TOWARDS *DE NOVO* SYNTHESIS OF STRUCTURE-DEFINED OLIGOSACCHARIDES WITH HEPARAN SULFATE BIOSYNTHETIC ENZYMES

Miao Chen

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 the School of Pharmacy.

Chapel Hill
2008

Approved by:

Jian Liu, Ph.D.

Kenneth Bastow, Ph.D.

Frank Church, Ph.D.

Anthony Hickey, Ph.D.

Harold Kohn, Ph.D.

ABSTRACT

Miao Chen: Towards *de novo* synthesis of structure-defined oligosaccharides with heparan sulfate biosynthetic enzymes (Under the direction of Jian Liu, Ph.D.)

Heparan sulfate (HS), a highly sulfated polysaccharide, plays fundamental roles in a wide range of physiological and pathophysiological processes. The specific structure of HS determines its biological function. Various HS biosynthetic enzymes dictate the final structure of HS products. As an attractive target for medical and pharmaceutical research, synthesis of structurally defined HS oligosaccharides is one of the major challenges in the field of glycobiology. Enzymatic synthesis of HS represents a valuable complementary approach to chemical synthesis. The access to the unsulfated, unepimerized HS backbones is an essential step towards enzymatic synthesis of HS. In this research, we developed a chemoenzymatic approach for *de novo* synthesis of HS backbones with two bacterial glycosyltransferase, KfiA and PmHS2. KfiA is an *N*-acetylglucosaminyltransferase in *Escherichia coli* strain K5, while PmHS2 is a bifunctional enzyme with both *N*-acetylglucosaminyltransferase and D-glucuronyltransferase in *Pasteurella multocida* Types A, D, and F. Both of the recombinant proteins were prepared in large quantities. Characterization of KfiA reveals that its enzymatic activity is independent of the size of the acceptor substrates but a particular structure in the natural donor UDP-*N*-acetylglucosamine is required to bind KfiA. UDP-*N*-trifluoroacetylglucosamine, a derivative of UDP-*N*-acetylglucosamine, was recognized by KfiA and employed to synthesize unnatural HS backbones containing *N*-trifluoroacetylglucosamine residues. The synthesized oligosaccharides range from

trisaccharide to undecasaccharide with a yield of 100 µg to milligram scales. The controlled *N*-sulfation on HS backbones was achieved by selective removal of *N*-trifluoroacetyl groups and a subsequent *N*-sulfation with *N*-sulfotransferase. The synthesized octasaccharide backbone with defined *N*-sulfation was further modified with a collection of HS biosynthetic enzymes and yielded a bioactive octasaccharide binding to antithrombin. The results of this study open a new approach for the *de novo* synthesis of structure-defined HS oligosaccharides.

ACKNOWLEDGEMENTS

enzymes in large quantities.

I must thank all the other members and ex-members of the excellent glycobiology group, Dr. Suzanne Edavettal, Dr. Michael Duncan, Dr. Ronald Copeland, Dr. Yongmei Xu, Tanya Scarlett, Sherket Peterson, Ryan Bullis and Xinan Lu, for their assistance. The enlightening discussions and the friendly atmosphere have made my stay pleasant and enjoyable. I am so fortunate to have been a part of this big family.

Finally, this dissertation is dedicated to my parents, Yiguang Chen and Huili Wang, and my big bear, Dundun, for their unconditional love and unwavering faith in me.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

ABBREVIATIONS

2OST	uronosyl 2-*O*-sulfotransferase
3OST	glycosaminyl 3-*O*-sulfotransferase
6OST	glycosaminyl 6-*O*-sulfotransferase
anMannose	2,5-anhydromannose
anMan$_R$	2,5-anhydromannitol
AST-IV	arylsulfotransferase IV
AT	antithrombin
ATP	adenosine triphosphate
*C*5-epi	glucoronyl *C*5-epimerase
CDSNS	chemically desulfated and *N*-sulfated
CE	capillary electrophoresis
CoASH	coenzyme A
CS	chondroitin sulfate
DTT	dithiothreitol
ECM	extracellular matrix
EST	estrogen sulfotransferase
EXT	exostosin
FGF(R)	fibroblast growth factor (receptor)
GAG	glycosaminoglycan
Gal	Galactose

GalNAc	*N*-acetylated galactosamine
GalTI/II	HS galactosyltransferase I/II
GB/C/D	HSV-1 glycoprotein B/C/D
Glc	glucose
GlcA	glucuronic acid
GlcATI/II	HS glucuronosyltransferase I/II
GlcNS	*N*-sulfated glucosamine
GlcN	*N*-unsubstituted glucosamine
GlcN-1-P	glucosamine-1-phosphate
GlcNAc	*N*-acetylated glucosamine
GlcNAcTI/II	HS GlcNAc transferase I/II
GlcNCOCF$_3$	*N*-trifluoroacetylglucosamine
GlmU	glucosamine-1-phosphate acetyltransferase / *N*-acetylglucosamine-1-phosphate uridyltransferase
HA	hyaluronic acid
HCV	hepatitis C virus
HIV	human immunodeficiency virus
HPLC	high performance liquid chromatography
HS	heparan sulfate
HSPG	heparan sulfate proteoglycan
HSV	herpes simplex virus
IdoA	iduronic acid
ITC	isothermal titration calorimetry

KS	keratan sulfate
LMWH	low molecular weight heparin
MS	mass spectrometry
NDST	*N*-deacetylase/*N*-sulfotransferase
NST	*N*-sulfotransferase
PAMN	polyamine anion exchange
PAP	3'-phosphoadenosine 5'-phosphate
PAPS	3'-phosphoadenosine 5'-phosphosulfate
PCR	polymerase chain reaction
RPIP	reverse-phase ion-pairing
UDP	uridine diphosphate
UDP-Gal	uridine diphosphogalactose
UDP-GalNAc	uridine diphospho-*N*-acetyl-galactose
UDP-Glc	uridine diphosphoglucose
UDP-GlcA	uridine diphosphoglucuronic acid
UDP-GlcNAc	uridine diphospho-*N*-acetyl-glucosamine
UDP-GlcNCOCF$_3$	uridine diphospho-*N*-trifluoroacetyl-glucosamine
UDP-Xyl	uridine diphosphoxylose
UTP	uridine triphosphate
VEGF	vascular endothelial growth factor
XT	xylosyltransferase
Xyl	xylose
ΔUA	Δ4,5-unsaturated uronic acid

Chapter I. Introduction

Section 1. Structure of heparan sulfate

1. Overall structure of heparan sulfate

Heparan sulfate (HS) is a linear polymer that belongs to a family of macromolecules known as glycosaminoglycans (GAGs). GAGs are ubiquitous macromolecules associated with the cell surface and in the extracellular matrix (ECM). They are classified into four subclasses: chondroitin sulfate (CS), keratan sulfate (KS), hyaluronic acid (HA) and heparan sulfate (HS).

HS consists of alternating hexuronic acid (D-glucuronic acid/GlcA or L-iduronic acid/IdoA) and D-glucosamine (GlcN) units carrying sulfo groups, attached via 1→4 linkages. HS is structurally diverse through various modifications on the repeating disaccharide unit. The hexuronic acid residue can be either GlcA or IdoA depending upon the configuration at the $C5$ position, and its $C2$ position can be O-sulfated. The GlcN residue can be N-acetylated (GlcNAc), N-sulfated (GlcNS) and/or O-sulfated at its $C3$ and $C6$ positions (figure 1).
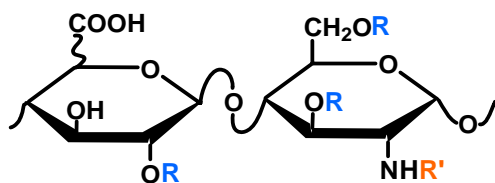
Figure 1. Disaccharide repeating unit of heparan sulfate
R= H or SO_3^-, R'= H, Ac or SO_3^-

Under physiological conditions, HS exists in the form of heparan sulfate proteoglycans (HSPGs). HSPG consists of a core protein attached by one or more HS chains. Based on their cellular locations, HSPGs are classified into two groups: membrane-bound HSPGs (syndecans and glypicans) positioned on the cell surface, and secreted HSPGs (perlecans, agrins and collagen XVIII) located in the ECM (Iozzo 2001).

2. Structural analysis of heparan sulfate

Unlike protein and nucleic acid preparative techniques, the methods for preparing and purifying HS are highly limited. Specifically, the heterogeneity of HS and lack of a fast automated method to sequence HS complicate the structure-function relationship study of HS. The most commonly used approach for analyzing the structure of HS is to degrade the polysaccharide into disaccharides or oligosaccharides. The resultant disaccharides or oligosaccharides are identified by various analytical techniques such as high performance liquid chromatography (HPLC), mass spectrometry (MS) and capillary electrophoresis (CE).

Depolymerization of HS can be achieved with enzymatic or chemical methods. Heparin lyases are endoglycosidases that have different substrate specificities (Desai 1993). As shown in figure 2, heparin lyase I cleaves the glycosidic linkage between 2-*O*-sulfated IdoA and GlcNS within HS where highly sulfated domains reside (Liu 2002). Heparin lyase III cleaves the glycosidic linkage between GlcA and either GlcNAc or GlcNS which belongs to the less sulfated domains within HS. Heparin lyase

II, however, demonstrates more flexibility and cleaves at the sites that either heparin lyase I or III recognizes. Using heparin lyases make it possible to directly analyze the degradation products. This may be explained by the generation of a $\Delta^{4,5}$-unsaturated uronic acid at the non-reducing end of the product that has a characteristic absorbance at 232 nm. On the other hand, it is impossible to assign GlcA or IdoA in the precursor structure, because the generation of this uronic acid residue leads to a loss of structural information on the C5 configuration of the hexuronic acid residue. It is also unlikely to reach completion for the enzymatic degradation of HS, therefore raising a concern of underestimation of some disaccharide species.
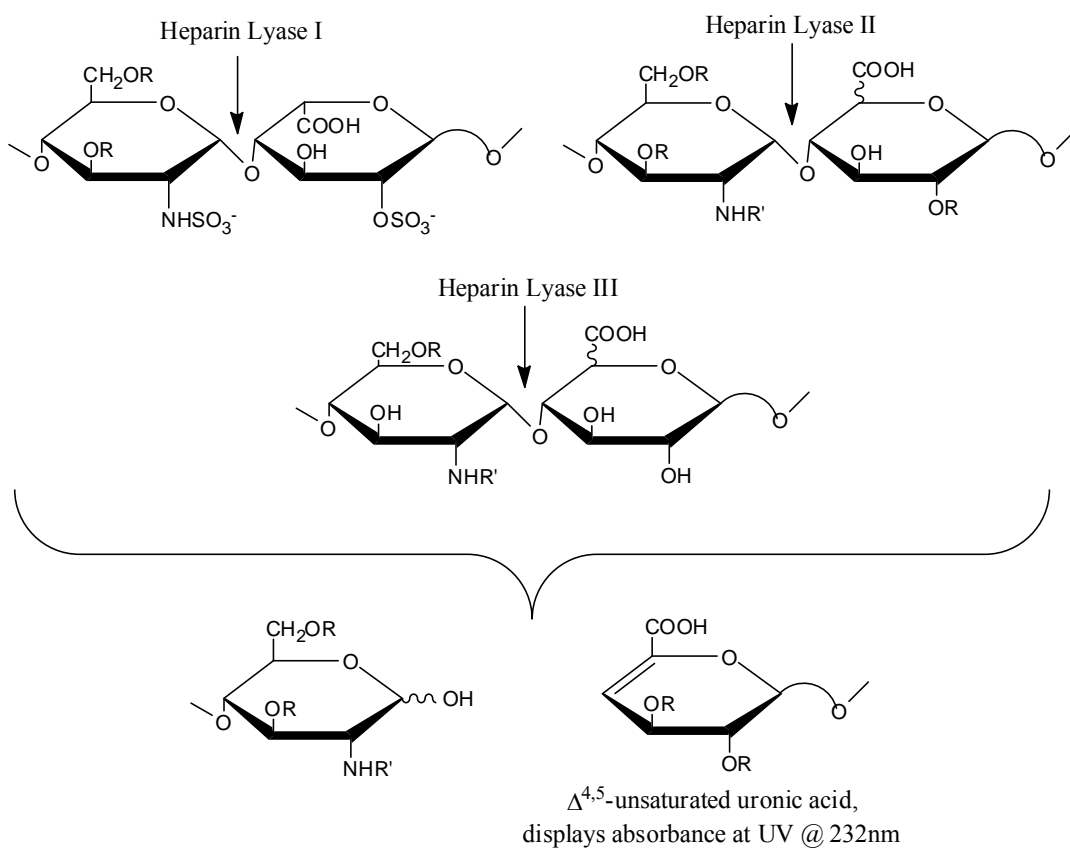
Figure 2. Heparin lyases degrade HS with different substrate specificities (Liu 2002)
R= H or SO$_3^-$, R'= H, Ac or SO$_3^-$

Nitrous acid degradation is commonly used as the chemical approach to depolymerize HS (figure 3). The reaction selectivity is pH dependent. At low pH (pH 1.5), nitrous acid predominantly reacts with GlcNS, while at high pH (pH 4.5-5.5) it primarily reacts with GlcN (Shively 1976). Either way the reaction will generate 2,5-anhydromannose which can be further reduced to 2,5-anhydromannitol. Notably, nitrous acid degradation retains the structure of GlcA/IdoA and, therefore, allows the determination of the $C5$ configuration of hexuronic acid on the HS chain. But the tradeoff is that additional radioactive or fluorescent label is required to monitor the resultant disaccharides. Liquid chromatography-mass spectrometry (LC-MS) is a recently developed approach that combines the physical separation capability of liquid chromatography with the mass analysis capability of mass spectrometry (Tomer 1994). It can be used with nanospray sources and has a high specificity and sensitivity. Most importantly, LC-MS can overcome the difficulty arising from the lack of radioactive or fluorescent labels.
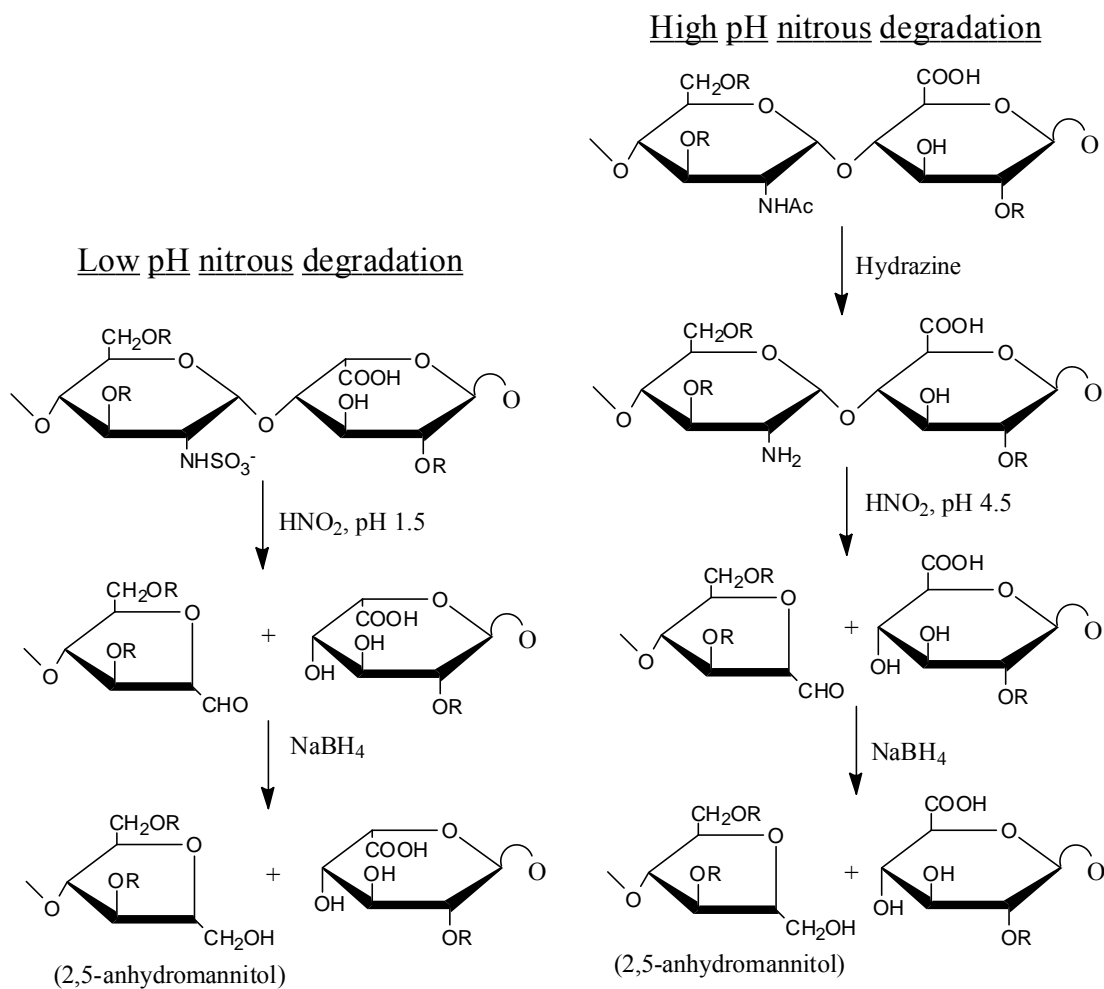
Figure 3. Nitrous acid degradation of heparan sulfate (Liu 2002)
R= H or SO$_3^-$

Section 2. Biological functions of heparan sulfate and the defined bioactive domains

HS is a charge-rich polymer widely distributed on the cell surface and in the ECM.   It plays crucial roles in numerous physiological and pathological processes through electrostatic interactions with proteins (Esko 2002).   Some biologically active oligosaccharide structures on the HS chain have been identified.

1. Antithrombotic effect and the antithrombin (AT)-binding domain

Heparin has been the top choice for the treatment of arterial and venous thrombotic disorders over a half century.   Blood coagulation is a cascade process (figure 4) (Broze 1995).   When the coagulant cascade is initiated intrinsically, the contact factors such as factor XII is exposed to the surface and converted to its active form, factor XIIa.   Factor XIIa is a serine protease and goes on to catalyze the activation of the next zymogen until prothrombin is converted to thrombin.   The extrinsic pathway is initiated by binding of factor VIIa to tissue factor which is followed by a series of proteolytic reactions, leading to the generation of thrombin.   In both pathways, thrombin cleaves fibrinogen into fibrin and the polymerization of fibrin leads to an insoluble clot.
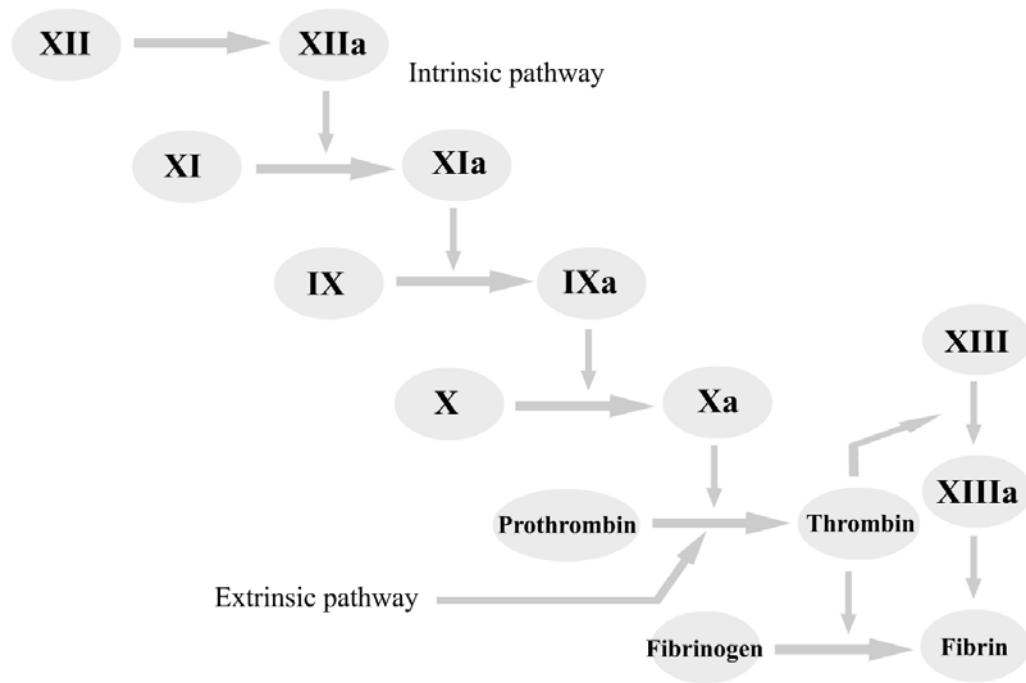
Figure 4. Blood coagulation cascade

The cascade is triggered by either intrinsic pathway or extrinsic pathway. The two routes converge on the generation of thrombin which leads to blood clotting.

The anticoagulant activity of heparin was first discovered in an extract of dog liver at Johns Hopkins University more than ninety years ago (McLean 1916). Heparin/HS modulates blood coagulation through the inhibition of the serine-proteases in the coagulation cascade, including factor Xa and thrombin. The inhibition is either direct or indirect through its binding with the protease inhibitor antithrombin (AT) (Lever 2002). Rosenberg and colleagues found that AT forms a 1:1 complex with thrombin (Rosenberg 1973). The research demonstrated that the addition of exogenous heparin markedly promotes the formation of this complex which decelerates blood coagulation. It was suggested that heparin binding with AT causes a conformational change of AT which leads to a more favorable exposure of the arginine reactive sites on this inhibitor, allowing a rapid interaction with thrombin.

A well-defined structure (figure 5) has been determined to be crucial for the AT-binding capability of heparin/HS (Lindahl 1980; Atha 1985). The AT-binding domain is a pentasaccharide sequence of -GlcNAc(or GlcNS)6S-GlcA-GlcNS3S±6S-IdoA2S-GlcNS6S- and is the essential motif conferring anticoagulant activity to heparin/HS. The critical 3-*O*-sulfation on the GlcNS3S±6S residue is catalyzed by 3-*O*-sulfotransferase 1 (3OST1) or 3-*O*-sulfotransferase 5 (3OST5) (Shworak 1999; Xia 2002).

The pentasaccharide alone can effectively regulate the activity of factor Xa, while a ternary complex of HS, AT and thrombin was proposed to inactivate thrombin. Therefore, a longer HS fragment is necessary to accommodate both AT and thrombin in

order to inhibit the activity of thrombin (Olson 1992).

R = Acetyl or SO₃⁻

R' = H or SO₃⁻
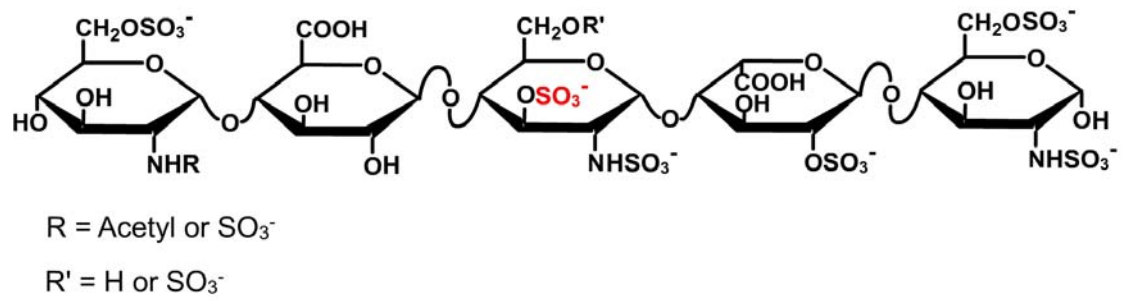
Figure 5. AT-binding pentasaccharide

The pentasaccharide contains a critical 3-*O*-sulfation (in red) in the middle glucosamine residue. Its AT-binding affinity is 30 nM, while the same pentasaccharide, except lacking this important 3-*O*-sulfation (in red), has a binding affinity of 500 μM (Atha 1985).

2. Cell proliferation/differentiation and the fibroblast growth factor (FGF)-binding domain

The FGF family of proteins form a family of approximately 20 structurally related polypeptides which are involved in a number of biological regulations in cell growth and development (Basilico 1992). The FGF signaling pathway is initiated by binding of FGF with its tyrosine kinase receptor (FGFR) on the cell surface, leading to the dimerization of FGFR and the phosphorylation/activation of downstream enzymes. The process is modulated through an interaction between FGF, FGFR and HS molecules. HS has been implicated to enhance the formation of FGF-FGFR complexes and stabilize FGFR oligomers (Brickman 1998; Kwan 2001).

FGF1 and FGF2 are the two mostly investigated FGF proteins. A protein crystallization technique has been employed to solve the structure of the complex of HS oligosaccharides and these two proteins. As shown in figure 6, the 2-*O*-sulfate on iduronate (IdoA2S) and *N*-sulfate on glucosamine (GlcNS) are critical for both FGF1 and FGF2 binding (Tabeur 1999), while additional 6-*O*-sulfate on glucosamine (GlcNS6S) is only necessary for FGF1 binding (Guerrini 2002). It is noteworthy that FGFR also contains a HS-binding site which interacts with GlcNS6S (Loo 2001), suggesting that, in order to trigger the dimerization of FGFR and the downstream cell signaling, the active HS fragment should be long enough (figure 6, n>5) to include two binding motifs for both FGF and FGFR. Further investigation of the HS structures involved in other

growth factor signaling pathway was also initiated. For example, a systematic investigation of the HS binding structures for various heparin-binding growth factors was made by subjecting an octasaccharide library to affinity chromatography (Ashikari-Hada 2004). Some octasaccharides were found capable of releasing FGF2 and FGF10 from their complex with HS, suggesting a new approach to specifically regulate bindings of the growth factors to HS.
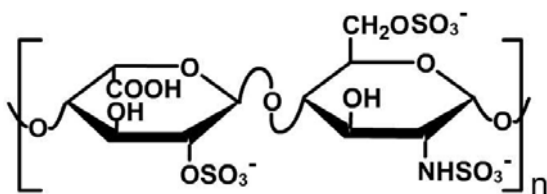
Figure 6. Repeating disaccharide unit –IdoA2S-GlcNS6S- implicated in FGF binding

n>2 needed for binding to FGF1 (-IdoA2S-GlcNS6S-) and FGF2 (-IdoA2S-GlcNS-).
n>5 needed for dimerization and cell signaling.

3. Viral infection and the herpes simplex virus (HSV)-1 glycoprotein D-binding domain

The recognition and binding of host cells is the first step for virus to establish an infection. Since HS is ubiquitously expressed on the surface of most animal cells, it is not surprising that HS often serves as a natural bridge for virion particles to attach and invade target cells. It has been demonstrated that HS is involved in the infection of many viruses including human immunodeficiency virus (HIV), hepatitis C virus (HCV), dengue virus and herpes simplex virus (HSV) (Liu 2002). As a result, the study on the structure and mechanism of HS-virus interaction has attracted considerable attention from both scientific and pharmaceutical communities. For example, the involvement of HS in HSV-1 infection has been well characterized in the last two decades.

HSVs are neurotropic and neuroinvasive viruses. The viral particles enter and reside in the human nervous system, accounting for their persistent existence in the human body. An infection by HSV is marked by watery blisters in the skin or mucous membranes of the mouth, lips or genitals. HSV-2 is often associated with genital herpes, whereas HSV-1 is commonly associated with herpes outbreaks of the face known as cold sores or fever blisters. After the primary infection, HSV becomes latent in the nerve cells. Some infected patients may experience sporadic episodes of viral reactivation when the virus travels back to the skin via the nerve's axon and replicates again (Ryan 2004).

As for the mechanism of HSV-1 infection, it is now believed that HSV-1 first attaches to target cells through the interaction between its envelope glycoprotein gC/gB and HS on

the surface of host cells (figure 7) (Trybala 1994).
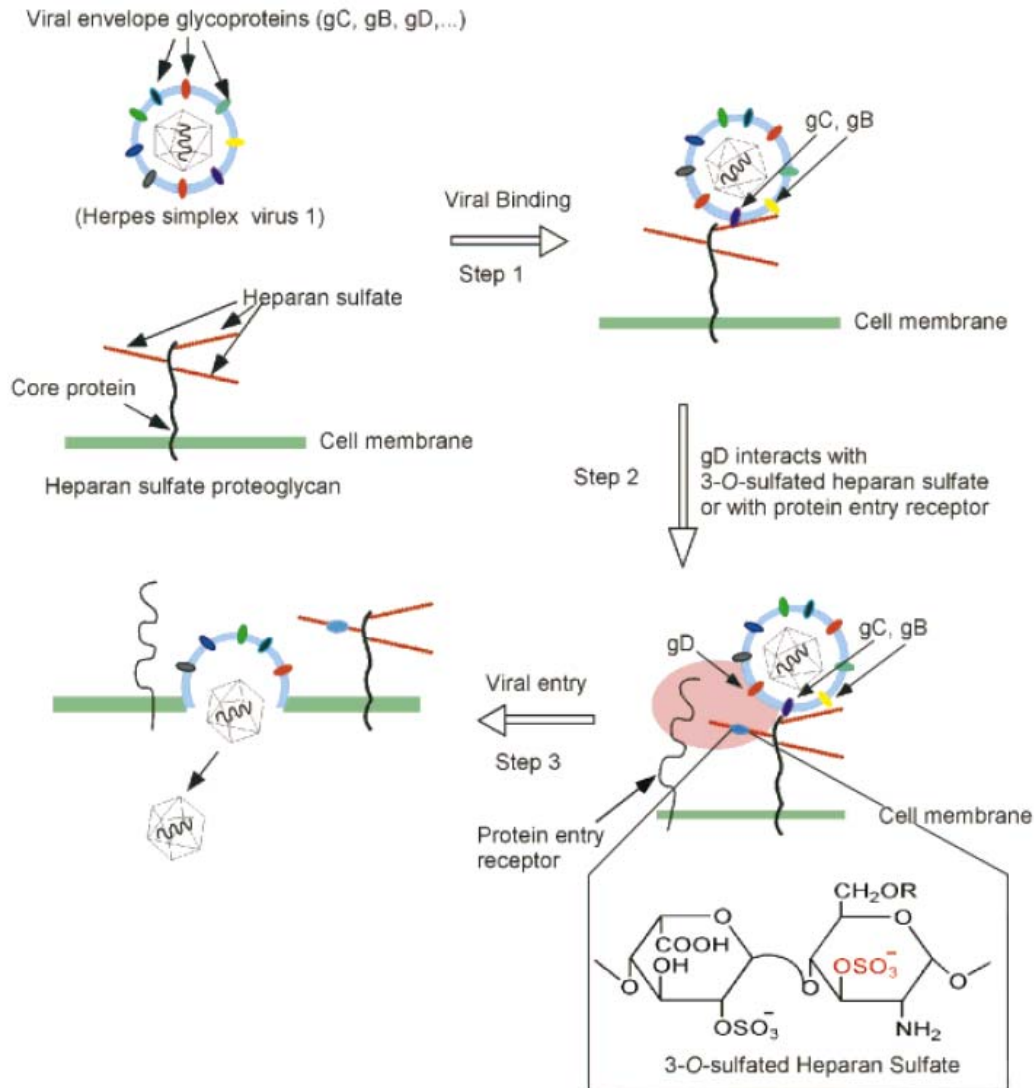
Figure 7. The binding and entry mechanism of HSV-1 virus (Liu 2002)

Step 1. Virus is attached to target cells via the interaction between viral glycoprotein B, C (gB, gC) and the HS on the surface of host cells.

Step 2. Viral glycoprotein D (gD) interacts with the entry receptor and/or 3-*O*-sulfated HS.

Step 3. The interaction in step 2 triggered uncharacterized fusion of cell membranes which leads to the entry of virus.

Once the contact is established, the binding between viral envelope glycoprotein gD and cell surface entry receptor such as 3-*O*-sulfated HS results in the fusion of the virus and cell membranes. This is evident by the observation that wild type CHO cells which are resistant to HSV-1 infection become susceptible to infection after transfected with the cDNA of 3OST3 (Shukla 1999). It is important to note that the transfection of the plasmid expressing 3OST1 did not lead to HSV-1 entry and 3OST1 modified HS does not bind to gD, suggesting that the 3-*O*-sulfated HS involved in HSV-1 infection contains a specific 3-*O*-sulfation pattern. A gD-binding octasaccharide was later isolated from a 3OST3-modified octasaccharide library with a binding affinity of 18 μM. Based on the results from nanoelectrospray ionization mass spectrometry and matrix-assisted laser desorption/ionization mass spectrometry, a proposed heptasulfated sequence of ΔUA-GlcNS-IdoA2S-GlcNAc-UA2S-GlcNS-IdoA2S-GlcN3S6S was assigned to this gD-binding octasaccharide (figure 8) (Liu 2002). Directed by this information, Copeland and colleagues synthesized a HS octasaccharide with the purified 3OST3 enzyme and a heparin octasaccharide *in vitro*. It has a structure of ΔUA2S-GlcNS6S-IdoA2S-GlcNS6S-IdoUA2S-GlcNS3S6S-IdoA2S-GlcNS6S. This 3-*O*-sulfated octasaccharide was demonstrated to have a binding affinity to gD of 19 μM (very similar to that of the previously characterized gD-binding octasaccharide in figure 8) and was effective in blocking HSV-1 infection in a cell-based assay (Copeland 2008).
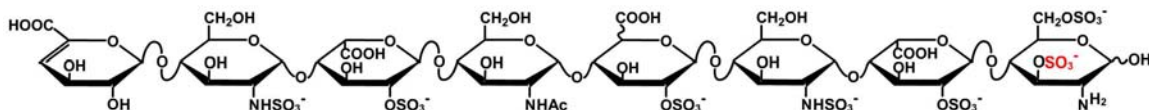
Figure 8. Structure of a gD-binding octasaccharide

The octasaccharide contains a critical 3-*O*-sulfation (in red) at the reducing-end glucosamine residue.   It has a binding affinity to gD of 18 µM (Liu 2002).

4. Inflammation

Inflammation is a protective response by the organism to harmful stimuli from the environment, such as pathogens, damaged cells, or irritants.   In the absence of inflammation, wounds and infections would never heal and progressive destruction of the tissue would compromise the survival of the organism.   However, inflammation which runs unchecked can also lead to a host of diseases.   It is for this reason that inflammation is normally tightly regulated by the body.

In a typical inflammation response, leukocytes migrate from circulation to the injured area by attaching to activated endothelial cells.   Leukocyte trafficking is believed to be directed by the chemokines immobilized on the luminal surface of the endothelium (Middleton 2002).   It has been demonstrated that endothelial heparan sulfate participates in multiple steps in this process, including L-selectin binding, chemokine presentation to its receptors, and the transportation of chemokines across endothelium (Middleton 1997). The   studies   with   mutant   mouse   models   demonstrated   that

N-deacetylase/*N*-sulfotransferase 1 (NDST1) deficient mice showed significantly impaired inflammatory response. The proposed mechanism is that the undersulfated HS in the endothelial cells of mutant mice led to reduction in the presentation of chemokines on their surface, the weakening binding with L-selectin on neutrophil, and ultimately, less recruitment of leukocytes to the inflammation sites *in vivo* (Wang 2005).

The multiple functions of HS in the process of inflammation suggest it has conceivable clinical applications in treating inflammatory diseases. As a potential anti-inflammatory drug, HS has been tested against inflammatory bowel disease (IBD) in an animal test. Heparin pretreatment greatly reduced the inflammatory response in rats elicited by TNF-α administration through a CD11b dependent mechanism, as preincubation with an anti-CD11b mAb significantly diminished heparin binding (Salas 2000). In another two-period, randomized, double-blinded clinical study on asthma, the inhaled heparin attenuated the early and late asthmatic response in patients by an average of 40% and 36%, respectively, which suggests heparin has potential as an anti-asthma therapy (DianMant 1996).

5. Tumor growth and metastasis

Malignant cancer is a class of diseases in which a group of cells display the traits of abnormal and uncontrolled growth, invasion and metastasis (Hanahan 2000). Because HSPGs and HSPG-degrading enzymes mediate cell-cell contact and modulate the

activities of growth and motility factors, they play a critical role in regulating tumor progression (Sasisekharan 2002).

One of the hallmarks of the tumor cells is the sustained abnormal growth. Cell surface HSPGs interact with numerous growth factors, facilitate the binding with their receptors, and therefore promote cell growth. The sulfate groups on HS are critical to form the growth-factor binding sites on HSPGs. Therefore, it is not surprising that the human sulfatase 1 (SULF1), which removes the important sulfate groups on HS and down-regulates cell growth signaling, is suppressed in breast, hepatocellular, pancreatic and many other cancers (Dai 2005).

On the other hand, cell surface HSPGs were also reported to inhibit cancer invasion by keeping tight cell-cell and cell-ECM adhesion. Low level of cell surface HSPGs was shown to correlate with high metastatic activity of tumor cells (Redini 1986) (Moczar 1993). Syndecan-1 on the cell surface was also determined to be a potent suppressor of metastasis by maintaining epithelial morphology (Leppa 1992). For instance, uterine carcinoma and multiple myeloma have lowered expression level of syndecan-1 during their transition to invasive carcinoma (Sanderson 2001). Furthermore, complementation of syndecan-1 in invasive B lymphoid cells enables their tight adhesion to type I collagen and blocks their ability to invade collagen gels (Liebersbach 1994), while transfection of another cell surface HSPG, glypican-1, could not achieve these effects (Liu 1998), indicating that different HSPGs have distinct anticancer capability.

Heparanase is an endo-beta-D-glycosidase acting at both the cell surface and within

the ECM to degrade HS molecules into the polysaccharides with a shorter size. Its overexpression was found in many human tumors examined, including carcinomas of the colon, thyroid, liver, pancreas, head and heck, salivary gland, nasopharyngs, leukemia, lymphoma and multiple myeloma (Ilana 2006). It is postulated that an effective inhibitor of heparanase could serve as a potential anticancer drug. Cell surface HSPGs help to the self-assembly and barrier function of ECM, as well as in cell adhesion and locomotion. Cleavage of HS by over-expressed heparanase may, therefore, lead to ECM remodeling and tumor metastasis. Direct evidence was provided in the studies in which transfection and over-expression of the heparanase gene led to promoted metastases of mouse melanoma, lymphoma and prostate carcinoma cells (Vlodavsky 1999) while heparanase gene silencing greatly suppressed tumor metastasis (Edovitsky 2004).

