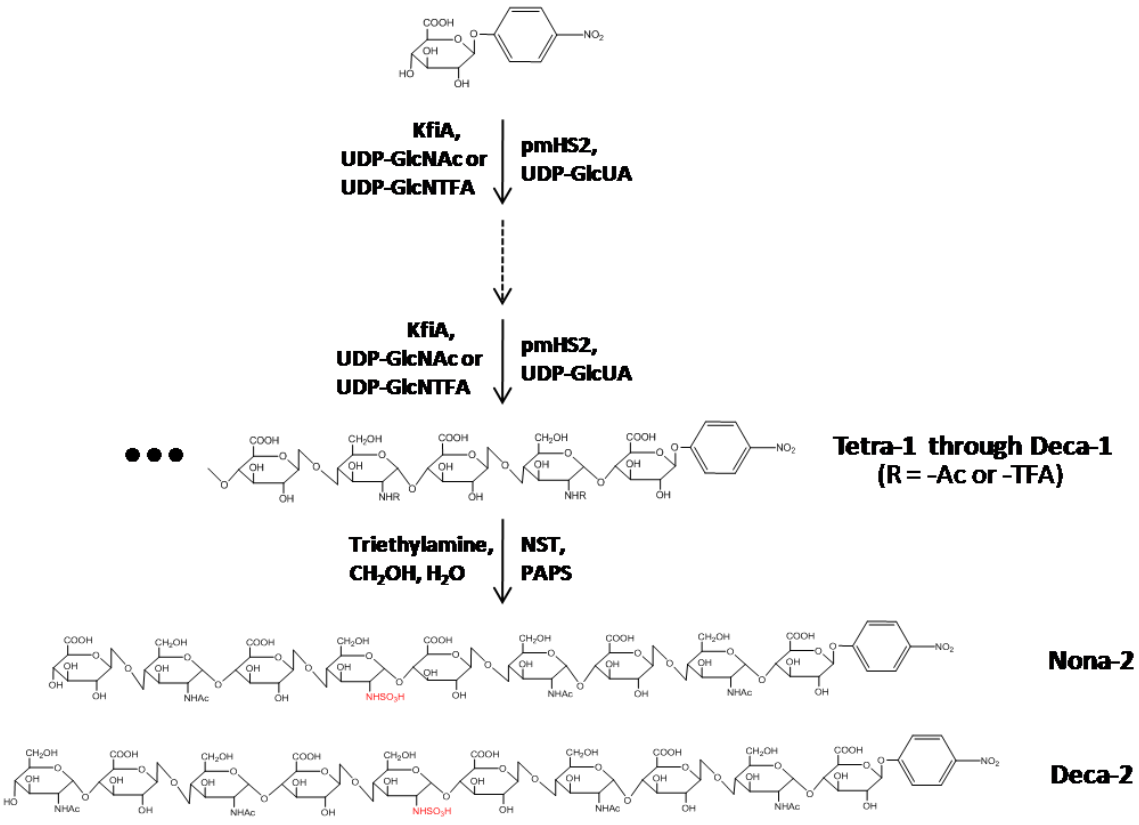

Figure 28. Chemoenzymatic synthesis of structurally defined oligosaccharide library. The synthesis begins by initiating the addition of a GlcNAc or GlcNTFA monosaccharide unit by the glycosyltransferase enzyme KfiA from their corresponding UDP-sugar donors to the pnp-GlcUA starting material. Next, a GlcUA monosaccharide unit is transferred from the UDP-sugar donor to the growing oligosaccharide chain by the glycosyltransferase enzyme pmHS2. This process is repeated to achieve the substrate of desired oligosaccharide length. The unnatural GlcNTFA monosaccharide unit is added to the chain at locations where *N*-sulfation is desired. Once the elongation of the chain reaches the desired length, it is treated with triethylamine, CH_2OH, and H_2O to achieve deacetotrifluoronation. Once the TFA is removed, the sulfotransferase enzyme NST is used to transfer a sulfo group from the sulfo donor PAPS to the selected *N*-positions.

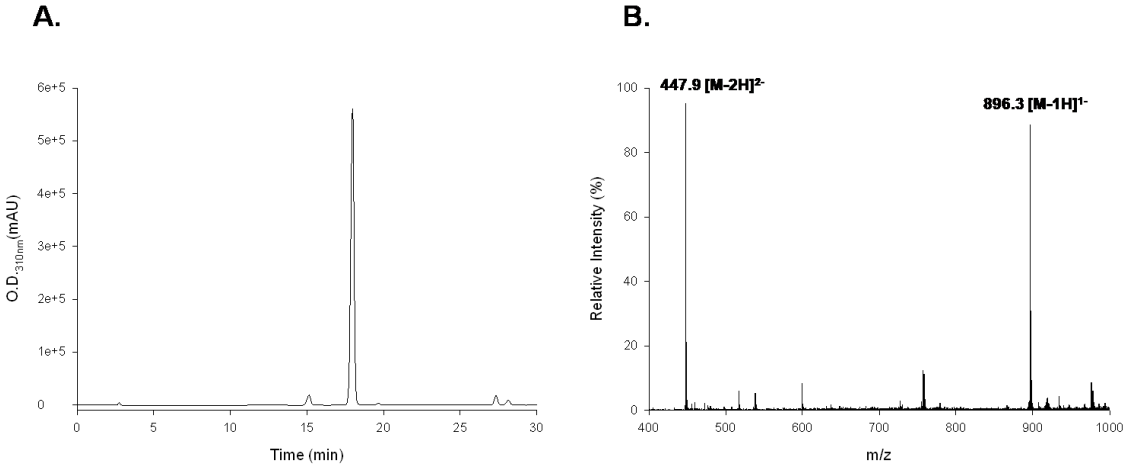

Figure 29. Purity and structural analysis of Tetra-1. (A) PAMN-HPLC chromatogram of Tetra-1 reveals 96% purity. **(B)** ESI-MS spectrum of Tetra-1 confirms structure.

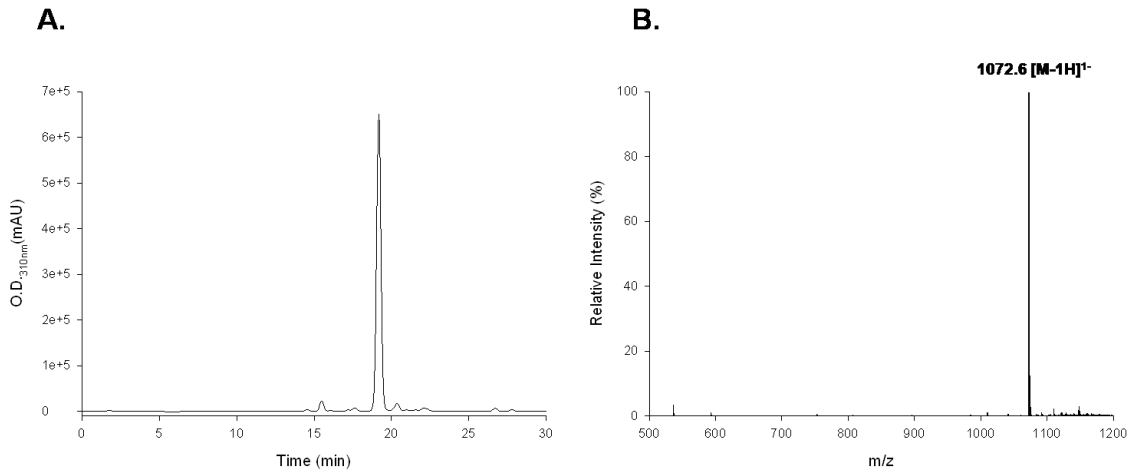

Figure 30. Purity and structural analysis of Penta-1. (A) PAMN-HPLC chromatogram of Penta-1 reveals 94% purity. (B) ESI-MS spectrum of Penta-1 confirms structure.

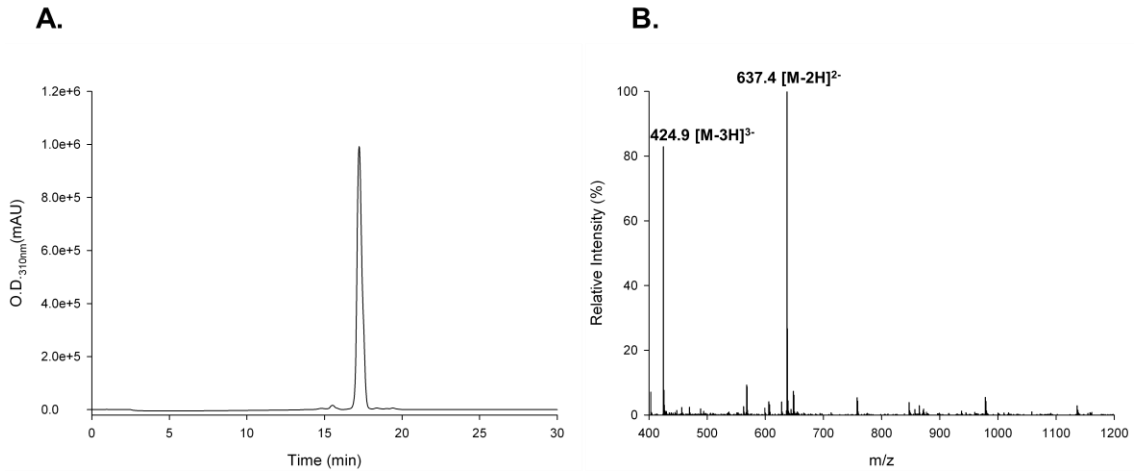

Figure 31. Purity and structural analysis of Hexa-1.(A) PAMN-HPLC chromatogram of Hexa-1 reveals 98% purity. (B) ESI-MS spectrum of Hexa-1 confirms structure.

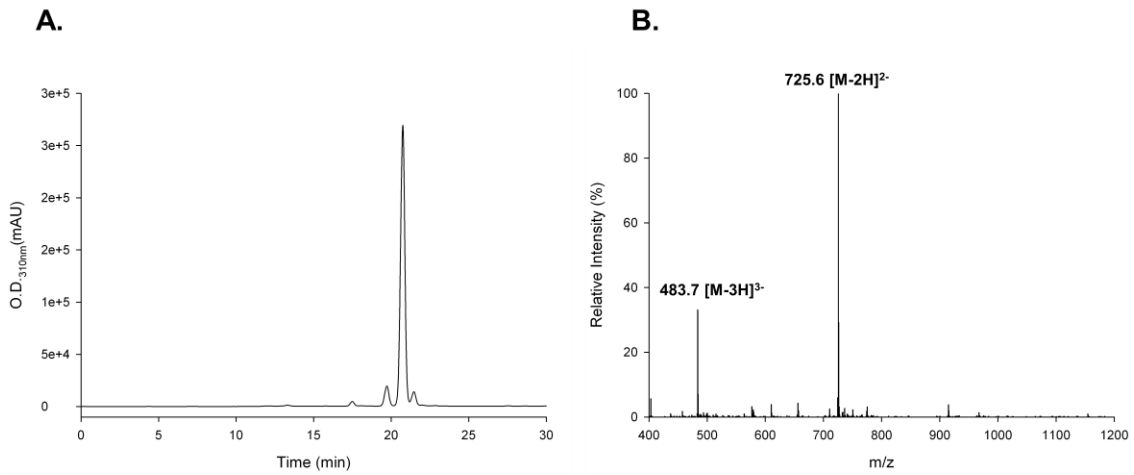

Figure 32. Purity and structural analysis of Hepta-1. (A) PAMN-HPLC chromatogram of Hepta-1 reveals 90% purity. (B) ESI-MS spectrum of Hepta-1 confirms structure.

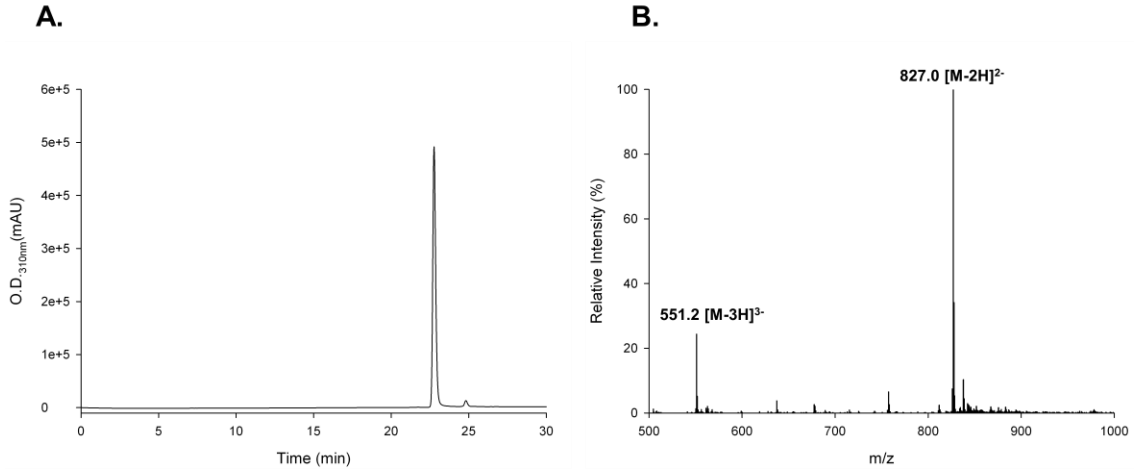

Figure 33. Purity and structural analysis of Octa-1. (A) PAMN-HPLC chromatogram of Octa-1 reveals 98% purity. (B) ESI-MS spectrum of Octa-1 confirms structure.

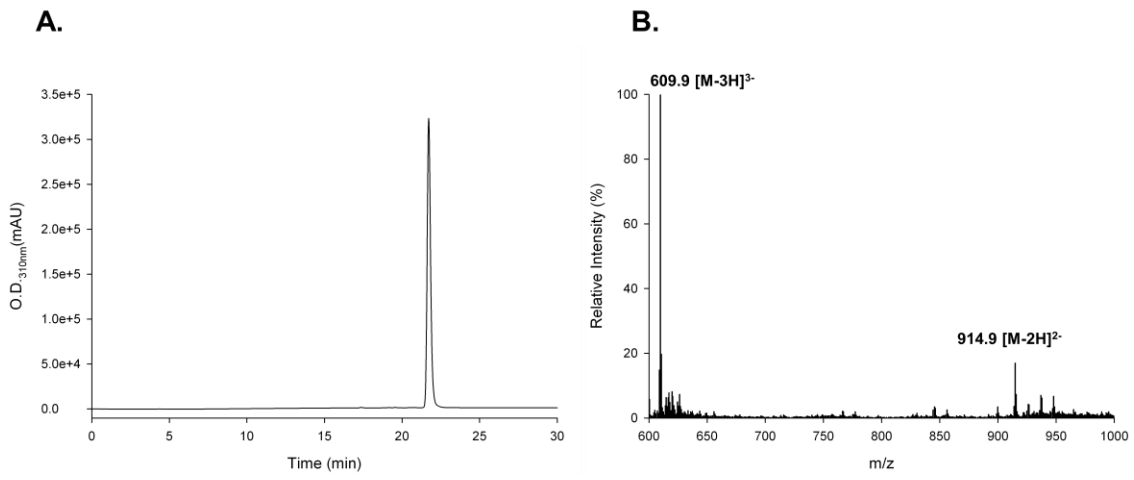

Figure 34. Purity and structural analysis of Nona-1. (A) PAMN-HPLC chromatogram of Nona-1 reveals 99% purity. (B) ESI-MS spectrum of Nona-1 confirms structure.

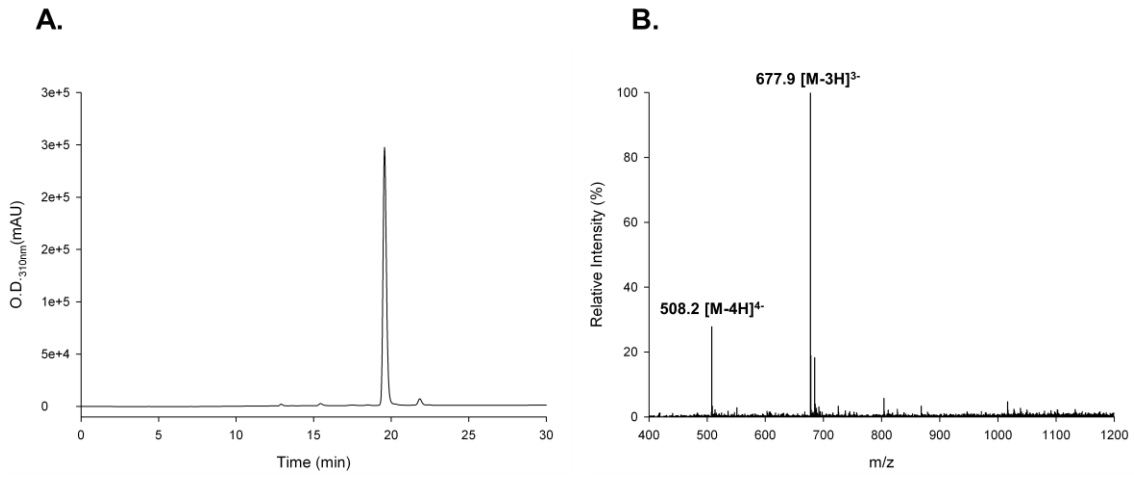

Figure 35. Purity and structural analysis of Deca-1. (A) PAMN-HPLC chromatogram of Deca-1 reveals 97% purity. (B) ESI-MS spectrum of Deca-1 confirms structure.

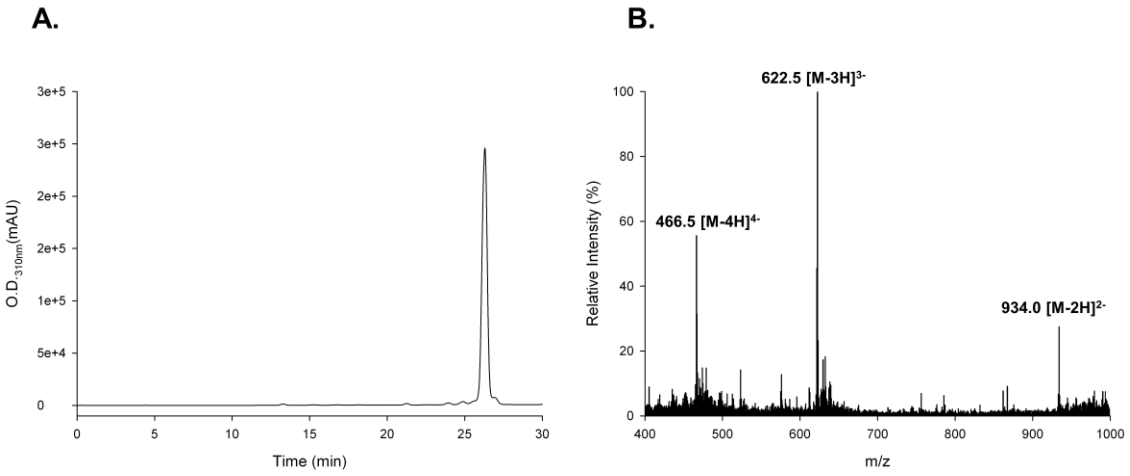

Figure 36. Purity and structural analysis of Nona-2. (A) PAMN-HPLC chromatogram of Nona-2 reveals 97% purity. (B) ESI-MS spectrum of Nona-2 confirms structure.

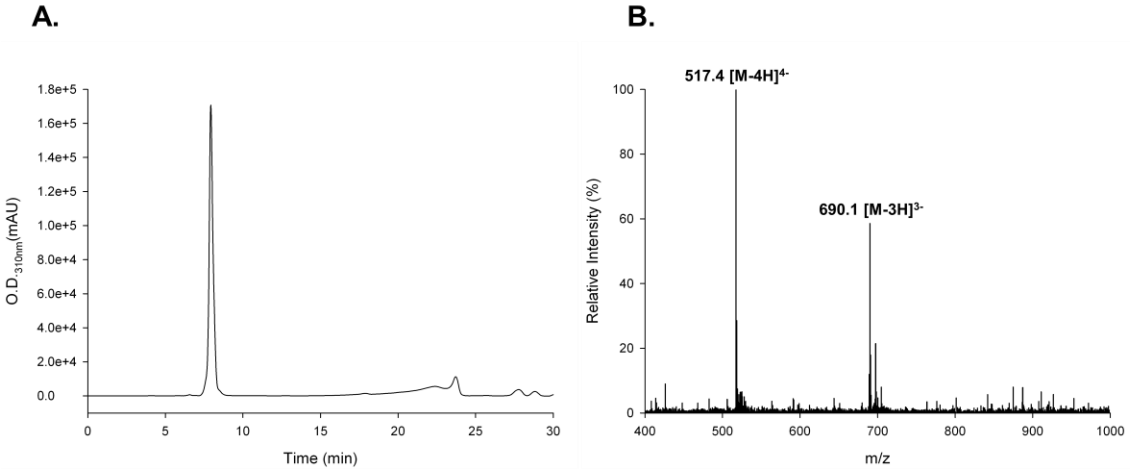

Figure 37. Purity and structural analysis of Deca-2. (A) PAMN-HPLC chromatogram of Deca-2 reveals 90% purity. (B) ESI-MS spectrum of Deca-2 confirms structure.

Determination of Minimal Size Substrate to be Modified by NDST-2

The initial characterization of the NDST-2 modification was to determine the minimal size oligosaccharide substrate that was capable of being modified by NDST-2. The availability of the oligosaccharide library that was presented in the previous section enabled us to complete this study. We began the study by incubating **Tetra-1** with NDST-2 and followed by purification on Q sepharose-HPLC, and subsequent analysis by PAMN-HPLC and ESI-MS. After purification on Q sepharose-HPLC we observed one peak corresponding with an elution time of 12 minutes (Figure 38A). This analysis indicated that there was either no reaction or one single product from the reaction. We isolated the peak and analyzed the pooled fractions by PAMN-HPLC (Figure 38B). Again we observed a single peak with an elution time at 19 minutes. Finally, the sample was analyzed by ESI-MS to reveal a molecular mass of 896.2 Da, corresponding the molecular weight of the starting material (Figure 38C). We determined that no reaction occurred between NDST-2 and **Tetra-1** indicating that this oligosaccharide is not large enough to be modified by NDST-1.

Next, we incubated NDST-2 and **Penta-1** and followed with the same methods of purification and analysis as before. The reaction was purified by Q sepharose-HPLC, whereby 2 peaks were observed eluting at 16 and 26 minutes respectively (Figure 39A). This indicated that a product was formed during this reaction. The peaks were collected separately and pooled (Pool A and Pool B),followed by PAMN-HPLC analysis (Figure 39B). Pool A was eluted at approximately 18 minutes, while Pool B was eluted at 26 minutes. Thus, it was determined the product yield was approximately 13%. Finally, both pools were analyzed by ESI-MS and it was determined that Pool A was indeed

starting material with a corresponding molecular weight of 1072.3 Da (Figure 39C). Analysis of Pool B corresponded with deacetylation and sulfation at a single GlcNAc position on **Penta-1** displaying an m/z value of 554.9 which corresponds very closely to a doubly charged molecule with the expected molecular weight of 1111.2 Da. These results have led us to the understanding that a pentasaccharide is the smallest oligosaccharide that is capable of being modified by NDST-2.

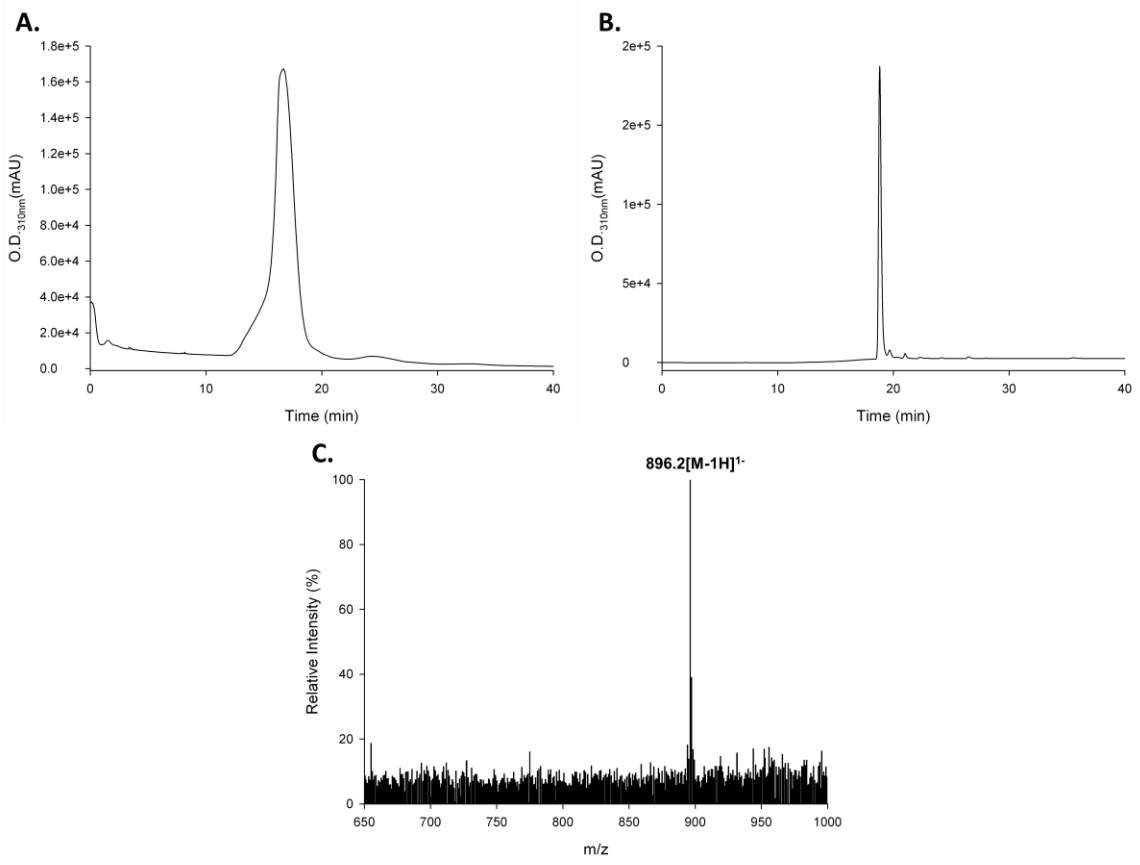

Figure 38. Tetra-1 + NDST-2 reaction analysis(A) Q sepharose-HPLC purification reveals a single peak eluted at 12 minutes. (B) PAMN-HPLC confirms the presence of a single peak eluted at 19 minutes. (C) ESI-MS confirmation of starting material remaining with a molecular mass of 896.2 Da.

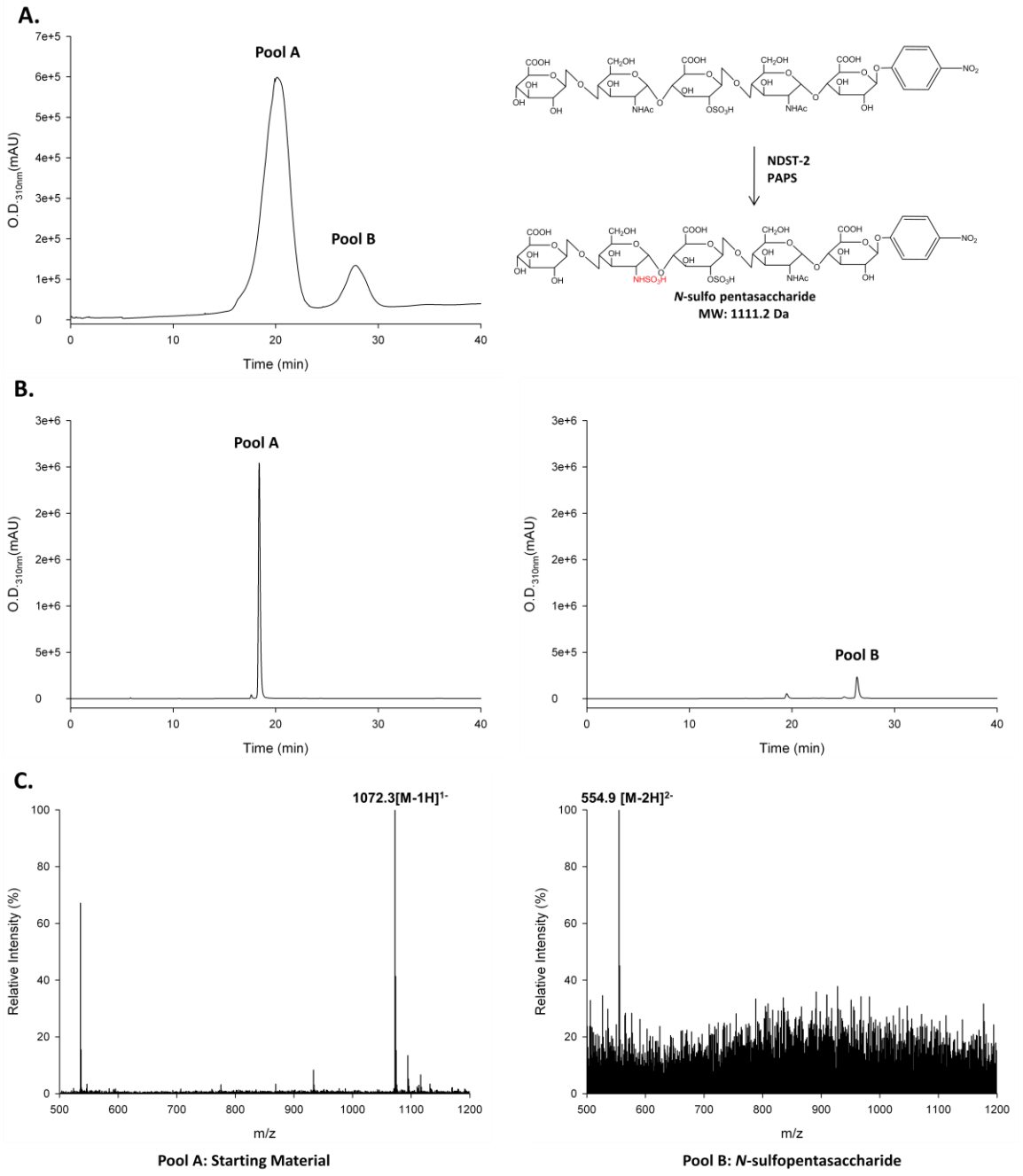

Figure 39. Penta-1 + NDST-2 reaction analysis. (A) Q sepharose-HPLC purification reveals two peaks eluted at 16 and 26 minutes respectively. (B) PAMN-HPLC confirms the presence of a single peak from Pool A eluted at 18 minutes and a single peak from Pool B eluted at 26 minutes. (C) ESI-MS confirmation of starting material remaining in Pool A with a molecular weight of 1072.3 Da and confirmation of a product in Pool B containing one GlcNS with an m/z value of 554.9, corresponding to a doubly charged ion with the expected molecular weight of 1111.2 Da.

Visualizing the NDST-2 Mode of Action

Using a combination of HPLC and mass spectrometric techniques we were able to determine the exact structures of the oligosaccharides modified by NDST-2. This has aided us in understanding the preferred substrate binding site, the patterns of modification, and the overall behavior of the NDST-2 enzyme when presented with various structurally defined oligosaccharides. As described in the previous section, the enzymatic reactions were purified and analyzed initially by Q sepharose-HPLC. After purification, each of the fractions under the peaks corresponding to UV absorbance at 310nm were collected and pooled separately. These samples were then analyzed by PAMN-HPLC to determine purity and overall recovery yield for each product from the reaction. The samples were then analyzed by ESI-MS to determine the overall molecular weight of each product. Following ESI-MS, the samples were analyzed by a tandem mass spectrometry technique (MS/MS) that allows for the identification of the precise monosaccharide units in the product that have been modified. MS/MS analysis makes use of a condition that allows the oligosaccharides to be fragmented at the glycosidic linkages. We were then able to piece together the fragments based on their m/z values and corresponding molecular weights to determine the overall structure of the product.

In the previous section it was noted that **Penta-1** was the smallest oligosaccharide to form a product (***N*-sulfo pentasaccharide**) after incubation with NDST-2. The *N*-sulfo pentasaccharide product possesses a molecular weight of 1111.2 Da, which is confirmed by the *m/z* value of 554.9 corresponding to a doubly charged ion (Figure 39C). This ion was then selected as the precursor ion for MS/MS analysis (Figure 40). Our initial MS/MS analysis revealed two major fragments with *m/z* values 485.3 and 971.1

(Figure 40A). Both of these values correspond with fragment B_5^{2-} and B_5^{1-} respectively, or with the loss of the pnp tag. The loss of pnp was shown to be a common occurrence across each of the oligosaccharide product MS/MS analyses, perhaps due to the relative weakness of this bond. We then decided to exclude these most common fragments by refocusing the scanning range from m/z 250-500. Following this scan we observed two characteristic fragmented ions, C_2 and Y_4^{2-} at the *m/z* values of 434.3 and 346.7 (Figure 40B). This provided structural evidence that the GlcNS residue is at the monosaccharide position 4 within the ***N*-sulfo pentasaccharide** product.

These data show that the initial site of modification must be at least 4 monosaccharide units from the reducing end of the oligosaccharide as there is no evidence of any ***N*-sulfo pentasaccharide** products carrying two GlcNS residues.

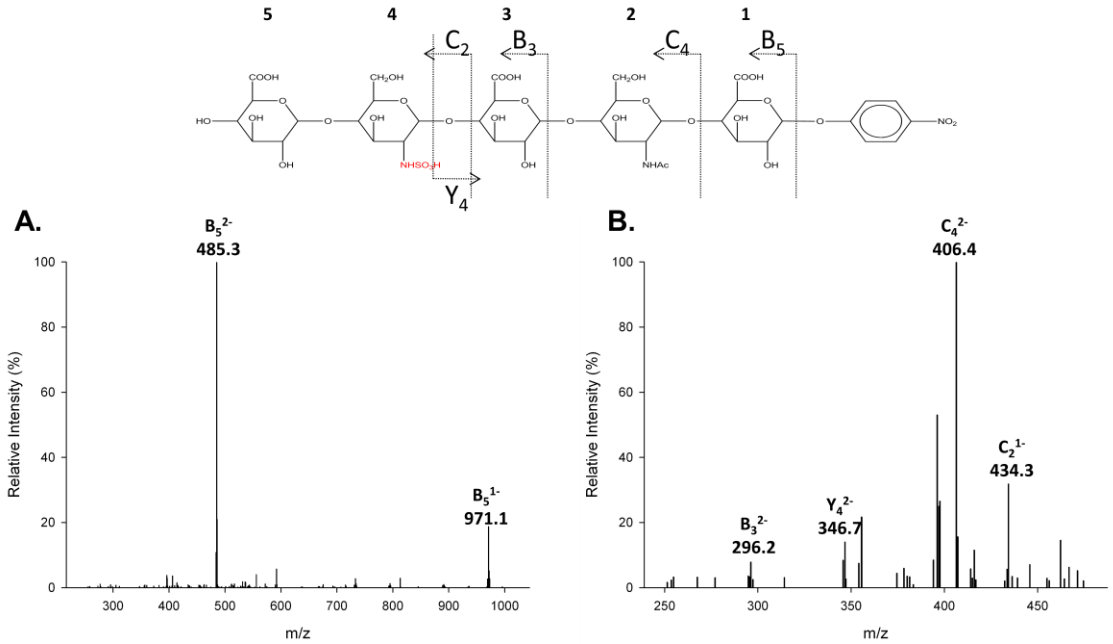

Figure 40. ***N*-sulfo pentasaccharide MS/MS analysis.** (A) Precursor ion selection was at $[M-2H]^{2-}$, *m/z* 554.9. MS/MS analysis across *m/z* value range 250-1000 shows the loss of pnp tag corresponding with the B_5 fragment. (B) MS/MS analysis across *m/z* value range 250-475 reveals *N*-sulfation at the glucosamine residue in position 4 (Monosaccharide units are numbered from reducing end toward nonreducing end).

After incubation of **Hexa-1** with NDST-2 we observed two peaks by purification on Q sepharose-HPLC, eluted at 18 minutes and 24 minutes respectively (Figure 41A). This data revealed that **Hexa-1** is a substrate for NDST-2. The fractions under each peak were collected and pooled individually (Pool A and PoolB). PAMN-HPLC analysis revealed single peaks for PoolA and PoolB, with elution times of 17 minutes and 23 minutes respectively (Figure 41B). Finally, each sample was analyzed by ESI-MS. PoolA displayed an *m/z* value of 637.4, representing a doubly charged ion that corresponds very closely to the molecular weight of the starting material, 1276.4 Da. PoolB displayed an *m/z* value of 656.3, representing a doubly charged ion that corresponds very closely to the molecular weight of ***N*-sulfo hexasaccharide** carrying one GlcNS residue, 1314.3 Da. PAMN-HPLC analysis shows that the yield of this product after NDST-2 modification is 71%. The sample was then analyzed by MS/MS to determine the location of the GlcNS residue (Figure 42). Two characteristic fragmented ions Y_5 and C_3, at the *m/z* values of 936.2 and 637.1, provided the structural evidence that the GlcNS residue is at position 4 within the ***N*-sulfo hexasaccharide** product.

These data reveal that the addition of a GlcNAc residue at the nonreducing end of the pentasaccharide increases the ability of the NDST-2 reaction to form a product by nearly 6 fold. In addition, the modification site is the same location between **Penta-1** and **Hexa-1**, thus the GlcNAc at the nonreducing end of the oligosaccharide remains unmodified.

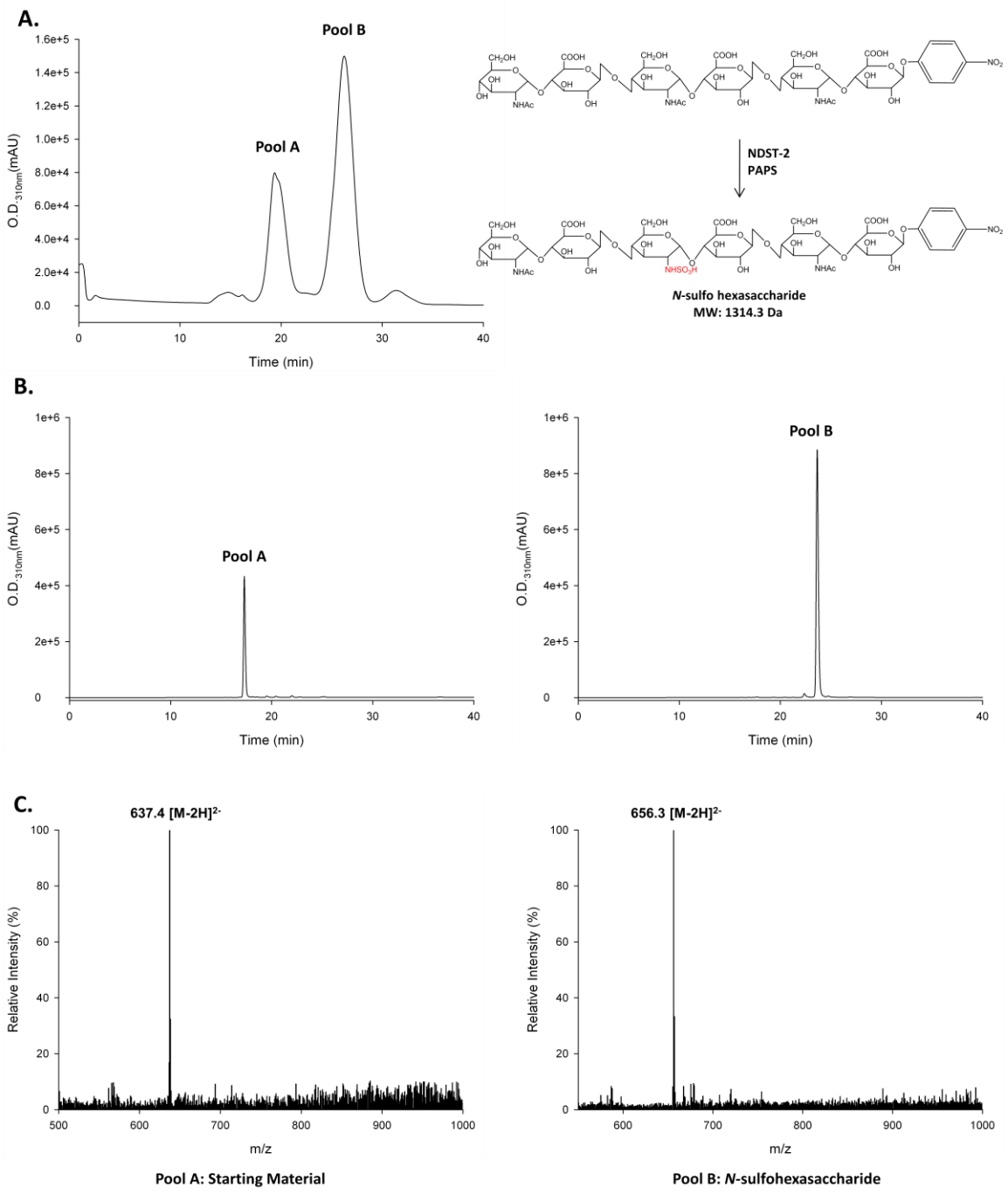

Figure 41. Hexa-1 + NDST-2 reaction analysis. (A) Q sepharose-HPLC purification reveals two peaks eluted at 18 and 24 minutes respectively. (B) PAMN-HPLC confirms the presence of a single peak from Pool A eluted at 17 minutes and a single peak from Pool B eluted at 23.5 minutes. (C) ESI-MS confirmation of starting material remaining in Pool A with m/z value of 637.4 corresponding with a doubly charged ion very close to the expected molecular weight of 1276.4 Da and confirmation of a product in Pool B containing one GlcNS with an m/z value of 656.3, corresponding to a doubly charged ion with the expected molecular weight of 1314.3 Da.

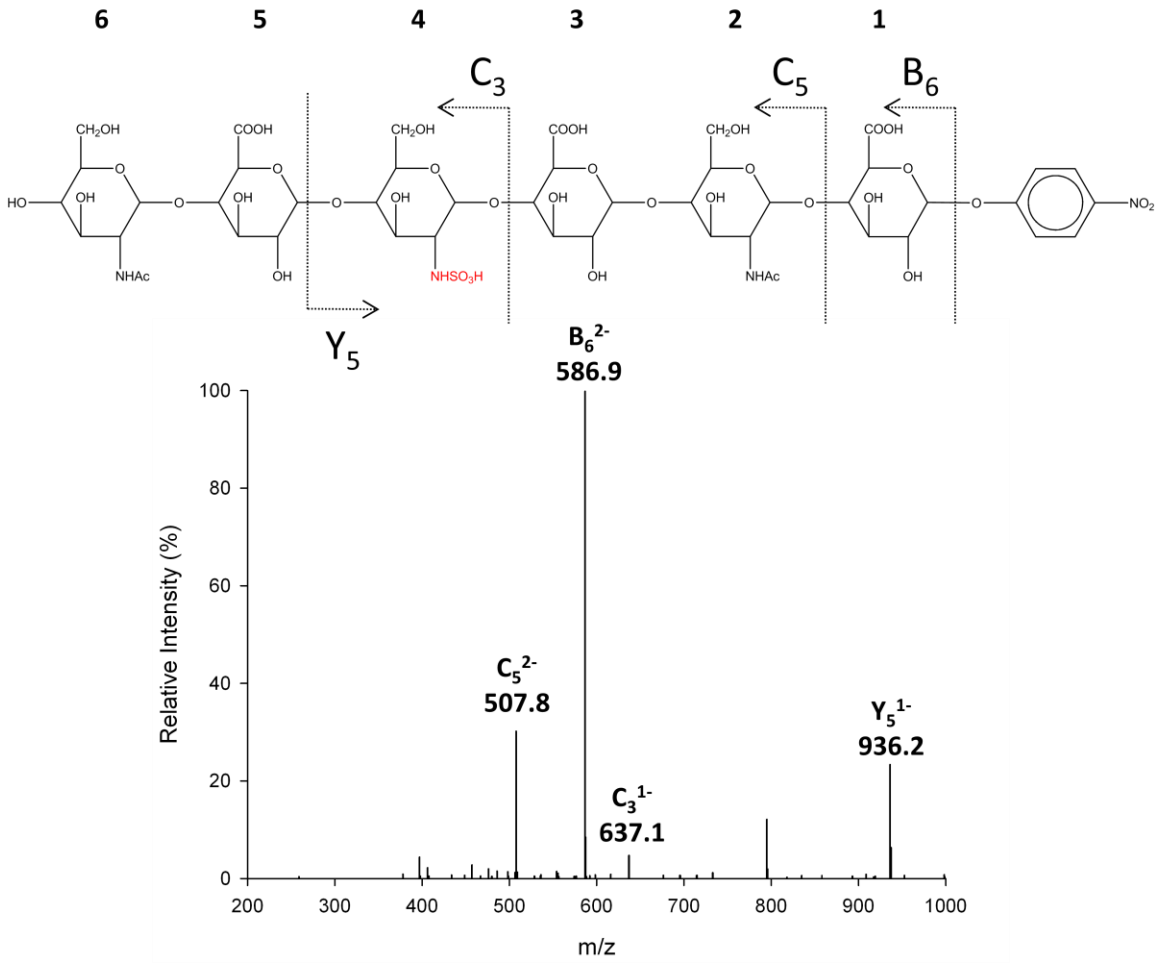

Figure 42. ***N*-sulfo hexasaccharide MS/MS analysis.** Precursor ion selection was at $[M-2H]^{2-}$, *m/z* 656.3. Analysis shows a GlcNS residue at monosaccharide position 4.

Incubation of **Hepta-1** with NDST-2 led to the observation of 3 peaks which were separated on Q sepharose-HPLC, eluted at 21 minutes, 26 minutes, and 31.5 minutes respectively (Figure 43A). This data revealed that there is more than one site of possible NDST-2 modification on the **Hepta-1** substrate. The fractions under each peak were collected and pooled individually (Pool A, Pool B, and Pool C). PAMN-HPLC analysis revealed single peaks for each sample, with PoolA eluting at 20 minutes, PoolB at 26 minutes, and Pool C at 30 minutes, indicating that there is most likely a sample with

multiple modifications by NDST-2 (Figure 43B). Each sample was then analyzed by ESI-MS (Figure 43C). Pool A displayed an *m/z* value of 725.3, representing a doubly charged ion that corresponds very closely to the molecular weight of the starting material, 1452.4 Da. Pool B displayed an *m/z* value of 496.1, representing a triply charged ion that corresponds very closely to the molecular weight of ***N*-sulfo heptasaccharide 1**, 1490.3 Da, carrying a single GlcNS monosaccharide unit. Pool C displayed an *m/z* value of 508.9, representing a triply charged ion that corresponds very closely to the molecular weight of ***N*-sulfo heptasaccharide 2**, 1528.3 Da, carrying two GlcNS monosaccharide units. PAMN-HPLC analysis shows the yield of ***N*-sulfo heptasaccharide 1** to be 29%, while the yield of ***N*-sulfo heptasaccharide 2** was 23%. The sample was then analyzed by MS/MS to determine the location of the GlcNS residues within each of the products (Figure 44). Analysis of ***N*-sulfo heptasaccharide 1** revealed three characteristic fragmented ions, C_2, C_4^{2-}, and B_5^{2-} at the *m/z* values of 396.7, 406.3, and 485.5, which provided the structural evidence that the GlcNS residue is located at position 4. Analysis of ***N*-sulfo heptasaccharide 2** revealed a single characteristic fragmented ion, C_4^{2-} at the *m/z* value of 425.1, which provided the structural evidence that the two GlcNS residues are located at positions 4 and 6 within the oligosaccharide product.

These data reveal that the addition of a GlcUA on the end of the hexasaccharide gave NDST-2 the ability to modify the GlcNAc residue at position 6. Position 6 remained unmodified in the reaction with **Hexa-1**. This data also indicates that NDST-2 prefers to modify GlcNAc residues when they are flanked by an adjacent GlcUA at the nonreducing end. In addition, the products from the **Hepta-1** reaction are nearly equal in yield distribution.

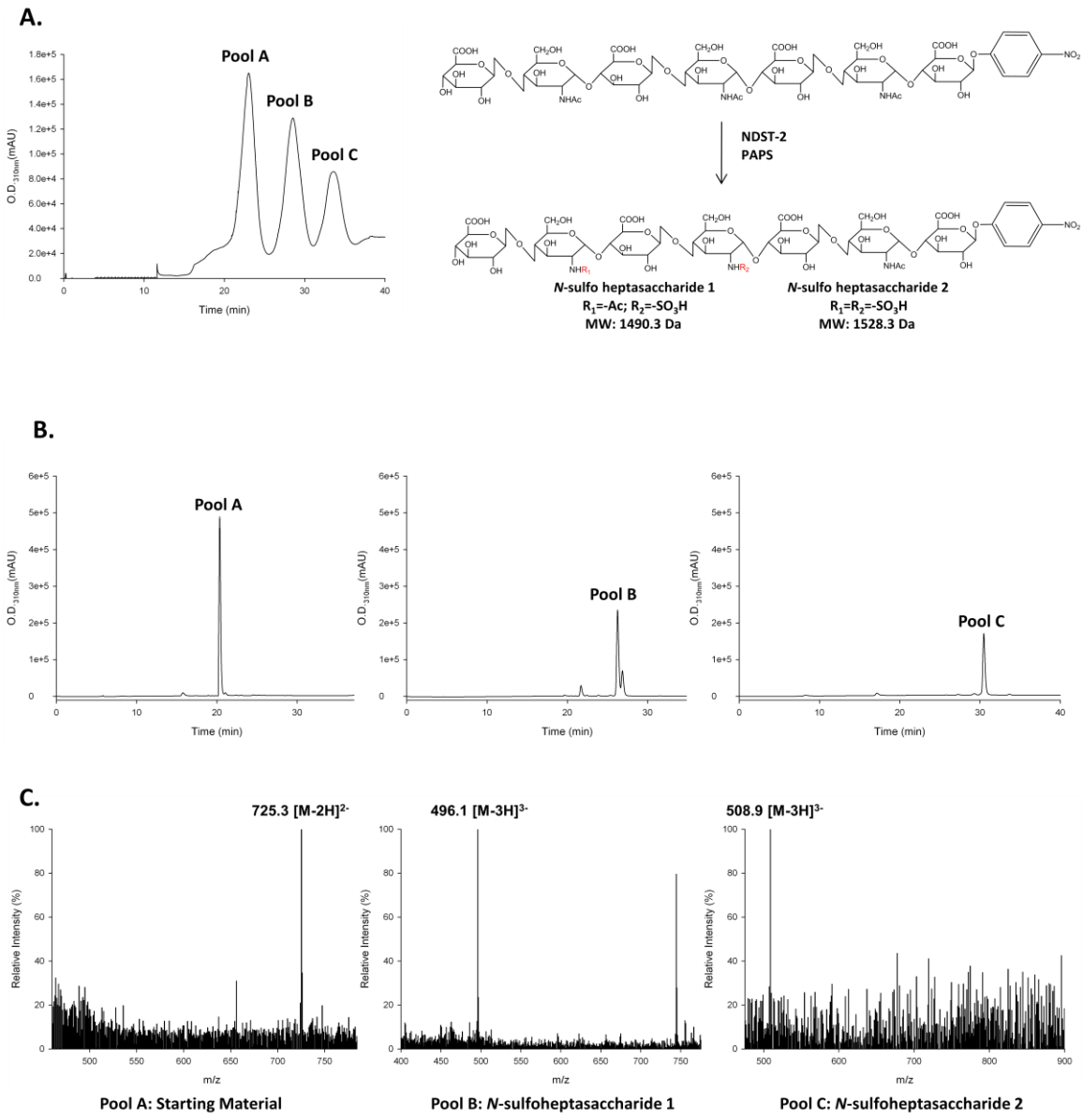

Figure 43. Hepta-1 + NDST-2 reaction analysis. (A) Q sepharose-HPLC purification reveals three peaks eluted at 21, 26, and 31.5 minutes respectively. (B) PAMN-HPLC confirms the presence of a single peak from Pool A eluted at 20 minutes, a single peak from Pool B eluted at 26 minutes, and a single peak from Pool C eluted at 30 minutes. (C) ESI-MS confirmation of starting material remaining in Pool A with *m/z* value of 725.3 corresponding with a doubly charged ion very close to the expected molecular weight of 1452.4 Da; confirmation of a product in Pool B containing one GlcNS with an *m/z* value of 496.1, corresponding to a triply charged ion with the expected molecular weight of 1490.3 Da; confirmation of a product in Pool C containing two GlcNS residues with an *m/z* value of 508.9, corresponding to a triply charged ion with the expected molecular weight of 1528.3.

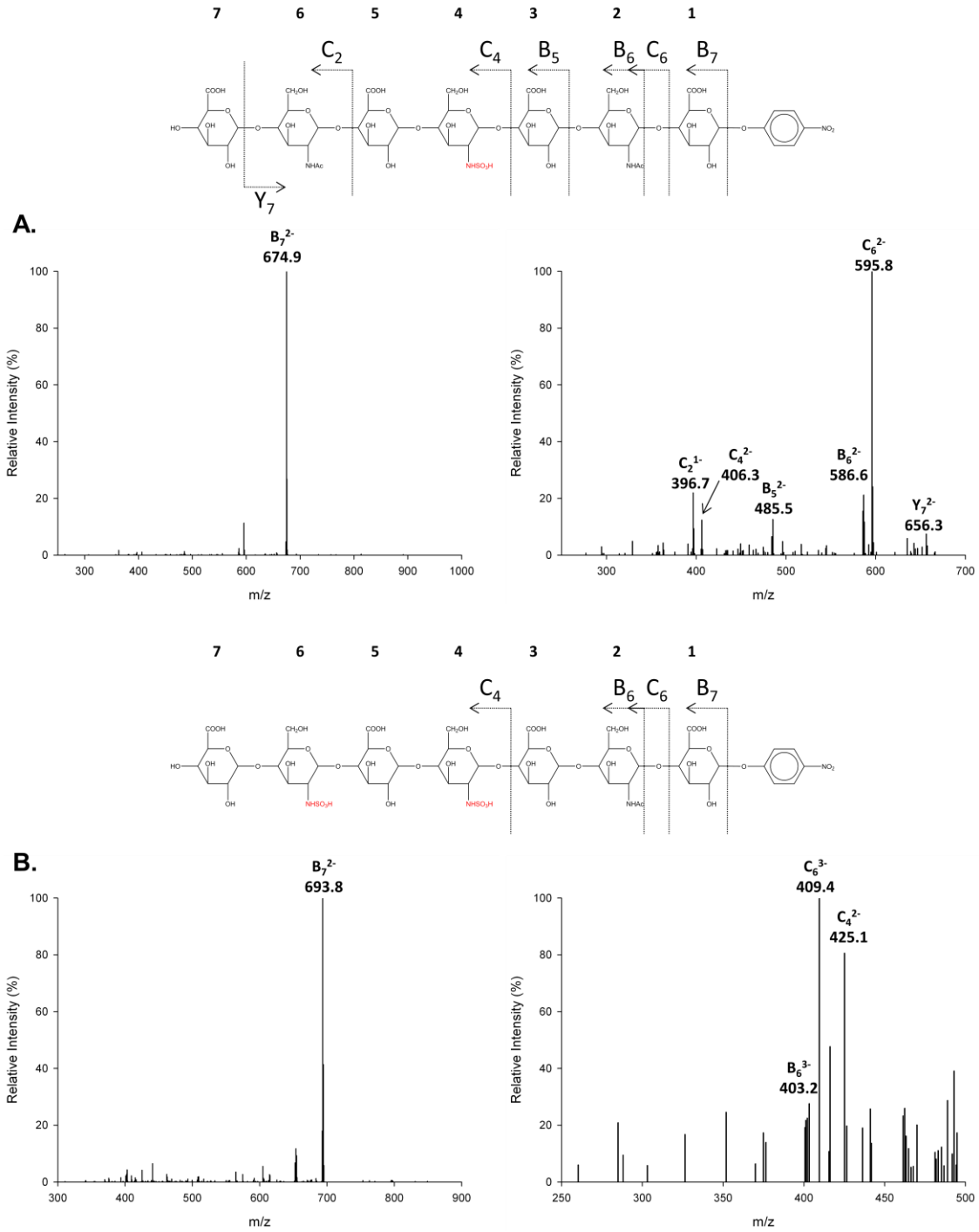

Figure 44. *N*-sulfo heptasaccharide MS/MS analysis. (A) *N*-sulfo heptasaccharide 1; Precursor ion selection was at $[M\text{-}3H]^{3-}$, *m/z* 496.1. Analysis shows a GlcNS residue at monosaccharide position 4. (B) *N*-sulfo heptasaccharide 2; Precursor ion selection was at $[M\text{-}3H]^{3-}$, *m/z* 508.9. Analysis shows two GlcNS residues at monosaccharide positions 4 and 6.

After incubation of **Octa-1** with NDST-2 we observed 3 minor peaks and one major peak by purification on Q sepharose-HPLC (Figure 45A). The first minor peak was eluted at 19 minutes, the second minor peak at 22.5 minutes, the major peak at 25 minutes, and the last minor peak at 28.5 minutes. This data revealed that **Octa-1** is the best substrate thus far for NDST-2 as it appears the majority of starting material has been converted into one major product. The fractions under each peak were collected and pooled individually (Pools A-D). PAMN-HPLC analysis revealed single peaks for Pools A-D, with elution times of 22.5 minutes for Pool A, 36.5 minutes for Pool B, 28 minutes for Pool C, and 31.5 minutes for Pool D (Figure 45B). The different elution times lead us to believe that each of these pools represent a unique set of modifications to the **Octa-1** starting material. Each sample was then analyzed by ESI-MS (Figure 46). Pool A displayed an *m/z* value of 550.9, representing a triply charged ion that corresponds very closely to the molecular weight of the starting material, 1655.5 Da. Pool B displayed an *m/z* value of 563.7, representing a triply charged ion that corresponds very closely to the molecular weight of ***N*-sulfo octasaccharide 1** carrying one GlcNS residue, 1693.4 Da. Pool C displayed an *m/z* value of 576.4, representing a triply charged ion that corresponds very closely to the molecular weight of ***N*-sulfo octasaccharide 2** carrying two GlcNS residues, 1731.4 Da. Pool D displayed an *m/z* value of 589.0, representing a triply charged ion that corresponds very closely to the molecular weight of ***N*-sulfo octasaccharide 3** carrying three GlcNS residues, 1769.3 Da. These data make it clear that the optimal modification on **Octa-1** is deacetylation and *N*-sulfation at two GlcNAc residue sites. PAMN-HPLC analysis shows that the product yield for the minor products ***N*-sulfo octasaccharide 1**and ***N*-sulfo octasaccharide 3** is 13% and 12% respectively.

While the yield for the major product, ***N*-sulfo octasaccharide 2**, is 68%. It is interesting to note the presence of a small product fraction maintaining three GlcNS residues. We have attempted to force the reaction by adjusting enzyme ratios and reaction time, however this product remains a minor product (10-15%) for each attempt (Data not shown). The sample was then analyzed by MS/MS to determine the location of the GlcNS residues within the oligosaccharide products (Figures 47-49). Analysis of ***N*-sulfo octasaccharide 1** revealed two characteristic fragmented ions Y_5 and C_5^{2-}, at the *m/z* values of 936.2 and 507.7, which provided the structural evidence that the GlcNS residue is at monosaccharide position 4 within the ***N*-sulfo octasaccharide 1** product (Figure 47). Analysis of ***N*-sulfo octasaccharide 2** revealed three characteristic fragmented ions Z_8^{2-}, Z_4, and B_6^{2-}, at the *m/z* values of 755.4, 676.9, and 606.1, which provided the structural evidence that the two GlcNS residues are located at monosaccharide positions 4 and 6 within the ***N*-sulfo octasaccharide 2** product (Figure 48). Analysis of ***N*-sulfo octasaccharide 3** revealed a single characteristic fragmented ion B_6^{3-}, at the *m/z* value of 416.3, which provided the structural evidence that the three GlcNS residues are located at monosaccharide positions 4, 6, and 8 within the ***N*-sulfo octasaccharide 3** product (Figure 49).

These data reveal that the addition of a GlcNAc residue at the nonreducing end of the heptasaccharide increases the ability of NDST-2 to form one major product. The remainder of less than 10% starting material means that nearly all of **Octa-1** is being modified by NDST-2 and it is an optimal substrate for the enzyme. Interestingly, there is a presence of a product carrying three GlcNS residues, but this may be an artifact that is

under more stringent cellular control *in vivo*, as it is not readily produced *in vitro*. For the most part, the GlcNAc residue at position 8 seems to prefer to remain unmodified.