Angiogenesis is a physiological process involving the growth of new blood vessels from pre-existing vessels. Angiogenesis is a normal process in development and wound healing, as well as in cancer progression for tumor cells to provide sufficient nutrient for their growth. Up-regulated heparanase expression in tumor cells may first directly destroy the subendothelial basement membrane (BM) by cleaving HS in the BM so that endothelial cells can invade and sprout in the early stage of angiogenesis. Heparanase distribution was indeed revealed in capillaries but not in mature blood vessels with the immunohistochemical approach (Elkin 2001). Furthermore, degradation of HS in the ECM may release HS-bound angiogenic growth factors, such as FGF2 and vascular endothelial growth factor (VEGF), from the ECM (Vlodavsky 1996). These molecules

may cooperatively increase the proliferation and migration of endothelial cells and promote neovascularization.

The multifaceted functions of HSPGs in cancer progression have considerable clinical implications for anti-cancer therapy. With the availability of recombinant heparanase and high-throughput screening methods, a variety of heparanase inhibitors have been developed. These inhibitors include the peptides competing with the HS binding domain of heparanase (Vlodavsky 2007), the small molecules targeting the catalytic domain of heparanase (Pan 2006), the monoclonal antibodies (Ferro 2004), and HS derivatives as mimetics of HS, such as maltohexaose sulfate (MHS) and phophomannopentaose sulfate, PI-88 (Parish 1999) (Ferro 2007).

Section 3. Biosynthesis of heparosan

1. Polysaccharide capsules of bacteria

Polysaccharide capsules are found on the surface of a broad spectrum of bacterial species. Such a structure surrounding the microbial cell may protect bacteria from desiccation (Roberson 1992), help bacteria adhere to each other and form a protective biofilm (Costerton 1987), mediate bacteria-plant interaction (Denny 1991), and protect bacteria from specific and nonspecific host immunity (Roberts 1989) (Moxon 1990).

Different from vertebrates, microbes produce the four types of polysaccharide (chondroitin, keratan, hyaluronan and heparason) without epimerization or sulfation

(table 1) (DeAngelis 2002).

Polysaccharide	Disaccharide repeat	Postpolymerization modifications	
		Vertebrates	Bacteria
Chondroitin	→3)GalNAc(β1→4)GlcA(β1→	O-sulfated, epimerized	none or fructose(β1→3)GlcA
Keratan	→4)GlcNAc(β1→3)Gal(β1→	O-sulfated	not reported
Hyaluronan	→3)GlcNAc(β1→4)GlcA(β1→	none	none
Heparason	→4)GlcNAc(α1→4)GlcA(β1→	N-, O-sulfated, epimerized	none

Table 1. Structures of the vertebrate and microbial GAG repeating backbones (DeAngelis 2002).

2. Heparosan glycosyltransferases in bacteria

In bacteria, heparosan glycosyltransferases utilize uridine diphospho (UDP)-monosaccharide precursors (UDP-GlcNAc and/or UDP-GlcA) to synthesize the alternating sugar repeat with the presence of metal cofactors (manganese and/or magnesium ion) near neutral pH (figure 9). Some of the microbial heparosan glycosyltransferases have been identified and characterized such as those in *Escherichia* and *Pasteurella*.
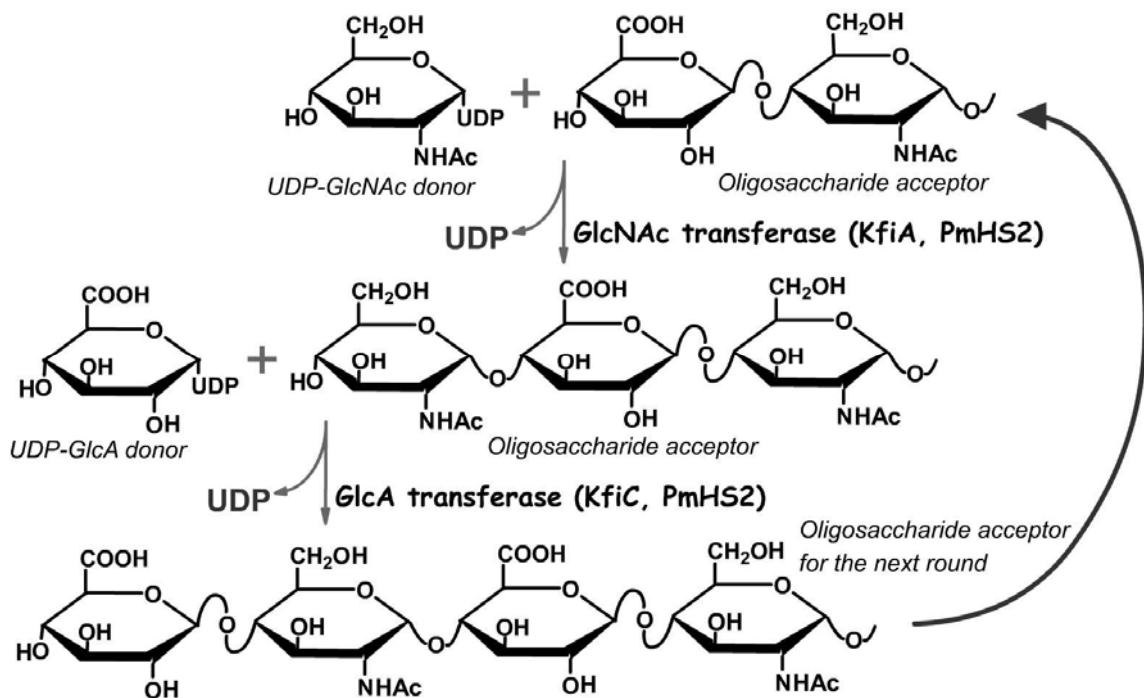
Figure 9. Biosynthesis of heparosan

GlcNAc and GlcA transferases incorporate GlcNAc and GlcA, step by step, into the growing heparosan backbone at its non-reducing end. In *E. coli* K5, the two transferase activities are separately carried by KfiA and KfiC, while in *P. multocida*, PmHS2 is one of the bifunctional enzymes that have both activities.

Escherichia coli K5 capsular polysaccharide is composed of an unepimerized, unsulfated *N*-acetyl heparosan (Vann 1981). It was reported that K5 capsule expression can enhance intestinal colonization by *E. coli* (Hérias 1997). The region 2 of the *E. coli* K5 capsule gene cluster contains four genes, kfiA-D, which have been determined to be essential for heparosan biosynthesis (Petit 1995). With biochemical analysis of the bacteria membrane and lysate, a 60-kDa protein, KfiC, was first identified to be a dual-function enzyme with both glucuronyltransferase and *N*-acetylglucosaminyltransferase activities (Griffiths 1998). However, the same group reported two years later that the *N*-acetylglucosaminyltransferase is actually another protein, KfiA, which is encoded in the same gene cluster (Hodson 2000). The recombinant KfiA protein was expressed and purified in large quantity by our group. Its *N*-acetylglucosaminyltransferase activity was confirmed and its substrate specificities was determined (Chen 2006). The glucuronyltransferase activity of KfiC has yet to be confirmed by the single sugar addition assay with the recombinant protein *in vitro*. The predicted amino acid sequence of KfiD has homology to a number of nicotinamide adenine dinucleotide (NAD)-dependent dehydrogenase enzymes and was demonstrated to be a UDP-glucose dehydrogenase that catalyses the formation of UDP-glucuronic acid from UDP-glucose (Petit 1995). The other deduced protein in the operon, KfiB, has not been assigned any enzymatic activity, but with Western blot analysis it has been suggested to stabilize the enzyme complex during the elongation process *in vivo* (Hodson 2000). Collectively, the four proteins KfiA-D encoded in the region 2 of the *E. coli* K5

capsule gene cluster work in concert to build the heparosan polysaccharide in *E. coli* K5.

Pasteurella multocida is a type of Gram-negative bacteria and its type D capsular polysaccharide is the same as that of *E. coli* K5 (DeAngelis 2002). The two heparosan synthases, PmHS1 (DeAngelis 2002) and PmHS2 (DeAngelis 2004), were identified to synthesize heparosan in *P. multocida* type D. PmHS1 and PmHS2 are bifunctional glycosyltransferases polymerizing both GlcNAc and GlcA *in vitro*. The mutagenesis study showed that PmHS1 possesses both a glucuronyl and a glucosaminyl transfer sites at different regions in one polypeptide chain (Kane 2006).

Besides these bacterial enzymes, heparosan glycosyltransferases have also been identified in mammals. They are known as HS polymerases and will be discussed in detail in Section 4.

Section 4. Biosynthesis of heparan sulfate

In mammals, HS backbone is built by multiple glycosyltransferases and further modified by additional HS biosynthetic enzymes, including *N*-deacetylase/*N*-sulfotransferases, epimerase, and *O*-sulfotransferases in the lumen of Golgi apparatus (Esko 2002). After years of the sustained studies on these enzymes, many of them have been cloned and purified. Much information about the pathway of HS biosynthesis and the substrate specificities of HS biosynthetic enzymes is also available.

1. Chain initiation

HS biosynthesis is initiated in the Golgi apparatus after transfer of xylose (Xyl) from UDP-Xyl to the specific serine residues of a core protein catalyzed by xylosyltransferases (XTs).   UDP-Xyl is derived from the activated glucose, UDP-Glucose (UDP-Glc), by UDP-GlcA decarboxylase.   The attachment sites recognized by XT typically consist of a Ser-Gly dipeptide with one or more flanking acidic amino acid residues (Sugahara 2000). XT has type II transmembrane topology consistent with its location in Golgi where HS biosynthesis occurs.   Without XT, Chinese hamster ovary (CHO) cells were unable to synthesize HS and CS, suggesting that XT is necessary for initiating the biosynthesis of HS and CS (Esko 1985).

Once the Xyl residue is linked to the core protein, a tetrasaccharide linkage (GlcA-β1,3-Gal-β1,3-Gal-β1,4-Xyl-β1-O-[Ser], figure 10) on the core protein will be formed after the sequential attachment of two D-galactose (Gal) residues by galactosyltransferases I and II (GalTI and GalTII) and a GlcA by glucuronosyltransferase I (GlcATI).   Both GalTI and GalTII use UDP-Galactose (UDP-Gal) as the donor substrate and they both show type II transmembrane topology (Sugahara 2000; Bai 2001). The general linkage region for GAG biosynthesis may undergo a series of sulfation and phosphorylation reactions which may regulate the biosynthesis of GAGs.   The Xyl residue can be phosphorylated at its $C2$ position, and the two Gal residues may be sulfated at their $C4$ or $C6$ positions (figure 10) (Gulberti 2005).   In an *in vitro* assay,

GalTI has no activity towards the *C*2-phosphorylated xyloside, suggesting this

modification at Xyl may arrest the biosynthesis of some GAG chains (Jacquinet 2004).

On the other hand, the sulfated Gal only occurs in the tetrasaccharide linkage of CS but

not HS biosynthesis.   The studies concerning the substrate specificity of GlcATI also

demonstrated that sulfation on the two Gal residues influenced the activity of

glucuronyltransferase (Sato 2003; Seko 2003; Gulberti 2005).   Therefore, it is

reasonable to propose that the sulfated Gal may provide a regulatory mechanism to

control the rate of HS biosynthesis and may modulate the direction of the GAG assembly.
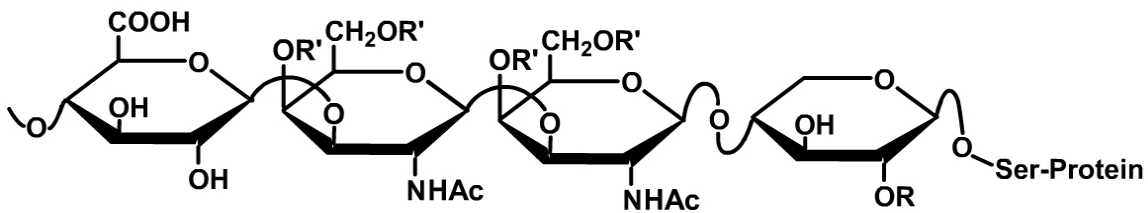
Figure 10. The tetrasaccharide linkage in the biosynthesis of heparan sulfate

R = H or PO$_3^-$, R' = H or SO$_3^-$ (SO$_3^-$ only occurs in CS)

The tetrasaccharide linkage domain (GlcA-β1,3-Gal-β1,3-Gal-β1,4-Xyl-β1-*O*-[Ser]) was sequentially built by four types of transferases.   XTs first place a Xyl to the core protein, followed by the attachment of two Gal residues by GalTI and GalTII, and a GlcA by GlcATI.

2. Chain elongation

Both HS and CS biosynthesis use the same tetrasaccharide linkage sequence that is attached to the core protein. However, the two routes diverge in the step in which the assembly of HS continues with the addition of GlcNAc by HS GlcNAc transferase I (GlcNAcTI) at the non-reducing end of the tetrasaccharide linkage. It was suggested that two mammalian enzymes, EXTL2 (Kitagawa 1999) and EXTL3 (Kim 2001), possessed GlcNAcTI activity. It is noteworthy that GlcNAcTI is an initiating GlcNAc transferase and different from HS GlcNAc transferase II (GlcNAcTII). GlcNAcTII adds GlcNAc to the growing HS chain instead of the tetrasaccharide linkage. *EXTL2* and *EXTL3* belong to the exostosin (*EXT*) gene family which includes *EXT1*, *EXT2* and three *EXT*-like genes *EXTL1*, *EXTL2* and *EXTL3*. Both EXTL2 and EXTL3 demonstrated GlcNAcTI activity towards a synthetic analogue (GlcA-β1,3-Gal-β1-*O*-naphthalenemethanol) which mimics the tetrasaccharide linkage and the hydrophobic domain in the core protein (Kim 2001). Interestingly, EXTL3 has also been shown to possess GlcNAcTII activity, indicating this enzyme may participate in both chain initiation and elongation processes (Kim 2001).

The D-GlcNAc residue is first attached to the tetrasaccharide linkage. This is followed by the alternate addition of GlcA and GlcNAc residues, which leads to the formation of HS backbone composed of 50-100 repeating disaccharide units (figure 9). The polymerization reactions may be catalyzed by EXT1 and EXT2 which are

bifunctional enzymes with both GlcATII and GlcNAcTII activities, noting that the GlcATII activity is distinct from the GlcATI activity in synthesizing the tetrasaccharide linkage.   Based on biochemical and genetic evidence, EXT1 and EXT2 are indispensable for each other to form a heterocomplex to maintain their enzymatic activities in the Golgi apparatus both *in vivo* and *in vitro* (McCormick 2000) (Kim 2003). Besides EXT1 and EXT2, it was also found that EXTL1 and EXTL3 have GlcNAcTII activity *in vitro* (Kim 2001).   One hypothesis is that EXTL1 and EXTL3 may regulate polymerization by competing with EXT1/EXT2 complex.   It is also possible that as there are a couple of redundant GlcNAcTII enzymes (EXT1/EXT2, EXTL1, EXTL3) available, the GlcATII reaction which adds GlcA to the growing sugar chain is a rate-determination step in HS elongation.

3. Backbone modifications

The length of HS backbone varies and is controlled by regulating the activities of EXT family enzymes.   The HS chain diversity is further complicated as the backbone can be modified by a series of HS biosynthetic enzymes, including *N*-deacetylase/*N*-sulfotransferase (NDST), *C*5-epimerase (C5-epi), 2-*O*-sulfotransferase (2OST), 6-*O*-sulfotransferase (6OST) and 3-*O*-sulfotransferase (3OST) as illustrated in figure 11.
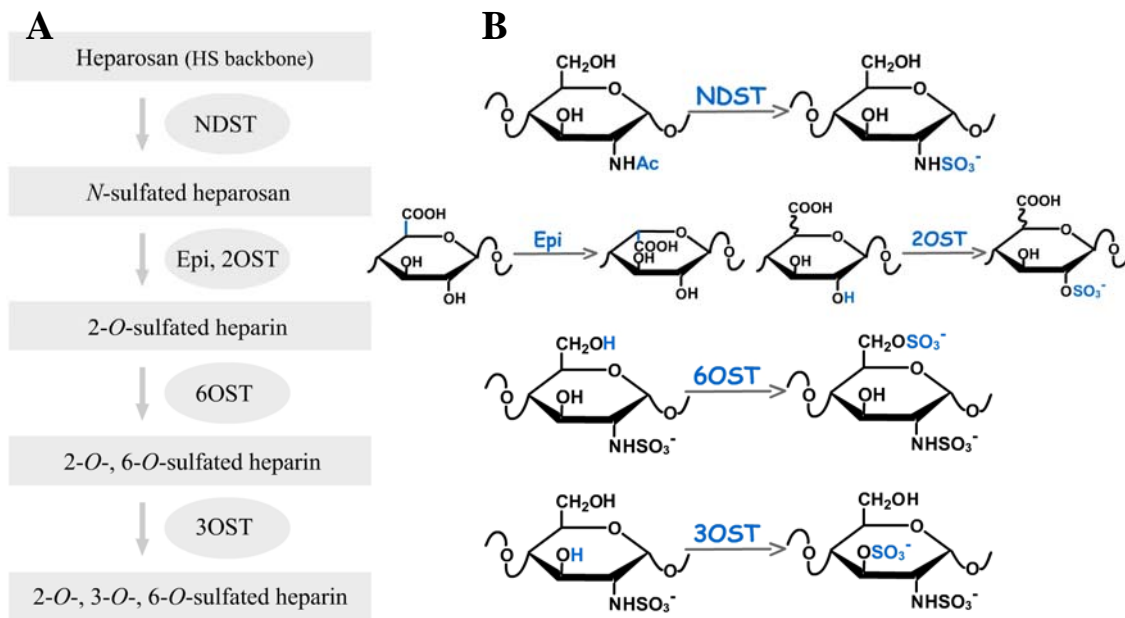
Figure 11. Heparan sulfate backbone modifications and its biosynthetic enzymes

A. HS backbone is modified by NDST, C5-epi, 2OST, 6OST and 3OST in sequence, leading to the HS structures with various sulfation patterns.

B. HS biosynthetic enzymes change the structure of HS with geometric specificities. The modified positions are shown in blue.

N-deacetylase/*N*-sulfotransferase

NDST converts the *N*-acetyl group to *N*-sulfate group in a GlcNAc residue. The modification by NDST is the first step of the backbone modifications. In fact, NDST is a dual-action enzyme with two distinct functions: *N*-deacetylation and *N*-sulfation (Perrimon 2000).

The conversion of GlcNAc to GlcNS is believed to be a prerequisite step for subsequent epimerization and *O*-sulfations, and is, therefore, a decisive step in HS modifications. Controlled by the activity of NDST, there are NS domains (GlcNS rich), NA domains (largely unmodified GlcNAc) and NS/NA (mixture of GlcNS and GlcNAc) domains in the HS polysaccharide chain. NDST has four isoforms, NDST1, NDST2, NDST3 and NDST4 with different enzymatic activity and expression patterns (Aikawa 2001). The redundancy and difference among NDST isoforms may lead to various sulfation patterns in the HS chain.

The polypeptide chain of NDST has a type II transmembrane topology and can be divided into four domains: a cytoplasmic region, a transmembrane region, a stem region, followed by the catalytic domain where the independent active sites for *N*-deacetylase and *N*-sulfotransferase activities are located. The truncated NDST enzymes with only *N*-deacetylase or *N*-sulfotransferase active site were constructed and they demonstrated respective enzymatic activities (Kakuta 1999; Duncan 2006). The transmembrane and stem regions of NDST are believed to help localize it in Golgi.

The crystal structure of the *N*-sulfotransferase domain of NDST1 was solved with adenosine-3', 5'-diphosphate (PAP) (figure 12) (Kakuta 1999). The overall structure of NDST1 is roughly spherical and contains an open cleft with a hydrophilic surface where a HS substrate as large as hexasaccharide can bind. The 5'-phosphate binding loop (PSB-loop) and α6 defines a cavity where adenosine 3'-phosphate 5'-phosphosulphate (PAPS) is bound. Superimposition of the crystal structures of NDST1 and estrogen sulfotransferase (EST) suggests that Lys-614 of NDST1 and Lys-48 of EST function similarly as a proton donor in the mechanism of the action of these two enzymes (Kakuta 1997) (Kakuta 1998). Lys-614 is believed to be critical in *N*-sulfotransferase reaction also because it is a conserved residue among other HS sulfotransferases and the results from site-directed mutagenesis confirmed that it is essential for NDST1 to be active (Sueyoshi 1998).
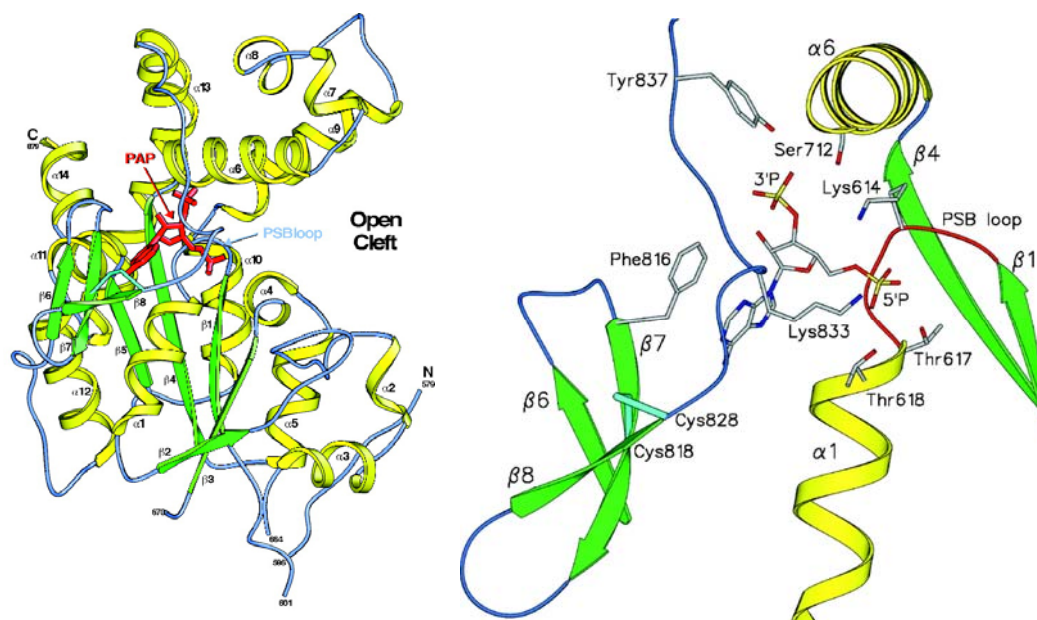
Figure 12. Overall structure and the PAPS-binding site of NDST1 (Kakuta 1999).

Unlike the *N*-sulfotransferase domain, less information about the *N*-deacetylase domain of NDST is known and this domain has hardly any homology to other carbohydrate deacetylases. Various isoforms of NDST are distinct in the activity of *N*-deacetylase, suggesting it may determine the substrate specificity and regulate the modifications of HS backbones. Indeed, various isoforms of NDST overexpressed in human embryonic kidney 293 cells individually resulted in the synthesis of HS with different sulfation patterns (Pikas 2000). Recently the region of NDST2 that carries the *N*-deacetylase activity was located to its *N*-terminal domain (Ala66-Pro604), which helps further elucidate the mechanism of the function of NDST (Duncan 2006).

Glucuronyl *C*5-epimerase

Following *N*-deacetylation and *N*-sulfation, glucuronyl *C*5-epimerase (*C*5-epi) acts on the GlcA residue and interconverts its *C*5 configuration between D-GlcA and L-IdoA in a fashion of a reversible two-step process (figure 13). A carbanion intermediate is first produced by deprotonation of *C*5 on the GlcA residue. This is followed by protonation again at the *C*5 position but stereochemistry is changed such that the carboxyl group is shifted across the plane of the sugar ring (Hagner-McWhirter 2000). As for substrate specificity, an adjacent GlcNS residue at the non-reducing end of the substrate (GlcA or IdoA) must be recognized by *C*5-epi (Esko 2001).
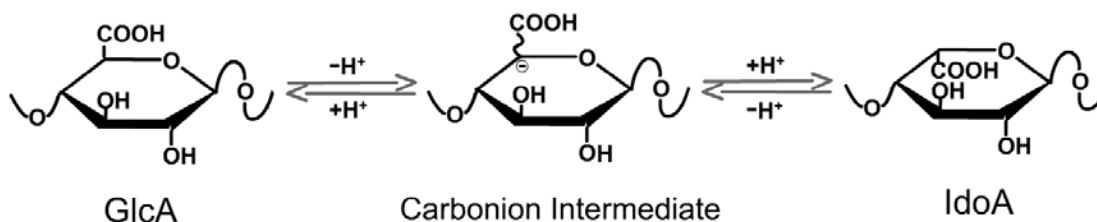
Figure 13. Reversible *C*5 epimerization between GlcA and IdoA catalyzed by *C*5-epi

IdoA residues promote the conformational flexibility of HS, because the pyranose ring of IdoA can adopt more conformations than that of GlcA does (Mulloy 2000). IdoA residues also complicate the structure diversity of HS. 2OST is believed to preferentially sulfate *C*2 on IdoA leading to the synthesis of HS with distinct sulfation patterns. Particularly, *C*5-epi and 2OST are proposed to work synergistically and form a complex *in vivo* (Pinhal 2001). In addition, 2-*O*-sulfated IdoA was resistant to conversion back to GlcA, indicating that *C*5-epi/2OST complex may achieve higher efficiency of modification (Jacobsson 1984).

To date only one isoform of *C*5-epi has been found in almost all species examined except for fish. Conserved C-terminal lysine residues in *C*5-epi may act as proton donors/acceptors in the reversible protonation/deprotonation reaction (Crawford 2001). Absence of IdoA in the HS chain in *C*5-epi deficient mice suggests that *C*5-epi is the only enzyme to produce HS containing IdoA and/or 2-*O*-sulfated IdoA (IdoA2S). *C*5-epi deficient mice afforded poorly developed lungs and died immediately after birth due to respiratory failure, suggesting that *C*5-epi is a critical enzyme involved in development (Li 2003).

Uronosyl 2-*O*-sulfotransferase

2-*O*-sulfation is the only sulfation occurring on the uronic acid residues in the HS chain and is important for the interactions between HS and a variety of proteins, such as, FGF (figure 6) (Tabeur 1999).   The reaction is catalyzed by the uronosyl 2-*O*-sulfotransferase (2OST) which acts with both IdoA and GlcA residues.   When provided with a substrate with both uronic acid units, 2OST, however, strongly favors the 2-*O*-sulfation of IdoA in line with its stronger affinity for IdoA-containing substrates (Rong 2001).

Similar to *C*5-epi, there is one isoform of 2OST in most species.   As the crystal structure of 2OST has not been solved, the mechanism of its action remains unclear. Nevertheless, its biological importance is clear from genetic studies in 2OST deficient mice which die due to defective kidney development (Merry 2001).   The different phenotypes between 2OST deficient mice and *C*5-epi deficient mice suggest that 2-*O*-sulfation in the HS chain has its particular biological effects (Li 2003).

Glycosaminyl 6-*O*-sulfotransferase

There are two positions (6-OH and 3-OH) on the glucosamine residues of HS chain available for *O*-sulfation.   Glycosaminyl 6-O-sulfotransferase (6OST) acts on 6-OH and transfers a sulfo group from PAPS to the glucosamine residue.   Three isoforms of 6OST (6OST1, 6OST2 and 6OST3) have been identified and they share 50-55% sequence homology.   Their sulfotransferase domains are more conserved and *N*- and *C*-termini are more variable (Habuchi 1998).   Some insights into the substrate specificity have been

gained from studies of recombinant enzymes using various *O*-desulphated poly- and oligo-saccharides as substrates (Habuchi 2000) (Smeds 2003). The HS containing -IdoA-GlcNS- seems to be a better substrate for 6OST1 than that containing -GlcA-GlcNS-, while 6OST2 and 6OST3 appear more promiscuous and utilize both disaccharide structures. Recent studies suggested that all the three isoenzymes can also modify GlcNAc units (Zhang 2001) (Smeds 2003). Although the three isoforms of 6OST have relatively similar substrate specificity, their expression patterns are variable. 6OST1 expression in liver is higher than other tissue, 6OST2 is predominantly found in brain and spleen, and 6OST3 is ubiquitous (Habuchi 2000).

6OST is important in HS modification as HS lacking 6-*O*-sulfation has no anticoagulant activity and loses its capability to trigger FGF signaling (Pye 1998). Knocking out both 2OST and 6OST disrupts the development of tracheal system in Drosophila, indicating 6OST is important for HS-mediated FGF activity and for maintaining the whole *O*-sulfation level of HS chain (Kamimura 2006).

Glucosaminyl 3-*O*-sulfotransferase

The final and most rare step in HS modifications is believed to be 3-*O*-sulfation catalyzed by 3-*O*-sulfotransferase (3OST). The 3-*O*-sulfation occurs predominantly in heavily sulfated regions of HS, because the enzyme interacts with numerous sulfate groups of the polysaccharide substrate (Liu 2002). As the last step in HS biosynthesis, it is conceivable that 3-*O*-sulfation directly determines the unique structures and functions of

HS, such as the AT-binding site and gD-binding site discussed previously.

There are totally seven isoforms of 3OST identified to date (Shworak 1999). 3OST1 and 3OST5 share 53% identity. 3OST2, 3OST3A, 3OST3B, 3OST4 and 3OST6 share about 70% identity. Particularly, 3OST3A and 3OST3B are nearly identical in their sulfotransferase domain (Liu 2002). The seven isoenzymes have different expression patterns. 3OST1 is highly expressed in heart, brain, lung and kidney. 3OST2 and 3OST4 are mainly expressed in brain. Both 3OST3A and 3OST3B are found in heart, lung, liver and kidney, but only 3OST3A exists in placenta and only 3OST3B is expressed in brain. 3OST5 is found in brain, spinal cord and skeletal muscle, while 3OST6 is expressed in liver and kidney (Shworak 1999).

3OST isoenzymes have respective substrate specificities with significant biological consequences (table 2). Particularly, 3OST1 generates GlcA-GlcNS3S±6S within the AT-binding structure (Shworak 1997) and 3OST2, 3, 4, and 6 produce IdoA2S-GlcNS3S±6S, the glycoprotein gD-binding structure to assist HSV infection (Liu 1999; Wu 2004). 3OST5 is promiscuous and generates a unique disaccharide structure IdoA-GlcNS3S±6S besides the two domains mentioned above (Xia 2002). The results from the disaccharide analysis of the 3OST2-modified HS also suggests that 3OST2 transfers sulfate to GlcA2S-GlcNS±6S (Liu 1999).

Substrate	Related 3OST	Product	Biological function
GlcA-GlcNS±6S	3OST1 and 3OST5	GlcA-GlcNS3S±6S	AT-binding/anticoagulant
IdoA2S-GlcNS±6S	3OST2,3,4,6 and 3OST5	IdoA2S-GlcNS3S±6S	gD-biding/HSV infection
IdoA-GlcNS±6S	3OST5	IdoA-GlcNS3S±6S	Unknown
GlcA2S-GlcNS±6S	3OST2	GlcA2S-GlcNS3S±6S	Unknown

Table 2. Substrate specificities of 3OSTs

The crystal structure of 3OST1 has been solved in a binary complex with PAP (figure 14) (Edavettal 2004). The complex is roughly spherical with a large open cleft and similar to the overall fold of the sulfotransferase domain of NST1 (Kakuta 1999). The large open cleft was suggested to be the HS-binding site. The structure also reveals residues in the PAPS binding site that are critical for binding with the 5'-phosphate and 3'-phosphate of PAPS via a number of hydrogen bonding interactions. In addition, site-directed mutagenesis analyses suggest that residues Arg-67, Lys-68, Arg-72, Glu-90, His-92, Asp-95, Lys-123, and Arg-276 are essential for the enzymatic activity of 3OST1 (Edavettal 2004).
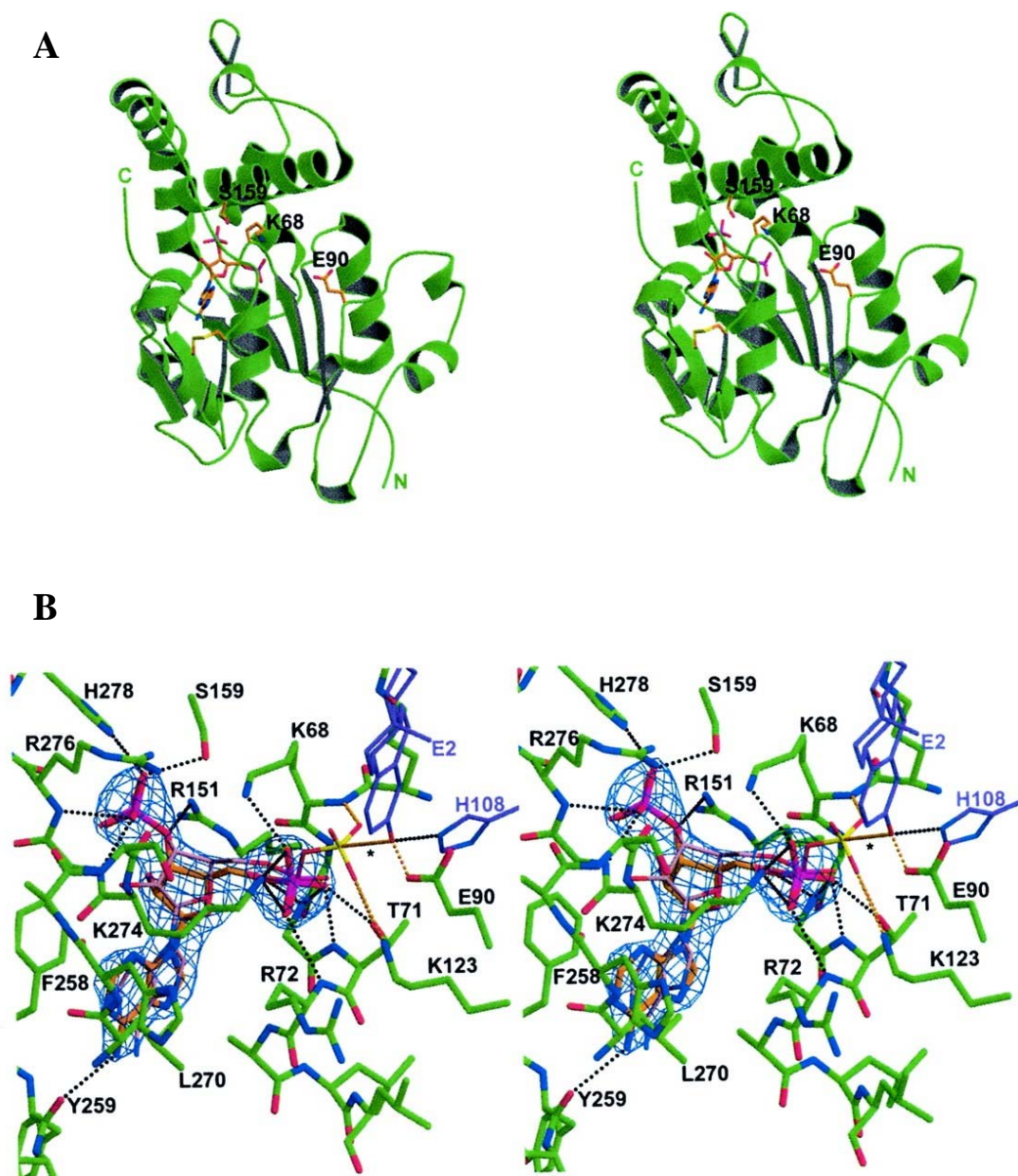
Figure 14. The crystal structure of 3OST1 (Edavettal 2004)

A. Stereo ribbon diagram of the structure of 3OST1

B. Stereo diagram of the PAPS-binding site of 3OST1 with PAP bound

The crystal structure of 3OST3 was also reported in a ternary complex with PAP and a tetrasaccharide substrate (Moon 2004) (figure 15A). The catalytic domain (Gly139-Gly406) of 3OST3 is spherical and remarkably similar to that of 3OST1. A major difference is the structure of the loop formed by the disulfide bond between Cys351 and Cys363 in strands 9 and 10 of the *C*-terminal sheet of 3OST3. The additional residues (Ala355-Gly357) in 3OST3 cause the loop to swing in the direction of the open cleft. However, replacing the disulfide bond loop of 3OST1 (Cys260-Cys269) with that of 3OST3 (Cys351-Cys363) does not alter its substrate specificity, indicating the difference in this loop does not directly contribute to the substrate recognition. Gln255 and Lys368 lie within the hydrogen bonding distances to the carboxyl and sulfate groups of the uronic acid of the tetrasaccharide acid. Mutational analysis suggests that these residues are critical for the substrate recognition of 3OST3. Glu184 is very close to the 3-OH group of the substrate and conserved among NSTs and 3OSTs. It is conceivable that it may function as a catalytic base to deprotonate the 3-OH for nucleophilic attack on the sulfate group (figure 15B). Indeed, mutations of this glutamate residue lead to a complete loss of the sulfotransferase activities for NST1 (Kakuta 2003) and 3OST3 (Edavettal 2004) and a 99.9% loss for 3OST3 (Moon 2004).