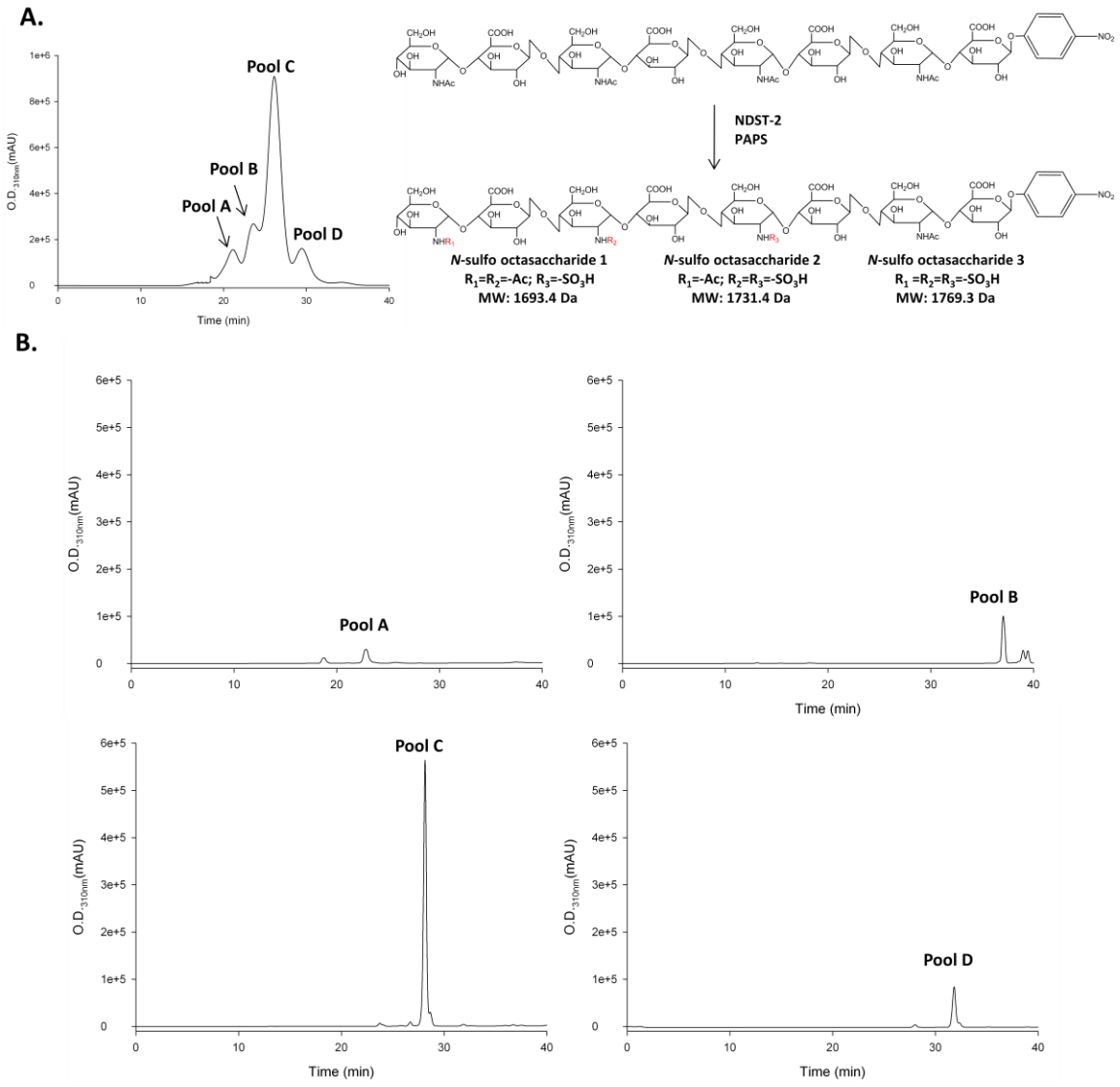

Figure 45. Octa-1 + NDST-2 HPLC reaction analysis. (A) Q sepharose-HPLC purification reveals four peaks eluted at 19, 22.5, 25 and 28.5 minutes respectively. (B) PAMN-HPLC confirms the presence of a single peak from Pool A eluted at 22.5 minutes, a single peak from Pool B eluted at 36.5 minutes, a single peak from Pool C eluted at 28 minutes, and a single peak from Pool D eluted at 31.5 minutes.

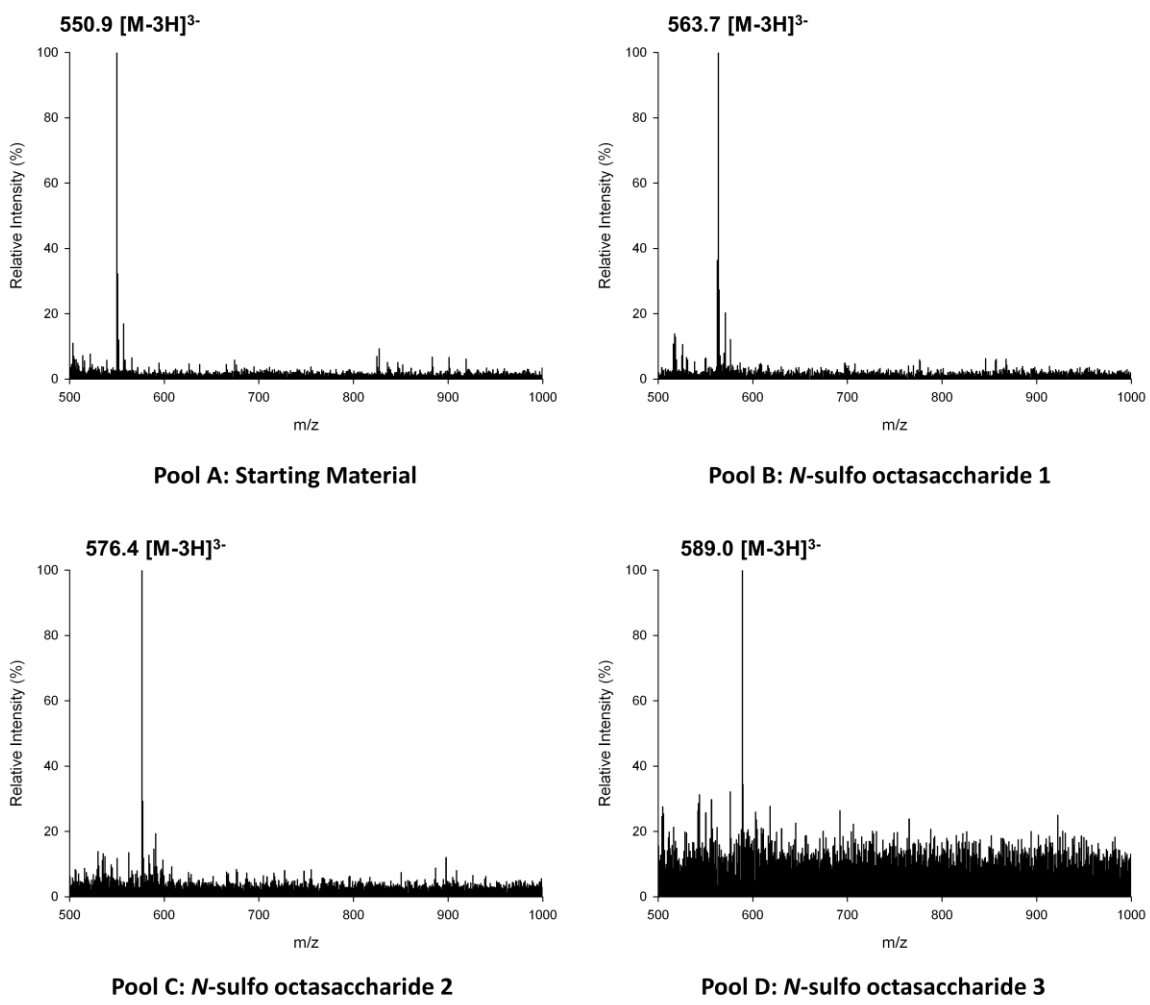

Figure 46. Octa-1 + NDST-2 ESI-MS reaction analysis. ESI-MS confirmation of starting material remaining in Pool A with *m/z* value of 550.9 corresponding with a triply charged ion very close to the expected molecular weight of 1655.5 Da; confirmation of a product in Pool B containing one GlcNS with an *m/z* value of 563.7, corresponding to a triply charged ion with the expected molecular weight of 1693.4 Da; confirmation of a product in Pool C containing two GlcNS residues with an *m/z* value of 576.4, corresponding to a triply charged ion with the expected molecular weight of 1731.4 Da; confirmation of a product in Pool D containing three GlcNS residues with an *m/z* value of 589.0, corresponding to a triply charged ion with the expected molecular weight of 1769.3 Da.

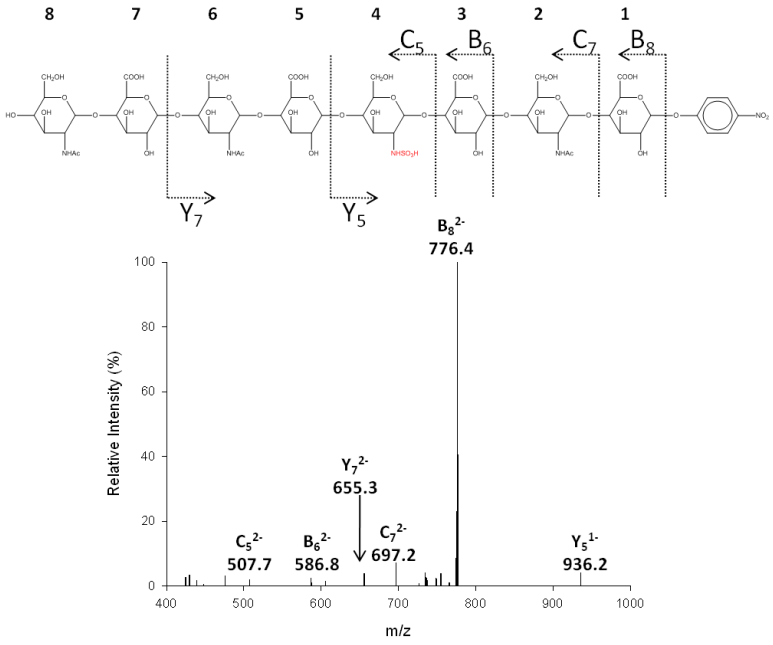

Figure 47. *N*-sulfo octasaccharide 1 MS/MS analysis. Precursor ion selection was at $[M-3H]^{3-}$, *m/z* 563.7. Analysis shows a GlcNS residue at monosaccharide position 4.

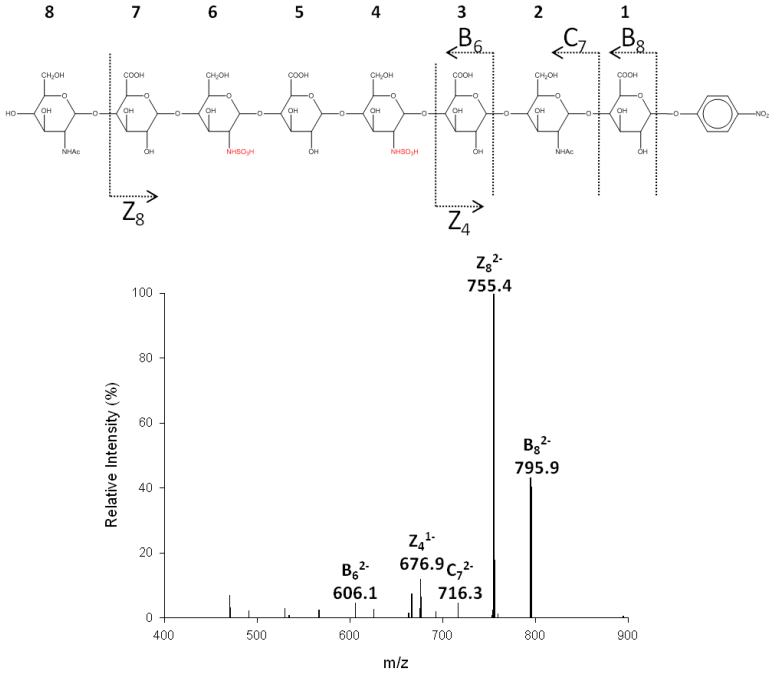

Figure 48. *N*-sulfo octasaccharide 2 MS/MS analysis. Precursor ion selection was at $[M-3H]^{3-}$, *m/z* 576.4. Analysis shows two GlcNS residues at monosaccharide positions 4 and 6.

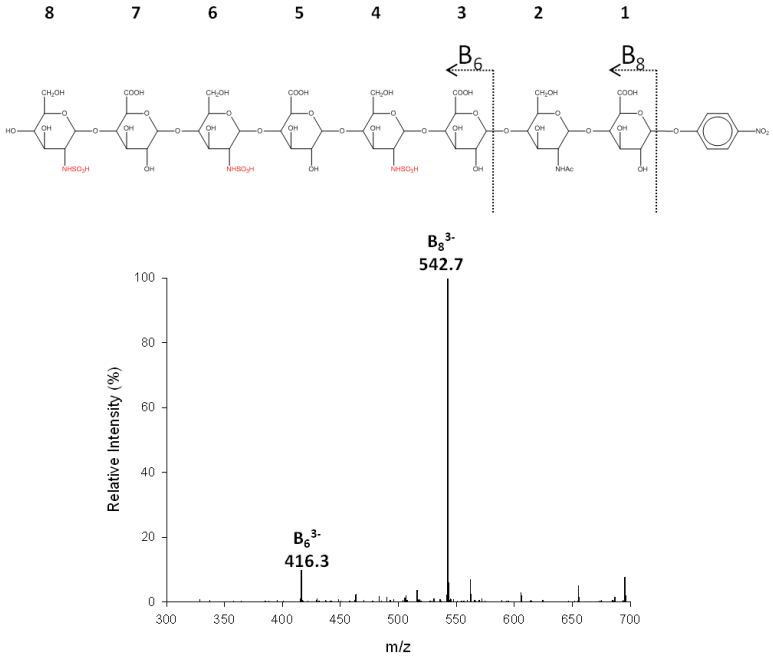

Figure 49. ***N*-sulfo octasaccharide 3 MS/MS analysis.** Precursor ion selection was at $[M-3H]^{3-}$, *m/z* 589.0. Analysis shows three GlcNS residues at monosaccharide positions 4, 6, and 8.

After incubation of **Nona-1** with NDST-2 we observed 4 major peaks by purification on Q sepharose-HPLC, eluted at 21 minutes, 26 minutes, 31 minutes, and 34.5 minutes respectively (Figure 50A). This data revealed that **Nona-1** is a substrate for NDST-2, but does not form a single major product as was observed with **Octa-1**. The fractions under each peak were collected and pooled individually (Pools A-D). PAMN-HPLC analysis revealed single peaks for Pools A, B, and D, with elution times of 22.5 minutes, 26 minutes, and 35.5 minutes (Figure 50B). Pool C was observed as two peaks with elution times of 31 and 32 minutes, most likely due to a pressure malfunction within the HPLC system. Each sample was analyzed by ESI-MS (Figure 51). Pool A displayed an *m/z* value of 609.8, representing a triply charged ion that corresponds very closely to the molecular weight of the starting material, 1831.5 Da. Pool B displayed *m/z* values of 466.8 and 622.5, representing both quadruply and triply charged ions that correspond

very closely to the molecular weight of ***N*-sulfo nonasaccharide 1** carrying one GlcNS residue, 1869.5 Da. Pool C displayed an *m/z* value of 635.1, representing a triply charged ion that corresponds very closely to the molecular weight of ***N*-sulfo nonasaccharide 2** carrying two GlcNS residues, 1907.4 Da. Pool D displayed *m/z* values of 485.7 and 647.7, representing both quadruply and triply charged ions that correspond very closely to the molecular weight of ***N*-sulfo nonasaccharide 3** carrying three GlcNS residues, 1945.3 Da. PAMN-HPLC analysis shows that the yield distribution of these products is relatively even with ***N*-sulfo nonasaccharide 1**at 21%,***N*-sulfo nonasaccharide 2** at 21%, and ***N*-sulfo nonasaccharide 3** garnering 38% of the total yield, with 20% of unmodified starting material. As seen in the **Hepta-1** reaction with NDST-2, the presence of GlcUA at the nonreducing end prevents the reaction from resulting in the formation of a single major product, instead a relatively even distribution is formed between products carrying either one, two, or three GlcNS residues. The samples were then analyzed by MS/MS to determine the location of the GlcNS residues within the oligosaccharide products (Figures 52-54). Analysis of ***N*-sulfo nonasaccharide 1** revealed two characteristic fragmented ions B_5^{2-} and B_6^{2-}, at the *m/z* values of 466.9 and 586.8, providing the structural evidence that the GlcNS residue is at monosaccharide position 4 within the ***N*-sulfo nonasaccharide 1** product (Figure 52). Analysis of ***N*-sulfo nonasaccharide 2** revealed two characteristic fragmented ions Z_8^{3-} and C_6^{2-}, at the *m/z* values of 754.4 and 614.5, providing the structural evidence that the two GlcNS residues are at monosaccharide positions 4 and 6 within the ***N*-sulfo nonasaccharide 2** product (Figure 53). Analysis of ***N*-sulfo nonasaccharide 3** revealed three characteristic fragmented ions B_4^{2-}, B_5^{2-}, and B_7^{3-}, at the *m/z* values of 416.4, 504.3, and 474.9,

providing the structural evidence that the three GlcNS residues are at monosaccharide positions 4, 6, and 8 within the ***N*-sulfo nonasaccharide 3** product.

Taken together, these data reveal that the modifications on **Nona-1** are occurring in a manner very closely resembling the previous reactions. The most important note here is the incomplete manner by which the reaction is carried out. Much like the reaction with **Hepta-1**, the products have a relatively even distribution. But when compared to the single major product for **Octa-1**, it appears that the GlcNAc residue on the nonreducing end is important for allowing the reaction to proceed fully and form a single major product. In both cases of **Hepta-1** and **Nona-1** it is evident that the even product distribution is a by-product of placing a GlcUA residue at the nonreducing end.

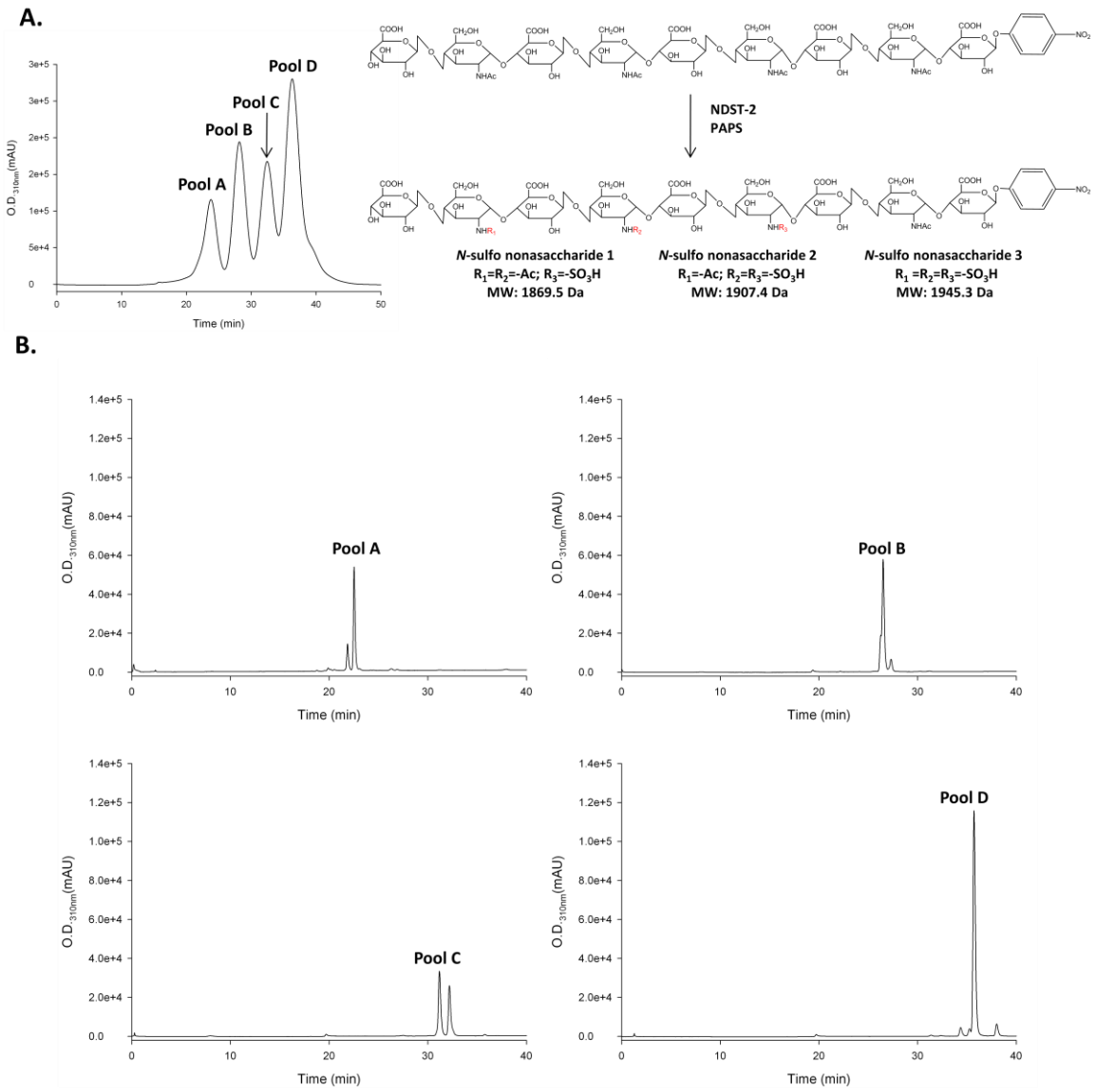

Figure 50. Nona-1 + NDST-2 HPLC reaction analysis. (A) Q sepharose-HPLC purification reveals four peaks eluted at 21, 26, 31 and 34.5 minutes respectively. (B) PAMN-HPLC confirms the presence of a single peak from Pool A eluted at 22.5 minutes, a single peak from Pool B eluted at 26 minutes, two peaks from Pool C eluted at 31 and 32 minutes, and a single peak from Pool D eluted at 35.5 minutes.

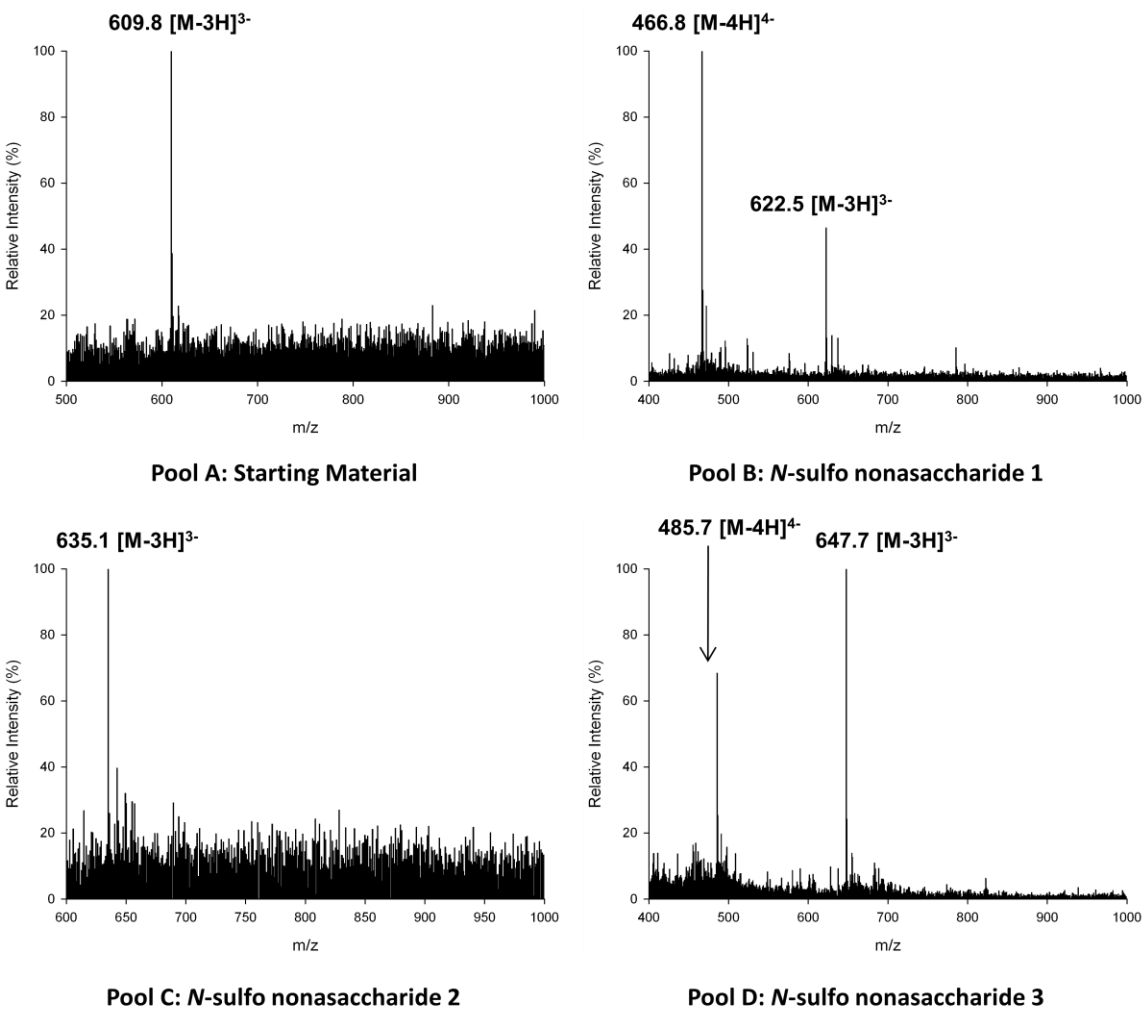

Figure 51. Nona-1 + NDST-2 ESI-MS reaction analysis. ESI-MS confirmation of starting material remaining in Pool A with *m/z* value of 609.8 corresponding with a triply charged ion very close to the expected molecular weight of 1831.5 Da; confirmation of a product in Pool B containing one GlcNS with an *m/z* value of 466.8, corresponding to a quadruply charged ion with the expected molecular weight of 1869.5 Da; confirmation of a product in Pool C containing two GlcNS residues with an *m/z* value of 635.1, corresponding to a triply charged ion with the expected molecular weight of 1907.4 Da; confirmation of a product in Pool D containing three GlcNS residues with an *m/z* value of 647.7, corresponding to a triply charged ion with the expected molecular weight of 1945.3 Da.

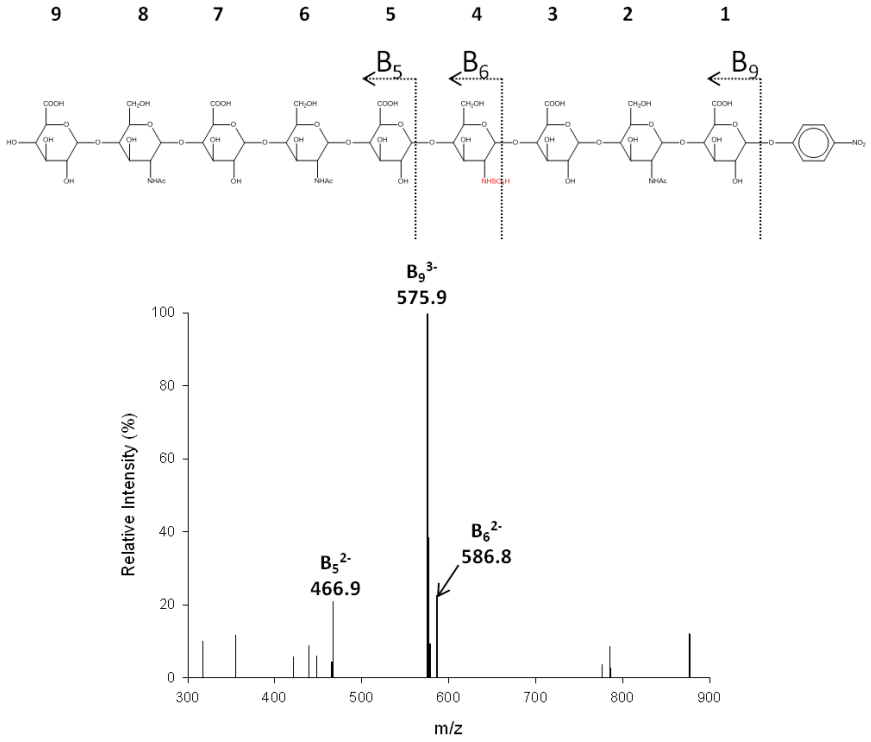

Figure 52. *N*-sulfo nonasaccharide 1 MS/MS analysis. Precursor ion selection was at $[M-4H]^{4-}$, *m/z* 466.8. Analysis shows a GlcNS residue at monosaccharide position 4.

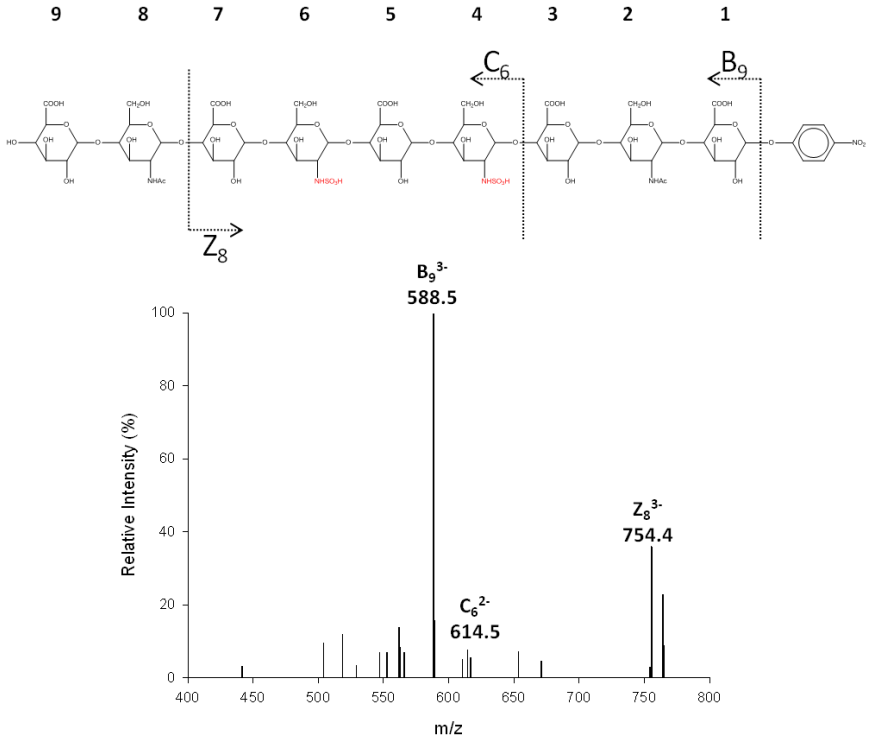

Figure 53. *N*-sulfo nonasaccharide 2 MS/MS analysis. Precursor ion selection was at $[M-3H]^{3-}$, *m/z* 635.1. Analysis shows two GlcNS residues at monosaccharide positions 4 and 6.

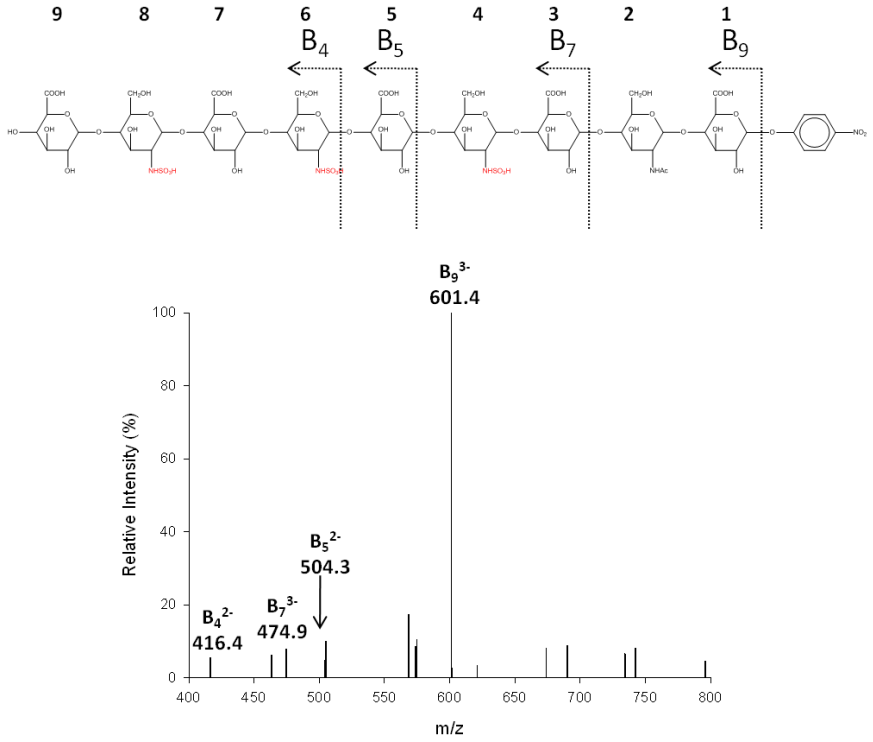

Figure 54. *N*-sulfo nonasaccharide 3 MS/MS analysis. Precursor ion selection was at $[M-3H]^{3-}$, *m/z* 647.7. Analysis shows three GlcNS residues at monosaccharide positions 4, 6, and 8.

After incubation of **Deca-1** with NDST-2 we observed three minor products and one major peak by purification on Q sepharose-HPLC (Figure 55A). There were two minor products that eluted at 26 minutes and 30 minutes, followed by the major peak at 33 minutes, and the final minor product at 38 minutes. This data revealed that **Deca-1** is an outstanding substrate for NDST-2, as one single major product has been formed. The fractions under each peak were collected and pooled individually (Pools A-D). PAMN-HPLC analysis revealed a very small single peak for Pool A eluting at 19 minutes, another very small single peak for Pool B eluting at 26 minutes, a large single peak for the major product Pool C eluting at 31 minutes, and a single peak for Pool D eluting at 34 minutes (Figure 55B). Each sample was analyzed by ESI-MS (Figure 56). Pool A displayed an *m/z* value of 1034.8, representing a doubly charged ion that corresponds

very closely to the molecular weight of ***N*-sulfo decasaccharide 1** carrying one GlcNS residue, 2072.5 Da. Pool B displayed an *m/z* value of 1054.1, representing a doubly charge ion that corresponds very closely to the molecular weight of ***N*-sulfo decasaccharide 2** carrying two GlcNS residues, 2110.5 Da. Pool C displayed *m/z* values of 536.5 and 715.5, representing both quadruply and triply charged ions that correspond very closely to the molecular weight of ***N*-sulfo decasaccharide 3** carrying three GlcNS residues, 2148.4 Da. Pool D displayed *m*/*z* values of 728.0 and 1092.1, representing both triply and doubly charged ions that correspond closely to the molecular weight of ***N*-sulfo decasaccharide 4** carrying four GlcNS residues, 2186.4 Da. PAMN-HPLC analysis shows that the yield distribution greatly favors the single major product, ***N*-sulfo decasaccharide 3**, with 83% yield. The minor products ***N*-sulfo decasaccharide 1** and ***N*-sulfo decasaccharide 2** garnered only 2% and 3% yield respectively. While ***N*-sulfo decasaccharide 4**, was produced at 12% yield. There was no detectable starting material remaining in the **Deca-1** reaction. Overall, these results prove that **Deca-1** is the best substrate to be tested thus far as all of the starting material has been consumed and the overwhelming majority of product was present as a single product containing three GlcNS residues. The samples were then analyzed by MS/MS to determine the location of the GlcNS residues within the oligosaccharide products (Figures 57-59). Analysis of ***N*-sulfo decasaccharide 1** revealed a mixture of two products carrying a single GlcNS residue (Figure 52A). The first product revealed two characteristic fragmented ions B_5 and B_8^{2-}, at the *m/z* values of 964.0 and 776.2, providing the structural evidence that the GlcNS residue is at the monosaccharide position 4 (Figure 57A). However, a second product carrying one GlcNS residue revealed two characteristic fragmented ions Y_9^{2-} and

Z_7^{2-}, at the *m/z* values of 845.2 and 629.6, providing the structural evidence that the GlcNS residue is at the monosaccharide position 8 (Figure 57B). This was the first time we had encountered this phenomenon, and it could indicate that once the oligosaccharide reaches a certain length (i.e. 10 monosaccharide units) it loses its direction and can modify the oligosaccharide in either direction. Analysis of ***N*-sulfo decasaccharide 2** revealed two characteristic fragmented ions B_8^{2-} and Z_8^{2-}, at the *m/z* values of 795.4 and 755.0, providing the structural evidence that the two GlcNS residues are at monosaccharide positions 4 and 6 (Figure 58). Analysis of the major product, ***N*-sulfo decasaccharide 3**, revealed two characteristic fragmented ions C_7^{2-} and Z_{10}^{3-}, at the *m/z* values of 734.4 and 642.5, providing the structural evidence that the three GlcNS residues are at monosaccharide positions 4,6, and 8 (Figure 59). Unfortunately, MS/MS analysis of ***N*-sulfo decasaccharide 4** was not accomplished as a clear fragmentation pattern was not observed.

These data seem to show that **Deca-1** is an outstanding substrate for NDST-2, as one single major product was produced and all of the starting material was consumed. As seen with **Octa-1**, the presence of a GlcNAc residue at the nonreducing end produces the effect of driving the reaction to completion. Whereas, when GlcUA is at the nonreducing end, in the cases of **Hepta-1** and **Nona-1**, we see a nearly even distribution of products.

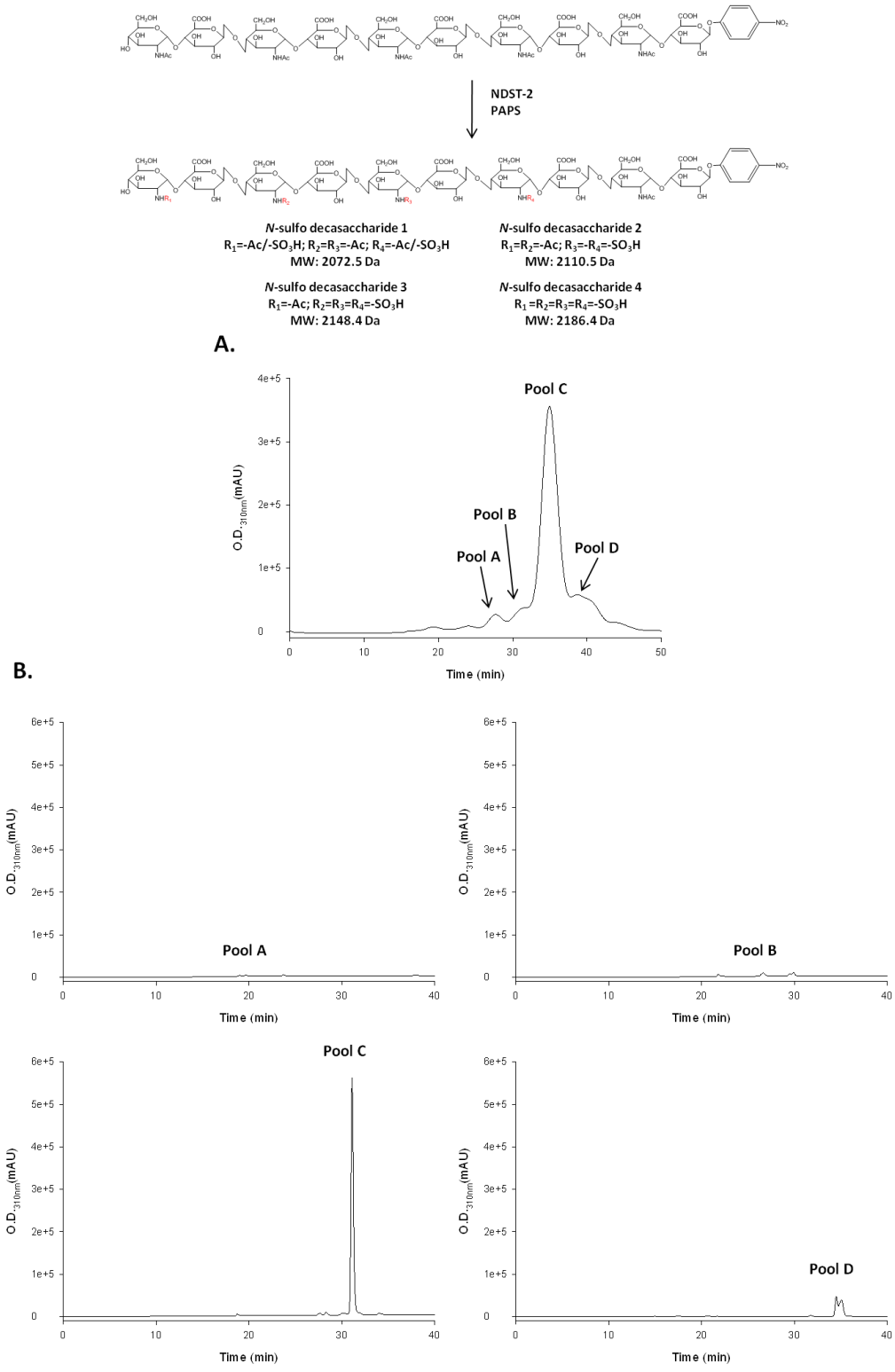

Figure 55. Deca-1 + NDST-2 HPLC reaction analysis. (A) Q sepharose-HPLC purification reveals four peaks eluted at 26, 30, 33 and 38 minutes respectively. (B) PAMN-HPLC confirms the presence of a single peak from Pool A eluted at 19 minutes, a single peak from Pool B eluted at 26 minutes, a single peak from Pool C eluted at 31 minutes, and a single peak from Pool D eluted at 34 minutes.

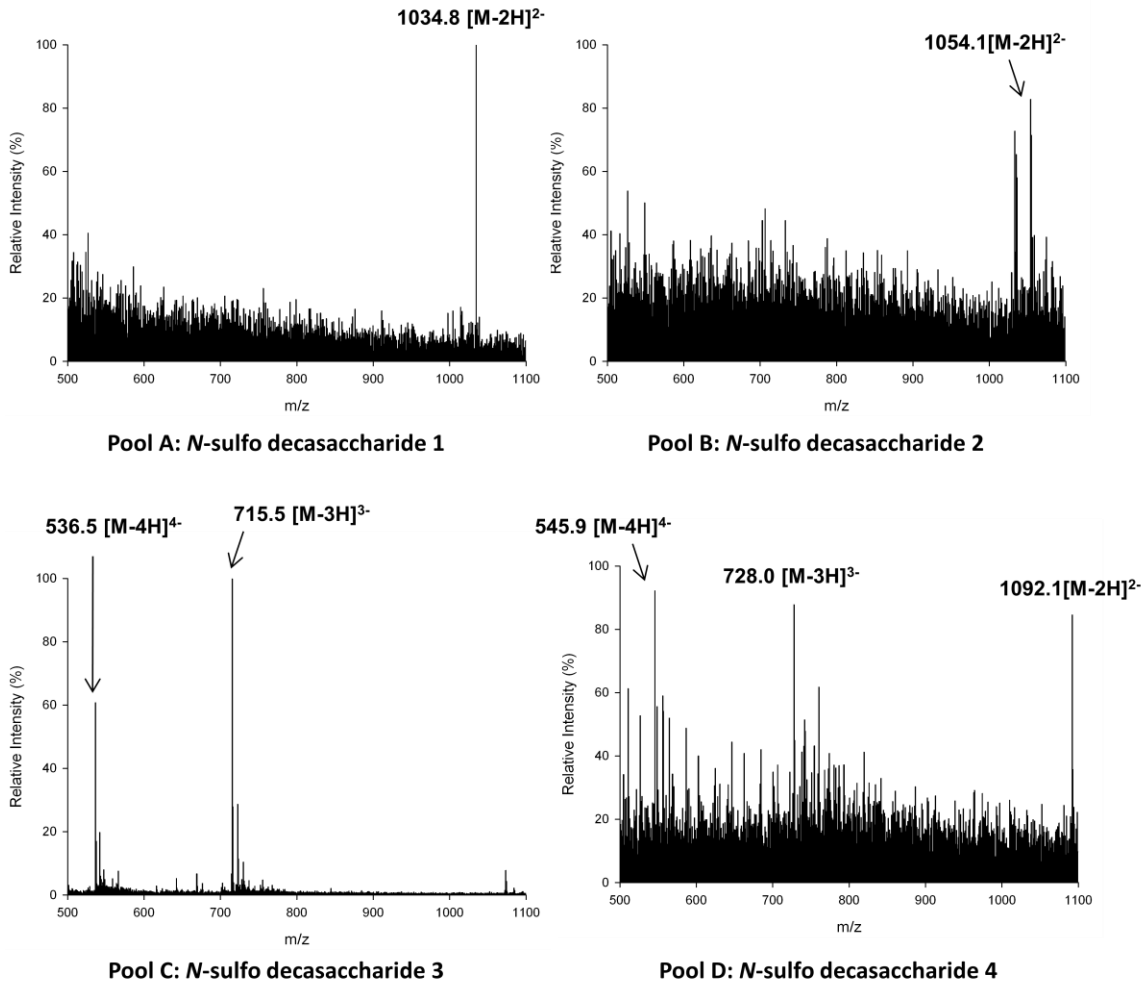

Figure 56. Deca-1 + NDST-2 ESI-MS reaction analysis. ESI-MS confirmation of a product in Pool A containing one GlcNS with *m/z* value of 1034.8 corresponding with a doubly charged ion very close to the expected molecular weight of 2072.5 Da; confirmation of a product in Pool B containing two GlcNS residues with an *m/z* value of 1054.1, corresponding to a doubly charged ion with the expected molecular weight of 2110.5 Da; confirmation of a product in Pool C containing three GlcNS residues with an *m/z* value of 715.5, corresponding to a triply charged ion with the expected molecular weight of 2148.4 Da; confirmation of a product in Pool D containing four GlcNS residues with an *m/z* value of 728.0, corresponding to a triply charged ion with the expected molecular weight of 2186.4 Da.

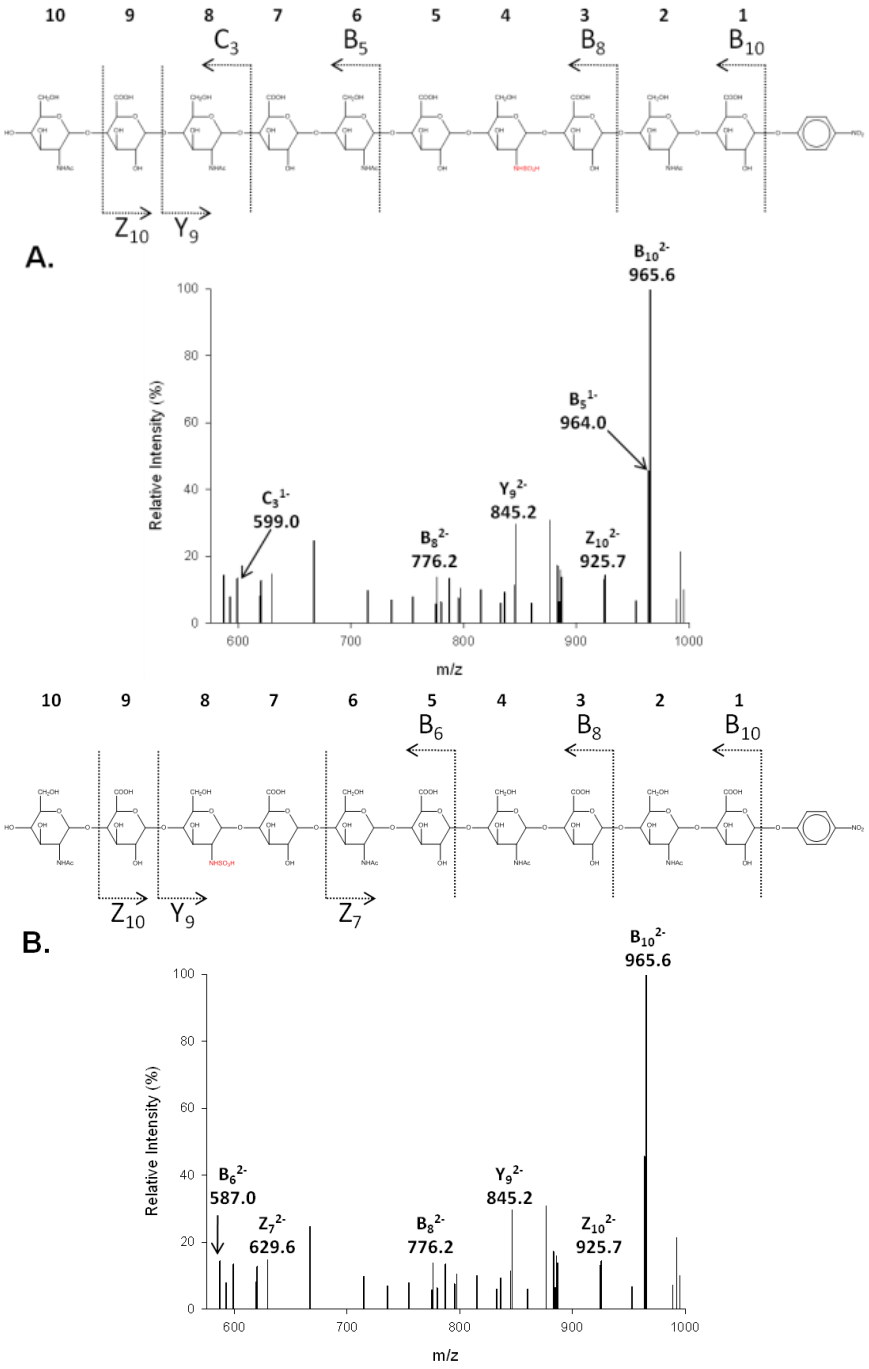

Figure 57. *N*-sulfo decasaccharide 1 MS/MS analysis. Precursor ion selection was at $[M-2H]^{2-}$, *m/z* 1034.8. Analysis shows a GlcNS residue at monosaccharide position 4 (A) or 8 (B).

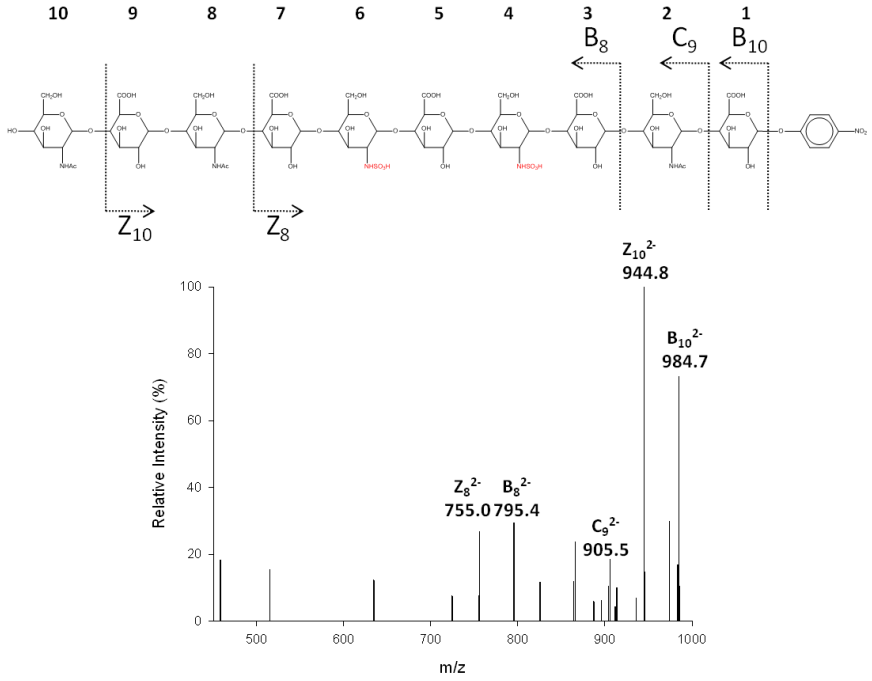

Figure 58. *N*-sulfo decasaccharide 2 MS/MS analysis. Precursor ion selection was at $[M-2H]^{2-}$, *m/z* 1054.1. Analysis shows two GlcNS residues at monosaccharide positions 4 and 6.

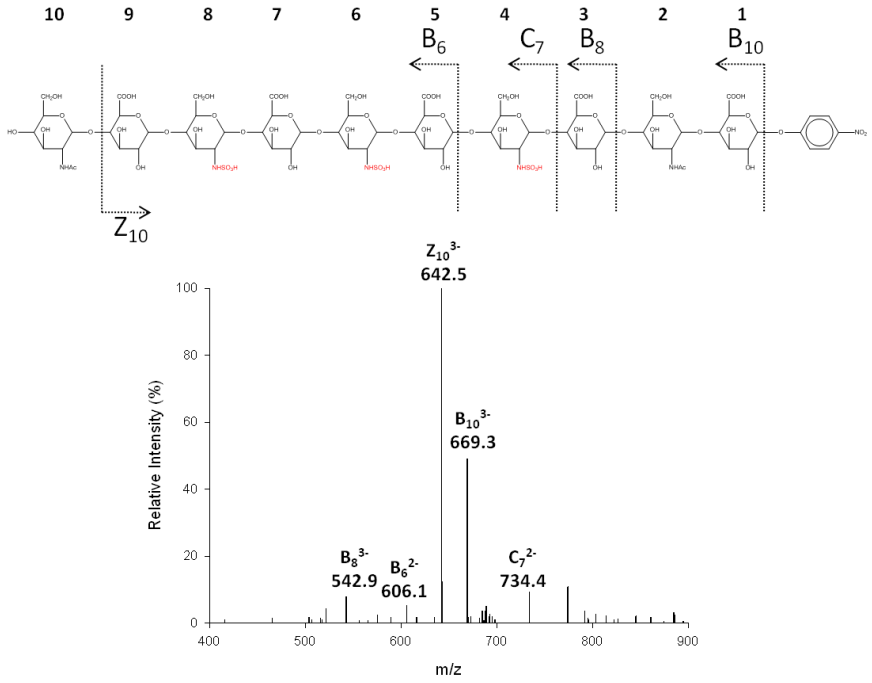

Figure 59. *N*-sulfo decasaccharide 3 MS/MS analysis. Precursor ion selection was at $[M-3H]^{3-}$, *m/z* 715.5. Analysis shows three GlcNS residues at monosaccharide positions 4, 6, and 8.

Understanding the effect of pre-existing oligosaccharide substrate *N*-sulfation on the modification of NDST-2

In order to determine any effect from a pre-existing *N*-sulfation within the structurally defined oligosaccharide substrate, we tested both a nonasaccharide (**Nona-2**) and decasaccharide (**Deca-2**) containing a GlcNS residue in the 6th monosaccharide position. After incubation of **Nona-2** with NDST-2 we observed three peaks, including one major peak, by purification on Q sepharose-HPLC. The minor peaks eluted at 27.5 minutes and 31.5 minutes, while the major peak eluted at 34.5 minutes (Figure 60A). This data reveals that **Nona-2** is an ideal substrate for NDST-2, as a single major product peak is observed after incubation. The fractions under each peak were collected and pooled individually (Pools A-C). PAMN-HPLC analysis revealed single peaks for Pools A-C, with elution times of 25 minutes, 30 minutes, and 34 minutes respectively (Figure 60B). Each sample was analyzed by ESI-MS (Figure 60C). Pool A displayed an *m/z* value of 933.9, representing a doubly charged ion that corresponds very closely to the molecular weight of the starting material carrying one GlcNS residue, 1869.5 Da. Pool B displayed an *m/z* value of 952.8, representing a doubly charged ion that corresponds very closely to the molecular weight of ***N*-sulfo nonasaccharide 4** carrying two GlcNS residues, 1907.4. The major product in Pool C displayed *m/z* values of 485.7, 647.7, and 971.8, representing a quadruply charged, triply charged, and doubly charged ion that corresponds very closely to the molecular weight of ***N*-sulfo nonasaccharide 5** carrying three GlcNS residues, 1945.3 Da. PAMN-HPLC analysis shows that the yield distribution of these products is nearly all one single product with ***N*-sulfo nonasaccharide 5** garnering 89% of the yield. ***N*-sulfo nonasaccharide 4** was

responsible for only 6% of the total yield and only 5% of the starting material remained. This result shows that the presence of a pre-existing *N*-sulfation on the oligosaccharide substrate has a positive effect on the ability of the reaction to run to completion and form a single major product. This result is in direct contrast to the even product distribution observed in the reaction with **Nona-1**. The samples were then analyzed by MS/MS to determine the location of the GlcNS residues resulting from the modification by NDST-2 (Figures 61-62). As was seen in the reaction with **Deca-1**, we observed a combination of two structures for ***N*-sulfo nonasaccharide 4** (Figure 61). The first structure revealed two characteristic fragmented ions B_4^{2-} and C_5^{2-}, at the *m/z* values of 416.4 and 512.2, providing the structural evidence that the newly modified GlcNS residue is in monosaccharide position 8 (Figure 61B). Meanwhile another structure was observed by three characteristic fragmented ions B_6^{2-}, Z_5^{2-}, and Z_8^{2-}, at the *m/z* values of 693.8, 458.3, and 755.3, providing the structural evidence that the newly modified GlcNS residue is at oligosaccharide position 4 within the ***N*-sulfo nonasaccharide 4** product (Figure 61A). Analysis of ***N*-sulfo nonasaccharide 5** revealed two characteristic fragmented ions B_5^{2-} and B_7^{3-}, at the *m/z* positions 504.3 and 475.0, providing the structural evidence that the newly modified GlcNS residues are in oligosaccharide positions 4 and 8 within the ***N*-sulfo nonasaccharide 5** product (Figure 62).

These data suggest that pre-existing *N*-sulfation within the oligosaccharide substrate enhances the ability of NDST-2 to finish the reaction and create one single major product. It is also evident that NDST-2 can begin modification on either end of this substrate due to the mixture of structures observed from ***N*-sulfo nonasaccharide 4** and most likely does not move in a specific direction.

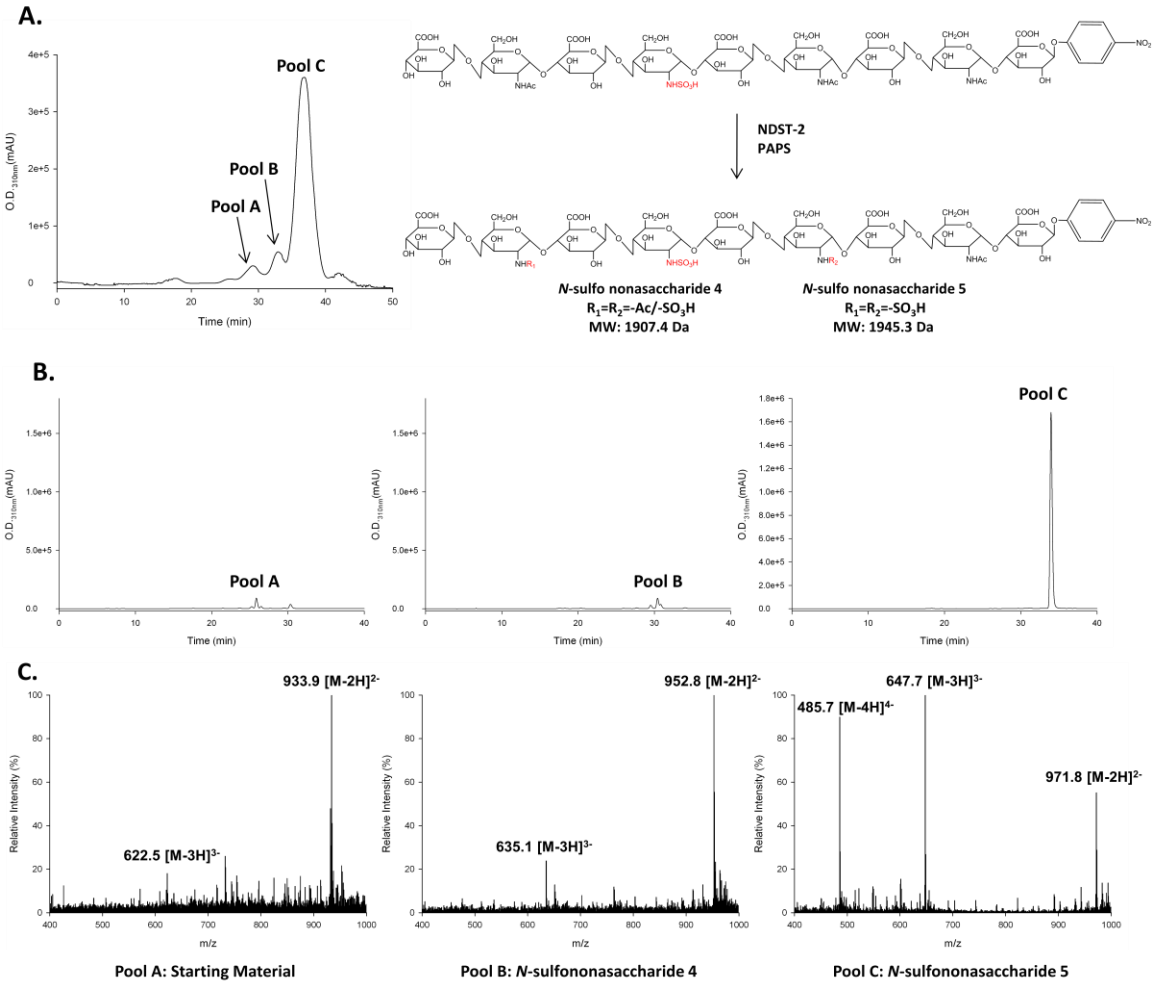

Figure 60. Nona-2 + NDST-2 reaction analysis. (A) Q sepharose-HPLC purification reveals three peaks eluted at 27.5, 31.5, and 34.5 minutes respectively. (B) PAMN-HPLC confirms the presence of a single peak from Pool A eluted at 25 minutes, a single peak from Pool B eluted at 30 minutes, and a single peak from Pool C eluted at 34 minutes. (C) ESI-MS confirmation of a product in Pool A containing one GlcNS with *m/z* value of 933.9 corresponding with a doubly charged ion very close to the expected molecular weight of 1869.5 Da; confirmation of a product in Pool B containing two GlcNS residues with an *m/z* value of 952.8, corresponding to a doubly charged ion with the expected molecular weight of 1907.4 Da; confirmation of a product in Pool C containing three GlcNS residues with an *m/z* value of 647.7, corresponding to a triply charged ion with the expected molecular weight of 1945.3 Da.