Figure 15. The crystal structure of 3OST3 (Moon 2004)

A. Ribbon diagram of the crystal structure of ternary complex 3OST3/PAP (blue)/tetrasaccharide (green) and sequence alignment of human 3OST3 and mouse 3OST1

B. Active site of 3OST3 with PAPS superimposed. Glu184 is shown in close contact with the 3-OH group of the glucosamine residue of the tetrasaccharide substrate

With site-directed mutagenesis, a recent study found that Ser120 and Ala306 in 3OST5 are critical to control its substrate specificity (Xu 2008). When one of these two residues or both were converted to the corresponding ones in 3OST1, 3SOT5 began to generate more AT-binding sites and less gD-binding sites, demonstrating the feasibility to tailor the substrate specificities of 3OST. In addition, the solved crystal structure of 3OST5 in this study (figure 16) (Xu 2008), along with the crystal structures of 3OST1 (Edavettal 2004) and 3OST3 (Moon 2004), have given insights in the mechanisms used by 3OST isoforms to dictate the substrate specificity. It has been proposed that the isoenzymes use two sites, the catalytic domain and the gate, to select appropriate sugar substrates. The catalytic domain is homologous among 3OST isoforms, while the gate residues are distinct and control the substrate specificity (Xu 2008).
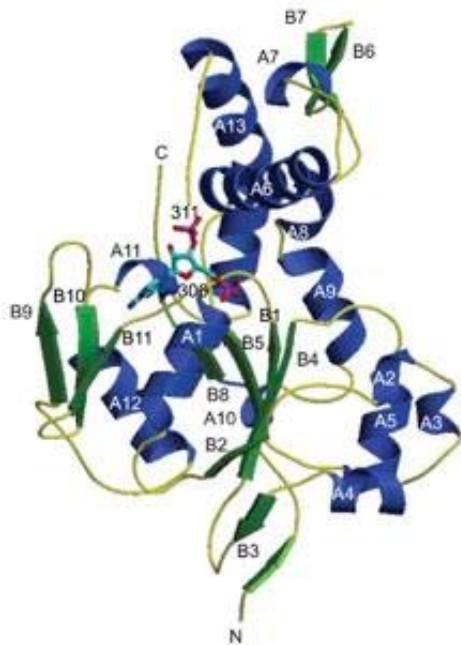
Figure 16. Ribbon diagram of the crystal structure of human 3OST5 with bound PAP (Xu 2008)

Using 3OST knock out mice is an important approach to study the biological functions of 3OST *in vivo*. The 3OST1 knock out mice was established showing profound phenotypes (HajMohammadi 2003). Although 3OST1 deficient mice suffered from postnatal lethality, they did not exhibit any different phenotype in anticoagulation, indicating that the anticoagulant HS is not a major hemostatic regulator like AT *in vivo*. Nevertheless, the specific postnatal lethality in 3OST1 deficient mice suggests that 3OST1 participates in some other profound biological processes.

Section 5.   Chemical and enzymatic synthesis of anticoagulant heparan sulfate

1. Chemical synthesis

The significant medical potential of HS makes it the most studied molecule among GAGs. Over the past twenty years, numerous syntheses of native HS, HS oligosaccharides and HS derivatives have been reported and reviewed (Gama 2005; Linhardt 2007 [a]; Linhardt 2007 [b]).   Chemical synthesis has been the major route to obtain structurally defined heparin and HS oligosaccharides.   Common strategies of chemical synthesis require efficient manipulation of protecting groups and stereoselective glycosylation, which relies on the appropriate leaving group at the anomeric (C1) position and the substituent at C2 position of the glycosyl donor (Karst 2003).   The synthesis of HS also depends on available tailor made monomers.   A "modular approach for HS synthesis" was

developed to prepare all combinations of monosaccharide building blocks (Haller 2001). Linhardt and colleagues used heparin lyase to depolymerize HS as an alternative chemoenzymatic approach for the preparation of various disaccharide building blocks (Yalpani 1996).

The most successful synthetic HS oligosaccharide has been Arixtra, a chemically synthesized AT-binding pentasaccharide (figure 5) marketed as an anticoagulant drug in both the United States and Europe (Vuillemenot 1999). Arixtra is a specific factor Xa inhibitor used in clinic to prevent venous thromboembolic incidents during surgery. Its derivative was also synthesized as a chemically more accessible analogue where *O*-sulfo groups replace the *N*-sulfo groups in the glucosamine residues and *O*-methyl groups are introduced to substitute hydroxyl groups (Westerduin 1994). However, Arixtra possesses lower anticoagulant activity than heparin because it is too short to harbor both anti-Xa and anti-thrombin activities. Longer oligosaccharides with more than sixteen sugar units are believed necessary to form an effective ternary complex between HS oligosaccharide, AT and thrombin (Lane 1984). Since the thrombin-binding domain is not restricted to any specific structures, the interaction between heparin and thrombin is believed to be non-specific and other linear oligosaccharides also have potential to bind thrombin (Olson 1991). Petitou and colleagues reported the synthesis of an HS analogue with seventeen saccharide units (figure 17) (Petitou 1999). To simplify the synthesis, hydroxyl groups were methylated and *N*-sulfo groups were replaced with *O*-sulfo groups in this compound. With regard to the concern about side effects, the

heptadecasaccharide does not interact with platelet factor 4 (PF4) which is related to heparin-induced thrombocytopaenia (HIT) (Petitou 1999).
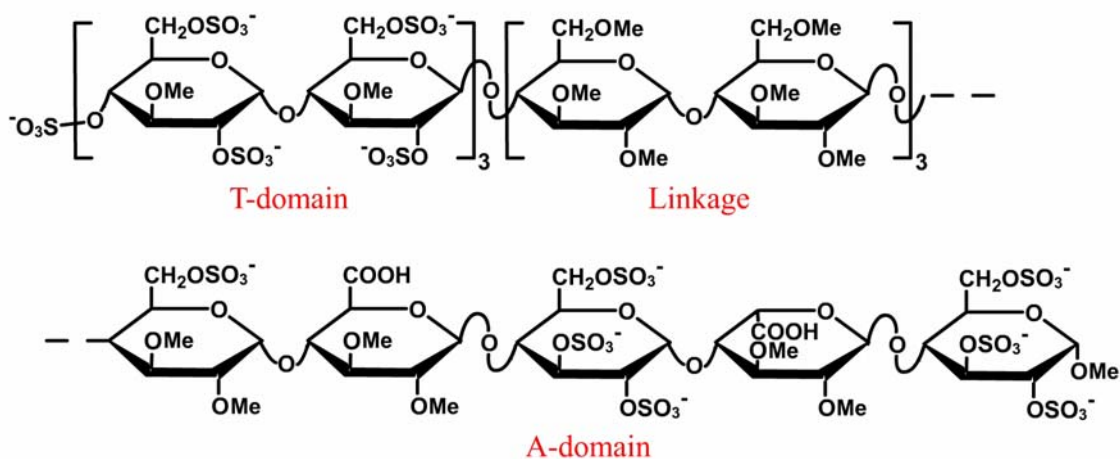
Figure 17. A heptadecasaccharide with both anti-Xa and anti-thrombin activities (Petitou 1999)

This heptadecasaccharide contains a highly sulfated thrombin-binding domain (T-domain), a neutral hexasaccharide linkage domain (Linkage) and an AT-binding domain (A-domain).

2. Enzymatic synthesis

Several mammalian and bacterial HS glycosyltransferases have been identified and harnessed as practical enzymatic tools to catalyze the polymerization of the activated sugar precursors.   These enzymes offer an alternative approach to synthesize HS backbone.   Busse and Kusche-Gullberg used recombinant soluble EXT1 and EXT1/2 complex on exogenous oligosaccharide acceptors derived from *Escherichia coli* K5 capsular polysaccharide (Busse 2003).   They demonstrated that EXT1 and EXT1/2 can act as HS polymerases without any other auxiliary proteins, while EXT2 incubation resulted in the addition of one GlcA residue but no further elongation.   Sismey-Ragatz and coworkers achieved *in vitro* polymerization of HS backbone with *Pasteurella Multocida* enzymes PmHS1 and PmHS2 (Sismey-Ragatz 2007).   Compared with PmHS2, PmHS1 was able to produce monodisperse polymers with large molecular mass (100-800 kDa) in a stoichiometrically controlled manner.   In contrast, PmHS1 was more promiscuous and was able to use many unnatural UDP-sugar analogs as donor substrates. Using bacterial glycosyltransferases to synthesize HS backbone becomes more attractive because these enzymes are readily expressed.

An enzymatic approach has been developed to synthesize anticoagulant HS from chemically desulfated *N*-sulfated (CDSNS) heparin (Chen 2005).   In this study, only three enzymatic modification steps were required to obtain the HS product with both anti-Xa and anti-thrombin activities in milligram scales.   A PAPS regeneration system

including arylsulfotransferase IV (AST-IV) was coupled with the sulfotransferase reactions catalyzed by 2OST, 6OST or 3OST (figure 18) (Burkart 2000; Chen 2005). This regeneration system is essential for practical syntheses of sulfated carbohydrate, because despite the high cost of PAPS, the product inhibition caused by PAP, even in low micromolar concentrations, is a major obstacle for the enzymatic synthesis with sulfotransferases (Lin 1995; Burkart 1999). With the PAPS regeneration system, the byproduct PAP produced from the sulfotransferase reaction is rapidly diminished and converted back to PAPS by AST-IV, meanwhile providing a sustained supply of PAPS.
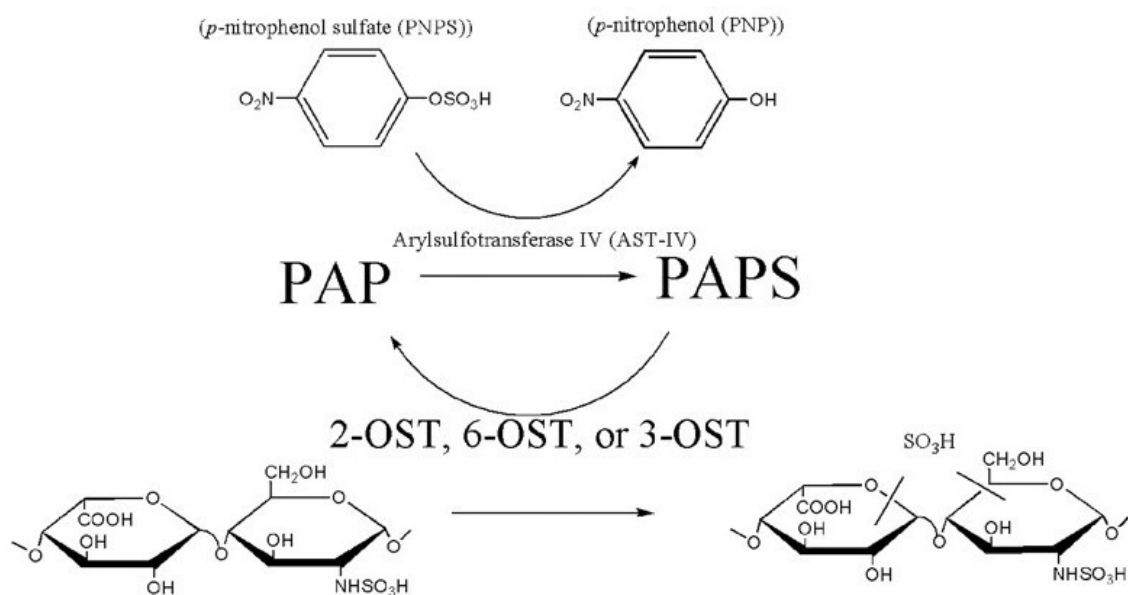
Figure 18. PAPS regeneration system (Chen 2005)

Arylsulfotransferase IV (AST-IV) uses PNPS as the sulfate donor to replenish the supply of PAPS for the sulfotransferase reactions catalyzed by HS sulfotransferases (2OST, 6OST or 3OST). The amount of PAP produced from the HS sulfotransferase reactions is also controlled as it is consumed by AST-IV to yield fresh PAPS.

The synthesis of structurally defined oligosaccharide has also been reported. Kuberan and colleagues proposed an enzymatic approach to simplify the synthesis of AT-binding pentasaccharide (Kuberan 2003). The method starts with heparitinase I-digested polysaccharide and requires five modification steps by *C*5-epi, 2OST, 6OSTs, glycuronidase and 3OST1. The resultant pentasaccharide was confirmed by LC-MS and gel mobility shift assay. In addition, Copeland and colleagues recently reported the synthesis of a 3-*O*-sulfated octasaccharide with 3OST3 (Copeland 2008). The octasaccharide blocks the entry and infection of HSV-1 by mimicking the gD-binding domain of the entry receptor. Collectively, these approaches demonstrate it feasible to employ current state-of-art enzymatic technologies in hand to assemble bioactive HS structures.

Section 6. Statement of problem

Heparin has been a commonly used anticoagulant drug for decades. Low-molecular-weight heparins (LMWHs) have also been approved for clinical use as anticoagulant agents with improved bioavailability and pharmacodynamics (Fareed 1998). However, as both heparin and LMWH are isolated from animal tissues, their structural heterogeneity may lead to multiple *in vivo* side effects, such as heparin-induced thrombocytopenia (HIT) and haemorrhages (Slaughter 1997). The synthesis of the

structure-defined anticoagulant HS is one of the major interests in the field of carbohydrate research.

Despite recent advances, the total chemical synthesis of HS-related structures remains challenging and has serious limitations (Linhardt 2007). For example, the chemical synthesis of Arixtra is a daunting task that requires as many as sixty steps with a yield of less than 0.5% (Sinaÿ 1984). With most of the important HS biosynthetic enzymes identified and cloned, enzymatic synthesis has become a valuable complement to the chemical approach. Enzymatic synthesis of AT-binding pentasaccharide was first accomplished by degrading K5 polysaccharides and utilizing a set of seven HS biosynthetic enzymes (Kuberan 2003). However, in this synthesis, the extent of polysaccharide *N*-deacetylation and digestion is hard to precisely control and only a small part of starting materials remained useful for downstream epimerization and *O*-sulfation. In addition, the substrate specificities of HS biosynthetic enzymes become a barrier to apply this method to HS oligosaccharides of other size and structures. In our study, we proposed another enzymatic route towards *de novo* synthesis of structurally defined HS oligosaccharides.

The primary goal of this project is to clone and express the heparosan polymerases in bacteria cell lines for the synthesis of HS backbone. Particularly, *E. coli* glucuronyl transferase KfiA is carefully characterized and tested for its potential in polymerizing unnatural sugar residues. With KfiA and PmHS2 available in large amount, both natural (-GlcA-GlcNAc-) and unnatural (-GlcA-GlcNCOCF$_3$-) HS backbones were synthesized

step by step from a starting disaccharide derived from *E. coli* K5 polysaccharide. Finally, we synthesized oligosaccharides with defined complete *N*-sulfation with the HS biosynthetic enzyme NST. We envision that, by driving each downstream enzymatic modification step (*C*5-epi/2OST/6OST/3OST) to completion, these oligosaccharides with controlled *N*-sulfation can be used as the starting substrates to synthesize bioactive HS oligosaccharides with defined structures and desired sulfation patterns.

Collectively, our research will help understand the biosynthesis of HS and simplify the syntheses of anticoagulant oligosaccharides. It may also open the access to the enzymatic synthesis of HS oligosaccharides with defined structures which should aid in establishing HS structure-function relationships in the future.

Chapter II. Materials and methods

Section 1. Expression and purification of proteins

<u>Preparation of bacterial genomic DNA (gDNA)</u>

The bacterial gDNA was prepared using Easy-DNA kit (Invitrogen).   A small overnight culture (2 mL) of *E. coli* K5, K12 or *P. multocida* type D (American Type Culture Collection, ATCC) was harvested and suspended in 200 μL of 1× PBS and mixed well with 350 μL of Solution A (Easy-DNA kit, Invitrogen).   After the incubation at 65 $^{\circ}$C for 10 minutes, 150 μL of Solution B (Easy-DNA kit, Invitrogen) was added and the mixture was vortexed until the sample was uniformly viscous.   This was followed by adding 500 μL of chloroform and a vigorous vortex until the sample was homogeneous.   The sample was then centrifuged at maximum speed for 15 min to separate phases.   The upper aqueous phase was recovered and gDNA was precipitated with 1 mL of 100% ethanol (-20 $^{\circ}$C).   The gDNA pellet was washed with 500 μL of 80% ethanol (-20 $^{\circ}$C) and air dried.   The gDNA was resuspended in 100 μL of sterile ddH$_2$O and treated with 2 μL of RNase (2 mg/mL, Easy-DNA kit, Invitrogen) at 37 $^{\circ}$C for 30 min.   The bacterial gDNA was then ready for further experiments.

<u>KfiA</u>

The KfiA DNA sequence was amplified from *E. coli* K5 (ATCC) genomic DNA using a

pair of primers: 5'-ATATATAA<u>GGATCC</u>GATGATTGTTGCAAATATGTC and 5'-ATAATATA<u>CTCGAG</u>CCCTTCCACATTATACAC (the cleavage sites of BamHI and XhoI are underlined).   The gene was cloned into the pET21b+ vector (Novagen) to yield the *C*-terminal 6×histidine-tagged protein.

Expression of KfiA was achieved in BL21 star (DE3) cells (Invitrogen) carrying pGro7 plasmid (Takara, Japan) that expresses chaperone proteins GroEL and GroES. Briefly, cells containing the plasmids expressing KfiA and GroEL/GroES were grown in LB medium supplemented with 50 μg/mL carbenicillin and 35 μg/mL chloramphenicol at 37 $^{\circ}$C.   When A_{600} reached 0.6, the temperature was decreased to 22 $^{\circ}$C, and arabinose (1 mg/mL) and isopropyl-β-D-thiogalactopyranoside (IPTG, 1mM) was added to induce the expression of chaperone and KfiA, respectively.   The culture was shaken overnight at 22 $^{\circ}$C.   The bacteria were harvested and lyased by sonication in a buffer containing 25 mM Tris-HCl (pH=7.5), 500 mM NaCl and 30 mM imidazole.   The protein was purified with Ni-Sepharose 6 Fast Flow (0.75 x 10 cm, Amersham), which was eluted with a linear gradient of 30 mM to 250 mM of imidazole in 500 mM NaCl and 25 mM Tris-HCl (pH 7.5) in 60 mL.   The purity of the resultant protein was analyzed by the precast 10% SDS-PAGE (BioRad).   Briefly, samples (10 μl) were diluted with an equal volume of sample buffer containing 200 mM Tris-HCl, pH=6.8, and 2% SDS (BioRad).   Gels were run at 110V for 1 hour, and then stained with coomassie blue (0.4%) for 30 mins.   Gels were destained using 10% acetic acid.

GlmU

The glucosamine-1-phosphate acetyltransferase / *N*-acetylglucosamine-1-phosphate uridyltransferase (GlmU) gene was amplified from *E. coli* K12 (ATCC) genomic DNA with the primers: 5'-ATATATAA<u>CATATG</u>TTGAATAATGCTATG and 5'-AATTATAA<u>GGATCC</u>TCACTTTTTCTTTACCGGAC and cloned into pET21b+ vector using NdeI and BamHI sites (underlined in the primer sequences) to prepare the *C*-terminal 6 × histidine-tagged GlmU protein. The expression was carried out in BL21 star (DE3) cells. The fusion protein was purified with Ni-Sepharose 6 Fast Flow resins as described above.

KfiA/KfiC coexpression

The full-length KfiA gene with *C*-his tag was first amplified from the pET21b+ vector prepared above with the primers: 5'-ATATTAT<u>CCATGG</u>ATGATTGTTGCAAATATGTC and 5'-ATATTTAAT<u>GCGGCCGC</u>TCAGTGGTGGTGGTGGTGGTGC. It was then cloned into the first multiple cloning site of the pCDFDuet-1 vector (Novagen) using NcoI and NotI sites (underlined in the primer sequences). The KfiC DNA sequence was amplified from *E. coli* K5 gDNA using a pair of primers: 5'-ATATATAA<u>GCGGCCGC</u>AATGAACGCAGAATATATA and 5'-AATTATAA<u>CTCGAG</u>TTGTTCAATTATTCCTGA (the cleavage sites of NotI and XhoI are underlined). The amplified KfiC gene was then cloned into the pET21b+ vector (Novagen).

The pET21b+ vector carrying KfiC gene was first transformed into the BL21 star (DE3) cells (Invitrogen) with chaperone proteins as used in the expression of KfiA.   A 50 mL culture of this cell was grown in the antibiotics-free LB medium until A_{600} reached 0.5.   The cell culture was chilled on ice for 20 min and harvested in sterile tubes.   The cell pellet was then resuspended in 5 mL of ice-cold sterile transformation and storage solution containing 10% polyethylene glycerol (PEG), 50 mM $MgCl_2$ and 5% dimethyl sulfoxide (DMSO) in LB medium.   Aliquots of the cell suspension were placed into small tubes by 50 μL each.   A second transformation of the pCDDuet-1 vector carrying KfiA gene into this competent cell was performed to yield the BL21 star cells coexpressing KfiA, KfiC and chaperone proteins.   The cell culture was prepared and harvested similarly as described above.   The *C*-His tagged KfiA/KfiC complex was again purified with Ni-Sepharose 6 Fast Flow (0.75 x 10 cm, Amersham).

PmHS2

PmHS2 was expressed as an *N*-terminal fusion to 6×histidine (DeAngelis 2004).   The PmHS2 DNA sequence was first amplified from *P. multocida* type D (P-3881, ATCC) genomic   DNA   using   a   pair   of   primers: 5'-ATGAAGAGAAAAAAAGAGATGACTCAAA   and 5'-TTATAAAAAATAAAAAGGTAAACAGGG without restriction enzyme sites.   The PCR product was purified by Qiaquick PCR purification kit (Qiagen) and used as the template for a secondary polymerase chain reaction (PCR) with another pair of primers

with restriction sites: 5'-ATATATAA<u>CATATG</u>ATGAAGAGAAAAAAAGAGATGACTC and 5'-AATTATAA<u>GGATCC</u>ATCATTATAAAAAATAAAAAGGTAAACAGG (the cleavage sites of NdeI and BamHI are underlined). The amplified PmHS2 gene was then cloned into the pET15b vector (Novagen) and PmHS2 was expressed in BL21 star (DE3) cells (Invitrogen) with chaperone proteins as described above.

The bacteria were harvested and lyased by sonication in a buffer containing 25 mM Tris-HCl (pH=7.5), 500 mM NaCl and 30 mM imidazole. The protein was purified with Ni-Sepharose 6 Fast Flow (0.75 x 10 cm, Amersham), which was eluted with a linear gradient of 30 mM to 300 mM of imidazole in 500 mM NaCl and 25 mM Tris-HCl (pH 7.5) in 60 mL. The purity of the resultant protein was analyzed by the precast 10% SDS-PAGE (BioRad) as described above.

NST

The expression of NST was reported previously (Sueyoshi 1998). A plasmid containing the sulfotransferase domain of human NDST-1 fused to glutathione S-transferase was obtained from Dr. Masahiko Negishi (NIH). The enzyme was expressed, bound to a glutathione column (Amersham Pharmacia) for purification, and cleaved from glutathione S-transferase using thrombin as described.

C5-epi

The catalytic domain of human *C*5-epimerase (E53-N609) was cloned into pMAL-c2X

vector (New England Biolabs) using the BamH1 and HindIII sites to generate a maltose-binding protein (MBP)-epimerase fusion protein (Muñoz 2006).   The full length cDNA of *C*5-epimerase was a gift from Dr. Rosenberg (Massachusetts Institute of Technology).   Expression of *C*5-epimerase was achieved in Origami-B cells (Novagen) carrying pGro7 plasmid (Takara, Japan) which expresses chaperone proteins GroEL and GroES.   The bacteria were grown in LB medium supplemented with 2 mg/ml glucose, 12.5 μg/ml tetracycline, 15 μg/ml kanamycin, 35 μg/ml chloramphenicol, and 50 μg/ml carbenicillin at 37 °C.   When the OD 600 nm reached 0.6–0.8, IPTG and L-arabinose were added to a final concentration of 0.1 mM and 1 mg/ml, respectively.   The protein was purified following the protocol provided by the manufacturer.

2OST, 6OST1 and 6OST3

As described in our previous publication (Chen 2005), the catalytic domains of 2OST of Chinese hamster ovary (Arg51–Asn356) and mouse 6OST1 (His53–Trp401) were cloned into the pMAL-c2X vector (New England BioLabs) using the BamHI and HindIII sites to generate maltose-binding protein (MBP)-2OST and MBP-6OST1 fusion proteins.   The full-length cDNAs of 2OST and 6OST1 were gifts from Dr. Rosenberg (Massachusetts Institute of Technology, Cambridge, MA) and Dr. Kimata (Aichi University, Japan), respectively.   Expression of 2OST and 6OST1 was achieved in Rosetta-gami B (DE3) cells (Novagen) using a standard procedure.   Briefly, cells containing the plasmid expressing 2OST or 6OST1 were grown in LB medium supplemented with 2 mg/ml

glucose, 15 µg/ml tetracycline, 15 µg/ml kanamycin, 35 µg/ml chloramphenicol, and 50 µg/ml carbenicillin at 37 °C.   When the A600 reached 0.6–0.8, IPTG (1 mM) was added, and the cells were shaken overnight at 22 °C.   The bacteria were harvested, and the proteins were purified by following a protocol from the manufacturer (New England Biolabs).

The expression of 6OST3 was also carried out in *E. coli* (Chen 2007).   In brief, the catalytic domain of mouse 6OST3 (Pro121–Pro450) was cloned into a pMalc2x vector (New England BioLab) from a mouse brain Quick clone cDNA library (BD Biosciences). The expression was carried out in Origami-B cells (Novagen) carrying pGro7 (Takara, Japan) plasmid expressing chaperone proteins GroEL and GroES of E. coli. Transformed cells were grown in 6 L LB medium supplemented with 2 mg/ml glucose, 12.5 µg/ml tetracycline, 15 µg/ml kanamycin, 35 µg/ml chloramphenicol, and 50 µg/ml carbenicillin at 37°C.   When the A600 reached 0.4–0.7, IPTG (0.15 mM) and L-arabinose (1 mg/ml) were added to induce the expression of 6OST3 and chaperone proteins, respectively. The cells were allowed to shake overnight at 22 °C.   Purification was carried out with an amylose-agarose (New England BioLab) column by following the protocols provided by manufacturer.

3OST1

As described previously (Edavettal 2004), the catalytic domain of 3OST1 (Gly48-His311) was amplified from m3OST1-pcDNA3 with a 5' overhang containing an NdeI site and a

3' overhang containing an EcoRI site.   This construct was inserted into the pET28a vector (Novagen) using the NdeI and EcoRI restriction sites to produce a His6-tagged protein. The resultant plasmid (b3OST1-pET28) was sequenced to confirm the reading frame and the lack of mutations within the coding region (University of North Carolina, DNA sequencing core facility).   The plasmid, b3OST1-pET28, was transformed into BL21(DE3)RIL cells (Stratagene) for the expression of 3OST1.   Cells containing the b3OST1-pET28 were grown in 12 2.8-liter Fernbach flasks containing 1 liter of LB media with 50 µg/ml kanamycin at 37 °C.   When the A600 reached 0.6-0.8, the temperature was lowered to 22 °C for 15 min.   IPTG was then added to a final concentration of 200 µM, and the cells were allowed to shake overnight. Cells were pelleted and resuspended in 120 ml of sonication buffer, 25 mM Tris, pH 7.5, 500 mM NaCl, and 10 mM imidazole.   Cells were disrupted by sonication then spun down. The supernatant was applied to nickel nitrilotriacetic acid-agarose resin (Qiagen) in batch and washed with sonication buffer.   The resin was loaded onto a column, and the protein was eluted with an imidazole gradient from 10 to 250 mM.   The protein was dialyzed then concentrated to 16 mg/ml in 20 mM Tris, pH 7.5, 100 mM NaCl, and 4 mM PAP.   A total of 28 mg of protein was obtained.

Section 2. Preparation of acceptor substrates

Preparation of heparosan from *E. coli* K5

One liter of an overnight culture of *E. coli* K5 was prepared and spun to recover the supernatant only, which was then mixed with equal volume of K5P buffer A containing 50 mM NaCl and 20 mM NaAcO at pH 4. After pH was adjusted to 4, the sample was loaded onto 10 mL of diethylaminoethyl (DEAE) column preequilibrated with K5P buffer A. The DEAE column was extensively washed with 100 mL of K5P buffer A and then eluted with K5P buffer B containing 1 M NaCl and 20 mM NaAcO at pH 4. K5 polysaccharide was precipitated from eluent after mixed with three volumes of ethanol at -20 °C overnight. The K5 polysaccharide was then dissolved in water and mixed with 1.7 mL of phenol/chloroform to purify the polysaccharide from the protein impurities. K5 polysaccharide was precipitated from the aqueous part again with the addition of three volumes of ethanol and incubation at -20 °C overnight. Typically 100 mg of K5 polysaccharide may be obtained from 1 L of the bacterial culture.

The concentration of K5 polysaccharide was determined by estimating the amount of the resultant disaccharide from a complete digestion of K5 polysaccharide with heparin lyase III. Briefly, K5 polysaccharide (10 µL) was mixed with heparin lyase III (10 µg) and incubated at 37 °C for 3 hours. The resultant disaccharide (ΔUA-GlcNAc) was resolved with reverse-phase ion-pairing HPLC (RPIP-HPLC) and quantified by comparing the area of a UV_{232} peak with the same disaccharide standard (Seikagaku) with known amount.

Preparation of K5 polysaccharide-derived oligosaccharides with various sizes

Even-numbered oligosaccharides were prepared from a purified capsular polysaccharide of *E. coli* K5, with the structure (GlcA-GlcNAc)$_n$. The K5 polysaccharide (5.5 mg) mixed with ^{3}H-labeled K5 polysaccharide (80,000 cpm) was partially *N*-deacetylated by treating it with 2 M NaOH at 68 °C for 60 min followed by deamination with nitrous acid at pH 4.5 (Shaklee 1984). The resultant oligosaccharides were reduced with sodium borohydride to yield a reducing terminal residue of 2, 5-anhydromannitol (anMan$_R$). The oligosaccharides (GlcA-[GlcNAc-GlcA]$_n$-anMan$_R$) were size-fractionated by a Bio-Gel P10 (Bio-Rad) column, which was eluted with 20 mM Tris-HCl (pH7.4) and 1 M NaCl. [^{3}H]radioactivity was determined. Those fractions corresponding to appropriate sizes of oligosaccharides ranging from disaccharides to octasaccharides were pooled separately (figure 16), dialyzed against double deionized water for 12 hours at 4 °C using 3,500 Da MWCO membrane (Fisher Scientific) and lyophilized. The amount of the purified oligosaccharides was determined based on the specific [^{3}H]radioactivity of the polysaccharide. Particularly, the disaccharide (GlcA-anMan$_R$) was synthesized by deamination of the completely *N*-deacetylated K5 polysaccharide by treating the K5 polysaccharide with 2 M NaOH at 68 °C for more than 3 hours.

Section 3. Preparation of donor substrates

UDP-GlcN[^{3}H]Ac

UDP-GlcN[^{3}H]Ac was prepared from sodium [^{3}H]acetate (500 mCi/mMole, MP Biomedicals) using GlmU following a method described previously (figure 18) (Leiting 1998). The reaction mixture contained Tris-HCl, pH 7.4 (50 mM), sodium [^{3}H]acetate (0.2 mM, 0.1 μCi/μL, 500 mCi/mMole), MgCl$_2$ (5 mM), dithiothreitol (DTT) (0.2 mM), UTP (2.5 mM), glucosamine-1-phosphate (GlcN-1-P) (2.5 mM), ATP (10 mM), CoASH (0.02 mM), acetyl coenzyme A synthetase (0.1 mg/mL, Sigma-Aldrich, St. Louis, MO), recombinant GlmU (0.4 mg/mL) and inorganic pyrophosphatase (0.01 U/μL, Sigma-Aldrich, St. Louis, MO). The reaction was incubated at 30 ^{o}C overnight with mild agitation. The yield of UDP-GlcN[^{3}H]Ac was monitored with RPIP-HPLC as demonstrated by a [^{3}H]-labeled peak coeluted with unlabeled UDP-GlcNAc standard (Sigma) that has absorbance at 256 nm. The specific [^{3}H]radioactivity of UDP-GlcN[^{3}H]Ac was estimated to be 172 mCi/mMole. UDP-[^{14}C]GlcA (100 μM, 20 μCi/mL, 200 mCi/mMole) was obtained from PerkinElmer Life and Analytic Sciences (Boston, MA).

UDP-GlcNCOCF$_3$

UDP-GlcNCOCF$_3$ was synthesized in a chemoenzymatic way as shown in figure 26. The reaction precursor GlcN-1-phosphate was bought from Sigma-Aldrich. 11 mg of GlcN-1-phosphate was dissolved in 200 μL of anhydrous methanol and mixed with 60 μL

of $(C_2H_5)_3N$ and 130 µL of $CF_3CO\text{-}SC_2H_5$. The reaction was incubated at room temperature for 24 hours and dried to recover the resultant $GlcNCOCF_3$-1-phosphate. About 0.5 mg of $GlcNCOCF_3$-1-phosphate was incubated with Tris-HCl 7.4 (46 µM), $MgCl_2$ (5 mM), dithiothreitol (DTT, 200 µM), UTP (2.5 mM), recombinant GlmU (1.5 mg/mL), and inorganic pyrophosphatase (0.012 U/µL) in a reaction of 200 µL. The product was purified by HPLC using a home-made amino column packed with YMC*GEL Amino 12 nm S-15µm (YMC). The column was eluted with a linear gradient of 0 mM to 1000 mM of KH_2PO_4 in 60 min at a flow rate of 1 mL/min. The synthesized $UDP\text{-}GlcNCOCF_3$ was eluted with 500 mM KH_2PO_4.

Section 4. Determination of the binding of KfiA to various donor substrates with Isothermal Titration Calorimetry (ITC)

ITC was performed on a MicroCal VP-ITC (MicroCal, Northampton, MA) (Edavettal 2004). Solutions were degassed under vacuum before use. Purified KfiA was dialyzed against a buffer containing 25 mM Tris, 250 mM NaCl, 10 mM $MnCl_2$, pH 7.5 (Tris buffer) at 4 °C overnight. Experiments were conducted in the Tris buffer using 50 µM KfiA (1.35 mg/ml). Titrations were performed by injecting 5 µL of 1.5 mM UDP and UDP-sugars (UDP-GlcNAc, UDP-GlcA, UDP-glucose and UDP-GalNAc) in 25 mM Tris buffer (250 mM NaCl, 10 mM $MnCl_2$, pH 7.5). The experimental data was analyzed with Origin Version 7 which was preinstalled in MicroCal VP-ITC.

Section 5. Glycosyltransferase assays on KfiA, KfiA/KfiC and PmHS2

KfiA

N-acetyl D-glucosaminyl transferase assays were carried out with purified KfiA and size-defined oligosaccharide acceptors.   The assay was designed to determine the incorporation of a GlcN[^{3}H]Ac residue to the defined oligosaccharide substrates.   The reaction mixture contained 15,000 cpm of UDP-GlcN[^{3}H]Ac, 100 µg/mL KfiA, 1% Triton X-100, 25 mM Tris-HCl 7.2, 10 mM MnCl$_2$, and the appropriate oligosaccharide acceptor (Hodson 2000).   Each reaction sample was mixed with 100 µL of 330 mM tetrabutylammonium phosphate monobasic (Fluka) and subject to the analysis by the reverse-phase ion-pairing HPLC (RPIP-HPLC).   A Beckman Coulter System Gold Microbore HPLC system with 32 Karat software (Beckman Coulter, Fullerton, CA) was used to perform all HPLC assays.   A C$_{18}$-column (0.46 × 25 cm) (Vydac) was used to resolve the resultant [^{3}H]oligosaccharides and preequilibrated in a buffer containing 38 mM ammonium phosphate monobasic, 2 mM phosphoric acid, and 1 mM tetrabutylammonium phosphate monobasic (Fluka).   The column was isocratically eluted at 0.5 mL/min with 0% acetonitrile at 0 min, 7.5% acetonitrile at 10 min, 15% acetonitrile at 25 min, and 30% acetonitrile at 40 min, and 50% acetonitrile at 50 min. The eluent was monitored by an on-line radioactive detector (β-ram, INUS) for measuring radioactivity and an on-line System Gold 166 detector (Beckman Coulter, Fullerton, CA) for measuring UV absorbance at the wavelength of 260 nM.

The enzyme kinetics of the *N*-acetyl D-glucosaminyl transferase activity of the

purified KfiA was evaluated by monitoring the incorporation of GlcN[^{3}H]Ac. Reactions were carried out at oligosaccharide acceptor substrate concentration varying from 3 μM to 800 µM in a buffer of enzymatic assay described above. To measure the K_m and K_{cat}/K_m of UDP-GlcNAc, we conducted the reactions with the disaccharide (700 µM) as an acceptor substrate. To study the influence of KfiC on KfiA kinetics, enzymatic reactions towards disaccharide substrates were also prepared with KfiA at 2.7 µM and KfiC at 1.5 µM. Because the reaction rate remained constant up to 6 min, the average reaction rate of the first 2 min represented the initial velocity. The reactions were incubated at 37 ^{o}C for 2 min and quenched by heating to 100 ^{o}C. The products were resolved by a C$_{18}$-column (0.46 × 25 cm) (Vydac) under the RPIP-HPLC elution condition (Liu 1999). Briefly, the column was isocratically eluted with different concentrations of acetonitrile depending on the size of oligosaccharide product in a solution containing 38 mM ammonium phosphate monobasic, 2 mM phosphoric acid, and 1 mM tetrabutylammonium phosphate monobasic (Fluka) at a flow rate of 0.5 ml/min. For the trisaccharide product, no acetonitrile was present in the elution buffer; for penta- and hepta-saccharide, 7.5% acetonitrile was present in the elution buffer; for the nonasaccharide, 15% acetonitrile was present in the mobile phase. The eluent was monitored by an on-line radioactive detector (β-ram, INUS) for measuring [^{3}H]radioactivity.

KfiA/KfiC

The glucuronyl transferase assay on KfiA/KfiC complex was carried out with KfiA/KfiC

and a [3H]-labeled heptasaccharide acceptor prepared from a KfiA reaction. The assay was designed to monitor the incorporation of a GlcA residue to [3H]-labeled heptasaccharide and the resultant [3H]-labeled octasaccharide with RPIP-HPLC using the same HPLC method as described above. The reaction mixture contained 15,000 cpm of [3H]-heptasaccharide, 50 μM UDP-GlcA, 100 μg/mL KfiA/KfiC, 1% Triton X-100, 25 mM Tris-HCl 7.2, 10 mM $MgCl_2$ and 10 mM $MnCl_2$ (Hodson 2000).