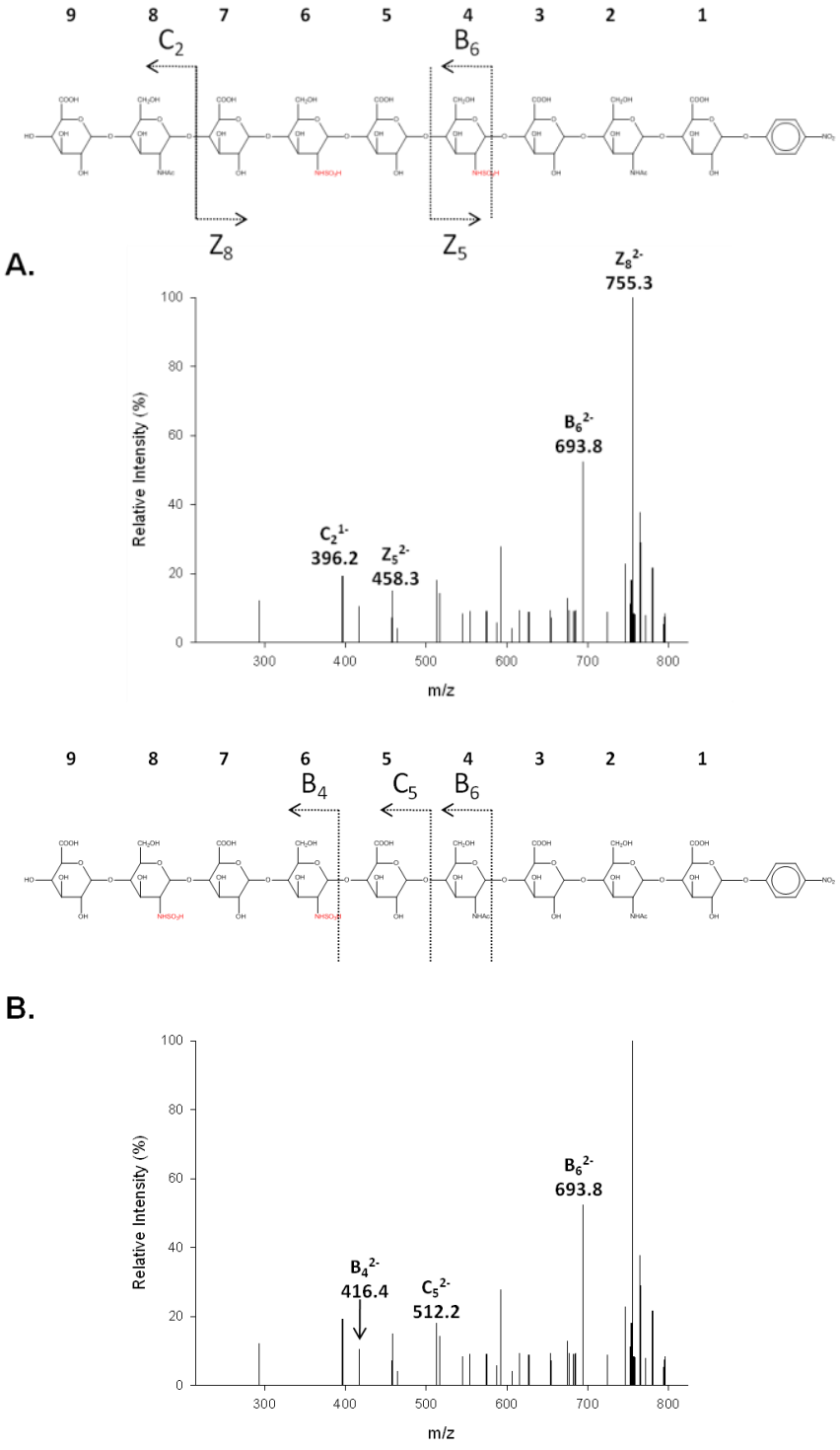

Figure 61. *N*-sulfo nonasaccharide 4 MS/MS analysis. Precursor ion selection was at $[M-2H]^{2-}$, *m/z* 952.8. Analysis shows a new GlcNS residue at either monosaccharide position 4 (A) or 8 (B). (B) *N*-sulfo nonasaccharide 5; Precursor ion selection was at $[M-3H]^{3-}$, *m/z* 647.7. Analysis shows two new GlcNS residues at monosaccharide positions 4 and 8.

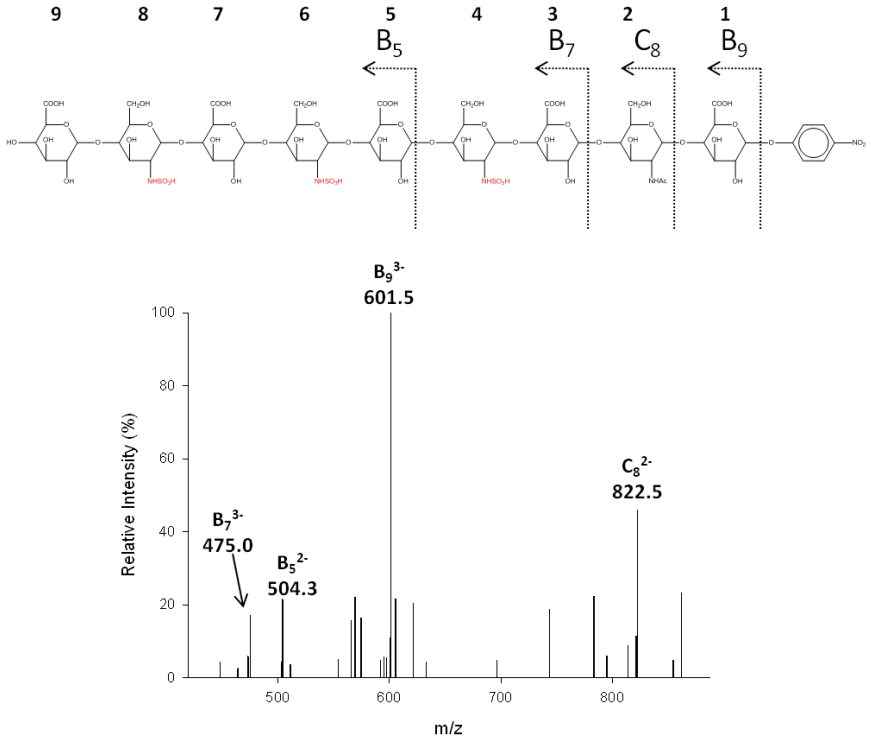

Figure 62. *N*-sulfo nonasaccharide 5 MS/MS analysis. Precursor ion selection was at $[M-3H]^{3-}$, *m/z* 647.7. Analysis shows two new GlcNS residues at monosaccharide positions 4 and 8.

After incubation of **Deca-2** with NDST-2 we observed two peaks, one major and one minor, by purification on Q sepharose-HPLC (Figure 63A). The minor peak eluted at 30 minutes and the major peak eluted at 33 minutes. This data reveals that **Deca-2** is likely the most ideal substrate of the entire library. The fractions under each peak were collected and pooled individually (Pool A and Pool B). PAMN-HPLC analysis revealed single peaks for both Pool A and Pool B with elution times of 28 minutes and 31 minutes respectively (Figure 63B). Each sample was then analyzed by ESI-MS (Figure 63C). Pool A displayed *m/z* values of 526.9 and 702.8, representing quadruply and triply charged ions that correspond very closely to the molecular weight of ***N*-sulfo decasaccharide 5** carrying two GlcNS residues, 2110.5 Da. The major product in Pool B displayed *m/z* values of 536.5 and 715.5, representing quadruply and triply charged ions

that correspond very closely to the molecular weight of ***N*-sulfo decasaccharide 6** carrying three GlcNS residues, 2148.4 Da. PAMN-HPLC analysis shows that the major product, ***N*-sulfo decasaccharide 6**, is formed with 88% yield. ***N*-sulfo decasaccharide 5** maintains 12% yield and there is no starting material remaining. The samples were then analyzed by MS/MS to determine the location of the newly modified GlcNS residues in both products (Figure 64). Analysis of ***N*-sulfo decasaccharide 5** provided two characteristic fragmented ions C_5^{2-} and Y_3, at the *m/z* values of 507.6 and 516.9, providing the structural evidence that the newly modified GlcNS residue is located at monosaccharide position 4 (Figure 64A). Analysis of the major product, ***N*-sulfo decasaccharide 6**, yielded two characteristic fragmented ions B_8^{3-} and Z_{10}^{3-}, at the *m/z* values of 542.7 and 642.9, providing the structural evidence that the newly modified GlcNS residues are located at monosaccharide positions 4 and 8 within the ***N*-sulfo decasaccharide 6** product (Figure 64B).

Interestingly, many factors seeming to have an enhancing effect on the NDST-2 modification of structurally defined oligosaccharides were combined in this substrate to produce an ideal substrate. From previous oligosaccharides it was noted that the addition of GlcNAc at the nonreducing end drives the reaction to form a single major product. It was also seen in the case of **Nona-2** that the pre-existing *N*-sulfation completely changed the product yield profile and produced a majority of one single major product. So, as we would expect, **Deca-2** was able to produce nearly 90% of a single product with no starting material remaining.

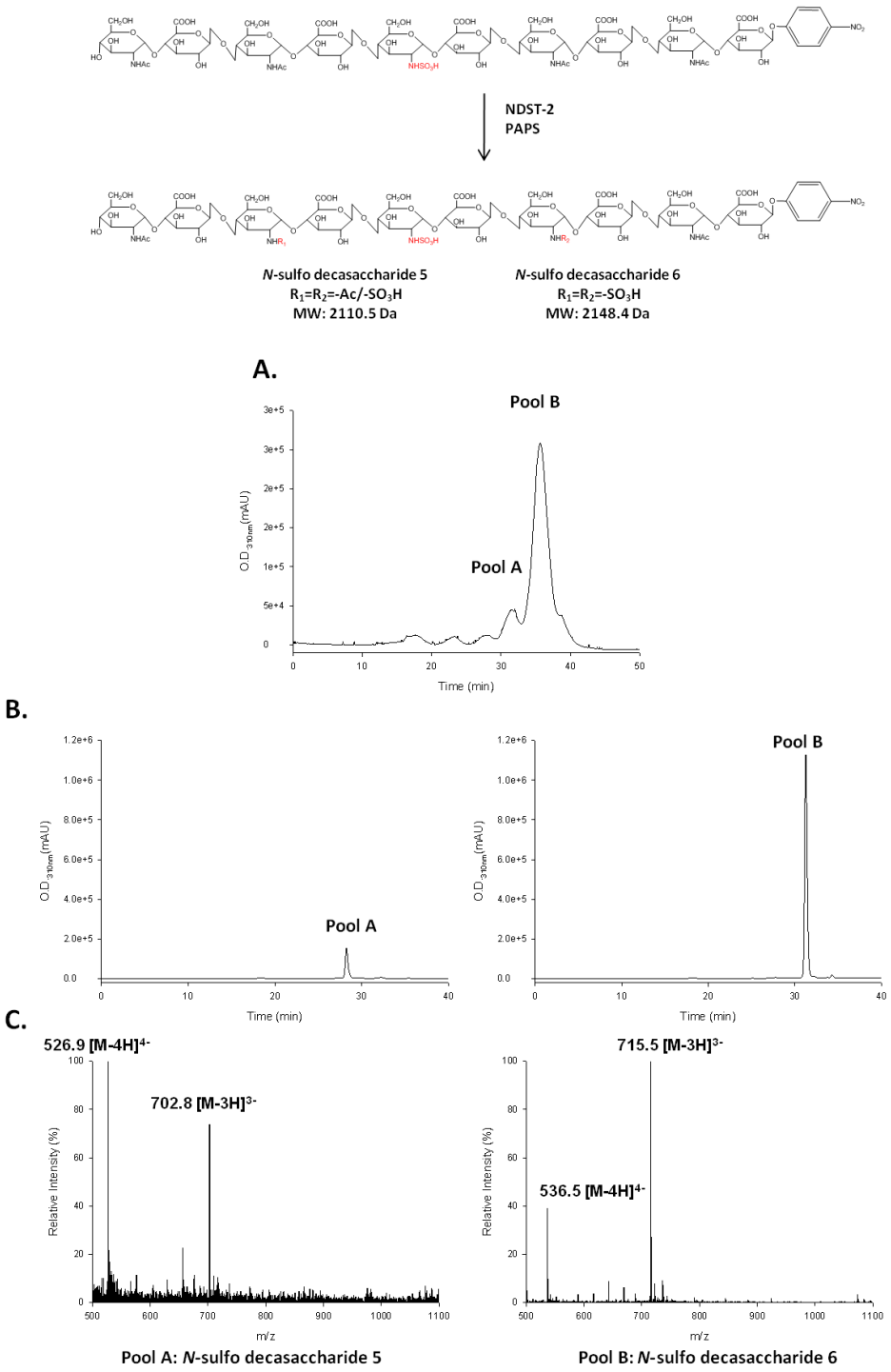

Figure 63. Deca-2 + NDST-2 reaction analysis. (A) Q sepharose-HPLC purification reveals two peaks eluted at 30 and 33 minutes respectively. (B) PAMN-HPLC confirms the presence of a single peak from Pool A eluted at 28 minutes and a single peak from Pool B eluted at 31 minutes. (C) ESI-MS confirmation of a product in Pool A containing two GlcNS residues with an *m/z* value of 702.8, corresponding to a triply charged ion with the expected molecular weight of 2110.5 Da; confirmation of a product in Pool B containing three GlcNS residues with an *m/z* value of 715.5, corresponding to a triply charged ion with the expected molecular weight of 2148.4 Da.

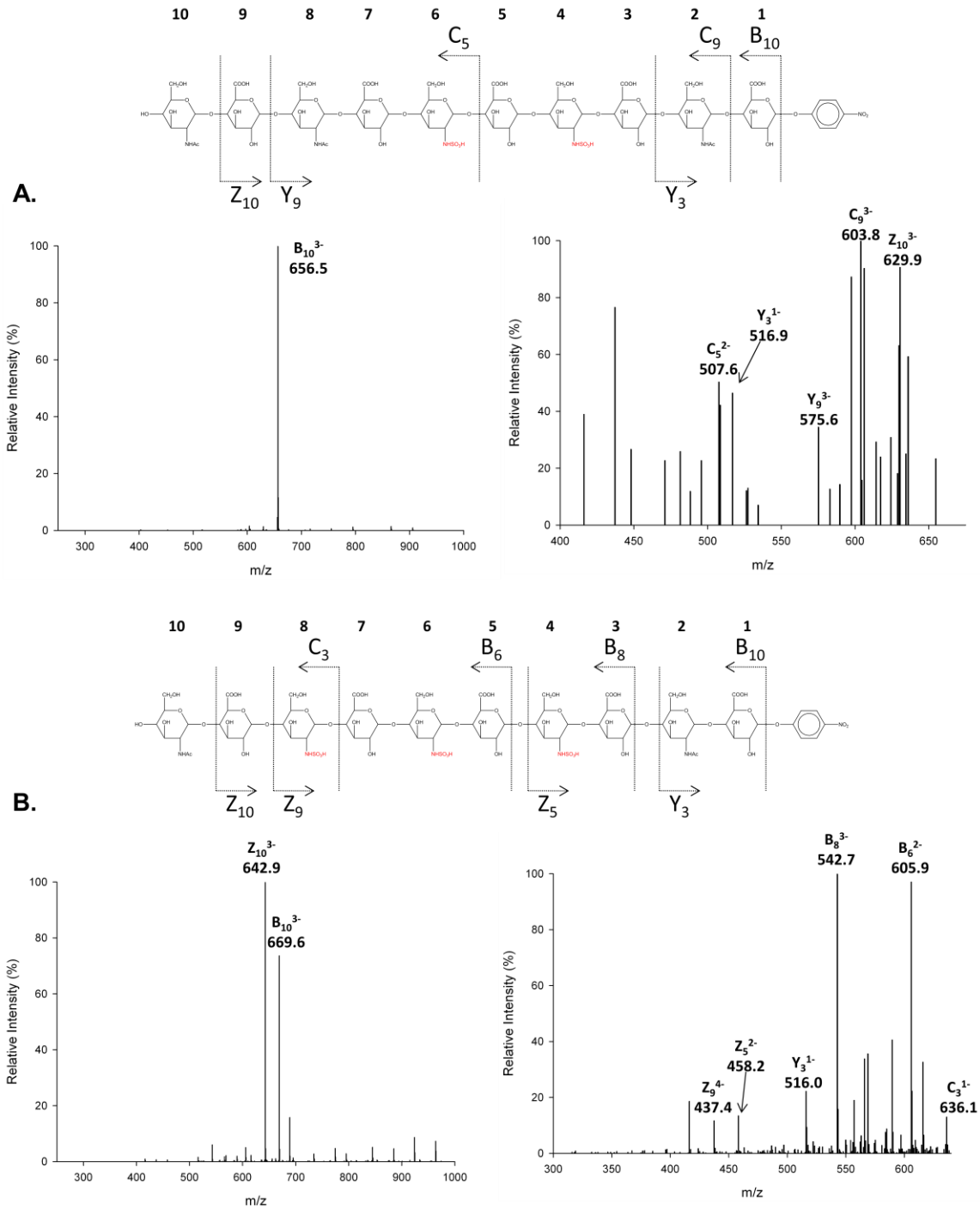

Figure 64. ***N*-sulfo decasaccharide MS/MS analysis.** (A) *N*-sulfo decasaccharide 5; Precursor ion selection was at $[M-3H]^{3-}$, *m/z* 702.8. Analysis shows a new GlcNS residue at monosaccharide position 4. (B) *N*-sulfo decasaccharide 6; Precursor ion selection was at $[M-3H]^{3-}$, *m/z* 715.5. Analysis shows two new GlcNS residues at monosaccharide positions 4 and 8.

Conclusions

The biosynthesis of HS and HP involves an array of specialized enzymes, including glycosyltransferases, an epimerase, and various sulfotransferases. *N*-deacetylation and subsequent *N*-sulfation is the most important sulfotransferase modification, as nearly every other sulfotransferase in the biosynthetic pathway relies on *N*-sulfated glucosamine for substrate recognition and modification. This is demonstrated by the fact that sulfate rich regions in HS structure that maintain a high density of negative charge, termed NS domains, are always characterized by long stretches of GlcNS residues mostly flanked by IdoUA2S (17). Both C_5-epimerase and 2-*O*-sulfotransferase recognize GlcNS residues and perform their modifications at the adjacent uronic acid (96, 102). The NS domains are the essential element of HS in cellular regulation and biological function (17). *In vivo* studies have implicated NDST-1 as the most important isozyme for the biosynthesis of HS. Mice lacking this enzyme underwent neonatal respiratory distress and embryonic death (82). These animals were characterized as having abnormal HS. HS is also characterized by regions of low sulfation density, known as NAc domains (17). Sheng et al. was able to demonstrate the ability for NDST-1 to recognize a pre-existing GlcNS residue, causing the enzyme to leave a region of unmodified GlcNAc residues before reinitiation of sulfation at a GlcNAc site that is at least four monosaccharide units toward the nonreducing end (26). This results in the formation of a NAc domain. In contrast, NDST-2 has been implicated *in vivo* to be responsible for HP biosynthesis. Mice lacking the NDST-2 enzyme developed normally, with the notable exception of abnormal connective tissue type mast

cells (134). HP resides in mast cells and is responsible for the sequestration of various proteases and histamines that are released at the site of an injury (87).

HS has a wide array of biological functions through many routes of cellular signaling, adhesion, and protein binding. Also, given the fact that the biosynthesis of HS is non-template driven, it takes a high degree of transcriptional and translational regulation and cooperation among the various biosynthetic enzymes in order to create the very unique sequences required to perform specific functions at targeted tissues throughout the body (10). This is why HS derived from the same tissue has a very similar sulfation pattern, in addition to expression of the same very specific isoforms of each biosynthetic enzyme. While HS derived from different tissues will vary dramatically, in addition to the expression of enzymatic isoforms. On the other hand, HP is only synthesized in a single cell type, and performs a very specific function. This results in the synthesis of a much more homogeneous HP structure. In fact, 70-90% of HP structure is composed of the single trisulfated disaccharide structure, IdoUA2S-GlcNS6S (21). This essentially makes HP one long NS domain.

NDST-1 has shown the ability to mostly form a distribution of products in structurally defined oligosaccharides, without any tendency to form a single product in high yield (26). This was to be expected because NDST-1 prefers to form a highly heterogeneous structure as is observed in HS. That said, we would expect NDST-2 to have the ability to form a single nearly homogeneous product as observed in HP structure. While this was not always the case, we made some very valuable observations that lead us to believe this enzyme is capable of doing just that. We observed that when the oligosaccharide substrate maintains a GlcNAc residue at the nonreducing end, the

reaction seemingly goes to completion and all of the starting material is consumed, resulting in one single fully *N*-sulfated major product. We also observed that when a single GlcNS residue is placed within the structure of the oligosaccharide substrate the reaction also goes to completion and all of the starting material becomes consumed. This data is consistent with an observation made over two decades ago, when the Lindahl group observed that *N*-sulfation seemed to enhance further modification by NDST in mice mastocytoma cells (89, 90). We now know that this enzyme was in fact NDST-2 and have confirmed this observation *in vitro*.

We also have observed a similarity between the two isoforms, NDST-1 and NDST-2. The first common thread is that the oligosaccharide substrate must be of a certain length before modification will occur. In the case of NDST-1, modification will only occur on the fifth monosaccharide unit from the reducing end, or the second GlcNAc residue (26). The first GlcNAc residue from the reducing end present in the oligosaccharide is never modified in any of the samples. This same effect holds true with NDST-2, as none of the substrates possessed modification at the monosaccharide position 2, or the first GlcNAc from the reducing end. Instead, the 4 position was always modified in every substrate that reacted with NDST-2. In the NDST-1 study octasaccharide was deemed to be the smallest substrate capable of modification. However, it was stated that the hexasaccharide did create a small radiolabeled peak after NDST-1 incubation, leading us to believe that this substrate did actually react, just not at an acceptable threshold (26). The structure of the oligosaccharide substrates utilized in the NDST-1 study were characterized by a 2,5-anhydromannitol in position 1 at the reducing end, which was linked to GlcUA. Thus, the structure of the hexasaccharide was

homologous to the structure of **Penta-1** in our NDST-2 study. Indeed, we observed a very small product after incubation of **Penta-1** with NDST-2. Perhaps this product was a bit more prominent than the hexasaccharide product in the NDST-1 study, but it appears as if both enzymes are able to begin modification at the same oligosaccharide length.

These data make clear that NDST-2 acts in many ways that are different than its counterpart NDST-1. But the observations that were different are easily explainable and expected based on the structures of the respective glycosaminoglycans that each enzyme is responsible for synthesizing. NDST-2 is clearly capable of synthesizing stretches of GlcNS residues and does not form NAc domains when presented with a pre-existing GlcNS, as is the case for NDST-1. This difference is key to understanding the role of each enzyme. The next step in uncovering the role of NDST-2 in HP biosynthesis is to ask the question: What type of interplay exists between NDST-2 and the other key enzymes responsible for creating the IdoUA2S-GlcNS6S trisulfated disaccharide that is so prevalent in HP structure? This question will be addressed in the following section.

Table 3. Summary of the structures of NDST-2 modified products.

Substrates	Structures of the major products	Major products	Measured MW (Da)	Calculated MW (Da)
Tetra-1	NONE	No Product	N/A	N/A
Penta-1	**GlcUA-GlcNS**-GlcUA-GlcNAc-GlcUA-PNP	*N*-sulfo pentasaccharide	1111.8±0.1	1111.9
Hexa-1	GlcNAc-GlcUA-**GlcNS**-GlcUA-GlcNAc-GlcUA-PNP	*N*-sulfo hexasaccharide	1314.6±0.3	1315.1
Hepta-1	**GlcUA-GlcNAc-GlcUA-GlcNS**-GlcUA-GlcNAc-GlcUA-PNP **GlcUA-GlcNS**-GlcUA-**GlcNS**-GlcUA-GlcNAc-GlcUA-PNP	*N*-sulfo heptasaccharide 1 *N*-sulfo heptasaccharide 2	1491.3±0.3 1529.7±0.5	1491.2 1529.2
Octa-1	GlcNAc-GlcUA-**GlcNS**-GlcUA-**GlcNS**-GlcUA-GlcNAc-GlcUA-PNP	*N*-sulfo octasaccharide 2	1732.2±0.2	1732.4
Nona-1	**GlcUA-GlcNAc-GlcUA-GlcNAc-GlcUA-GlcNS**-GlcUA-GlcNAc-GlcUA-PNP **GlcUA-GlcNAc-GlcUA-GlcNS**-GlcUA-**GlcNS**-GlcUA-GlcNAc-GlcUA-PNP **GlcUA-GlcNS**-GlcUA-**GlcNS**-GlcUA-**GlcNS**-GlcUA-GlcNAc-GlcUA-PNP	*N*-sulfo nonasaccharide 1 *N*-sulfo nonasaccharide 2 *N*-sulfo nonasaccharide 3	1870.9±0.4 1908.3±0.1 1946.5±0.4	1870.5 1908.6 1946.6
Deca-1	GlcNAc-GlcUA-**GlcNS**-GlcUA-**GlcNS**-GlcUA-**GlcNS**-GlcUA-GlcNAc-GlcUA-PNP	*N*-sulfo decasaccharide 3	2149.8±0.3	2149.8
Nona-2	**GlcUA-GlcNS**-GlcUA-GlcNS-GlcUA-**GlcNS**-GlcUA-GlcNAc-GlcUA-PNP	*N*-sulfo nonasaccharide 5	1946.2±0.5	1946.6
Deca-2	GlcNAc-GlcUA-**GlcNS**-GlcUA-GlcNS-GlcUA-**GlcNS**-GlcUA-GlcNAc-GlcUA-PNP	*N*-sulfo decasaccharide 6	2149.8±0.3	2149.8

[a] The boldface font represents pre-existing *N*-sulfation within the substrate
[b] The red font represents modifications made by NDST2

Chapter IV

Establishing the Cooperative Roles of NDST-2, C_5-epimerase, and 2OST in Heparin Biosynthesis

Since its discovery nearly a century ago, heparin has been widely accepted for its anticoagulant properties. It has been used clinically since the 1930's to treat various thrombotic disorders. Much attention has been given over the years to delineate the exact structure of the polysaccharide and the mechanism by which it confers its anticoagulant effect. However, less attention has been given to the biosynthesis of HP and its biological role *in vivo*. Heparin is known to be synthesized and stored in connective tissue type mast cells, whereby its high negative charge allows it to sequester various proteases and histamines that are released from the mast cell at the site of an injury. Much like its structural analog heparan sulfate, the biosynthesis of HP has always been presented in a linear, stepwise fashion, composed of various enzymes involved in chain initiation, elongation, and modification. However, there are many examples of enzymatic interactions that may challenge this notion and suggest that the process occurs in a much more dynamic fashion. For example, the glycosyltransferase enzymes EXT1 and EXT2 are known to form a complex to achieve full enzymatic potential (65, 66). The initiation enzymes XylT and GalT have been observed as a physical complex, and the same is true for C_5-Epi and 2OST (48, 232). It has also been observed that EXT2 expression can alter the enzymatic activity of NDST-1 (69). With so many examples of enzymatic

interaction, it seems unlikely that biosynthesis occurs in a stepwise fashion. In other words, elongation and modification most likely occur simultaneously.

In 2002, Esko and Selleck introduced the idea of a GAGosome, where the biosynthetic enzymes are contained within specialized compartments inside the golgi apparatus. This close association of enzymes allows for rapid production of HS/HP molecules and helps to explain the interplay involved in the biosynthesis of the highly regulated enzymatic biosynthesis. While a GAGosome has never been observed, the most important aspect of this line of thinking is the dynamic way by which the HS biosynthetic enzymes interact with each other to build HS/HP glycosaminoglycans.

In context of the various enzymes that are responsible for the biosynthesis of HP versus HS, there are a few distinctions as to which isoforms are produced in mast cells. The linkage enzymes are most likely the same as epithelial cells, since there are single isoforms responsible for building the linkage region. For the same reasons, EXT1/EXT2 must also be present in the mast cells. A critical difference is in the expression of NDST-1 and NDST-2. NDST-2 is abundantly expressed in mast cells, while NDST-1 is barely detected (79). The necessity for NDST-2 for HP production in mast cells was also demonstrated in a mouse model where NDST-2 was silenced, resulting in a phenotype of irregular mast cell production and reduction in mast cell proteases and histamines (87). There are only single isoforms of both 2OST and C_5-Epi in the human genome, so they must also be present. 6-OST-2/3 have been shown to be the most important for providing 6-*O*-sulfation in a mouse model (131). Finally, 3OST-1 has been identified in a mouse mastocytoma cell line, but other isoforms have not been studied (111). This specific

combination of HS biosynthetic enzymes is a critical point to begin to understand what sets HP apart from HS.

As stated previously, HP is essentially one long NS domain that is present in HS structure. It is composed of 70-90% of the repeating trisulfated disaccharide unit IdoUA2S-GlcNS6S. There is very little variability in the structure. This fact allows HP to maintain an extremely high density of sulfation and negative charge. In addition to the charge, the flexibility of the IdoUA residues allow this structure to conform for the binding of a wide array of proteins. These factors led us to the theory that there must be some interplay between NDST-2, 2OST, and C_5-Epi in order for HP biosynthesis to occur.

In this chapter, we describe a straightforward one pot reaction using the enzymes NDST-2, C_5-epimerase, and 2OST in order to demonstrate a benchtop method to prove our theory of how HP is synthesized *in vivo*. Three key observations from the previous section made us believe this was possible. The first was the observation that NDST-2 always formed one single major product when a GlcNAc residue was at the nonreducing end of our structurally defined oligosaccharide. The second observation was that NDST-2 preferred not to modify the GlcNAc residue at the nonreducing end, thus preferring to act on GlcNAc residues that were flanked by GlcUA at the nonreducing end. The final observation was that a pre-existing GlcNS residue on the oligosaccharide substrate did not impede the action of NDST-2 or form any NAc domains as seen with NDST-1. Instead, the GlcNS residue seemed to enhance the ability of NDST-1 to form a single major product. When we considered these facts, combined with a thorough

understanding of C_5-Epi mode of irreversibility (discussed below), a viable alternative to the stepwise layout of HS/HP biosynthesis came to light.

Understanding the interdependency of NDST, C_5-Epi, and 2OST

In the context of substrate specificities, it is quite obvious that *N*-sulfation, epimerization, and 2-*O*-sulfation are all entwined together. First of all, we know that *N*-sulfation is a prerequisite for both enzyme specificities. C_5-Epi will only modify a GlcUA if it is adjacent to a GlcNS residue at the reducing end. In fact, IdoUA has only been found adjacent to GlcNS residue in naturally derived HS (99). Likewise, 2OST will only modify a uronic acid residue if it is located adjacent to a GlcNS residue at the nonreducing end. It has also been reported that epimerization will not take place if 2-*O*-sulfation has already occurred at the uronic acid residue, meaning that the GlcUA must be epimerized and followed by sulfation (95-97). On the other hand, epimerization will only occur if 2OST is present. A study in CHO cells showed that these enzymes form a complex before moving into the golgi apparatus to perform their functions (48). Without 2OST, there was no epimerization in mutant cells. In addition, 2OST has the capability to sulfate both GlcUA and IdoUA, however it greatly prefers IdoUA. To take the idea full circle, recent evidence has shown that C_5-Epi relies on a code of *N*-sulfation in order to impede the reversibility of the enzyme, thus creating an irreversible IdoUA site. This discovery will be further discussed in the next section. So it is clear that these enzymes each rely on one another to perform their functions and create a biologically functional HP molecule.

Examining the Influence of Glucosamine Sulfation State on the Reversibility of C_5epi

C_5-epimerase has long been understood to catalyze a reversible enzymatic reaction. In the case of C_5-Epi, this means that the enzyme has the ability to convert a GlcUA to IdoUA by epimerization at the C_5-position, but also has the ability to convert back to GlcUA. This has been a hurdle for scientists to overcome for quite some time in regards to fully characterizing this enzyme. So the key question becomes, if this enzyme catalyzes a reversible reaction, how are long stretches of IdoUA synthesized for NS domains in HS or the whole of HP structure? The answer seems to come from a "code" of *N*-sulfation, or lack thereof, which directs C_5-Epi into an irreversible state. Sheng et al. published data which shows the sulfation state of the glucosamine residue located three monosaccharide units toward the nonreducing end (-3 site) of the GlcUA to be modified determines whether this site is reversible or not (Figure 65) (100). The glucosamine responsible for the determination is termed the mode of reaction recognition site (MRRS), while the GlcUA to be modified is known as the epimerization product site (EPS). When the MRRS site is *N*-sulfated or *N*-unsubstituted, the EPS site remains reversible, converting back and forth between GlcUA and IdoUA. But when the MRRS site remains as a GlcNAc residue, the EPS site is locked into an IdoUA conformation and subsequently undergoes 2-*O*-sulfation.

With the knowledge of this data in mind, we began thinking about our results from the previous chapter. Each structurally defined oligosaccharide that maintained a GlcNAc residue at the nonreducing end not only was modified to near completion, forming a single major product, but also mostly left the GlcNAc residue unmodified. At this point we wanted to ask the question, is this observation directly related to the

irreversible mode of epimerization? Using the major product from the **Octa-1**reaction with NDST-2, we present a hypothetical example below.

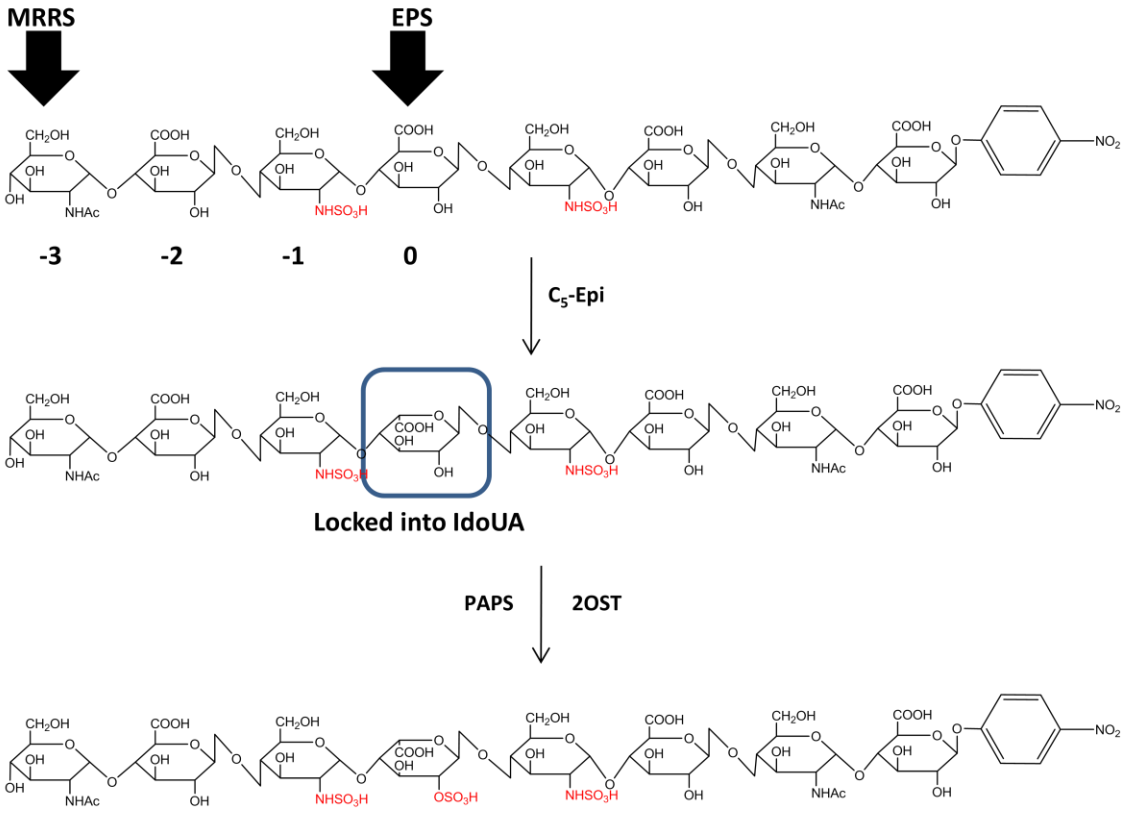

Figure 65. Demonstration of C_5-epimerase irreversibility with *N*-sulfo octasaccharide 2. In this example the epimerization site is labeled EPS, while the mode of reaction recognition site at -3 is labeled MRRS. When the MRRS is a GlcNAc residue the EPS site becomes locked into IdoUA conformation. This conformation is greatly favored by 2OST and will become subsequently sulfated (26).

Macromolecular Complex Theory of Heparin Biosynthesis

With the newfound mechanism for irreversible IdoUA synthesis, we have developed a model by which elongation and modification occur in a cooperative, dynamic fashion in order to design a HP-like IdoUA2S-GlcNS domain through a multi-step cycle (Figure 66). The cycle begins with a growing polysaccharide chain containing 2SIdoUA residues, such as the octasaccharide product in Figure 65 above. In the first step of the cycle, this oligosaccharide would be elongated by the glycosyltransferase

enzymes EXT1/EXT2 to add an additional disaccharide unit. Next, NDST-2 would act on the internal GlcNAc residue, converting it to GlcNS. This conversion to GlcNS creates an optimal site for epimerization and 2-*O*-sulfation of the GlcUA residue. Once epimerization and 2-*O*-sulfation takes place, the oligosaccharide is elongated by another disaccharide unit and the cycle repeats itself. Based on the irreversibility mechanism of C_5-Epi and the substrate specificities of both C_5-Epi and 2OST, this model seems to be a feasible scenario by which the IdoUA2S-GlcNS domain could be formed.

Based on the data presented in chapter one, we know that NDST-2 is capable of modifying the internal GlcNAc residues, leaving the GlcNAc on the nonreducing end unaffected. We also know that pre-existing *N*-sulfation does not disturb the enzymatic activity of NDST-2. With these observations in hand, we set off to develop an *in vitro* method for testing our proposed model.

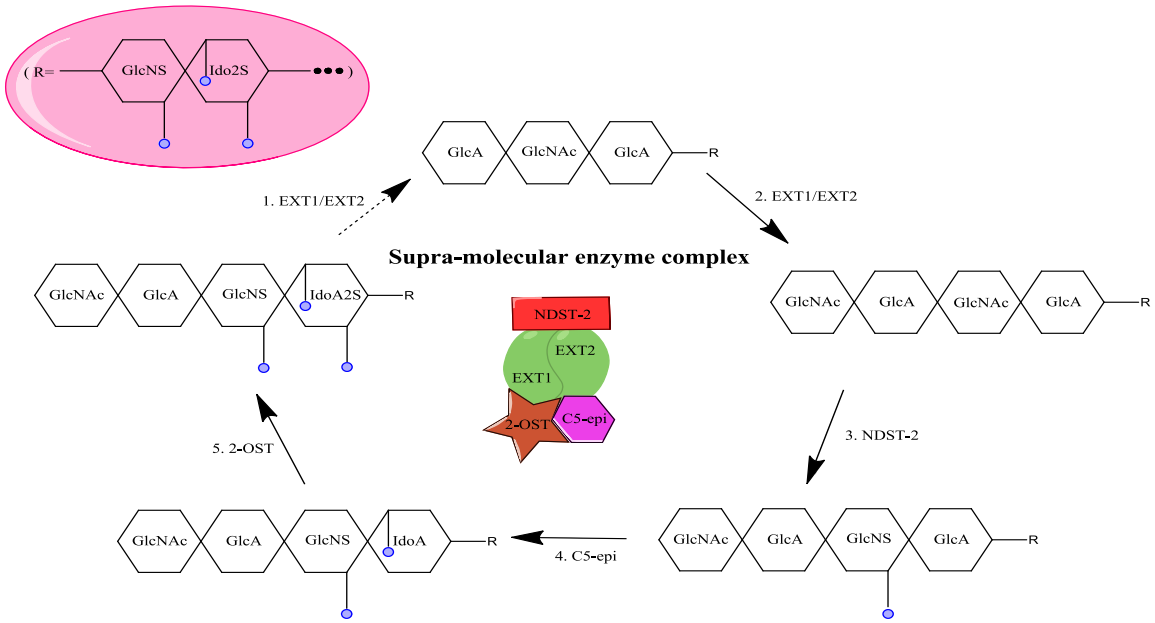

Figure 66. Proposed model for synthesis of IdoUA2S-GlcNS domain of heparin. A model of heparin biosynthesis by which polysaccharide chain elongation and modification occur in a cooperative manner simultaneously. The cycle begins with a growing oligosaccharide chain containing GlcNS-IdoUA2S. This chain is first elongated by EXT1/EXT2 by one disaccharide. NDST-2 modifies the resultant internal GlcNAc residue, leaving the GlcNAc residue at the nonreducing end in place. This results in an optimal site for epimerization and 2-*O*-sulfation. Once these reactions occur, elongation will take place once more and the cycle will repeat.

C_5-Epi/2OST Modification of Fully N-sulfated Octasaccharide

In order to provide an example of our reasoning behind the model of IdoUA2S-GlcNS domain biosynthesis, we incubated a fully *N*-sulfated octasaccharide containing only GlcUA with C_5-Epi and 2OST (Figure 67). The structure of this oligosaccharide was GlcNS-(GlcUA-GlcNS)$_3$-GlcUA-pnp (Figure 67A). This oligosaccharide should provide the evidence to show that biosynthesis of the IdoUA2S-GlcNS domain of HP does not occur in a stepwise fashion. We expected that this substrate would not maintain control of the C_5-epimerase reversibility mechanism due to a lack of GlcNAc residues, resulting in a complicated mixture of products. If biosynthesis did occur in a stepwise fashion, this substrate would represent the structure of the oligosaccharide post NDST-2 modification, readied for modification by C_5-Epi and 2OST. After incubation of our

fully *N*-sulfated substrate with C_5-Epi and 2OST, we analyzed the reaction by PAMN-HPLC and observed multiple peaks (Figure 67B). We observed a minor peak eluting at 38 minutes, followed by three major products eluting at 40 minutes, 43.5 minutes, and 48 minutes respectively. This result provides the proof that *N*-sulfation cannot occur first, followed by epimerization and 2-*O*sulfation. If that was the case, we would see a single product from this reaction. Instead, we see a mixture of products which cannot lead to the creation of a homogeneous repeating IdoUA2S-GlcNS domain structure.

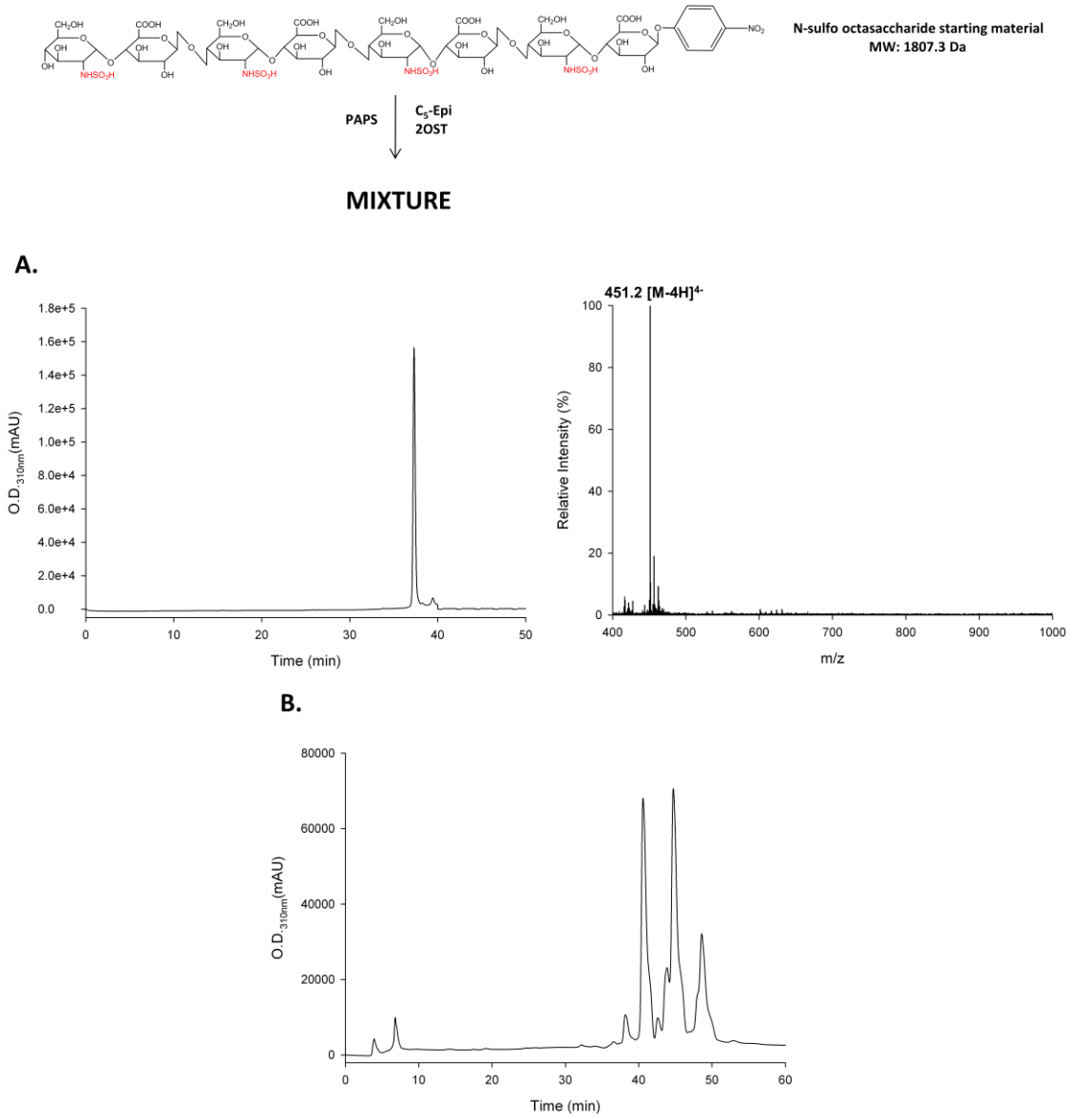

Figure 67. Analysis of fully *N*-sulfated octasaccharide + C_5-Epi/2OST. (A) PAMN-HPLC analysis of the fully *N*-sulfated octasaccharide starting material reveals 96% purity. ESI-MS confirms shows an *m/z* value of 451.2, corresponding to a quadruply charged ion very close to the expected molecular weight of 1807.3 Da. (B) PAMN-HPLC analysis after incubation of fully *N*-sulfated octasaccharide with C_5-Epi and 2OST shows a mixture of 3 major product peaks, and one minor peak. The minor peak is eluted at 38 minutes. The major peaks are eluted at 40, 43.5, and 48 minutes respectively.

Analyzing NDST-2 Modification of an Ideal Oligosaccharide Substrate Containing 2-*O*-Sulfated Iduronic Acid

At the start of our study on the cooperative roles of NDST-2, C_5-Epi, and 2OST, we first needed to make sure NDST-2 would actually modify a substrate containing an IdoUA2S residue. We did this by employing an ideal structure-defined octasaccharide substrate with the structure GlcNAc-GlcUA-GlcNAc-GlcUA-GlcNS-IdoUA2S-GlcNS-GlcUA-pnp (**I2S-octa**) (Figure 68A). This substrate would allow us to ensure there are only two possible GlcNAc sites for NDST-2 modification to take place, both very close in proximity to the IdoUA2S monosaccharide unit. After incubation of **I2S-octa** with NDST-2 we observed the reaction by PAMN-HPLC analysis (Figure 68B). We noted two peaks resulting from the reaction, the first peak eluted at 33 minutes, while the second peak eluted at 37 minutes. These peaks were then isolated by Q sepharose-HPLC and analyzed by ESI-MS (Figure 68C). Analysis of the first peak revealed a mixture of two products. The first product in the peak 1 mixture displayed an *m/z* value of 461.3, representing a quadruply charged ion that corresponds very closely to the molecular weight resulting from the addition of a single *N*-sulfation to our substrate, 1849.3 Da. The second product in the peak 1 mixture displayed an *m/z* value of 451.1, representing a quadruply charged ion that corresponds very closely to the molecular weight resulting from the addition of a single *N*-sulfation and deacetylation at the second site resulting in a free amine, 1807.3 Da. Analysis of the second peak revealed a single product which displayed an *m/z* value of 471.1, representing a quadruply charged ion that corresponds very closely to the molecular weight resulting from the addition of two *N*-sulfations.

This discovery was both pleasing and unexpected. First, we were pleased that NDST-2 was able to modify this substrate and seemingly consume the starting material. However, this was the first instance where we witnessed the GlcNAc at the nonreducing end being modified to such a significant extent. This result gave us pause, because if both GlcNAc sites are being modified, then the mechanism of irreversibility for C_5-Epi will not be available in the form of a nonreducing end GlcNAc residue. Still, we were able to witness NDST-2 modification of a substrate carrying a IdoUA2S residue so we decided to carry the experiment further.

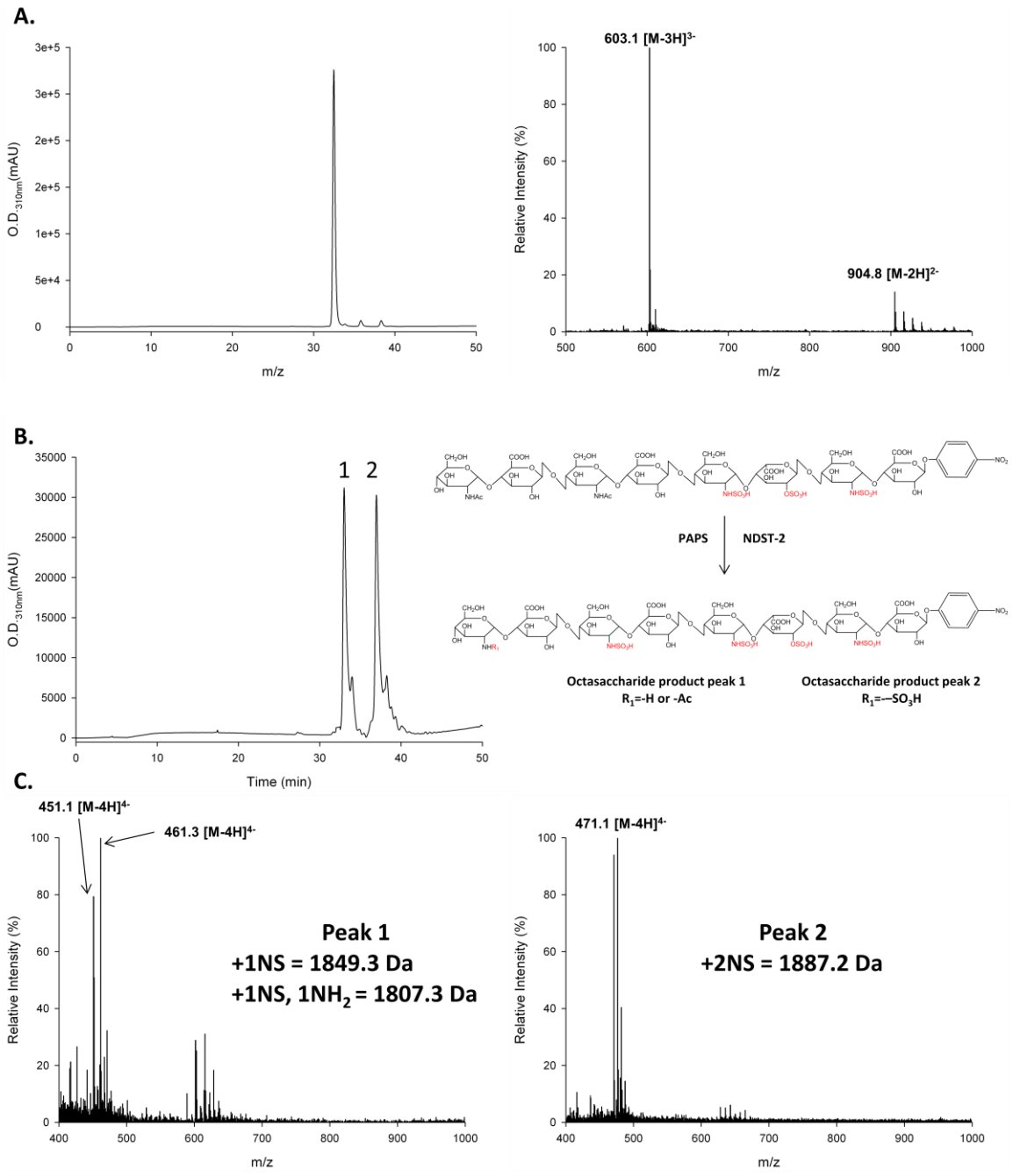

Figure 68. Analysis of NDST-2 modification of I2S-octa. (A) PAMN-HPLC analysis shows I2S-octa starting material as a single peak with 97% purity. ESI-MS analysis reveals an *m/z* value of 603.1 $[M-3H]^{3-}$, corresponding very closely to the expected molecular weight of 1811.3 Da. (B) PAMN-HPLC analysis of the reaction between I2S-octa and NDST-2 reveals two product peaks eluting at 33 and 37 minutes respectively. (C) ESI-MS analysis of the NDST-2 modified products reveals a mixture of two products in the first peak. An *m/z* value of 451.1 $[M-4H]^{4-}$ corresponds with the addition of a single *N*-sulfation, 1849.3 Da. An *m/z* value of 461.3 $[M-4H]^{4-}$, corresponds with the addition of a single *N*-sulfation and deacetylation at the second GlcNAc site, 1807.3 Da. Analysis of the second peak revealed an *m/z* value of 471.1, corresponding with the addition of 2 *N*-sulfations, 1887.2 Da.

Enzymatic Synthesis of IdoUA2S-GlcNS Domain in Ideal Oligosaccharide Substrates Using a One Pot Approach with NDST-2, C_5-Epi, and 2OST

Based on the results in the previous section, we recognized that **I2S-octa** was a substrate for NDST-2. However, our results also demonstrated that NDST-2 had the ability to modify the GlcNAc residue at the nonreducing end over 50% of the time. Nonetheless, we decided to carry on with an attempt at creating an IdoUA2S-GlcNS domain by utilizing the **I2S-octa** oligosaccharide substrate in a one pot reaction approach with NDST-2, C_5-Epi, and 2OST. We did this with the understanding that if NDST-2 modified the nonreducing end GlcNAc residue, then we would observe a complicated mixture of products because the epimerization site would remain reversible and maintain a 50% chance of remaining as GlcUA or converting to IdoUA. In addition, modification of the nonreducing end GlcNAc would also create a second epimerization and 2-*O*-sulfation site, thus further complicating our product. Yet, we reserved the belief that if our model holds true, there should be interplay between the enzymes which will prevent the modification of the nonreducing end GlcNAc, all the while creating our desired IdoUA2S-GlcNS repeating domain structure.

After incubation of **I2S-octa** with NDST-2, C_5-Epi, and 2OST, we observed 2 peaks by PAMN-HPLC analysis (Figure 69A). The first peak was eluted at 32 minutes, while the second peak eluted at 40 minutes. This result confirmed that a reaction had occurred by the one pot approach. In addition, we only observed a single major product with sufficient separation from the suspected starting material to represent the addition of at least one *N*-sulfation and 2-*O*-sulfation. The reaction was then purified by Q sepharose-HPLC, whereby the fractions under each peak were pooled and isolated

individually. The samples from each peak were then analyzed by ESI-MS to confirm their structures (Figure 69B). The first peak displayed *m/z* values of 452.2 and 603.1, representing a quadruply and triply charged ion corresponding with the molecular weight of the starting material, 1811.3 Da (Figure 69B). The second peak resulting from the one pot reaction displayed *m/z* values of 481.6 and 642.3, corresponding very closely to the molecular weight of an octasaccharide product containing a new GlcNS residue in addition to a single 2-*O*-sulfation, 1929.2 Da (Figure 69B). These data confirm that we have indeed proven that *N*-sulfation, epimerization, and 2-*O*-sulfation have the ability to work simultaneously in the biosynthesis of the IdoUA2S-GlcNS domain (Figure 69C). In addition, we have not observed any products containing an additional GlcNS residue. We postulate that the reliance of C_5-Epi on the GlcNAc residue at the nonreducing end may have a regulatory effect on the ability of NDST-2 to modify this position. This could also help to explain why NDST-2 does not prefer to modify that position in the data set from chapter one, as it has evolved to only modify the internal GlcNAc residues in order to maintain an irreversible epimerization substrate that is capable of forming the homogeneous IdoUA2S-GlcNS repeating disaccharide.

As a negative control for this experiment, we incubated **I2S-octa** with C_5-Epi and 2OST only, excluding NDST-2. PAMN-HPLC analysis of this reaction revealed that, as expected, there was indeed no reaction and the chromatogram looks identical to the starting material alone (Figure 70).

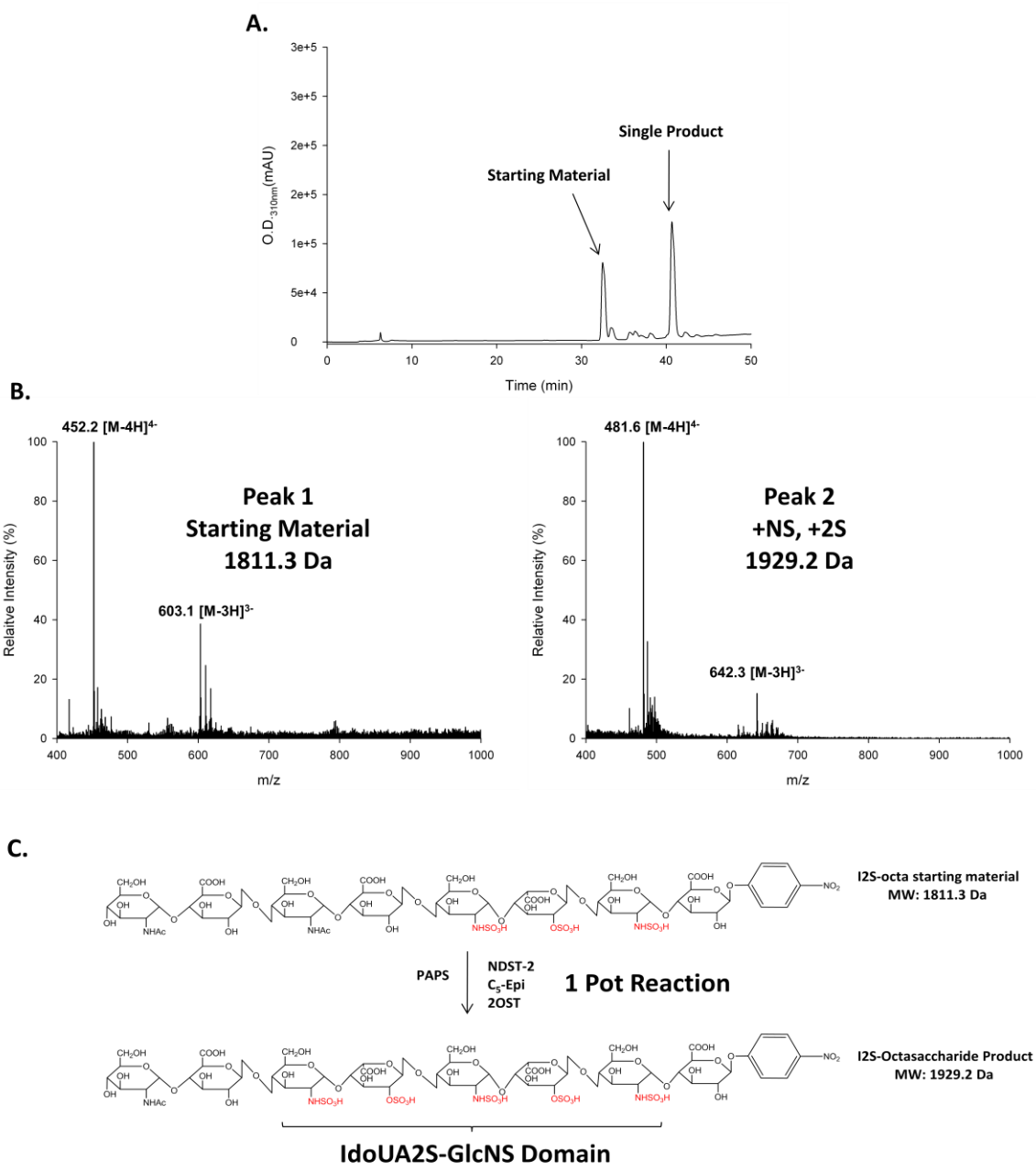

Figure 69. Analysis of I2S-octa + NDST-2/C_5-Epi/2OST one pot reaction. (A) PAMN-HPLC analysis reveals two peaks after incubation of **I2S-octa** with NDST-2, C_5-Epi, and 2OST. The first peak eluted at 32 minutes, while the second peak eluted at 40 minutes. (B) ESI-MS analysis of peak 1 reveals an *m/z* value of 452.2 $[M-4H]^{4-}$, corresponding to the molecular weight of the starting material, 1811.3 Da. ESI-MS analysis of peak 2 reveals an *m/z* value of 481.6 $[M-4H]^{4-}$, corresponding to the molecular weight of a product carrying an additional *N*-sulfation and 2-*O*-sulfation, 1929.2 Da. (C) Based on the PAMN-HPLC and ESI-MS data, combined with the well characterized substrate specificity of both C_5-Epi and 2OST, we have confirmed the structure of an IdoUA2S-GlcNS domain.

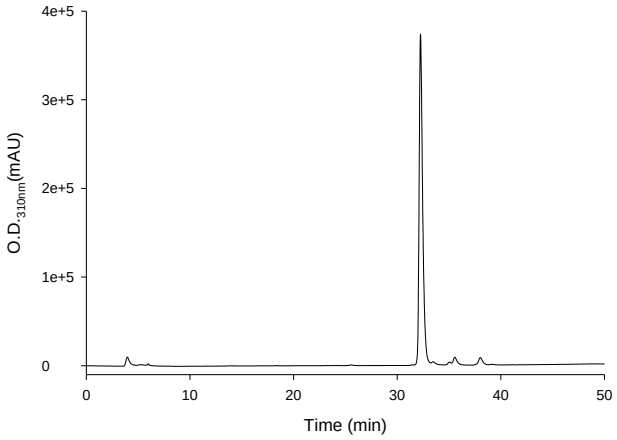

Figure 70. PAMN-HPLC analysis of I2S-octa + C_5-Epi/2OST negative control. Analysis confirms no reaction in the absence of NDST-2 from the reaction mixture.