PmHS2

The glycosyltransferase activities of PmHS2 were tested in a different fashion. The reactions contained 15,000 cpm of UDP-GlcN[3H]Ac (for *N*-acetyl D-glucosaminyl transferase test) or UDP-[14C]GlcA (for glucuronyl transferase test), 100 μg/mL PmHS2, 1% Triton X-100, 25 mM Tris-HCl 7.2, 10 mM $MnCl_2$, and an even-numbered oligosaccharide acceptor with a GlcA at the non-reducing end (such as a disaccharide, for *N*-acetyl D-glucosaminyl transferase test) or an odd-numbered oligosaccharide with a GlcNAc at the non-reducing end (such as a trisaccharide, for glucuronyl transferase test). The resultant [3H]-labeled trisaccharide or [14C]-labeled tetrasaccharide were determined on a Bio-Gel P10 (Bio-Rad) column, which was eluted with 20 mM Tris-HCl (pH7.4) and 1 M NaCl.

Section 6. Synthesis and purification of heparan sulfate backbones

Disaccharides were prepared from the complete deacetylation and degradation of K5

polysaccharides as described above and were used as the starting materials in the *de novo* synthesis of HS backbones. A small scale radioactive KfiA reaction with disaccharide as acceptor and UDP-GlcN[^{3}H]Ac as donor was incubated at 37 ^{o}C for 1 hour and then loaded onto a Bio-Gel P10 column. The P10 column was used to purify the synthesized oligosaccharides based on their size. The elution time/position of the [^{3}H]trisaccharide on the P10 column was recorded as a standard. A large scale KfiA reaction containing 4.5 μmole (1.52 mg) of disaccharide as acceptor in excess and 3.9 μmole (2.58 mg) of UDP-GlcNCOCF$_3$ as donor was also prepared in a buffer with 25 mM Tris-HCl (pH 7.2), 10 mM MnCl$_2$, 10 mM MgCl$_2$ in the presence of KfiA (0.1 mg/mL). The reaction product was loaded onto the P10 column where the elution fractions corresponding to the trisaccharide standard were pooled, dialyzed against double deionized water for four hours (MWCO 1000, Spectrum), and dried.

The purified trisaccharide was used as donor for the next radioactivity-free PmHS2 reaction containing 2.87 μmole (1.71 mg) of trisaccharide, 3 μmole (1.74 mg) of UDP-GlcA in excess, 25 mM Tris-HCl (pH 7.2), 10 mM MnCl$_2$, 10 mM MgCl$_2$ and 0.1 mg/mL PmHS2. The trisaccharide was also used in a small scale radioactive PmHS2 reaction with UDP-[^{14}C]GlcA as donor. The resultant [^{14}C]-tetrasaccharide was again resolved on a Bio-Gel P10 column and its elution time/position was recorded as the standard to blindly purify radioactivity-free tetrasaccharide from the large scale cold synthesis.

Additional rounds of KfiA and PmHS reactions were repeated until the synthesized

oligosaccharides reached the size we desired.   The enzymatic synthesis of natural HS backbones containing GlcNAc was carried out in the same way as mentioned above.

Section 7. Mass spectrometry

Electrospray ionization mass spectrometry (ESI-MS)

The purified oligosaccharides were further desalted by dialysis against doubled deionized water.   The samples were then dried with a vacuum concentrator, reconstituted in doubled deionized water (50 µM, 200 µL) and introduced by direct infusion (10 µL/min) to ESI-MS (Agilent 1100 MSD-Trap at the Mass Spectroscopy Core in the UNC School of Pharmacy).   Experiments were performed at negative ionization mode with a source temperature of 350 $^{\circ}$C.   Nitrogen was used as a drying (5 liters/min) and nebulizing gas (15 p.s.i.).

Liquid chromatography-mass spectrometry (LC-MS)

LC-MS analyses were performed on Agilent 1100 LC/MSD instrument (Agilent Technologies, Inc. Wilmington, DE, U.S.A.) equipped with an ion trap, binary pump and a UV detector at the Rensselaer Polytechnic Institute.   The column was a 5 µm Agilent Zorbax SB-C18 (0.5 × 250 mm) from Agilent Technologies. Eluent A was water/acetonitrile (85:15), v/v and eluent B was water/acetonitrile (35:65) v/v.   Both eluents contained 12 mM tributylamine (TBA) and 38 mM NH_4OAc and their pH was

adjusted to 6.5 with HAcO. The sample was diluted 10-fold with water and 1 μl of the resulting analyte was injected by auto-sampler. A gradient of 0% B for 15 min, and 0-100% B over 85 min. was used at a flow rate of 10 μl/min. Mass spectra were obtained using an Agilent 1100 series Classic G2445D LC/MSD trap (Agilent Technologies, Inc. Wilmington, DE, U.S.A.). The electrospray interface was set in negative ionization mode with the skimmer potential -40.0 V, capillary exit -120 .5 V and a source temperature of 325 °C to obtain maximum abundance of the ions in a full scan spectra (150–1500 Da, 10 full scans/s). Nitrogen was used as a drying (5 liters/min) and nebulizing gas (20 p.s.i.). Auto tandem mass spectrometry (MS/MS) was turned on in these experiments using an estimated cycle time of 0.07 min. Total ion chromatograms (TIC) and mass spectra were processed using Data Analysis 2.0 (Bruker software).

Section 8.   Modifications of heparan sulfate backbones

_N_-deacetylation (removal of _N_-trifluoroacetyl group)

200 μg of the purified _N_-trifluoroacetyl oligosaccharides were dried and resuspended in a solution of 200 μL containing methanol, ddH$_2$O and triethylamine (v/v/v=2:2:1). The reaction was incubated at room temperature for 20 hours. The sample was then dried by a vacuum concentrator (Contrivap Concentrator, Labconco) and reconstituted in ddH$_2$O to recover the _N_-deacetylated oligosaccharides.

Partial and complete *N*-sulfation of [³H]heptasaccharide by NST

PAPS was directly used as the sulfate donor in a 50-µL NST reaction which contained 16 µM [³H]heptasaccharide, 50 mM MES (pH 7), 0.1% Triton X-100, 10 mM $MnCl_2$, 500 µM PAPS and 0.36 mg/mL NST.   The reaction was incubated at 37 °C and sampled at 2 hours and overnight.   3 µL of each sample was analyzed with RPIP-HPLC which was preequilibrated with buffer A (38 mM ammonium phosphate monobasic, 2 mM phosphoric acid and 1 mM tetrabutylammonium) and eluted at 0.5 mL/min.   The isocratic elution started at 100% of buffer A (or 0% of buffer B) (38 mM ammonium phosphate monobasic, 2 mM phosphoric acid, 1 mM tetrabutylammonium, and 50% acetonitrile), switched to 30% of buffer B at 10 min, 60% at 25 min, and finally changed to 100% of buffer B at 40 min.   The monosulfated and disulfted heptasaccharide were eluted at 27 min and 35 min, respectively.

Sulfotransferase reactions coupled with a PAPS regeneration system

A PAPS regeneration system was used to scale up the sulfation reactions (figure 18) (Burkart 2000).   Briefly, a 1-mL regeneration reaction consisted of 50 mM MES (2-(*N*-morpholino)ethanesulfonic acid) (pH 7), 40 µM PAP, 1 mM paranitrophenol sulfate (PNPS), 50 µl purified arylsulfotransferase-IV (AST-IV), and 10 µg of the purified HS sulfotransferase (NST, 2OST, 6OST1, 6OST3 or 3OST1) was incubated at room temperature for 15 mins.   The coupling efficiency was monitored by measuring the concentration of PNP at A_{410}.   Once equilibrium was reached, 50 µg of the oligosaccharide substrate was added and the mixture was rotated at room temperature

overnight.

Purification of the fully N-sulfated [^{3}H]oligosaccharides with HPLC

The synthesized NS-[^{3}H]heptasaccharide was purified with a C-18 column (Vydac).

The same RPIP-HPLC method was used as for analysis of the partially N-sulfated

[^{3}H]heptasaccharide mentioned above, except that 35% of buffer B was applied at 10 min.

NS-[^{3}H]heptasaccharide was eluted at around 30% acetonitrile (60% buffer B).

The synthesized NS-[^{3}H]undecasaccharide was purified with a polyamine II column

(YMC), which was eluted with a linear gradient of 0 mM to 1000 mM KH_2PO_4 in 60 min

at 0.5 mL/min on HPLC. NS-[^{3}H]undecasaccharide was eluted at 680 mM KH_2PO_4.

Nitrous acid degradation of the fully N-sulfated [^{3}H]oligosaccharides

The extent of N-sulfation of the synthesized NS-[^{3}H]heptasaccharide and

NS-[^{3}H]undecasaccharide was examined by high pH and low pH nitrous acid

degradation.

In the nitrous acid degradation at high pH (pH 4.5), about 150 ng (10000 cpm) of the

[^{3}H]-labeled NS-oligosaccharide was dried and redissolved in 100 µL of the mixture of

0.5 M H_2SO_4 and 5.5 M $NaNO_2$ (v/v=2:5). The sample was incubated on ice for 30 min

and occasionally vortexed. The degradation was stopped by adding 60 µL of the

solution containing 1 M Na_2CO_3, 1 M $NaHCO_3$ and ddH_2O (v/v/v=3:5:5). The sample

was then mixed with 0.5 µL of 5% phenol red and loaded onto a Bio-Gel P10 column for

analysis.

In the nitrous acid degradation at low pH (pH 1.5), about 150 ng (10000 cpm) of the [^{3}H]-labeled NS-oligosaccharide was dried and reconstituted in 20 µL of the ddH$_2$O. 20 µL of 0.5 M H$_2$SO$_4$ and 20 µL of 0.5 M Ba(NO$_2$)$_2$ was mixed and spun. The supernatant was added to the oligosaccharide sample, which was then incubated on ice for 30 min with occasional vortex. A 20-µL solution containing 1 M 1 M Na$_2$CO$_3$ and 1 M NaHCO$_3$ (v/v=7:3) was added to quench the degradation. The sample was then mixed with 0.5 µL of 5% phenol red and loaded onto a P10 column for analysis.

Modification of the *N*-sulfated octasaccharide in a one-pot reaction

The *N*-sulfated octasaccharide backbone was modified by HS biosynthetic enzymes in a one-pot 100-µL reaction. Briefly, 200 ng of the *N*-sulfated octasaccharide was first treated with C5-epi (0.1 mg/mL) in the presence of 25 mM MES (pH 7), 0.5% Triton X-100, and 1 mM CaCl$_2$ at 37 ^{o}C for 1 hour. The other HS biosynthetic enzymes were then added, including 2OST (0.4 mg/mL), 6OST1 (0.3 mg/mL), 6OST3 (0.5 mg/mL), 3OST1 (0.1 mg/mL), along with PAP[^{35}S] (25 µM, 19000 cpm/µL). A negative control reaction was also prepared with every enzyme included except for 3OST1. Both reactions were incubated at 37 ^{o}C overnight and boiled for 1 min to stop the reaction.

Determination of the binding of the modified octasaccharides to antithrombin (AT)

Both the 3-*O*-sulfated octasaccharide and the modified octasaccharide without

3-*O*-sulfation were tested for AT-binding with a concanavalin A (ConA)-based assay.

Approximately 260 ng of [^{35}S]-labeled compound were incubated with 5 µg of human AT (Cutter Biological) in 50 µl of binding buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM Mn^{2+}, 1 mM Mg^{2+}, 1 mM Ca^{2+}, 15 µM dextran sulfate, 0.003% Triton X-100 for 1 hour at room temperature. Concanavalin A-Sepharose (Sigma; 50 µl of 1:1 slurry) was then added, and the reaction was shaken at room temperature for 4 h. The beads were then washed by 3 × 1 ml of binding buffer, and the bound polysaccharide was eluted with a buffer containing 20 mM Tris-HCl (pH 7.5) and 1 M NaCl.

Chapter III. Determination of the substrate specificities of

N-acetyl-D-glucosaminyl transferase KfiA

Section 1. Introduction

E. coli K5 capsular polysaccharide consists of GlcNAc and GlcA repeating disaccharide, an analogue of unsulfated and unepimerized HS backbone (Vann 1981). Using the enzymes involved in the biosynthesis of K5 polysaccharide offers an alternative approach to prepare HS backbones (figure 9). Previous studies (Hodson 2000) have indicated that the biosynthesis of K5 polysaccharide requires an *N*-acetyl-D-glucosaminyl transferase, KfiA. This enzyme, together with D-glucuronyl transferase KfiC (Griffiths 1998), alternatively incorporates GlcNAc and GlcA residues, from UDP-GlcNAc and UDP-GlcA to the nonreducing ends of the oligo- or polysaccharides (figure 9). Although KfiA is of interest to synthesize HS backbones, the substrate specificity of this enzyme is still poorly understood. The lack of information about its biochemical features is primarily due to the unavailability of the purified enzyme. In this chapter, an efficient way to prepare purified KfiA in multi-milligram quantities is described. Substrate specificities of KfiA were also characterized. The results from this study demonstrated the possibility of the synthesis of HS backbone from a disaccharide structure using KfiA.

Section 2. Expression and purification of KfiA

KfiA was previously identified as carrying *N*-acetyl glucosaminyl transferase activity responsible for the synthesis of capsular polysaccharide in *E. coli* K5 using a site-directed mutagenesis approach in a cell based assay (Hodson 2000).   However, the biochemical character of this enzyme has not been fully elucidated because of the unavailability of purified enzyme.   The potential application of KfiA in synthesizing defined size HS backbone structures prompted us to investigate its donor and acceptor substrate specificities.   The full length KfiA gene was cloned from *E. coli* K5 genomic DNA and expressed in a format of *C*-terminal 6×His fusion protein as described under "Materials and Methods".   Recombinant KfiA was purified by affinity chromatography using a nickel column.   The purified KfiA predominantly migrated as one major band at ~27 KDa on a 10% SDS-PAGE (figure 19), which is nearly identical to the calculated molecular weight of KfiA (27.3 KDa).   We estimated that the purity of KfiA was greater than 90%.   Approximately 10 mg of purified KfiA was obtained from 1 liter of bacterial culture.
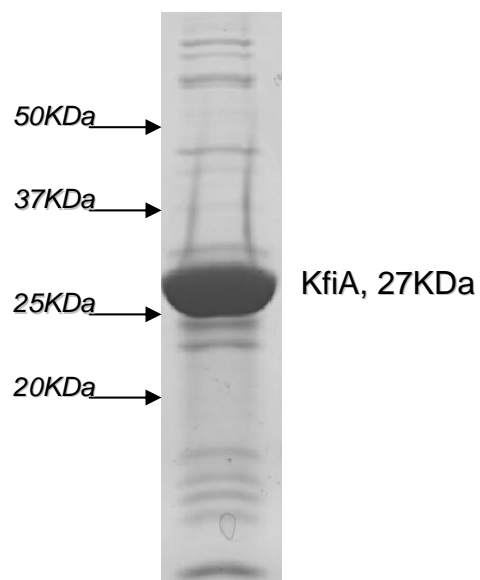
Figure 19. The purified KfiA on SDS-PAGE

Affinity chromatography purified KfiA protein was resolved on precasted 10% SDS-PAGE.   The gel was stained with coomassie blue.   Migration positions of molecular markers (from BioRad) are indicated.   The apparent mass was approximately 27 KDa. The predicted molecular mass of KfiA is 27.3 KDa.

Section 3. Preparation of even-numbered oligosaccharide acceptors

To study the size specificity of the acceptor for KfiA, even-numbered oligosaccharides were prepared from K5 polysaccharide by partial *N*-deacetylation followed by deaminative cleavage. To monitor the polysaccharide during the purification, unlabeled K5 polysaccharide was mixed with [^{3}H]metabolically labeled K5 polysaccharide. The polysaccharide was deacetylated under a basic condition, and the resultant polysaccharide was subject to nitrous acid (pH 4.5) degradation. The oligosaccharide products with the structures of GlcA-[GlcNAc-GlcA]$_n$-anMan$_R$ were resolved on Bio-Gel P10 column based on the size of the oligosaccharides, such as di-, tetra-, hexasaccharide, etc (figure 20).
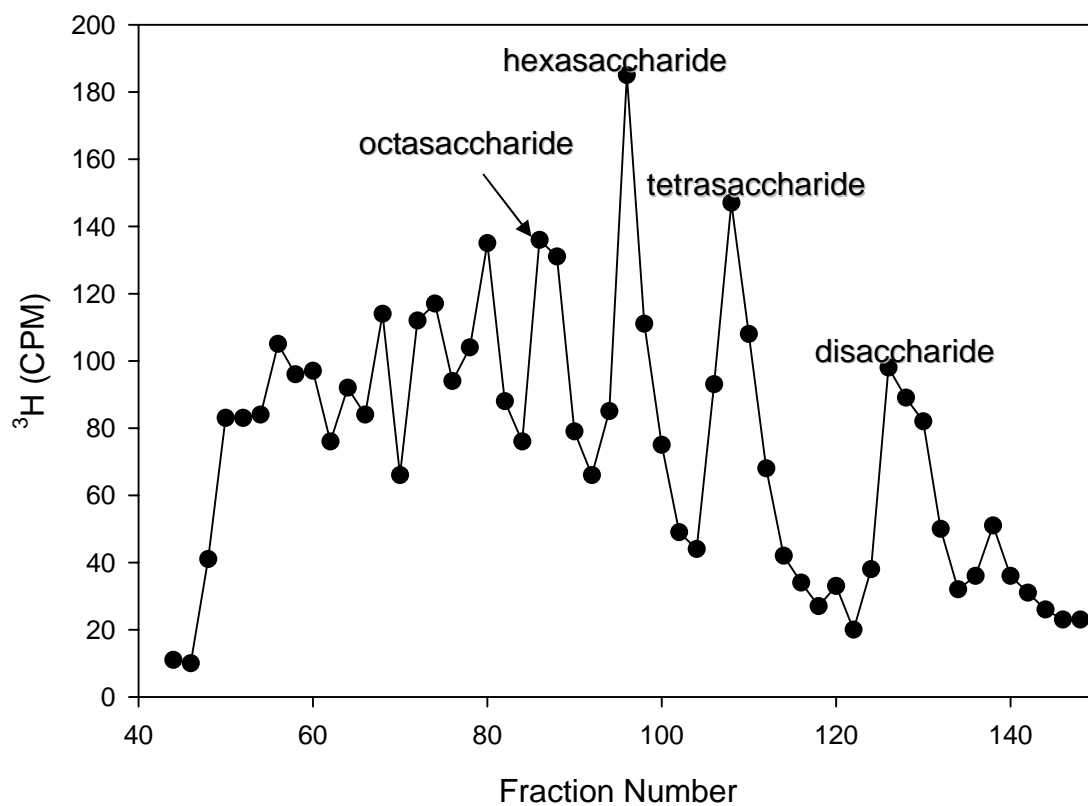
Figure 20. Fractionation of partially depolymerized K5 polysaccharide on P10 column.

K5 polysaccharide was partially digested as described under "Materials and Methods". The products were fractionated (9 drops/fraction) by gel chromatography on a Bio-Gel P10 column eluted with 20 mM Tris-HCl 7.4 and 1 M NaCl. The resultant oligosaccharides (from disaccharide to octasaccharide) were desalted and reconstituted in double deionized water for mass spectrometry analysis and enzymatic assays.

Due to the resolution of the BioGel P-10 column, we were limited to obtaining up to octasaccharide with sufficient purity (>90%). The purified oligosaccharides were also subjected to the analysis of electrospray ionization mass spectrometry (ESI-MS) to confirm the molecular weight. The MS spectra of those analyzed oligosaccharides are shown in figure 21. Furthermore, we did not observe the oligosaccharide with different sizes in a given sample, suggesting that the purified oligosaccharides are sufficiently pure (with purity >90%) for the kinetic studies as described below. The molecular weights of the purified oligosaccharides were consistent with the anticipated size of the oligosaccharide (table 3).

Sample	Formula[a]	Calculated Mr (Da)	Detected Mr (Da)	Detected molecular ion (m/z)
disaccharide	$C_{12} H_{20} O_{11}$	340.28	340.01	$[M-H]^-$ 339.0
tetrasaccharide	$C_{26} H_{41} O_{22} N$	719.60	719.25 ± 0.05^b	$[M-2H]^{2-}$ 358.6 $[M-2H+Na]^-$ 740.3
hexasaccharide	$C_{40} H_{62} O_{33} N_2$	1098.91	1098.84 ± 0.40^b	$[M-3H+Na]^{2-}$ 559.6 $[M-3H+2Na]^-$ 1141.4
octasaccharide	$C_{54} H_{83} O_{44} N_3$	1478.23	1477.46 ± 0.01^b	$[M-4H+2Na]^{2-}$ 759.7 $[M-4H+3Na]^-$ 1542.4

Table 3. Molecular weight (Mr) of the purified oligosaccharides determined by ESI-MS

a. Molecular masses were calculated by Molecular Weight Calculator (http://ncrr.pnl.gov/software/).
b. The value is the average ± range of the detected molecular ions.
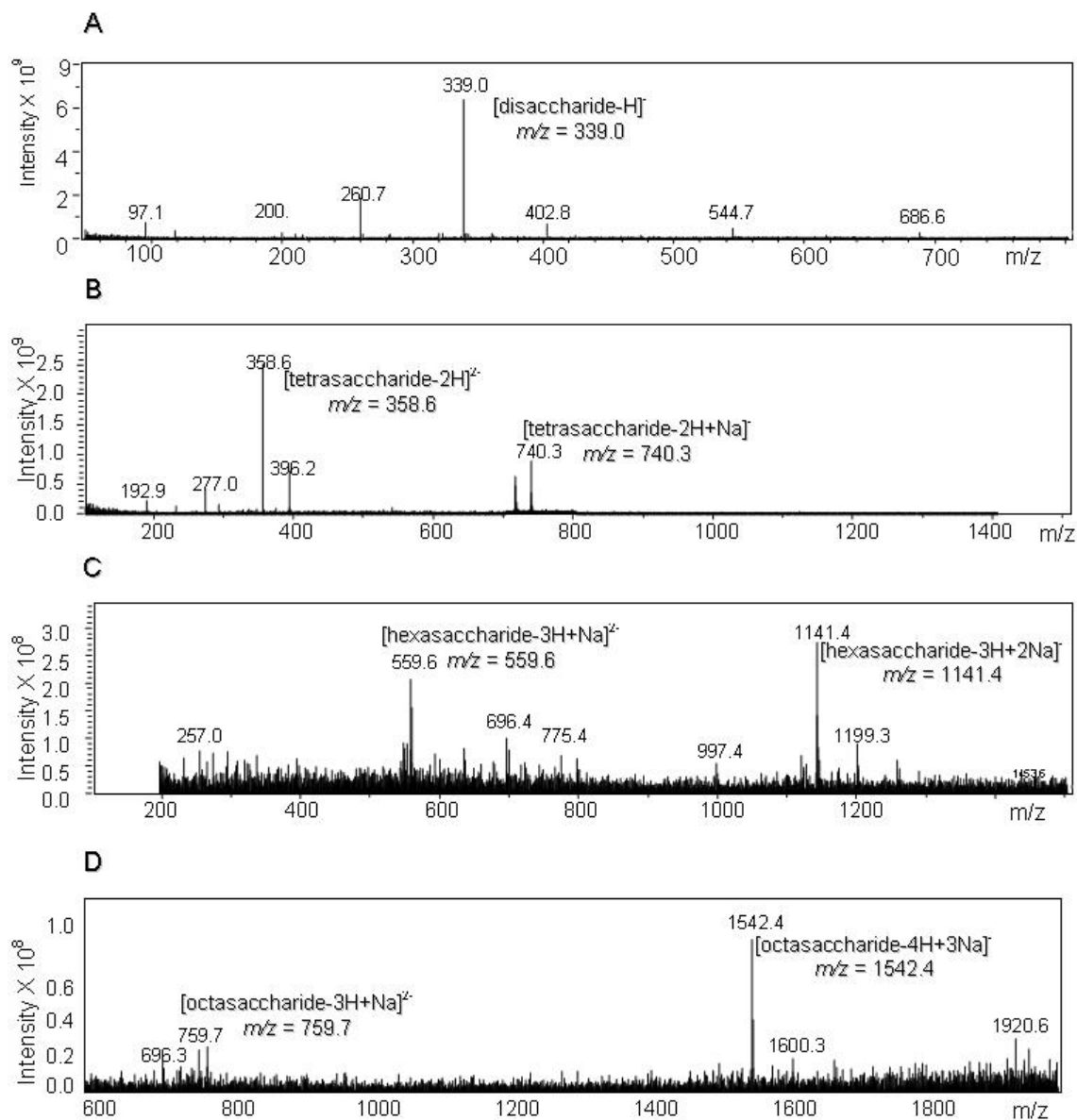
Figure 21. Electrospray ionization mass spectra of purified oligosaccharides.

Panel A, B, C, and D shows the ESI-MS spectrum of the disaccharide, tetrasaccharide, hexasaccharide, and octasaccharide, respectively.   The samples were introduced by direct infusion (10 μL/min) to ESI-MS.   The expected anions are indicated.

Section 4. Synthesis of [³H]-labeled UDP-*N*-acetylglucosamine (UDP-GlcN[³H]Ac)

Since radio-labeled UDP-sugars are prohibitively expensive and the stability of commercial UDP-GlcN[³H]Ac was not reliable, I synthesized UDP-GlcN[³H]Ac using a method modified from the previous publication (Leiting 1998) (figure 22). [³H]Acetate was used to produce [³H]acetyl CoA by acetyl coenzyme A synthetase. Bifunctional enzyme GlmU was cloned from *E. coli* genome and expressed in bacteria to synthesize GlcN[³H]Ac-1-phosphate and finally produce UDP-GlcN[³H]Ac.
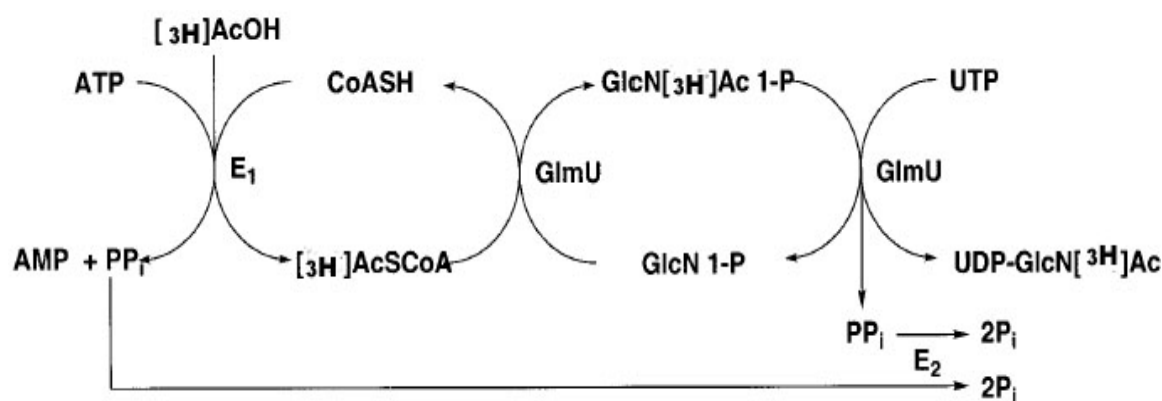
Figure 22. Enzymatic synthesis of UDP-GlcN[³H]Ac (Leiting 1998)

E1, acetyl coenzyme A synthetase
E2, inorganic pyrophosphatase

Section 5. Determination of the enzymatic activity of KfiA

The activity of KfiA was determined using a purified tetrasaccharide as an acceptor, and the product was resolved on RPIP-HPLC. A [3H]-labeled peak eluted at 18 min was observed (solid line, figure 23A), while this [3H]-labeled peak was absent in the control (without KfiA enzyme, dotted line, figure 23A), suggesting that a 3H-labeled pentasaccharide was generated due to the action of KfiA. It should be noted that more than 90% of the UDP-GlcN[3H]Ac had been converted to the 3H-labeled pentasaccharide within 30 min of initiating the reaction under the standard assay condition, demonstrating that recombinant KfiA was highly active. This pentasaccharide was purified by RPIP-HPLC, and its molecular mass was determined to be 922.6 Da ([M-2H]$^{2-}$ = 460.3, Mr = 922.6 Da) by ESI-MS (figure 23B), which is almost identical to the calculated molecular mass of the anticipated pentasaccharide (922.8 Da).
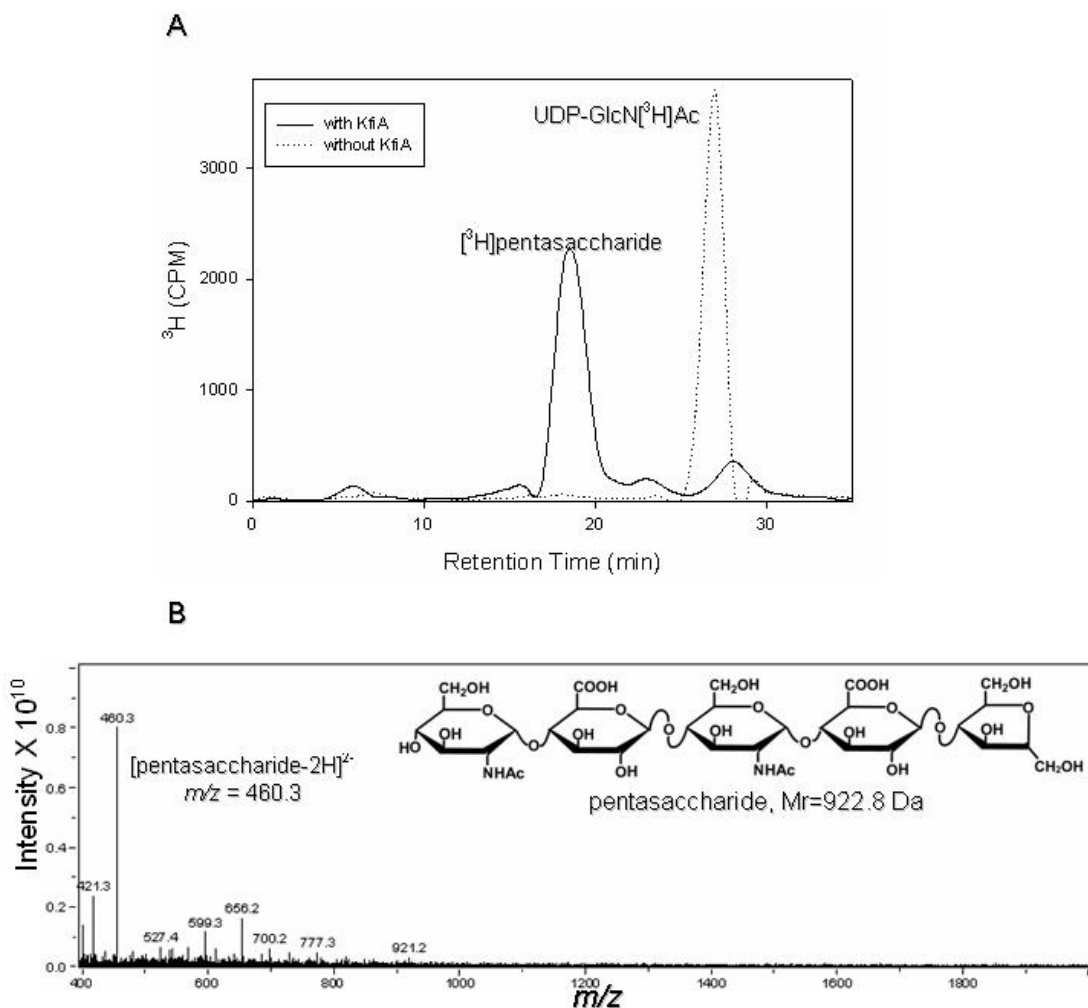
Figure 23. RPIP-HPLC chromatograms and ESI-MS spectrum of KfiA-modified tetrasaccharide acceptor.

Panel A shows the RPIP-HPLC chromatograms of KfiA-modified tetrasaccharide. UDP-GlcN[³H]Ac was incubated with tetrasaccharide acceptors as described in "Materials and Methods" with (solid line) or without KfiA (dotted line).   The reaction products were separated on RPIP-HPLC eluted with 50% acetonitrile in a solution containing 38 mM ammonium phosphate monobasic, 2 mM phosphoric acid, and 1 mM tetrabutylammonium phosphate monobasic at a flow rate of 0.5 ml/min.

Panel B shows the ESI-MS spectrum of KfiA-modified product.   The pentasaccharide was prepared by incubating the tetrasaccharide, KfiA and unlabeled UDP-GlcNAc.   The product was purified by RPIP-HPLC, knowing the eluted position of [3H]pentasaccharide.   The resultant pentasaccharide was dialyzed and analyzed by ESI-MS.   A doubly charged anion (m/z = 460.3), corresponding to the pentasaccharide molecule, was detected.   Thus, the determined Mr was 922.6 (Da), which is consistent with the calculated Mr of the pentasaccharide (922.8 Da).

Section 6. KfiA transfers *N*-acetyl glucosamine to the acceptors with different sizes

The requirement for the size of the substrate for KfiA was investigated by incubating the enzyme and the acceptor substrates with various sizes in the presence of UDP-GlcN[^{3}H]Ac, and the [^{3}H]-labeled products were analyzed by RPIP-HPLC (figure 24). It shows that KfiA is capable of transferring the GlcNAc residue to the acceptor as small as disaccharide.
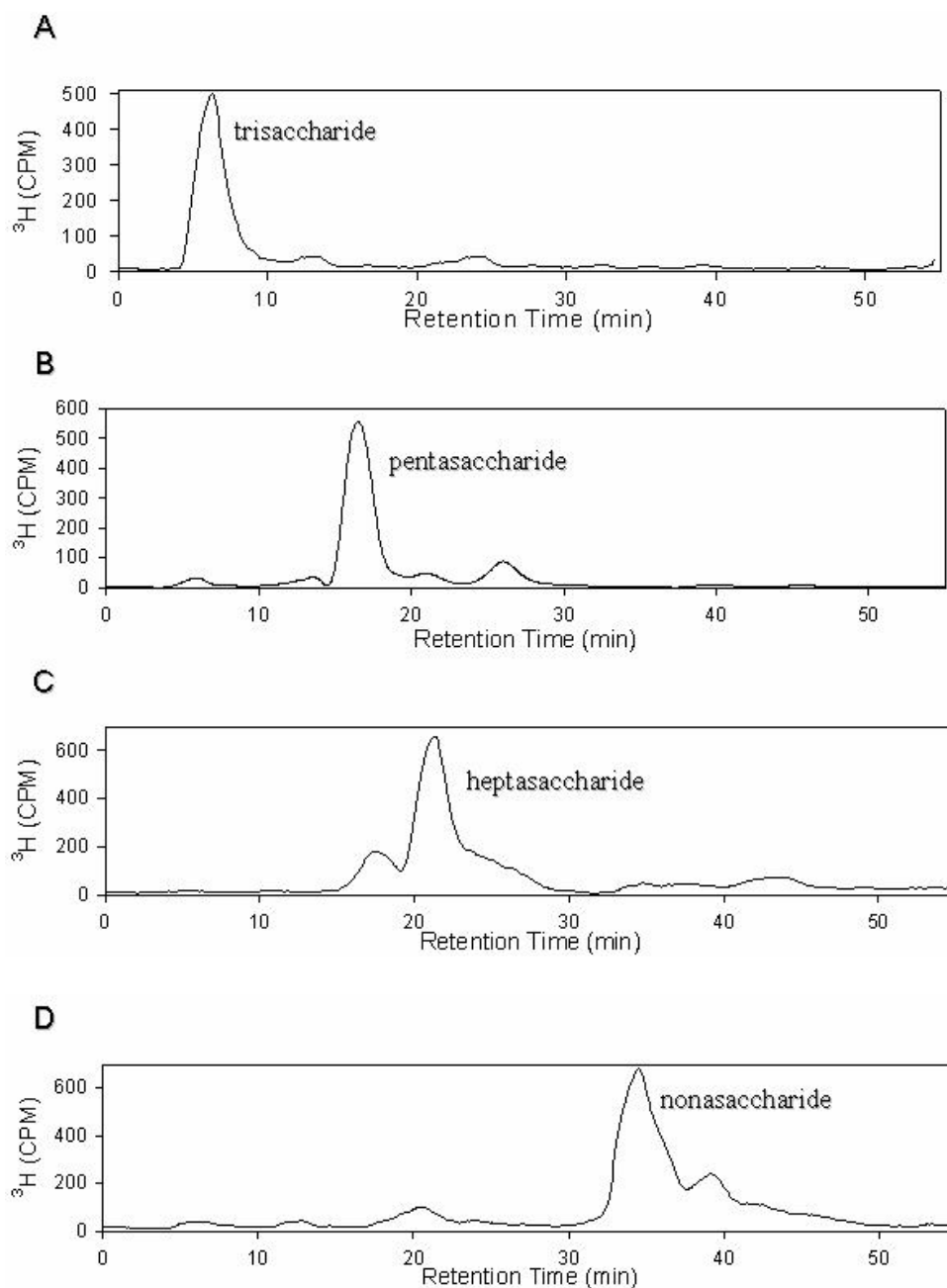
Figure 24. RPIP-HPLC chromograms of KfiA catalyzed *N*-acetyl glucosaminyl reactions with different sizes of the acceptor substrate.

Various oligosaccharide substrates were incubated with purified KfiA and UDP-GlcN[^{3}H]Ac. The reaction was terminated by heating at 100°C for 2 min.   The resultant [^{3}H]oligosaccharides were resolved by RPIP-HPLC eluted with different concentrations of acetonitrile as described in "Materials and Methods".   Panel A, B, C, and D shows the chromatogram of using disaccharide, tetrasaccharide, hexasaccharide and octasaccharide substrate, respectively.