In addition to demonstrating our one pot approach to building an IdoUA2S-GlcNS domain using NDST-2, C_5-Epi, and 2OST on a medium sized oligosaccharide substrate, we also sought out to demonstrate our method using a longer oligosaccharide substrate with the same basic elements. We designed a tetradecasaccharide containing a stretch of five GlcNS residues interceded by four IdoUA2S residues (**I2S-tetradeca)** (Figure 71C). This would allow us to prove that our method would continue to hold true independent of oligosaccharide/IdoUA2S-GlcNS chain length. **I2S-tetradeca** would represent the structure of our oligosaccharide after three full cycles of disaccharide elongation and subsequent modifications by NDST-2, C_5-Epi, and 2OST from our previous **I2S-octa** substrate.

After incubation of **I2S-tetradeca** with NDST-2, C_5-Epi, and 2OST, we observed one major peak by DEAE-HPLC analysis (Figure 71B). This peak was eluted at 32 minutes, approximately 2 minutes later than our starting material (Figure 71A), thus convincing us that we have achieved a single major product. After purification of the

reaction mixture by Q sepharose-HPLC, the major product peak was analyzed by ESI-MS (Figure 71B). ESI-MS analysis revealed *m/z* values of 488.0 $[M-7H]^{7-}$, 569.3 $[M-6H]^{6-}$, and 683.5 $[M-5H]^{5-}$, corresponding very closely to the molecular weight of a product resulting from the addition of one *N*-sulfation and one 2-*O*-sulfation, 3420.3 Da. These data show that the reaction behaves even better and runs to completion for a longer oligosaccharide. Thus, once the oligosaccharide reaches a critical length, the one pot reaction proceeds seamlessly to produce a homogeneous IdoUA2S-GlcNS domain in our oligosaccharide product.

Low ph NO_2 Degradation of I2S-octa product

In order to definitively prove that the uronic acid residue at the epimerization site in the **I2S-octa** reaction has indeed been converted from GlcUA to IdoUA2S, we employed low pH nitrous acid degradation. The reaction was repeated using [^{35}S]PAPS to place the radiolabel on the IdoUA2S-GlcNS disaccharide within the substrate after incubation with NDST-2, C_5-Epi, and 2OST. The radiolabeled product was then subjected to nitrous acid degradation at pH 1.5 to yield 2-*O*-sulfated disaccharide products followed by resolving by reverse phase ion pairing-HPLC. We observed a single radiolabeled peak eluted at 45 minutes that corresponds very closely to the elution time for the disaccharide standard, IdoUA2S-AnMan (Figure 72). This data confirmed that the uronic acid at the epimerization site has been fully converted to IdoUA followed by subsequent 2-*O*-sulfation.

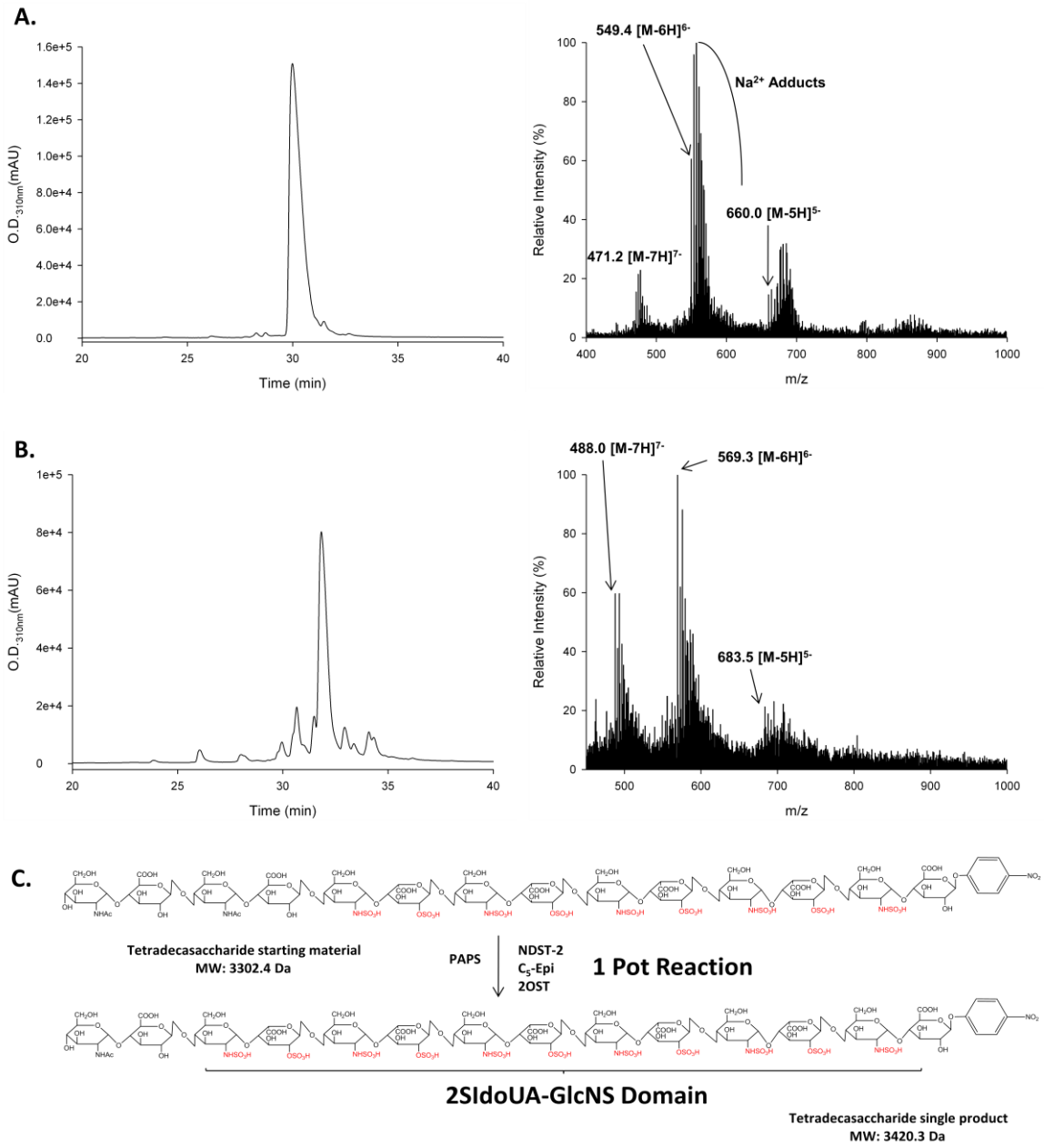

Figure 71. Analysis of I2S-tetradeca + NDST-2/C_5-Epi/2OST reaction. (A) DEAE-HPLC analysis reveals 98% purity for **I2S-tetradeca** substrate. ESI-MS analysis reveals an *m/z* value of 549.4 $[M-6H]^{6-}$, corresponding very closely to the expected molecular weight of the starting material, 3302.4 Da. (B) DEAE-HPLC analysis reveals a single major product peak eluted at 32 minutes. ESI-MS analysis reveals an *m/z* value of 569.3 $[M-6H]^{6-}$, corresponding very closely to the molecular weight of a product carrying an additional *N*-sulfation and 2-*O*-sulfation. (C) Taken together, these data reveal the one pot reaction has worked to add an additional GlcNS-IdoUA2S unit to the growing domain.

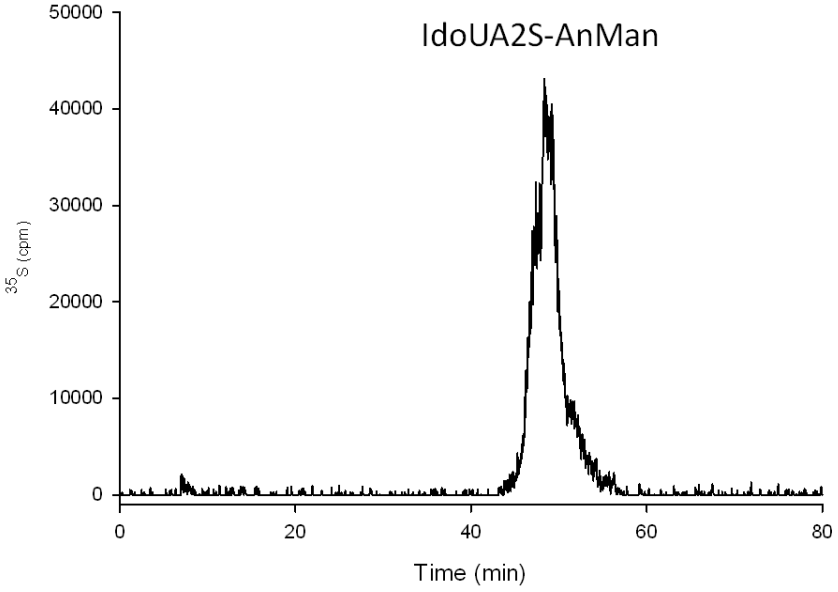

Figure 72. RPIP-HPLC disaccharide analysis profile of low pH NO_2 degraded I2S-octa product. I2S-octa was incubated with NDST-2, C_5-Epi, 2OST, and [^{35}S]PAPS. The product was purified by BioGel P10 size exclusion chromatography and subjected to nitrous acid degradation at pH 1.5 to yield the radiolabeled disaccharide products. The products were purified by BioGel P2 size exclusion chromatography and analyzed by RPIP-HPLC analysis. One peak was observed with an elution time of 45 minutes. This peak corresponded very closely to the elution time of the disaccharide standard, IdoUA2S-AnMan, confirming the complete conversion of the uronic acid at the epimerization site to IdoUA, followed by 2-*O*-sulfation.

Conclusions

In this chapter we have presented a model of HP biosynthesis based on construction of the IdoUA2S-GlcNS domain that characterizes the majority of HP structure. The model was based on a combination of substrate binding capabilities and modification patterns among the various biosynthetic enzymes involved. First, we recognized the enzymatic capabilities of NDST-2 based on data from the previous chapter. NDST-2 was observed to modify the internal GlcNAc residues of substrates, while mostly leaving the nonreducing end GlcNAc residue unmodified. Additionally, NDST-2 had high affinity for substrates maintaining a pre-existing GlcNS residue. Secondly, based on previous data from our group, we understood that C_5-Epi displayed a mode of irreversibility when presented with a GlcNAc residue located three monosaccharide units towards the reducing end, allowing it to lock the uronic acid into an IdoUA conformation. Finally, with the mode of NDST-2 action providing *N*-sulfation to all of the internal GlcNAc residues, an optimal binding site would be presented to 2OST based on its preference for a GlcNS residue at the adjacent monosaccharide to the nonreducing end of the uronic acid moiety to be modified. In addition, 2OST activity would be further enhanced by its preference for IdoUA, which would now be locked into this conformation by C_5-Epi.

We began by testing NDST-2 activity on a substrate containing an IdoUA2S unit, flanked by two GlcNS units (**I2S-octa**). We observed more activity than in previous studies as NDST-2 had modified the nonreducing end GlcNAc residue in over 50% of the starting material. Nonetheless, we had shown that NDST-2 was able to react with such a substrate. Next, we decided to combine NDST-2, C_5-Epi, and 2OST into a one pot

reaction with **I2S-octa**. We observed a single product resulting from the one pot reaction approach, with a molecular weight corresponding to the expected product carrying a new single *N*-sulfation and 2-*O*-sulfation. This revealed to us that not only is there cooperation between the enzymes, there also seems to be a mode of regulatory control. We did not observe a product consistent with the addition of *N*-sulfation at the nonreducing end GlcNAc residue, thus we believe that C_5-Epi may be occupying this position due to its mode of irreversibility. We then decided to attempt the same reaction using an **I2S-tetradeca** substrate containing four pre-existing IdoUA2S units and five GlcNS units. The reaction proceeded better than the octasaccharide reaction, as only one single major product was formed and the starting material was consumed. The product also had a molecular weight consistent with the addition of one *N*-sulfation and one 2-*O*-sulfation. This result showed that the one pot reaction approach using NDST-2, C_5-Epi, and 2OST was capable of forming the IdoUA2S-GlcNS domain *in vitro* using both shorter and longer oligosaccharides.

In the future, we would like to combine this method with the human glycosyltransferase enzymes EXT1/EXT2. With a controlled approach, we believe there is a great possibility that we could synthesize an entire IdoUA2S-GlcNS domain *in vitro*, including the elongation steps. In addition, it would be very interesting to perform a fluorescence experiment using the same three enzymes to understand if they actually form a macromolecular complex *in vivo*. And if the complex does form, we would like to co-crystallize the complex with an oligosaccharide substrate, such as **I2S-octa**, in order to gain a more thorough understanding of how they cooperate and regulate each other.

Chapter V

Conclusions

Heparan sulfate is a highly sulfated glycosaminoglycan molecule consisting of a repeating disaccharide unit of uronic acid, either glucuronic or iduronic, linked to glucosamine (12). HS is expressed abundantly and ubiquitously throughout the body and has a wide range of biological effects such as anticoagulation, cell growth and differentiation, and mediating inflammatory responses (3). In addition, HS plays a role in various pathological events such as cancer growth and metastasis and aiding in viral infection (3). Each of these events is the direct result of heparan sulfate binding to various physiological mediators, from chemokines/cytokines to lectins on cell membranes or binding directly to proteins (11). The binding of HS is a direct result of the specific sequence, provided by various sulfations and epimerization of the uronic acid, which is presented by the HS polysaccharide chain. The binding sites within HS structure that confer biological activity are maintained within highly sulfated regions of the polysaccharide, known as NS domains due to their abundance of GlcNS (16). The NS domains are flanked by regions of low sulfation density, known as NAc domains (17). The biosynthesis of HS is a non-template driven process, as HS sequence is determined by a strictly regulated interplay between HS biosynthetic enzymes. Many of the various enzymes are present in multiple isoforms and display unique substrate

binding capabilities and modification patterns (70). The result of the interplay between the various enzymes is a heterogeneous molecule maintaining very specific binding sequences within the whole of its structure. Many of the isoforms of biosynthetic enzymes are selectively expressed across various tissues and cell types. The result of this selective expression is HS polysaccharides derived from the same tissue having very similar structures and sequences, while HS from different tissues maintaining very different structures (10). This is a direct result of the biological function of HS at very specific physiological sites.

Heparin is a structural analog of HS, and maintains the same repeating disaccharide structure. HP, however, is much different than HS in many ways. While HS is located in nearly every tissue and cell type in the body, HP resides only in the intracellular vesicles of connective tissue type mast cells (1,9). It is responsible for the sequestration of various proteases and histamines that are released by the mast cell upon arrival at the site of injury (87,88). Thus, HP naturally plays a key role in immune response. HP is able to sequester the enzymes within the mast cell through its extremely high density of negative charge. HP maintains about 2.6 sulfo groups per disaccharide unit, while HS only maintains 0.8 (3). This is a direct result of the trisulfated disaccharide unit IdoUA2S-GlcNS6S which comprises between 70-90% of the entire HP structure (21). The repeating trisulfated disaccharide unit resembles the highly sulfated NS domains in HS structure. This results in a much more homogeneous molecule than its counterpart HS.

The focus of this dissertation was on the activity of *N*-deacetylase/*N*-sulfotransferase isoform 2. NDST-2 has been identified by mouse gene knockout studies

to be responsible for the biosynthesis of HP in mast cells. Mice deficient in NDST-2 displayed a deficiency and abnormalities in mast cell production (134). By contrast, mice deficient in NDST-1 underwent neonatal distress and embryonic death (82). The NDST-1 knockout mice were shown to have undersulfated, shorter HS fragments, thus implicating NDST-1 in the biosynthesis of HS. Although both mRNA constructs are expressed abundantly across most cell types, NDST-2 is under strict transcriptional and translational control in all cells except mast cells (135). In mast cells the exact opposite is true, as NDST-2 is expressed abundantly, while NDST-1 expression is barely detected (79). Thus, NDST-1 is the key player that provides *N*-sulfation in the context of HS, while NDST-2 is responsible for performing this action in mast cells to synthesize HP.

Previous research in our group revealed the role of NDST-1 in creating the domain structures that are prevalent in HS structure. Sheng et al. observed that when a pre-existing *N*-sulfation was present on the oligosaccharide substrate, NDST-1 would only reinitiate modification at a site that was at least four monosaccharide units away from the GlcNS site (26). This was a key discovery in the understanding of the role NDST-1 plays in the biosynthesis of HS. But an important point to this study was that the *N*-sulfation was already in place on the oligosaccharide substrate when approached by NDST-1, meaning that elongation of the polysaccharide must be occurring simultaneously in order to form the domain long enough for NDST-1 to reinitiate modification.

In contrast, HP does not possess NAc domains, instead consisting of mostly highly sulfated domains carrying the characteristic trisulfated disaccharide unit IdoUA2S-GlcNS6S (21). This highly sulfated HP domain has been observed up to 40

monosaccharide units in length, comprising the vast majority of the structure. So a key question we wanted to ask was whether or not pre-existing *N*-sulfation on the oligosaccharide substrate would affect the enzymatic modification of NDST-2 in such a way as was seen with NDST-1. In fact, quite the contrary was observed as the NDST-2 reaction was enhanced by the presence of *N*-sulfation and no domain structures were observed. This was quite an interesting and distinct observation from NDST-1, but based on the structure of HP not an unexpected one. There were also other points of contrast observed between the two isoforms. The smallest oligosaccharide to be modified by NDST-1 was an octasaccharide, while for NDST-2 was a pentasaccharide (26). In addition, our data shows that NDST-2 does not seem to modify in one specific direction, instead modifying each internal GlcNAc residue in no particular order. Finally, we noted that when a GlcNAc residue was located at the nonreducing end of the oligosaccharide substrate the NDST-2 reaction always formed a single major product and preferred to leave the nonreducing end GlcNAc residue unmodified. This was a very intriguing observation that got us thinking about the role of C_5-epimerase and 2OST in creating the characteristic IdoUA2S-GlcNS domain of HP.

The mechanism of C_5-epimerase in creating the IdoUA2S-GlcNS repeating disaccharide domain in HP has long been elusive to researchers due the reversibility of the reaction that this enzyme catalyzes. It is responsible for converting GlcUA to IdoUA through abstraction and subsequent addition of a proton at the C_5 position, but this reaction can also be reverted back to reform GlcUA. However, recent research from the Liu group has uncovered a mode of irreversibility of C_5-Epi (100). It was observed that when there is a GlcNAc residue three monosaccharide units from the epimerization site,

the IdoUA residue would lock into position and could not revert back to GlcUA. This was an important discovery for understanding how IdoUA2S-GlcNS domains are formed. In addition, this also added to the understanding that the biosynthesis of HS/HP is a dynamic cooperation between elongation enzymes and modification enzymes. This is due to the fact that there must always be a GlcNAc located three units towards the nonreducing end for the IdoUA to stay in conformation and form a single homogeneous domain. That means for long stretches of IdoUA you would need a GlcNAc residue to be constantly added to the growing domain, thus elongation has to take place alongside epimerization. To demonstrate that the process could not be a linear stepwise synthesis, we used a fully *N*-sulfated octasaccharide, representing the step after NDST modification, and treated it with C_5-Epi and 2OST. The result was exactly what we expected, as multiple peaks were observed, most likely due to the reversibility of C_5-Epi when no GlcNAc is available to deem the reaction irreversible.

All of these observations led to the development of a dynamic model of HP IdoUA2S-GlcNS domain biosynthesis. This model shows that elongation and modification occur simultaneously. Elongation must occur after each cycle to provide a fresh GlcNAc residue at the nonreducing end to allow for the C_5-Epi mode of irreversibility. Following elongation, NDST modifies the substrate, providing substrate binding sites for both C_5-Epi and 2OST. After these modifications take place the cycle is renewed with another disaccharide unit, repeating multiple rounds to form the IdoUA2S-GlcNS domain.

In context of HP biosynthesis, this brought us back to the key piece of the puzzle: would NDST-2 modify a substrate with a pre-existing IdoUA2S residue, in addition to

multiple GlcNS residues. The answer was yes, but not exactly in the way we had hoped. Using our **I2S-octa** substrate we observed *N*-sulfation at the nonreducing end GlcNAc in a large portion of our product. This initially was concerning to us because we knew that C_5-Epi relied on that nonreducing end GlcNAc residue to perform in an irreversible mode. However, we hoped that by combining these enzymes, we would achieve some level of cooperation and form a single homogeneous product resembling the IdoUA2S-GlcNS domain of HP.

Upon incubating the **I2S-octa** substrate with NDST-2, C_5-Epi, and 2OST, we indeed observed one single product. The molecular weight of the product was shown to be consistent with deacetylation at the *N*-position of a single glucosamine followed by *N*-sulfation, and with the addition of a single 2-*O*-sulfation. Low pH nitrous acid degradation also was used to prove that the GlcUA had now been converted to IdoUA2S. This result also showed that NDST-2 no longer modified the GlcNAc residue at the nonreducing end of the oligosaccharide substrate, as had been seen using NDST-2 alone. We believe that the presence of C_5-Epi and 2OST has a regulatory effect on NDST-2 and does not allow it to modify this position. But as we have seen in the first data set, NDST-2 already prefers not to modify this site, which is most likely an evolutionary effect of this mode of cooperation. Based on this result, we also sought to show that the one pot approach would also work with a longer substrate containing multiple IdoUA2S-GlcNS units, **I2S-tetradeca**. This would represent the substrate structure after several rounds of elongation and modification. Once again, the reaction showed one single product. But unlike the **I2S-octa** reaction, this reaction went to completion as there was no starting material remaining. This result showed us that the reaction is more complete as the

oligosaccharide grows longer. This most likely facilitates the rapid pace by which these polysaccharides are synthesized *in vivo*.

Taken together, these results have shown that NDST-2 is a key player in the synthesis of the HP structure, which is characterized by the repeating IdoUA2S-GlcNS disaccharide domain. In addition, we have shown that NDST-2, C_5-Epi, and 2OST can cooperate and regulate each other during the synthesis of a HP polysaccharide.

APPENDIX I

Curriculum Vitae

RYAN MATTHEW BULLIS

1910 White Dogwood Road, Apex NC 27502 • (336) 816-5252 • bullis@email.unc.edu

EDUCATION

University of North Carolina at Chapel Hill
Ph.D. in Pharmaceutical Sciences, October 2013
Concentration: Chemical Biology & Medicinal Chemistry
Dissertation: "Investigating the Role of *N*-deacetylase/*N*-sulfotransferase in Heparin Biosynthesis"
Advisor: Professor Jian Liu, Ph.D.

North Carolina State University
B.S. in Biological Sciences, cum laude, May 2005
GPA 3.415

RESEARCH EXPERIENCE

Ph.D. Dissertation Research Assistant
August 2007-October 2013
University of North Carolina at Chapel Hill, Division of Chemical Biology & Medicinal Chemistry
Mentor: Professor Jian Liu, Ph.D.

- Investigated the substrate specificity and modification patterns of *N*-deacetylase/*N*-sulfotransferase 2 using a library of structurally defined oligosaccharides.
- Designed an *in vitro* one pot reaction using *N*-deacetylase/*N*-sulfotransferase 2, C_5-epimerase, and 2-*O*-sulfotransferase to synthesize a characteristic IdoUA2S-GlcNS domain of heparin.

Research and Development Technician
September 2005-August 2007
Nomacorc, LLC.
Mentor: Eric Dunkelberg

- Tested various polymer materials for wine cork extraction force and torque.
- Designed experiments for sensory optimization of raw polymer materials.

TEACHING EXPERIENCE

Head Teaching Assistant, Spring 2008
University of North Carolina at Chapel Hill, School of Pharmacy
PHCY422, Biochemistry II for Pharm D candidates
- Coordinated the distribution of work for fellow TAs
- Proctored and graded exams, graded problem sets
- Lectured during recitation three times each semester
- Held office hours one hour per week

Teaching Assistant, Fall 2007, Fall 2008 & Spring 2009
University of North Carolina at Chapel Hill, School of Pharmacy
PHCY421, Biochemistry I for Pharm D candidates
MOPH864, Advances in Drug Delivery
MCNP 899, Seminar in Medicinal Chemistry & Natural Products
- Proctored and graded exams, graded problem sets
- Lectured during recitation three times each semester
- Held office hours one hour per week
- Managed online access to course resources

ACADEMIC HONORS

Native American Incentive Grant, May 2009-May 2013
Awarded by the University of North Carolina at Chapel Hill Graduate School to Native American Ph.D. candidates.

PUBLICATIONS

Bullis, RM, Xu, Y, and Liu, J. (2013) Control of heparin biosynthesis through cooperation among *N*-deacetylase/*N*-sulfotransferase 2, C_5-epimerase, and 2-*O*-sulfotransferase. *Manuscript in preparation.*

PRESENTATIONS

Bullis, RM, *Expression and scale-up of NDST-2 using the baculovirus expression system*. BRB startup meeting, Rensselaer Polytechnic Institute, Troy, NY, Oral presentation, October 6, 2009.

REFERENCES

1. Gandhi, N.S., Mancera, R.L. (2008) *Chem. Biol. Drug Des.***72**(6): 455-482.

2. Smith, G.E., Ju, G., Ericson, B.L., Moshera, J., Lahm, H., Chizzonite, R., and Summers, M.D. (1985) *Proc. Nat. Acad. Sci.***82**: 8404.

3. Peterson, S.B., Frick, A., Liu, J. (2009) *Nat. Prod. Rep.***26**: 610-627.

4. Bernfield, M., Gotte, M., Park, P.W., Reizes, O., Fitzgerald, M.L., Lincecum, J., and Zako, M. (1999). **68**: 729-777.

5. Iozzo, R.V. (1998) *Ann. Rev. Biochem.* **67**: 609-652.

6. Iozzo, R.V. (2001) *J. Clin. Invest.***108**(2): 165-167.

7. Liu, J., and Thorpe, S.C. (2002) *Med. Res. Rev.***22**(1): 1-25.

8. Park, P.W., Reizes, O, and M. Bernfield. (2000) *J. Biol. Chem.***275**: 29923-29926.

9. Abrink, M., Grujic, M., and Pejler, G. (2004) *J. Biol. Chem.***279**: 40897-40905.

10. Esko, J.D., Selleck, S.B. (2002) *Ann. Rev. Biochem.***71**(1): 435-471.

11. Kreuger, J., Kjellen, L. (2012) *J. Histochem. Cytochem.***60**(12): 898-907.

12. Kjellen, L., Lindahl, U. (1991) *Ann. Rev Biochem.***60**: 443-475.

13. Toida, T., Yoshida, H., Toyoda, H., Koshiishi, T., Imanari, T., Hileman, R.E., Fromm, J.R., Linhardt, R.J. (1997) *Biochem J.* **322**: 499-506.

14. Mulloy, B., and Forster, M.J. (2000) *Glycobiology.* **10**(11): 1147-1156.

15. Murphy, K.J., McLay, N., and Pye, D.A. (2008) *J. Am. Chem. Soc*. **130**(37): 12435-12444.

16. Kreuger, J., Spillmann, D., Li, J.-P., and Lindahl, U. (2006) *J. Cell Biol.***174**(3): 323-327.

17. Murphy, K.J., Merry, C.L.R., Lyon, M., Thompson, J.E., Roberts, I.S., and Gallagher, J.T. (2004) *J. Biol. Chem.***279**: 27239-27245.

18. Esko, J.D., and Lindahl, U. (2001) *J. Clin. Invest.***108**(2): 169-173.

19. Fareed J., H., D. A., Fareed, D., Demir, M., Wahi, R., Clarke, M., Adiguzel, C., and Bick, R. (2008) *Semin. Thromb. Hemostasis*. **34**(1): 58-73.

20. Qian, Y., Pan, J., Zhou, X., Weiser, P., Lu, H., Shih, F. F., Porche-Sorbet, R., Eby, C., and Zhang, L. (2010) *Clin. Appl. Thromb. Hemost.***16**(3): 251-260.

21. Hileman, R.E., Fromm, J.R., Weiler, J.M., and Linhardt, R.J. (1998) *Bioessays*. **20**(2): 156-167.

22. Rabenstein, D. (2002) *Nat. Prod. Rep.***19**: 312-331.

23. Lyon M., D.J.A., and Gallagher J.T. (1994) *J. Biol. Chem.***269**: 11208–11215.

24. Robinson H.C., H.A.A., Hook M., Ogren S., and Lindahl U. (1978) *J. Biol. Chem.***253**: 6687–6693.

25. Xu, Y., Masuko, S., Takkieddin, M., Xu, H., Liu, R., Jing, J., Mousa, S., Linhardt, R.J., and Liu, J. (2011) *Science* **334**(6055): 498-501.

26. Sheng, J., Liu, R., Xu, Y., and Liu, J. (2011) *J. Biol. Chem.***286**: 19768-19776.

27. Guo, Y.C., and Conrad, H.E. (1989) *Anal. Biochem.***176**: 96-104.

28. Koketsu, M., and Linhardt, R.J. (2000) *Anal. Biochem.***283**: 136-145.

29. Desai, U.R., Wang, H.M., and Linhardt, R.J. (1993) *Arch. Biochem. Biophys.***306**: 461-468.

30. Shworak, N.W., Shirakawa, M., Colliec-Jouault, S., Liu, J., Mulligan, R.C., Birinyi, L.K., and Rosenberg, R.D. (1994) *J. Biol. Chem.***269**: 24941–24952.

31. Liu, J., Shworak, N.W., Sinay, P., Schwartz, J.J., Zhang, L., Fritze, L.M.S., and Rosenberg, R.D. . (1999) *J. Biol. Chem.* . **274**: 5185–5192.

32. Shukla, D., Liu, J., Blaiklock, P., Shworak, N.W., Bai, X., Esko, J.D., Cohen, G.H., Eisenberg, R.J., Rosenberg, R.D., and Spear, P.G. (1999) *Cell*. **99**: 13-22.

33. Lindahl, U., Backstrom, G., Thunberg, L., and Leder, I.G. (1980) *Proc. Nat. Acad. Sci.***77**: 6551-6555.

34. Rosenberg, R.D., Oosta, G.M., Jordan, R.E., and Gardner, W.T. (1980) *Biochem. Biophys. Res. Commun.***96**: 1200-1208.

35. Shively, J.E., and Conrad, H.E. (1976) *Biochemistry*. **15**: 3932–3942.

36. Colliec-Jouault, S., Shworak, N.W., Liu, J., De Agostini, A.I., and Rosenberg, R.D. (1994) *J. Biol. Chem.***271**: 24953–24958.

37. Kinoshita, A., and Sugahara, K. . (1999) *Anal. Biochem.***269**: 367-378.

38. Kazuyuki, S., and Hiroshi, K. (2002) *IUBMB Life*. **54**(4): 163-175.

39. Ueno, M., Yamada, S., Zako, M., Bernfield, M., and Sugahara, K. (2001) *J. Biol. Chem.***276**: 29134–29140.

40. Schon, S., Prante, C., Bahr, C., Kuhn, J., Kleesiek, K., and Gotting, C. (2006) *J. Biol. Chem.***281**: 14224–14231.

41. Gotting, C., Kuhn, J., and Kleesiek, K. (2007) *Cell Mol. Life Sci.***64**: 1498-1517.

42. Esko, J.D., and Zhang, L. (1996) *Curr. Opin. Struct. Biol.***6**(5): 663-670.

43. Gotting, C., Kuhn, J., Zahn, R., Brinkmann, T., and Kleesiek, K. (2000) *J. Mol. Biol.***304**(4): 517-528.

44. Esko, J.D., Stewart, T.E., and Taylor, W.H. (1985) *Proc. Nat. Acad. Sci.***82**(10): 3197-3201.

45. Almeida, R., Amado, M., David, L., Levery, S.B., Holmes, E.H., Merkx, G., Geurtz van Kessel, A., et al. (1997) *J. Biol. Chem.***272**: 31979-31991.

46. Almeida, R., Amado, M., David, L., Levery, S.B., Holmes, E.H., Nomoto, M., Hollingsworth, M.A., et al. (1998) *J. Biol. Chem.***273**: 12770-12778.

47. Bai, X., Zhou, D., Brown, J.R., Crawford, B.E., Hennet, T., and Esko, J.D. (2001) *J. Biol. Chem.***276**: 48189–48195.

48. Pinhal, M.A., Smith, B., Olson, S., Aikawa, J., Kimata, K., and Esko, J.D. (2001) *Proc. Nat. Acad. Sci.***98**(23): 12984–12989.

49. Esko, J.D., Weinke, J.L., Taylor, W.H., Ekborg, G., Roden, L., Anantharamaiah, G., and Gawish, A. (1987) *J. Biol. Chem.***262**: 12189-12195.

50. Seiki, T., Oka, S., Tearyama, K., Imiya, K., and Kawasaki, T. (1999) *Biochem. Biophys. Res. Commun.***255**(1): 182-187.

51. Terayama, K., Oka, S., Seiki, T., Miki, Y., Nakamura, A., Kozutsumi, Y., Takio, K., et al. (1997) *Proc. Nat. Acad. Sci.***94**(12): 6093-6098.

52. Bai, X.M., Wei, G., Sinha, A., and Esko, J.D. (1999) *J. Biol. Chem.***274**: 13017-13024.

53. Gulberti, M., Lattard, V., Fondeur, M., Jacquinet, J.C., Mulliert, G., Netter, P., Magdalou, J., et al. (2005) *J. Biol. Chem.***280**(2): 1417-1425.

54. Sato, T., Gotoh, M., Kiyohara, K., Akashima, T., Iwasaki, H., Kameyama, A., Mochizuki, H., et al. (2003) *J. Biol. Chem.***278**: 3063-3071.

55. Seko, A., Dohmae, N., Takio, K., and Yamashita, K. (2003) *J. Biol. Chem.***278**: 9150-9158.

56. Jacquinet, J.C. (2004) *Carbohydrate Research*. **339**: 349-359.

57. Kim, B.T., Kitagawa, H., Tamura, J., Saito, T., Kusche-Gullberg, M., Lindahl, U., and Sugahara, K. (2001) *Proc. Nat. Acad. Sci.***98**: 7176-7181.

58. Kitagawa, H., Shimakawa, H., and Sugahara, K. (1999) *J. Biol. Chem.***274**: 13933-13937.

59. Zak, B.M., Crawford, B.E., and Esko, J.D. (2002) *Biochem. Biophys. Acta.***1573**: 346-355.

60. Chen, R.L., and Lander, A.D. (2001) *J. Biol. Chem.***276**: 7507-7517.

61. Dolan, M., Horchar, T., Rigatti, B., and Hassell, J.R. (1998) *J. Biol. Chem.***272**: 4316-4322.

62. Zhang, L., and Esko, J.D. (1994) *J. Biol. Chem.***269**: 19295-19299.

63. Zhang, L., David, G., and Esko, J.D. (1995) *J. Biol. Chem.***270**: 27127-27135.

64. Lin, X., Wei, G., Shi, Z., Dryer, L., Esko, J.D., Wells, D.E., and Matzuk M.M. (2000) *Dev. Biol.***24**: 299-311.

65. Senay, C., Lind, T., Muguruma, K., Tone, Y., Kitagawa, H., Sugahara, K., Lidholt, K., et al. (2000) *EMBO Reports*. **1**(3): 282-286.

66. McCormick, C., Duncan, G., Goutsos, K.T., and Tufaro, F. (2000) *Proc. Nat. Acad. Sci.***97**(2): 668-673.

67. Lidholt, K., Kjellen, L., and Lindahl, U. (1989) *Biochem. J.***261**: 999-1007.

68. Lidholt, K., and Lindahl, U. (1992) *Biochem. J.***287**: 21-29.

69. Presto, J., Thuveson, M., Carlsson, P., Busse, M., Wilen, M., Eriksson, I., Kusche-Gullberg, M., et al. (2008) *Proc. Nat. Acad. Sci.***105**: 4751-4756.

70. Feyerabend, T., Li, J., Lindahl, U., and Rodewald, H. (2006) *Nat. Chem. Biol.***2**: 195-196.

71. Holmborn, K., Ledin, J., Smeds, E., Eriksson, I., Kusche-Gullberg, M., and Kjellen, L. (2004) *J. Biol. Chem.***279**: 42355-42358.

72. Grobe, K., Ledin, J., Ringvall, M., Holmborn, K., Forsberg, E., Esko, J.D., and Kjellen, L. (2002) *Biochem. Biophys. Acta.***1573**: 209-215.

73. Aikawa, J., Grobe, K., Tsujimoto, M, and Esko, J.D. (2001) *J. Biol. Chem.***276**: 5876-5882.

74. Berninsone, P., and Hirschberg, C.B. (1998) *J. Biol. Chem.***273**: 25556-25559.

75. Duncan, M.B., Liu, M., Fox, C., and Liu, J. (2006) *Biochem. Biophys. Res. Commun.***339**: 1232-1237.

76. Kakuta, Y., Sueyoshi, T., Negishi, M., and Pedersen, L.C. (1999) *J. Biol. Chem.***274**: 10673-10676.

77. Wei, Z., and Swiedler, S.J. (1999) *J. Biol. Chem.***274**: 1966-1970.

78. Negishi, M., Pedersen, L.G., Petrotchenko, E., Shevtsov, S., Gorokhov, A., Kakuta, Y., and Pedersen, L.C. (2001) *Arch. Biochem. Biophys.***390**(2): 149-157.

79. Kusche-Gullberg, M., Eriksson, I., Pikas, D.S., and Kjellen, L. (1998) *J. Biol. Chem.***273**: 11902-11907.

80. Pallerla, S.R., Lawrence, R., Lewejohann, L., Pan, Y., Fischer, T., Schlomann, U., Zhang, X., et al. (2008) *J. Biol. Chem.***283**(24): 16885-16894.

81. Grobe, K., and Esko, J.D. (2002) *J. Biol. Chem.***277**: 30699-30706.

82. Fan, G., Xiao, L., Cheng, L., Wang, X., Sun, B., and Hu, G. (2000) *FEBS Letters*. **467**: 7-11.

83. Crawford, B.E., Garner, O.B., Bishop, J.R., Zhang, D.Y., Bush, K.T., Nigam, S.K., and Esko, J.D. (2010) *PLOS One*. **5**: 1-10.

84. Fuster, M.M., Wang, L., Castagnola, J., Sikora, L., Reddi, K., Lee, P.H., Radek, K., et al. (2007) *J. Cell. Biol.***177**: 539-549.

85. MacArthur, J.M., Bishop, J.R., Stanford, K.I., Wang, L., Bensadoun, A., Witztum, J.L., and Esko, J.D. (2007) *J. Clin. Invest.***117**: 153-164.

86. Wang, L., Fuster, M., Sriramarao, P., and Esko, J.D. (2005) *Nat. Immunol.***6**: 902-910.

87. Forsberg, E., Pejler, G., Ringvall, M., Lunderius, C., Tomasini-Johansson, B., Kusche-Gullberg, M.,Eriksson, I., et al. (1999) *Nature*. **400**: 773-776.

88. Humphries, D.E., Wong, G.W., Friend, D.S., Gurish, M.F., Qiu, W.T.,Huang, C., Sharpe, A.H., et al. (1999) *Nature*. **400**: 769-772.

89. Riesenfeld, J., Hook, M., and Lindahl, U. (1982) *J. Biol. Chem.***257**: 7050-7055.

90. Riesenfeld, J., Hook, M., and Lindahl, U. (1982) *J. Biol. Chem*. **257**: 421-425.

91. Li, J., Hagner-McWhirter, A., Kjellen, L., Palgi, J., Jalkanen, N., and Lindahl, U. (1997) *J. Biol. Chem.***272**: 28158-28163.

92. Hagner-McWhirter, A., Lindahl, U., and Li, J. (2000) *Biochem. J*.**347**: 69-75.

93. Prihar, H.S., Campbell, P., Feingold, D.S., Jacobsson, I., Jensen, J.W., Lindahl, U., and Roden, L. (1980) *Biochemistry*. **19**: 495-500.

94. Crawford, B.E., Olson, S.K., Esko, J.D., and Pinhal, M.A.S. (2001) *J. Biol. Chem.***276**(24): 21538-21543.

95. Backstrom, G., Hook, M., Lindahl, U., Feingold, D.S., Malmstrom, A., Roden, L., and Jacobsson, I. (1979) *J. Biol. Chem.***254**: 2975-2982.

96. Hagner-McWhirter, A., Hannesson, H.H., Campbell, P., Westly, J., Roden, L., Lindahl, U., and Li, J. (2000) *Glycobiology*. **10**(2): 159-171.

97. Jacobsson, I., Lindahl, U., Jensen, J.W., Roden, L., Prihar, H., and Feingold, D.S. (1984) *J. Biol. Chem.***259**: 1056-1063.

98. Lindahl, U., Jacobsson, I., Hook, M., Backstrom, G., and Feingold, D.S. (1976) *Biochem. Biophys. Res. Commun.***70**: 492-499.

99. Maccarana, M., Sakura, Y., Tawada, A., Yoshida, K., and Lindahl, U. (1996) *J. Biol. Chem.***271**: 17804-17810.

100. Sheng, J., Xu, Y., Dulaney, S.B., Huang, X., and Liu, J. (2012) *J. Biol. Chem.***287**(25): 20996-21002.

101. Rong, J., Habuchi, H., Kimata, K., Lindahl, U., and Kusche-Gullberg, M. (2001) *Biochemistry*. **40**(18): 5548-5555.

102. Kobayashi, M., Habuchi, H., Habuchi, O., Saito, M., and Kimata, K. (1996) *J. Biol. Chem.***271**(13): 7645-7653.

103. Bullock, S.L., Fletcher, J.M., Beddington, R.S.P., and Wilson, V.A. (1998) *Genes Dev.***12**(12): 1894-1906.

104. Kinnunen, T., Huang, Z., Townsend, J., Gatdula, M.M., Brown, J.R., Esko, J.D., and Turnbull, J.E. (2005) *Proc. Nat. Acad. Sci.***102**(5): 1507-1512.

105. Bai, X.M., and Esko, J.D. (1996) *J. Biol. Chem.***271**: 17711-17717.

106. Habuchi, H., Tanaka, M., Habuchi, O., Yoshida, K., Suzuki, H., Ban, K., and Kimata, K. (2000) *J. Biol. Chem.***275**(4): 2859-2868.

107. Brickman, Y.G., Ford, M.D., Gallagher, J.T., Nurcombe, V., Bartlett, P.F., and Turnbull, J.E. (1998) *J. Biol. Chem.***273**: 4350-4359.

108. Jemth, P., Smeds, E., Do, A.T., Habuchi, H., Kimata, K., Lindahl, U., and Kusche-Gullberg, M. (2003) *J. Biol. Chem.***278**(27): 24371-24376.

109. Smeds, E., Habuchi, H., Do, A.T., Hjertson, E., Grundberg, H., Kimata, K., Lindahl, U., et al. (2003) *Biochem. J.***372**(2): 371-380.

110. Zhang, L., Beeler, D.L., Lawrence, R., Lech, M., Liu, J., Davis, J.C., Shriver, Z. et al. (2001) *J. Biol. Chem.***276**(45): 42311-42321.

111. Shworak, N.W., Liu, J., Fritze, L.M., Schwartz, J.J., Zhang, L., Logeart, D., and Rosenberg, R.D. (1997) *J. Biol. Chem.***272**: 28008-28019.

112. Yabe, T., Shukla, D., Spear, P.G., Rosenberg, R.D., Seeberger, P.H., and Shworak, N.W. . (2001) *Biochem. J.***359**: 235-241.

113. Shworak, N.W., Liu, J.A., Petros, L.M., Zhang, L.J., Kobayashi, M., Copeland, N.G., Jenkins, N.A. et al. (1999) *J. Biol. Chem.* **274**: 5170-5184.

114. Xia, G., Chen, J., Tiwari, V., Ju, W., Li, J., Malmstrom, A., Shukla, D. et al. (2002) *J. Biol. Chem.***277**(40): 37912-37919.

115. Xu, D., Tiwari, V., Xia, G., Clement, C., Shukla, D., and Liu, J. (2005) *Biochem. J.***385**: 451-459.

116a. Edvettal, S.C., Lee, K.A., Negishi, M., Linhardt, R.J., Liu, J., and Pedersen, L.C. (2004) *J. Biol. Chem.***279**(24): 25789-25797.

116b. Moon, A.F., Xu, Y., Woody, S.M., Krahn, J.M., Linhardt, R.J., Liu, J., and Pedersen, L.C. (2012) *Proc. Nat. Acad. Sci.* **109**(14): 5265-5270.

117. Cadwaller, A.B., and Yost, H.J. (2006) *Dev. Dyn.***235**: 3423-3431.

118. Petitou, M., Casu, B., and Lindahl, U. (2003) *Biochimie*. **85**: 83-89.

119. Vanpouille, C., Deligney, A., Delehedde, M., Denys, A., Melchior, A., Lienard, X., Lyon, M. et al. (2007) *J. Biol. Chem.***282**: 24416-24429.

120. Zhang, L., Lawrence, R., Schwartz, J.J., Bai, X., Wei, G., Esko, J.D., and Rosenberg, R.D. (2001) *J. Biol. Chem.***276**: 28806-28813.

121. Shworak, N.W., HajMohammadi, S., de Agostini, A.I., and Rosenberg, R.D. (2002) *Glycoconjugate J.***19**(4): 355-361.

122. Liu, J.A., Shriver, Z., Blaiklock, P., Yoshida, K., Sasisekharan, R., and Rosenberg, R.D. (1999) *J. Biol. Chem.***274**: 38155-38162.

123. Borjigin, J., Deng, J., Sun, X., DeJesus, M., Liu, T., and Wang, M.M. (2003) *J. Biol. Chem.***278**(18): 16315-16319.

124. Kuberan, B., Lech, M., Bojigin, J., and Rosenberg, R.D. (2004) *J. Biol. Chem.***279**: 5053-5054.

125. Kamimura, K., Rhodes, J.M., Ueda, R., McNeely, M., Shukla, D., Kimata, K., Spear, P.G. et al. (2004) *J. Cell. Biol.***166**: 1069-1079.

126. Tiwari, V., O'Donnell, C.D., Myung-Jin, O., Valyi-Nagy, T., and Shukla, D. (2005) *Biochem. Biophys. Res. Commun.***338**(2): 930-937.

127. Xu, D., Moon, A.F., Song, D., Pedersen, L.C., and Liu, J. (2008) *Nat. Chem. Biol.***4**(3): 200-202.

128. Galagher, J.T. (2001) *J. Clin. Invest.***108**: 357-361.

129. Ledin, J., Staatz, W., Li, J.P., Gotte, M., Selleck, S., Kjellen, L., and Spillman, D. (2004) *J. Biol. Chem.***279**(41): 42732-42741.

130. Lindahl, B., Eriksson, L., and Lindahl, U. (1995) *Biochem. J.***306**(1): 177-184.

131. Habuchi, H., Nagai, N., Sugaya, N., Atsumi, F., Stevens, R.L., and Kimata, K. (2007) *J. Biol. Chem.***282**: 15578-15588.

132. Aikawa, J., and Esko, J.D. (1999) *J. Biol. Chem.***274**: 2690-2695.

133. Bame, K.J., and Esko, J.D. (1989) *J. Biol. Chem.***264**: 8059-8065.

134. Forsberg, E., and Kjellen, L. (2001) *J. Clin. Invest.***108**: 175-180.

135. Morii, E., Ogihara, H., Oboki, K., Sawa, C., Sakuma, T., Nomura, S., Esko, J.D. et al. (2001) *Blood*. **97**: 3032-3039.

136. Bornemann, D.J., Park, S., Phin, S., and Warrior, R. (2008) *Development*. **135**: 1039-1047.

137. Carlsson, P.a.K., L. (2012) *Handbook of Experimental Pharmacology*. **207**: 24-36.

138. Yusa, A., Kitajima, K., and Habuchi, O. (2005) *Biochem. J.***388**: 115-121.

139. Schwartz, N.B., Roden, L., and Dorfman, A. (1974) *Biochem. Biophys. Res. Commun.***56**: 717-724.

140. Ledin, J., Ringvall, M., Thuveson, M., Eriksson, I., Wilen, M., Kusche-Gullberg, M., Forsberg, E. et al. (2006) *J. Biol. Chem.***281**(47): 35727-35734.

141. Davie, E.W., Fujikawa, K., Kisiel, W. (1991) *Biochemistry*. **30**(43): 10363-10371.

142. Gailani, D., and Renne, T. (2007) *Arterioscler. Thromb. Vasc. Biol.***27**(12): 2507-2513.

143. Huntington, J.A., Read, R.J., and Carrell, R.W. (2000) *Nature*. **407**: 923-926.

144. Silverman, G.A., Whisstock, J.C., Bottomley, S.P., Huntington, J.A., Kaiserman, D., Luke, C.J., Pak, S.C. et al. (2010) *J. Biol. Chem.***285**: 24299-24305.

145. Whisstock, J.C., Silverman, G.A., Bird, P.I., Bottomley, S.P., Kaiserman, D., Luke, C.J., Pak, S.C. et al. (2010) *J. Biol. Chem.***285**: 24307-24312.

146. Huntington, J.A., Kjellberg, M., and Stenflo, J. (2003) *Structure*. **11**: 205-215.

147. Olson, S.T., Richard, B., Izaguirre, G., Schedin-Weiss, S., and Gettins, P.G. (2010) *Biochimie*. **92**(11): 1587-1596.

148. Ishiguro, K., Kojima, T., Kadomatsu, K., Nakayama, Y., Takagi, A., Suzuki, M., Takeda, N., et al. (2000) *J. Clin. Invest.***106**(7): 873-878.

149. Whisstock, J.C., Pike, R.N., Jin, L., Skinner, R., Pei, X.Y., Carrell, R.W., and Lesk, A.M. (2000) *J. Mol. Biol.***301**(5): 1287-1305.

150. Lindahl, U., Thunberg, L., Backstrom, G., Riesenfeld, J., Nordling, K., and Bjork, I. (1984) *J. Biol. Chem.***259**: 12368-12376.

151. Atha, D.H., Lormeau, J., Petitou, M., Rosenberg, R.D., and Choay, J. (1985) *Biochemistry*. **24**(23): 6723-6729.

152. Kuberan, B., Lech, M., Beeler, D.L., Wu, Z.L., and Rosenberg, R.D. (2003) *Nat. Biotech.***21**(11): 1343-1346.

153. Streusand, V.J., Bjork, I., Gettins, P.G., Petitou, M., and Olson, S.T. (1995) *J. Biol. Chem.***270**: 9043-9051.

154. Jin, L., Abrahams, J.P., Skinner, R., Petitou, M., Pike, R.N., and Carrell, R.W. . (1997) *Proc. Nat. Acad. Sci.***94**: 14683-14688.

155. Al Dieri, R., Wagenvoord, R., van Dedem, G.W., Beguin, S., and Hemker, H.C. (2003) *J. Thromb. Haemost.***1**: 907-914.

156. Gettins, P.G.W., and Olson, S.T. (2009) *J. Biol. Chem.***284**: 20441-20445.

157. Johnson, D.J., Langdown, J., and Huntington, J.A. (2010) *Proc. Nat. Acad. Sci.***107**: 645-650.

158. Yang, L., Manithody, C., and Rezaie, A.R. (2002) *J. Biol. Chem.***277**: 50756-50760.

159. Yang, L., Sun, M.F., Gailani, D., and Rezaie, A.R. (2009) *Biochemistry*. **48**: 1517-1524.

160. Yang, L., Manithody, C., Qureshi, S.H., and Rezaie, A.R. (2010) *Thromb. Haemost.***103**: 277-283.

161. O'Keefe, D., Olson, S.T., Gasiunas, N., Gallagher, J., Baglin, T.P., and Huntington, J.A. (2004) *J. Biol. Chem.***279**: 50267-50273.

162. Baglin, T.P., Carrell, R.W., Church, F.C., Esmon, C.T., and Huntington, J.A. (2002) *Proc. Nat. Acad. Sci*.**99**: 11079-11084.

163. Tollefsen, D.M., and Blank, M.K. (1981) *J. Clin. Invest.***68**: 589-596.

164. Sun, W., Eriksson, A.S., and Schedin-Weiss, S. (2009) *Biochemisty*. **48**: 1094-1098.

165. Van Walderveen, M.C., Berry, L.R., and Chan, A.K. (2010) *J. Biochem.***148**: 255-260.

166. Broze, G.J., Warren, L.A., Novotny, W.F., Higuchi, D.A., Girard, J.J., and Miletich, J.P. (1988) *Blood*. **71**: 335-343.

167. Givol, D., and Yayon, A. . (1992) *FASEB J.***6**: 3362-3369.

168. Jaye, M., Schlessinger, J., and Dionne, C.A. (1992) *Biochim. Biophys. Acta.***1135**: 185-199.

169. Lee, P.L., Johnson, D.E., Cousens, L.S., Fried, V.A., and Williams, L.T. (1989) *Science*. **245**: 57-60.

170. Ornitz, D.M., and Itoh, N. . (2001) *Genome Biol*. **2**: 3005.

171. Givol, D., Eswarakumar, V.P., and Lonai, P. (2003) *Oxford University Press*: 367-379.

172. Lin, X., Buff, E.M., Perrimon, M., and Michelson, A.M. . (1999) *Development*. **126**: 3715-3723.

173. Spivak-Kroizman, T., Lemmon, M.A., Dikic, I., Ladbury, J.E., Pinchasi, D., Huang, J. et al. (1994) *Cell*. **79**: 1015-1024.

174. Yayon, A., Klagsbrun, M., Esko, J.D., Leder, P., and Ornitz, D.M. (1991) *Cell*. **64**: 841-848.

175. Schlessinger, J., Plotnikov, A.N., Ibrahimi, O.A., Eliseenkova, A.V., Yeh, B.K., Yayon, A. et al. (2000) *Mol. Cell*. **6**: 743-750.

176. Maccarana, M., Casu, B., and Lindahl, U. (1993) *J. Biol. Chem.***268**: 23898-23905.

177. Kreuger, J., Prydz, K., Pettersson, R., Lindahl, U., and Salmivirta, M. (1999) *Glycobiology*. **9**: 723-729.

178. Sugaya, N., Habuchi, H., Nagai, N., Ashikara-Hada, S., and Kimata, K. (2008) *J. Biol. Chem.***283**: 10366-10376.

179. Shukla, D., and Spear, P.G. (2001) *J. Clin. Invest.***108**: 503-510.

180. Liu, J., Shriver, Z., Pope, R.M., Thorp, S.C., Duncan, M.B., Copeland, R.J., Raska, C.S. et al. (2002) *J. Biol. Chem.***277**: 33456-33467.

181. DeLarco, J.E., Wuertz, B.R., and Furcht, L.T. (2004) *Cancer Res.***10**(15): 4895-4900.

182. Jacobs, L., Nawrot, T.S., de Geus, B., Meeusen, R., Degraeuwe, B., Bernard, A., Sughis, M. et al. (2010) *Environmental Health*. **9**(64).

183. Waugh, D.J., and Wilson, C. (2008) *Cancer Res.***14**(21): 6735-6741.

184. Li, J., and Vlodovsky, I. (2009) *Thromb. Haemost.***102**: 823-828.

185. Massena, S., Christoffersson, G., Hjertstrom, E., Zcharia, E., Vlodovsky, I., Ausmees, N., Rolny, C. et al. (2010) *Blood*. **116**(11): 1924-1931.

186. Wang, L., Brown, J.R., Varki, A., and Esko, J.D. (2002) *J. Clin. Invest.***110**(1): 127-136.

187. Butcher, E.C. (1991) *Cell*. **67**: 1033-1036.

188. Springer, T.A. (1994) *Cell*. **76**: 301-314.

189. Najjam, S., Mulloy, B., Theze, J., Gordon, M., Gibbs, R., and Rider, C. (1998) *Glycobiology*. **8**(5): 509-516.

190. Salek-Ardakani, S., Arrand, J.R., Shaw, D., and Mackett, M. (2000) *Blood*. **96**(5): 1879-1888.

191. Spillman, D., Witt, D., and Lindahl, U. (1998) *J. Biol. Chem.***273**: 15487-15493.

192. Vives, R.R., Sadir, R., Imberty, A., Rencurosi, A., and Lortat-Jacob, H. (2002) *Biochemistry*. **41**(50): 14779-14789.

193. Witt, D.P., and Lander, A.D. (1994) *Curr. Biol.***4**: 394-400.

194. Tanaka, Y., Kimata, K., Wake, A., Mine, S., Morimoto, I., Yamakawa, N., Habuchi, H. et al. (1996) *J. Exp. Med.***184**: 1987-1997.

195. Norgard-Sumnicht, K., and Varki, A. (1995) *J. Biol. Chem.***270**: 12012-12024.

196. Luo, J., Kato, M., Wang, H., Bernfield, M., and Bischoff, J. (2001) *J. Cell. Biochem.***80**(4): 522-531.

197. Ahmed, T., Garrigo, J., and Danta, I. (1993) *N. Engl. J. Med.***329**: 90-95.

198. Saliba, M.J. (2001) *Burns*. **27**: 349-358.

199. Torkvist, L., Thorlacius, H., Sjoqvist, U., Bohman, L., Lapidus, A., Flood, L., Agren, B. et al. (1999) *Aliment. Pharmacol. Ther.***13**: 1323-1328.

200. Li, R.W., Freeman, C., Yu, D., Hindmarsh, E.J., Tymms, K.E., Parish, C.R., and Smith, P.N. (2008) *Arthritis Rheum.***58**: 1590-1600.

201. Waterman, M., Ben-Izhak, O., Eliakim, R., Groisman, G., Vlodovsky, I., and Ilan, N. (2007) *Mod. Pathol.***20**: 8-14.

202. Yang, Y., MacLeod, V., Miao, H.Q., Theus, A., Zhan, F., Shaughnessy, J.D., Sawyer, J. et al. (2007) *J. Biol. Chem.***282**: 13326-13333.

203. Gong, F., Jemth, P., Escobar Galvis, M.L., Vlodovsky, I., Horner, A., Lindahl, U., and Li, J. (2003) *J. Biol. Chem.***278**: 35152-35158.

204. Pikas, D.S., Li, J.P., Vlodovsky, I., and Lindahl, U. (1998) *J. Biol. Chem.***273**: 18770-18777.

205. Escobar Galvis, M.L., Jia, J., Zhang, X., Jastrebova, N., Spillman, D., Gottfridsson, E., van Kuppevelt, T.H. et al. (2007) *Nat. Chem. Biol.***3**: 773-778.

206. Shinn, K., Nigrovic, P.A., Crish, J., Boilard, E., McNeil, H.P., Larabee, K.S., Adachi, R. et al. (2009) *J. Immunol.***182**: 647-656.

207. Liu, D., Shriver, Z., Qi, Y., Venkataraman, G., and Sasisekharan, R. (2002) *Semin. Thromb. Hemost.***28**: 67-78.

208. Varki, N.M., and Varki, A. (2002) *Semin. Thromb. Hemost.***28**: 53-66.

209. Liu, D., Shriver, Z., Venkataraman, G., El Shabrawi, Y., and Sasisekharan, R. . (2002) *Proc. Nat. Acad. Sci*.**99**: 568-573.

210. Blackhall, F.H., Merry, C.L., Davies, E.J., and Jayson, G.C. (2001) *Br. J. Cancer*. **85**: 1094-1098.

211. Xiang, Y.Y., Ladeda, V., and Filmus, J. (2001) *Oncogene*. **20**: 7408-7412.

212. Filmus, J. (2001) *Glycobiology*. **11**: 19R-23R.

213. Sanderson, R.D. (2001) *Semin. Cell Dev. Biol.***12**: 89-98.

214. Sasisekharan, R., Shriver, G., Venkataraman, G., and Narayanasami, U. (2002) *Nature Rev.***2**: 521-528.

215. Iozzo, R.V., and San Antonio, R.D. (2001) *J. Clin. Invest.***108**: 349-355.

216. Freeman, C., and Parish, C.R. (1998) *Biochem. J.***330**: 1341-1350.

217. Okada, Y., Yamada, S., Toyoshima, M., Dong, J., Nakajima, M., and Sugahara, K. (2002) *J. Biol. Chem.***277**: 42488-42495.

218. Elkin, M., Ilan, N., Ishai-Michaeli, R., Friedmann, Y., Papo, O., Pecker, I., and Vlodovsky, I. (2001) *FASB J.***15**: 1661-1663.

219. Simizu, S., Ishida, K., and Osada, H. (2004) *Cancer Sci.***95**: 553-558.

220. Tessier, D.C., Thomas, D.Y., Khouri, H.E., Laliberie, F., and Vernet, T. (1991) *Gene*. **98**(2): 177-183.

221. Vaughn, J.L., Goodwin, R.H., Tompkins, G.L., and McCawley, P. (1977) *In Vitro*. **12**: 213.

222. Lockow, V.A., and Summers, M.D. (1988) *Bio/Technology*. **6**: 47.

223. Smith, G.E., Summers, M.D., and Fraser, M.J. (1983) *Mol. Cell. Biol.***3**: 2156.

224. Chen, M., Bridges, A., and Liu, J. (2006) *Biochemistry*. **45**(40): 12358-12365.

225. Sismey-Ragatz, A.E., Green, D.E., Otto, N.J., Rejzek, M., Field, R.A., and DeAngelis, P.L. (2007) *J. Biol. Chem.***282**: 28321-28327.

226. Liu, R., Xu, Y., Chen, M., Weiwer, M., Zhou, X., Bridges, A.S., DeAngelis, P.L. et al. (2010) *J. Biol. Chem.***285**: 34240-34249.

227. Domon, B., and Costello, C.E. (1988) *Glycoconj. J.***5**: 397-409.

228. Pettersson, I., Kusche, M., Unger, E., Wlad, H., Nyland, L., Lindahl, U., and Kjellen, L. . (1991) *J. Biol. Chem.***266**(13): 8044-8049.

229. Raman, K., Nguyen, T.K.N., and Kuberan, B. (2011) *FEBS Letters*. **585**: 3420-3423.

230. Mollay, C., Vilas, U., and Kreil, G. (1982) *Proc. Nat. Acad. Sci.***79**(7): 2260-2263.

231. Peterson, S., and Liu, J. (2012) *J. Biol. Chem.***287**: 34863-34843.

232. Schwartz, N.B. (1975) *FEBS Letters*. **49**: 342-345.