To further quantify the susceptibility of the acceptors to KfiA modification, we determined K_m and K_{cat}/K_m of KfiA towards various acceptors. The amount of the ^{3}H-labeled oligosaccharide product was determined by integrating the corresponding [^{3}H]-peak from RPIP-HPLC. As shown in table 4, the binding affinities (as estimated by K_m) of various acceptors to KfiA are similar (between 256 and 674 µM), while the catalytic efficiency of KfiA (K_{cat}/K_m) moderately increases as the acceptor elongates (from 1.2 to 4.1). This suggested that elevated GlcNAc transfer efficiency could occur as the saccharide chain is elongated. We attempted to determine the K_m and K_{cat}/K_m values of KfiA for K5 polysaccharide, but failed to obtain a consistent value partly because the impurities in the polysaccharide substrate interfered with the enzymatic reaction. We also examined the potential effect of purified recombinant KfiC on the K_m and K_{cat}/K_m of KfiA towards a disaccharide substrate. We found that the K_m and K_{cat}/K_m for KfiA (in the presence of 1.5 µM of KfiC) is 315 µM and 1.4 ($M^{-1}S^{-1}$) respectively, while the K_m for K_{cat}/K_m for KfiA (in the absence of KfiC) is 256 µM and 1.2 ($M^{-1}S^{-1}$) respectively. This observation suggests that purified KfiC has no significant effect on the kinetic character of KfiA in this system.

Substrates	K_m (μM)	K_{cat}/K_m $(M^{-1}S^{-1})$	R^{2a}
UDP-GlcNAc[b]	94	3.7	0.9935
Disaccharide[c]	256	1.2	0.9968
Tetrasaccharide	324	1.2	0.9976
Hexasaccharide	674	2.7	0.9959
Octasaccharide	254	4.1	0.9931

Table 4. K_m and K_{cat}/K_m values of KfiA towards different acceptors

a. R^2 represents the coefficient determination for the data to fit Michaelis-Menten equation.
b. Disaccharide was used as an acceptor substrate as explained under "Materials and Methods".
c. The concentration of UDP-GlcN[^{3}H]Ac used in the assays was 1270 μM.

Section 7. Donor specificities of KfiA

It is believed that KfiA utilizes UDP-GlcNAc as a donor substrate *in vivo*. We further studied the donor specificity of KfiA. Isothermal Titration Calorimetry (ITC) was employed to measure the binding affinities between KfiA and different UDP-sugars (table 5). As the natural donor of KfiA, the dissociation constant (K_d) for UDP-GlcNAc was determined to be 62 μM, while we could not detect bindings between KfiA and UDP-GlcA, UDP-glucose or UDP-*N*-acetyl-galactosamine (GalNAc). This suggested that KfiA recognizes specific functional groups of GlcNAc residue, including the *N*-acetyl group (the difference of glucose and GlcNAc is at the *C*2 position) and the orientation of the hydroxyl group (the difference between galactosamine and glucosamine is at the *C*4 position). UDP demonstrated a comparable binding affinity (27 μM) to that of UDP-GlcNAc, which implies that UDP might exhibit product inhibition of KfiA.

Indeed, the IC_{50} of UDP on the *N*-acetyl D-glucosaminyl transferase activity of KfiA was measured to be 150 μM. The observation of the binding of UDP to KfiA, although somewhat surprising, is consistent with the previous studies on UDP-*N*-acetylhexosaminyltransferase, EXTL2 (Saobhany 2005). The K_m of KfiA towards UDP-GlcNAc was determined to be 94 μM, which is very similar to the dissociation constant (K_d) (62 μM).
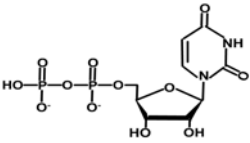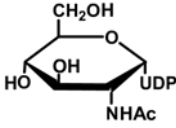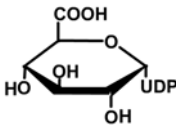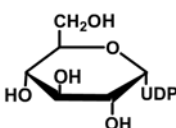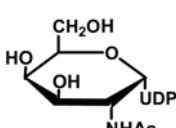
Donors	*Structures of donor substrates*	*K_d (μM)*	*N^a*
UDP		27	1.25
UDP-GlcNAc		62	1.04
UDP-GlcA		ND[b]	ND[b]
UDP-Glucose		ND[b]	ND[b]
UDP-GalNAc		ND[b]	ND[b]

Table 5. Binding affinities of various UDP-sugars

a, binding stoichiometry
b, not detected

Section 8. Conclusion and discussion

Previous reports demonstrated that at least four proteins are involved in the synthesis of K5 polysaccharide, including KfiA, KfiB, KfiC and KfiD (Griffiths 1998) (Hodson 2000). Although KfiA has been overexpressed in *E. coli* (Hodson 2000), the purified form has not been reported. Furthermore, the donor and acceptor substrate specificity of KfiA with purified protein has not been extensively studied. In this chapter, the successful expression and purification of KfiA in a multi-milligram scale was achieved with *E. coli* coexpressing chaperone proteins. The access of highly purified KfiA allowed us to conduct substrate specificities using established approaches.

The substrate specificities of KfiA are unique for its application in the synthesis of HS backbone. Incubation of KfiA with oligosaccharide acceptors of various sizes all yielded extended oligosaccharide products. Our results suggested that KfiA can be employed in the *de novo* synthesis of HS backbones starting from as short as disaccharide, provided that a disaccharide is readily available from complete depolymerization of K5 polysaccharide with nitrous acid. It is interesting to note that KfiA activity is inhibited by UDP. Whether UDP competes with UDP-GlcNAc in the same binding pocket on KfiA and regulate the extent of the polymerization of capsular polysaccharide of *in vivo* remains to be investigated. We do not rule out the possibility that KfiA transfers a GlcNAc residue towards other glycosaminoglycan substrates, including hyaluronic acid oligosaccharides, due to the unavailability of those substrates. Clearly, an efficient

expression of KfiA and detailed substrate specificity information will serve as an

important tool to conduct *in vitro* enzymatic synthesis of HS and to probe the biosynthetic

mechanism of HS.

Chapter IV. *De novo* synthesis of heparan sulfate oligosaccharide backbones

Section 1. Introduction

Structurally defined HS can be used to probe the functions of HS in a given biological system. The availability of structurally defined HS oligosaccharides is a major challenge to study the structure-function relationship of HS.   Although the chemical synthesis has proven to be the most efficient way to prepare the structurally defined HS oligosaccharides in sizes smaller than hexasaccharides (Avci 2003), it has become clear that chemical synthesis alone is currently incapable of generating majority of larger oligosaccharide structures.

Our laboratory and others began to explore the use of a collection of HS biosynthetic enzymes to synthesize HS with desired biological activities (Kuberan 2003; Kuberan 2003; Chen 2005; Lindahl 2005; Muñoz 2006).   However, it is unclear that the enzymatic approach offers the HS polysaccharide products with defined sequences in part due to the structural heterogeneity of starting backbone structures.   Furthermore, the modification of the backbone structure could determine the preference for the subsequent enzymatic modifications, and consequently regulate the synthesis of the polysaccharide with specific saccharide sequences (Esko 2002).   Availability of backbone structures

with defined size and *N*-sulfation serves as an essential step towards the *de novo* synthesis of HS and a critical reagent to delineate the biosynthetic mechanism of HS.

Section 2. Preparation of the glucuronyl transferases

The biosynthesis of HS backbone is a repeating process of the alternating addition of GlcNAc and GlcA monosaccharide units (figure 9) to the growing polysaccharide chain by two glycosyltransferases. With the glucosaminyl (GlcNAc) transferase KfiA available in large quantities, the next prerequisite for the *in vitro* synthesis of HS backbone is the glucuronyl (GlcA) transferase.

1. KfiA/KfiC complex

KfiC was previously identified as carrying the GlcA transferase activity in the elongation of *E. coli* K5 polysaccharide in a cell-based assay (Griffiths 1998). The GlcA transferase activity of the recombinant KfiC was tested by a procedure modified from the previous report (Griffiths 1998). Briefly, [^{3}H]-labeled heptasaccharide was first prepared by adding a GlcN[^{3}H]Ac residue to the nonreducing end of hexasaccharide through a KfiA reaction. The resultant [^{3}H]heptasaccharide was used as the acceptor substrate and the unlabeled UDP-GlcA was used as the donor substrate in the enzymatic assay described in figure 25A. Purified proteins were incubated with acceptors and donors under the standard assay conditions. The GlcA transferase activity was

measured by the resultant [3H]-labeled heptasaccharide on RPIP-HPLC (figure 25B).

To our surprise, the purified KfiC did not show any detectable GlcA transferase activity (dotted line, figure 25B).   We, therefore, coexpressed full-length KfiA and KfiC in *E. coli* BL21 star cells and found that glucuronyl transferase activity was readily detectable in this enzyme complex (dashed line, figure 25B).   It is noteworthy that mixing KfiA and KfiC *in vitro* did not restore this activity, suggesting that the active form of KfiA/KfiC complex is established immediately after their biosynthesis.   Another possibility is that KfiA helps KfiC correctly fold into its active forms and that KfiC is still the real GlcA transferase.   However, our method cannot distinguish which possible scenario is correct.
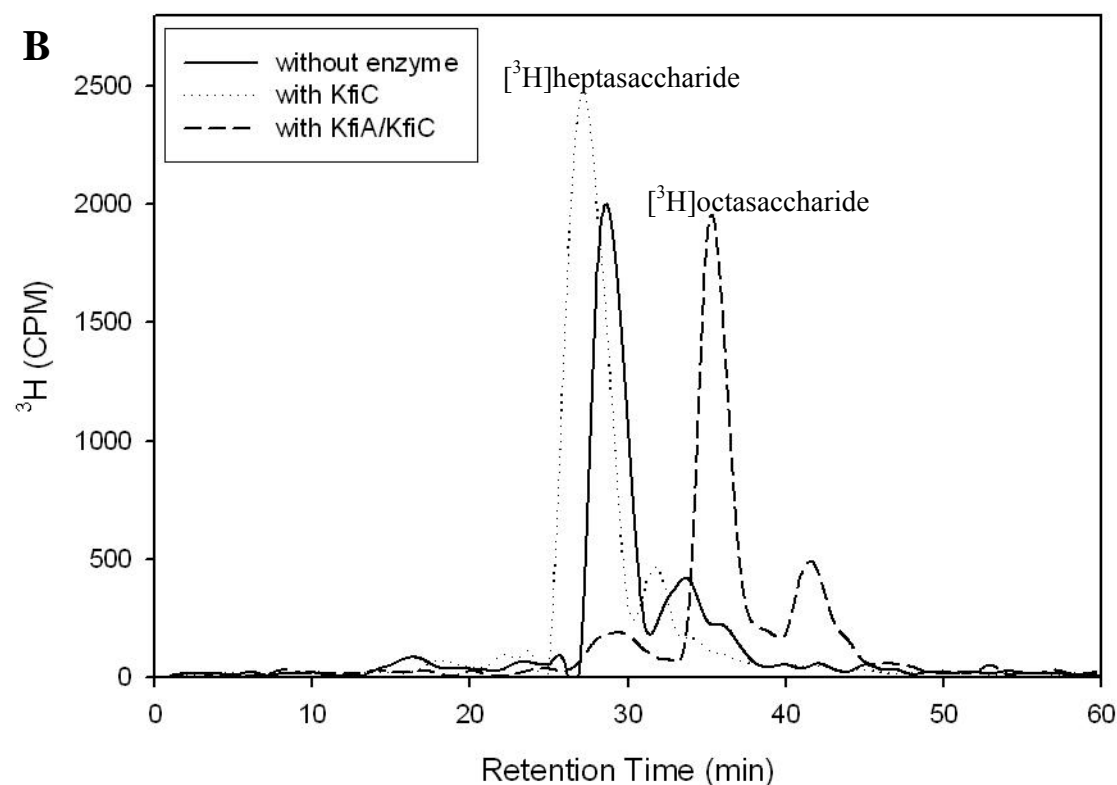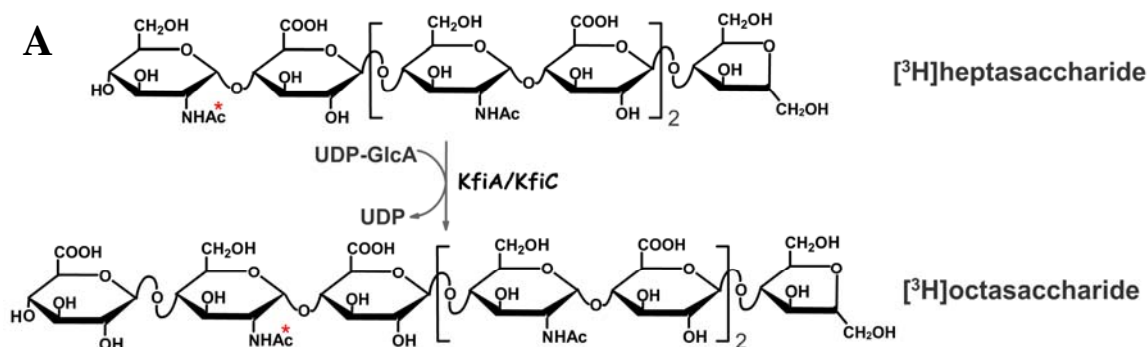
Figure 25. GlcA transferase assay on KfiC and KfiA/KfiC

A. The enzymatic assay used [³H]heptasaccharide as the acceptor and UDP-GlcA as the donor. The [³H]acetyl group is labeled with a red star.

B. The resultant [³H]octasaccharide was resolved on RPIP-HPLC.
Solid line: without any enzyme; dotted line: with KfiC; dashed line: with KfiA/KfiC complex.
The starting [³H]heptasaccharide was eluted around 28 min and the resultant [³H]octasaccharide was eluted around 36 min.

The reactivity of KfiA/KfiC complex towards acceptor size is also critical for the *de novo* synthesis of HS backbones. Although KfiA/KfiC is the only enzyme currently identified as carrying the GlcA transferase activity in *E. coli* K5, we failed in using this enzyme complex to work on trisaccharide with a GlcNAc residue at the non-reducing end. This suggests KfiA/KfiC may be only active on longer oligosaccharides or the growing K5 polysaccharide.   Other factors responsible for the GlcA activity in the initiation stage of the biosynthesis of K5 polysaccharide in *E. coli* cannot be ruled out.

2. PmHS2

PmHS1 and PmHS2 are the two bifunctional glycosyltransferases in *P. multocida* with both GlcNAc and GlcA transferase activities (DeAngelis 2002; DeAngelis 2004).   The full length PmHS2 gene was cloned from *P. multocida* type D genomic DNA and expressed in a format of 6×histidine tagged protein as described in "Materials and Methods".   Recombinant PmHS2 was purified by affinity chromatography using a nickel column.   The purified PmHS2 predominantly migrated as one major band at ~73 KDa on a 10% SDS-PAGE (figure 26), which is very close to the calculated molecular weight of PmHS2 (75 KDa).   We estimated that the purity of PmHS2 was greater than 90%.   Approximately 15 mg of purified PmHS2 was obtained from 1 liter of bacterial culture.

100 KDa

75 KDa — PmHS2, 75 KDa

50 KDa

37 KDa
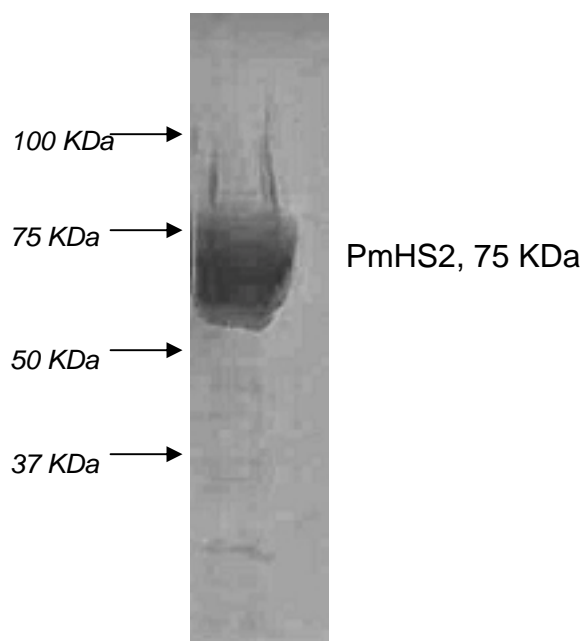
Figure 26. SDS-PAGE of the purified PmHS2

Affinity chromatography purified PmHS2 protein was resolved on precasted 10% SDS-PAGE. The gel was stained with coomassie blue. Migration positions of molecular markers (from BioRad) are indicated. The apparent mass was approximately 73 KDa. The predicted molecular mass of PmHS2 is 75 KDa.

The GlcA transferase activity of PmHS2 was tested in a reaction with trisaccharide (GlcNAc-GlcA-anMan$_R$) as acceptor substrates and UDP-[^{14}C]GlcA as donor substrates as described in "Materials and Methods". The reaction mixture was loaded onto a Bio-Gel P10 column and the resultant [^{14}C]-labeled tetrasaccharide was identified based on its molecular size and by comparing its elution position/time (P10 fraction number 66) with those of [^{3}H]-labeled trisaccharide and [^{3}H]-labeled pentasaccharide (figure 27). The elution positions of the [^{3}H]-labeled trisaccharide (P10 fraction number 68) and pentasaccharide (P10 fraction number 61) prepared from radioactive KfiA reactions had been previously recorded as the standards. A negative control reaction was also carried out unders standard conditions except for lacking the trisaccharide acceptor. The results support that PmHS2 is active when a trisaccharide is used as its substrate.

It is worthwhile noting that some of the UDP-[^{14}C]GlcA degraded into [^{14}C]GlcA even in the enzymatic assay with trisaccharide included (figure 27, fraction number 79) in the reaction catalyzed by PmHS2, raising the question of the catalytic efficiency of PmHS2 in the reaction of single sugar (GlcA) addition. With this concern in mind, we found that PmHS2 can use oligosaccharides of various sizes as its acceptor (as short as a trisaccharide which could not be recognized by the complex of KfiA/KfiC). This suggests that PmHS2 is indispensable in the *de novo* synthesis of HS backbones. My work, therefore, has focused on using PmHS2 to synthesize the backbone of HS.
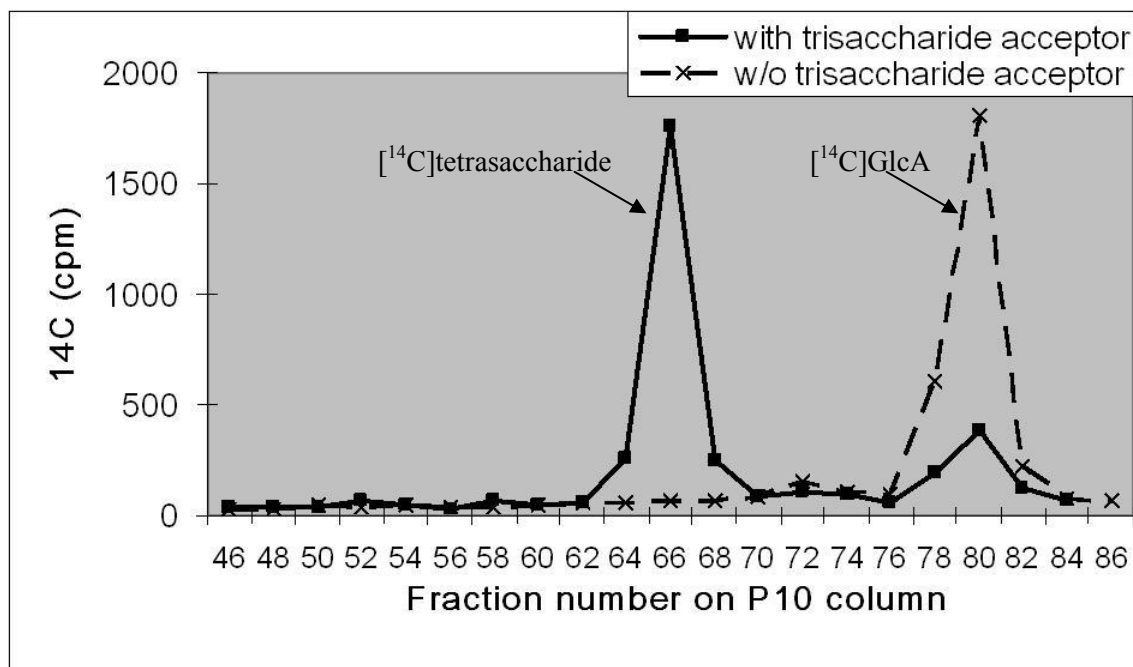
Figure 27. PmHS2 enzymatic assay on a Bio-Gel P10 column

PmHS2-synthesized [^{14}C]tetrasaccharide ([^{14}C]GlcA-GlcNAc-GlcA-anMan$_R$) is shown in solid line and was eluted around fraction number 66 which is between the elution positions of trisaccharide (fraction number 68) and pentasaccharide (fraction number 61).
The dotted line shows the product from the negative control reaction where no trisaccharide acceptor was included.

Section 3. Synthesis and test of UDP-*N*-trifluoroacetylglucosamine (UDP-GlcNCOCF$_3$)

The purpose of my research is to position the *N*-sulfated glucosamine (GlcNS) units in

the backbone of HS.   It appears that using UDP-GlcNS as a donor is a direct approach.

However, direct transfer of GlcNS is excluded at this point due to the unavailability of

UDP-GlcNS.   To this end, we prepared UDP-GlcNAc derivatives, including

UDP-glucosamine   (UDP-GlcN)   and   UDP-*N*-trifluoroacetylglucosamine

(UDP-GlcNCOCF$_3$) as potential alternative donor substrates for KfiA reaction.

Although the direct transfer of GlcN was not successful, we found that KfiA can

recognize UDP-GlcNCOCF$_3$ and transfer GlcNCOCF$_3$ to oligosaccharide acceptors.   In

this experiment, [^{3}H]-labeled hexasaccharide was incubated with KfiA in the presence of

UDP-GlcNCOCF$_3$.   The resultant GlcNCOCF$_3$-[^{3}H]hexasaccharide was resolved by

polyamine anion exchange-HPLC (PAMN-HPLC) (figure 28).   This finding is

consistent with a previous study where UDP-GlcNCOCF$_3$ was used as an unnatural donor

of the "core 2" GlcNAc transferase (Sala 1998).   In the same study, the *N*-trifluoroacetyl

groups were selectively removed from the sugar chain in a mild base condition without

affecting the *N*-acetyl groups, providing a potential strategy to control the GlcNS

positions on the backbone of HS.   Furthermore, compared with the *N*-acetyl groups, the

N-trifluoroacetyl groups were more susceptible to the deacetylation treatment.

Therefore, using UDP-GlcNCOCF$_3$ is advantageous when we plan to prepare the

completely *N*-unsubstituted HS backbones.
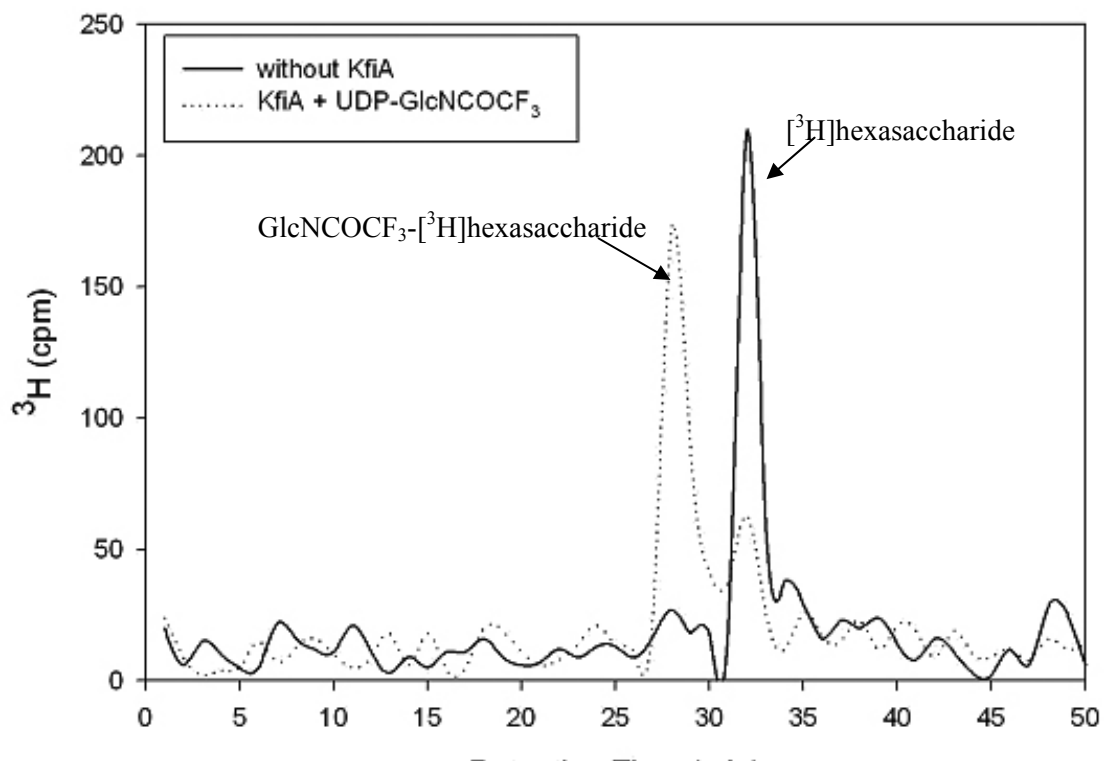
Figure 28. KfiA reaction using UDP-GlcNCOCF$_3$.

The resultant GlcNCOCF$_3$-[^{3}H]hexasaccharide (eluted around 28 min) was shifted left to the starting [^{3}H]hexasaccharide (eluted around 32 min) on PAMN-HPLC.

We further demonstrated that KfiA/KfiC complex indeed couples a GlcA to the GlcNCOCF$_3$ end of oligosaccharide with high efficiency. Here we prepared a heptasaccharide that contains a GlcNCOCF$_3$ at the nonreducing end by incubating a hexasaccharide with UDP-GlcNCOCF$_3$ in the presence of KfiA. Further extension of GlcNCOCF$_3$-hexasaccharide was achieved by incubating UDP-[^{14}C]GlcA and GlcNCOCF$_3$-hexasaccharide in the presence of KfiA/KfiC complex. PAMN-HPLC profiles showed a ~90% yield of [^{14}C]GlcA-GlcNCOCF$_3$-hexasaccharide (a [^{14}C]-labeled octasaccharide eluted around 33 min, figure 29). It should be noted that PmHS2 can also transfer GlcA to the oligosaccharide acceptors that contain the GlcNCOCF$_3$ residue at the nonreducing end as described below in Section 4.
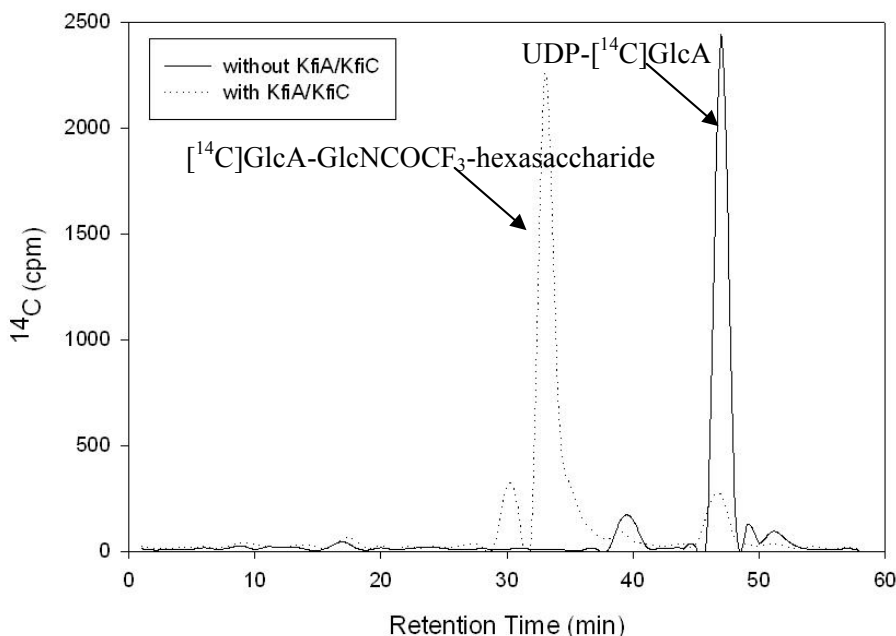
Figure 29. The complex of KfiA/KfiC transfers GlcA residue to GlcNCOCF$_3$.
Solid line: reaction without enzyme; Dotted line: reaction with KfiA/KfiC.
UDP-[^{14}C]GlcA was eluted around 47 min and the resultant [^{14}C]-labeled octasaccharide was eluted around 33 min.

As UDP-GlcNCOCF$_3$ is not commercially available, we developed a chemoenzymatic approach to synthesize this compound using *N*-acetylglucosamine-1-phosphate uridyltransferase (GlmU), in a similar way to that previously described in the synthesis of UDP-GlcN[^{3}H]Ac (Leiting 1998) (figure 30).

This two-step reaction uses GlcN-1-phosphate as the starting material. GlcN-1-phosphate first reacts with CF$_3$CO-SC$_2$H$_5$ and yields GlcNCOCF$_3$-1-phosphate, which is followed by the GlmU-catalyzed attachment of UDP to the *C*1 position on the GlcNCOCF$_3$ sugar. We obtained milligram-scale UDP-GlcNCOCF$_3$ with this approach as described in "Materials and Methods". In addition, the Linhardt group at Rensselaer Polytechnic Institute (RPI) used a chemical approach and synthesized 100 mg quantities of UDP-GlcNCOCF$_3$ for us as a large scale source of this UDP-sugar.
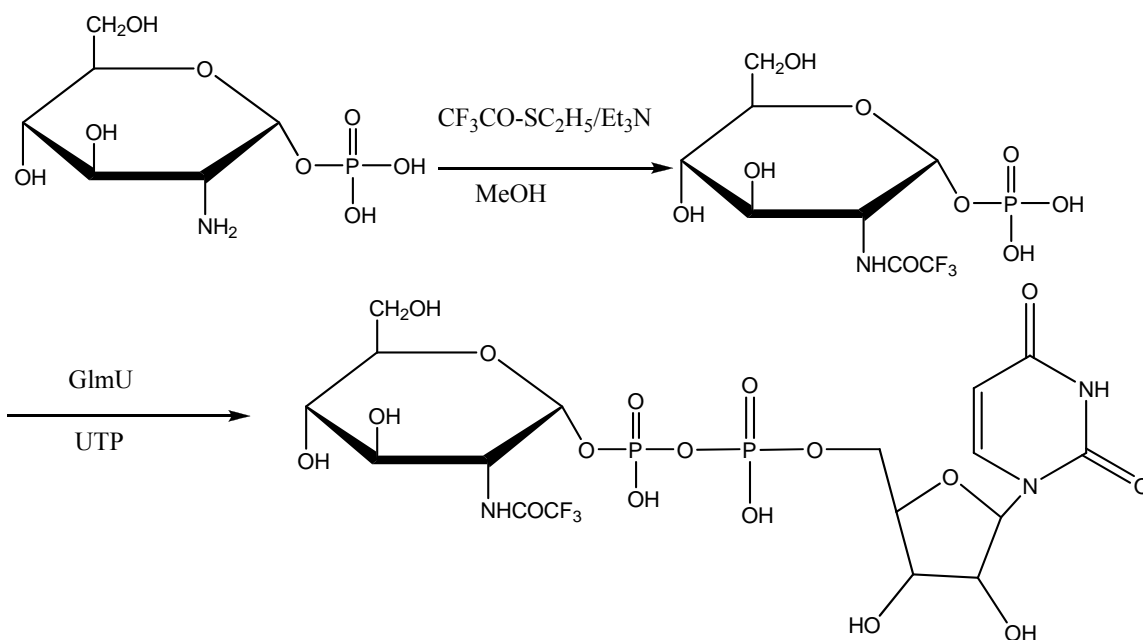
Figure 30. Chemoenzymatic synthesis of UDP-GlcNCOCF$_3$ (courtesy of Renpeng Liu)

Section 4. *De novo* synthesis and purification of heparosan

1. Preparation of the starting material-disaccharide

In order to control the *N*-sulfation and, therefore, the *O*-sulfation pattern of HS, it is necessary to synthesize the HS backbone from disaccharide. The starting disaccharide, GlcA-anMan$_R$, was prepared by the complete *N*-deacetylation and depolymerization of K5 polysaccharide (figure 31A). The [^{14}C]-labeled K5 polysaccharide was first subject to complete degradation and the resultant [^{14}C]disaccharide was loaded onto a Bio-Gel P10 column as a standard (figure 31B). The nonradioactive disaccharide was, thereafter, prepared in a large scale and purified with a P10 column by blindly collecting all the P10 fractions between 70 and 78 that was decided by the elution position of the [^{14}C]disaccharide standard. Approximately 100 mg of K5 polysaccharide was degraded to disaccharide using nitrous acid for my study.
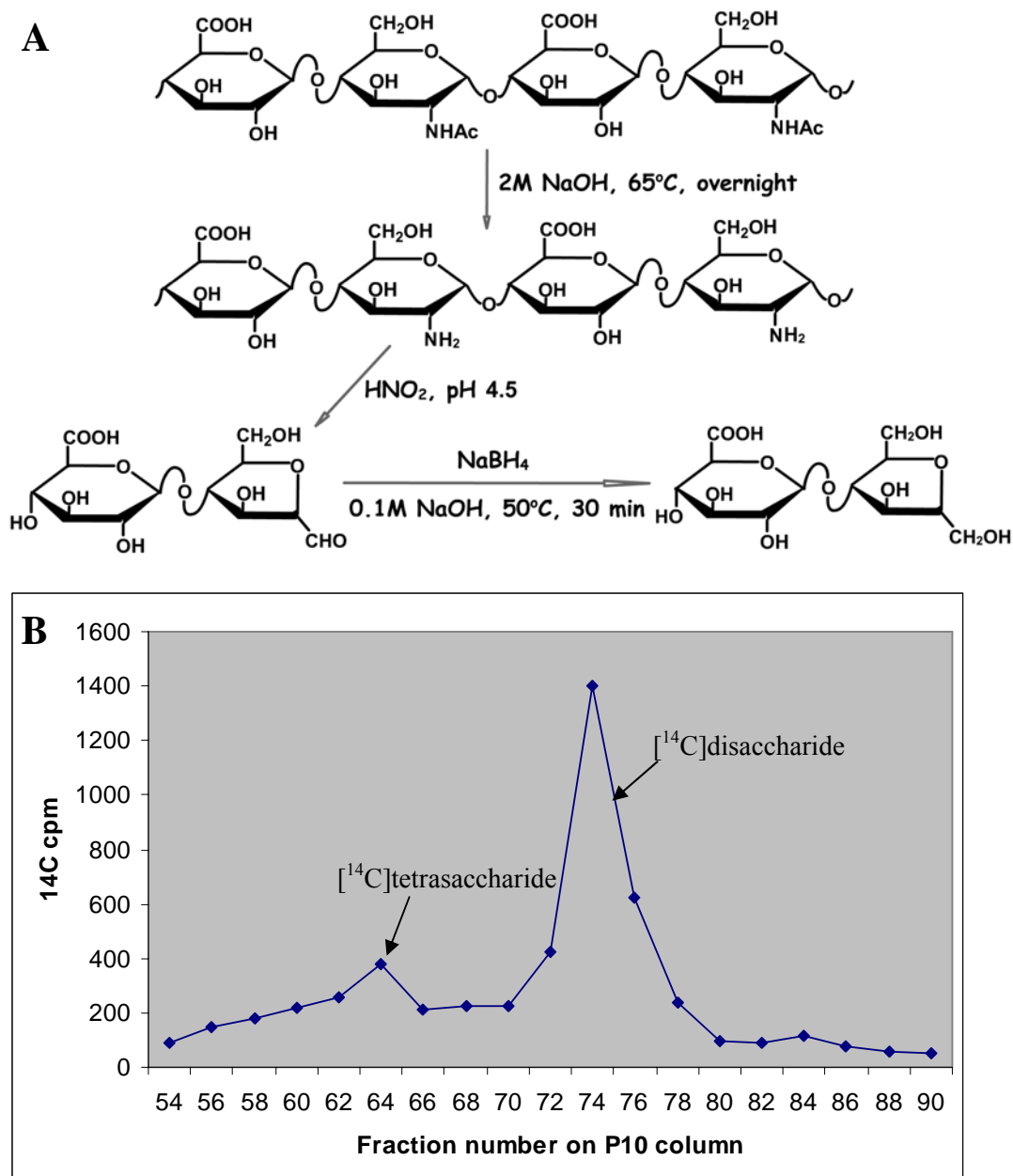
Figure 31. Preparation of disaccharide

A. Complete degradation of K5 polysaccharide into disaccharide

The incubation time of K5 polysaccharide with 2 M NaOH is critical for the extent of *N*-deacetylation and it should be more than 3 hours to reach completion.

B. The [^{14}C]disaccharide used as a standard on a Bio-Gel P10 column

Most of the [^{14}C]-labeled K5 polysaccharide was completely depolymerized into [^{14}C]disaccharide (GlcA-anMan$_R$) which was eluted around fraction number 74 (ranging from number 70 to number 78) on the P10 column (15 drops/fraction). There was a minor peak at fraction number 64 believed to be [^{14}C]tetrasaccharide (GlcA-GlcNAc-GlcA-anMan$_R$) coming from the partial degradation of K5 polysaccharide.

2. Synthesis of trisaccharide

Trisaccharide, GlcNAc-GlcA-anMan$_R$, was synthesized with disaccharide (GlcA-anMan$_R$) as acceptor and UDP-GlcNAc as donor in a KfiA reaction (figure 32A). Similar to the strategy we used to purify disaccharide, the [^{3}H]trisaccharide was prepared from a radioactive KfiA reaction using unlabeled disaccharide and UDP-GlcN[^{3}H]Ac. The product was loaded onto a Bio-Gel P10 column to locate the elution position of a trisaccharide (figure 32B). A nonradioactive trisaccharide was then prepared by KfiA in the presence of unlabeled UDP-GlcNAc.

Complete removal of UDP-GlcNAc from the reaction product is essential for the subsequent PmHS2-catalyzed synthesis of tetrasaccharide. If UDP-GlcNAc is present with UDP-GlcA, PmHS2, a bifunctional enzyme with both GlcNAc and GlcA transferase activities, is able to extend the trisaccharide to polysaccharide, which will reduce the yield of tetrasaccharide significantly. However, trisaccharide and UDP-GlcNAc cannot be resolved very well on P10 column. To overcome this problem, a limited amount of UDP-GlcNAc was incubated with excess disaccharide acceptor to exhaust UDP-GlcNAc under the standard conditions. The resultant trisaccharide was purified with P10 column. Whether the trisaccharide was produced was determined by the level of the transfer of [^{14}C]GlcA using PmHS2.
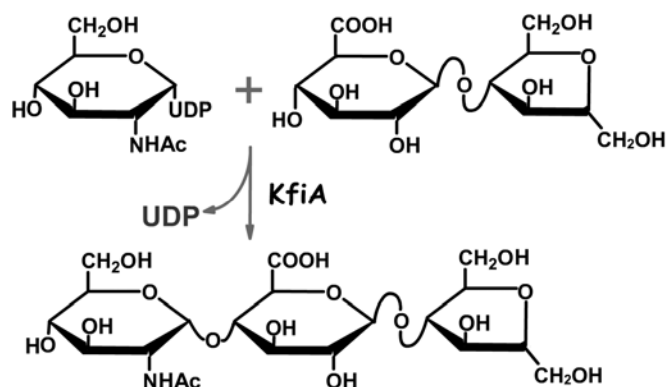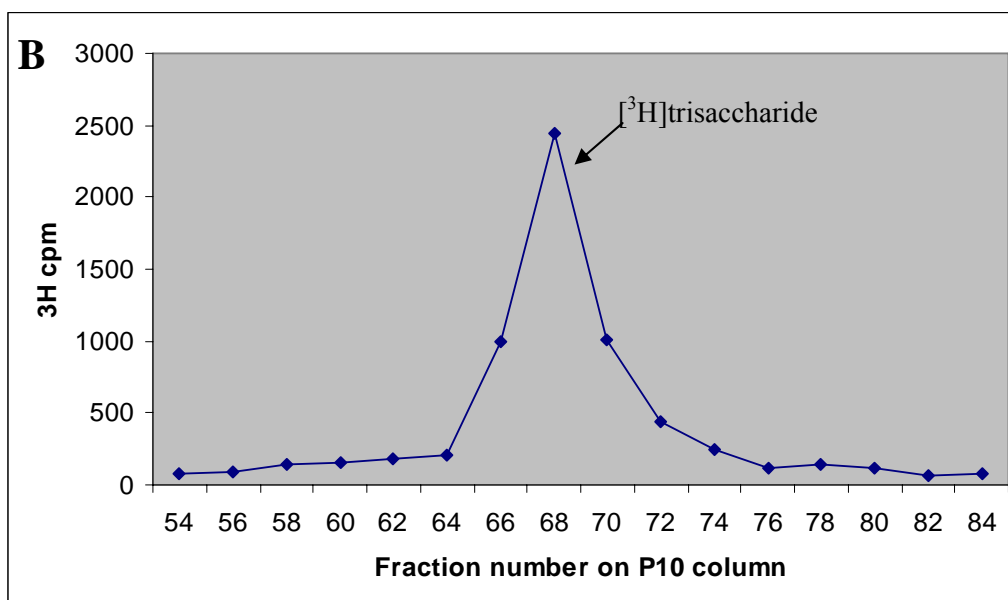
Figure 32. Synthesis of trisaccharide

A. KfiA transfers the GlcNAc residue to disaccharide

B. The elution profile of the [^{3}H]trisaccharide standard on a Bio-Gel P10 column
[^{3}H]trisaccharide standard was prepared from a KfiA reaction containing UDP-GlcN[^{3}H]Ac and it was eluted from fraction number 65 to 71 (peaked at number 68) on the P10 column (15 drops/fraction).

3. Synthesis of unnatural oligosaccharide backbones

As mentioned previously, UDP-GlcNCOCF$_3$ was used by KfiA as well as its natural donor substrate UDP-GlcNAc.   Therefore, the unnatural HS backbones with repeating -GlcA-GlcNCOCF$_3$- disaccharide unit can also be synthesized with KfiA and PmHS2 as illustrated in figure 33.   Then we focused our attention on the synthesis of the unnatural backbone with GlcNCOCF$_3$ other than the natural one containing GlcNAc, in order to control the *N*-sulfation of HS oligosaccharides.
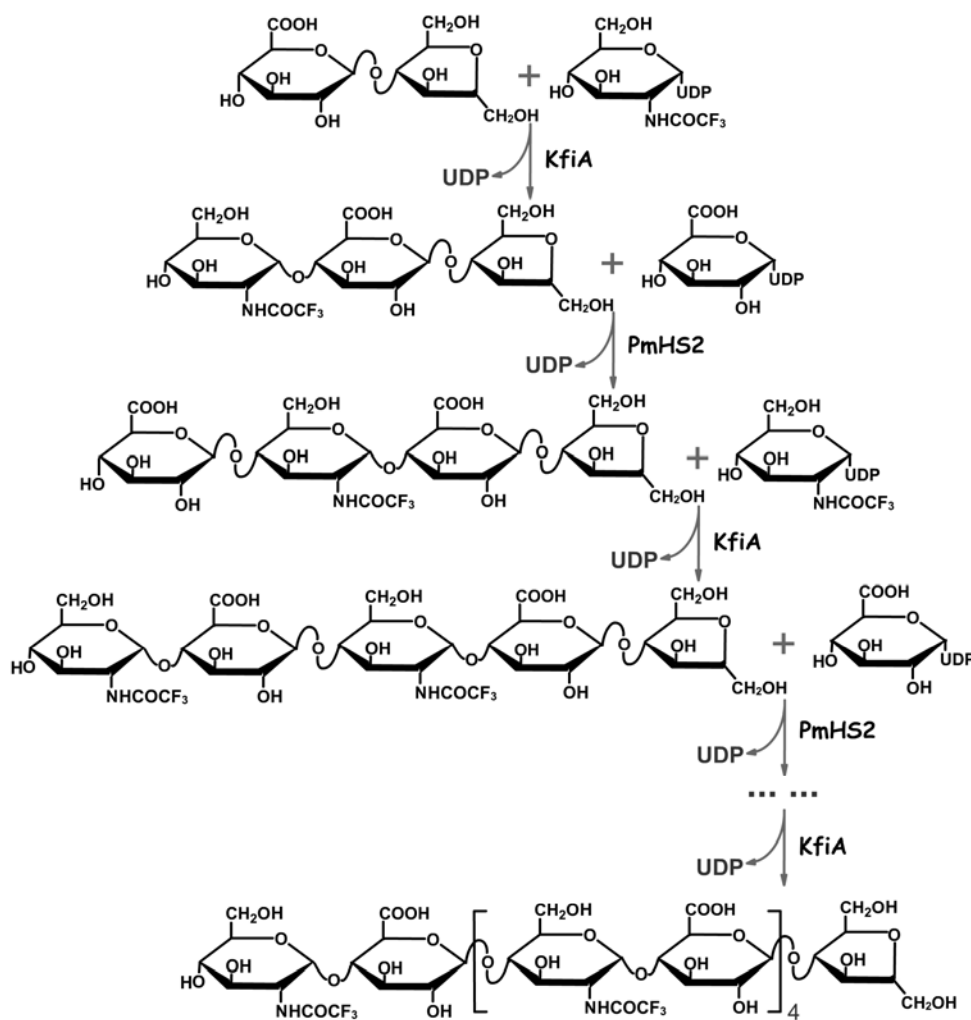
Figure 33. Synthesis of HS unnatural backbones with various size

Synthesis of unnatural trisaccharide (GlcNCOCF$_3$-GlcA-anMan$_R$)

In the first step, we started with the disaccharide (GlcA-anMan$_R$) (4.5 μmole, 1.52 mg) as the acceptor and UDP-GlcNCOCF$_3$ (3.9 μmole, 2.58 mg) as the unnatural donor substrate in the presence of KfiA as described in "Materials and Methods" (figure 34A). The resultant unnatural trisaccharide (GlcNCOCF$_3$-GlcA-anMan$_R$) was purified with a Bio-Gel P10 column in a similar fashion as we prepared the natural trisaccharide (GlcNAc-GlcA-anMan$_R$). The fractions from P10 column corresponding to the synthesized unnatural trisaccharide were blindly recovered according to the elution position of the [^{3}H]trisaccharide standard (GlcN[^{3}H]Ac-GlcA-anMan$_R$). The purified unnatural trisaccharide (GlcNCOCF$_3$-GlcA-anMan$_R$) was dissolved in ddH$_2$O and subjected to a test to check its quantity and purity. This test used PmHS2 and excess UDP-[^{14}C]GlcA with known concentration (100 μM) and specific activity (200 mCi/mMole). The resultant tetrasaccharide ([^{14}C]GlcA-GlcNCOCF$_3$-GlcA-anMan$_R$) was analyzed on the P10 column (figure 34B) and its quantity was calculated based on the specific activity of UDP-[^{14}C]GlcA. There was no other oligosaccharide peak besides the expected [^{14}C]tetrasaccharide, indicating that the synthesized trisaccharide was pure. There was a notable [^{14}C]-labeled unknown species that eluted very late (around fraction number 77) on the P10 column. Its molecular weight was close to monosaccharide ([^{14}C]GlcA, figure 34B), which is consistent with our speculation on the degradation of UDP-[^{14}C]GlcA into [^{14}C]GlcA by PmHS2 in the reaction. The amount of the synthesized trisaccharide was 2.87 μmole (1.71 mg). The yield of the synthesis of

trisaccharide (GlcNCOCF$_3$-GlcA-anMan$_R$) was about 73 % based on the amount of

UDP-GlcNCOCF$_3$ (3.9 μmole, 2.58 mg) used in the reaction.
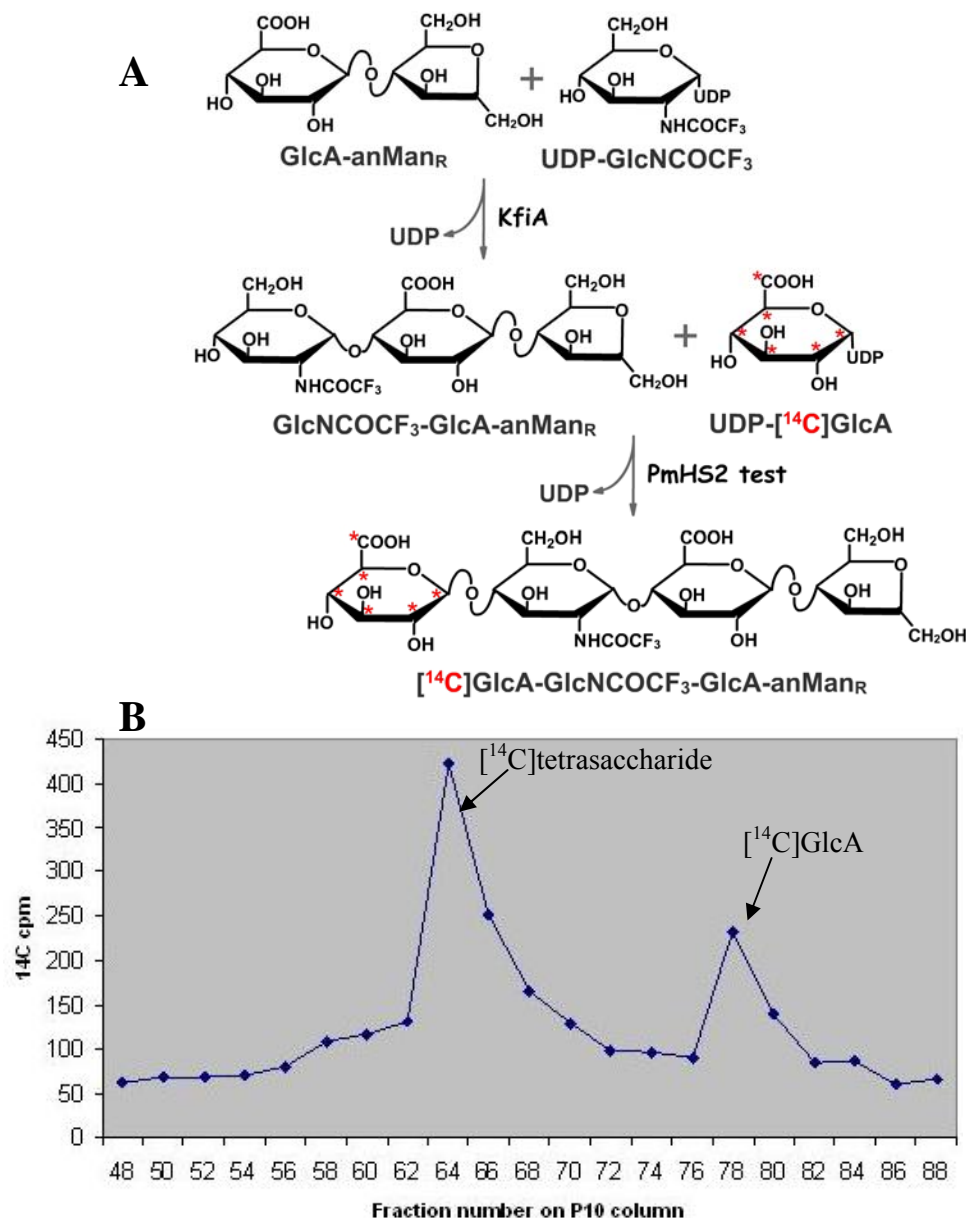
Figure 34. Synthesis of unnatural trisaccharide
A. KfiA transfers the GlcNCOCF$_3$ residue to disaccharide.
The resultant trisaccharide was tested in a subsequent radioactive reaction catalyzed by PmHS2.
B. Test on the quality and quantity the synthesized unnatural trisaccharide
The tested trisaccharide and UDP-[^{14}C]GlcA was incubated with PmHS2.   The resultant [^{14}C]tetrasaccharide was eluted around fraction number 65 on the P10 column.   The second peak (fraction number 79) is believed to be [^{14}C]GlcA degraded from UDP-[^{14}C]GlcA.

Synthesis of unnatural tetrasaccharide (GlcA-GlcNCOCF$_3$-GlcA-anMan$_R$)

The synthesized trisaccharide was further elongated to an unnatural tetrasaccharide (GlcA-GlcNCOCF$_3$-GlcA-anMan$_R$) in a subsequent reaction catalyzed by PmHS2 as described in "Materials and Methods" (figure 35A). Because the elution positions of trisaccharide and the resultant tetrasaccharide were too close to each other, it was difficult to separate tetrasaccharide from the unreacted trisaccharide. To overcome this problem, a limited amount of trisaccharide was incubated with excess UDP-GlcNCOCF$_3$ under standard conditions. Trisaccharide was exhausted and the unreacted UDP-GlcNCOCF$_3$ was completely removed from the resultant tetrasaccharide with a P10 column.

In details, as shown in figure 34, we knew where the radio-labeled tetrasaccharide ([^{14}C]GlcA-GlcNCOCF$_3$-GlcA-anMan$_R$) eluted and use its elution position (fraction number 65 on P10 column) as the standard to blindly recover the purified unnatural tetrasaccharide (GlcA-GlcNCOCF$_3$-GlcA-anMan$_R$). Similarly, the purity and quantity of the purified GlcA-GlcNCOCF$_3$-GlcA-anMan$_R$ was subject to a KfiA-catalyzed test with the donor substrate UDP-GlcN[^{3}H]Ac (407 μM, 172 mCi/mMole) in excess. The analysis with P10 column showed that the reaction product contained only [^{3}H]pentasaccharide (GlcN[^{3}H]Ac-GlcA-GlcNCOCF$_3$-GlcA-anMan$_R$) and the unreacted UDP-GlcN[^{3}H]Ac without any other sugar of different size. This suggested that the tested tetrasaccharide was pure (figure 35B). The quantity of the synthesized tetrasaccharide was estimated to be 1.79 μmole (1.38 mg) based on the specific activity of UDP-GlcN[^{3}H]Ac. The yield of the KfiA-catalyzed elongation reaction (from

trisaccharide to tetrasaccharide) was around 62 %.
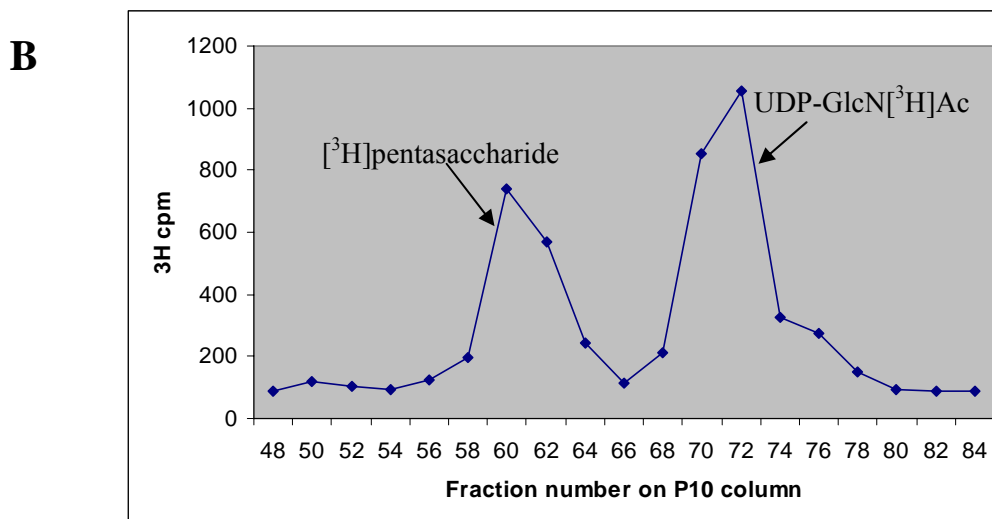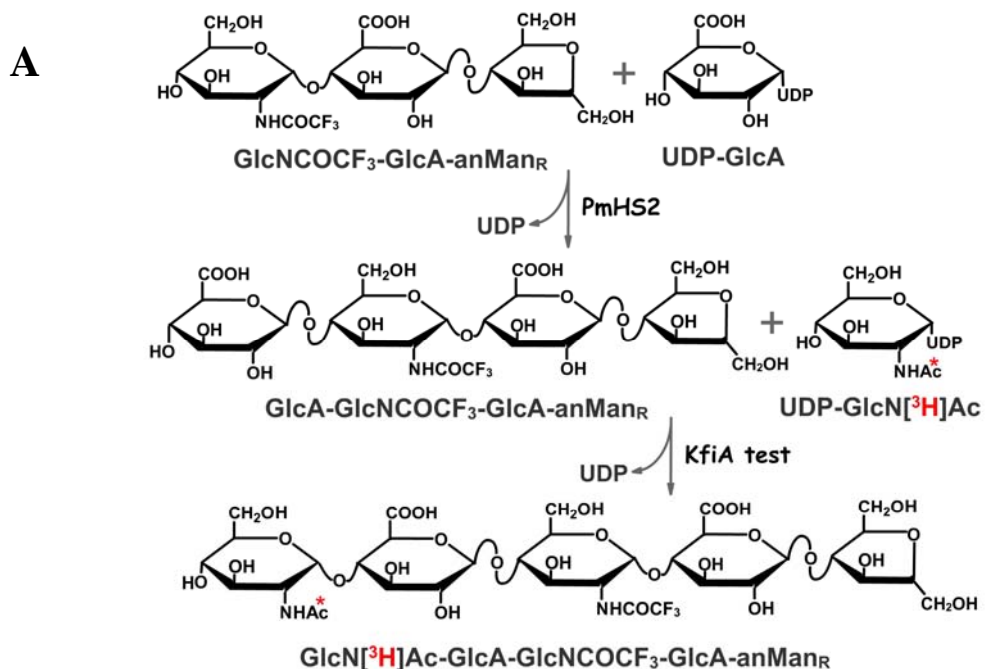
Figure 35. Synthesis of unnatural tetrasaccharide

A. PmHS2 transfers the GlcA residue to trisaccharide
The resultant tetrasaccharide was tested in a subsequent radioactive reaction catalyzed by KfiA.
B. Test on the quality and quantity the synthesized unnatural tetrasaccharide
[³H]pentasaccharide was produced from a KfiA-catalyzed test reaction containing UDP-GlcN[³H]Ac and was eluted around fraction number 58 to 64 (peaked at number 61) on the P10 column (15 drops/fraction). The second peak (around number 72) is believed to be unreacted UDP-GlcN[³H]Ac.

Section 5. Synthesis of longer unnatural oligosaccharides

The alternate KfiA and PmHS2 elongation reactions were repeated stepwise.   In each step, the quality and quantity of the resultant products were closely controlled in an approach described above, until the sugar chain reached the desired size.   In summary, we prepared the unnatural HS backbones at 100 μg to milligram scales (table 6).   All of these oligosaccharides contain $GlcNCOCF_3$ only without GlcNAc residues.   The sequence of the synthesized tetrasaccharide, hexasaccharide and octasaccharide were confirmed by mass spectrometry by Linhardt's group at the Rensselaer Polytechnic Institute (RPI).   As an example, 1 μg of the unnatural octasaccharide, $(GlcA\text{-}GlcNCOCF_3)_3$-$GlcA$-$anMan_R$, was examined by LC-MS and MS/MS as described in "Materials and Methods".   The formula weight of the octasaccharide was 1640.1 $(C_{54}H_{74}O_{44}N_3F_9)$ which is very close to the molecular weight (1640.0) calculated based on its MS spectrum $([M\text{-}2H]^{2-}$ and $[M+Na\text{-}2H]^{2-}$, figure 36A).   The subsequent MS/MS spectrum further confirmed its structure by identifying various fragmentations (figure 36B).

    In summary, our results demonstrated that it is feasible to synthesize HS backbones with the two bacterial glycosyltransferases (KfiA and PmHS2).   These results also support our hypotheses that $UDP\text{-}GlcCOCF_3$ can be employed to build unnatural HS backbones and that the size of HS oligosaccharides can be precisely controlled by regulating the access of KfiA, PmHS2 and the two UDP-sugar donor substrates.

Structure of oligosaccharides	Number of sugar residues	Amount (µg)	Molecular weight (Da)	
			Calculated	Determined[a]
GlcNCOCF$_3$-GlcA-anMan$_R$	3	1710	--	Not determined
GlcA-GlcNCOCF$_3$-GlcA-anMan$_R$	4	1380	--	Not determined
GlcA-GlcNCOCF$_3$-GlcA-anMannose[b]	4	160	771	771
GlcA-GlcN-GlcA-GlcN-GlcA-anMan$_R$[c]	6	150	1013	1014
(GlcA-GlcNCOCF$_3$)$_3$-GlcA-anMan$_R$	8	280	1639	1640
(GlcA-GlcNCOCF$_3$)$_4$-GlcA-anMan$_R$	10	250	--	Not determined
(GlcNCOCF$_3$-GlcA)$_5$-anMan$_R$	11	400	--	Not determined

Table 6. List of the synthesized oligosaccharides

a. Molecular weight was determined by LC-MS and MS/MS in Linhardt's group at RPI.

b. The reducing end of this tetrasaccharide contains anhydromannose, instead of anhydromannitol.

c. The hexasaccharide was completely *N*-deacetylated (*N*-unsubstituted).
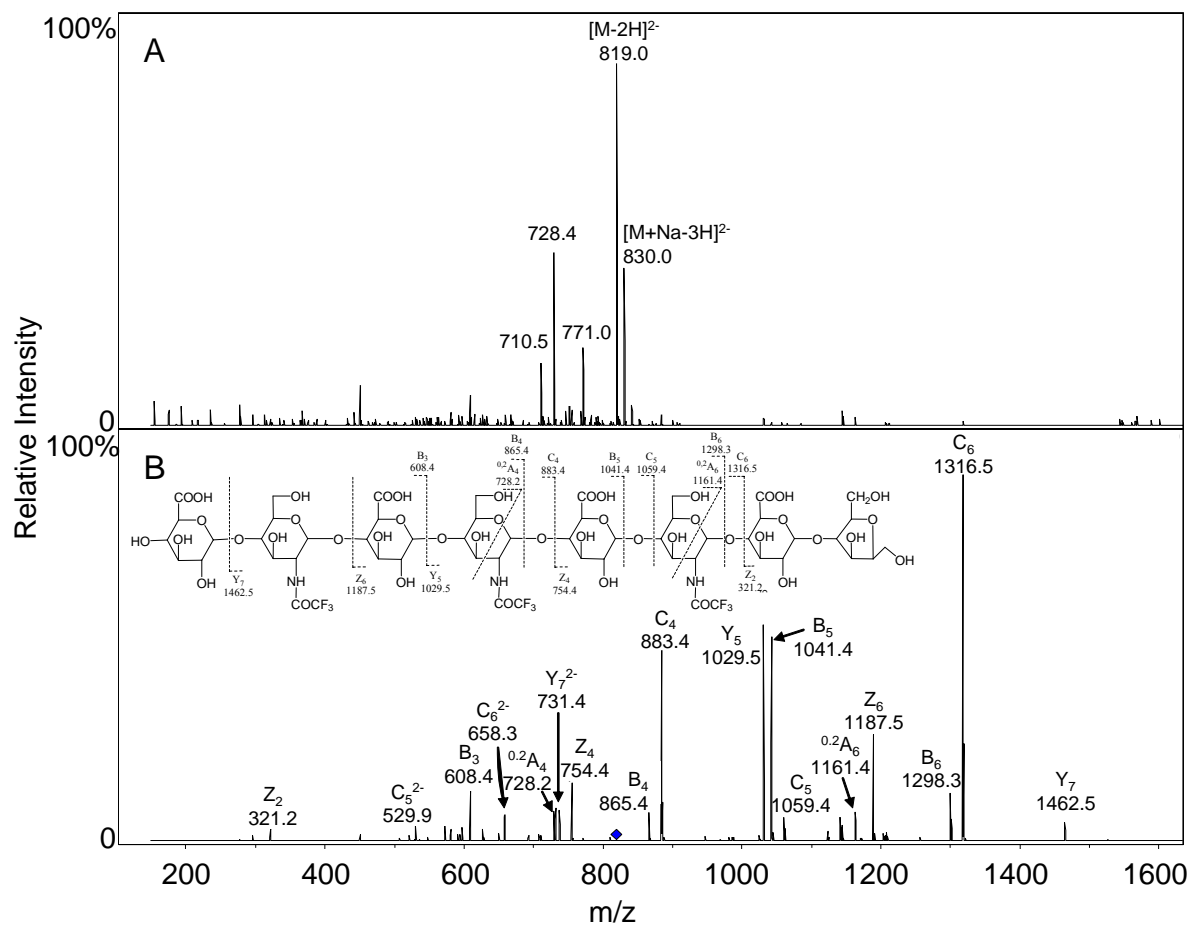
Figure 36. MS and MS/MS performed on 1 μg of octasaccharide

A. MS spectrum of octasaccharide
B. MS/MS spectrum and scheme of fragmentation of octasaccharide

Chapter V. Modifications of heparan sulfate backbones

Section 1. Introduction

A set of HS modification enzymes, including *N*-sulfotransferase (NST), *C*5-epimerase (C5-epi), 2-*O*-sulfotransferase (2OST), 6-*O*-sulfotransferase isoforms 1 and 3 (6OST1 and 6OST3), and 3-O-sulfotransferase isoform 1 (3OST1) have been cloned and expressed in large quantities in our laboratory.   The availabilities of these biosynthetic enzymes and the HS backbones allow us to prepare the HS oligosaccharides with desired sizes and sulfation patterns.   Directed by the knowledge of the biosynthesis of HS and the structures of several known anticoagulant HS oligosaccharides (Vuillemenot 1999; Chen 2007), we synthesized HS oligosaccharides with more defined structures and with antithrombin (AT)-binding activity enzymatically.

Section 2. *N*-deacetylation and *N*-sulfation of heparan sulfate backbones

In the biosynthesis of HS, the downstream epimerization and various *O*-sulfations preferentially occur at the disaccharide units having *N*-sulfations (Jacobsson 1984). As NST only reacts at the glucosamine residue with a free amine group at the *N* position, *N*-deacetylation is, thus, a decisive step in the modifications of HS backbones. Compared with the natural HS backbones with GlcNAc residues, the unnatural

backbones we prepared contain GlcNCOCF$_3$ residues.   The incorporated GlcNCOCF$_3$

residues can be deemed as the protected forms of GlcN, because the selective removal of

the *N*-trifluoroacetyl group can be achieved under mild conditions while the *N*-acetyl

group will be retained.   The selective *N*-deacetylation is, therefore, feasible by placing

GlcNCOCF$_3$ and GlcNAc at desired places on the sugar chain.   The complete

N-deacetylation of HS backbones containing GlcNCOCF$_3$ is also less demanding than

those with natural GlcNAc residues.

1. Complete *N*-deacetylation/*N*-sulfation of [^{3}H]-labeled heptasaccharide

A small amount of [^{3}H]-heptasaccharide (figure 37D, compound a) was first used to test

the efficiency of our approach for complete *N*-deacetylation/*N*-sulfation as described in

figure 37D.   The *N*-unsubstituted [^{3}H]-heptasaccharide (compound b), and the 2-hour

and overnight NST reactions were analyzed by RPIP-HPLC (figure 37, panel A, B, C).

Compared with compound b, compounds c and d were eluted later on RPIP-HPLC.

Since there are only two susceptible *N*-sulfation sites for compound b, we believe that

they were partially and completely *N*-sulfated [^{3}H]-heptasaccharides, respectively.

Compound c (figure 37B, 27 min) represents a monosulfated heptasaccharide and

compound d (figure 37B and C, 35min) represents a disulfated heptasaccharide.   Based

on this analysis, we estimated that the purity of the fully *N*-sulfated [^{3}H]-heptasaccharide

(compound d) in the overnight NST reaction was greater than 90%.   This experiment

also demonstrated that when the heptasaccharide was treated with triethylamine for

N-deacetylation, only its two *N*-trifluoroacetyl groups were susceptible and the

[³H]-labeled *N*-acetyl group was intact on the backbone. The result proves that the

selective *N*-deacetylation can be achieved with triethylamine.
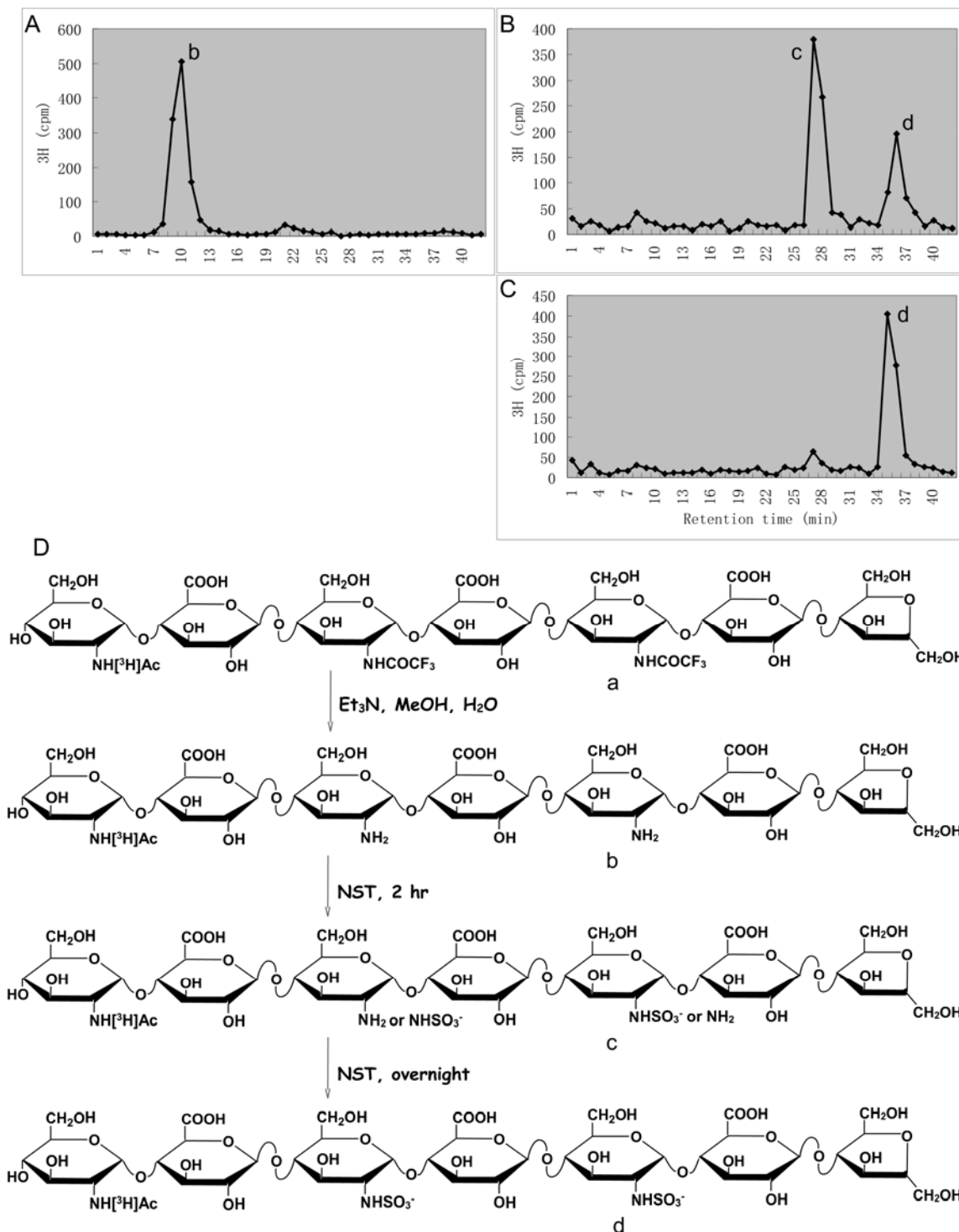
Figure 37. Complete *N*-deacetylation/*N*-sulfation of [³H]-labeled heptasaccharide

A. The *N*-unsubstituted [³H]-heptasaccharide analyzed on RPIP-HPLC

B. The RPIP-HPLC profile of the products from 2-hour NST reaction

C. The RPIP-HPLC profile of the product from overnight NST reaction

D. Synthetic scheme of *N*-deacetylation/*N*-sulfation

To prepare NS-[^{3}H]-heptasaccharide in a larger scale for further structural analysis, the hexasaccharide backbone (figure 38A) was first prepared with alternate KfiA and PmHS2 reactions and purified with a Bio-Gel P10 column as described in Chapter V. The complete *N*-deacetylation of its GlcNCOCF$_3$ residues with triethylamine in aqueous methanol yielded an *N*-unsubstituted hexasaccharide which was confirmed by mass spectrometry (table 6).   An additional GlcN[^{3}H]Ac monosaccharide was then added to the non-reducing end of the hexasaccharide as a radioactive label in a KfiA reaction using UDP-GlcN[^{3}H]Ac.   The product was purified by a P10 column.   This was followed by an NST reaction coupled with PAPS regeneration system leading to the *N*-sulfation of [^{3}H]-labeled heptasaccharide, which was purified by RPIP-HPLC (figure 38B).   The majority of NS-[^{3}H]heptasaccharide bound with the RPIP column tightly by forming a hydrophobic ion paired complex with tetrabutylammonium (TBA) and was eluted at around 31 min.   The earlier and smaller peak (eluted at around 7 min) represented the unbound product since when the load of NS-[^{3}H]heptasaccharide decreased, this peak disappeared.

Figure 38. Preparation of *N*-deacetylated/*N*-sulfated [^{3}H]-heptasaccharide

A. synthetic route

B. purification of *N*-sulfated [^{3}H]heptasaccharide

To examine the extent of *N*-deacetylation and *N*-sulfation, both high pH and low pH nitrous acid degradation tests were performed on [³H]-labeled heptasaccharide before and after the NST reaction (figure 39C).   At high pH (pH 4.5-5.5), nitrous acid only cleaves the oligosaccharide at the *N*-deacetylated glucosamine residues with free amine groups. In contrast, at low pH (pH 1.5), nitrous acid predominantly reacts at the *N*-sulfated glucosamine residues.   The products from the four tests were resolved with a Bio-Gel P10 column based on their molecular size (figure 39A, 39B).   Our results demonstrated that before the NST reaction, the [³H]-labeled heptasaccharide was resistant to low pH degradation and remained intact as a heptasaccharide, while it was sensitive to high pH degradation and was completely degraded to a [³H]-labeled trisaccharide without any partial degradation (figure 39A).   This suggests that after *N*-deacetylation, the *N*-trifluoroacetyl groups were completely removed from the HS backbone.   In contrast, after the NST reaction, the [³H]-labeled heptasaccharide became resistant to high pH degradation and could be completely degraded into [³H]-labeled trisaccharide by low pH degradation (figure 39B), indicating that *N*-sulfation was complete and there was no free amine groups left on the glucosamine residues of the HS backbone.

Figure 39. Nitrous acid degradation of [3H]heptasaccharide

A. low pH (dashed line) and high pH (solid line) nitrous acid degradation of *N*-deacetylated [3H]heptasaccharide

B. low pH (dashed line) and high pH (solid line) nitrous acid degradation of *N*-sulfated [3H]heptasaccharide

C. *N*-deacetylated [3H]heptasaccharide is susceptible to high pH degradation and *N*-sulfated [3H]heptasaccharide is susceptible to low pH degradation.   The resultant [3H]trisaccharides were detected as shown in panel A and B.   The low pH treatment of *N*-deacetylated [3H]heptasaccharide and high pH treatment of *N*-sulfated [3H]heptasaccharide did not cause any degradation and thus are not drawn.

119

2. Complete *N*-deacetylation/*N*-sulfation of [³H]-labeled undecasaccharide

As there were only two GlcNCOCF₃ residues in the [³H]-labeled heptasaccharide to be *N*-deacetylated and sulfated, we further used [³H]-labeled undecasaccharide, to test if our approach was still effective for longer HS backbones.

The decasaccharide backbone (figure 40A) was *N*-deacetylated by triethylamine. GlcN[³H]Ac was again attached to the non-reducing end of the decasaccharide as a radioactive label, followed by purification with a Bio-Gel P10 column and a subsequent NST reaction. The resultant *N*-sulfated and [³H]-labeled undecasaccharide was purified by PAMN-HPLC (figure 40B). The major radioactive population corresponding to the elution time 51-54 min on PAMN-HPLC was pooled and saved for further modification.
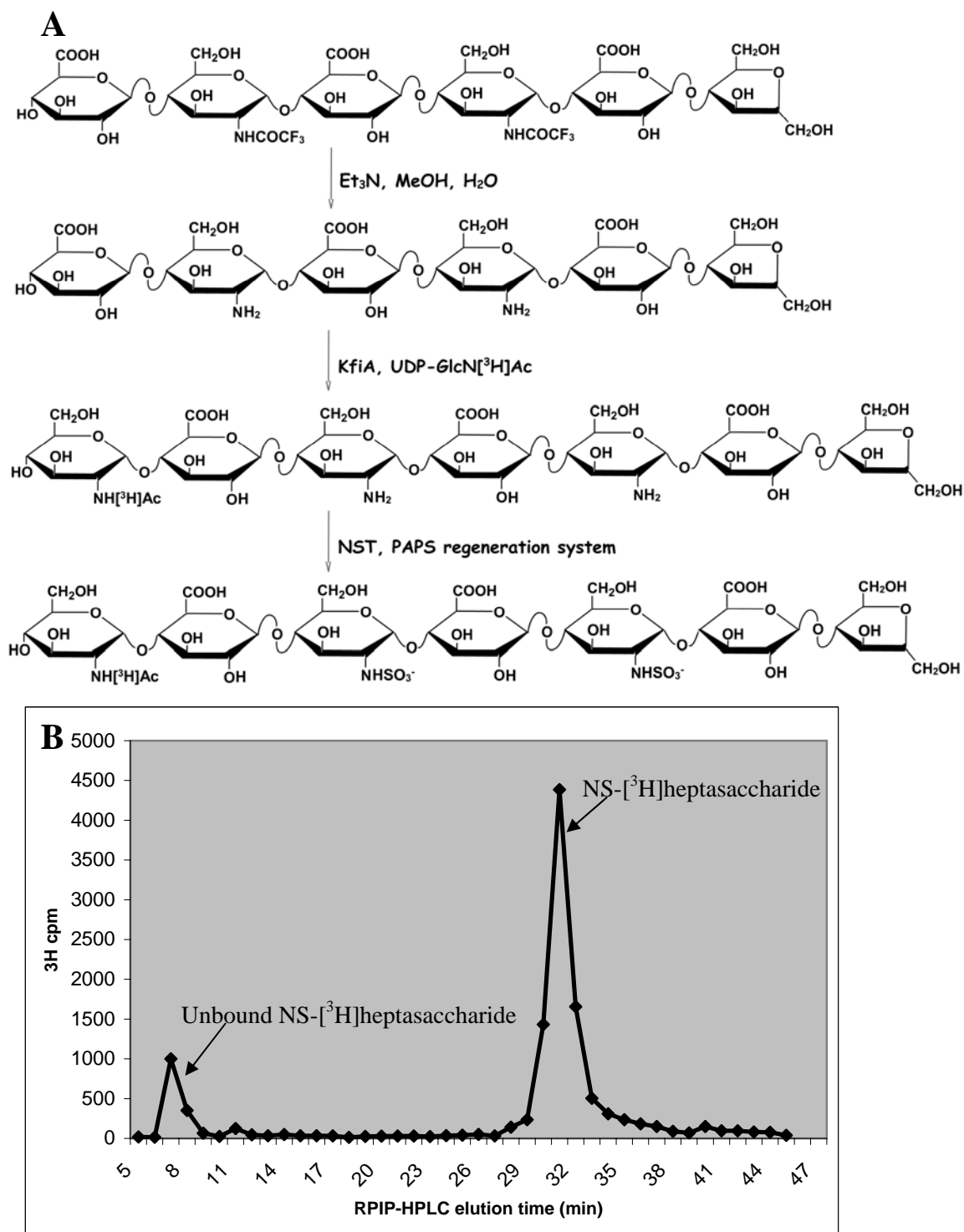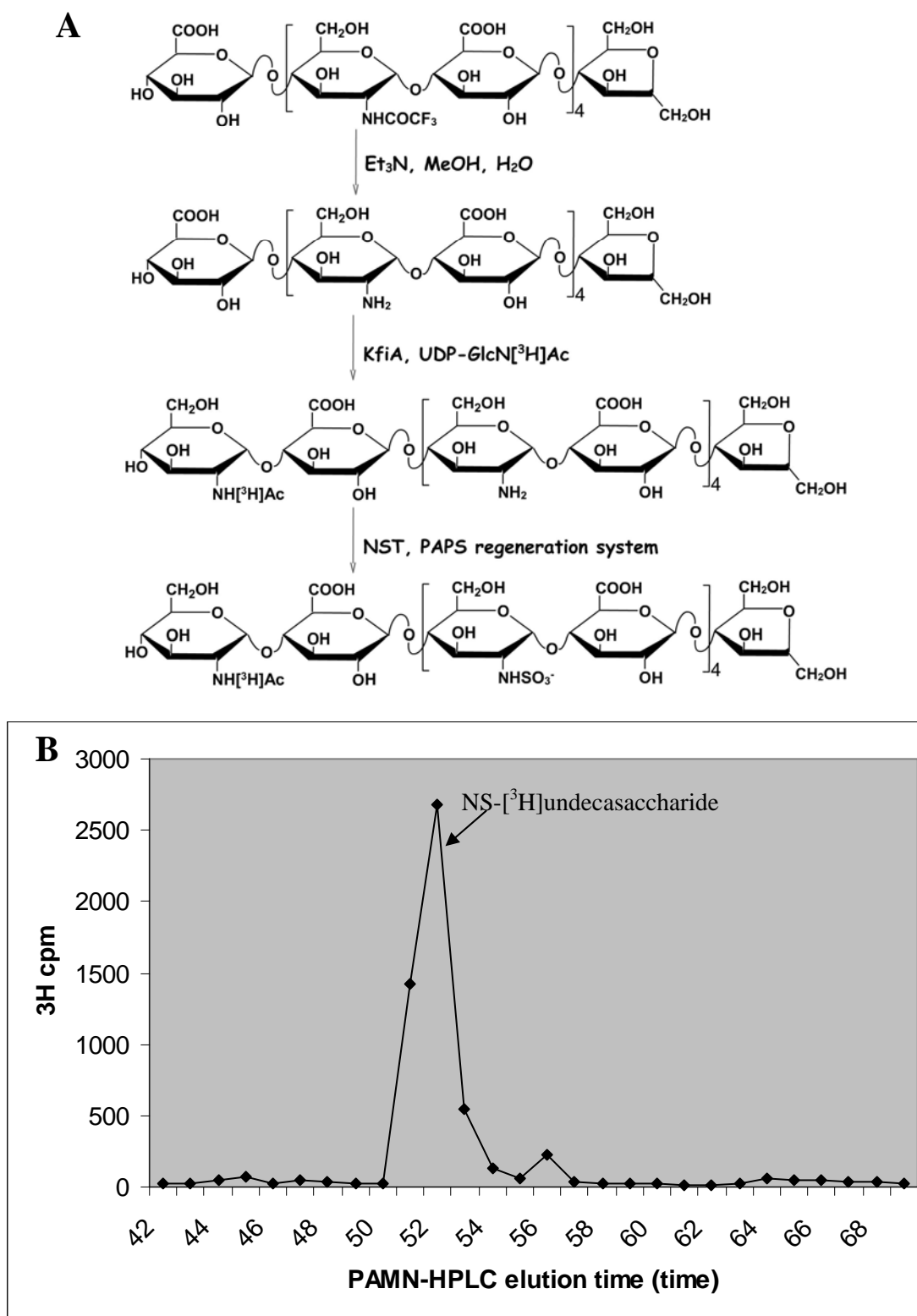
Figure 40. Preparation of *N*-deacetylated/*N*-sulfated [^{3}H]-undecasaccharide

A. Synthetic route

B. Purification of the *N*-sulfated [^{3}H]undecasaccharide (51-54min)

The extent of *N*-sulfation on the resultant undecasaccharide was examined by high pH and low pH nitrous acid degradation before and after the NST reaction (figure 41A, B). Similar to previous observations, the unsulfated undecasaccharide was intact when treated with nitrous acid at low pH, but was completely degraded to a [^{3}H]-labeled trisaccharide at high pH (figure 41A).   On the contrary, the *N*-sulfated undecasaccharide became completely sensitive to low pH nitrous acid degradation and resistant to high pH nitrous acid depolymerization (figure 41B).   The results indicate that the synthesized NS-[^{3}H]undecasaccharide was indeed fully *N*-deacetylated/*N*-sulfated.   Meanwhile, a partially *N*-sulfated dodecasaccharide with a [^{14}C]GlcA residue at its nonreducing end was also treated with nitrous acid at low pH.   The degradation yielded both [^{14}C]tetrasaccharide and [^{14}C]disaccharide (figure 41C).   The result was consistent with our expected outcome.   Some of the glucosamine residues on this [^{14}C]dodecasaccharide do not have *N*-sulfation.   Therefore, some of the [^{14}C]dodecasaccharide are resistant to the complete degradation by nitrous acid at low pH.   The result also confirms that nitrous acid degradation is able to probe the incompletion of *N*-deacetylation/*N*-sulfation.
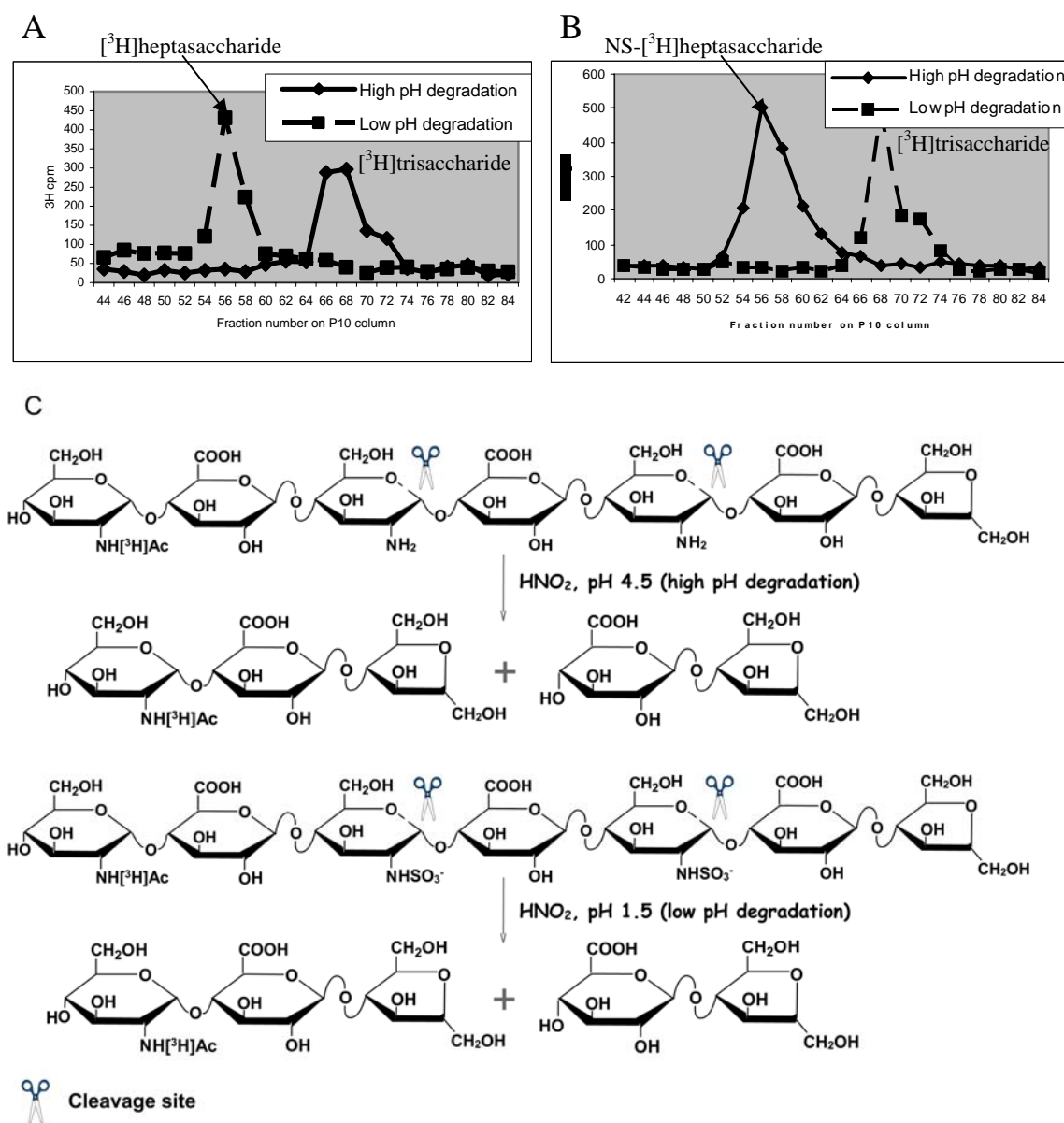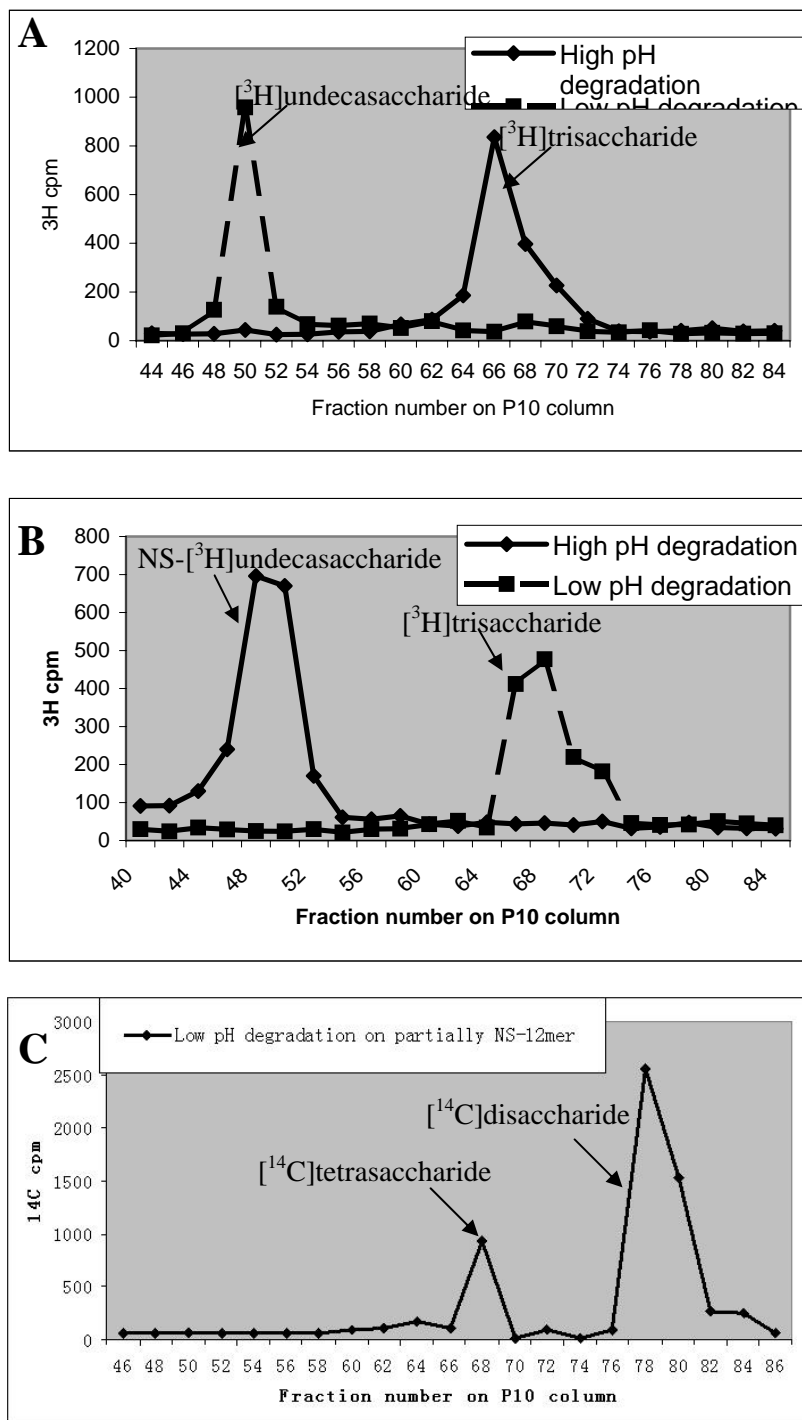
Figure 41. Nitrous acid degradation of [³H]undecasaccharide

A. low pH (dashed line) and high pH (solid line) nitrous acid degradation of *N*-deacetylated [³H]undecasaccharide

B. low pH (dashed line) and high pH (solid line) nitrous acid degradation of *N*-sulfated [³H]undecasaccharide

C. low pH nitrous acid degradation of partially *N*-sulfated [¹⁴C]dodecasaccharide

Section 3. Synthesis of heparan sulfate octasaccharide with AT-binding affinity

The Arixtra pentasaccharide (figure 5) is an AT-binding structure which works as a template for our synthesis. Using HS biosynthetic enzymes and the appropriately *N*-sulfated HS backbone, we synthesized HS octasaccharide containing this AT-binding domain by a route described in figure 42.
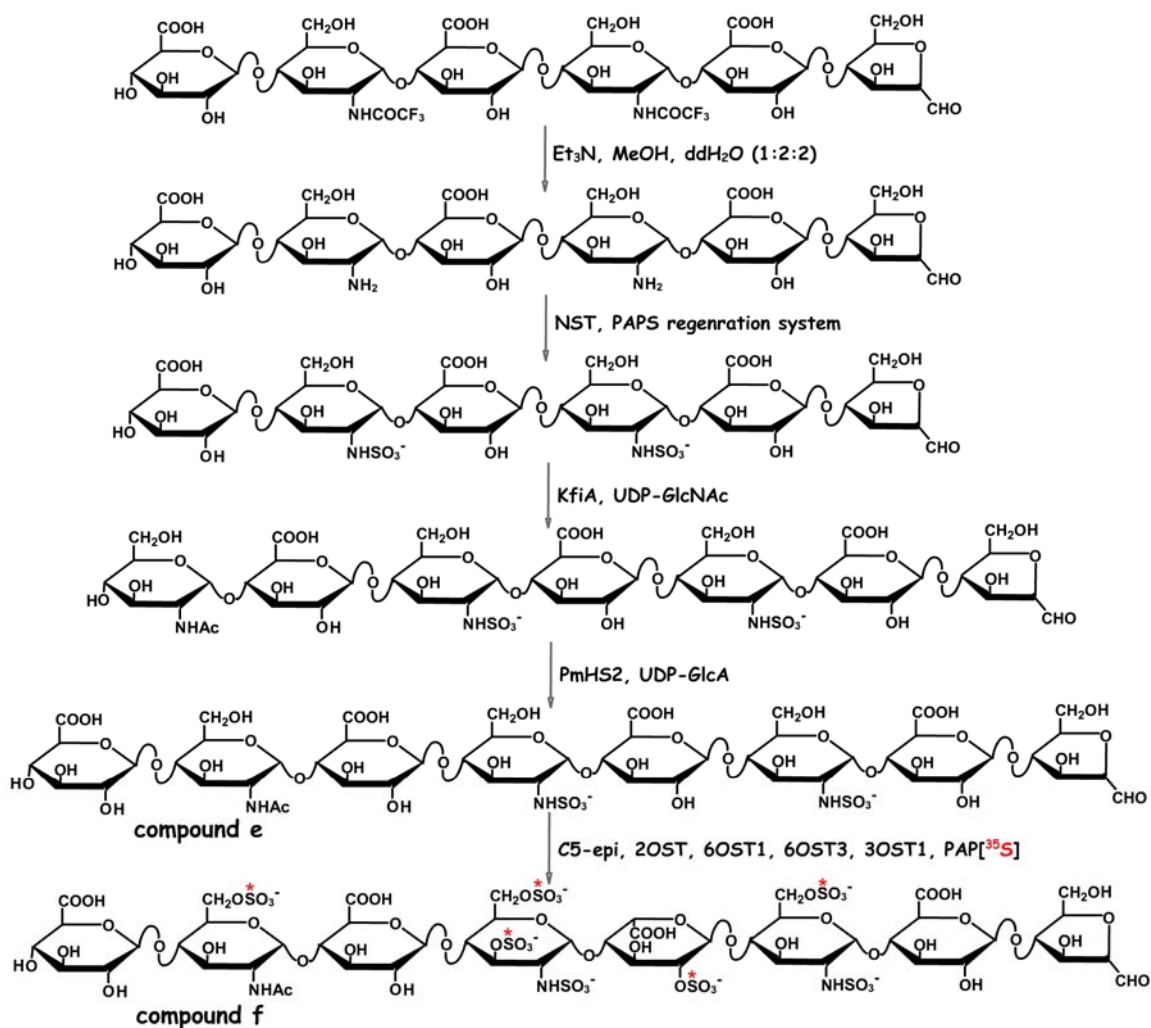
Figure 42. Synthetic route to the HS octasaccharide with AT-binding affinity

The synthesis started with an unnatural hexasaccharide containing GlcNCOCF$_3$ residues, which was followed by the complete *N*-deacetylation by triethylamine and complete *N*-sulfation by NST (figure 42).   The reaction yielded a completely *N*-sulfated hexasaccharide without radioactivity which was purified with a Bio-Gel P10 column. Its elution time/position on P10 column was already calibrated with an N[^{35}S]-hexasaccharide standard prepared from an NST reaction using PAP[^{35}S] (figure 43).   PAP[^{35}S] was eluted very late and far away from N[^{35}S]-hexasaccharide, suggesting that P10 column has sufficient capability to purify NS-hexasaccharide from other reaction reagent which was predominantly unreacted PAPS.
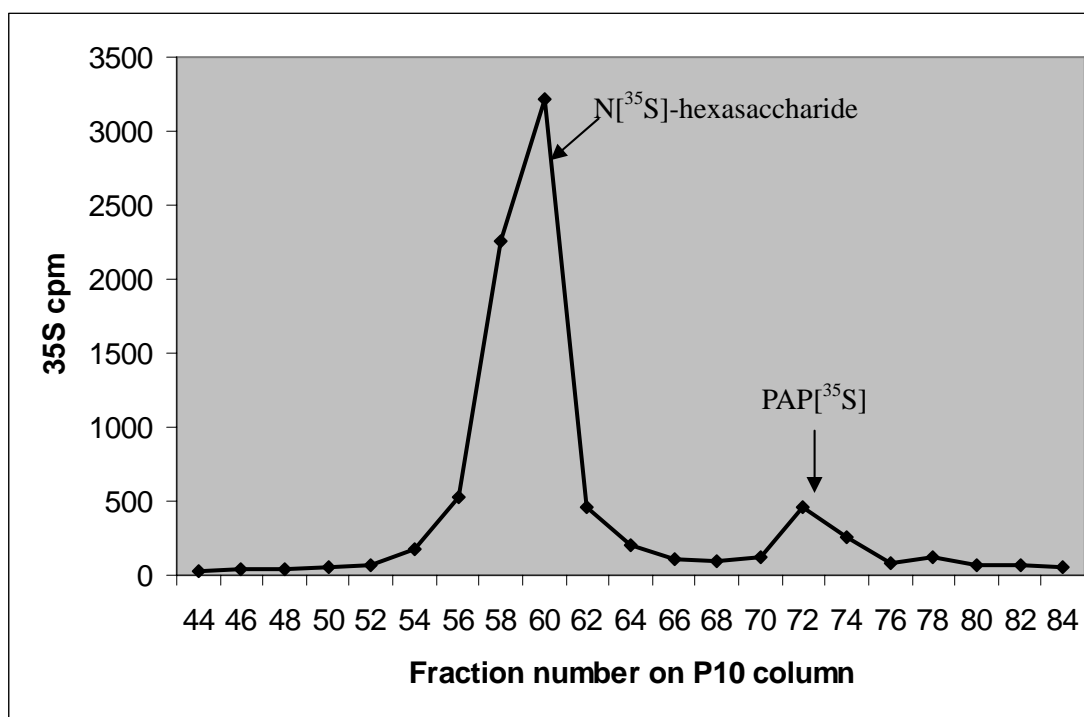
Figure 43. Elution time/position of N[^{35}S]-hexasaccharide on P10 column

N[^{35}S]-hexasaccharide (eluted from fraction number 56 to 62) was used as a standard to calibrate where NS-hexasaccharide is eluted on P10 column.   Unreacted PAP[^{35}S] was eluted around fraction number 73.

Considering the HS biosynthetic enzymes may need flanking region to efficiently react with the pentasaccharide domain, we added two more sugar residues to the non-reducing end of the *N*-sulfated hexasaccharide. These two extension reactions were catalyzed by KfiA and PmHS2, and the respective products were blindly purified with a P10 column based on their molecular size. The synthesized *N*-sulfated octasaccharide (figure 42, compound e) was subject to further modifications (epimerization, 2-*O*-sulfation, 6-*O*-sulfation, and 3-*O*-sulfation) in a one-pot reaction using [^{35}S]-labeled PAPS. To test if the resultant octasaccharide (figure 52, compound f) binds to AT, an assay was carried out with concanavalin A (ConA)-sepharose as described in "Materials and Methods". The synthesized octasaccharide had a reasonable binding affinity to AT. In contrast, the octasaccharide lost its AT-binding capability in the negative control even if only the last-step modification of 3-*O*-sulfation was neglected (table 7).

Starting substrate	Enzymatic modifications				AT-binding Octasaccharide (cpm)
	C5-epi	2OST	6OST1 6OST3	3OST1	
NS-octasaccharide	Yes	Yes	Yes	Yes	26220
(compound e)	Yes	Yes	Yes	No	350

Table 7. The synthesized HS octasaccharide binds to AT in a ConA-based assay

Collectively, our data demonstrated that the synthesized HS backbones with GlcNCOCF$_3$ residues can be fully *N*-deacetylated/*N*-sulfated in a chemoenzymatic approach. In addition, the resultant *N*-sulfated oligosaccharides can be effectively modified by HS biosynthetic enzymes (C5-epi, 2OST, 6OSTs and 3OST1). The modified final product binds to AT.

Chapter VI. Conclusion

Widely distributed on the cell surface and in the extracellular matrix, HS plays important roles in various biological processes, including cell differentiation, signaling, inflammatory response, blood coagulation and pathogen invasion (Bishop 2007). The significant medical potential of HS makes it an attractive target for pharmaceutical research. Heparin, a specialized form of HS, has been a commonly used anticoagulant drug for decades and HS is also recently employed in treating cancers and inflammatory disorders (McLean 1916; Fuster 2005; Parish 2005). However, as a natural product extracted from animal tissues, HS may lead to risky side effects due to its heterogeneous structure and potential contamination. A recent medical accident was reported that heparin distributed by Baxter International was contaminated and has caused 81 deaths in the United States according to the report by The New York Times (The New York Times April 28, 2008).

To improve the homogeneity of heparin-based drugs and their quality control, one of the major focuses in the field of carbohydrate research has been the development of synthetic substitute for heparin. Chemical synthesis has been an important approach to make small molecules but remains inefficient in the synthesis of oligosaccharides larger

than hexasaccharide.   Our group and others have begun to synthesize bioactive HS with a collection of its biosynthetic enzymes (Kuberan 2003; Chen 2005; Lindahl 2005; Muñoz 2006), although these approaches were at the polysaccharide level and the synthesized HS were still heterogeneous.

Busse and Kusche-Gullberg used mammalian HS polymerases EXT1 and EXT1/2 to synthesize HS polysaccharide backbones *in vitro* (Busse 2003).   In my research a chemoenzymatic method has been developed first time to synthesize HS oligosaccharides with desired size and defined *N*-sulfation.   My approach exploits the two bacterial glycosyltransferases, KfiA (Hodson 2000) and PmHS2 (DeAngelis 2004), to build the HS backbones *de novo* using a K5 polysaccharide-derived disaccharide (GlcA-anMan$_R$) as the staring substrate.   The enzymatic assays in Chapter III and IV demonstrated that both KfiA and PmHS2 are active on small oligosaccharides (disaccharide and trisaccharide) and capable of initiating the polymerization of sugar chain.   The size of the synthesized oligosaccharides was precisely controlled in the stepwise elongation by KfiA and PmHS2.   In addition, long oligosaccharides larger than decasaccharide can also be easily synthesized with this method, providing a valuable alternative for chemical synthesis.   Besides KfiA and PmHS2, the recombinant KfiC also demonstrates a notable GlcA transferase activity when it was coexpressed with KfiA in *E. coli* and a heptasaccharide substrate was used in Chapter IV.   However, it cannot work on trisaccharide, suggesting that KfiC may only be responsible for the elongation of the initiated sugar chain *in vivo*.

Another advantage of the *de novo* synthesis is that it facilitates the substitution of the unnatural sugar residue for the natural one on the HS backbones. Although KfiA can not directly incorporate GlcNS into the growing backbone, we found that KfiA can use UDP-GlcNCOCF$_3$ as an unnatural donor in the GlcNAc transferase reaction as efficiently as UDP-GlcNAc. In the *de novo* synthesis, the GlcNCOCF$_3$ residue can be placed at any desired glucosamine position on HS backbones and is labile to *N*-deacetylation with the treatment of triethylamine. The natural GlcNAc residue is, however, resistant to this treatment and remains intact. Because NST only reacts with the *N*-unsubstituted glucosamine, the selective *N*-deacetylation results in the selective *N*-sulfation. In this way, *N*-sulfation on the HS backbones can be precisely regulated. The subsequent epimerization and *O*-sulfations dictated by *N*-sulfation may also be controlled as demonstrated by the AT-binding octasaccharide in Chapter V (figure 52, compound f).

The specific function of HS relies on its particular sulfation patterns which results from the modifications catalyzed by various HS biosynthetic enzymes. Unlike DNA and protein, the biosynthesis of HS is not a template-driven process. A complete picture on the substrate specificities of HS modification enzymes has not been obtained due to the limited availability of structure-defined HS substrates. The lack of this knowledge prevents the researchers from clarifying the regulating mechanisms of the biosynthesis of HS. This dissertation proposed a feasible chemoenzymatic approach to synthesize oligosaccharide backbones with desired size and controlled *N*-sulfation. These compounds may serve as potential substrates to study the substrate specificities of HS

biosynthetic enzymes, including *C*5-epi, 2-*O*, 6-*O*, and 3-*O*-sulfotransferases, at the oligosaccharide level. This type of research will help us understand how the size of HS backbones and the distribution of its *N*-sulfations control the extent of downstream epimerization and *O*-sulfations. This will ultimately help to answer how the specific HS structures are generated and regulated *in vivo*.

The chemoenzymatic method developed in this dissertation can also be used to prepare an oligosaccharide library with various size and diverse *N*-sulfation patterns. In a recent study, Chen and colleagues found that the size of the 3-*O*-sulfated HS oligosaccharides is critical for their anticoagulant characters (Chen 2007). Therefore, such a library will be valuable to probe the structure-function relationship of bioactive HS. As an example, a target library containing pentasaccharides, hexasaccharides and heptasaccharides can be summarized in table 8.

Size	Structure	Number of NS
Pentasaccharide	GlcNAc-GlcA-GlcNAc-GlcA-anMan$_R$	0
	GlcNS-GlcA-GlcNAc-GlcA- anMan$_R$	1
	GlcNAc-GlcA-GlcNS-GlcA-anMan$_R$	1
	GlcNS-GlcA-GlcNS-GlcA-anMan$_R$	2
Hexasaccharide	GlcA-GlcNAc-GlcA-GlcNAc-GlcA-anMan$_R$	0
	GlcA-GlcNS-GlcA-GlcNAc-GlcA- anMan$_R$	1
	GlcA-GlcNAc-GlcA-GlcNS-GlcA-anMan$_R$	1
	GlcA-GlcNS-GlcA-GlcNS-GlcA-anMan$_R$	2
Heptasaccharide	GlcNAc-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-anMan$_R$	0
	GlcNS-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-anMan$_R$	1
	GlcNAc-GlcA-GlcNS-GlcA-GlcNAc-GlcA-anMan$_R$	1
	GlcNAc-GlcA-GlcNAc-GlcA-GlcNS-GlcA-anMan$_R$	1
	GlcNS-GlcA-GlcNS-GlcA-GlcNAc-GlcA-anMan$_R$	2
	GlcNS-GlcA-GlcNAc-GlcA-GlcNS-GlcA-anMan$_R$	2
	GlcNAc-GlcA-GlcNS-GlcA-GlcNS-GlcA-anMan$_R$	2
	GlcNS-GlcA-GlcNS-GlcA-GlcNS-GlcA-anMan$_R$	3

Table 8. The list of the oligosaccharides included in a target library
The *N*-sulfated glucosamine (GlcNS) residues are shown in red.

Our approach yielded the structure-defined HS oligosaccharides at 100 μg level.   In future studies, the synthesis should be scaled up with larger quantities of acceptor and donor substrates.   In addition, a fluorous tagging strategy is being developed in our group to simplify the purification of the synthesized oligosaccharides (Zhang 2004).   We expect the yield of each elongation step may be improved with the aid of this phase-tag method.   More importantly, the reduced purification time with the fluorous tagging strategy makes it less demanding to prepare the oligosaccharide library mentioned in table 8.

Besides the AT-binding octasaccharide, the chemoenzymatic method described in this dissertation may be used to synthesize other bioactive HS oligosaccharides as a general approach.   This exploratory study is the foundation for *de novo* synthesis of the structure-defined HS oligosaccharides with various biological functions, which may become novel therapeutic agents for cardiovascular diseases, cancers and inflammatory disorders.

APPENDIX 1

Curriculum Vitae

Miao Chen

School of Pharmacy, CB #7360   Work: (919)-962-0065

University of North Carolina at Chapel Hill   Fax: (919)-843-5432

Chapel Hill, NC 27599-7360   Email: mchen@unc.edu

Education

August 2003-August 2008 (expected), **Ph.D. in Pharmaceutical Sciences**
Dissertation title: Towards *de novo* synthesis of structure-defined oligosaccharides with heparan sulfate biosynthetic enzymes,   Advisor: Dr. Jian Liu
University of North Carolina (UNC), Chapel Hill, NC

September 1998–July 2003, **B.S. in Molecular Biology**
University of Science and Technology of China (USTC), Hefei, Anhui, China

Experiences

January 2004-Present, **Research Assistant**, UNC School of Pharmacy, Chapel Hill, NC
Projects: 1) Expression and characterization of the *N*-acetylglucosaminyltransferase KfiA;
2) Expression of the glucuronyltransferase PmHS2; 3) Synthesis of the heparan sulfate (HS) backbones containing either *N*-acetylglucosamine or *N*-trifluoroacetylglucosamine;
4) Modification of the HS backbones with its biosynthetic enzymes.

July 2002–June 2003, **Research Intern**, Chinese Academy of Sciences, Shanghai, China
Project: Phospholipase A_2 involved in the signal transduction pathway of *Arabidopsis Thaliana*

Awards and Honors

1999–2002, Outstanding Student Scholarship (OSS), USTC

Publications

Miao Chen, Arlene Bridges, and Jian Liu (2006). Determination of the substrate specificities of *N*-Acetyl-D-glucosaminyltransferase. *Biochemistry*, 45: 12358-12365.

Jinghua Chen, Fikri Y. Avci, Eva M. Muñoz, Lynda M. McDowell, **Miao Chen**, Lars C. Pedersen, Lijuan Zhang, Robert J. Linhardt, and Jian Liu (2005). Enzymatic redesigning of biologically active heparan sulfate. *Journal of Biological Chemistry*. 280: 42817-42825.

Zhenqing Zhang, Scott A. McCallum, Jin Xie, Lidia Nieto, Francisco Corzana, Jesús Jiménez-Barbero, **Miao Chen**, Jian Liu, and Robert J. Linhardt (2008). Solution structures of chemoenzymatically synthesized heparin and its precursors. *Journal of the American Chemical Society* (submitted).

REFERENCES

Aikawa, J. I., Grobe, K., Tsujimoto, M., and Esko, J. D. (2001). "Multiple isozymes of heparan sulfates/heparin GlcNAc N-deacetylase/GlcN N-sulfotransferase: Structure and activity of the fourth member, NDST4." Journal of Biological Chemistry **276**: 5876-5882.

Ashikari-Hada, S., Habuchi, H., Kariya, Y., Itoh, N., Reddi, A.H., and Kimata, K. (2004). "Characterization of Growth Factor-binding Structures in Heparin/Heparan Sulfate Using an Octasaccharide Library." Journal of Biological Chemistry **279**(13): 12346-12354.

Atha, D. H., Lormeau, J.-C., Petitou, M., Rosenberg, R. D., and Choay, J. (1985). "Contribution of monosaccharide residues in heparin binding to antithrombin III." Biochemistry **24**(23): 6723-6729.

Avci, F. Y., Karst, N. A., and Linhardt, R. J. (2003). "Synthetic Oligosaccharides as Heparin-Mimetics Displaying Anticoagulant Properties." Current Pharmaceutical Design **9**(28): 2323-2335.

Bai, X., Zhou, D., Brown, J. R., Crawford, B. E., Hennet, T., and Esko, J. D. (2001). "Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth member of the b1,3-galactosyltransferase family (b3GalT6)." Journal of Biological Chemistry **276**: 48189-48195.

Basilico, C., and Moscatelli, D. (1992). "The FGF family of growth factors and oncogenes." Advances in Cancer Research **59**: 115-165.

Bishop, J., Schuksz, M., and Esko, J. D. (2007). "Heparan sulfate proteoglycans fine-tune mammalian physiology." Nature **446**: 1030-1037.

Brickman, Y. G., Ford, M.D., Gallagher, J.T., Nurcombe, V., Bartlett, P.F. and Turnbull, J.E. (1998). "Structural Modification of Fibroblast Growth Factor-binding Heparan Sulfate at a Determinative Stage of Neural Development." Journal of Biological Chemistry **273**(8): 4350-4359.

Broze, G. J. J. (1995). "Tissue factor pathway inhibitor and the revised theory of coagulation." Annual Review of Medicine **46**: 103-112.

Burkart, M. D., Izumi, M., and Wong, C. H. (1999). "Enzymatic regeneration of 3'-phosphoadenosine-5'-phosphosulfate using aryl sulfo-transferase for the preparative enzymatic synthesis of sulfated carbohydrates." Angewandte Chemie, International

Edition **38**(18): 2747-2750.

Burkart, M. D., Izumi, M., Chapman, E., Lin, C. H., and Wong, C. H. (2000). "Regeneration of PAPS for the enzymatic synthesis of sulfated oligosaccharides." Journal of Organic Chemistry **65**: 5565-5574.

Busse, M., and Kusche-Gullberg, M. (2003). "In Vitro Polymerization of Heparan Sulfate Backbone by the EXT Proteins." Journal of Biological Chemistry **278**(42): 41333-41337.

Chen, J., Avci, F. Y., Muñoz, E. M., McDowell, L. M., Chen, M., Pedersen, L. C., Zhang, L., Linhardt, R. J. and Liu, J. (2005). "Enzymatic redesigning of biological active heparan sulfate." Journal of Biological Chemistry **280**: 42817-42825.

Chen, J., Jones, C., and Liu, J. (2007). "Using an Enzymatic Combinatorial Approach to Identify Anticoagulant Heparan Sulfate Structures." Chemistry and Biology **14**(9): 986-993.

Chen, M., Bridges, A., and Liu, J. (2006). "Determination of the substrate specificities of N-acetyl-D-glucosaminyltransferase." Biochemistry **45**(40): 12358-12365.

Copeland, R., BalasubranManiam, A., Tiwari, V., Zhang, F., Bridges, A., Linhardt, R. J., Shukla, D., and Liu, J. (2008). "Using a 3-O-Sulfated Heparin Octasaccharide To Inhibit the Entry of Herpes Simplex Virus Type 1." Biochemistry **47**(21): 5774-5783.

Costerton, J. W., Cheng, K. J., Geesey, G. G., Ladd, T. I., Nickel, J. C., Dasgupta, M., Marrie, T. J. (1987). "Bacterial biofilms in nature and disease." Annual Review of Microbiology **41**: 435-464.

Crawford, B. E., Olson, S. K., Esko, J. D., and Pinhal, M. A. S. (2001). "Cloning, Golgi Localization, and Enzyme Activity of the Full-length Heparin/Heparan Sulfate-Glucuronic Acid C5-epimerase." Journal of Biological Chemistry **276**: 21538-21543.

Dai, Y., Yang, Y., MacLeod, V., Yue, X., Rapraeger, A. C., Shriver, Z., VenkataranMan, G., Sasisekharan, R., and Sanderson, R. D. (2005). "HSulf-1 and HSulf-2 Are Potent Inhibitors of Myeloma Tumor Growth in Vivo." Journal of Biological Chemistry **280**(48): 40066-40073.

DeAngelis, P. L. (2002). "Microbial glycosaminoglycan glycosyltransferases." Glycobiology **12**(1): 9R-16R.

DeAngelis, P. L., and Carissa L. White (2002). "Identification and Molecular Cloning of a Heparosan Synthase from Pasteurella multocida Type D." Journal of Biological Chemistry **277**(9): 7209-7213.

DeAngelis, P. L., and White, C. L. (2004). "Identification of a Distinct, Cryptic Heparosan Synthase from Pasteurella multocida Types A, D, and F." Journal of Bacteriology **186**(24): 8529-8532.

Denny, T. P., and Baek, S. R. (1991). "Genetic evidence that extracellular polysaccharide is a virulence factor of Pseudomonas solanacearum." Molecular Plant-Microbe Interactions **4**: 198-206.

Desai, U. R., Wang, H. M., and Linhardt, R. J. (1993). "Substrate specificity of the heparin lyases from Flavobacterium heparinum." Archives of Biochemistry and Biophysics **306**(2): 461-468.

DianMant, Z., Timmers, M.C., van der Veen, H., Page, C.P., van der Meer, F.J., Sterk, P.J. (1996). "Effect of inhaled heparin on allergen-induced early and late asthmatic responses in patients with atopic asthma." American Journal of Respiratory and Critical Care Medicine **153**: 1790-1795.

Dixon, J., Loftus, S. K., Gladwin, A. J., Scambler, P. J., Wasmuth, J. J., and Dixon, M. J. (1995). "Cloning of the human heparan sulfate-N-deacetylase/N-sulfotransferase gene from the Treacher Collins syndrome candidate region at 5q32-q33.1." Genomics **26**: 239-244.

Duncan, M. B., Liu, M., Fox, C., and Liu, J. (2006). "Characterization of the N-deacetylase domain from the heparan sulfate N-deacetylase/N-sulfotransferase 2." Biochemical and Biophysical Research Communications **339**(4): 1232-1237.

Edavettal, S. C., Lee, K. A., Negishi, M., Linhardt, R. J., Liu, J., and Pedersen, L. C. (2004). "Crystal structure and mutational analysis of heparan sulfate 3-O-sulfotransferase isoform 1." Journal of Biological Chemistry **279**: 25789-25797.

Edovitsky, E., Elkin, M., Zcharia, E., Peretz, T., and Vlodavsky, I. (2004). "Heparanase Gene Silencing, Tumor Invasiveness, Angiogenesis, and Metastasis." Journal of the National Cancer Institute **96**: 1219-1230.

Elkin, M., Ilan, N., Ishai-Michaeli, R., Friedmann, Y., Papo, O., Pecker, I., and Vlodavsky, I. (2001). "Heparanase as mediator of angiogenesis: mode of action." The FASEB Journal **15**(9): 1661-1663.

Esko, J. D., and Lindahl, U. (2001). "Molecular diversity of heparan sulfate." <u>Journal of Clinical Investigation</u> **108**: 169-173.

Esko, J. D., and Selleck, S.B. (2002). "Order out of chaos: Assembly of ligand binding sites in heparan sulfate." <u>Annual Review of Biochemistry</u> **71**: 435-471.

Esko, J. D., Stewart, T. E., and Taylor, W. H. (1985). "Animal cell mutants defective in glycosaminoglycan biosynthesis." <u>Proceedings of the National Academy of Sciences</u> **82**: 3197-3201.

Fareed, J., Jeske, W., Hoppensteadt, D., Clarizio, R, Walgenga, J. M. (1998). "Low molecular weight heparins: pharmacologic profile and product differentiation." <u>American Journal of Cardiology</u> **82**: 3L-10L.

Ferro, V., Dredge, K., Liu, L., Hammond, E., Bytheway, I., Li, C., Johnstone, K., Karoli, T., Davis, K., Copeman, E., and Gautam, A. (2007). "PI-88 and novel heparan sulfate mimetics inhibit angiogenesis." <u>Seminars in Thrombosis and Hemostasis</u> **3**: 557-568.

Ferro, V., Hammond, E., and Fairweather, J. K. (2004). "The development of inhibitors of heparanase, a key enzyme involved in tumour metastasis, angiogenesis and inflammation." <u>Mini Reviews in Medicinal Chemistry</u> **4**: 693-702.

Fuster, M. M., and Esko, J. D. (2005). "The sweet and sour of cancer: Glycans as novel therapeutic targets." <u>Nature Reviews Cancer</u> **5**: 526-542.

Gama, C. I., and Hsieh-Wilson, L. C. (2005). "Chemical approaches to deciphering the glycosaminoglycan code." <u>Current Opinion in Chemical Biology</u> **9**(6): 609-619.

Griffiths, G., Cook, N. J., Gottfridson, E., Lind, T., Lidholt, K., and Roberts, I. S. (1998). "Characterization of the glycosyltransferase enzyme from the Escherichia coli K5 capsule gene cluster and identification and characterization of the glucuronyl active site." <u>Journal of Biological Chemistry</u> **273**: 11752-11757.

Guerrini, M., Agulles, T., Bisio, A., Hricovini, M., Lay, L., Naggi, A., Poletti, L., Sturiale, L., Torri, G., and Casu, B. (2002). "Minimal heparin/heparan sulfate sequences for binding to fibroblast growth factor-1." <u>Biochemical and Biophysical Research Communications</u> **292**(1): 222-230.

Gulberti, S., Lattard, V., Fondeur, M., Jacquinet, J. C., Mulliert, G, Netter, P., Magdalou, J., Ouzzine, M., and Fournel-Gigleux, S. (2005). "Phosphorylation and Sulfation of

Oligosaccharide Substrates Critically Influence the Activity of Human {beta}1,4-Galactosyltransferase 7 (GalT-I) and {beta}1,3-Glucuronosyltransferase I (GlcAT-I) Involved in the Biosynthesis of the Glycosaminoglycan-Protein Linkage Region of Proteoglycans." Journal of Biological Chemistry **280**(2): 1417-1425.

Hérias, M. V., Midtvedt, T., Hanson, L. A., and Wold, A. E. (1997). "Escherichia coli K5 capsule expression enhances colonization of the large intestine in the gnotobiotic rat." Infection and Immunity **65**(2): 531-536.

Habuchi, H., Kobayashi, M., and Kimata, K. (1998). "Molecular characterization and expression of heparan-sulfate 6-sulfotransferase: complete cDNA cloning in human and partial cloning in Chinese hamster ovary cells." Journal of Biological Chemistry **273**: 9208-9213.

Habuchi, H., Tanaka, M., Habuchi, O., Yoshida, K., Suzuki, H., Ban, K., and Kimata, K. (2000). "The occurrence of three isoforms of heparan sulfate 6-O-sulfotransferase having different specificities for hexuronic acid adjacent to the targeted N-sulfoglucosamine." Journal of Biological Chemistry **275**(4): 2859-2868.

Hagner-McWhirter, A., Lindahl, U., and Li, J. (2000). "Biosynthesis of heparin/heparan sulphate: mechanism of epimerization of glucuronyl C-5." Biochemical Journal **347**: 69-75.

HajMohammadi, S., Enjyoji, K., Princivalle, M., Christi, P., Lech, M., Beeler, D. L., Rayburn, H., Schwartz, J. J., Barzegar, S., de Agostini, A. I., Post, M. J., Rosenberg, R. D., and Shworak, N. W. (2003). "Normal levels of anticoagulant heparan sulfate are not essential for normal hemostasis." Journal of Clinical Investigation **111**(7): 989-999.

Haller, M. a. B., G. J. (2001). "Towards a modular approach for heparin synthesis." Journal of Chemical Society, Perkin Transactions **1**: 814-822.

Hanahan, D., and Weinberg, R.A. (2000). "The hallmarks of cancer." Cell **100**(1): 57-70.

Hodson, N., Griffiths, G., Cook, N., Pourhossein, M., Gottfridson, E., Lind, T., Lidholt, K., and Roberts, I. S. (2000). "Identification that KfiA, a protein essential for the biosynthesis of the Escherichia coli K5 capsular polysaccharide, is an alpha -UDP-GlcNAc glycosyltransferase. The formation of a membrane-associated K5 biosynthetic complex requires KfiA, KfiB, and KfiC." Journal of Biological Chemistry **275**: 27311-27315.

Ilana, N., Elkinb, M., and Vlodavsky, I. (2006). "Regulation, function and clinical

significance of heparanase in cancer metastasis and angiogenesis." International Journal of Biochemistry and Cell Biology **38**(12): 2018-2039.

Iozzo, R. V. (2001). "Series Introduction: Heparan sulfate proteoglycans: intricate molecules with intriguing functions." Journal of Clinical Investigation **108**(2): 165-167.

Jacobsson, I., Lindahl, U., Jensen, J. W., Roden, L., Prihar, H., and Feingold, D. S. (1984). "Biosynthesis of heparin. Substrate specificity of heparosan N-sulfate D-glucuronosyl 5-epimerase." Journal of Biological Chemistry **259**: 1056-1063.

Jacquinet, J. C. (2004). "An expeditious preparation of various sulfoforms of the disaccharide ß-D-Galp-(1-->3)-D-Galp, a partial structure of the linkage region of proteoglycans, as their 4-methoxyphenyl ß-D-glycosides." Carbohydrate Research **339**: 349-359.

Kakuta, Y., Li, L., Pedersen, L. C., Pedersen, L. G., and Negishi, M. (2003). "Heparan sulphate N-sulphotransferase activity: reaction mechanism and substrate recognition." Biochemical Society Transactions **31**(2): 331.

Kakuta, Y., Pedersen, L. G., Carter, C. W., Negishi, M., and Pedersen, L. C. (1997). "Crystal structure of estrogen sulphotransferase." Nature Structural Biology **4**: 904-908.

Kakuta, Y., Petrotchenko, E. V., Pedersen, L. C., and Negishi, M. (1998). "The sulfuryl transfer mechanism: crystal structures of a vanadate complex of estrogen sulfotransferase and mutational analysis." Journal of Biological Chemistry **273**: 27325-27330.

Kakuta, Y., Sueyoshi, T., Negishi, M., and Pedersen, L. C. (1999). "Crystal Structure of the Sulfotransferase Domain of Human Heparan Sulfate N-Deacetylase/ N-Sulfotransferase 1." Journal of Biological Chemistry **274**(16): 10673-10676.

Kamimura, K., Koyama, T., Habuchi, H., Ueda, R., Masu, M., Kimata, K., and Nakato, H. (2006). "Specific and flexible roles of heparan sulfate modifications in Drosophila FGF signaling." Journal of Cell Biology **174**(6): 773-778.

Kane, T. A., White, C. L., and DeAngelis, P. L. (2006). "Functional Characterization of PmHS1, a Pasteurella multocida Heparosan Synthase." Journal of Biological Chemistry **281**(44): 33192-33197.

Karst, N. A., and Linhardt, R. J. (2003). "Recent chemical and enzymatic approaches to the synthesis of glycosaminoglycan oligosaccharides." Current Medicinal Chemistry **10**: 1993-2031.

Kim, B. T., Kitagawa, H., Tamura, J.-i., Saito, T., Kusche-Gullberg, M., Lindahl, U., and Sugahara, K. (2001). "Human tumor suppressor EXT gene family members EXTL1 and EXTL3 encode alpha 1,4- N-acetylglucosaminyltransferases that likely are involved in heparan sulfate/ heparin biosynthesis." Proceedings of the National Academy of Sciences **98**(13): 7176-7181.

Kim, B. T., Kitagawa, H., Tanaka, J., Tamura, J. I., and Sugahara, K. (2003). "In Vitro Heparan Sulfate Polymerization. CRUCIAL ROLES OF CORE PROTEIN MOIETIES OF PRIMER SUBSTRATES IN ADDITION TO THE EXT1-EXT2 INTERACTION." Journal of Biological Chemistry **278**(43): 41618-41623.

Kitagawa, H., Shimakawa, H., and Sugahara, K. (1999). "The Tumor Suppressor EXT-like Gene EXTL2 Encodes an alpha 1, 4-N-Acetylhexosaminyltransferase That Transfers N-Acetylgalactosamine and N-Acetylglucosamine to the Common Glycosaminoglycan-Protein Linkage Region. THE KEY ENZYME FOR THE CHAIN INITIATION OF HEPARAN SULFATE." Journal of Biological Chemistry **274**(20): 13933-13937.

Kuberan, B., Beeler, D. L., Lawrence, R., Lech, M., and Rosenberg, R. (2003). "Rapid two-step synthesis of mitrin from heparosan: a replacement for heparin." Journal of the American Chemical Society **125**: 12424-12425.

Kuberan, B., Lech, M. Z., Beeler, D. L., Wu, Z. L. and Rosenberg, R. D. (2003). "Enzymatic synthesis of antithrombin III−binding heparan sulfate pentasaccharide." Nature Biotechnology(21): 1343-1346.

Kuberan, B., Lech, M.Z., Beeler, D.L., Wu, Z.L., and Rosenberg, R.D. (2003). "Enzymatic synthesis of antithrombin III-binding heparan sulfate pentasaccharide." Nature Biotechnology **21**: 1343-1346.

Kwan, C. P., VenkataranMan, G., Shriver, Z., RanMan, R., Liu, D., Qi, Y., Varticovski, L., and Sasisekharan, R. (2001). "Probing Fibroblast Growth Factor Dimerization and Role of Heparin-like Glycosaminoglycans in Modulating Dimerization and Signaling " Journal of Biological Chemistry **276**(26): 23421-23429.

Lane, D. A., Denton, J., Flynn, A. M., Thunberg, L., and Lindahl, U. (1984). "Anticoagulant activities of heparin oligosaccharides and their neutralization by platelet factor 4." Biochemical Journal **218**: 725-732.

Leiting, B., Pryor, K. D., Eveland, S. S., and Anderson, M. S. (1998). "One-day

enzymatic synthesis and purification of UDP-N-[1-14C]acetyl-glucosamine." <u>Analytical Biochemistry</u> **256**: 185-191.

Leppa, S., Mali, M., Miettinen, H. M., and Jalkanen, M. (1992). "Syndecan Expression Regulates Cell Morphology and Growth of Mouse Mammary Epithelial Tumor Cells." <u>Proceedings of the National Academy of Sciences</u> **89**: 932-936.

Lever, R., and Page, C.P. (2002). "Novel drug development opportunities for heparin." <u>Nature Reviews Drug Discovery</u> **1**(2): 140-148.

Li, J. P., Gong, F., Hagner-McWhirter, A., Forsberg, E., Abrink, M., Kisilevsky, R., Zhang, X., and Lindahl, U. (2003). "Targeted Disruption of a Murine Glucuronyl C5-epimerase Gene Results in Heparan Sulfate Lacking L-Iduronic Acid and in Neonatal Lethality." <u>Journal of Biological Chemistry</u> **278**(31): 28363-28366.

Liebersbach, B. F., and Sanderson, R.D. (1994). "Expression of syndecan-1 inhibits cell invasion into type I collagen." <u>Journal of Biological Chemistry</u> **269**: 20013-20019.

Lin, C. H., Shen, G. J., Garcia-Junceda, E., and Wong, C. H. (1995). "Enzymic Synthesis and Regeneration of 3'-Phosphoadenosine 5'-Phosphosulfate (PAPS) for Regioselective Sulfation of Oligosaccharides." <u>Journal of American Chemical Soceity</u> **117**(30): 8131-8132.

Lindahl, U., Backstrom, G., Thunberg, L., and Leder, I. G. (1980). "Evidence for a 3-O-sulfated d-glucosamine residue in the antithrombin-binding sequence of heparin." <u>Proceedings of the National Academy of Sciences</u> **77**: 6551-6555.

Lindahl, U., Li, J., Kusche-Gullberg, M., Salmivirta, M., Alaranta, S., Veromaa, T., Emies, J., Roberts, I., Taylor, C., Oreste, P., Zoppetti, G., Naggi, A., Torri, G., and Casu, B. (2005). "Generation of "neoheparin" from E. coli K5 capsular polysaccharide." <u>Journal of Medicinal Chemistry</u> **48**: 349-352.

[a]Linhardt, R., and Kim, J. (2007). "Combinatorial Enzymatic Synthesis of Heparan Sulfate." <u>Chemistry and Biology</u> **14**(9): 972-973.

[b]Linhardt, R., Dordick, J., Deangelis, P., and Liu, J. (2007). "Enzymatic synthesis of glycosaminoglycan heparin." <u>Seminars in Thrombosis and Hemostasis</u> **33**(5): 453-465.

Liu, J., and Rosenberg, R. D. (2002). "Heparan sulphate D-glucosaminyl 3-O-sulfotransferase-1, -2, -3, and -4." <u>Handbook of glycosyltransferases and their related genes</u>: 475-483.

Liu, J., and Thorp, S. C. (2002). "Cell surface heparan sulfate and its roles in assisting viral infections." <u>Medicinal Research Reviews</u> **22**(1): 1-25.

Liu, J., Z. Shriver, et al. (1999). "Heparan sulfate D-glucosaminyl 3-O-sulfotransferase-3A sulfates N-unsubstituted glucosamine residues." <u>J. Biol. Chem.</u> **274**: 38155-38162.

Liu, J., Shriver, Z., Pope, R. M., Thorp, S. C., Duncan, M. B., Copeland, R. J., Raska, C. S., Yoshida, K., Eisenberg, R. J., Cohen, G., Linhardt, R. J., and Sasisekharan, R. (2002). "Characterization of a Heparan Sulfate Octasaccharide That Binds to Herpes Simplex Virus Type 1 Glycoprotein D." <u>Journal of Biological Chemistry</u> **277**: 33456-33467.

Liu, J., Shworak, N. W., Sinaÿ, P., Schwartz, J. J., Zhang, L., Fritze, L. M. S., and Rosenberg, R. D. (1999). "Expression of Heparan Sulfate D-Glucosaminyl 3-O-Sulfotransferase Isoforms Reveals Novel Substrate Specificities." <u>Journal of Biological Chemistry</u> **274**(8): 5185-5192.

Liu, W., Litwack, E. D., Stanley, M. J., Langford, J. K., Lander, A. D., and Sanderson, R. D. (1998). "Heparan Sulfate Proteoglycans as Adhesive and Anti-invasive Molecules." <u>Journal of Biological Chemistry</u> **273**: 22825-22832.

Loo, B. M., Kreuger, J., Jalkanen, M., Lindahl, U., and Salmivirta, M. (2001). "Binding of Heparin/Heparan Sulfate to Fibroblast Growth Factor Receptor 4." <u>Journal of Biological Chemistry</u> **276**(20): 16868-16876.

McCormick, C., Duncan, G., Goutsos, K.   T., and Tufaro, F. (2000). "The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate." <u>Proceedings of the National Academy of Sciences</u> **97**(2): 668-673.

McLean, J. (1916). "The thromboplastic action of cephalin." <u>American Journal of Physiology</u> **41**: 250.

Merry, C. L. R., Bullock, S. L., Swan, D. C., Backen, A. C., Lyon, M., Beddington, R. S. P., Wilson, V. A., and Gallagher, J. T. (2001). "The Molecular Phenotype of Heparan Sulfate in the Hs2st-/- Mutant Mouse." <u>Journal of Biological Chemistry</u> **276**: 35429-35434.

Middleton, J., Neil, S., Wintle, J., Clark-Lewis, I., Moore, H., Lam, C., Auer, M., Hub, E, and Rot, A. (1997). "Transcytosis and Surface Presentation of IL-8 by Venular

Endothelial Cells." Cell **91**: 385-395.

Middleton, J., Patterson, A.M., Gardner, L., Schmutz, C., and Ashton, B.A. (2002). "Leukocyte extravasation: chemokine transport and presentation by the endothelium." Blood **100**: 3853-3860.

Moczar, M., Caux, F., Bailly, M., Berthier, O, and Doré, J-F (1993). " Accumulation of heparan sulfate in the culture of human melanoma cells with different metastatic ability." Clinical and Experimental Metastasis **11**: 462-471.

Moon, A., Edavettal, S. C., Krahn, J. X., Munoz, E. M., Negishi, M., Linhardt, R. J., Liu, J., and Pedersen, L. C. (2004). "Structural analysis of the sulfotransferase (3-OST-3) involved in the biosynthesis of an entry receptor of herpes simplex virus 1." Journal of Biological Chemistry **279**: 45185-45193.

Moxon, E., and Kroll, J. (1990). "The role of bacterial polysaccharide capsules as virulence factors." Current Topics in Microbiology and Immunology **150**: 65-85.

Muñoz, E., Xu, D., Avci, F., Kemp, M., Liu, J., and Linhardt, R. J. (2006). "Enzymatic synthesis of heparin related polysaccharides on sensor chips: Rapids screening of heparin-protein interactions." Biochemical and Biophysical Research Communications **339**: 597-602.

Mulloy, B., and Forster, M. J. (2000). "Conformation and dynamics of heparin and heparan sulfate." Glycobiology **10**(11): 1147-1156.

Olson, S. T., and Bjork, I. (1991). "Predominant contribution of surface approximation to the mechanism of heparin acceleration of the antithrombin-thrombin reaction. Elucidation from salt concentration effects." Journal of Biological Chemistry **266**(10): 6353-6364.

Olson, S. T., Bjork, I., Sheffer, R., Craig, P.A., Shore, J.D., and Choay, J. (1992). "Role of the antithrombin-binding pentasaccharide in heparin acceleration of antithrombin-proteinase reactions. Resolution of the antithrombin conformational change contribution to heparin rate enhancement " Journal of Biological Chemistry **267**(18): 12528.

Pan, W., Miao, H-Q, Xu, Y-J, Navarro, E. C., Tonra, J. R., Corcoran, E., Lahiji, A., Kussie, P., Kiselyov, A. S., Wong, W. C., and Liu, H. (2006). "1-[4-(1H-Benzoimidazol-2-yl)-phenyl]-3-[4-(1H-benzoimidazol-2-yl)-phenyl]-urea derivatives as small molecule heparanase inhibitors." Bioorganic and Medicinal Chemistry Letters **16**: 409-412.

Parish, C. R. (2006). "The role of heparan sulfate in inflammation." <u>Nature Reviews Immunology</u> **6**: 633-643.

Parish, C. R., Freeman, C., Brown, K. J., Francis, D. J., and Cowden, W. B. (1999). "Identification of sulfated oligosaccharide-based inhibitors of tumor growth and metastasis using novel in vitro assays for angiogenesis and heparanase activity." <u>Cancer Research</u> **59**: 3433-3441.

Perrimon, N., and Bernfield, M. (2000). "Specificities of heparan sulfate proteoglcans in developmental processes." <u>Nature</u> **404**: 725-728.

Petit, C., Rigg, G. P., Pazzani, C., Smith, A., Sieberth, V., Stevens, M., Boulnois, G., Jann, K., and Roberts, I. S. (1995). "Region 2 of the Escherichia coli K5 capsule gene cluster encoding proteins for the biosynthesis of the K5 polysaccharide." <u>Molecular Microbiology</u> **17**(4): 611-620.

Petitou, M., Herault, L. P., Bernat, A., Driguez, P. A., Duchaussoy, P., Lormeau, J. C. and Herbert, J. M. (1999). "Synthesis of thrombin-inhibiting heparin mimetics without side effects." <u>Nature</u> **398**: 417-422.

Pikas, D. S., Eriksson, I., and Kjellen, L. (2000). "Overexpression of Different Isoforms of Glucosaminyl N-Deacetylase/N-Sulfotransferase Results in Distinct Heparan Sulfate N-Sulfation Patterns." <u>Biochemistry</u> **39**: 4552-4558.

Pinhal, M. A. S., Smith, B., Olson, S., Aikawa, J. I., Kimata, K., and Esko, J. D. (2001). "Enzyme interactions in heparan sulfate biosynthesis: Uronosyl 5-epimerase and 2-O-sulfotransferase interact in vivo." <u>Proceedings of the National Academy of Sciences</u> **98**: 12984-12989.

Pye, D. A., Vives, R. R., Turnbull, J. E., Hyde, P., and Gallagher, J. T. (1998). "Heparan sulfate oligosaccharides require 6-O-sulfation for promotion of basic fibroblast growth factor mitogenic activity." <u>Journal of Biological Chemistry</u> **273**: 22936-22942.

Redini, F., Moczar, E., Poupon, M. F. (1986). "Cell surface glycosaminoglycans of rat rhabdomyosarcoma lines with different metastatic potentials and of non-malignant rat myoblasts." <u>Biochimica et Biophysica Acta</u> **883**: 98-105.

Roberson, E., and Firestone, M. (1992). "Relationship between desiccation and exopolysaccharide production in soil Pseudomonas sp." <u>Applied and Environmental Microbiology</u> **58**: 1284-1291.

Roberts, I. S., Saunders, F. K., and Boulnois, G. J. (1989). "Bacterial capsules and interactions with complement and phagocytes." <u>Biochemical Society Transactions</u> **17**: 462-464.

Rong, J., Habuchi, H., Kimata, K., Lindahl, U., and Kusche-Gullberg, M. (2001). "Substrate specificity of the heparan sulfate hexuronic acid 2-O-sulfotransferase." <u>Biochemistry</u> **40**: 5548-5555.

Rosenberg, R. D., and Damus, P.S. (1973). "The purification and mechanism of action of human antithrombin-heparin cofactor." <u>Journal of Biological Chemistry</u> **248**(18): 6490-6505.

Ryan, K. J., and Ray, C. G. (editors) (2004). <u>Sherris Medical Microbiology</u>, McGraw Hill.

Sala, R. F., MacKinnon, S. L., Palcic, M. M., and Tanner, M. E. (1998). "UDP-N-trifluoroacetylglucosamine as an alternative substrate in N-acetylglucosaminyltransferase reactions." <u>Carbohydrate Research</u> **306**: 127-136.

Salas, A., Sans, M., Soriano, A., Reverter, J.C., Anderson, D.C., Piqué, J.M., and Panés, J. (2000). "Heparin attenuates TNF-alpha induced inflammatory response through a CD11b dependent mechanism." <u>Gut</u> **47**: 88-96.

Sanderson, R. (2001). "Heparan sulfate proteoglycans in invasion and metastasis." <u>Seminars in Cell and Developmental Biology</u> **12**: 89-98.

Saobhany, M., Dong, J., and Negishi, M. (2005). "Two-step mechanism that determines the donor binding specificity of human UDP-N-acetylhexosaminyltransferase." <u>Journal of Biological Chemistry</u> **280**: 23441-23445.

Sasisekharan, R., Shriver, Z., VenkataranMan, G., and Narayanasami, U. (2002). "Roles of heparan-sulphate glycosaminoglycans in cancer." <u>Nature Reviews Cancer</u> **2**: 521-528.

Sato, T., Gotoh, M., Kiyohara, K., Akashima, T., Iwasaki, H., Kameyama, A., Mochizuki, H., Yada, T., Inaba, N., Togayachi, A., Kudo, T., Asada, M., Watanabe, H., Imamura, T., Kimata, K., and Narimatsu, H. (2003). "Differential roles of two N-acetylgalactosaminyltransferases, CSGalNAcT-1, and a novel enzyme, CSGalNAcT-2. Initiation and elongation in synthesis of chondroitin sulfate." <u>Journal of Biological Chemistry</u> **278**: 3063-3071.

Seko, A., Dohmae, N., Takio, K., and Yamashita, K. (2003). "ß1,4-Galactosyltransferase

(beta 4GalT)-IV is specific for GlcNAc 6-O-sulfate. ß4GalT-IV acts on keratan sulfate-related glycans and a precursor glycan of 6-sulfosialyl-Lewis X." Journal of Biological Chemistry **278**: 9150-9158.

Shaklee, P. N., and Conrad, H. E. (1984). "Hydrazinolysis of heparin and other glycosaminoglycans." Biochemical Journal **217**: 187-197.

Shively, J. E., and Conrad, H. E. (1976). "Formation of anhydrosugars in the chemical depolymerization of heparin." Biochemistry **15**(18): 3932-3942.

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). "A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry." Cell **99**: 13-22.

Shworak, N. W., Liu, J., Fritze, L. M. S., Schwartz, J. J., Zhang, L., Logeart, D., and Rosenberg, R. D. (1997). "Molecular Cloning and Expression of Mouse and Human cDNAs Encoding Heparan Sulfate D-Glucosaminyl 3-O-Sulfotransferase." Journal of Biological Chemistry **272**(44): 28008-28019.

Shworak, N. W., Liu, J., Petros, L., Zhang, L., Kobayashi, M., Copeland, N., Jenkins, N., and Rosenberg, R. (1999). "Multiple isoforms of heparan sulfate D-glucosaminyl 3-O-sulfotransferase. Isolation, characterization, and expression of human cdnas and identification of distinct genomic loci." Journal of Biological Chemistry **274**: 5170-5184.

Shworak, N. W., Liu, J., Petros, L.M., Zhang, L., Kobayashi, M., Copeland, N.G., Jenkins, N.A., and Rosenberg, R.D. (1999). "Multiple Isoforms of Heparan Sulfate D-Glucosaminyl 3-O-Sulfotransferase. ISOLATION, CHARACTERIZATION, AND EXPRESSION OF HUMAN cDNAs AND IDENTIFICATION OF DISTINCT GENOMIC LOCI." Journal of Biological Chemistry **274**(8): 5170-5184.

Sinaÿ, P., Jacquinet, J. C., Petitou, M., Duchaussoy, P., Lederman, I., Choay, J., and Torri, G. (1984). "Total synthesis of a heparin pentasaccharide fragment having high affinity for antithrombin III." Carbohydrate Research **132**: C5-C9.

Sismey-Ragatz, A. E., Green, D. E., Otto, N. J., Rejzek, M., Field, R. A., and DeAngelis, P. L. (2007). "Chemoenzymatic synthesis with distinct Pasteurella heparosan synthases: monodisperse polymers and unnatural structures." Journal of Biological Chemistry **282**(39): 28321-28327.

Slaughter, T. F., and Greenberg, C. S. (1997). "Heparin-associated thrombocytopenia and thrombosis: implications for perioperative management." Anesthesiology **83**: 667-675.

Smeds, E., Habuchi, H., Do, A.-T., Hjertson, E., Grundberg, H., Kimata, K., Lindahl, U., and Kusche-Gullberg, M. (2003). "Substrate specificities of mouse heparan sulphate glucosaminyl 6-O-sulphotransferases." Biochemical Journal **372**(Pt 2): 371-380.

Sueyoshi, T., Kakuta, Y., Pedersen, L. C., Wall, F. E., Pedersen, L. G., and Negishi, M. (1998). "A role of Lys614 in the sulfotransferase activity of human heparin sulfate N-deacetylase/N-sulfotransferase." FEBS Letters **433**: 211-214.

Sugahara, K., and Kitagawa, H. (2000). "Recent advances in the study of the biosynthesis and functions of sulfated glycosaminoglycans." Current Opinion in Structural Biology **10**(5): 518-527.

Tabeur, C., Mallet, J.M., Bono, F., Herbert, J.M., Petitou, M., and Sinaÿ, P. (1999). "Oligosaccharides corresponding to the regular sequence of heparin: chemical synthesis and interaction with FGF-2." Bioorganic and Medicinal Chemistry **7**(9): 2003-2012.

Tomer, K. B., Moseley, M. A., Deterding, L. J., and Parker, C. E. (1994). "Capillary liquid chromatography/mass spectrometry." Mass Spectrometry Reviews **13**: 431-457.

The New York Times, E. (April 28, 2008). "The Frightening Heparin Case." Bioorganic and Medicinal Chemistry **7**(9): 2003-2012.

Trybala, E., Bergström, T., Svennerholm, B., Jeansson, S., Glorioso, J.C., Olofsson, S. (1994). "Localization of a functional site on herpes simplex virus type 1 glycoprotein C involved in binding to cell surface heparan sulphate." Journal of General Virology **75**: 743-752.

Vann, W. F., Schmidt, M. A., Jann, B., and Jann, K. (1981). "The structure of the capsular polysaccharide (K5 antigen) of urinary-tract-infective Escherichia coli 010:K5:H4. A polymer similar to desulfo-heparin." European Journal of Biochemistry **116**: 359-364.

Vlodavsky, I., Friedmann, Y., Elkin, M., Aingorn, H., Atzmon, R., Ishai-Michaeli, R., Bitan, M., Pappo, O., Peretz, T., Michal, I., Spector, L., and Pecker, I. (1999). "Mammalian heparanase: Gene cloning, expression and function in tumor progression and metastasis." Nature Medicine **5**: 793-802.

Vlodavsky, I., Ilan, N., Naggi, A., and Casu, B. (2007). "Heparanase: Structure, Biological Functions, and Inhibition by Heparin-Derived Mimetics of Heparan Sulfate." Current Pharmaceutical Design **13**(20): 2057-2073.

Vlodavsky, I., Miao, H-Q, Medalion, B., Danagher, P., and Ron, D. (1996). "Involvement of heparan sulfate and related molecules in sequestration and growth promoting activity of fibroblast growth factor." Cancer and Metastasis Reviews **15**: 177-186.

Vuillemenot, A., Schiele, F., Meneveau, N., Claudel, S., Donat, F., Fontecave, S., Cariou, R., Samama, M. M., and Bassand. J. P. (1999). "Efficacy of a synthetic pentasaccharide, a pure factor Xa inhibitor, as an antithrombotic agent--a pilot study in the setting of coronary angioplasty." Thrombosis and Haemostasis **81**(2): 214-220.

Wang, L., Fuster, M., Sriramarao, P., and Esko, J.D. (2005). "Endothelial heparan sulfate deficiency impairs L-selectin- and chemokine-mediated neutrophil trafficking during inflammatory responses." Nature Immunology **6**(9): 902-910.

Westerduin, P., van Boeckel, C. A. A, Basten, J. E. M., Broekhoven, M. A., Lucas, H., Rood, A., van der Heijden, H., van Amsterdam, R. G. M., van Dinther, T. G., Meuleman, D. G., Visser, A., Vogel, G. M. T., Damm, J. B. L., and Overklift, G. T. (1994). "Feasible synthesis and biological properties of six 'non-glycosamino' glycan analogues of the antithrombin III binding heparin pentasaccharide." Bioorganic and Medicinal Chemistry **2**: 1267-1280.

Wu, Z. L., Lech, M., Beeler, D. L., and Rosenberg, R. D. (2004). "Determining Heparan Sulfate Structure in the Vicinity of Specific Sulfotransferase Recognition Sites by Mass Spectrometry." Journal of Biological Chemistry **279**: 1861-1866.

Xia, G., Chen, J., Tiwari, V., Ju, W., Li, J. P., Malmström, A., Shukla, D., and Liu, J. (2002). "Heparan Sulfate 3-O-Sulfotransferase Isoform 5 Generates Both an Antithrombin-binding Site and an Entry Receptor for Herpes Simplex Virus, Type 1." Journal of Biological Chemistry **277**: 37912-37919.

Xu, D., Moon, A. F., Song, D., Pedersen, L. C, Liu, J. (2008). "Engineering sulfotransferases to modify heparan sulfate." Nature Chemical Biology **4**: 200-202.

Yalpani, M. e. (1996). "Biomedical Functions and Biotechnology of Natural and Artificial Polymers." ATL Press: 45-61.

Zhang, L. J., Beeler, D. L., Lawrence, R., Lech, M., Liu, J., Davis, J. C., Shriver, Z., Sasisekharan, R., and Rosenberg, R. D. (2001). "6-O-Sulfotransferase-1 Represents a Critical Enzyme in the Anticoagulant Heparan Sulfate Biosynthetic Pathway." Journal of Biological Chemistry **276**(45): 42311-42321.

Zhang, W. (2004). "Fluorous Tagging Strategy for Solution-Phase Synthesis of Small Molecules, Peptides and Oligosaccharides." <u>Current Opinion in Drug Discovery and Development</u> **7**: 784-797.